August 2014

Estimating the Effective Number of Breeders of Brook Trout, Salvelinus fontinalis, Over Multiple Generations in Two Stream Systems

Matthew R. Cembrola
University of Massachusetts Amherst

ESTIMATING THE EFFECTIVE NUMBER OF BREEDERS OF BROOK TROUT, *SALVELINUS FONTINALIS*, OVER MULTIPLE GENERATIONS IN TWO STREAM SYSTEMS

A Thesis Presented

By

MATTHEW R. CEMBROLA

Submitted to the Graduate School of the

University of Massachusetts Amherst in partial fulfillment

of the requirements for the degree of

MASTER OF SCIENCE

May 2014

Wildlife and Fisheries Conservation

ESTIMATING THE EFFECTIVE NUMBER OF BREEDERS OF BROOK TROUT, *SALVELINUS FONTINALIS*, OVER MULTIPLE GENERATIONS IN TWO STREAM SYSTEMS

A Thesis Presented

By

MATTHEW R. CEMBROLA

Approved as to style and content by:

_____

Andrew R. Whiteley, Chair

_____

Benjamin H. Letcher, Member

_____

David B. Kittredge, Member

_____

Curtice R. Griffin, Department Head

Department of Environmental Conservation

ACKNOWLEDGEMENTS

I want to thank my advisor, Andrew Whiteley, for all of his encouragement and support throughout my graduate studies: for taking me on as a student, supporting me as a research assistant, helping me understand the concepts described in this thesis and patiently guiding me through this research. I also want to express my sincere thanks to my committee members Ben Letcher and Dave Kittredge for their feedback and encouragement. Special thanks to Matt O'Donnell for all of his help with the Stanley Brook data and field work.

This work is ongoing in both WB and SB as part of long-term research headed by the USGS Conte Lab with collaboration from the USFS Northern Research Station and UMass Amherst. Collaborators for the SB study also include the National Park Service (NPS) and USGS Cooperative Fish and Wildlife Research Unit at the University of Maine.

Todd Dubreuil, Doug Sigourney, Bruce Connery, Bic Wheeler, Joe Zydlewski, Keith Nislow, and many others were of great help to me and the larger studies. Jason Coombs was extremely helpful with genotyping questions and providing insight into brook trout biology. I thank Maili Page, Steve Jane, Matt Burak, Zak Robinson, Morgan Lindemeyer, Katherine Terkanian, and everyone in the conservation genetics lab at UMass. I deeply appreciate the friendly environment provided by the ECo department and my fellow grad students throughout my graduate studies. Tim Lens, whose undergraduate honors thesis helped get my own research going, was a great help, as were Elizabeth Crowley and Meaghan Horak. Lastly I want to thank my family for their love and support, and I am especially grateful towards my parents for their encouragement and understanding.

ABSTRACT

ESTIMATING THE EFFECTIVE NUMBER OF BREEDERS OF BROOK TROUT, *SALVELINUS FONTINALIS*, OVER MULTIPLE GENERATIONS IN TWO STREAM SYSTEMS

MAY 2014

MATTHEW R. CEMBROLA, B.S. WORCESTER POLYTECHNIC INSTITUTE

M.S., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Andrew Whiteley

The use of genetic markers in monitoring populations has become increasingly important for conservation purposes, and can take many forms. While effective population size (N_e) is of great interest to conservation genetics, it may be much easier and more practical to estimate the effective number of breeders (N_b) per cohort as a tool for genetic monitoring of populations. Few studies have estimated N_b for the same species over long periods of time in comparison with demographic or environmental variables. I estimated N_b of the eastern brook trout, *Salvelinus fontinalis*, as part of long-term studies of two stream systems: West Brook (WB) in Massachusetts and Stanley Brook (SB) in Maine. I used eight microsatellite loci for all available young of the year (YOY) from each cohort in WB and a random subset of YOY distributed evenly throughout SB to obtain genetic-based estimates. I estimated adult abundance (N_C) from mark-recapture data, and used seasonal stream flow as an environmental variable. I performed linear models with N_b as the response variable and family structure (number of families and variance in family size), N_C, and seasonal stream flow as predictor variables. I found that both the number of families and variance in family size had a strong influence on N_b. Compared to abundance of adults and YOY, N_b was relatively stable

over time. Stream flow in both autumn and spring showed a quadratic relationship with N_b in WB, suggesting that intermediate flows are optimal for maintaining a higher N_b. SB, with fewer years of data, did not show these relationships. If incorporated into monitoring programs, N_b can be a useful tool for detecting changes in population status and for informing management decisions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

Within the past decade genetic monitoring has become a useful tool for the conservation and management of species (Schwartz et al. 2007). Genetic monitoring is defined as the use of molecular markers to quantify changes in population metrics over time and, in some cases, space (Schwartz et al. 2007). Genetic monitoring provides an opportunity to observe a study site over time and see trends, rather than a "snapshot" or single observation (Schwartz et al. 2007). Monitoring genetic metrics has become increasingly useful and reliable and complementary to estimates of numerical abundance (Schwartz et al. 2007) and in some cases, genetic metrics may provide improved estimates of population trend, compared to numerical abundance estimates (Antao et al. 2011). Genetic metrics provide complementary information about population resilience that, when added to demographic monitoring efforts, promise to provide a more comprehensive picture of population resilience to environmental change. Additionally, it is possible that genetic metrics are highly responsive to management actions such as habitat improvement and sensitive to habitat degradation, however this assertion requires further testing.

There are two primary categories of genetic metrics that can be used for genetic monitoring. Genetic metrics that quantify within-population genetic diversity fall in the first category. These include heterozygosity and number of alleles, which can help predict a population's ability to adapt to future changes. For example, if a population bottleneck were detected, the decreased genetic diversity makes the population more susceptible to inbreeding depression and disease. However, genetic diversity is not

sensitive enough to short-term population fluctuations to be an effective metric for monitoring population trends over longer periods of time (Schwartz et al. 2007).

Genetic metrics in the second category measure the effective size of a population (N_e). N_e is one of the most fundamental parameters in evolutionary biology (Hare et al. 2011). It is defined as the size of a theoretical (and imaginary) ideal population affected by genetic drift at the same rate per generation as the population under consideration (Wright 1931). N_e is important because it mediates the influence of genetic drift and natural selection for a given population. N_e allows prediction of a population's adaptive potential in response to environmental change because it is closely related to a population's vulnerability to genetic drift (Hare et al. 2011). For iteroparous organisms with overlapping generations, it is challenging to reliably estimate generational N_e (Waples & Yokota 2007; Waples & Do 2010); however, it is more straightforward to estimate the effective number of breeders per cohort (N_b) (Whiteley et al. 2012). Estimation of N_b uses the genetic information and the same methods developed for generational N_e but applies them to a single cohort, yielding an estimate of the effective number of breeders that gave rise to the cohort. Whiteley et al. (2012) demonstrated that N_b can be reliably and precisely estimated for brook trout (*Salvelinus fontinalis*), based on a single cohort (young of the year, YOY) and with a sufficient sample size that avoids family over-representation.

The effective number of breeders (N_b) appears to provide information about reproductive output and success (Waples & Do 2010; Hare et al. 2011; Whiteley et al. 2012). This metric combines information from the number of families produced by the parents of a given cohort, the variance in reproductive success among those parents, and

early family-dependent survival of the offspring produced (Waples & Do 2010; Christie et al. 2012). Extremes of family structure are important to consider. If a cohort consists of few families that vary substantially in size, N_b will be low. At the other extreme, a cohort that contains many small families will have a large N_b. Early juvenile family-dependent mortality could have opposing effects on N_b. Family-dependent mortality that targets entire families, particularly small families, can lead to increased skew in reproductive success and reduce N_b. However, loss of entire large families could result in less variance in reproductive success, which would increase N_b.

N_b appears to be an estimable genetic metric that is habitat-dependent in a way that is directly related to recruitment (Whiteley et al. 2013). N_b might also serve to rank population risk and as the foundation for genetic monitoring efforts. Simulation studies suggest that N_b may be more effective than population abundance for monitoring population trend (Tallmon et al. 2010). There remains, however, an incomplete understanding of variation in N_b in natural populations. A number of critical information gaps exist regarding factors that influence N_b in wild populations. Further understanding of these issues is needed before N_b can be widely used for monitoring wild populations. These include: 1) the time frame to which N_b estimates apply, 2) temporal variation in N_b within populations, and 3) demographic and environmental drivers of N_b.

First, uncertainty remains regarding the time frame to which N_b estimates apply. N_b estimates may apply to a combination of 1) the single reproductive event that gives rise to a cohort and 2) "legacy" genetic effects that persist over the past one or few generations. Small populations are likely to have relatively few parents contributing to each reproductive event. Genetic drift that results from few parents mating generates

nonrandom associations among alleles at different loci, or linkage (gametic) disequilibrium (Waples 1991; Luikart et al. 2010; Waples 2010). The signal of linkage disequilibrium (LD) is used to estimate N_b, at least for the most commonly used estimator (Waples & Do 2008). LD, even among unlinked loci (loci on different chromosomes), can take several generations to decay and therefore may persist for one to few generations (Waples 1991). However, I hypothesize that family structure (defined as the number of families and variance in family size) within a cohort will have a predominant effect on N_b. Strong family structure (i.e. relatively few families with high variance in family size) limits the number of associations among alleles at different loci in progeny and therefore creates a cohort-specific LD signal. A strong relationship between a cohort's family structure and N_b would favor the hypothesis that N_b applies largely to a given cohort and not to a legacy LD effect.

Second, we have little understanding of how much N_b varies over time within populations. If N_b varies substantially over time, single point estimates may not be useful for genetic monitoring because they would not be representative of the effective number of breeders in that population. Instead, genetic monitoring may require an assessment of variation in N_b over time. Several studies have estimated variation in N_b over time within long-term studies. Duong et al. (2013) found that N_b varied approximately three-fold (ranged from 47 - 167) over 10 years in a single population of lake sturgeon (*Acipenser fulvescens*). Ardren & Kapuscinski (2003), in an 18-year study of rainbow trout (*Oncorhynchus mykiss*), found that N_b (estimated based on demographic parameters) varied 17-fold (range 5.8 - 101.3). Single cohort N_b ranged between 4 and ∞ (294 if infinite estimates excluded) over 31 years in a lake-dwelling population of brown trout

(*Salmo trutta*) (Charlier et al. 2012), though it should be noted that estimates lacked precision because low resolution genetic markers were used.

Third, demographic and environmental drivers responsible for variation in N_b have received little attention. Most attempts have involved examination of the relationship between N_b and the number of adults (N_C) present at the time of reproduction (Ardren & Kapuscinski 2003; Duong et al. 2013). N_b and N_C of reproducing adults might exhibit a positive relationship if the number of adults corresponds to the production of more families. However, there are a number of reasons why this relationship might be weak. More adults could correspond to elevated variance in family size (lowering N_b) if reproductive success is positively density dependent; larger male fish might be better able to monopolize access to females under higher densities. A similar phenomenon, termed genetic compensation, has been proposed for the relationship between variance in reproductive success and N_e at low population densities (Ardren & Kapuscinski 2003). The presence of few spawning individuals may cause reduced variance in reproductive success and therefore lead to relatively high N_b at low densities.

