Assessing the Impacts of Imperfect Detection in Stream Fish Communities through Multispecies Occupancy Modelling

by

David Benoit

A thesis submitted in conformity with the requirements for the degree of Master of Science

Ecology & Evolutionary Biology
University of Toronto

© Copyright by David Benoit 2017
Assessing the Impacts of Imperfect Detection in Stream Fish Communities through Multispecies Occupancy Modelling

David Benoit

Master of Science

Ecology & Evolutionary Biology
University of Toronto

2017

Abstract

Regardless of sampling effort, it is rare to detect all individuals or species in a given survey. This issue, more commonly known as imperfect detection, can have negative impacts on data quality and interpretation, most notably leading to false absences for rare or difficult-to-detect species. In this study, I set out to determine the impacts of imperfect detection on estimates of species richness and community structure in a stream fish assemblage. Multi-species occupancy modelling was used to estimate species-specific occurrence probabilities while accounting for imperfect detection, thus creating a more informed dataset. This dataset was then compared to the original to see where differences occurred. In my analyses, I demonstrated that imperfect detection can lead to large changes in estimates of species richness at the site level and summarized differences in the community structure and sampling locations, represented through correspondence analyses.
Acknowledgements

I would like to thank my supervisors, Professor Donald Jackson and Professor Mark Ridgway, who have inspired me to continue my studies in the field of aquatic ecology. Without your support and guidance, I would not have been able to complete this project. Don, your thoughtful mentorship and open-door policy were crucial in the development of these ideas. Mark, your feedback and interest in this work were always motivating and inspired me to ask critical questions. I would also like to thank the members of my advisory committee, Dr. Nick Mandrak and Dr. Marty Krkosek. Your questions and comments contributed greatly to this thesis.

To my office mates (Abby and Darren) and Jackson Lab members: Thank you for being great friends and colleagues. This experience would not have been the same without you and I am incredibly grateful to have worked alongside you. Your support and friendship kept me sane through all the ups and downs of grad school. To the Harkness Crew: Thanks for making my time in Algonquin Park such an unforgettable experience. Colin, without your help in the field, this project would not have been possible.

Special thanks to Gerald from the Maps and Data Library at U of T and Elise Zipkin from Michigan State University. Your guidance and advice were crucial to the completion of this project.

I would like to thank my family for their unwavering support and encouragement. Thank you for always pushing me to do my best and supporting me throughout the years. Without you, I would not be where I am today.
Contents

Abstract ............................................................................................................................................ ii
Acknowledgements ....................................................................................................................... iii
List of Tables ................................................................................................................................... v
List of Figures .................................................................................................................................. vi
List of Appendices ........................................................................................................................ vii
Introduction ..................................................................................................................................... 1
Methods .......................................................................................................................................... 3
  Study Area and Data Collection ................................................................................................... 3
  Addressing Assumptions: Independence among sites ................................................................. 5
  Multispecies Occupancy Modelling ............................................................................................. 7
  Data Analysis ............................................................................................................................... 8
Results ............................................................................................................................................. 9
  Occupancy Modelling .................................................................................................................. 9
  Species Richness .......................................................................................................................... 16
  Community Structure .................................................................................................................. 19
Discussion ....................................................................................................................................... 27
  Occupancy Modelling .................................................................................................................. 27
  Species Richness .......................................................................................................................... 29
  Community Structure .................................................................................................................. 29
  Limitations .................................................................................................................................. 30
  Significance .................................................................................................................................. 31
  Conclusions ................................................................................................................................. 32
References ...................................................................................................................................... 33
Appendix 1 .................................................................................................................................... 37
Appendix 2 .................................................................................................................................... 40
List of Tables

**Table 1**: Species observed during at least one of the two July surveys in Costello Creek. *Species removed from the modelling dataset*.................................10

**Table 2**: Occupancy probability values produced by the model for each species at each site.................12

**Table 3**: Mean detection probability values produced by the model for each species during the July 2009 and July 2015 surveys.................................15
List of Figures

**Figure 1:** Costello Creek with Habitat A (draining a bog area) depicted in red and Habitat B (clear, fast flowing) depicted in blue, with arrows representing the direction of flow. ......................................................... 4

**Figure 2:** Home range of the largest fish caught in the system surrounding each site. Home ranges in the upper stream are depicted in blue and ranges in the lower stream are depicted in green. .............................. 6

**Figure 3:** Number of times each total number of species was predicted to be present in Costello Creek by the occupancy model........................................................................................................... 11

**Figure 4:** Species richness at each site when comparing between the standard dataset and: a.) A 95% occupancy probability threshold; b.) An 85% occupancy probability threshold; c.) A 75% occupancy probability threshold; and, d.) A 50% occupancy probability threshold ................................................................. 18

**Figure 5:** Results of the correspondence analysis using: a.) The standard dataset of fish species presence-absence in Costello Creek; b.) A dataset informed at the 95% occupancy threshold; c.) A dataset informed at the 85% occupancy threshold; d.) A dataset informed at the 75% occupancy threshold; and, e.) A dataset informed at the 50% occupancy threshold ............................................................... 22

**Figure 6:** Procrustes residuals from analysis comparing the correspondence analysis results from the standard dataset to one informed at: a.) The 95% occupancy threshold for both species and sites; b.) The 85% occupancy threshold for both species and sites; c.) The 75% occupancy threshold for both species and sites; and, d.) The 50% occupancy threshold for both species and sites. ................................................................. 26
List of Appendices

Appendix 1.................................................................................................................................................37

Figure 1: Custom minnow traps used to collect samples. Traps were rectangular in shape, had three conical entrances, and were baited with a handful of dry dog food before each use. .......................... 37

Table 1: Catch data for the July 2015 survey of Costello Creek. Catch counts were calculated by summing the total number of individuals caught in each of the three temporal replicates. .........................38

Table 2: Catch data for the July 2009 survey of Costello Creek. Catch counts were calculated by summing the total number of individuals caught in each of the three temporal replicates. .........................39

Appendix 2 (R code)...................................................................................................................................40
Introduction

Regardless of sampling effort, it is rare to detect all individuals or species in a given survey (Iknayan et al., 2014; Royle et al., 2005). This is true whether the target organism is a species of insect (Dorazio et al., 2011), bird (Ruiz-Gutierrez et al., 2010), mammal (Burton et al., 2012), or fish (Jackson & Harvey, 1997). This issue, more commonly known as imperfect detection, is caused by a number of factors. Variability in abundance is arguably the largest driver of variability in detection, as abundant species are more likely to be encountered within a survey than those that are less common (Royle & Nichols, 2003; Royle et al., 2005). Variability in detection is also heavily influenced by differences among species, as well as differences among individuals of the same species. For example, traits such as size and colour may affect the ability of an observer to detect an organism in its natural habitat (Boulinier et al., 1998). Additionally, variation in behaviour within a species may make some individuals easier to detect than others. Site- and survey-specific factors have also been shown to influence detectability (Iknayan et al., 2014). For instance, habitat structure at a survey site or inclement weather may impede an observer’s ability to detect all present individuals. Rates of detection can also be influenced by variation among observers, with less-skilled observers being more likely to miss or falsely identify an organism (McClintock et al., 2010). With such a diverse set of causes, it can easily be understood that imperfect detection is prevalent among ecological studies. Unfortunately, this issue is often left unaddressed.

Ignoring imperfect detection can have a number of negative impacts, most notably on data quality and interpretation. When species are abundant, errors in detection can lead to underestimation of population and range size (Iknayan et al., 2014). This can have serious consequences for those interested in population management, including fisheries and wildlife managers working with economically important species. When species are rare and characterized by low population sizes, imperfect detection leads to false absences and can, ultimately, reduce the accuracy of distribution models and diversity estimates, such as species richness (Dorazio et al., 2011). This is especially concerning when considering
the importance of these estimates. Species richness estimates are not only used in the development of novel ecological theory (Dorazio et al., 2006), but are also frequently used as a variable on which conservation and management decisions can be based (Yoccoz et al., 2001). It is therefore crucial that the accuracy of these estimates be addressed.

