Sampling Effort Required to Detect Fishes at Risk in Ontario

A.J. Dextrase[1,2], N.E. Mandrak[1], J. Barnucz[1], L.D. Bouvier[1], R. Gaspardy[1], and S.M Reid[3]

[1] Great Lakes Laboratory for Fisheries and Aquatic Sciences
Central and Arctic Region
Department of Fisheries and Oceans
P.O. Box 5050, 867 Lakeshore Road
Burlington, Ontario
L7R 4A6

[2] Current address: Biodiversity Branch
Ontario Ministry of Natural Resources
300 Water Street
Peterborough, Ontario
K9J 8M5

[3] Aquatic Research and Monitoring Section
Ontario Ministry of Natural Resources
Trent University – DNA Building
2140 East Bank Drive
Peterborough, Ontario
K9J 7B8

2014

Canadian Manuscript Report of Fisheries and Aquatic Sciences 3024

Fisheries and Oceans Canada   Pêches et Océans Canada

Canada

Canadian Manuscript Report of Fisheries and Aquatic Sciences

Manuscript reports contain scientific and technical information that contributes to existing knowledge but which deals with national or regional problems. Distribution is restricted to institutions or individuals located in particular regions of Canada. However, no restriction is placed on subject matter, and the series reflects the broad interests and policies of Fisheries and Oceans Canada, namely, fisheries and aquatic sciences.

Manuscript reports may be cited as full publications. The correct citation appears above the abstract of each report. Each report is abstracted in the data base *Aquatic Sciences and Fisheries Abstracts*.

Manuscript reports are produced regionally but are numbered nationally. Requests for individual reports will be filled by the issuing establishment listed on the front cover and title page.

Numbers 1-900 in this series were issued as Manuscript Reports (Biological Series) of the Biological Board of Canada, and subsequent to 1937 when the name of the Board was changed by Act of Parliament, as Manuscript Reports (Biological Series) of the Fisheries Research Board of Canada. Numbers 1426 - 1550 were issued as Department of Fisheries and Environment, Fisheries and Marine Service Manuscript Reports. The current series name was changed with report number 1551.

Rapport manuscrit canadien des sciences halieutiques et aquatiques

Les rapports manuscrits contiennent des renseignements scientifiques et techniques qui constituent une contribution aux connaissances actuelles, mais qui traitent de problèmes nationaux ou régionaux. La distribution en est limitée aux organismes et aux personnes de régions particulières du Canada. Il n'y a aucune restriction quant au sujet; de fait, la série reflète la vaste gamme des intérêts et des politiques de Pêches et Océans Canada, c'est-à-dire les sciences halieutiques et aquatiques.

Les rapports manuscrits peuvent être cités comme des publications à part entière. Le titre exact figure au-dessus du résumé de chaque rapport. Les rapports manuscrits sont résumés dans la base de données *Résumés des sciences aquatiques et halieutiques.*

Les rapports manuscrits sont produits à l'échelon régional, mais numérotés à l'échelon national. Les demandes de rapports seront satisfaites par l'établissement auteur dont le nom figure sur la couverture et la page du titre.

Les numéros 1 à 900 de cette série ont été publiés à titre de Manuscrits (série biologique) de l'Office de biologie du Canada, et après le changement de la désignation de cet organisme par décret du Parlement, en 1937, ont été classés comme Manuscrits (série biologique) de l'Office des recherches sur les pêcheries du Canada. Les numéros 901 à 1425 ont été publiés à titre de Rapports manuscrits de l'Office des recherches sur les pêcheries du Canada. Les numéros 1426 à 1550 sont parus à titre de Rapports manuscrits du Service des pêches et de la mer, ministère des Pêches et de l'Environnement. Le nom actuel de la série a été établi lors de la parution du numéro 1551.

Canadian Manuscript Report of
Fisheries and Aquatic Sciences 3024

2014

Sampling Effort Required to Detect Fishes at Risk in Ontario

by

A.J. Dextrase[1,2], N.E. Mandrak[1], J. Barnucz[1], L.D. Bouvier[1], R. Gaspardy[1], and S.M Reid[3]

Fisheries and Oceans Canada
867 Lakeshore Road
Burlington, Ontario
L7R 4A6

[1] Great Lakes Laboratory for Fisheries and Aquatic Sciences, Central and Arctic Region, Department of Fisheries and Oceans, P.O. Box 5050, 867 Lakeshore Road, Burlington, Ontario, L7R 4A6
[2] Current address: Biodiversity Branch, Ontario Ministry of Natural Resources, 300 Water Street, Peterborough, Ontario, K9J 8M5, e-mail: alan.dextrase@ontario.ca
[3]Aquatic Research and Monitoring Section, Ontario Ministry of Natural Resources, Trent University – DNA Building, 2140 East Bank Drive, Peterborough, Ontario, K9J 7B8

HTML version: Cat. No. Fs23-176/2013E-HTML ISBN 0-662-35572-5
PDF version: Cat. No. FS23-176/2013E-PDF ISBN 0-662-35609-8

Correct citation for this publication:

Dextrase, A.J., Mandrak, N.E., Barnucz, J., Bouvier, L.D., Gaspardy, R., Reid, S.M. 2014. Sampling Effort Required to Detect Fishes at Risk in Ontario. Can. Manuscr. Rep. Fish. Aquat. Sci. 3024: v + 50 p.

TABLE OF CONTENTS

ABSTRACT

This report examines species-specific detection probabilities to complement the general guidance on sampling effort to detect Ontario fish species at risk developed in 2008. Thirty-six occupancy model candidate sets were analyzed for six fish species at risk (Silver Shiner, Pugnose Shiner, Lake Chubsucker, Northern Madtom, Eastern Sand Darter, and Channel Darter) detected using a variety of sampling gears and protocols in 10 southern Ontario waterbodies. Estimates of detection probability were used to identify the sampling effort required to detect species at the site and reach scales. Estimated detection probability ranged from 0.06 to 0.89 and varied between and within species. Within-species detection probability differed between watersheds and was positively correlated with relative abundance. Guidance on required sampling effort can be used to help design detection sampling for select Ontario fish species at risk within a defined area using specific gears.

RÉSUMÉ

Ce rapport examine les probabilités de détection propres aux espèces dans le but de compléter les directives générales sur les échantillonnages visant à détecter les espèces de poissons en péril en Ontario, élaborées en 2008. Trente-six ensembles de modèles d'occupation candidats ont été analysés pour six espèces de poissons en péril (méné miroir, méné camus, sucet de lac, chat-fou du nord, dard de sable et fouille-roche gris) détectées à l'aide de divers instruments et protocoles d'échantillonnage dans dix plans d'eau du sud de l'Ontario. On a utilisé des estimations des probabilités de détection afin de cerner l'effort d'échantillonnage requis pour détecter les espèces sur le site atteindre les échelles. Les probabilités de détection variaient entre 0,06 et 0,89 et différaient d'une espèce à l'autre. À l'intérieur des espèces, les probabilités variaient entre les bassins hydrographiques et étaient corrélées de manière positive avec une relative abondance. On peut utiliser les directives concernant l'échantillonnage requis pour faciliter l'échantillonnage de détection nominal pour certaines espèces de poissons en péril en Ontario à l'intérieur d'une zone bien définie au moyen d'instruments particuliers.

1.0 INTRODUCTION

Many activities have the potential to affect aquatic species at risk and their habitats. Species listed as Extirpated, Endangered, and Threatened under Canada's Species at Risk Act (SARA) and their residences and critical habitats are protected from harm. When a proposed activity has the potential to affect a listed fish species or its residence or critical habitat, the project proponent must work with Fisheries and Oceans Canada to determine if impacts can be avoided, or if an authorization under SARA may be required. Impacts to species at risk (including species listed as Special Concern) and their habitats must also be considered for projects subject to approval under the Canadian Environmental Assessment Act. Information on the biology of the species at risk as well as the nature, area and duration of the direct and indirect impacts of the proposed activity can be used to avoid or mitigate impacts.

In some cases, the presence of listed fish species at risk within the footprint or zone of impact of a proposed project or activity may already be known from recent fish sampling information. In other situations, there will be the potential that one or more fish species at risk occurs within the zone of impact, but there are no recent fish collection data for the area. In these cases, an appropriate sampling program needs to be designed to determine the presence or probable absence of fish species at risk that could potentially be affected by the proposed activity. While the presence of a particular species can be determined with certainty, its absence cannot. However, with a properly designed sampling program, it can be demonstrated that the presence of a species within an area is quite unlikely. The elements of a properly designed sampling program include knowledge of the biology of the target species and the deployment of the appropriate gear under the direction of experienced personnel (Portt et al. 2008). Sufficient sampling effort also needs to be deployed in suitable habitat at the appropriate time of year.

Portt et al. (2008) developed general protocols and methods for the detection of fish species at risk in Ontario to help determine if proposed projects may have an impact on species at risk and, therefore, require review under SARA. For each Ontario fish species listed under SARA at the time of development, the document identifies appropriate gear types for sampling in different habitat types (e.g., lake vs. stream vs. wetland, depth, gradient) as well as general guidance on the amount of sampling effort that would be required with different gears in different habitats to demonstrate probable absence if a species is not detected. In addition to the guidance on sampling design, the protocol provides good advice on determining the need for sampling, permit requirements, documentation of sampling, and the identification of fishes.

The recommendations for required effort in the sampling protocols of Portt et al. (2008) are expressed in terms of the amount of area that needs to be sampled without catching an individual of a species to be reasonably confident that the species is absent (Table 1). In the absence of species-specific information on capture probabilities, the recommended levels of effort were based on previous studies that examined species accumulation curves and the amount of effort required to capture a high percentage (e.g., > 95%) of the species present (e.g., Lyons 1992; Angermeier and Smogor 1995; Dauwater and Pert 2003). Because sampling effort recommendations were developed based on studies of species richness, the amount of effort required is driven by the species in the community with the lowest probability of capture. Therefore, the guidance in the protocol is conservative in that it should detect species at risk when they are present at low densities or have low capture probabilities, or both (Portt et al. 2008). Although the minimum sample areas identified in the protocol are based on single-pass sampling (sampling each portion of the sample area only once), the value of **repeat surveys**[1] is reviewed and the protocol identifies that re-sampling may be appropriate when the area of investigation or the amount of suitable habitat for the target species is small.

When repeat surveys are available, site **occupancy** (% of sites occupied by a species) and **detection probability** (probability that a species is detected at a site in an individual **survey**) can be easily estimated for individual species using the multinomial likelihood occupancy model of Mackenzie et al. (2002; 2006). This approach also allows for site occupancy and detection probability to be modelled independently as a function of **environmental and survey covariates**. When occupancy or distribution models are developed without accounting for **imperfect detection**, estimates of occupancy are biased low (Mackenzie et al. 2002; 2006) and the importance of environmental covariates may be masked or overstated (Tyre et al. 2003; Gu and Swihart 2004). Estimates of species- and gear-specific detection probability may also be affected by the suite of covariates that are included in candidate models (Dextrase 2013). Therefore, it is important to determine which candidate models have the most empirical support when developing such estimates. When estimates of site-specific detection probability are available, it is possible to determine the amount of effort that is required to detect a species within a particular area given the area is occupied by the species. This effort can be expressed as the number of repeat surveys within a site (Fig. 1) or the number of sub-sites that need to be sampled within a larger area (e.g., reach).