Environmental factors could influence variation in N_b if they increase the prevalence of family-dependent survival (Christie et al. 2012). Assuming no size-dependent fecundity, family-independent survival should not influence N_b because it should not change the family size distribution. However, family-dependent survival could either increase or decrease N_b, depending on the effect on variance in reproductive success. Mortality of large families could reduce variance in reproductive success and therefore increase N_b relative to a situation without family-dependent survival. Mortality

of small families could increase variance in reproductive success and therefore decrease N_b relative to a situation without family-dependent survival.

For stream-dwelling fishes, stream discharge is the environmental factor most likely to cause family-dependent survival, though little work has attempted to understand these relationships. For fall spawning species, low fall flows might limit spawning habitat and create more competition for spawning sites. Less reproduction with higher variance in success would lower N_b. On the other hand, high fall flows might also limit successful reproduction and could destroy nests (redds) in a family-dependent manner. The nature of family-dependent survival would then determine if the relationship between fall flow and N_b is positive or has an intermediate optimum. Spring flow might also cause family-dependent survival. Low flows could limit habitat and possibly food availability in a family-dependent manner. High flows could flush entire families out of stream systems. Again, an intermediate optimum spring flow is possible.

Addressing these factors that influence N_b requires detailed long-term studies of natural populations. Variation of N_b over time, the time frame of inference for N_b, and influence of demographic and environmental drivers of N_b can only be understood by comprehensive sampling within a population. Such sampling must allow estimates of N_b, family structure, and abundance. In addition, the time frame needs to be long enough to contain yearly environmental variation within seasons. Here, I used information from two long-term studies of brook trout populations, one in Massachusetts and one in Maine to ask the following questions:

1) What is the relationship between family structure and N_b?

2) How much does N_b vary over time?

3) What is the influence of environmental and demographic factors on N_b?

It should be noted that these issues are generally applicable to all organisms, but here I focus specifically on stream fish populations. In eastern North America, there is a need to monitor the population status of brook trout (*Salvelinus fontinalis*) (Hudy et al. 2008). The brook trout is an important indicator species of ecosystem health (Eastern Brook Trout Joint Venture 2014); however, competition with non-native species, and habitat fragmentation and degradation threaten its persistence (Hudy et al. 2008). Additionally, as a species dependent upon cold water, it is vulnerable to the effects of climate change (McKenna & Johnson 2011). Long-term studies within the native eastern distribution of brook trout have performed detailed demographic analyses (Letcher et al. 2007), and provide an opportunity to examine N_b over time in multiple populations of the same species for the first time. I addressed all three questions above in one long-term study in Massachusetts. I addressed questions 2 and 3 in another study in Maine.

CHAPTER 2

METHODS

2.1 Study sites

Brook trout were sampled from two long-term study sites in Massachusetts (West Brook) and Maine (Stanley Brook; Fig. 2). West Brook (hereafter WB) is a headwater stream in western Massachusetts, described in detail by Letcher et al. (2007). Sampling has been conducted four times a year since 2001. The WB study is ongoing. The WB mainstem consists of forty-seven sections, each approximately 20m long. Two-pass electrofishing was performed with block nets at the top and bottom of the sampling reach. Three tributaries were also sampled, but I only used genetic data from the mainstem. After capture, fish were measured, weighed, and an adipose or anal fin clip was taken as a tissue sample

Stanley Brook (hereafter SB) is a coastal stream in Acadia National Park in Maine. Sampling occurred twice a year (spring and autumn) from 2006 through 2012. This stream was sampled using two-pass electrofishing in fifty-three 40m sections divided with block nets. The first approximately 240m are estuarine. The mainstem branches after approximately 1320m, and the east branch is sampled for approximately 120m beyond the branch. The west branch is sampled approximately 680m upstream from where it joins the mainstem. After capture, fish were measured, weighed, and a fin clip was taken as a tissue sample. Data for this study were collected before and after the closure of the stream to fishing in October 2009.

2.2 Genetic analysis

I examined variation at eight microsatellite loci in 1,922 brook trout from WB and in 1,179 brook trout from SB (Table 1). DNA was extracted from fin clip tissue samples following a standard salt precipitation procedure. Polymerase chain reaction (PCR) was used to amplify the microsatellite loci *Sfo*-C113, *Sfo*-C88, *Sfo*-D75, *Sfo*-D100, *Sfo*-C24, *Sfo*-C115, *Sfo*-C129 (King et al. 2003), and *Ssa*-D237 (King et al. 2005). Four additional loci were added for the WB samples to test for effects on precision of genetic metrics. These four loci were *Sfo*-C38, *Sfo*-C86, *Sfo*-B52, and *Sfo*-D91a (King et al. 2012). We followed protocols for DNA extraction and amplification detailed in Whiteley et al. (2013). Loci were electrophoresed on either an ABI Prism 3100-Avant or an ABI Prism 3130xl genetic analyzer (Applied Biosystems Inc., Foster City, California), and alleles were hand-scored using GENEMAPPER version 3.2 and PEAK SCANNER version 1.0 software (Applied Biosystems Inc.). Positive controls of brook trout with known genotypes were used for each set of PCR and electrophoresis to ensure correct scoring of genotypes. For the 2007 WB cohort and 2007 SB cohort, genetic data were unavailable because the DNA did not amplify.

For WB, I used entire cohort samples for genetic analyses (Table 1). For SB, I randomly sampled in R (R Development Core Team, 2006) approximately 200 fish representing an even spatial distribution throughout each study site. Random sampling was conducted to avoid family over-representation effects (Whiteley et al. 2012). SB has 53 sections, thus I took a random sample of four fish per section. If fewer than four fish were available, I used all available fish from that section. For WB autumn samples with fewer than 200 individuals, I supplemented those cohorts with individuals from the

previous summer sample (YOY from same year) and following winter and spring samples (same cohort, fish then at age-1).

I used CREATE version 1.33 (Coombs et al. 2008) to create input files for FSTAT version 2.9.3.2 (Goudet 2002), GENEPOP version 4.0.10 (Rousset 2008), LDNE version 1.31 (Waples & Do 2008), and COLONY ver 1.2 (Wang 2004). I used FSTAT to calculate allele frequencies in each cohort and calculate the mean number of alleles per cohort (A_O); mean allelic richness, standardized to the cohort with the lowest number of individuals (A_R); mean expected heterozygosity (H_S); and a measure of departure from Hardy-Weinberg proportions (F_{IS}) for each cohort (Table 2). Tests for departures from Hardy-Weinberg (HW) proportions within each cohort across loci were performed with GENEPOP. I corrected for multiple tests with a Bonferroni procedure (Rice 1989) and sequential Bonferonni procedure (Narum 2006). I tested for significant gametic (linkage) disequilibrium (LD) with GENEPOP. I again corrected for multiple tests with the Bonferroni and sequential Bonferroni procedures.

I used the program LDNE (Waples & Do 2008) to obtain N_b estimates (Table 2). Based on the amount of linkage disequilibrium that occurs within a cohort, this is the most extensively tested single-sample estimator of N_b (Luikart et al. 2010). For small populations, genetic drift can induce linkage disequilibrium among pairs of unlinked loci (Waples & Do 2010). I used the monogamy model of LDNE and estimated 95% confidence intervals based on jackknifing. I used an allele frequency cutoff (P_{crit}) of 0.02, as this has been suggested to be the best balance between precision and bias (Waples & Do 2008; Whiteley et al. 2012).

In SB, I also split the random subsets of YOY into the mainstem and east branch as one group and the west branch as a second group, and estimated N_b for both, to test if breeders were using one site more than the other, and if this site selection could be influenced by environmental conditions.

2.3 Family structure

To understand the time frame to which single-cohort N_b estimates apply, I tested the relationship between N_b and aspects of family structure. These include the number of families produced and the variance in the size of those families. I predicted a strong positive relationship for both. To test the relative effects of these two factors, I constructed a model with both as predictors. A weak relationship would suggest that N_b is determined by factors other than family structure. For the estimator of N_b I used, this would likely be a signal of LD from past generations (Luikart et al. 2010).

Robust estimates of the number of full-sibling families and family size were only possible for WB. Robust estimation of family-level structure requires that the majority of a given cohort has been sampled and genotyped. The approximately 200 randomly selected individuals per cohort in SB did not meet this requirement. I used COLONY ver. 1.2 (Wang 2004) to estimate the number of full-sib families in each WB cohort, using all available individuals for a cohort. A previous study used empirically-parameterized simulations to confirm high accuracies of sibship reconstruction in WB (Letcher et al. 2011). These simulations revealed that for reconstructed full-sib families composed of at least two individuals, the rate of correct family inference was 91.2% (0.7% SE). For full-sib families composed of at least five individuals, the rate of correct family inference was

97.7% (0.4% SE). I estimated family evenness (*FE*) as a measure of variance in family size. *FE* was calculated using the equations: $FE = \frac{H\prime}{H\prime_{Max}}$, where $H' = -\sum_1^S p_i ln(p_i)$ and $H'_{Max} = \ln(S)$ (Mulder et al. 2004). *S* refers to the number of families and p_i refers to the proportion of the i^{th} family.

While incomplete genotyping of cohorts in SB precluded analyses to test the influence of family-level structure on N_b, I estimated family-level structure in SB to test whether randomly subsampling approximately 200 individuals resulted in relatively similar family representation across cohorts. I again used COLONY ver 1.2 and calculated *FE* within each SB cohort.

2.4 Factors that influence variation in N_b over time

My goal was to test the relative effects of a number of factors that may influence N_b over time within WB and SB (Fig. 1).

2.5 Adult abundance (N_C)

Census population size (N_C) is standardly defined in conservation genetics as the number of potentially reproducing adults in a population (Luikart et al. 2010). The number of adults at the time of reproduction could be a major driver of variation in N_b.

Estimates of N_C were based on abundance estimates from each autumn preceding the spring-defined cohort. These abundance estimates were obtained from long-term individual-based mark-recapture data. Abundance (N_C) and 95% confidence intervals were estimated as the count of age-1 and older fish divided by the probability of capture (*p*). The probability of capture (*p*) is the probability of detection given that an individual

is alive at the time of sampling and available for capture. I obtained estimates of p from ongoing demographic analyses of much larger data sets for both the WB and SB (Letcher et al. in review). Briefly, p was modeled as a function of body size with a logistic regression (Letcher et al. in review). I extracted estimates of p as the intercept coefficient (β_0) specific to the desired season (autumn), river (WB or SB), and age class (age-1 and older) for separate models performed for the WB and SB (Letcher et al. in review).