A number of statistical methods have been used in the past to estimate species richness while accounting for errors in detection. The most popular of these approaches are the Jackknife and Chao estimators (Iknayan et al., 2014). Despite their popularity, these approaches have been criticized for confounding occurrence and detection and, ultimately, estimating only the apparent occurrence of species (Kery, 2010). This criticism is due to their failure to address all factors contributing to variation in detectability. A number of previous studies have shown that these estimators underestimate species richness at low sampling effort and overestimate richness at high effort, due to their strong dependence on sample size, highlighting their inability to provide accurate estimates of species richness (Melo, 2004; Wagner & Wildi, 2002; Hellmaan & Fowler, 1999). An alternative approach that has been put forth to tackle this issue and fully address imperfect detection is that of multispecies occupancy modelling. This form of modelling, developed by Dorazio et al. (2005), can be used to estimate species-specific occurrence probabilities, while accounting for variability in detection from numerous sources (MacKenzie et al., 2006; Zipkin et al., 2009). To use this approach, sites must be sampled repeatedly and the duration of the survey must be kept short enough that species richness can be assumed constant (Dorazio & Royle, 2005). Other major assumptions for this form of modelling are that sample sites remain independent, implying that the detection of a species at one site is independent of detecting that species at other sites and that all species are correctly identified (MacKenzie et al., 2006). One of the major strengths of this form of modelling is its ability to consider rare or difficult-to-detect species that may otherwise be ignored due to limited data. This is made possible through a Bayesian technique known as “borrowing strength”, whereby, parameter estimates for each species are drawn from a common, community-level distribution, thus, allowing for more precise estimates for rare species (Broms et al., 2016; Zipkin et al.,
such a community approach can also provide benefits to conservationists and resource managers, as it is much more feasible than conducting numerous single-species assessments due to reduced costs and time requirements (Dewan & Zipkin, 2010).

Despite their potential, multispecies occupancy models have seen limited use. This is particularly true for freshwater systems, where previous applications have focused on estimating sampling effort (Holtrop et al., 2010) and developing species-habitat relationships (Midway et al., 2014; Kirsch & Peterson, 2014). Because of this lack of consideration, the impacts of imperfect detection on our understanding of freshwater diversity and ecology remain largely unknown. This is of high concern considering the threatened state of many freshwater ecosystems. Due to a combination of stressors, including habitat degradation, overexploitation, and the presence of invasive species, freshwater fauna have been predicted to experience extinction rates much higher than their terrestrial counterparts (Ricciardi, 1999). To better conserve these species and the integrity of freshwater ecosystems, it is crucial that we better comprehend how variability in detection may alter our understanding of these communities. In this study, I set out to determine the impacts of imperfect detection on estimates of diversity and community structure in an assemblage of freshwater fishes.

Methods

Study Area and Data Collection

Presence-absence data of fish communities in Algonquin Provincial Park were used as an exemplar of aquatic communities. Data collection took place in May and July of 2009, and July of 2015, in Costello Creek. Sampling consisted of baiting and setting custom triple-entry minnow traps (See image in Appendix 1) to be left overnight. The following morning, the traps were retrieved and the captured species were identified, allowing for both abundance and presence-absence data to be recorded. Species that proved difficult to identify were brought back to the lab for identification. A total of 36 sites were sampled in this manner. Sites were randomly chosen along the creek edge and kept constant across all
three surveys. Sampling was conducted using a triple-pass system, wherein, each site was sampled three times within a survey period to allow for the calculation of detection probabilities. The creek was divided into two distinct habitat types, denoted as A and B for ease in modelling. Habitat A was largely made up of a spruce bog with turbid, slow-moving water. Sites in Habitat B were characterized by more habitat heterogeneity, but generally exhibited clear, fast-flowing water (Figure 1).

Figure 1: Costello Creek with Habitat A (draining a bog area) depicted in red and Habitat B (clear, fast flowing) depicted in blue, with arrows representing the direction of flow.
Addressing Assumptions: Independence among sites

As previously mentioned, one of the assumptions concerning occupancy modelling is that sites remain independent. Because of the large spatial component inherent in this style of sampling, it was important to test for independence among sites. This was done by estimating the home range of the largest fish caught, approximately 150 mm, using an equation described by Minns (1995).

\[
\log_e H = -2.91 + 1.65 \log_e L
\]

H = Home range (m²)

L = Body length (mm)

Once the home range area was calculated, it was converted into a linear distance according to creek width. Width values were determined using the measuring tool on Google Maps. The creek was divided into upper and lower segments, as the width seemed to remain relatively constant in each segment. The average width of the upper and lower segments were 5 m and 22 m, respectively. This led to a linear home range of 42.44 m in the upper stream and 9.6 m in the lower stream. To increase the probability that sites retained their independence, the full home range was mapped on each side of the sampling points. All mapping was done in ArcMAP 10.3.1 using the Network Analyst Service Area function (ESRI, 2011; Figure 2). Where overlap occurred between two sites, the second site was removed. This was done to standardize the removal process and keep the samples random. This led to the removal of 5 sites, bringing the total number of sites to 31. Assessing site independence in such a manner also led to the removal of one species from the 2015 dataset, as the Brassy Minnow (*Hybognathus hankinsoni*) was only found at removed sites.
Figure 2: Home range of the largest fish caught in the system surrounding each site. Home ranges in the upper stream are depicted in blue and ranges in the lower stream are depicted in green.
Multispecies Occupancy Modelling

The hierarchical community-model framework used in this study was developed by Dorazio and Royle (2005) and includes data augmentation, a process that allows for completely unobserved species to be considered in the analysis. This is achieved by adding a fixed number of all-zero encounter histories to the dataset, creating a zero-inflated version of a model where the actual number of species in the community is known. A uniform (0, M) prior for N, the true number of species present, is assumed, where M represents the sum of observed species and the all-zero encounter histories (Zipkin et al., 2010). Code made available by Zipkin et al (2010) was used as a base for modelling and modified to suit the study system. To promote model convergence, data from different surveys needed to be pooled. Data from July 2009 and July 2015 were combined to create one dataset. Despite a gap of several years between surveys, I assumed that there was a quasi-equilibrium of species at the various sites from Costello Creek. Data from the May 2009 survey were left out, as their inclusion was recognized to create a seasonal bias in the data. Additionally, any species that were not present in both surveys were removed from the dataset, as their inclusion would have led to underestimations of detection probabilities. Finescale Dace (*Chrosomus neogaeus*), detected at a total of five sites in July 2009, was removed as it was not detected in the July 2015 survey. The final dataset to be used for modelling, thus, consisted of 12 species and six sampling replicates at each of the 31 sites.

Modelling was conducted under the assumption that the occurrence ($\Psi_{i,j}$) and detection ($p_{j,k,i}$) probabilities varied among species and were influenced by habitat characteristics and survey year, respectively. The occupancy probabilities of each species were modelled dependent on habitat type (A or B) and creek width on the logit-probability scale, as such:

$$\text{logit} (\Psi_{i,j}) = u.A[i] \times (1 - \text{Ind}[j]) + u.B[i] \times \text{Ind}[j] + a1[i] \times \text{width1}[j]$$

In this scenario, $u.A[i]$ and $u.B[i]$ represent occurrence probabilities for species $i$ at points $j$ in each of the habitat types. This separation is made possible through the use of the indicator function $\text{Ind}[j]$, which is
used to determine what habitat type each of the $j$ sites corresponds to. The coefficient $a_1[i]$ represents the linear effect of creek width on the occurrence of species $i$ and $width[j]$ is a vector containing standardized width values for each of the $j$ sampled sites. Detection was modelled with no covariates, yet still incorporated the year of survey into the analysis:

$$\logit(p_{j,k,i}) = v.A[i] \times (1 - Year[k]) + v.B[i] \times Year[k]$$

In the above equation, $v.A[i]$ and $v.B[i]$ reflect the detection probability of species $i$ in the $k$th survey replicate in year 2015 and 2009, respectively. This separation is made possible by the indicator function $Year[k]$, which is used to determine which year each of the $k$ survey replicates corresponds to. Modelling occurrence and detection in such a manner allowed for the calculation of occupancy probabilities for each species at each site (Refer to R script in Appendix 2). Bayesian shrinkage was incorporated into the model by drawing all species parameters from common, community-level distributions characterized by uninformative priors (Zipkin et al., 2009). All analyses were performed using the programs R (R Core Team, 2015) and WinBUGS (Spiegelhalter et al., 2003). Model convergence was assessed using the R-hat statistic (Gelman & Hill, 2007).