The purpose of this report is to identify sampling effort required to detect select Ontario fish species at risk at various spatial scales to inform the development of sampling programs to detect species when proposed activities have the potential to adversely affect species at risk and their habitat. Over the last decade, the Great Lakes Laboratory for Fisheries and Aquatic Sciences (GLLFAS) has conducted sampling with repeat surveys on several southern Ontario waterbodies that support species at risk. This information is used for the development of occupancy models and estimates of detection probabilities to identify appropriate levels of sampling effort. Factors affecting occupancy and detection probability are also examined to assist in the design of species-specific sampling programs. Guidance on collecting information for the development of additional occupancy models is provided based on a meta-analysis of occupancy models. It is not the intent of this report to replace the general guidance on sampling effort provided in the protocols of Portt et al. (2008). Rather, it complements these protocols by providing guidance on sampling effort needed to detect specific species using specific gears within a defined area.

2.0 METHODS

2.1 Field surveys

Between 2002 and 2012, 10 waterbodies in southern Ontario that support populations of fish species at risk were sampled with repeat surveys at each site (Fig. 2). Data for six fish species at risk captured in these field surveys were used for modelling occupancy and detection probability: Silver Shiner (*Notropis photogenis*); Pugnose Shiner (*Notropis anogenus*); Lake Chubsucker (*Erimyzon sucetta*); Northern Madtom (*Noturus stigmosus*); Eastern Sand Darter (*Ammocrypta pellucida*); and, Channel Darter (*Percina copelandi*). The majority of field

[1] Terms defined in the glossary are shown in bold face the first time they appear in the body of the text.

sampling was conducted between May and October by GLLFAS and graduate students associated with the lab. The United States Fish and Wildlife Service conducted Northern Madtom minnow trap surveys in the Detroit River between April and August. Individual projects focused on one or more species at risk. In some cases, projects targeting a single species incidentally captured other species at risk allowing additional modelling. For most projects, calculating detection probability was not the primary objective, but repeat surveys were conducted so that imperfect detection could be specifically accounted for. A variety of gears and sampling protocols were used for the different projects (Table 2; Appendix 1). Sites were selected randomly, systematically, using a targeted approach (sites likely to support species of interest), or by using a combination of these approaches (Table 2). For most projects, repeat surveys of sites were conducted within the same sampling visit. Repeat surveys were done in the same location for most projects, but three projects used spatial replicates of a site as repeat surveys. These spatial replicates were either stratified based on depth, or divided into consecutive transects along a linear sampling site. For three trawling projects (two on the St. Clair River and one on the Detroit River), each site was subjected to the same repeated survey design on two separate dates. Projects that involved more than one gear type (Sydenham River, Old Ausable Channel) only included one survey with each gear type and in the Old Ausable Channel, each gear was used in a different quadrant of the larger site (i.e., spatially replicated). Two projects (Eastern Sand Darter seining on the Grand and Thames rivers, and Silver Shiner seining) were designed with sample sites nested within larger reaches allowing estimation of occupancy and detection probability at both scales.

Fishes captured in each survey within a site were placed into separate holding bins and counted so that the catches from each survey could be distinguished for the calculation of detection probabilities. For most projects, fishes were held or released outside of the study site prior to further repeated surveys. For the Sydenham River gear comparison project, fishes were released back into the study site and the site was re-sampled with the second gear 1 h later. Environmental covariates measured at most sites included water velocity, depth, water clarity/turbidity, substrate composition, and macrophyte cover. A variety of detection covariates were recorded depending on the gear type and project. For projects where fishes were removed between surveys, removal was considered in **candidate sets** as a binary detection covariate.

2.2 Development of occupancy models

Occupancy models and estimates of detection probability were developed in the context of environmental and survey covariates since these factors can affect the outcome of modelling. Candidate sets of models for each species and gear type were developed *a priori* based on hypothesized relationships with environmental and survey covariates. Five employees from GLLFAS and the Ontario Ministry of Natural Resources, who have significant knowledge and field experience working on Ontario fish species at risk and a variety of sampling gears, separately ranked potential occupancy and detection covariates for each species and gear type. Candidate models were based on the top-ranked covariates from this exercise (Appendix 2). Correlations between individual detection and between individual occupancy covariates were assessed using Spearman rank-order correlations. Correlated variables were not used in the same model, but were considered in competing models. Therefore, models were not nested within the candidate set. All covariates were standardized using **z-scores** prior to modelling.

Site occupancy models that incorporate imperfect detection were developed based on the **detection histories** from repeat surveys at each site using the multinomial likelihood function of MacKenzie et al. (2002; 2006) that jointly considers the probability of occupancy (ψ) and detection probability (*p*). Through the use of a logit link function, the probability of occupancy was modelled as a function of environmental covariates and detection probability was modelled as a function of environmental covariates and survey covariates for each species and gear. Where data were available from the same project for multiple years, modelling was done separately for each year. For projects where sites were nested within larger sample reaches, reach occupancy was modelled using sites as spatially replicated surveys within reaches. To address the issue of closure violation when repeat surveys are spatial replicates (see Kendall and White 2009), multi-scale occupancy models (Nichols et al. 2008; Pavlacky et al. 2012) were used to model occupancy at the reach scale. Multi-scale occupancy models incorporate an additional parameter into the likelihood, **small-scale occupancy** (θ), that addresses the violation of closure by specifically modelling the probability that the species is present at the immediate sample site given that the larger sample unit is occupied (Nichols et al. 2008). Detection probability at the immediate sample site was modelled based on the repeat surveys at the site level forming a hierarchical model. All occupancy modelling was conducted with the program PRESENCE 5.8 (Hines 2006a, http://www.mbr-pwrc.usgs.gov/software/presence.html).

Candidate models were compared with an information-theoretic approach using **Akaike's Information Criterion (AIC)** corrected for small sample size (AICc) (Burnham and Anderson 2002). The number of sites or reaches was used as the sample size. Quasi-AIC corrected for small sample size (QAICc) was used in cases where there was **overdispersion** in the data ($\hat{c} > 1$). Goodness of fit of occupancy models within each candidate set was assessed using the Pearson chi-square statistic and the parametric bootstrap test of MacKenzie and Bailey (2004) by performing 10,000 bootstraps on the most parameter-rich (most global) model. Models in the candidate set that contained "**pretending variables**" were removed *post hoc* prior to conducting additional analysis (Anderson 2008).

Candidate occupancy models were developed in a staged fashion. First, detection models were compared using combinations of the suite of possible detection covariates while assuming constant occupancy. The best detection models were then used to examine occupancy. Small sample sizes limited the number of covariates and candidate models for many of the projects. Multi-model inference was used to create **model-averaged** parameter estimates and standard errors from all models in each candidate set (Burnham and Anderson 2002). The level of support for individual covariates was assessed by summing the **AIC weights** of the models containing that particular variable (combined model weight).

2.3 Spatial analysis

Presences and absences of species are often spatially autocorrelated (neighbouring sites are likely to have the same occupancy status) violating the assumption of independence between samples (Moore and Swihart 2005). **Spatial autocorrelation** (SAC) can result in artificially narrow confidence intervals on parameter estimates and lead to incorrect conclusions regarding the importance of predictor variables (Legendre 1993). Spatial autocorrelation in the standardized Pearson residuals for occupancy models was assessed by calculating Moran's *I* for linear river distance classes. Significance of SAC in all analyses was assessed through a randomization of 10,000 permutations and progressive Bonferroni correction at an experiment-wise error rate of $\alpha = 0.05$. All spatial analyses were conducted using PASSaGE version 2 software (Rosenberg and Anderson 2011, http://www.passagesoftware.net/). Spatial **autocovariates** were developed in an attempt to remove spatial patterns in residuals for models that displayed significant SAC. Autocovariates for individual sites and reaches were calculated based on observed occupancy of corresponding sample units within a specified lag distance (Betts et al. 2006). Autocovariates were calculated according to the following formula:

$$\text{autocov}_i = \frac{\Sigma\,[w_{ij}y_j]}{\Sigma\,[w_{ij}]},$$

where y_j is the occupancy status of the j^{th} site within the lag distance (if species is present at site j then $y_j = 1$, otherwise $y_j = 0$), and w_{ij} is the weight for site j determined by the formula:

$$w_{ij} = \frac{1}{h_{ij}},$$

where h_{ij} is the linear river distance (m) between sites i and j. Autocovariates were calculated separately for various lag distances depending on the scale of sampling. Occupancy models that displayed significant residual SAC were re-analyzed by including spatial autocovariates. The effectiveness of this approach was determined by examining AIC values of the spatial models and by assessing SAC in the residuals of the spatial models. Spatial models were used in subsequent analyses when they improved model fit and reduced residual SAC; otherwise, non-spatial models were considered.

2.4 Survey effort required to detect species

Model-averaged estimates of detection probability (p) were used to determine the amount of sampling effort required to detect a species at a site given that the site is occupied by the species using the formula:

$$\%\text{ confidence in detecting species at site} = 1 - [1 - p]^j,$$

where p is the probability of detecting the species in a single survey and j is the number of surveys conducted at a site (Stauffer et al. 2002). The numbers of repeat surveys required (J) to achieve 95% and 99% confidence in site detection were calculated based on the estimate of p and the lower 95% confidence limit of the estimate.

For reach models, model-averaged estimates of small-scale occupancy (θ) were used in conjunction with the site formula above to determine the amount of sampling effort required to detect a species in a reach given that the reach is occupied by the species using the formula:

$$\%\text{ confidence in detecting species in reach} = 1 - [1 - (Pj \times \theta)]^J,$$

where, Pj is the probability of detecting species at the site level in j surveys of the site given the site is occupied ($= 1 - [1 - p]^j$), θ is the probability that an individual site is occupied given the reach is occupied, and J is the number of sites sampled within a reach with a sampling effort of j surveys/site. The numbers of sample sites required (J) to achieve 95% and 99% confidence in reach detection were calculated based on the estimates of Pj and θ, and the lower 95% confidence limits of these estimates. To assess the implications of using site occupancy models to determine the number of sites that need to be sampled to detect species at larger scales when larger scale occupancy models are not available, model-averaged estimates of site occupancy (ψ) were substituted for estimates of small-scale occupancy (θ) in the reach detection formula. The results, in terms of required sampling effort, were compared between modelling approaches.

3.0 RESULTS

3.1 Species summaries

Thirty-six occupancy model candidate sets were analyzed for the six fish species at risk detected in 10 southern Ontario waterbodies. Summaries of key findings for each species are provided below with the results of modelling for each project included in Appendix 3. Unless specifically noted, there was no significant residual SAC in models. An example of the detailed analysis performed for an individual project is shown in Appendix 4. All estimates of occupancy and detection probabilities and associated error are based on model-averaged values.

Silver Shiner – Occupancy and detection probability were estimated at the site and reach level for four streams (Bronte Creek, Grand River [separate estimates for lower and middle portions of the river], Sixteen Mile Creek, and Thames River) that were sampled using the same seining protocol. Estimates of site occupancy ranged from 0.29 in the middle Grand River to 0.57 in Bronte Creek. There was limited support for occupancy covariates in most candidate sets with the **null model** of constant occupancy across sites ranking highest or within 1 AIC unit of the best model. The only exceptions were the Thames River and the middle Grand River where models with reach occupancy index (proportion of sites within reach where Silver Shiner was detected) as an occupancy covariate clearly outperformed other models. Estimates of reach occupancy ranged from 0.53 in the middle Grand River to 1.00 in Sixteen Mile Creek. There was limited support for covariates in reach occupancy models. There was significant residual SAC in lower Grand River site models and in the combined reach models for streams sampled in 2011. Including a spatial autocovariate eliminated residual SAC and improved model fit in both of these candidate sets (Appendix 3).

Estimates of detection probability ranged from 0.42 in the lower Grand River to 0.81 in Sixteen Mile Creek and had relatively small standard errors (9-32% of estimate; Appendix 3). A different suite of detection covariates was supported in occupancy models for each stream including removals in previous surveys (negative), depth (positive for upper Grand River and negative for the Thames River), and substrate size (negative).