Coefficient estimates were back-transformed with the formula $1/(1+e^{-\beta})$. Abundance estimates included all age-1 fish and older because age at maturity is variable in brook trout, but both sexes tend to start spawning at age-1 (Hutchings 1994). We used one p for WB. SB was divided into three sections: the first 240m, the rest of the mainstem plus the east branch, and the west branch, each with its own capture probability. For SB, N_C was calculated for each stream section, then summed for a total N_C estimate for that sample. The last fall sample for SB was in 2012, so abundance and detection probability estimates were confounded. For that sample, I divided fish counts for each of the three stream sections by the average capture probability for that section across all other years of the study, then summed these values for the N_C estimate reported.

To estimate the number of YOY (N_{YOY}) present at the time of sampling, I divided the YOY counts by the adult capture probabilities for each sample. For WB, I used fall YOY counts divided by fall adult probabilities of detection. Estimates of p are not available for YOY in the fall because it is the first time they are large enough to tag upon capture. Therefore, I used fall YOY and p for adults (age-1 and older), though this will lead to an underestimate of YOY abundance because p for YOY is lower than p for adults

(Letcher et al. in review). YOY counts from 2003 to 2009 were included for WB (Table 1). As with N_C estimates in SB, I used YOY counts from each of the three stream sections and used section-specific capture probabilities to estimate YOY abundance. As with adults, counts and p for the last sample in SB were confounded, so I again used the average of p for each section across all other years of the study (Table 4).

2.6 Seasonal stream discharge

I tested the influence of stream flow during reproduction (autumn), egg incubation (winter), and the early juvenile phase (spring) on variation in N_b. Autumn stream flow was used as a surrogate for spawning habitat quality and availability. I used mean autumn discharge in the window from 1 October to 31 December. Brook trout spawn in approximately late October to mid-November in WB. This three-month window was meant to capture the spawning period and early winter flows. I also calculated mean discharge from 1 November to 30 November and 15 October to 31 December, to test for sensitivity of the length of this time window. Results did not differ and therefore the results for the three-month window were presented. As a surrogate for early rearing habitat quality, I used estimates of both winter and spring discharge. The two-month window from 1 January to 28 February was used for winter flow. The time period used for spring discharge was 1 March to 31 May. I also tested the time windows 1 February to 31 May and 1 February to 30 June, but again, results did not differ and only the three-month spring window was used for subsequent analyses.

For WB, discharge (cubic meters per second; cms) was estimated from a flow extension model (Nielsen 1999) based on discharge from a USGS gaging station located

on the Mill River, into which West Brook flows (Xu et al. 2010). I used the mean of daily mean flow over each window of time. For the period 6 February 2007 to 1 March 2007, ice prevented the gage from taking a daily mean, so I took the mean daily flow of available days for that period. For SB, data from a nearby USGS stream gage were not available; instead we used a stream depth data from a depth logger (Solinst Barologger) located approximately 440m upstream from the mouth as a surrogate for stream discharge. I present mean of the daily means for each window of time. Environmental data were not available for the 2006 SB cohort, as this was before data collection began.

2.7 Statistical modeling

I tested relative direction and magnitude of relationships in the conceptual diagram (Fig. 1) with linear models. For WB, I constructed a linear model with N_b as the response variable and the number of full-sib families and family evenness (*FE*) as the explanatory factors. However, because the number of families is generally higher with a larger sample size, in this case number of fish genotyped per cohort, I used the residuals from a regression of number of estimated families on number of fish genotypes per cohort. I used the logit transformation of *FE* per Warton & Hui (2011). I also performed linear models, applied separately to WB and SB, with N_C, autumn discharge, winter discharge, and spring discharge as explanatory factors and N_b as the response variable. For each of these three seasons, I fit a model with and without a quadratic term to test for intermediate optima. Statistical analyses were performed with R version 2.15.2 (R Development Core Team, 2006).

CHAPTER 3

RESULTS

3.1 West Brook

3.1.1 Genetic variation within cohorts

The WB summary statistics reflect the sampling strategy and effects of family structure per cohort. Mean A_O per cohort ranged from 7.8 to 10.5, mean A_R (standardized to lowest sample size per cohort) ranged from 7.3 to 8.5, mean H_S ranged from 0.593 to 0.639. Prior to correction for multiple tests, 40 of 64 (63%) tests for deviations from Hardy-Weinberg (HW) proportions were significant ($P < 0.05$), where 3.2 were expected by chance ($\alpha = 0.05$). Following sequential Bonferroni correction for 64 tests ($\alpha = 0.05$; initial nominal P value = 0.00625), 36 tests (56%) remained significant. Prior to correction for multiple tests, 158 of 224 (71%) tests for LD were significant ($P < 0.05$) for the entire data set, where 11.2 were expected by chance ($\alpha = 0.05$). Following sequential Bonferroni correction for 224 tests ($\alpha = 0.05$; initial nominal P value = 0.001786), 113 (50%) tests remained significant and the mean number of significant LD tests per population was 14.1 (range 6 to 25). The 2002, 2006, 2008, and 2009 cohorts, those with the lowest N_b and family evenness estimates, together comprise 53% of the number of total significant tests for HW ($P < 0.05$) and 56% of significant tests for LD ($P < 0.05$).

3.1.2 Influence of family structure on $\widehat{N}_b$

The mean estimated number of full-sibling families was 91, and ranged from 33 to 132. Mean family size ranged from 1.9 (in 2005) to 3.7 (in 2002). Mean family evenness (*FE*) was 0.925 and ranged from 0.855 (in 2002) to 0.969 (in 2004). The relationship between $\widehat{N}_b$ and standardized number of families was significant ($P < 0.05$), as was the relationship between $\widehat{N}_b$ and the logit transformation of *FE* (Table 3, Fig. 6). *FE* explained slightly more of the variation in $\widehat{N}_b$ (Table 3). Evenness and standardized number of families were highly positively correlated ($r = 0.84$, $P = 0.01$).

3.1.3 Variation in $\widehat{N}_b$ over time

The harmonic mean of $\widehat{N}_b$ across cohorts was 75.7. Point estimates of $\widehat{N}_b$ ranged from 48.9 (in 2002) to 127.5 (in 2004; Table 2, Fig. 3a). The coefficient of variation (CV) of $\widehat{N}_b$ was 0.34. $\widehat{N}_b$ varied over time with no positive or negative trend (Fig. 3a). Estimates of $\widehat{N}_b$ with eight and 12 loci were significantly correlated ($r = 0.96$, $P = 0.0001$).

3.1.4 Influence of demographic and environmental factors on $\widehat{N}_b$

Mean $\widehat{N}_C$ (age-1 and older) was 301.3 (range 82.1 - 750.8, CV = 0.72; Table 1; Fig 3a). The ratio of $\widehat{N}_b/\widehat{N}_C$ ranged from 0.097 to 0.624, again with no temporal trend (Table 2, Fig. 5a). The relationship between $\widehat{N}_b$ and $\widehat{N}_C$ was not significant ($P < 0.05$; Table 3), nor was the relationship with the added quadratic term. The number of YOY ($\widehat{N}_{YOY}$) estimated to be present in the autumn for each cohort ranged from 138.9 (in 2007) to 628.7 (in 2009) with a CV of 0.49. Variation in both $\widehat{N}_C$ and $\widehat{N}_{YOY}$ was substantially greater than $\widehat{N}_b$. Stream flow, as an environmental factor, explained substantial variation in $\widehat{N}_b$. Autumn flow with a quadratic term explained 49% of the variation in $\widehat{N}_b$ (Table 3). Spring flow with a quadratic term explained 84% of the variation in $\widehat{N}_b$ (Table 3; Fig. 7b).

3.2 Stanley Brook

3.2.1 Genetic variation within cohorts

The SB summary statistics reflect the sampling strategy, which appears to have removed the effects of family structure per cohort. Mean A_O per cohort ranged from 5.8 to 6.3, mean A_R (standardized to lowest sample size per cohort) ranged from 5.7 to 6.2, mean H_S ranged from 0.498 to 0.518. Prior to correction for multiple tests, 7 of 48 (15%) tests for deviations from HW proportions were significant ($P < 0.05$), where 2.4 were expected by chance ($\alpha = 0.05$). Following sequential Bonferroni correction for 48 tests ($\alpha = 0.05$; initial nominal P value = 0.00625), 4 tests remained significant. Prior to correction for multiple tests, 40 of 168 (24%) tests for LD were significant ($P < 0.05$) for the entire data set, where 8.4 were expected by chance ($\alpha = 0.05$). Following sequential Bonferroni correction for 168 tests ($\alpha = 0.05$; initial nominal P value = 0.00179), 13 (8%) tests remained significant and the mean number of significant LD tests per population was 2.2 (range 0 to 6).

3.2.2 Influence of family structure on $\widehat{N}_b$

In SB, I randomly sampled YOY from each stream reach. I could not test the relationship between $\widehat{N}_b$ and fully evaluated family structure. However, I used estimates of family structure obtained from these subsetted data to test the effect of my randomized sampling procedure. The mean estimated number of full-sibling families was 83, and ranged from 75 to 90. Mean family size ranged from 2.2 (in 2009) to 2.6 (in 2010). Mean FE was 0.967 and ranged from 0.961 (in 2008) to 0.972 (in 2006). Thus, I effectively avoided family overrepresentation with my sampling protocol.

3.2.3 Variation in $\widehat{N}_b$ over time

The harmonic mean of $\widehat{N}_b$ across cohorts was 186.9. Point estimates of $\widehat{N}_b$ ranged from 131.9 (in 2011) to 388.2 (in 2012; Table 5). The coefficient of variation (CV) of $\widehat{N}_b$ was 0.45. $\widehat{N}_b$ varied over time with no positive or negative trend (Fig. 3b). Mainstem-only $\widehat{N}_b$ estimates were not consistently higher than $\widehat{N}_b$ estimates for the west branch (Table 5).

3.2.4 Influence of demographic and environmental factors on $\widehat{N}_b$

Mean $\widehat{N}_C$ (age-1 and older) was 1380.3 (range 709.0 - 1902.4, CV = 0.30; Table 4, Fig. 3b). The ratio of $\widehat{N}_b/\widehat{N}_C$ ranged from 0.073 to 0.337 again with no trend (Fig. 5b). The relationship between $\widehat{N}_b$ and $\widehat{N}_C$ was not significant ($P < 0.05$), nor was the relationship with $\widehat{N}_C{}^2$ added (Table 6). The number of YOY ($\widehat{N}_{YOY}$) estimated to be present in the autumn for each cohort ranged from 110.3 (in 2008) to 1476.1 (in 2007) with a CV of 0.67. Variation in $\widehat{N}_{YOY}$ was substantially greater than $\widehat{N}_b$, though variation in $\widehat{N}_C$ was lower. Autumn flow with a quadratic term explained 79% of the variation in $\widehat{N}_b$ (Table 6, Fig. 8a).