**Data Analysis**

To assess the impacts of imperfect detection, the original dataset was compared to one that took imperfect detection into account. Occupancy probability values for each species at each site were used to create this more informed dataset. This was done by selecting a probability threshold of 95% and comparing it with probability values produced by the model. Any lack of detection for a species at a site with an occupancy probability of 95% or higher were thus considered false absences. These absences, represented by zeroes in the dataset, would then be converted to ones to mark occurrence. Species richness at the site level was then compared between the two datasets. This was done by comparing the total number of species observed at each site in the original dataset with the total number of species estimated to be present at each site in the informed dataset.
Correspondence analyses were conducted on each dataset to assess community structure (Greenacre, 2007). The two ordinations were then compared using a resistant-fit Procrustes analysis. Resistant-fit methods were chosen over the ordinary least-squares procedures, as they are less likely to show misleading representations of shape differences when one or a few landmarks display large changes in position (Claude, 2008; Marcus et al., 2013). Procrustes residuals were plotted to determine which sites and species exhibited the most change when imperfect detection was taken into account. This process was then repeated for a number of arbitrarily chosen probability thresholds (50%, 75%, and 85%). All analyses were performed using the program R (R Core Team, 2015).

Results

*Occupancy Modelling*

As previously mentioned, a total of 12 species were included in the dataset used for modelling. The included species spanned a number of families, although the majority were in the family Cyprinidae (Table 1). Catch data for each species at each site can be found for both the July 2009 and July 2015 surveys in Appendix 1 (Tables 4-5).

Multispecies occupancy modelling predicted a mean of 12.273 (Standard deviation = 0.623) species in the system across all posterior estimates, with a maximum of 23 (Figure 4). Occupancy probability values produced by the model varied greatly among species and, to a lesser degree, within species, with values ranging from 8.39x10⁻⁵ to 0.999 (Table 2). The mean probability of detection across all sites varied greatly among species, ranging from 0.066 to 0.978 (Table 3). Variation in detection among species was generally much larger than the variation observed between survey years; however, some species did exhibit large changes in detection values between the two years. The largest change was seen in the Northern Redbelly Dace, where the probability of detection varied by almost 40% between 2009 and 2015.
**Table 1:** Species observed during at least one of the two July surveys in Costello Creek. *Species removed from the modelling dataset.

<table>
<thead>
<tr>
<th>Species</th>
<th>Scientific Name</th>
<th>Family</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Sucker</td>
<td><em>Catostomus commersonii</em></td>
<td>Catostomidae</td>
</tr>
<tr>
<td>Pumpkinseed</td>
<td><em>Lepomis gibbosus</em></td>
<td>Centrarchidae</td>
</tr>
<tr>
<td>Smallmouth Bass</td>
<td><em>Micropterus dolomieu</em></td>
<td>Centrarchidae</td>
</tr>
<tr>
<td>Northern Redbelly Dace</td>
<td><em>Chrosomus eos</em></td>
<td>Cyprinidae</td>
</tr>
<tr>
<td>Finescale Dace*</td>
<td><em>Chrosomus neogaeus</em></td>
<td>Cyprinidae</td>
</tr>
<tr>
<td>Brassy Minnow*</td>
<td><em>Hybognathus hankinsoni</em></td>
<td>Cyprinidae</td>
</tr>
<tr>
<td>Northern Pearl Dace</td>
<td><em>Margariscus nachtriebi</em></td>
<td>Cyprinidae</td>
</tr>
<tr>
<td>Common Shiner</td>
<td><em>Luxilus cornutus</em></td>
<td>Cyprinidae</td>
</tr>
<tr>
<td>Golden Shiner</td>
<td><em>Notemigonus crysoleucas</em></td>
<td>Cyprinidae</td>
</tr>
<tr>
<td>Creek Chub</td>
<td><em>Semotilus atromaculatus</em></td>
<td>Cyprinidae</td>
</tr>
<tr>
<td>Brook Stickleback</td>
<td><em>Culaea inconstans</em></td>
<td>Gasterosteidae</td>
</tr>
<tr>
<td>Brown Bullhead</td>
<td><em>Ameiurus nebulosus</em></td>
<td>Ictaluridae</td>
</tr>
<tr>
<td>Yellow Perch</td>
<td><em>Perca flavescens</em></td>
<td>Percidae</td>
</tr>
<tr>
<td>Brook Trout</td>
<td><em>Salvelinus fontinalis</em></td>
<td>Salmonidae</td>
</tr>
</tbody>
</table>
Figure 3: Number of times each total number of species was predicted to be present in Costello Creek by the occupancy model.
Table 2: Occupancy probability values produced by the model for each species at each site.