The number of repeated seine surveys required to detect Silver Shiner at 92 m^2 sites and the number of sites required to detect the species in reaches (10 x stream width) based on estimates of detection probability differed between streams and was relatively small (Tables 3, 4). For most streams, the required detection effort approximately doubled when the lower confidence limit of p was considered. Silver Shiner detection probability was positively correlated with **naïve relative abundance** (Fig. 3).

Pugnose Shiner – Occupancy and detection probability were estimated from a 2002 gear comparison survey in the Old Ausable Channel of the Ausable River drainage. Four gears (boat seine, hoop net, Windermere/minnow traps, and boat electrofisher) were used at each of 16 sites in this survey. **Naïve site occupancy** was high (0.81) and modelled site occupancy was estimated at 1.00. The model that assumed constant detection probability across gear types did not reach **numerical convergence**. Models that included environmental covariates also did not reach numerical convergence, probably due to the small sample size. Estimates of detection probability for the different gear types varied widely, ranging from 0.06 for hoop nets and Windermere/minnow traps to 0.80 for boat electrofishing.

The number of repeat surveys required to detect Pugnose Shiner at sites similarly ranged from a small amount using a boat electrofisher to a prohibitive amount using the passive trapping gears with low detection probabilities (Table 5). Caution should be used in interpreting detection probabilities for Pugnose Shiner due to the small sample size, the use of spatial replicates (may underestimate detection probability), and the fact that each gear was fished at each site only once.

Lake Chubsucker – Occupancy and detection probability were estimated from the same gear comparison study used for Pugnose Shiner discussed above. Naïve Lake Chubsucker site occupancy was 0.62 and modelled site occupancy was estimated at 1.00. The model that assumed constant detection probability across gear types did not reach numerical convergence. Models that included environmental covariates also did not reach numerical convergence. The highest estimated detection probability was for boat electrofishing (0.47) while no Lake Chubsucker were caught in the Windermere/minnow trap combination. Detection probabilities for the boat seine and hoop nets were low (0.25 and 0.12, respectively). There was significant residual SAC, but models that included a spatial autocovariate did not reach numerical convergence.

Required sampling effort to detect Lake Chubsucker at a site was lower for boat electrofishing than for the other gear types (Table 6). Similar to the warning for Pugnose Shiner, caution should be used in interpreting detection probabilities for Lake Chubsucker from this study.

Northern Madtom – Ten site occupancy models for Northern Madtom were developed based on trawl surveys in the St. Clair (4) and Detroit (1) rivers, minnow trap surveys in the Detroit River (4), and seining in the Thames River (1). A variety of sampling protocols were used for trawling associated with different assessment programs so the estimates of detection probability are not directly comparable between trawling projects. Three projects (3 x 50 m trawls on the St. Clair River, 10 x 100 m consecutive transect trawls on the St. Clair River, and 2 x [5 x 100 m consecutive transect trawls] on the Detroit River) were repeated on two occasions. Occupancy and detection probability for all three of these projects were analyzed based on the repeat surveys within the first occasion as well as by pooling observations from each occasion and treating each occasion as an individual survey. Attempts to model occupancy for a single sampling occasion for the linear transect surveys by treating transect segments as repeated surveys in occupancy models with spatial correlation (see Hines et al. 2010) were unsuccessful. These models did not reach numerical convergence, likely due to the small sample size and the extra parameters included in the models. Minnow trap data for the Detroit River were analyzed separately for each survey year. Because site locations and effort (traps/survey) varied between years, estimates of detection probability are not directly comparable. The candidate set for the 2008 minnow trapping was discarded as the data were highly overdispersed ($\hat{c}$ = 6.06) and there was a lack of model fit (p-value = 0.03)

There were few Northern Madtom detected in each project with five of the ten occupancy models having naïve occupancies of less than 0.10. Estimates of site occupancy ranged from 0.12 for minnow traps in the Detroit River in 2011 to 0.94 for 10 x 100 m trawl sets (a trawl set represents an individual haul along a transect) in the St. Clair River. There was limited support for occupancy covariates with the null model of constant occupancy across sites ranking highest in all candidate sets. Habitat information was not available for the Detroit minnow trapping so occupancy covariates were not considered. There was significant residual SAC in three of the Northern Madtom candidate sets for the St. Clair River, but models that included a spatial autocovariate did not reach numerical convergence.

Estimates of detection probability ranged from 0.07 for 50 m trawl sets in the St. Clair River to 0.67 for 1 km trawl sets (2 x [5 x 100 m]) in the Detroit River. The standard errors of most estimates were relatively large and, in three of nine cases, were larger than the estimates of detection probability (37-136% of estimate; Appendix 3). Across streams, detection probability was highest for those projects with larger sample sites (*p* of 0.41 and 0.67 for 1 km trawl sites versus *p* of 0.07 for 50 m trawl sites and 0.09 for 92 m^2 seine sites). The same pattern was seen for trawling within the St. Clair River with progressively larger detection probabilities at site lengths of 50 m, 100 m and 1 km (Fig. 4). It should be noted that these projects were not conducted at the same locations within the St. Clair River. There was no support for detection covariates in most candidate sets. The only exceptions were the 50 m trawl set project in the St. Clair River and the minnow trap projects in the Detroit River. Detection probability was positively related to trawl time (inversely related to trawling speed) for the trawling project and was positively related to water temperature and trap soak time in some years of the minnow trap project.

The number of repeat surveys required to detect Northern Madtom at sites based on estimates of detection probability varied widely and tended to be large with the exception of the 1 km trawl sets for the Detroit and St. Clair rivers and the 2003 minnow traps in the Detroit River (Table 7). Due to the large standard errors of the detection probability estimates, the required detection effort was extremely large when the lower confidence limit of *p* was considered. Only the Thames River seining study was designed to allow occupancy modelling of Northern Madtom at the larger reach scale. Despite the predicted reach occupancy of 1.00, the extremely low estimate of detection probability (0.02) meant that a large number of sites would need to be sampled in an individual reach to detect the species (Table 7).

Eastern Sand Darter - Occupancy and detection probability were estimated at the site level for three streams (Grand River, Sydenham River, and Thames River) using a variety of gear types and sampling protocols. Reach occupancy was also estimated for the Grand and Thames rivers. Estimates of site occupancy using comparable seining techniques in the three rivers ranged from 0.23 in the Grand River to 0.67 in the Thames River. There was strong support for substrate as an occupancy covariate in almost all candidate sets (the null model ranked highest for the Sydenham seining project with a small sample size [*n* = 12]). For the Sydenham gear comparison study that sampled larger areas of stream, there was a positive relationship between occupancy and mean substrate size contrary to the hypothesized relationship. For all other projects, the relationship between Eastern Sand Darter occupancy and substrate was consistent with the hypothesized relationship (occupancy positively related to smaller substrate and sand and fine gravel). Other important occupancy covariates included water clarity in **additive models** for the Grand and Thames rivers (positive relationship with occupancy), and water velocity (positive) and river distance (greater probability of occupancy further downstream) in the Sydenham River. Reach occupancy was estimated at 0.52 in the Grand River and 0.98 in the Thames River. In the Grand River, reach occupancy was positively related to the amount of sand and fine gravel at sites and distance upstream of dams. There was little support for occupancy covariates in the Thames River as virtually all reaches were occupied.

There was significant residual SAC in site models for the Grand River (seining and 2010 trawling), Sydenham River (2003), and Thames River. Including a spatial autocovariate eliminated residual SAC and improved model fit for the Grand River candidate sets. Residual SAC remained in spatial models for the Thames River and model fit was not improved. Spatial models for the Sydenham River did not reach numerical convergence.

Estimates of detection probability for seining projects ranged from 0.41 in the Thames River to 0.71 in the Grand River. Naïve relative abundance of Eastern Sand Darter was highest in the Grand River (Fig. 5). For the Grand River trawling projects, detection probability was high in both years (0.71, 0.76) similar to the Grand River seining project. Eastern Sand Darter detection probabilities for backpack electrofisher and seine were similar in the first year of the Sydenham River gear comparison project (0.53 and 0.52, respectively), with limited support for sampling gear as a detection covariate. In the second year of the project, detection probability was higher for electrofishing (0.62) than for seining (0.45), and the highest ranked model had sampling gear as a detection covariate. Substrate was an important detection covariate for three of the five seining projects and for both trawling projects (detection probability negatively related to substrate size).

The number of repeat surveys required to detect Eastern Sand Darter at sites based on estimates of detection probability differed between streams and gear types, but tended to be relatively small (Table 8). When the lower confidence limit of *p* was considered, the required detection effort increased substantially for the Sydenham River where sample sizes were small and the standard errors of detection probability estimates were relatively large(32-49% of estimate; Appendix 3). Despite the higher detection probability at the site level in the Grand River versus the Thames River, more sampling effort was required to detect Eastern Sand Darter in Grand River reaches associated with the lower likelihood of occupancy of sites within occupied reaches (Table 9).

Channel Darter - Occupancy and detection probability were estimated at the site level based on trawling projects in the Ottawa and Detroit rivers that used different trawling protocols (100 m vs. 1 km sets). Models for the Lake St. Clair trawling project did not reach numerical convergence as Channel Darter were only detected at one of thirty-six sites sampled. Estimates of site occupancy were high for the Ottawa (0.97) and Detroit (0.84) rivers. There was little support for occupancy covariates in both rivers as most sites were occupied. Estimates of detection probability were 0.68 for the Ottawa River and 0.89 for the Detroit River. There was limited support for detection covariates in site occupancy models for both rivers.

The number of repeat trawl surveys required to detect Channel Darter was low for both rivers when estimates of detection probability and their lower confidence limits were considered (Table 10). It should be noted that both of these surveys targeted areas that were likely to support Channel Darter and, consequently, the probabilities of occupancy and detection were high.

3.2 Meta-analysis of occupancy models

Predicted site occupancy was always higher than the observed or naïve occupancy (Fig. 6). The ratio of predicted occupancy/naïve occupancy ranged from 1.01 (Silver Shiner in Sixteen Mile Creek) to 12.68 (Northern Madtom in Thames River) with a mean value of 2.22. Predicted site occupancies tended to be closer to naïve occupancies when detection probabilities were high (Fig. 7). Naïve site occupancy and predicted site occupancy were not strongly correlated with estimates of detection probability (Figs. 8, 9). Relative errors of detection probability estimates (SE of p/p) were examined in the context of sample size, detection probability, and the number of repeat surveys per site. The relative error was low for projects that included more than 50 sample sites, with the exception of the Northern Madtom seining project on the Thames River (n = 53) that had an extremely low p estimate of 0.09

(Fig. 10). There was also a tendency of decreasing relative error as estimated detection probability increased (Fig. 11). There was no apparent relationship between relative error and the number of repeat surveys at each site (Figs. 10, 11).

There was significant residual SAC in the top-ranked models in 10 of the 36 model candidate sets. Inclusion of the spatial autocovariate removed residual SAC and improved model fit in four of the candidate sets. Inclusion of the spatial autocovariate did not appreciably change most estimates of detection probability and required sampling effort. The only exception was the spatial Silver Shiner model for sites in the lower Grand River where the required sampling effort to detect the species based on the lower confidence limit of p increased in spatial models that had a larger standard error for the estimate of detection probability.

Similar to the site occupancy models, predicted reach occupancy was higher than naïve reach occupancy for all projects with the exception of the Silver Shiner seining project on Sixteen Mile Creek, where Silver Shiner were detected in the field at all reaches sampled. The ratio of predicted to naïve occupancy ranged from 1.0 to 4.0 with a mean value of 1.4. The probability of small-scale occupancy (θ) for reach models was higher than the estimate of occupancy (ψ) based on site models for the same data with the exception of the Silver Shiner model for Sixteen Mile Creek where these two values were equal. The difference between estimates of small-scale occupancy from reach models and estimated occupancy from site models decreased as estimated reach occupancy approached 1 (Fig. 12). When the estimate of site occupancy (ψ) was substituted for small-scale occupancy (θ), the number of sites that need to be sampled to detect a species at risk at the reach scale increased (Fig. 13). On average, the number of sites that need to be sampled was 1.5 x higher based on estimates of p and ψ, and 2.0 x higher when based on the lower confidence limits of these estimates. The amount of required sampling effort was only similar between the two approaches when estimated reach occupancy was near 1.