CHAPTER 4

DISCUSSION

This is one of the first studies to compare N_b with demographic and environmental variables for relatively long time periods in two different sites for the same species. Three major results emerge from this study. First, N_b was strongly influenced by family structure, that is, both the number of families and variance in family size. Second, N_b was relatively stable over time, especially when compared to variance in adult and YOY abundance, with the exception of lower variability of adult abundance in SB. Third, I identified stream flow as a possible environmental driver, of variance in N_b.

4.1 Influence of family structure

My results confirm that the number of families and variance in family size are primary drivers of variation in N_b. It can take several generations for LD to decay, so N_b estimates based on LD can contain information from one or more previous generations (Waples 1991; Luikart et al. 2010). It has been generally assumed that N_b would contain this 'legacy' effect, however this assumption has not been tested to date. The strong relationships between N_b estimates in WB for both number of families and *FE* support the hypothesis that family structure, rather than a legacy effect, has a greater influence on N_b. This is important because a legacy effect of one to several generations would be confounded with the influence of family structure on N_b. That N_b estimates are cohort-specific and appear to largely reflect family structure at the time of spawning will make N_b a more useful metric for genetic monitoring, especially when the goal is to monitor trend in N_b over multiple successive cohorts.

4.2 Demographic and environmental drivers of N_b

In WB, N_b was more consistent over time than N_C and N_{YOY}. In SB, N_b was more stable than N_{YOY}, though N_C had a lower CV (mean N_C = 1380.3, SD = 412.2; mean N_b = 216.0, SD = 97.6). In WB, over-yearling abundance decreased over the study period. Despite this decrease in abundance, N_b was relatively stable. In SB, over-yearling abundance increased in the second half of the study, but again, N_b remained relatively stable. This suggests that N_b is relatively insensitive to demographic fluctuations of this magnitude. Stable N_b despite increasing N_C suggests that spawning habitat may be limited in SB. That is, changing numbers of spawners might not translate to more effective spawners because of limited spawning sites. Stable N_b despite declining N_C in WB is consistent with genetic compensation (Ardren & Kapuscinski 2003). It is possible that lower adult density leads to relatively less variance in reproductive success than at higher densities, which would translate to similar N_b estimates across adult densities. However, an additional regression of logit FE on NC was not significant ($F = 0.026$, $R^2 = 0.004$, $P = 0.88$). Instead, spawning site availability in WB might be saturated across the adult densities examined here. The more than doubled harmonic mean of SB (186.9) compared to WB (75.7) also suggests that there might be greater availability of spawning habitat in SB.

It is possible that variance across years within sites in mean adult body size, rather than simply N_C, could have an important influence on N_b. I did not formally include mean adult body size in the statistical analyses, but preliminary tests revealed no relationship with N_b. However, to verify if larger adults contribute disproportionately to the next generation would require a pedigree.

Serbezov et al. (2012) suggested that as N_C is often the only demographic parameter available, reporting the ratio of N_b/N_C may be useful for genetic monitoring. They report a range of N_b/N_C ratios of (0.16-0.28) over three years in a population of brown trout. This range is close to that of SB (0.07-0.38), but in WB this extends higher (0.10-0.62). Charlier et al. (2012) found that estimates of N_b fluctuated over a 23-year period, but with an overall increase that also corresponds with an increase in N_C. High variability and lack of patterns in N_b/N_C ratios across studies suggest that this ratio may have little utility for genetic monitoring.

I found evidence for stream flow influencing N_b in WB. The relationship between spring flow and N_b was stronger than autumn flow and N_b. The effect of autumn flow might be related to spawning habitat availability. Low autumn discharge might indicate low spawning habitat availability and therefore may be associated with low N_b. Higher autumn discharge may indicate the opposite: greater spawning habitat availability associated with higher N_b. An alternative is that extremely high autumn discharge may lead to suboptimal spawning conditions and lower N_b. For example, autumn floods could destroy particular families, lowering the number of available families, therefore lowering N_b as well.

Spring discharge may also have an important effect on N_b, and the direction of the relationship could be either positive or negative. Brook trout have been shown to remain in family groups after emergence (Hudy et al. 2010). The pattern of an intermediate optimum for spring flow suggests better juvenile survival that can lead to more family representation and higher N_b. Low flows were associated with low evenness of the full-sib family distributions. Many of the singleton families (full-sib families within only one

member) were not present. However, a small number of larger families still occurred. I hypothesize that larger families, resulting from larger fish, occurred in more optimal habitat. During low spring flows these larger families are able to survive. Small families, likely issuing from suboptimal spawning locations, may have lower survival under these conditions. High spring flows appear to have a similar (negative) effect on family-dependent survival. Timing of emergence in relation to high flows could also drive this effect.

SB showed less of a relationship with environmental drivers. The highest N_b was associated with the lowest autumn flows, but it is unclear why. Spawning conditions may be better in the mainstem than in the west branch under low flows. Splitting N_b estimates for the mainstem and west branch did not indicate any pattern related to autumn flow. This could therefore be a spurious relationship, and we have not identified the true drivers of variation in N_b in SB. Furthermore, spring flow did not show a significant relationship with N_b in SB.

4.3 Estimating N_b

Estimates of N_b from this study are likely to be unbiased. Sample sizes were generally above 100 and approached 200 for many cohorts. Whiteley et al. (2012) demonstrated that sample sizes greater than 75 generally provide unbiased and precise estimates, provided that samples were spaced out enough to avoid issues of family structure. Precision was also generally high. Several of the cohorts had wide confidence intervals, but none of the confidence intervals contained infinity, a problem that often arises with estimation of effective population size (Palstra & Ruzzante 2008). I used eight

microsatellite loci, which is on the lower end of a study like this. However, N_b estimates for WB based on 8 rather than 12 loci were highly similar. It appears that large sample size compensated for relatively few loci in this analysis. This tradeoff has been shown elsewhere (Waples & Do 2010). I observed a high rate of deviations from Hardy-Weinberg proportions and a strong signal of LD in both sites. This is likely due to family structure effects such as over-representation of certain families.

4.4 Conclusions

Monitoring N_b along with demographic and environmental factors can help to clarify management actions that could serve to increase N_b and maintain more genetic diversity. If intermediate autumn or spring discharge is optimal, managing flows could help to provide adequate habitat available for spawning. Because of the positive relationship between N_b and family structure, protecting juvenile survival in the spring may maintain a higher number of families, which should also serve to increase N_b. If this survival is evenly distributed throughout a stream system, increased family evenness will also benefit N_b. If habitat availability remains relatively constant, or is managed well, N_b may also remain stable, whether or not N_C fluctuates. N_b could serve as a valuable metric of stream-specific spawning and early rearing habitat quantity and quality. No other metric that encapsulates these characteristics is available. As such, this metric could serve to prioritize populations for conservation and management. It appears that N_b might not track small demographic fluctuations, but I conjecture that large positive fluctuations in response to management for habitat improvement, or large negative fluctuations in response to environmental degradation, will cause a response in N_b. Therefore, N_b estimation could serve as a valuable way to monitor these types of population response.

Table 1: Demographic and environmental data for the West Brook. N_G = number of genotyped fish used to estimate $\widehat{N}_b$. $\widehat{N}_{YOY}$ = estimated number of YOY at time of sampling. $\widehat{N}_C$ = estimated number of adults at time of sampling. Autumn = mean stream flow (in cms) for the time window 1 October - 31 December of the year preceding the cohort, reflecting the spawning period of the parents. Winter = mean stream flow 1 January - 28 February. Spring = mean flow flow 1 March - 31 May. N_{YOY} counts were unavailable for the 2001 and 2002 cohorts. Genotypes were unavailable for the 2007 cohort.

Cohort	N_G	$\widehat{N}_{YOY}$	$\widehat{N}_C$	Autumn	Winter	Spring
2001	332	-	750.8 (480.9 – 1790.0)	0.19	0.11	0.77
2002	197	-	503.1 (443.3 - 561.5)	0.04	0.05	0.31
2003	296	461.3 (403.0 - 555.0)	419.4 (385.1 - 462.3)	0.24	0.19	0.73
2004	157	519.1 (450.3 - 638.2)	276.9 (255.4 - 308.8)	0.75	0.34	0.64
2005	177	275.2 (238.7 - 334.0)	164.2 (149.3 - 184.8)	0.34	0.67	0.73
2006	79	177.8 (145.9 - 240.4)	165.3 (142.6 - 193.5)	0.92	1.06	0.37
2007	-	138.9 (113.2 - 186.5)	82.1 (76.5 - 90.5)	0.39	0.37	0.96
2008	358	544.8 (460.3 - 775.3)	149.0 (138.5 - 163.5)	0.07	0.68	0.90
2009	326	628.7 (528.6 - 874.4)	200.4 (187.1 - 217.7)	0.50	0.21	0.50

Table 2: WB genetic summary statistics. Note that 2007 cohort is not included, as genetic data were unavailable. HW = number of significant tests for departures from Hardy-Weinberg proportions following sequential Bonferroni correction. LD = number of significant tests for linkage disequilibrium following sequential Bonferroni correction. A_O = mean number of alleles. A_R = mean allelic richness. H_S = mean expected heterozygosity. F_{IS} = measure of departure from HW proportions. Num. fam. = number of families. Mean FS = mean family size. *FE* = family evenness. The ratio of $\widehat{N}_b$ to $\widehat{N}_C$ was based on the 8-locus $\widehat{N}_b$ estimate.

Cohort	HW	LD	A_O	A_R	H_S	F_{IS}	Num. fam.	Mean FS	*FE*	$\widehat{N}_b$ based on 8 loci	$\widehat{N}_b$ based on 12 loci	$\widehat{N}_b/\widehat{N}_C$ ratio
2001	8	16	10.5	8.5	0.617	0.063	123	2.7	0.943	102.2 (79.8-132.0)	95.2 (81.4-111.8)	0.136
2002	4	21	8.1	7.3	0.615	0.019	53	3.7	0.855	48.9 (40.7-58.4)	34.3 (29.1-40.2)	0.097
2003	3	8	9.6	8.3	0.631	0.015	132	2.2	0.940	109.5 (86.5-139.7)	127.1 (102.7-158.7)	0.261
2004	1	6	8.5	8.0	0.625	0.055	76	2.1	0.969	127.5 (91.6-184.1)	132.6 (107.3-166.3)	0.460
2005	3	8	9.6	8.4	0.639	0.052	94	1.9	0.963	102.5 (80.9-131.3)	126.7 (106.8-151.4)	0.624
2006	4	10	7.8	7.8	0.593	-0.016	33	2.4	0.918	49.1 (31.2-80.7)	45.4 (34.4-60.5)	0.297
2008	7	25	9.6	8.2	0.627	0.014	105	3.4	0.886	65.6 (52.5-81.5)	55.3 (46.8-65.0)	0.440
2009	6	19	9.4	8.1	0.625	0.026	113	2.9	0.923	76.7 (56.7-104.3)	81.9 (66.4-101.3)	0.383

Table 3: Linear models with $\widehat{N}_b$ as the dependent variable for WB. Number of families is standardized with residuals from a regression of number of estimated number of families on number of fish genotypes per cohort. Family evenness is logit transformed. $\widehat{N}_C$ refers to estimated abundance of fish age-1 and older at the time of sampling.