<table>
<thead>
<tr>
<th>Site</th>
<th>Species</th>
<th>Brown Bullhead</th>
<th>Brook Stickleback</th>
<th>Brook Trout</th>
<th>Creek Chub</th>
<th>Common Shiner</th>
<th>Golden Shiner</th>
<th>Northern Redbelly Dace</th>
<th>Pumpkinseeds</th>
<th>Northern Pearl Dace</th>
<th>Smallmouth Bass</th>
<th>White Sucker</th>
<th>Yellow Perch</th>
</tr>
</thead>
<tbody>
<tr>
<td>A01</td>
<td>0.9999782</td>
<td>0.1786021</td>
<td>8.39E-05</td>
<td>0.9999782</td>
<td>0.9864971</td>
<td>0.9999719</td>
<td>0.4422729</td>
<td>0.9999786</td>
<td>0.3808691</td>
<td>8.21E-05</td>
<td>0.9995903</td>
<td>0.9999789</td>
<td></td>
</tr>
<tr>
<td>A02</td>
<td>0.9999504</td>
<td>0.1703105</td>
<td>6.11E-05</td>
<td>0.9999552</td>
<td>0.9682768</td>
<td>0.9999787</td>
<td>0.2512965</td>
<td>0.9999563</td>
<td>0.16387104</td>
<td>3.46E-05</td>
<td>0.9991437</td>
<td>0.9999567</td>
<td></td>
</tr>
<tr>
<td>A03</td>
<td>0.9999484</td>
<td>0.1699329</td>
<td>6.02E-05</td>
<td>0.9999537</td>
<td>0.9670076</td>
<td>0.999979</td>
<td>0.2438573</td>
<td>0.9999548</td>
<td>0.15672024</td>
<td>3.32E-05</td>
<td>0.9991138</td>
<td>0.9999552</td>
<td></td>
</tr>
<tr>
<td>A04</td>
<td>0.9999562</td>
<td>0.1715635</td>
<td>6.42E-05</td>
<td>0.9999599</td>
<td>0.9721448</td>
<td>0.9999778</td>
<td>0.2769395</td>
<td>0.9999608</td>
<td>0.18937629</td>
<td>3.95E-05</td>
<td>0.9992353</td>
<td>0.9999612</td>
<td></td>
</tr>
<tr>
<td>A05</td>
<td>0.9999543</td>
<td>0.1711204</td>
<td>6.31E-05</td>
<td>0.9999583</td>
<td>0.970835</td>
<td>0.9999781</td>
<td>0.267707</td>
<td>0.9999593</td>
<td>0.1800429</td>
<td>3.77E-05</td>
<td>0.9992041</td>
<td>0.9999597</td>
<td></td>
</tr>
<tr>
<td>A06</td>
<td>0.999945</td>
<td>0.1692946</td>
<td>5.87E-05</td>
<td>0.999951</td>
<td>0.9647442</td>
<td>0.9999794</td>
<td>0.2315964</td>
<td>0.9999522</td>
<td>0.14518604</td>
<td>3.10E-05</td>
<td>0.9990609</td>
<td>0.9999526</td>
<td></td>
</tr>
<tr>
<td>A07</td>
<td>0.9998814</td>
<td>0.1618435</td>
<td>4.37E-05</td>
<td>0.9999038</td>
<td>0.923553</td>
<td>0.9999841</td>
<td>0.1187076</td>
<td>0.9999068</td>
<td>0.05497205</td>
<td>1.38E-05</td>
<td>0.9981276</td>
<td>0.9999073</td>
<td></td>
</tr>
<tr>
<td>A08</td>
<td>0.9999356</td>
<td>0.1677434</td>
<td>5.53E-05</td>
<td>0.9999437</td>
<td>0.9585698</td>
<td>0.9999805</td>
<td>0.2034911</td>
<td>0.9999452</td>
<td>0.11996298</td>
<td>2.63E-05</td>
<td>0.98918</td>
<td>0.9999456</td>
<td></td>
</tr>
<tr>
<td>A09</td>
<td>0.9999711</td>
<td>0.1757534</td>
<td>7.53E-05</td>
<td>0.9999722</td>
<td>0.9819265</td>
<td>0.9999744</td>
<td>0.3719854</td>
<td>0.9999727</td>
<td>0.29441657</td>
<td>6.12E-05</td>
<td>0.9994738</td>
<td>0.999973</td>
<td></td>
</tr>
<tr>
<td>A10</td>
<td>0.9999799</td>
<td>0.1794538</td>
<td>8.66E-05</td>
<td>0.9999797</td>
<td>0.9876185</td>
<td>0.9999711</td>
<td>0.4637111</td>
<td>0.99998</td>
<td>0.40836251</td>
<td>8.96E-05</td>
<td>0.9996196</td>
<td>0.9999803</td>
<td></td>
</tr>
<tr>
<td>A11</td>
<td>0.9999744</td>
<td>0.1769621</td>
<td>7.89E-05</td>
<td>0.9999749</td>
<td>0.9840341</td>
<td>0.9999734</td>
<td>0.4014425</td>
<td>0.9999754</td>
<td>0.32988048</td>
<td>6.94E-05</td>
<td>0.999527</td>
<td>0.9999757</td>
<td></td>
</tr>
<tr>
<td>A12</td>
<td>0.9999282</td>
<td>0.1666777</td>
<td>5.30E-05</td>
<td>0.9999381</td>
<td>0.9537101</td>
<td>0.9999812</td>
<td>0.1855964</td>
<td>0.9999397</td>
<td>0.10481804</td>
<td>2.34E-05</td>
<td>0.9988067</td>
<td>0.9999402</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A13</td>
<td>0.9999894</td>
<td>0.1861346</td>
<td>1.11E-04</td>
<td>0.9999884</td>
<td>0.993676</td>
<td>0.9999641</td>
<td>0.6277166</td>
<td>0.9999885</td>
<td>0.62662576</td>
<td>1.75E-04</td>
<td>0.9997855</td>
<td>0.9999887</td>
<td></td>
</tr>
<tr>
<td>A15</td>
<td>0.999992</td>
<td>0.1861346</td>
<td>1.24E-04</td>
<td>0.999991</td>
<td>0.9953106</td>
<td>0.9999605</td>
<td>0.6939469</td>
<td>0.999991</td>
<td>0.7133766</td>
<td>2.36E-04</td>
<td>0.9998336</td>
<td>0.9999912</td>
<td></td>
</tr>
<tr>
<td>A16</td>
<td>0.9999926</td>
<td>0.1900146</td>
<td>1.27E-04</td>
<td>0.9999916</td>
<td>0.9956887</td>
<td>0.9999595</td>
<td>0.7113275</td>
<td>0.9999916</td>
<td>0.73546604</td>
<td>2.57E-04</td>
<td>0.9998451</td>
<td>0.9999918</td>
<td></td>
</tr>
<tr>
<td>A18</td>
<td>0.9999929</td>
<td>0.1904953</td>
<td>1.30E-04</td>
<td>0.9999919</td>
<td>0.9958871</td>
<td>0.9999588</td>
<td>0.7208001</td>
<td>0.9999919</td>
<td>0.74734908</td>
<td>2.69E-04</td>
<td>0.9998511</td>
<td>0.9999921</td>
<td></td>
</tr>
<tr>
<td>B01</td>
<td>0.9437737</td>
<td>0.6985026</td>
<td>4.68E-01</td>
<td>0.9876718</td>
<td>0.9430085</td>
<td>0.9430085</td>
<td>0.6800479</td>
<td>0.9875235</td>
<td>0.774325</td>
<td>3.15E-01</td>
<td>0.9173124</td>
<td>0.987777</td>
<td></td>
</tr>
<tr>
<td>B02</td>
<td>0.952069</td>
<td>0.7009846</td>
<td>4.85E-01</td>
<td>0.9893478</td>
<td>0.9519026</td>
<td>0.9519026</td>
<td>0.717153</td>
<td>0.9892024</td>
<td>0.8126874</td>
<td>3.54E-01</td>
<td>0.9280903</td>
<td>0.9894246</td>
<td></td>
</tr>
<tr>
<td>B03</td>
<td>0.9665176</td>
<td>0.7064528</td>
<td>5.21E-01</td>
<td>0.9923051</td>
<td>0.9671618</td>
<td>0.9671618</td>
<td>0.789513</td>
<td>0.9921722</td>
<td>0.8795917</td>
<td>4.48E-01</td>
<td>0.9475346</td>
<td>0.9923474</td>
<td></td>
</tr>
<tr>
<td>B04</td>
<td>0.9679274</td>
<td>0.7070995</td>
<td>5.25E-01</td>
<td>0.9925976</td>
<td>0.9686313</td>
<td>0.9686313</td>
<td>0.7971517</td>
<td>0.9924665</td>
<td>0.886004</td>
<td>4.60E-01</td>
<td>0.9494868</td>
<td>0.9926368</td>
<td></td>
</tr>
<tr>
<td>B05</td>
<td>0.9247824</td>
<td>0.6938825</td>
<td>4.39E-01</td>
<td>0.983859</td>
<td>0.9223762</td>
<td>0.9223762</td>
<td>0.6053567</td>
<td>0.9837131</td>
<td>0.6897549</td>
<td>2.48E-01</td>
<td>0.8934565</td>
<td>0.9840101</td>
<td></td>
</tr>
<tr>
<td>B07</td>
<td>0.9691534</td>
<td>0.7076839</td>
<td>5.29E-01</td>
<td>0.9928527</td>
<td>0.9699058</td>
<td>0.9699058</td>
<td>0.8038831</td>
<td>0.9927234</td>
<td>0.8915472</td>
<td>4.70E-01</td>
<td>0.9511936</td>
<td>0.9928892</td>
<td></td>
</tr>
<tr>
<td>B08</td>
<td>0.9496448</td>
<td>0.7002209</td>
<td>4.80E-01</td>
<td>0.9888569</td>
<td>0.9493122</td>
<td>0.9493122</td>
<td>0.7059975</td>
<td>0.9887103</td>
<td>0.8014286</td>
<td>3.42E-01</td>
<td>0.9249149</td>
<td>0.9889398</td>
<td></td>
</tr>
<tr>
<td>B09</td>
<td>0.9433791</td>
<td>0.698393</td>
<td>4.68E-01</td>
<td>0.9875923</td>
<td>0.9425834</td>
<td>0.9425834</td>
<td>0.6783559</td>
<td>0.9874439</td>
<td>0.7725145</td>
<td>3.13E-01</td>
<td>0.9168057</td>
<td>0.9876916</td>
<td></td>
</tr>
<tr>
<td>B10</td>
<td>0.9686107</td>
<td>0.7074225</td>
<td>5.27E-01</td>
<td>0.9927397</td>
<td>0.969342</td>
<td>0.969342</td>
<td>0.8008928</td>
<td>0.9926096</td>
<td>0.8890972</td>
<td>4.66E-01</td>
<td>0.9504369</td>
<td>0.9927774</td>
<td></td>
</tr>
<tr>
<td>B11</td>
<td>0.9700276</td>
<td>0.708114</td>
<td>5.32E-01</td>
<td>0.9930351</td>
<td>0.9708127</td>
<td>0.9708127</td>
<td>0.8087341</td>
<td>0.9929071</td>
<td>0.8954789</td>
<td>4.78E-01</td>
<td>0.952416</td>
<td>0.9930697</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B13</td>
<td>0.9232209</td>
<td>0.693551</td>
<td>4.37E-01</td>
<td>0.983546</td>
<td>0.9206659</td>
<td>0.5919139</td>
<td>0.5997776</td>
<td>0.9834007</td>
<td>0.6830836</td>
<td>2.44E-01</td>
<td>0.8915388</td>
<td>0.9837017</td>
<td></td>
</tr>
<tr>
<td>B15</td>
<td>0.8593947</td>
<td>0.6833689</td>
<td>3.74E-01</td>
<td>0.9705881</td>
<td>0.8496431</td>
<td>0.6457984</td>
<td>0.4244937</td>
<td>0.9705165</td>
<td>0.4563273</td>
<td>1.37E-01</td>
<td>0.8173274</td>
<td>0.9709525</td>
<td></td>
</tr>
<tr>
<td>B16</td>
<td>0.9540118</td>
<td>0.7016227</td>
<td>4.89E-01</td>
<td>0.9897422</td>
<td>0.953973</td>
<td>0.5467461</td>
<td>0.7262874</td>
<td>0.9895978</td>
<td>0.8217253</td>
<td>3.65E-01</td>
<td>0.9306516</td>
<td>0.9898141</td>
<td></td>
</tr>
<tr>
<td>B17</td>
<td>0.9700276</td>
<td>0.708114</td>
<td>5.32E-01</td>
<td>0.9930351</td>
<td>0.9708127</td>
<td>0.5092958</td>
<td>0.8087341</td>
<td>0.9929071</td>
<td>0.8954789</td>
<td>4.78E-01</td>
<td>0.952416</td>
<td>0.9930697</td>
<td></td>
</tr>
<tr>
<td>B18</td>
<td>0.9232209</td>
<td>0.693551</td>
<td>4.37E-01</td>
<td>0.983546</td>
<td>0.9206659</td>
<td>0.5919139</td>
<td>0.5997776</td>
<td>0.9834007</td>
<td>0.6830836</td>
<td>2.44E-01</td>
<td>0.8915388</td>
<td>0.9837107</td>
<td></td>
</tr>
</tbody>
</table>
Table 3: Mean detection probability values produced by the model for each species during the July 2009 and July 2015 surveys.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Brown Bullhead</td>
<td>0.8044464</td>
<td>0.0412245</td>
<td>0.6209534</td>
<td>0.0505856</td>
<td>0.1834930</td>
</tr>
<tr>
<td>Brook Stickleback</td>
<td>0.0663796</td>
<td>0.0558128</td>
<td>0.2320031</td>
<td>0.1319449</td>
<td>-0.1656235</td>
</tr>
<tr>
<td>Brook Trout</td>
<td>0.2296301</td>
<td>0.1451114</td>
<td>0.1753991</td>
<td>0.1238113</td>
<td>0.0542310</td>
</tr>
<tr>
<td>Creek Chub</td>
<td>0.9050042</td>
<td>0.0300042</td>
<td>0.9775264</td>
<td>0.0151700</td>
<td>-0.0725222</td>
</tr>
<tr>
<td>Common Shiner</td>
<td>0.7296776</td>
<td>0.0470803</td>
<td>0.7766607</td>
<td>0.0443122</td>
<td>-0.0469831</td>
</tr>
<tr>
<td>Golden Shiner</td>
<td>0.7436326</td>
<td>0.0507872</td>
<td>0.8139368</td>
<td>0.0453049</td>
<td>-0.0703042</td>
</tr>
<tr>
<td>Northern Redbelly Dace</td>
<td>0.1746583</td>
<td>0.0552348</td>
<td>0.5588116</td>
<td>0.0821313</td>
<td>-0.3841533</td>
</tr>
<tr>
<td>Pumpkinseed</td>
<td>0.4940624</td>
<td>0.0513037</td>
<td>0.5693613</td>
<td>0.0508198</td>
<td>-0.0752989</td>
</tr>
<tr>
<td>Northern Pearl Dace</td>
<td>0.2926874</td>
<td>0.0713019</td>
<td>0.3346882</td>
<td>0.0761511</td>
<td>-0.0420008</td>
</tr>
<tr>
<td>Smallmouth Bass</td>
<td>0.1521381</td>
<td>0.1297467</td>
<td>0.2412743</td>
<td>0.1743811</td>
<td>-0.0891362</td>
</tr>
<tr>
<td>White Sucker</td>
<td>0.1442313</td>
<td>0.0435121</td>
<td>0.1088084</td>
<td>0.0375968</td>
<td>0.0354229</td>
</tr>
<tr>
<td>Yellow Perch</td>
<td>0.7477128</td>
<td>0.0443271</td>
<td>0.8655151</td>
<td>0.0348345</td>
<td>-0.1178023</td>
</tr>
</tbody>
</table>
Species Richness