4.0 DISCUSSION

Detection of fish species at risk was imperfect for all projects in this study with estimates of the probability of detecting species in a single survey at a site ranging from 0.06 to 0.89. Detection probabilities of less than 1 have implications for designing surveys to determine the presence or absence of a target species and the development of occupancy models to predict the distribution of species. Presence/absence surveys need to include repeat surveys to have confidence in detection within a defined area, and occupancy models need to specifically account for detection probability to avoid underestimating occupancy and to identify important occupancy covariates.

There are four assumptions of the multinomial likelihood models used to model occupancy and detection probability in this study: 1) the occupancy state of the sites does not change during the period of surveying (i.e., sites are closed); 2) the probability of occupancy is constant across sites, or differences in occupancy are modelled using covariates; 3) the probability of detection is constant across sites and surveys or differences in detection probability are modelled using site and survey covariates; and, 4) the detection of species and detection histories at each location are independent (MacKenzie et al. 2006). Because the repeat surveys conducted at each site were conducted over a relatively short time frame, it is reasonable to assume closure for most surveys. Exceptions were the Northern Madtom trawling project on the St. Clair River (100 m sets) and the Old Ausable Channel gear comparison study, which used spatial replicates as repeat surveys of each site. Violation of the closure assumption at the reach level (using sites as spatial replicates) was addressed through the use of multi-scale models. The probabilities of occupancy and detection were both modelled using site-specific covariates based on hypothesized relationships and survey-specific covariates were also used to model detection probability. The fourth assumption that the detection of species and detection histories at sites are independent was likely violated for some surveys as indicated by significant SAC in model residuals. This was addressed by developing spatial models. The small sample sizes associated with some of the projects (e.g., Detroit River trawling, OAC gear comparison) meant that the number of covariates that could be included was limited, standard errors associated with estimates were large, and several candidate models did not reach numerical convergence. Despite these drawbacks, results of modelling were included when stable maximum likelihood outcomes were achieved due to the lack of alternative information on detection probability for several species.

The probability of detecting a species is related to the abundance of the species and the probability of individual capture, both of which may be influenced by habitat (Bayley and Peterson 2001). Detection probability may also vary depending on the sampling method, amount of sampling effort, and the time of year or time of day that sampling takes place and, therefore, it can vary among survey sites and within the sampling season for the same species (Gu and Swihart 2004; MacKenzie et al. 2006; Hayer and Irwin 2008). Ecological variables can affect detection probability when they are related to density, behaviours that influence detection, or the efficiency of a sampling technique (Gu and Swihart 2004). A variety of site and survey covariates were important detection covariates for the species at risk in this study and these varied between and within species. The only consistent detection covariate was a negative relationship between substrate size and seining detection probability for Eastern Sand Darter. Even when identical gear and sampling protocols were used, there was evidence of large differences in detection probabilities for populations of the same species in different streams (Silver Shiner, Eastern Sand Darter). These differences mean that the required effort to detect the species is specific to each stream/population. For Silver Shiner and Eastern Sand Darter, the required detection effort was two to three times higher for populations with lower detection probability. The differences in detection probability within these species may, in part, be driven by differences in abundance and density in each stream. For Northern Madtom where surveys were conducted at sample sites of different sizes, detection probability was also higher for the larger sites. This supports sampling larger areas to detect rare species with low detection probabilities.

There is a higher likelihood of larger areas being occupied than smaller areas. The probability that a species will be detected in an area larger than an individual sample site when several sites are used to sample the broader scale, depends on the probability of detection at an individual site and the probability that individual sites are occupied given that the larger area is occupied. This means that more effort will be required to detect species at broader scales in systems where they are patchily distributed than in systems where they are more continuously distributed, even when detection probabilities are identical. Paradoxically, more effort is required to detect Eastern Sand Darter in reaches in the Grand River than in the Thames River, despite the fact that less effort would be required to detect the species at sites in the Grand River. When occupancy surveys have been specifically designed to assess occupancy at the site level as well as at larger scales (e.g., reach), the probability of small-scale (site) occupancy can be specifically modelled and used with estimates of detection probability to determine the number of sites that need to be sampled to detect the species at the broader scale. When surveys are designed to evaluate only site-level occupancy, required detection effort at larger scales can still be calculated, but bias is introduced due to the assumption of constant site-level occupancy within each larger scale sample unit. This inflates the number of sites that need to be sampled within the larger sample unit, unless the occupancy of the larger units is close to 1.

An important caveat regarding the guidance on survey effort provided by the analysis in this report is the need to consider the context of models used to estimate occupancy and detection probability. For each species and project, the way that sample sites were selected, the size of sites, the sample gear, and the time of year could all affect detection probability estimates. Models built based on sites that were randomly or systematically selected are more likely to represent detection probabilities relevant to **detection sampling** associated with projects that could potentially impact species at risk. Targeted surveys (e.g., Channel Darter in the Ottawa River) will have higher than average occupancy, abundance, and detection probability and, therefore, underestimate the amount of effort required to detect the species at randomly selected sites. Detection probability may also vary with different sizes of sample sites. For Northern Madtom in the St. Clair River, detection probability increased with increasing site size. The use of different gears can also affect detection probability. Active gears (electrofishing, seining) had higher detection probabilities for Pugnose Shiner and Lake Chubsucker than passive gears (minnow traps and hoop nets) and, consequently, can be used to detect these species with less effort. On the Sydenham River, the amount of effort required to detect Eastern Sand Darter with electrofishing and seining was comparable. Most of the surveys used to model detection probability were conducted during low flow conditions during the late spring and summer months. Detection probability may be quite different during colder months or during periods of high flow. The detection probability of Northern Madtom in minnow traps was positively related to increasing water temperatures in the spring. Because abundance of species and habitat conditions may change from year to year, detection probability may also change between years. In this study, there were only two surveys repeated over two years (Eastern Sand Darter trawling on Grand River and gear comparison on the Sydenham River). Although these surveys did not sample all of the same sites in both years and there were some slight differences in estimates of occupancy and detection probability, the required sampling effort to detect the species was similar for both years in both surveys. This study also did not distinguish between different life stages of individual species. Different life stages and sizes of fish (e.g., juvenile versus adult) may vary in their abundance and susceptibility of capture in particular gears. All of these factors need to be considered when using information from this report as guidance for designing detection sampling.

Guidance on the number of repeat surveys required to detect a species at risk within a specified area (site, reach) is based both on the estimate of detection probability and the lower confidence limit of the estimate. When sample sites in detection sampling have characteristics that would lead to high detection probability (e.g., small substrate size for Eastern Sand Darter seining), it would make sense to use the lower level of sample effort based on the estimate of detection probability. If models predict lower detection probability or if the site characteristics are unknown, then survey effort could be based on the lower confidence limit of the estimate of p that generally involves higher sample effort. Detection sampling should also be designed to account for the expected abundance of the species within the system or specific site. When a species would be expected to be abundant because of site-specific habitat conditions or the nature of the system as a whole, less survey effort would be required. If the level of abundance is expected to be low or is unknown, a greater level of survey effort would be required. Although the number of repeat surveys may appear prohibitive in some cases, especially when detection probability is low, it should be noted that detection sampling efforts can be suspended as soon as one individual of the target species is detected. Also, the proponent of a proposed activity has the option of simply assuming that the species at risk is present and not doing any survey work (Portt et al. 2008).

The analysis of detection probability and required sampling effort presented in this report provides insights into the design of future projects to determine detection probabilities for additional species and gear types. The relative error of detection probability estimates was consistently low only for surveys with greater than 50 sample sites and non-negligible detection probabilities (> 0.10). Surveys with high detection probability also had lower relative error. Mackenzie and Royle (2005) recommended that occupancy surveys should consist of at least three repeat surveys/site when detection probability is high (> 0.50) and more repeat surveys at lower values. There was no obvious relationship between the number of repeat surveys and relative error of *p* in this study. Mackenzie and Royle (2005) examined trade-offs between the total number of sites and the number of repeat surveys when developing occupancy models. They demonstrated that for rare species it is better to survey more sites less intensively (fewer repeat surveys) and that, for common species, it is more efficient to sample fewer sites with more repeat surveys/site. However, when detection probability is a quantity of interest, it is more efficient to increase the number of surveys at each site compared to when occupancy is of primary concern (Guillera-Arroita et al. 2010). If there is prior knowledge of potential occupancy rates and detection probabilities, the freely available software GENPRES can be used to identify optimal allocation of effort to sites and repeat surveys to minimize error when developing occupancy models and estimating detection probabilities (Hines 2006b, http://www.mbr-pwrc.usgs.gov/software/presence.html; Bailey et al. 2007). It is recommended that a minimum of 50 sample sites and three repeat surveys/site be used as a starting point for designing studies to estimate occupancy and detection probability for Ontario fish species at risk. In addition to reducing error, larger sample sizes allow for the consideration of covariates that can result in better occupancy models and help inform the design of detection sampling. It is important that hypothesized covariates be identified *a priori* and that they are measured at each sample site. The size of sample sites should be relevant to the biology of the species (e.g., seasonal home range) and should be selected randomly from within a larger area expected to be occupied by the population. Sample sites can also be nested within larger sample units that are randomly selected to assess effort required to detect species at larger scales. If possible, repeated surveys should be conducted on separate site visits if these can be conducted in a short enough time frame to avoid violation of the closure assumption (same occupancy status for each repeat survey). Removal designs (where repeat surveys are terminated at a site once a species is detected) should be avoided because of the reduced information provided on factors that affect detection probability (MacKenzie and Royle 2005).

Guidance on required sampling effort in this report can be used to help design detection sampling for select Ontario fish species at risk within a defined area using specific gears. It is recognized that comparable information for other species at risk and for other appropriate sampling gears for the species examined in this study (e.g., electrofishing for Channel Darter in wadeable streams) is not currently available. For these situations, it is recommended that the general guidance for designing detection sampling found in Portt et al. (2008) be followed. The development of occupancy and detection models for additional species and gear types will be useful to build upon the guidance provided by this study.

5.0 ACKNOWLEDGMENTS

The authors are grateful to James Boase and Justin Chiotti of the United States Fish and Wildlife Service for sharing their Northern Madtom minnow trapping data from the Detroit River. The authors would also like to thank two internal reviewers for their comments that resulted in improvements to earlier drafts of the report.

6.0 GLOSSARY

Additive model: a model that includes two or more covariates without considering interactions between covariates.

AIC (Akaike's Information Criterion): a measure of the relative quality of a statistical model for a given set of data that represents the amount of information lost for any particular model. Lower AIC values within the same *candidate set* of models represent better models. AIC = -2 x [log likelihood] + 2*K*, where *K* is the number of model parameters.

AIC weight: level of support for each model in the *candidate set*. For a given model, the AIC weight is between 0 and 1 and the sum of AIC weights for all models in the *candidate set* totals 1.

Autocovariate: a covariate in a statistical model that represents the relationship between sample points (temporal or spatial) to remove autocorrelation in model residuals.

Candidate set: a set of competing statistical models that are assessed to determine how well they explain observed data.

Detection history: a record of detections (1's) and non-detections (0's) in *repeat surveys* at an individual site. If a species was detected in only the second of three *repeat surveys* at a site, the detection history would be: [010].