Model	*N*	*F*	*P*	R^2 (multiple)	R^2 (adjusted)
$\widehat{N}_C$	8	0.23	0.65	0.04	-0.12
$\widehat{N}_C + \widehat{N}_C{}^2$	8	0.12	0.89	0.05	-0.33
Number of families	8	14.89	0.008	0.71	0.66
Family evenness	8	21.0	0.004	0.78	0.74
Autumn flow	8	0.06	0.82	0.0096	-0.16
Autumn flow + autumn flow^2	8	2.44	0.18	0.49	0.29
Winter flow	8	0.68	0.44	0.10	-0.05
Winter flow + winter flow^2	8	2.03	0.23	0.45	0.23
Spring flow	8	2.77	0.15	0.32	0.20
Spring flow + spring flow^2	8	13.31	0.00995	0.84	0.78

Table 4: SB demographic and environmental data. N_G = number of genotyped fish used to estimate $\widehat{N}_b$. $\widehat{N}_{YOY}$ = estimated number of YOY at time of sampling. $\widehat{N}_C$= estimated number of adults at time of sampling. Autumn = mean stream depth (in cm) for the time window 1 October - 31 December of the year preceding the cohort, reflecting the spawning period of the parents. Winter = mean stream depth 1 January - 28 February. Spring = mean flow depth 1 March - 31 May. Stream depth data were not collected until July 2007. Genotypes were unavailable for the 2007 cohort.

Cohort	N_G	$\widehat{N}_{YOY}$	$\widehat{N}_C$	Autumn	Winter	Spring
2006	192	1106.3 (983.2 - 1266.1)	1411.0 (1258.1-1607.9)	-	-	-
2007	-	1476.1 (1283.1 - 1743.8)	1135.2 (1008.6-1298.5)	28.5	28.3	29.2
2008	203	110.3 (102.0 - 120.4)	1196.8 (1094.8-1326.1)	23.6	26.0	16.7
2009	195	519.7 (480.8 - 568.2)	709.0 (643.1-791.7)	22.8	18.2	22.5
2010	192	552.5 (516.6 - 596.3)	1493.0 (1360.3-1679.4)	20.9	20.0	19.2
2011	196	480.8 (427.6 - 540.0)	1814.6 (1607.6-2055.5)	20.3	17.5	19.9
2012	201	503.2 (450.7 - 569.5)	1902.4 (1589.7-2368.5)	18.8	18.0	17.8

Table 5: SB genetic summary statistics. Note that 2007 cohort is not included, as genetic data were unavailable. HW = number of significant tests for departures from Hardy-Weinberg proportions following sequential Bonferroni correction. LD = number of significant tests for linkage disequilibrium following sequential Bonferroni correction. A_O = mean number of alleles. A_R = mean allelic richness. H_S = mean expected heterozygosity. F_{IS} = measure of departure from HW proportions. Num. fam. = number of families. Mean FS = mean family size. *FE* = family evenness. $\widehat{N}_b$ mainstem refers to estimates of $\widehat{N}_b$ based on the genotyped fish from the mainstem and east branch only; $\widehat{N}_b$ west branch refers to estimates of $\widehat{N}_b$ based on genotyped fish from the west branch only. The ratio of $\widehat{N}_b$to $\widehat{N}_C$ was based on the $\widehat{N}_b$ estimate for the entire stream.

Cohort	HW	LD	A_O	A_R	H_S	F_{IS}	Num. fam.	Mean FS	*FE*	$\widehat{N}_b$	$\widehat{N}_b/\widehat{N}_C$ ratio	$\widehat{N}_b$ mainstem	$\widehat{N}_b$ west branch
2006	0	0	5.9	5.9	0.501	0.044	85	2.3	0.972	243.2 (141.4-508.8)	0.172	153.8 (98.3 - 270.5)	211.2 (96.7 - 2189.0)
2008	3	6	6.3	6.2	0.518	0.052	81	2.5	0.961	139.7 (97.9-207.4)	0.117	204.6 (123.2 - 435.9)	52.4 (38.8 - 71.7)
2009	0	1	6.3	6.2	0.498	-0.009	90	2.2	0.969	238.7 (158.0-399.0)	0.337	158.5 (98.5 - 293.4)	175.8 (101.1 - 440.1)
2010	0	2	5.9	5.9	0.511	0.005	75	2.6	0.966	154.1 (95.1-272.9)	0.103	163.0 (95.7 - 336.8)	79.8 (40.2 - 218.6)
2011	1	3	6.1	6.1	0.512	0.004	82	2.4	0.971	131.9 (98.8-179.8)	0.073	103.6 (75.0 - 148.0)	139.4 (63.9 - 777.1)
2012	0	1	5.8	5.7	0.506	0.001	87	2.3	0.962	388.2 (185.3-1674.3)	0.204	256.0 (129.2 - 860.9)	321.1 (116.6 – ∞)

Table 6: Linear models with $\widehat{N}_b$ as the dependent variable for SB. $\widehat{N}_C$ refers to estimated abundance of fish age-1 and older at the time of sampling.

Model	*N*	*F*	*P*	R^2 (multiple)	R^2 (adjusted)
$\widehat{N}_C$	6	0.2	0.68	0.05	-0.19
$\widehat{N}_C + \widehat{N}_C{}^2$	6	0.72	0.56	0.33	-0.12
Autumn flow	5	1.33	0.33	0.31	0.08
Autumn flow + autumn flow^2	5	3.98	0.20	0.80	0.60
Winter flow	5	0.63	0.48	0.17	-0.10
Winter flow + winter flow^2	5	0.22	0.82	0.18	-0.64
Spring flow	5	0.004	0.96	0.001	-0.33
Spring flow + spring flow^2	5	0.03	0.97	0.03	-0.94

Figure 1: Relationships examined among demographic and environmental variables and their potential effects on N_b. N_C refers to estimated abundance of fish age-1 and older at the time of sampling.

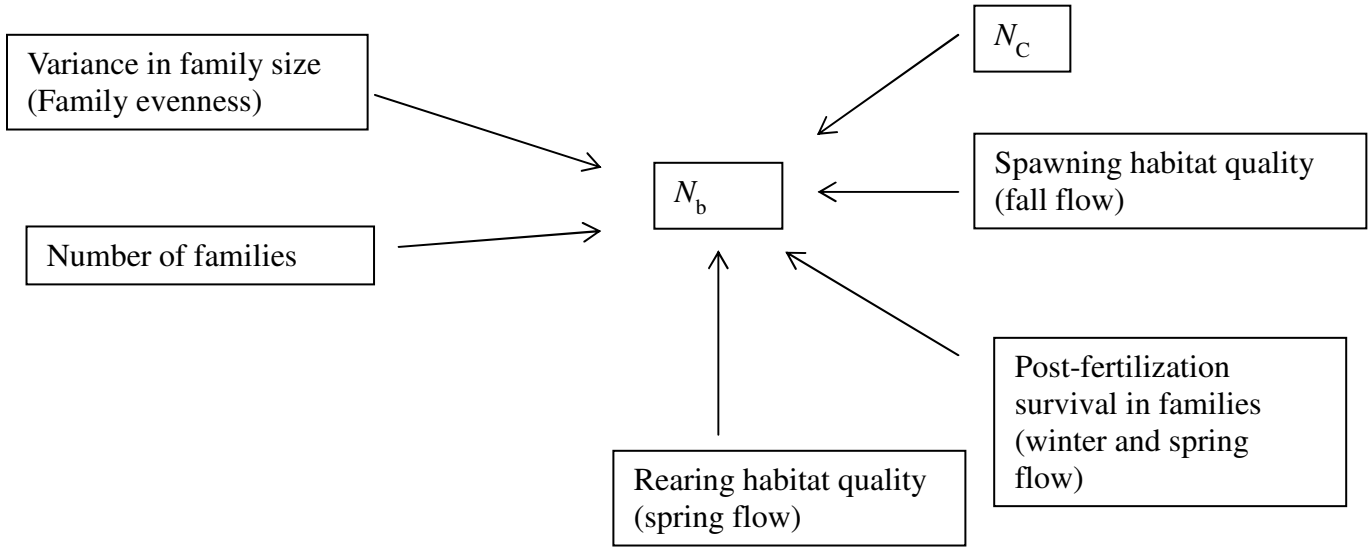


Figure 2: Map of historical eastern U.S. brook trout range, adapted from (Hudy et al. 2008). SB depicted in red, WB in yellow.