Estimates of species richness at the site level underwent changes at all threshold comparisons. When the standard dataset was compared to the 95% threshold, estimated species richness increased by one at 11 of the 31 sites, with most changes occurring in sites found in Habitat A, a bog habitat (Figure 4a). At the 85% threshold, estimates of species richness increased by one at 18 sites and by two at one site. In this scenario, increases in species richness appeared to be balanced between the two habitat types (Figure 4b). When the standard dataset was compared to the 75% threshold, an increasing number of sites displayed richness increases of two species, with a total of 20 sites experiencing augmentation (Figure 4c). In this scenario, increases in richness were once again balanced between the two habitat types. At the 50% threshold, 24 of the 31 sites experienced changes in species richness. Increases in richness varied across sites; however, all sites in Habitat B displayed changes in richness and at much higher levels than at sites in Habitat A. The largest change in richness was an addition of 5 species (Figure 4d).

a.)
d.) Figure 4: Species richness at each site when comparing between the standard dataset and: a.) A 95% occupancy probability threshold; b.) An 85% occupancy probability threshold; c.) A 75% occupancy probability threshold; and, d.) A 50% occupancy probability threshold.
Community Structure

Community structure, summarized through correspondence analyses, also underwent changes in each threshold scenario. When comparing the standard ordination to one conducted at the 95% threshold, small changes were seen in the position of White Sucker and Northern Redbelly Dace (Figure 5a & Figure 5b). When comparing the standard ordination to one conducted at the 85% threshold, changes became much more noticeable, with Smallmouth Bass and Brook Stickleback displaying large changes of position (Figure 5a & Figure 5c). The same species appeared to show the largest changes when comparing the standard ordination to one conducted at the 75% threshold (Figure 5a & Figure 5d). When comparing the standard ordination to one conducted at the 50% threshold, Brook Trout appeared to show the most obvious change in position (Figure 5a & Figure 5e). The total variation explained by dimensions 1 and 2 differed greatly across the threshold scenarios. Using the standard dataset (i.e. no occupancy-modelling adjustment), a total of 49.9% of the variation was explained by these two dimensions. In contrast, a total of 76.2% of the variation was explained by dimensions 1 and 2 when analyzing the dataset informed at the 50% occupancy probability threshold.