Detection probability (*p*): the probability that at least one individual of a target species is detected in a single *survey* of a site given the site is occupied by the species.

Detection sampling: field work conducted to determine if a species at risk is present in the area affected by a proposed activity.

Environmental (site) covariate: a covariate used in *occupancy* models that describes the environmental condition of sample sites. Environmental covariates are the same for all *repeat surveys* at a site.

Imperfect detection: the condition when *detection probability* is less than 1. It implies that sampling efforts will not detect a species at some sites where it is actually present.

Model-averaged: parameter estimates that use information from each model in the candidate set based on their level of support as indicated by *AIC weight*.

Naïve occupancy: the proportion of sites sampled where the species was detected.

Naïve relative abundance: the mean number of individuals detected at sample sites.

Null model: a model that does not include any covariates. For *occupancy* models, null models assume that the probability of *occupancy* and *detection probability* are constant at all sites.

Numerical convergence: condition achieved when parameters are successfully estimated using the iterative maximum likelihood approach. Parameters cannot be estimated for models that do not reach numerical convergence.

Occupancy (ψ): probability that a randomly selected site or sampling unit in an area of interest is occupied by a species.

Overdispersion (*ĉ*): occurs when the observed variance in the data is larger than the predicted variance. It is necessary to account for overdispersion in the data (i.e., ĉ > 1) when calculating *AIC* values (Quasi-AIC) and unconditional variances of *model-averaged* parameter estimates.

Predicted occupancy: estimate of *occupancy* from models that incorporate detection probability. When there is *imperfect detection*, predicted occupancy > *naïve occupancy.*

Pretending variable: a variable that has no effect on the deviance (-2 x [log likelihood]) and results in an *AIC* value approximately 2 *AIC* units higher than the previous model. It has no effect on the predictive value of the model (as indicated by the lack of change in deviance) and simply increases the *AIC* value by 2 units through adding an additional variable to the *AIC* parameter penalty term.

Repeat surveys: refers to individual *surveys* that are conducted more than once at an individual site within one sample season.

Small-scale occupancy (θ): probability that an individual sample site is occupied by a species given that the larger scale sample unit is also occupied.

Spatial autocorrelation (SAC): occurs when characteristics at locations close together are negatively or positively correlated violating the assumption of independence of observations.

Survey: in *occupancy* models, a survey is an individual sampling event at a site. For the *repeat surveys* required in *occupancy* modelling, there are multiple surveys at each site.

Survey covariate: a covariate used in *occupancy* models that describes factors affecting *detection probability*. Survey covariates can either remain constant for *repeat surveys* at a site (e.g., substrate size) or vary from *survey* to *survey* (e.g., trawl speed, removals in previous *surveys*).

z-score: transformation that identifies the distance of each observation from the mean relative to the standard deviation. $\text{z-score}_i = [x_i - \mu] / s$.

7.0 REFERENCES

Anderson, D.R. 2008. Model based inference in the life sciences: a primer on evidence. Springer, New York. 184 p.

Angermeier, P.L., and Smogor, R.A. 1995. Estimating number of species and relative abundances in stream-fish communities: effects of sampling effort and discontinuous spatial distributions. Can. J. Fish. Aquat. Sci. 52: 936-949.

Bailey, L.L., Hines, J.E., Nichols, J.D., and MacKenzie, D.I. 2007. Sampling design trade-offs in occupancy studies with imperfect detection: examples and software. Ecol. Appl. 17: 281-290.

Bayley, P.B., and Peterson, J.T. 2001. An approach to estimate probability of presence and richness of fish species. T. Am. Fish. Soc. 130: 620-633.

Betts, M.G., Diamond, A.W., Forbes, G.J., Villard, M.-A., and Gunn, J. 2006. The importance of spatial autocorrelation, extent and resolution in predicting forest bird occurrence. Ecol. Model. 191: 197–224.

Burnham, K.P., and Anderson, D.R. 2002. Model selection and multimodel inference: a practical information-theoretic approach, second edition. Springer-Verlag, New York. 488 p.

Dauwalter, D.C., and Pert, E.J. 2003. Electrofishing effort and fish species richness and relative abundance in Ozark Highland streams of Arkansas. N. Am. J. Fish. Manage. 23: 1152–1166.

Dextrase, A.J. 2013. Modelling occupancy and abundance of Eastern Sand Darter (*Ammocrypta pellucida*) while accounting for imperfect detection. Thesis (Ph.D.) Trent University, Peterborough, ON. xxi + 352 p.

Gu, W., and Swihart, R.K. 2004. Absent or undetected? Effects of non-detection of species occurrence on wildlife-habitat models. Biol. Cons. 116: 195-203.

Guillera-Arroita, G., Ridout, M.S., and Morgan, B.J.T. 2010. Design of occupancy studies with imperfect detection. Methods Ecol. Evol.1: 131-139.

Hayer, C.-A., and Irwin, E.R. 2008. Influence of gravel mining and other factors on detection probabilities of Coastal Plain fishes in the Mobile River basin, Alabama. T. Am. Fish. Soc. 137: 1606-1620.

Hines, J. E. 2006a. PRESENCE 5.7- Software to estimate patch occupancy and related parameters. United States Geological Survey – Patuxent Wildlife Research Centre. http://www.mbr-pwrc.usgs.gov/software/presence.html.

Hines, J. E. 2006b. GENPRES- Software to Generate patch occupancy data and analyze using programs MARK or PRESENCE. United States Geological Survey – Patuxent Wildlife Research Centre. http://www.mbr-pwrc.usgs.gov/software/presence.html.

Hines, J.E., Nichols, J.D., Royle, J.A., MacKenzie, D.I., Gopalaswamy, A.M., Samba Kumar, N., and Karanth, K.U. 2010. Tigers on trails: Occupancy modeling for cluster sampling. Ecol. Appl. 20: 1456-1466.

Kendall, W.L., and White, G.C. 2009. A cautionary note on substituting spatial subunits for repeated temporal sampling in studies of site occupancy. J. Appl. Ecol. 46: 1182-1188.

Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? Ecology 74: 1659-1673.

Lyons, J. 1992. The length of stream to sample with a towed electrofishing unit when fish species richness is estimated. N. Am. J. Fish. Manage. 12: 198-203.

MacKenzie, D.I., and Bailey, L.L. 2004. Assessing the fit of site occupancy models. J. Agric. Biol. Envir. S. 9: 300-318.

MacKenzie, D., Nichols, J.D., Lachman, G., Droege, S., Royle, A., and Langtimm, C. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83: 2248–2255.

MacKenzie, D.I., Nichols, J.D., Royle, J.A., Pollock, K.H., Hines, J.E., and Bailey, L.L. 2006. Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Elsevier, San Diego, CA. 324 p.

MacKenzie, D.I., and Royle, J.A. 2005. Designing occupancy studies: general advice and allocating survey effort. J. Appl. Ecol. 42: 1105-1114.

Moore, J.E., and Swihart, R.K. 2005. Modelling patch occupancy by forest rodents: incorporating detectability and spatial autocorrelation with hierarchically structured data. J. Wildl. Manage. 69: 933-949.

Nichols, J.D., Bailey, L.L., O'Connell, A.F., Talancy, N.W., Grant, E.H.C., Gilbert, A.T., Annand, E.M., Husband, T.P., and Hines, J.E. 2008. Multi-scale occupancy estimation and modelling using multiple detection methods. J. Appl. Ecol. 45: 1321-1329.

Pavlacky, D.C., Jr., Blakesley, J.A., White, G.C., Hanni, D.J., and Lukacs, P.M. 2012. Hierarchical multi-scale occupancy estimation for monitoring wildlife populations. J. Wildl. Manage. 76: 154–162.

Portt, C.B., Coker, G.A., Mandrak, N.E., and Ming, D.L. 2008. Protocol for the detection of fish species at risk in Ontario Great Lakes Area (OGLA). DFO Can. Sci. Advis. Sec. Res. Doc. 2008/026. v + 31 p.

Rosenberg, M.S., and Anderson, C.D. 2011. PASSaGE: Pattern analysis, spatial statistics, and geographic exegesis. Version 2. Method. Ecol. Evol. 2: 229–232.

Stauffer, H.B., Ralph, C.J., and Miller, S.L. 2002. Incorporating detection uncertainty into presence–absence surveys for marbled murrelet. In Predicting Species Occurrences: Issues of Accuracy and Scale. Edited by J.M. Scott, P.J. Heglund, M.L. Morrison, J.B. Haufler, M.G. Rafael, W.A. Wall, and F.B. Samson. Island Press, Washington, DC. pp. 357–365.

Tyre, A.J., Tenhumberg, B., Field, S.A., Niejalke, D., Parris, K., and Possingham, H.P.. 2003. Improving precision and reducing bias in biological surveys: estimating false-negative error rates. Ecol. Appl. 13: 1790-1801.

Table 1. Preferred gears and minimum sampling effort required in order for failure to capture a species at risk to demonstrate that the species is probably not present in the study area (from Portt et al. 2008).

Habitat	Gear[1]	Sampling distance/area
wadeable streams	backpack electrofisher	greater of 250 m or 50 stream widths
	seine (where conditions suitable)	greater of 250 m or 75 stream widths
non-wadeable rivers and streams	boat electrofisher	50 stream widths
nearshore or littoral habitats	backpack electrofisher	shoreline length of 50 times wadeable distance from shore
	seine (where conditions suitable)	shoreline length of 75 times wadeable distance from shore
	boat electrofisher	shoreline length of 50 times width of target depth-defined habitat, as measured perpendicular to shore
offshore shallow habitats (< 3 m)	multiple gears - trap nets, boat electrofisher, seine (from boat)	30 stations
wetlands	multiple gears - hoop nets, trap nets, seines	50 sampling events/stations with appropriate gear(s)
deep habitats (lake or river)	video or other observation methods, gill nets, trawls, traps	Reasonable effort decided on a site-specific basis

[1] Conductivity must be considered when determining if electrofishing is a suitable method.

Table 2. Summary of sampling projects used to develop occupancy models and estimates of detection probability (*n* represents the number of sample sites repeatedly sampled; see Appendix 1 for specifications of netting gear).

Species	Location (*n*)	Gear	Site selection	Sampling protocol
Eastern Sand Darter	Grand River (151), Thames River (131)	9.2 m bag seine	stratified, random	3 repeated surveys at each site (92 m^2) during the same visit, 1-9 sites/reach[1] (May-Sep.)
	Sydenham River (12)	9.2 m bag seine	targeted	3 repeated surveys at each site (92 m^2) during the same visit (Aug.)
	Sydenham River (28, 26)	8.2 m bag seine, BP electrofisher	systematic	~ 60 m length of stream sampled with each gear type during the same visit, 2 sample years (Jun.-Aug.)
	Grand River (33)	Missouri trawl	targeted + systematic	3 repeated 100 m trawl sets at each site during the same visit, 2 sample years (Jul.-Aug.)
Silver Shiner	Lower Grand River (73)	9.2 m bag seine	stratified, random	3 repeated surveys at each site (92 m^2) during the same visit, 3-8 sites/reach[1] (May-Jul.)
	Middle Grand River (95), Thames River (83), Bronte Creek (30), Sixteen Mile Creek (24)	9.2 m bag seine	targeted + systematic with stratification within reaches	3 repeated surveys at each site (92 m^2) during the same visit, 3-5 sites within different habitats/reach[1] (May-Jul.)
Northern Madtom	St. Clair River (17)	Missouri trawl	targeted + systematic	3 spatial replicate 100 m trawl sets at each site stratified by depth (Oct.)
	St. Clair River (31)	Siamese trawl	targeted	3 repeated 50 m trawl sets at each site repeated after Sea Lamprey treatment (Jun.)
	St. Clair River (8)	Siamese trawl	targeted	10 consecutive 100 m trawl sets per site, each site sampled twice (Jul.-Oct.)
	Detroit River (6)	Missouri trawl	targeted (historic RAP sites)	2 x [5 x 100 m consecutive trawl sets] per site stratified by depth, each site sampled twice (Jul.-Sep.)
	Detroit River (12, 25, 15, 35)	Minnow trap	targeted	1.8-2.6 repeated surveys at each site with 2-4 traps, 4 sample years (Apr.-Aug)
	Thames River (53)	9.2 m bag seine	stratified, random	3 repeated surveys at each site (92 m^2) during the same visit, 2-7 sites/reach[1] (May-Sep.)
Channel Darter	Ottawa River (15)	Missouri trawl	targeted	3 repeated 100 m trawl sets at each site during the same visit (Aug.)
	Lake St. Clair (36)	Siamese trawl	targeted	3 repeated 100 m trawl sets at each site during the same visit (Aug.-Oct.)
	Detroit River (6)	Missouri trawl	targeted (historic RAP sites)	2 x [5 x 100 m consecutive trawl sets] per site stratified by depth, each site sampled twice (Jul.-Sep.)
Pugnose Shiner	Old Ausable Channel (16)	61 m boat seine, hoop net, Windermere/minnow trap, Boat electrofisher	systematic	each 250 m site sampled with one spatial replicate of each gear type randomly assigned within the site (Sep.)
Lake Chubsucker	Old Ausable Channel (16)	61 m boat seine, hoop net, Windermere/minnow trap, Boat electrofisher	systematic	each 250 m site sampled with one spatial replicate of each gear type randomly assigned within the site (Sep.)