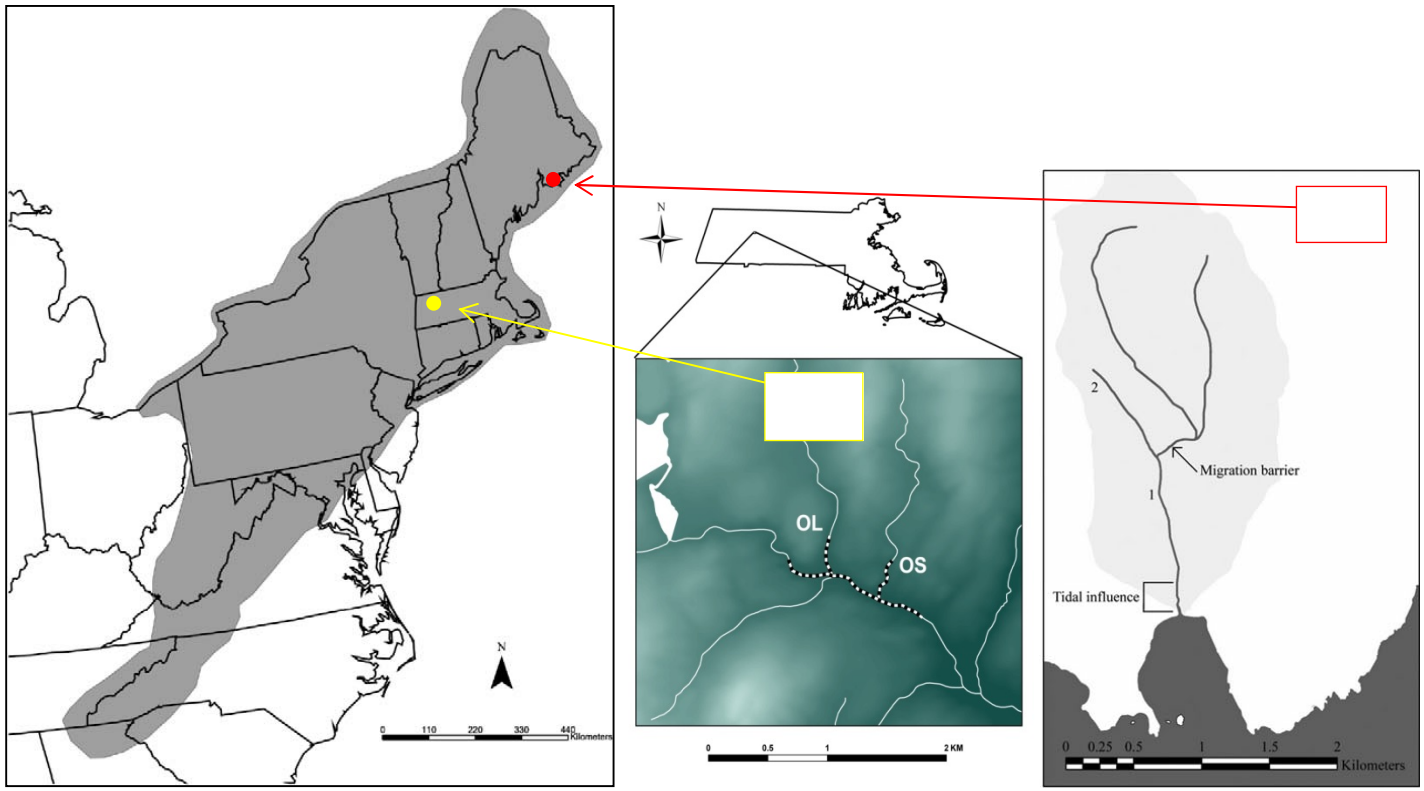
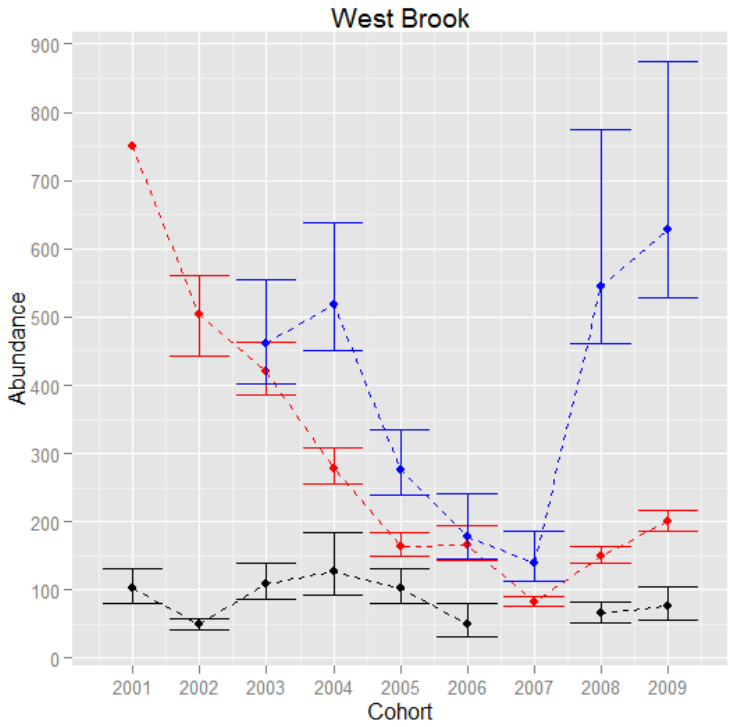
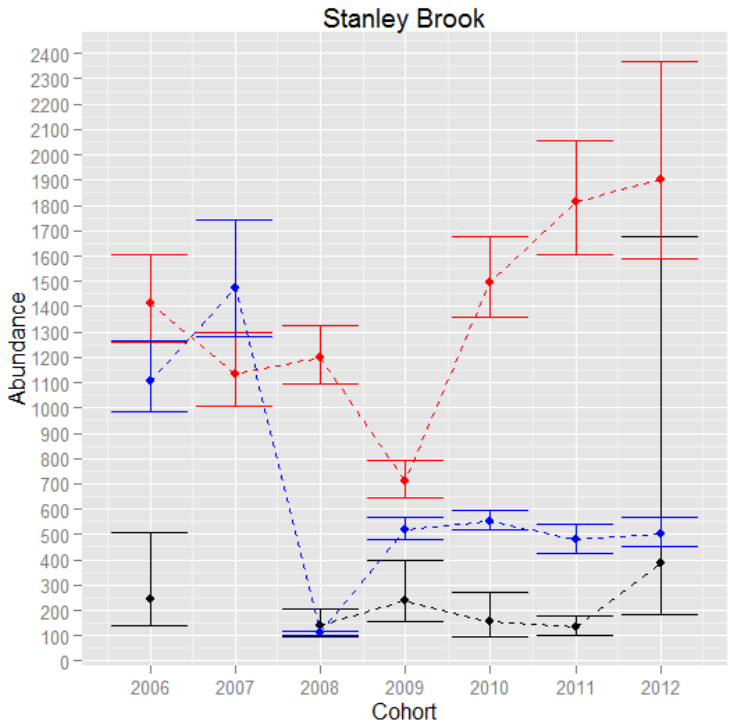


Figure 3: $\widehat{N}_b$ (black), $\widehat{N}_C$ (red), and $\widehat{N}_{YOY}$ (blue) over time in a) WB and b) SB. 95% CI not shown for $\widehat{N}_C$ of 2001 WB cohort.



a)



b)

Figure 4: Histograms of family size distribution per cohort in WB.

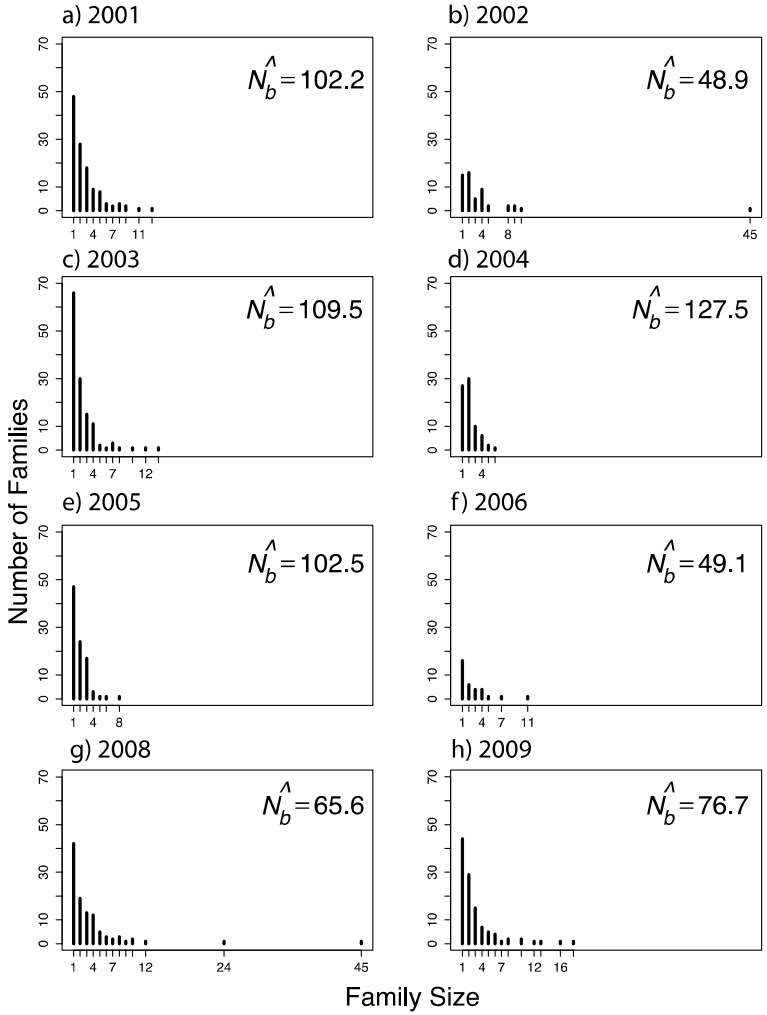
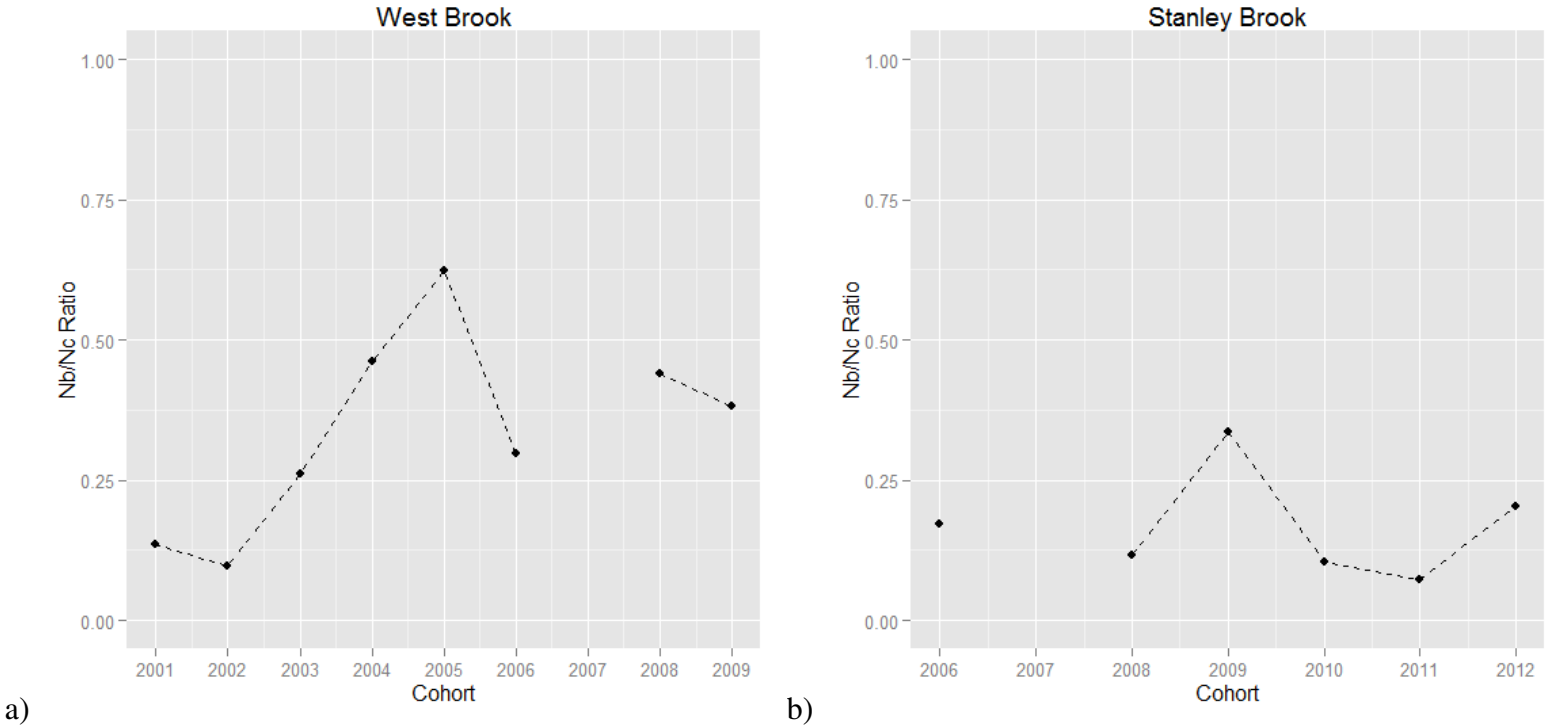
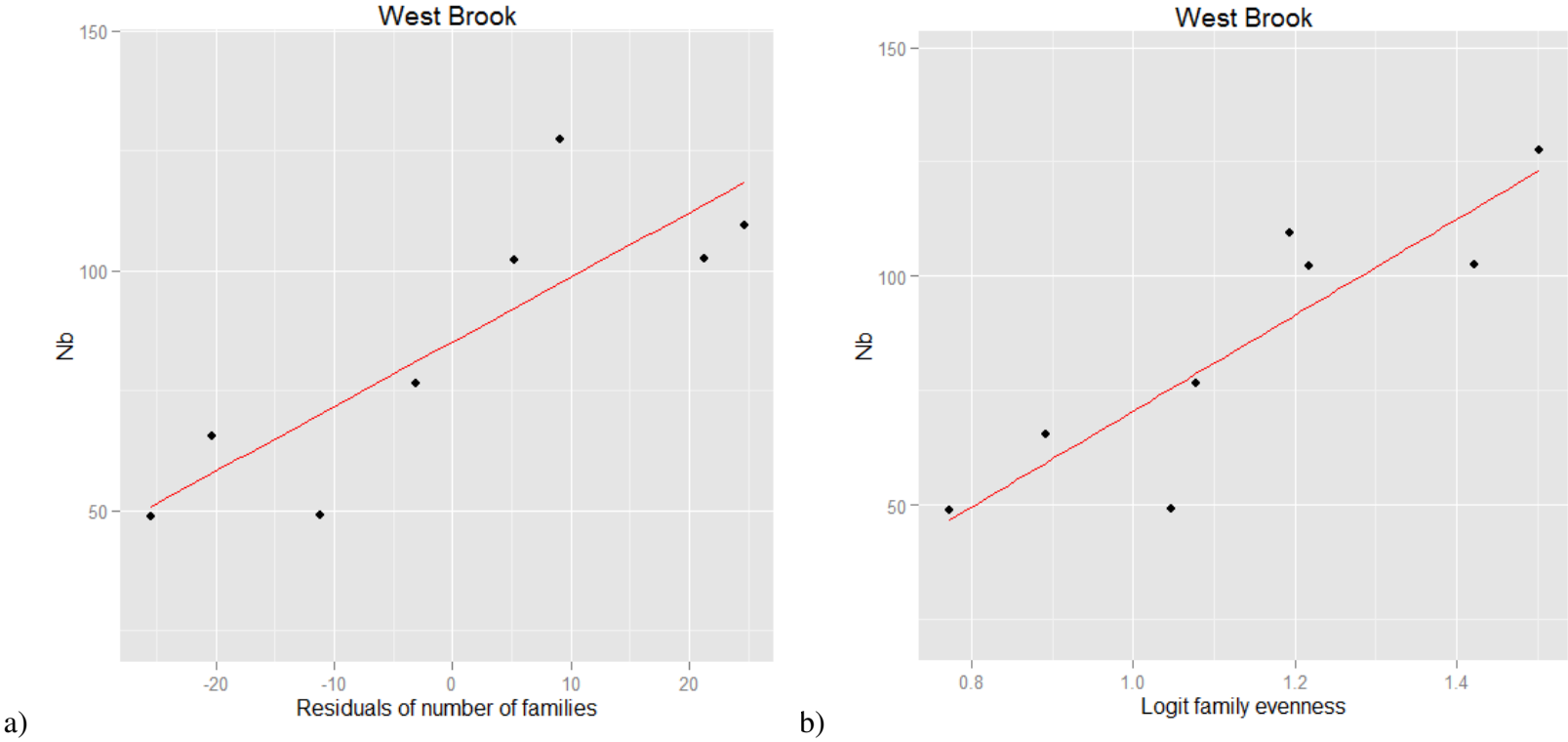