The species and sites responsible for these changes became evident through resistant-fit Procrustes analyses (Figure 6). When comparing the correspondence analyses of the standard and 95% informed datasets (Figure 6a), the largest changes were seen in Brook Trout and White Sucker. A number of sites also demonstrated large changes. When comparing the standard dataset to the 85% informed dataset (Figure 6b), the largest changes were seen in Smallmouth Bass and White Sucker. The relative proportion of change among sites remained relatively constant; however, the residual vector for site B17 was considerably longer than found for all other sites. When comparing the standard dataset to the 75% informed dataset (Figure 6c), the largest change among species was seen in the Brook Stickleback. Further, there appeared to be larger changes at sites located in Habitat B, with site B17 again expressing the largest change. This trend continued when comparing the standard dataset to the 50% informed dataset (Figure 6d); however, in this scenario, Smallmouth Bass showed the largest change.
Figure 5: Results of the correspondence analysis using: a.) The standard dataset of fish species presence-absence in Costello Creek; b.) A dataset informed at the 95% occupancy threshold; c.) A dataset informed at the 85% occupancy threshold; d.) A dataset informed at the 75% occupancy threshold; and, e.) A dataset informed at the 50% occupancy threshold.
Procrustes: Standard vs 95%
Figure 6: Procrustes residuals from analysis comparing the correspondence analysis results from the standard dataset to one informed at: a.) The 95% occupancy threshold for both species and sites; b.) The 85% occupancy threshold for both species and sites; c.) The 75% occupancy threshold for both species and sites; and, d.) The 50% occupancy threshold for both species and sites.
Discussion

Although imperfect detection is prevalent among ecological studies, its impacts have rarely been considered. In this study, I set out to address this issue using data from a stream fish assemblage in Costello Creek, Algonquin Provincial Park. My results, obtained through multispecies occupancy model use, demonstrate that imperfect detection can impact estimates of species richness and community structure in aquatic systems. This was found to be true at all threshold levels, highlighting the need to consider variability in detection in ecological studies. Additionally, this study illustrates the benefits of multispecies occupancy model application in aquatic community ecology, as it highlights the ability of these models to determine species-habitat relationships and consider species that may otherwise be ignored due to limited data.

Occupancy Modelling

The multispecies occupancy model used in this study predicted a mean number of species equivalent to that included in the dataset, despite using a data augmentation framework. This may be unexpected, considering that not all observed species were included in the dataset. However, these results are consistent with other studies focusing on communities characterized by a number of relatively common species. In such communities, repeated sampling can substantially increase the chances of detecting most species, regardless of detection probabilities (Dorazio et al., 2006). Further evidence for this idea is provided by the high occupancy probabilities of most species in Costello Creek (Table 2). Although a number of species displayed relatively constant occupancy probability values throughout the entirety of the creek, approximately 50% showed high variation in occupancy across the habitat types. In general, occupancy probability values tended to be much higher in sites characterized by clear, fast-flowing water (Habitat B). This trend was most noticeable for small-bodied species, including Northern Redbelly Dace, Pearl Dace, and Brook Stickleback. The same was also true for Brook Trout, a species known to prefer clear, oxygenated streams and lakes (Scott & Crossman, 1973). These findings highlight
the importance of habitat and environmental factors in structuring stream fish communities, and the ability of multispecies occupancy models to determine these relationships. These ideas have been discussed in further detail in previous studies (Kirsch & Peterson, 2014). Variation in occurrence may have also been driven by interactions between species, such as competition or predation; however, for simplicity, modelling was conducted under the assumption that species occurrences were independent of each other. As seen in Table 3, detection probabilities estimated by the model varied drastically among species. These findings, which exemplify the species-specific nature of detectability, are consistent with previous work done on reef fish, where it was shown that detection probabilities varied substantially among different species and family groups (MacNeil et al., 2008). Some of this variation is likely attributed to differences in abundance, as commonly observed species in the system tended to have higher probabilities of detection. This trend is to be expected, as both probabilities of occurrence and detection are expected to increase as the abundance of a species increases (Dorazio & Royle, 2005). Variation in detection among species may also be attributed to differences in behaviour. For example, fishes have been shown to display differing levels of boldness, a trait that may influence the likelihood of an individual to enter a trap (Bell, 2004). For most species, the mean probability of detection across sites showed little change between survey years; however, this was not the case for Northern Redbelly Dace. The large change in detection for this species is due to a large discrepancy in the number of sites where it was observed between the two survey years. A number of factors could have led to this result, including population growth or variation in survey-specific characteristics. This discrepancy could also be due to a failure to distinguish between Finescale Dace and Northern Redbelly Dace in the 2015 survey, as the increase in Northern Redbelly Dace coincides with the loss of Finescale Dace. This scenario is not unlikely, as the two species are similar in appearance and have been known to hybridize to produce fertile offspring (Scott & Crossman, 1973).
Species Richness

Species richness estimates underwent changes at all threshold levels. At the strictest level, an occupancy probability of 95%, richness estimates changed by only one species; however, as threshold values decreased, an increasing number of species were added to an increasing number of sites (Figure 4). At the lowest level, an occupancy probability of 50%, over 75% of sites displayed changes in species richness. Although these threshold values were chosen arbitrarily, their consideration is important as they demonstrate the variability imperfect detection can introduce into estimates of species richness. As previously mentioned, these estimates are central to both the development of ecological theory and conservation biology and policy. It is, therefore, crucial that threshold values be given serious thought, as they have the ability to significantly alter richness estimates, and ultimately the development and testing of theory, as well as, policy development and implementation.

In Costello Creek, the largest changes in richness estimates were seen in sites characterized by clear, fast-flowing water (Habitat B). This is related to the increased occupancy probabilities of many species in this habitat, as species that expressed this relationship were more likely to cross thresholds in clear-water sites (Habitat B) without doing so in the sites within the bog area (Habitat A). These results demonstrate that, even in a relatively small system, variation in habitat can lead to largely different levels of species richness at the site level. Additionally, these findings suggest that certain habitats may be more prone to issues of detection. This is consistent with the idea that site-specific characteristics can influence detectability relative to comparable samples at other sites (Iknayan et al., 2014).

Community Structure

Community structure also underwent changes at each threshold level, with overall representation undergoing increasing change as threshold values decreased. Procrustes analyses revealed that the largest changes were consistently seen in a limited number of species: Smallmouth Bass, Brook Trout, White Sucker, and Brook Stickleback. These species, although unrelated and spanning various families, were all
characterized by low detection probabilities. Interestingly, occupancy probabilities were highly variable among these species, ranging from close to zero to just below one. This suggests that difficult-to-detect species may be the most influential in determining our understanding of community structure, regardless of occurrence or rarity. Procrustes analyses also revealed a number of sites that exhibited large changes across the thresholds. In general, sites in the clear, fast-flowing waters (Habitat B) underwent larger changes than those in the bog habitat (Habitat A), with sites B05, B11, and B17 displaying the most notable changes. Each of these sites contained at least one species characterized by low detection. Site B17, which showed the largest change across the thresholds, contained three of the four difficult-to-detect species. This suggests that sites or habitats with many species characterized by low detection probabilities may have the potential to contribute more significantly to our overall understanding of community structure than previously thought.

Limitations

Despite conducting multiple surveys, the major limitation of this study was the need for additional data. This was due to the limited number of sites available for sampling within the creek. To promote model convergence, surveys from different years needed to be pooled. This was done under the assumption that the system was closed to changes in species richness; however, not all species observed in the 2009 survey were detected in 2015. This may be a result of local extirpation or false identification, both scenarios that would violate the closure assumption. For this reason, species not present in both surveys were removed. Additionally, the complexity of the multispecies occupancy model used was determined by the available data. This led to few, broad covariates, ultimately affecting the accuracy with which the system could be represented. Further, the results of this project are contextually specific, as detection probabilities estimated by the model are highly dependent on both the type of gear used in sampling and the number of sampling replicates. The detection and occupancy probabilities of each species may, therefore, not be consistent with other locations or sampling styles. Despite these
shortcomings, this study has a number of important implications, both within and beyond the field of ecology.

Significance

These findings suggest that standard datasets, obtained through ecological sampling, are likely missing information about species occurrences across sampling locations and underestimate the composition of species at any given location. This issue may be especially pertinent for aquatic communities, as these systems present challenges in sampling that are often not encountered in terrestrial systems. For example, visual surveys are often made impossible in aquatic systems by environmental factors specific to these ecosystems, such as water clarity and depth. Other factors, such as water flow, can make sampling in these systems logistically more difficult. Imperfect detection may, thus, be playing a larger role in driving our understanding of aquatic communities than previously thought. This is particularly true for studies focusing on communities that contain large numbers of difficult-to-detect species. Imperfect detection may also have impacts beyond the field of aquatic community ecology. In this study, variability in detection was shown to largely impact estimates of species richness. This has serious implications for conservationists, as these estimates are frequently used as a basis for both decision and policy making (Yoccoz et al., 2001). Errors in these estimates may also affect population and game management, as they would lead to reduced accuracy in distribution modelling.