[1]Length of sample reach = 10 x stream width.

Table 3. Number of seine hauls required to detect Silver Shiner at 9.2 x 10 m sites in southwestern Ontario streams based on model-averaged estimate of detection probability ($\hat{p}$) and its lower 95% confidence limit.

Location	Number of repeat surveys required to detect Silver Shiner			
	confidence based on $\hat{p}$		confidence based on LCL of $\hat{p}$	
	95%	99%	95%	99%
Thames River	4	6	7	11
Lower Grand River	6	9	14	22
Middle Grand River	4	6	7	11
Bronte Creek	5	7	9	14
Sixteen Mile Creek	2	3	4	5

Table 4. Number of 9.2 x 10 m sites (sampled with three repeated seine hauls) required to detect Silver Shiner in reaches of southwestern Ontario streams based on model-averaged estimates of detection probability ($\hat{p}$) and small-scale occupancy ($\hat{\theta}$ – probability that site is occupied given reach occupancy) and their lower 95% confidence limits.

Location	**Number of sites required to detect Silver Shiner in reach**			
	confidence based on $\hat{p}$ and $\hat{\theta}$		**confidence based on LCL's of $\hat{p}$ and $\hat{\theta}$**	
	95%	**99%**	**95%**	**99%**
Thames River	5	8	13	19
Lower Grand River	6	9	15	23
Middle Grand River	5	7	10	15
Bronte Creek	4	6	12	18
Sixteen Mile Creek	5	8	10	16

Table 5. Number of repeated surveys required to detect Pugnose Shiner at sites in the Old Ausable Channel using various gears based on estimated detection probability ($\hat{p}$) and its lower 95% confidence limit.

Gear	No. of repeat surveys required to detect Pugnose			
	confidence based on $\hat{p}$		confidence based on LCL of $\hat{p}$	
	95%	99%	95%	99%
61 m boat seine	5	7	14	21
Hoop net	47	71	343	528
Windermere/minnow traps	47	71	343	528
Boat electrofisher (500 s)	2	3	4	7

Table 6. Number of repeated surveys required to detect Lake Chubsucker at sites in the Old Ausable Channel using various gears based on estimated detection probability ($\hat{p}$) and its lower 95% confidence limit.

Gear	No. of repeat surveys required to detect Lake Chubsucker			
	confidence based on $\hat{p}$		confidence based on LCL of $\hat{p}$	
	95%	99%	95%	99%
61 m boat seine	10	17	47	71
Hoop net	23	35	94	145
Windermere/minnow traps[1]	—	—	—	—
Boat electrofisher (500 s)	5	8	11	17

[1]No Lake Chubsucker were detected by Windermere/minnow trap combinations.

Table 7. Number of repeat surveys required to detect Northern Madtom using trawls and seines at sites in southwestern Ontario streams based on model-averaged estimate of detection probability ($\hat{p}$) and its lower 95% confidence limit.

Location	Gear and protocol	No. of repeat surveys required to detect Northern Madtom			
		confidence based on $\hat{p}$		confidence based on LCL of $\hat{p}$	
		95%	**99%**	**95%**	**99%**
St. Clair River	50 m trawl set	44	67	597	918
St. Clair River	3 x 50 m trawl sets	26	39	473	727
St. Clair River	100 m adjacent trawl sets	11	16	94	144
St. Clair River	10 x 100 m trawl sets	6	9	17	26
Detroit River	2 x [5 x 100 m] trawl sets	3	5	26	40
Detroit River	Minnow trap (3/site - 2003)	6	9	32	49
Detroit River	Minnow trap (4/site - 2009)	13	20	83	127
Detroit River	Minnow trap (3 /site - 2011)	10	15	62	96
Thames River	9.2 x 10 m seine haul	31	47	463	712
Thames River	[3 x 92 m^2] seine sites[1]	55	85	166	256

[1]Number of 92 m^2 sites sampled with three repeated seine hauls required to detect Northern Madtom in a reach.

Table 8. Number of repeat surveys required to detect Eastern Sand Darter using various gears at sites in southwestern Ontario streams based on model-averaged estimate of detection probability ($\hat{p}$) and its lower 95% confidence limit.

Location	Gear and protocol	No. of repeat surveys required to detect Eastern Sand Darter			
		confidence based on $\hat{p}$		confidence based on LCL of $\hat{p}$	
		95%	99%	95%	99%
Grand River	9.2 x 10 m seine haul	3	4	5	8
Thames River	9.2 x 10 m seine haul	6	9	8	12
Sydenham River	9.2 x 10 m seine haul	6	8	14	21
Sydenham River	~ 60 m BP efisher – 2002[1]	4	7	13	20
Sydenham River	~ 60 m BP efisher – 2003[1]	4	5	18	28
Sydenham River	~ 60 m seine - 2002	5	7	14	21
Sydenham River	~ 60 m seine - 2003	6	8	23	35
Grand River	100 m trawl set - 2010	3	4	5	7
Grand River	100 m trawl set - 2011	3	4	4	6

[1]BP efisher – backpack electrofisher.

Table 9. Number of 9.2 x 10 m sites (sampled with three repeated seine hauls) required to detect Eastern Sand Darter in reaches of Grand and Thames rivers based on model-averaged estimates of detection probability ($\hat{p}$) and small-scale occupancy ($\hat{\theta}$ – probability that site is occupied given reach occupancy) and their lower 95% confidence limits.

Location	**Number of sites required to detect Eastern Sand Darter in reach**			
	confidence based on $\hat{p}$ and $\hat{\theta}$		**confidence based on LCL's of $\hat{p}$ and $\hat{\theta}$**	
	95%	**99%**	**95%**	**99%**
Grand River	5	7	10	15
Thames River	3	5	5	8

Table 10. Number of repeat surveys required to detect Channel Darter using trawls at sites based on model-averaged estimate of detection probability ($\hat{p}$) and its lower 95% confidence limit.

Location	Length of trawl set	No. of repeat surveys required to detect Channel Darter			
		confidence based on $\hat{p}$		confidence based on LCL of $\hat{p}$	
		95%	99%	95%	99%
Ottawa River	100 m	3	4	5	7
Detroit River	1000 m (2 x [5 x 100 m])	2	3	5	8

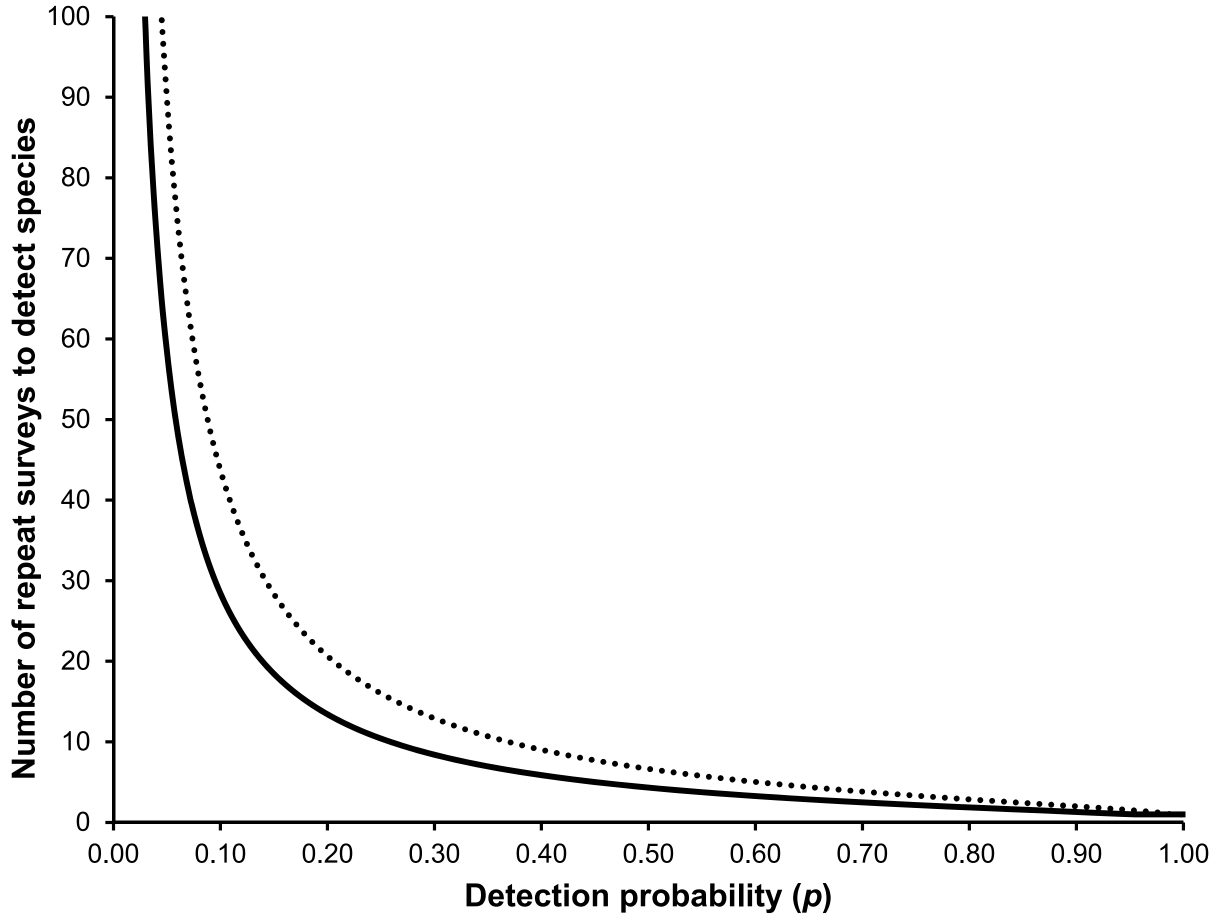


Fig.1. Relationship between detection probability (*p*) in a single survey at a site and the number of repeat surveys required to detect the species given the site is occupied by the species (solid line – 95% confidence in detecting species; dashed line – 99% confidence in detecting species).