Figure 5: Ratio of $\widehat{N}_b$ to $\widehat{N}_C$ over time in a) WB and b) SB.



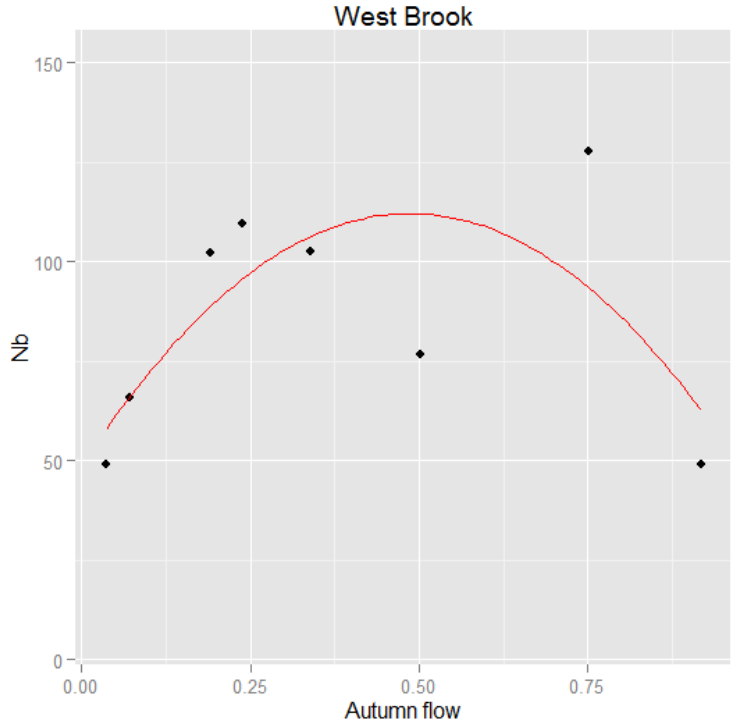
a) b)

Figure 6: Relationships between $\widehat{N}_b$ and a) residuals of number of families and b) logit family evenness in WB. A positive relationship is seen for both measures of family structure.

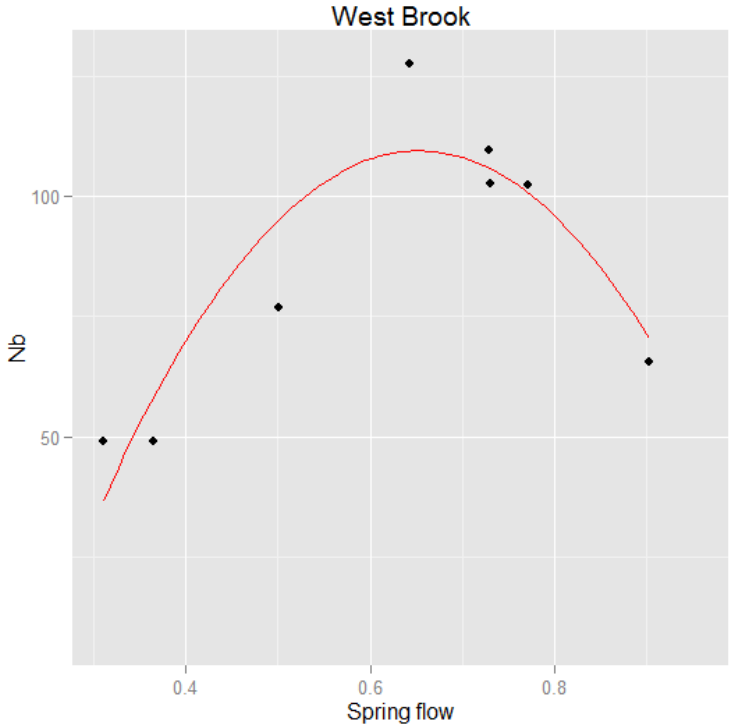


a) b)

Figure 7: Relationships between $\widehat{N}_b$ and stream flow (cubic meters per second) in WB a) in autumn and b) in spring. A quadratic relationship is seen in both seasons. The time window used for autumn was 1 October to 31 December of the year preceding each cohort, during the spawning period of the parents. The time window used for spring was 1 March to 31 May.

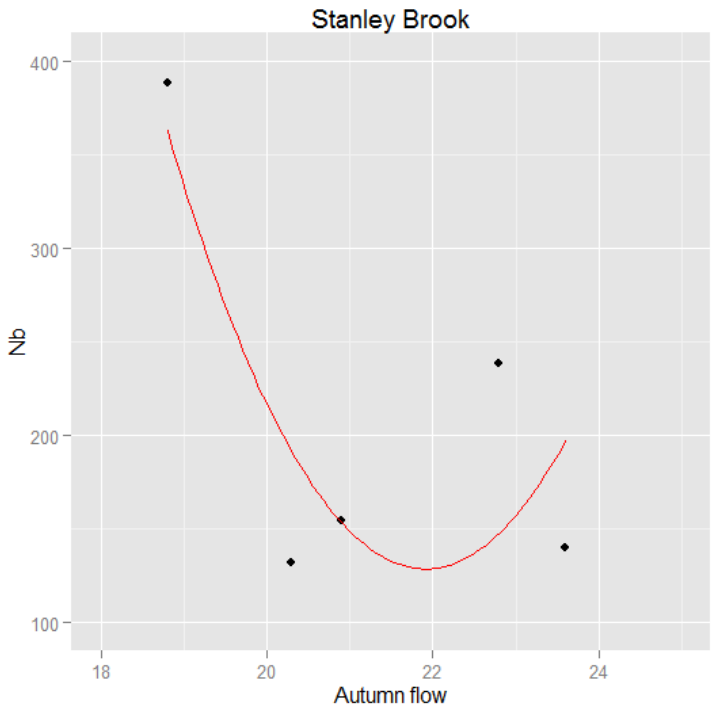


a)