Incorporating multispecies occupancy models into the field of aquatic community ecology may provide a number of benefits. Firstly, it may lead to more accurate representations of diversity and community structure in these systems, ultimately, having impacts beyond the field of ecology. Additionally, incorporating model use may provide insight into species-habitat relationships, particularly for rare or difficult-to-detect species that may have been previously ignored due to limited data. Multispecies occupancy models may also have benefits that have yet to be explored. For instance, these models could be used to combat invasive species in aquatic communities. This could be done by
predicting where invasions may occur based on habitat structure or by estimating where an invasive species may be present but undetected due to low numbers.

Conclusions

In this study, imperfect detection was shown to significantly alter estimates of species richness and community structure of stream fishes. This work highlights the importance of accounting for variability in detection, particularly in communities with difficult-to-detect species. Additionally, these findings illustrate the benefits of multispecies occupancy models and, thus, promote their incorporation into the field of community ecology. Future studies may benefit by following a sampling design that is conducive to this form of modelling, as this would allow them to avoid facing the same limitations seen in this work. As these models have seen limited use, their full potential has yet to be realized and warrants further research.
References

Bell, A.M. 2004. Behavioural differences between individuals and two populations of stickleback 


Burton, A. C., Sam, M. K., Balangtaa, C., and Brashares, J.S. 2012. Hierarchical multi-species modeling 
of carnivore responses to hunting, habitat and prey in a West African protected area. \textit{PLoS ONE,} 


DeWan, A.A. and Zipkin, E.F. 2010. An integrated sampling and analysis approach for improved 

modeling the occurrence of species. \textit{Journal of the American Statistical Association, 100\textit{(470):} 389-398.}

Dorazio, R., Royle, J., Soderstrom, B., and Glimskar, A. 2006. Estimating species richness and 

Dorazio, R., Gotelli, N., and Ellison, A. 2011. Modern methods of estimating biodiversity from presence-


Appendix 1

Figure 1: Custom minnow traps used to collect samples. Traps were rectangular in shape, had three conical entrances, and were baited with a handful of dry dog food before each use.
Table 1: Catch data for the July 2015 survey of Costello Creek. Catch counts were calculated by summing the total number of individuals caught in each of the three temporal replicates.

<table>
<thead>
<tr>
<th>Site</th>
<th>Brown Bullhead</th>
<th>Brook Stickleback</th>
<th>Brook Trout</th>
<th>Creek Chub</th>
<th>Common Shiner</th>
<th>Golden Shiner</th>
<th>Northern Redbelly Dace</th>
<th>Pumpkinseed</th>
<th>Northern Pearl Dace</th>
<th>Smallmouth Bass</th>
<th>White Sucker</th>
<th>Yellow Perch</th>
</tr>
</thead>
<tbody>
<tr>
<td>A01</td>
<td>33</td>
<td>0</td>
<td>0</td>
<td>37</td>
<td>1</td>
<td>35</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>A02</td>
<td>55</td>
<td>0</td>
<td>0</td>
<td>33</td>
<td>2</td>
<td>12</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>A03</td>
<td>43</td>
<td>1</td>
<td>0</td>
<td>12</td>
<td>4</td>
<td>10</td>
<td>0</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>A04</td>
<td>99</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>2</td>
<td>11</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>A05</td>
<td>119</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>A06</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td>3</td>
<td>13</td>
<td>0</td>
<td>40</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>A07</td>
<td>22</td>
<td>0</td>
<td>0</td>
<td>40</td>
<td>1</td>
<td>16</td>
<td>1</td>
<td>12</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>A08</td>
<td>43</td>
<td>0</td>
<td>0</td>
<td>29</td>
<td>4</td>
<td>45</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>A09</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>75</td>
<td>14</td>
<td>31</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>A10</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>105</td>
<td>35</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>A11</td>
<td>81</td>
<td>0</td>
<td>0</td>
<td>35</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>A12</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>68</td>
<td>7</td>
<td>29</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>A13</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>45</td>
<td>6</td>
<td>9</td>
<td>5</td>
<td>26</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>A14</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>115</td>
<td>8</td>
<td>29</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>51</td>
</tr>
<tr>
<td>A15</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>9</td>
<td>46</td>
<td>1</td>
<td>14</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>A18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>139</td>
<td>1</td>
<td>16</td>
<td>6</td>
<td>13</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>37</td>
</tr>
<tr>
<td>B01</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>78</td>
<td>29</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>B02</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>92</td>
<td>17</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>B03</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>87</td>
<td>18</td>
<td>5</td>
<td>0</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>37</td>
</tr>
<tr>
<td>B04</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>156</td>
<td>9</td>
<td>0</td>
<td>50</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>B05</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>162</td>
<td>33</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>27</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>B07</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>305</td>
<td>9</td>
<td>0</td>
<td>33</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>B08</td>
<td>2</td>
<td>8</td>
<td>0</td>
<td>54</td>
<td>3</td>
<td>0</td>
<td>95</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>B09</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>100</td>
<td>8</td>
<td>53</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>26</td>
</tr>
<tr>
<td>B10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>116</td>
<td>50</td>
<td>0</td>
<td>21</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>B11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>105</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>36</td>
</tr>
<tr>
<td>B13</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>126</td>
<td>51</td>
<td>12</td>
<td>6</td>
<td>2</td>
<td>10</td>
<td>0</td>
<td>3</td>
<td>24</td>
</tr>
<tr>
<td>B15</td>
<td>13</td>
<td>1</td>
<td>0</td>
<td>127</td>
<td>9</td>
<td>9</td>
<td>64</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>B16</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>182</td>
<td>6</td>
<td>0</td>
<td>18</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>19</td>
</tr>
<tr>
<td>B17</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>71</td>
<td>16</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>B18</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>34</td>
<td>0</td>
<td>0</td>
<td>41</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 2: Catch data for the July 2009 survey of Costello Creek. Catch counts were calculated by summing the total number of individuals caught in each of the three temporal replicates.

<table>
<thead>
<tr>
<th>Site</th>
<th>Brown Bullhead</th>
<th>Brook Stickleback</th>
<th>Brook Trout</th>
<th>Creek Chub</th>
<th>Common Shiner</th>
<th>Golden Shiner</th>
<th>Northern Redbelly Dace</th>
<th>Northern Pumpkinseed</th>
<th>Northern Pearl Dace</th>
<th>Smallmouth Bass</th>
<th>White Sucker</th>
<th>Yellow Perch</th>
</tr>
</thead>
<tbody>
<tr>
<td>A01</td>
<td>83</td>
<td>0</td>
<td>0</td>
<td>27</td>
<td>0</td>
<td>43</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>A02</td>
<td>130</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A03</td>
<td>95</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>1</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A04</td>
<td>84</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>1</td>
<td>25</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>A05</td>
<td>92</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>A06</td>
<td>118</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>A07</td>
<td>113</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>A08</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>43</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A09</td>
<td>60</td>
<td>0</td>
<td>0</td>
<td>36</td>
<td>7</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>A10</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>104</td>
<td>33</td>
<td>7</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>A11</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>25</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>A12</td>
<td>27</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>9</td>
<td>101</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>A13</td>
<td>42</td>
<td>0</td>
<td>0</td>
<td>36</td>
<td>39</td>
<td>29</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>A15</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>71</td>
<td>28</td>
<td>14</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>A16</td>
<td>34</td>
<td>0</td>
<td>0</td>
<td>45</td>
<td>31</td>
<td>27</td>
<td>0</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>A18</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>88</td>
<td>25</td>
<td>95</td>
<td>6</td>
<td>56</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>52</td>
</tr>
<tr>
<td>B01</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>100</td>
<td>15</td>
<td>15</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>B02</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>66</td>
<td>15</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>B03</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>54</td>
<td>29</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>B04</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>11</td>
<td>8</td>
<td>1</td>
<td>8</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>B05</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>176</td>
<td>23</td>
<td>0</td>
<td>35</td>
<td>9</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>35</td>
</tr>
<tr>
<td>B07</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>126</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>29</td>
</tr>
<tr>
<td>B08</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>86</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>B09</td>
<td>39</td>
<td>0</td>
<td>0</td>
<td>51</td>
<td>57</td>
<td>7</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>B10</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>193</td>
<td>232</td>
<td>0</td>
<td>6</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>B11</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>42</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>B13</td>
<td>14</td>
<td>0</td>
<td>0</td>
<td>154</td>
<td>28</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>B15</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>45</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>B16</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>77</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>B17</td>
<td>17</td>
<td>1</td>
<td>0</td>
<td>14</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>B18</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>60</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Appendix 2

# Model designed to estimate static species-specific occupancy and detection
# with site specific and sampling covariates using the community model.