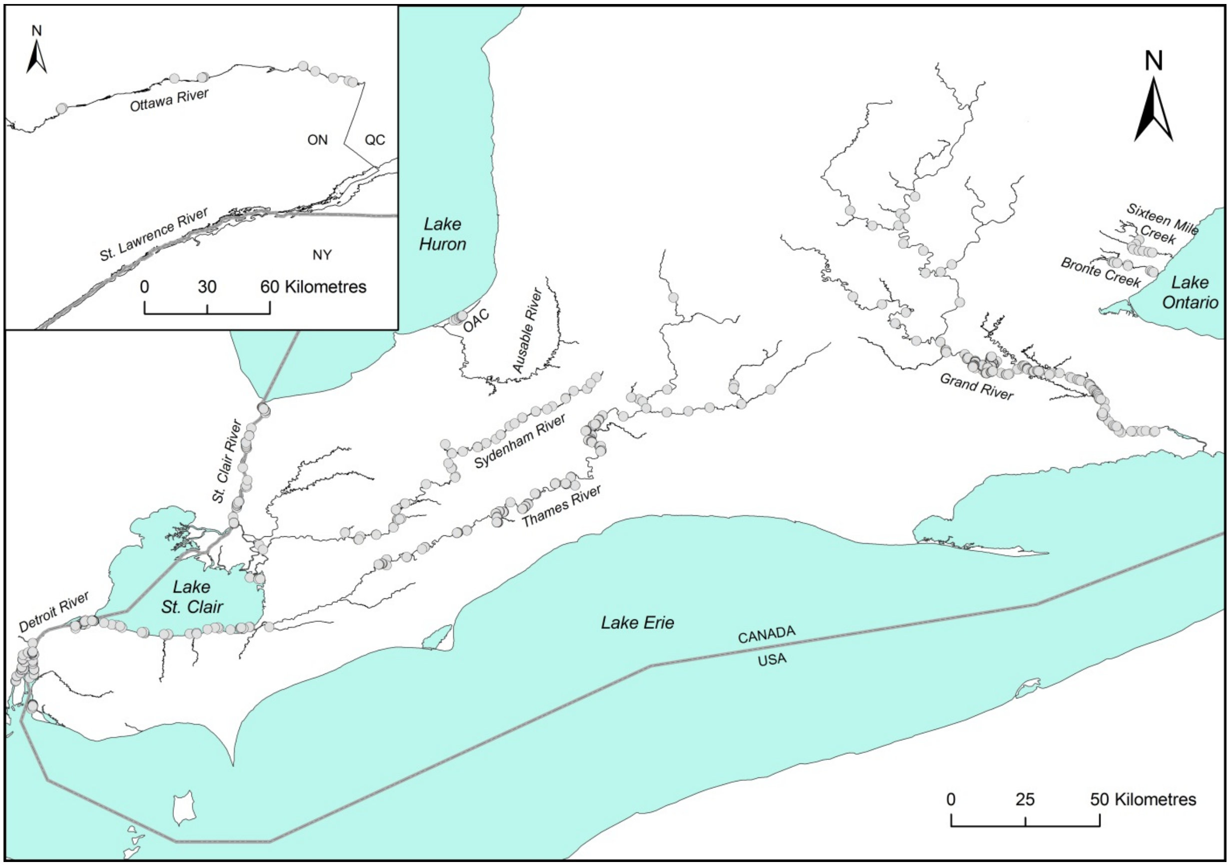


Fig. 2. Locations of sites sampled in southern Ontario for projects used to develop occupancy models for fish species at risk (OAC – Old Ausable Channel).

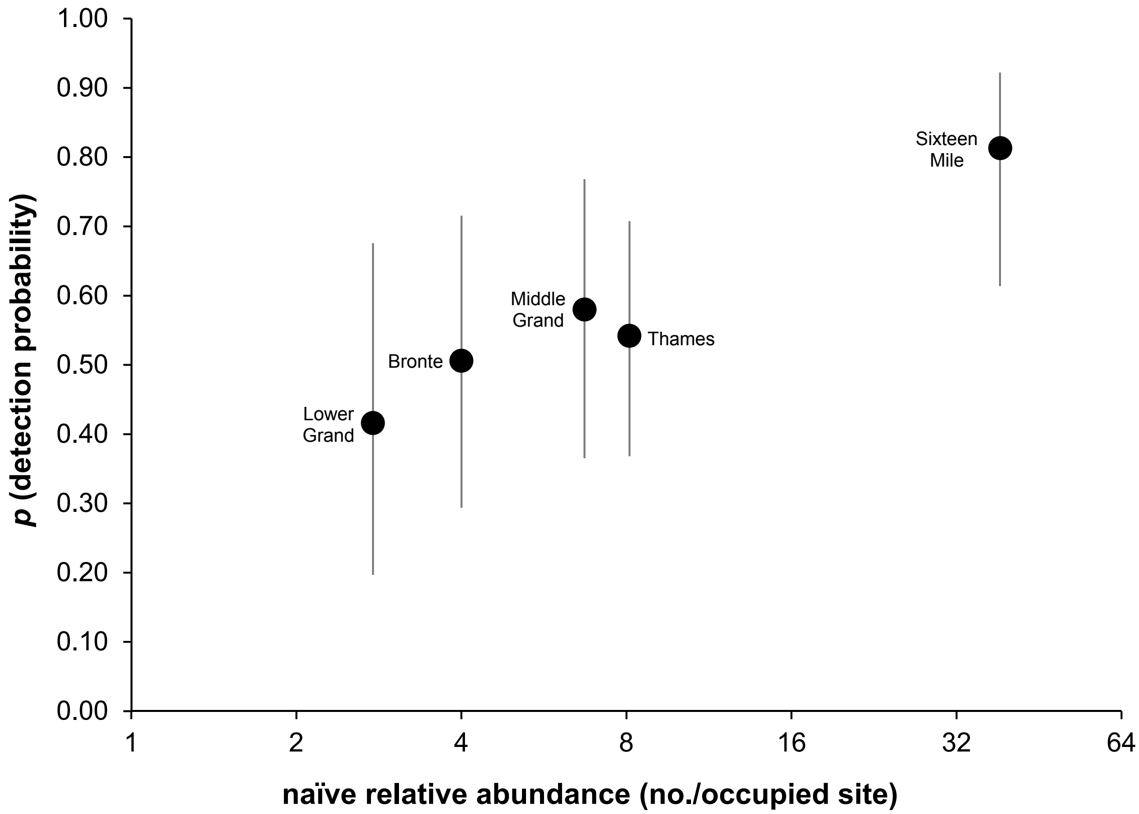


Fig. 3. Plot of naïve relative abundance and detection probability (*p*) for Silver Shiner sampled using seine nets at 9.2 x 10 m sites in southwestern Ontario streams (error bars are 95% confidence limits of estimate).

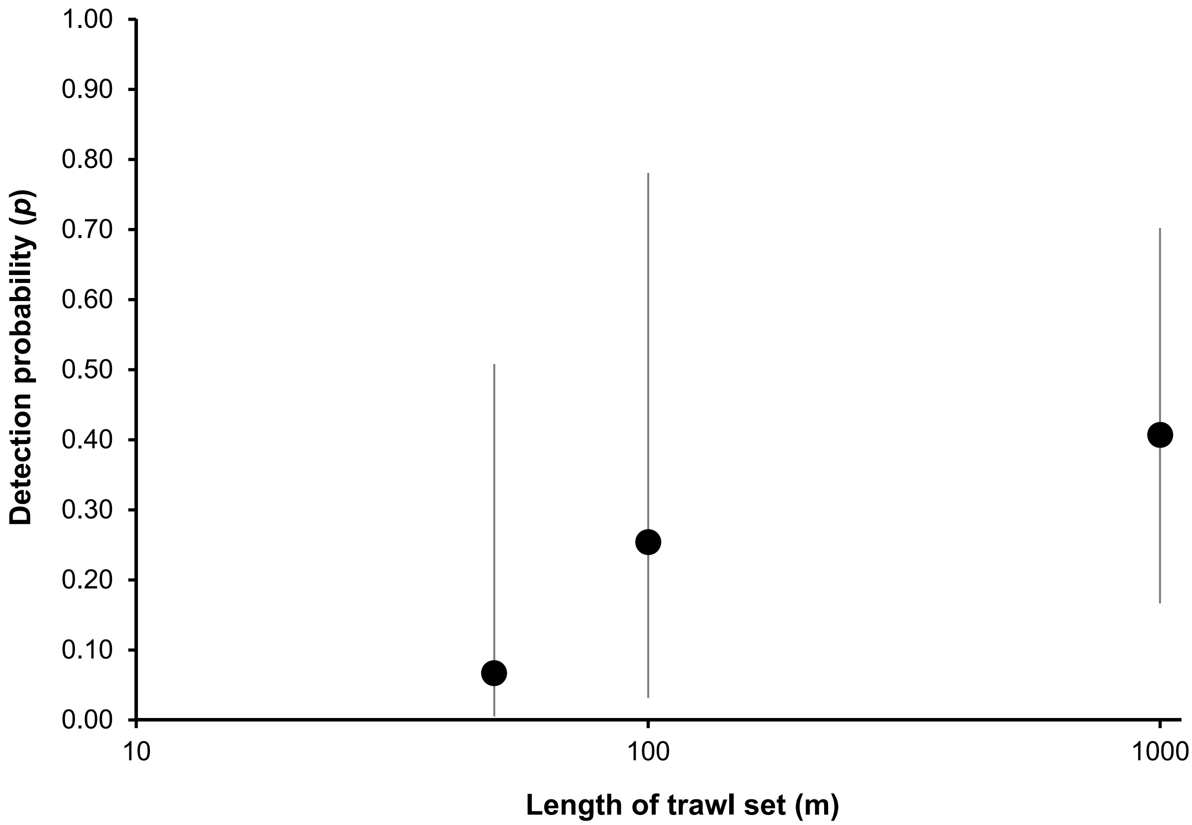


Fig. 4. Detection probability (*p*) of Northern Madtom in trawl sets of different lengths in the St. Clair River (error bars are 95% confidence limits of estimate).

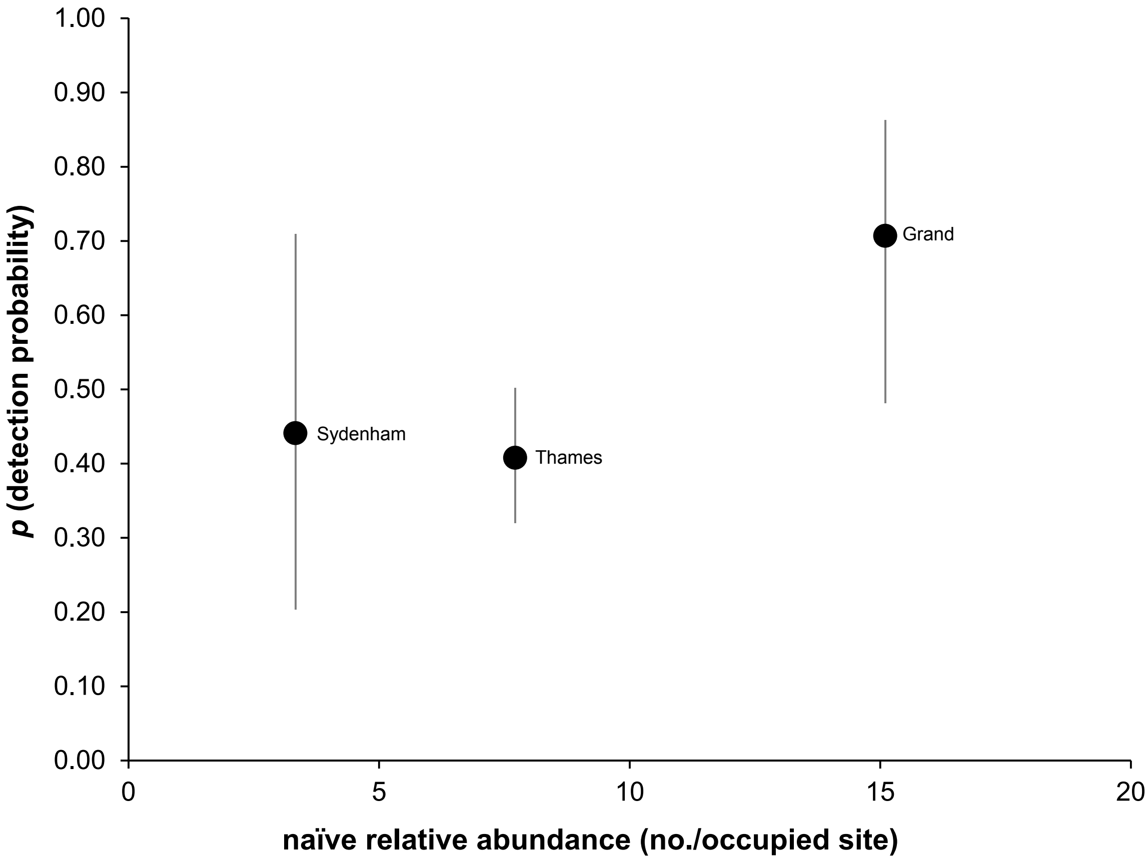


Fig. 5. Plot of naïve relative abundance and detection probability (*p*) for Eastern Sand Darter at 9.2 x 10 m sites in southwestern Ontario streams (error bars are 95% confidence limits of estimate).