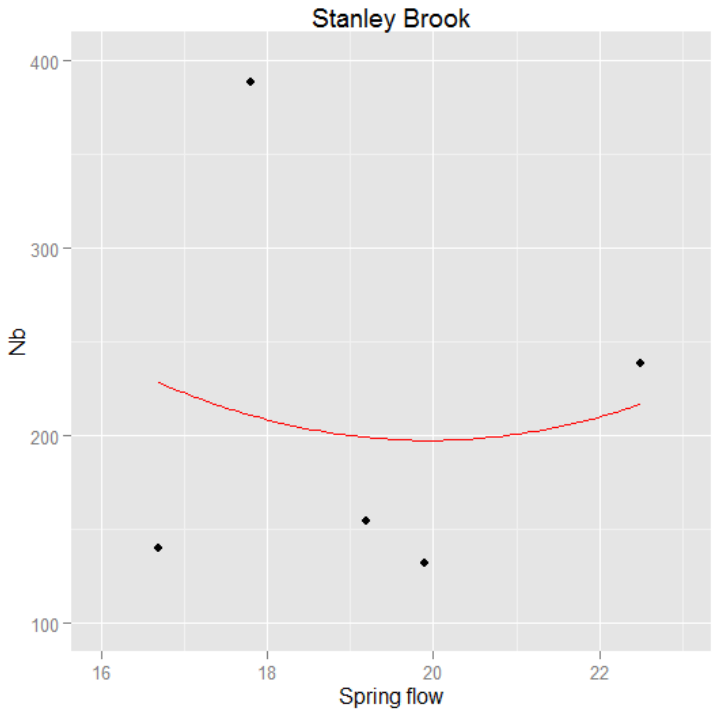


b)

Figure 8: Relationships between $\widehat{N}_b$ and stream depth in SB a) in autumn and b) in spring. No significant relationship is apparent in either season. The time window used for autumn was 1 October to 31 December of the year preceding each cohort, during the spawning period of the parents. The time window used for spring was 1 March to 31 May. Stream depth data were not available for the 2006 cohort.



a)



b)

LITERATURE CITED

Antao, T., Perez-Figueroa, A., & Luikart, G. (2011). Early detection of population declines: High power of genetic monitoring using effective population size estimators. *Evolutionary Applications, 4*(1), 144-154. doi:10.1111/j.1752-4571.2010.00150.x

Ardren, W., & Kapuscinski, A. (2003). Demographic and genetic estimates of effective population size (N-e) reveals genetic compensation in steelhead trout. *Molecular Ecology, 12*(1), 35-49. doi:10.1046/j.1365-294X.2003.01705.x

Charlier, J., Laikre, L., & Ryman, N. (2012). Genetic monitoring reveals temporal stability over 30 years in a small, lake-resident brown trout population. *Heredity, 109*(4), 246-253. doi:10.1038/hdy.2012.36

Christie, M. R., Marine, M. L., French, R. A., Waples, R. S., & Blouin, M. S. (2012). Effective size of a wild salmonid population is greatly reduced by hatchery supplementation. *Heredity, 109*(4), 254-260. doi:10.1038/hdy.2012.39

Coombs, J. A., Letcher, B. H., & Nislow, K. H. (2008). CREATE: A software to create input files from diploid genotypic data for 52 genetic software programs. *Molecular Ecology Resources, 8*(3), 578-580. doi:10.1111/j.1471-8286.2007.02036.x

Duong, T.Y., Scribner, K. T., Forsythe, P. S., Crossman, J. A., & Baker, E. A. (2013). Interannual variation in effective number of breeders and estimation of effective population size in long-lived iteroparous lake sturgeon (acipenser fulvescens). *Molecular Ecology, 22*(5), 1282-1294. doi:10.1111/mec.12167

Eastern Brook Trout Joint Venture. (2014). Eastern brook trout joint venture: A fish habitat partnership. Retrieved 3/19, 2014, from http://easternbrooktrout.org/

Goudet, J. (2002). *FSTAT version 2.9.3.2, A program to estimate and test gene diversities andfixation indices*. Lausanne, Switzerland: Lausanne University.

Hare, M. P., Nunney, L., Schwartz, M. K., Ruzzante, D. E., Burford, M., Waples, R. S., . . . Palstra, F. (2011). Understanding and estimating effective population size for practical application in marine species management. *Conservation Biology, 25*(3), 438-449. doi:10.1111/j.1523-1739.2010.01637.x

Hudy, M., Coombs, J. A., Nislow, K. H., & Letcher, B. H. (2010). Dispersal and within-stream spatial population structure of brook trout revealed by pedigree reconstruction analysis. *Transactions of the American Fisheries Society, 139*(5), 1276-1287. doi:10.1577/T10-027.1

Hudy, M., Thieling, T. M., Gillespie, N., & Smith, E. P. (2008). Distribution, status, and land use characteristics of subwatersheds within the native range of brook trout in the eastern united states. *North American Journal of Fisheries Management, 28*(4), 1069-1085. doi:10.1577/M07-017.1

Hutchings, J. (1994). Age-specific and size-specific costs of reproduction within populations of brook trout, salvelinus-fontinalis. *Oikos, 70*(1), 12-20. doi:10.2307/3545693

King, T., Julian, S., Coleman, R., & Burnham-Curtis, M. (2003). Isolation and characterization of novel tri- and tetranucleotide microsatellite DNA markers for brook trout Salvelinus fontinalis: GenBank submission numbers AY168187, AY168192,
AY168193, AY168194, AY168195, AY168197, AY168199. available: Ncbi:Nlm.nih.gov/nucleotide/

King, T.L., Eackles, M., & Letcher, B. (2005). Microsatellite DNA markers for the study of atlantic salmon (salmo salar) kinship, population structure, and mixed-fishery analyses. *Molecular Ecology Notes, 5*(1), 130-132. doi:10.1111/j.1471-8286.2005.00860.x

King, T. L., Lubinski, B. A., Burnham-Curtis, M. K., Stott, W., & Morgan,R.P.,,II. (2012). Tools for the management and conservation of genetic diversity in brook trout (salvelinus fontinalis): Tri- and tetranucleotide microsatellite markers for the assessment of genetic diversity, phylogeography, and historical demographics. *Conservation Genetics Resources, 4*(3), 539-543. doi:10.1007/s12686-012-9603-z

Letcher, B. H., Coombs, J. A., & Nislow, K. H. (2011). Maintenance of phenotypic variation: Repeatability, heritability and size-dependent processes in a wild brook trout population. *Evolutionary Applications, 4*, 602-615.

Letcher, B. H., Nislow, K. H., Coombs, J. A., O'Donnell, M. J., & Dubreuil, T. L. (2007). Population response to habitat fragmentation in a stream-dwelling brook trout population. *Plos One, 2*(11), e1139. doi:10.1371/journal.pone.0001139

Luikart, G., Ryman, N., Tallmon, D. A., Schwartz, M. K., & Allendorf, F. W. (2010). Estimation of census and effective population sizes: The increasing usefulness of DNA-based approaches. *Conservation Genetics, 11*(2), 355-373. doi:10.1007/s10592-010-0050-7

McKenna, J. E., Jr., & Johnson, J. H. (2011). Landscape models of brook trout abundance and distribution in lotic habitat with field validation. *North American Journal of Fisheries Management, 31*(4), 742-756. doi:10.1080/02755947.2011.593940

Mulder, C., Bazeley-White, E., Dimitrakopoulos, P., Hector, A., Scherer-Lorenzen, M., & Schmid, B. (2004). Species evenness and productivity in experimental plant communities. *Oikos, 107*(1), 50-63. doi:10.1111/j.0030-1299.2004.13110.x

Narum, S. R. (2006). Beyond bonferroni: Less conservative analyses for conservation genetics. *Conservation Genetics, 7*(5), 783-787. doi:10.1007/s10592-005-9056-y

Nielsen, J. P. (1999). *Record extension and streamflow statistics for the pleasant river, maine.* (No. WRIR 99-4078).USGS.

Palstra, F. P., & Ruzzante, D. E. (2008). Genetic estimates of contemporary effective population size: What can they tell us about the importance of genetic stochasticity for wild population persistence? *Molecular Ecology, 17*(15), 3428-3447. doi:10.1111/j.1365-294X.2008.03842.x

R Development Core Team. (2006). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rice, W. (1989). Analyzing tables of statistical tests. *Evolution, 43*(1), 223-225. doi:10.2307/2409177

Rousset, F. (2008). GENEPOP ' 007: A complete re-implementation of the GENEPOP software for windows and linux. *Molecular Ecology Resources, 8*(1), 103-106. doi:10.1111/j.1471-8286.2007.01931.x

Schwartz, M. K., Luikart, G., & Waples, R. S. (2007). Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology & Evolution, 22*(1), 25-33. doi:10.1016/j.tree.2006.08.009

Serbezov, D., Jorde, P. E., Bernatchez, L., Olsen, E. M., & Vollestad, L. A. (2012). Life history and demographic determinants of effective/census size ratios as exemplified by brown trout (salmo trutta). *Evolutionary Applications, 5*(6), 607-618. doi:10.1111/j.1752-4571.2012.00239.x

Tallmon, D. A., Gregovich, D., Waples, R. S., Baker, C. S., Jackson, J., Taylor, B. L., . . . Schwartz, M. K. (2010). When are genetic methods useful for estimating contemporary abundance and detecting population trends? *Molecular Ecology Resources, 10*(4), 684-692. doi:10.1111/j.1755-0998.2010.02831.x

Wang, J. (2004). Sibship reconstruction from genetic data with typing errors. *Genetics, 166*(4), 1963-1979. doi:10.1534/genetics.166.4.1963

Waples, R. S. (1991). *Genetic methods for estimating the effective size of cetacean populations.* (Contract Report). Seattle, WA: Northwest Fisheries Center, National Marine Fisheries Service.

Waples, R. S. (2010). Spatial-temporal stratifications in natural populations and how they affect understanding and estimation of effective population size. *Molecular Ecology Resources, 10*(5), 785-796. doi:10.1111/j.1755-0998.2010.02876.x

Waples, R. S., & Do, C. (2008). LDNE: A program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources, 8*(4), 753-756. doi:10.1111/j.1755-0998.2007.02061.x

Waples, R. S., & Do, C. (2010). Linkage disequilibrium estimates of contemporary N-e using highly variable genetic markers: A largely untapped resource for applied conservation and evolution. *Evolutionary Applications, 3*(3), 244-262. doi:10.1111/j.1752-4571.2009.00104.x

Waples, R. S., & Yokota, M. (2007). Temporal estimates of effective population size in species with overlapping generations. *Genetics, 175*(1), 219-233. doi:10.1534/genetics.106.065300

Warton, D. I., & Hui, F. K. C. (2011). The arcsine is asinine: The analysis of proportions in ecology. *Ecology, 92*(1), 3-10. doi:10.1890/10-0340.1

Whiteley, A. R., Coombs, J. A., Hudy, M., Robinson, Z., Colton, A. R., Nislow, K. H., & Letcher, B. H. (2013). Fragmentation and patch size shape genetic structure of brook

trout populations. *Canadian Journal of Fisheries and Aquatic Sciences, 70*(5), 678-688. doi:10.1139/cjfas-2012-0493

Whiteley, A. R., Coombs, J. A., Hudy, M., Robinson, Z., Nislow, K. H., & Letcher, B. H. (2012). Sampling strategies for estimating brook trout effective population size. *Conservation Genetics, 13*(3), 625-637. doi:10.1007/s10592-011-0313-y

Wright, S. (1931). Evolution in mendelian populations. *Genetics, 16*(2), 0097-0159.

Xu, C., Letcher, B. H., & Nislow, K. H. (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.