# The occurrence date are in the file "occ data.csv".

# The covariate data are in the files "habitat.csv" (occurrence) and "date.csv" (detection).

# See Zipkin et al. 2010 (Biological Conservation) for more context and details on the model.

# Read in the occurrence data

data1 <- read.table("C:/Users/David/Desktop/Occupancy_Modeling/2year_data.csv", header = TRUE, sep = ",", na.strings = TRUE)

data1$Occ <- rep(1, dim(data1)[1])

# See the first one hundred lines of data

data1[1:100,]

# How many citings for each species

total.count = tapply(data1$Occ, data1$Species, sum)

# Find the number of unique species

uspecies = as.character(unique(data1$Species))

# n is the number of observed species

n = length(uspecies)

# Find the number of unique sampling locations

upoints = as.character(unique(data1$Site))

# J is the number of sampled points

J = length(upoints)

# Reshape the data using the R package "reshape"

install.packages("reshape2")

library(reshape2)

# The detection/non-detection data is reshaped into a three dimensional
#array X where the first dimension, j, is the point (site);
#the second dimension, k, is the rep; and the last dimension, i, is the species.

junk.melt = melt(data1, id.var = c("Species", "Site", "Rep"), measure.var = "Occ")

X = acast(junk.melt, Site ~ Rep ~ Species)

#Add in the missing lines with NAs
for (i in 1: dim(X)[3]){
  b = which(X[,i] > 0)
  X[,i][b] = 1
  X[,i][-b] = 0
}

#Create all zero encounter histories to add to the detection array X
#as part of the data augmentation to account for additional
#species (beyond the n observed species)

#nzeroes is the number of all zero encounter histories to be added
nzeroes = 25

#X.zero is a matrix of zeroes, including the NAs for when a point has not been sampled
X.zero = matrix(0,nrow = 31, ncol = 6)

#Xaug is the augmented version of X. The first n species were actually observed
#and the n+1 through nzeroes species are all zero encounter histories.
Xaug <- array(0, dim = c(dim(X)[1], dim(X)[2], dim(X)[3] + nzeroes))
Xaug[,,(dim(X)[3]+1):dim(Xaug)[3]] = rep(X.zero, nzeroes)
dimnames(X) = NULL
Xaug[,1:dim(X)[3]] <- X

#K is a vector of length J indicating the number of reps at each point j
KK <- X.zero
a = which(KK==0); KK[a] <- 1
K = apply(KK,1,sum, na.rm = TRUE)
K = (as.vector(K))

#Read in the distance data
distance <- read.table("C:/Users/David/Desktop/Occupancy_Modeling/distance.csv", header = TRUE, sep = ",", na.strings = c("NA"))
#Standardize the distance data
distance <- as.vector(distance$meters)
mdistance <- mean(distance, na.rm = TRUE)
sddistance <- sd(distance, na.rm = TRUE)
distance1 <- as.vector((distance - mdistance)/sddistance)

#Read in the habitat data
habitat <- read.table("C:/Users/David/Desktop/Occupancy_Modeling/habitat.csv", header = TRUE, sep = ",", na.strings = c("NA"))
#Standardize the width data
width <- as.vector(habitat$Width)
mwidth <- mean(width, na.rm = TRUE)
sdwidth <- sd(width, na.rm = TRUE)
width1 <- as.vector((width - mwidth)/sdwidth)
width2 <- as.vector(width1 * width1)

#Create a vector to indicate which habitat type each point is in (A = 0, B = 1)
#Create vector for A sites
A <- as.vector(matrix(rep(0,16), ncol = 1, nrow = 16))
#Create vector for B sites
B <- as.vector(matrix(rep(1,15), ncol = 1, nrow = 15))
#Combine the vectors
Ind = c(A,B)

#Create a vector to indicate which year each sample was taken in (2015 = 0, 2009 = 1)
#Create vector for 2015
YR15 <- as.vector(matrix(rep(0,3), ncol=1, nrow=3))

#Create a vector for 2009
YR09 <- as.vector(matrix(rep(1,3), ncol=1, nrow=3))

#Combine the vectors
Year = c(YR15,YR09)

#Write the model code to a text file
sink("covarmodel.txt")
cat(""
    model{

    #Define prior distributions for community-level model parameters

    omega ~ dunif(0,1)

    A.mean ~ dunif(0,1)
    mu.uA <- log(A.mean) - log(1 - A.mean)

    B.mean ~ dunif(0,1)
    mu.uB <- log(B.mean) - log(1 - B.mean)

    A2.mean ~ dunif(0,1)
    mu.vA <- log(A2.mean) - log(1 - A2.mean)

    B2.mean ~ dunif(0,1)
    mu.vB <- log(B2.mean) - log(1 - B2.mean)

    mua1~dnorm(0,0.001)

    tau.uA ~ dgamma(0.1,0.5)
    tau.uB ~ dgamma(0.1,0.5)"
\[
\begin{align*}
tau.vA & \sim \text{dgamma}(0.1, 0.5) \\
tau.vB & \sim \text{dgamma}(0.1, 0.5) \\
tau.a1 & \sim \text{dgamma}(0.1, 0.5) \\
\end{align*}
\]

for (i in 1:(n+nzeroes))

# Create priors for species i from the community level prior distributions

\[
\begin{align*}
w[i] & \sim \text{dbin}(\omega) \\
u.A[i] & \sim \text{dnorm}(mu.uA, tau.uA) \\
u.B[i] & \sim \text{dnorm}(mu.uB, tau.uB) \\
v.A[i] & \sim \text{dnorm}(mu.vA, tau.vA) \\
v.B[i] & \sim \text{dnorm}(mu.vB, tau.vB) \\
a1[i] & \sim \text{dnorm}(mu.a1, tau.a1) \\
\end{align*}
\]

# Create a loop to estimate the Z matrix (true occurrence for species i at point j)

for (j in 1:J){

logit(psi[j,i]) <- u.A[i]*(1-Ind[j]) + u.B[i]*Ind[j] + a1[i]*width1[j]

mu.psi[j,i] <- psi[j,i]*w[i]

Z[j,i] ~ \text{dbin}(mu.psi[j,i])

# Create a loop to estimate detection for species i at point k during sampling period k

for (k in 1:K[j]){ 

logit(p[j,k,i]) <- v.A[i]*(1-Year[k]) + v.B[i]*Year[k]

mu.p[j,k,i] <- p[j,k,i]*Z[j,i]

X[j,k,i]~\text{dbin}(mu.p[j,k,i])

}
#Sum all species observed (n) and unobserved species (n0) to find
#the total estimated richness
n0 <- sum(w[(n+1):(n+nzeroes)])
N <- n + n0

#Create a loop to determine point level richness estimates for the whole community
for(j in 1:J){
  Nsite[j] <- inprod(Z[j,1:(n+nzeroes)],w[1:(n+nzeroes)])
}

#Finish writing the text file into a document we call covarmodel.txt


u.A=rnorm(n+nzeroes),  u.B=rnorm(n+nzeroes), 

v.A=rnorm(n+nzeroes),  v.B=rnorm(n+nzeroes), 

Z = matrix(rbinom((n+nzeros)*J, size = 1, prob = psi.meanGuess),
          nrow=J, ncol=(n+nzeros)),
a1=rnorm(n+nzeros)

)

)

library(R2WinBUGS)

bugs.dir <- "c:/Programs/WinBUGS14"

#Run the model and call the results fit
fit = bugs(sp.data, sp.inits, sp.params, "covarmodel.txt", debug = TRUE, n.chains = 2, n.iter = 300000,
          n.burnin = 10000, n.thin = 5, bugs.directory = bugs.dir, working.directory = getwd())