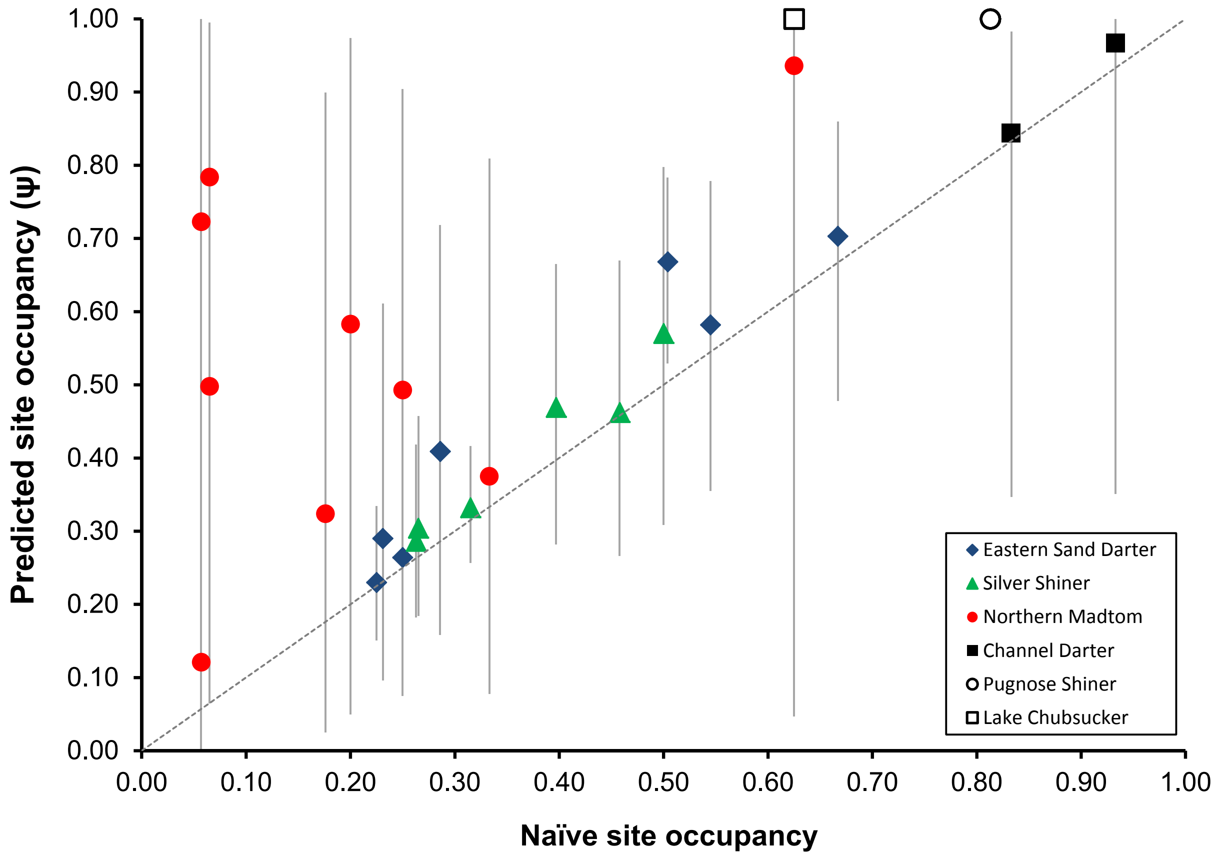


Fig. 6. Relationship between observed (naïve) site occupancy and model-averaged predicted site occupancy based on multinomial likelihood models (error bars represent 95% confidence limit of estimates; dashed line represents equivalency of naïve and predicted site occupancies).

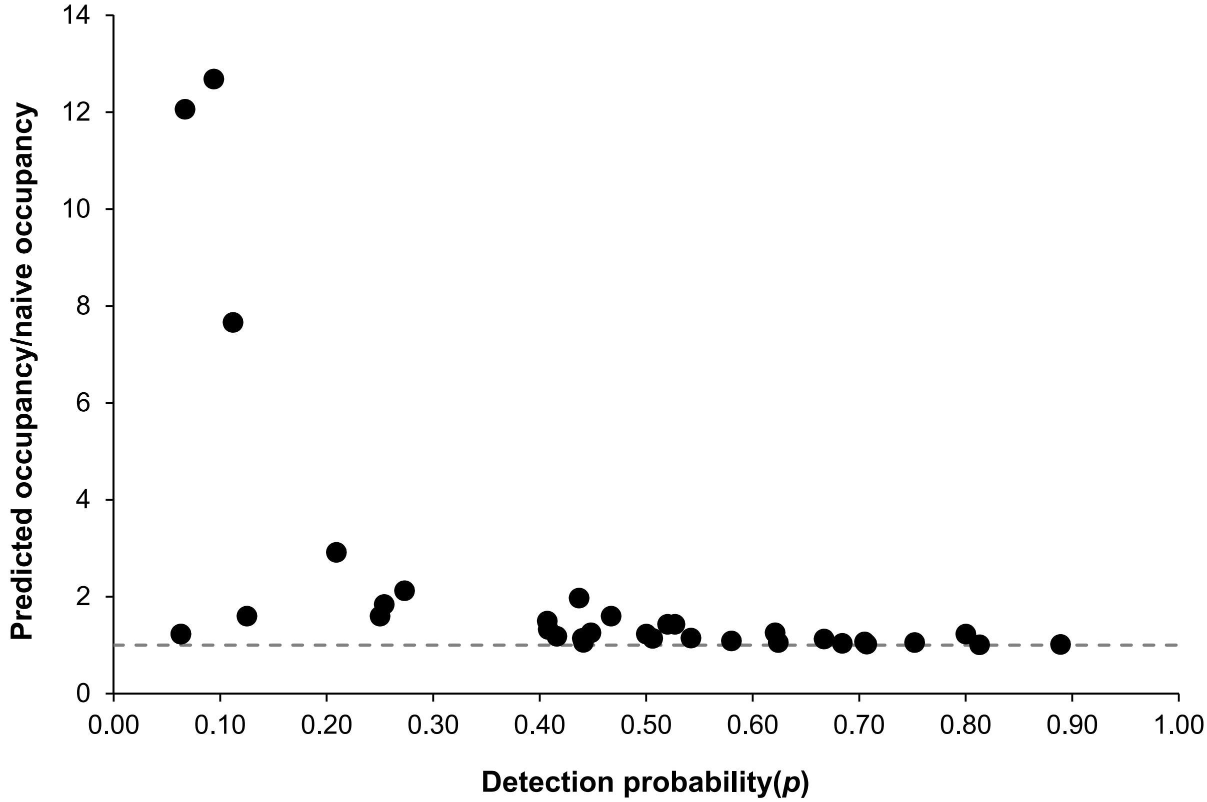


Fig. 7. Relationship between detection probability and the ratio of predicted occupancy to observed (naïve) occupancy (dashed line represents equivalence of predicted and observed occupancies).

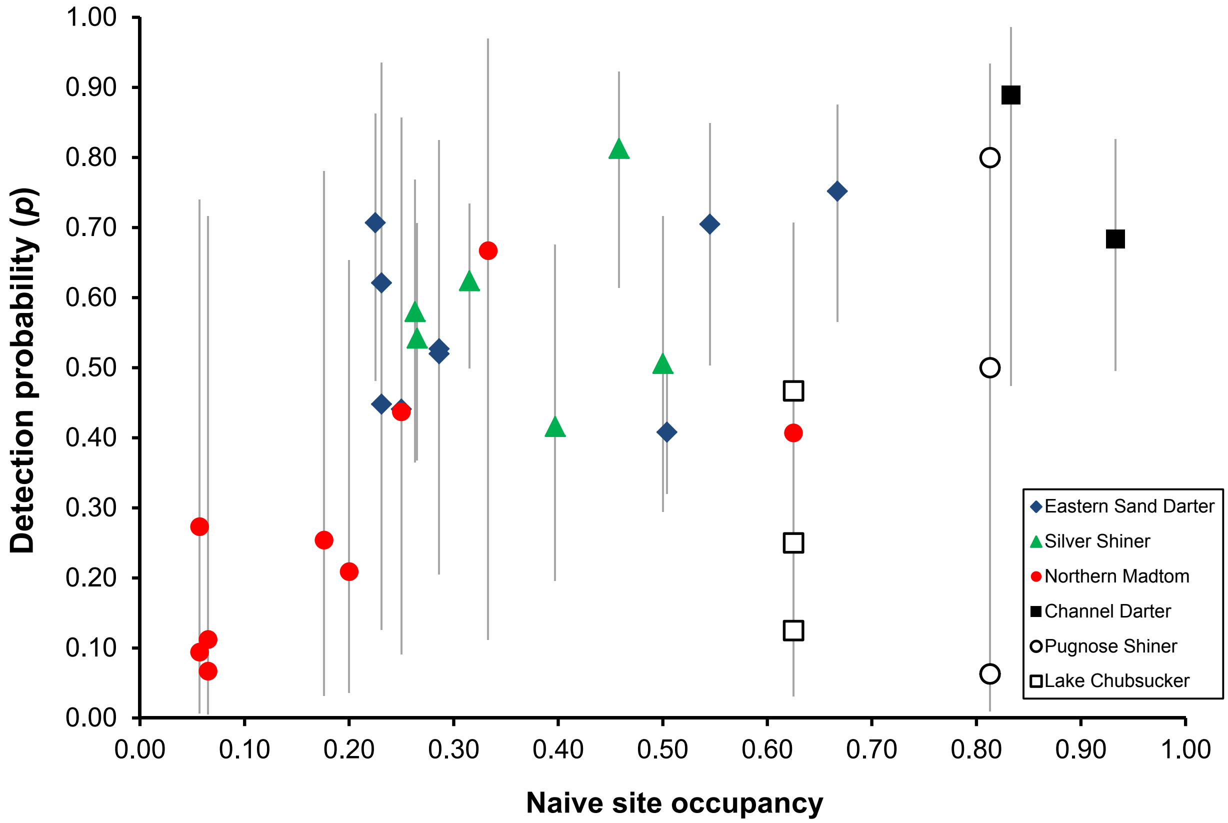


Fig. 8. Relationship between observed (naïve) site occupancy and model-averaged detection probability based on multinomial likelihood models (error bars represent 95% confidence limit of estimates).

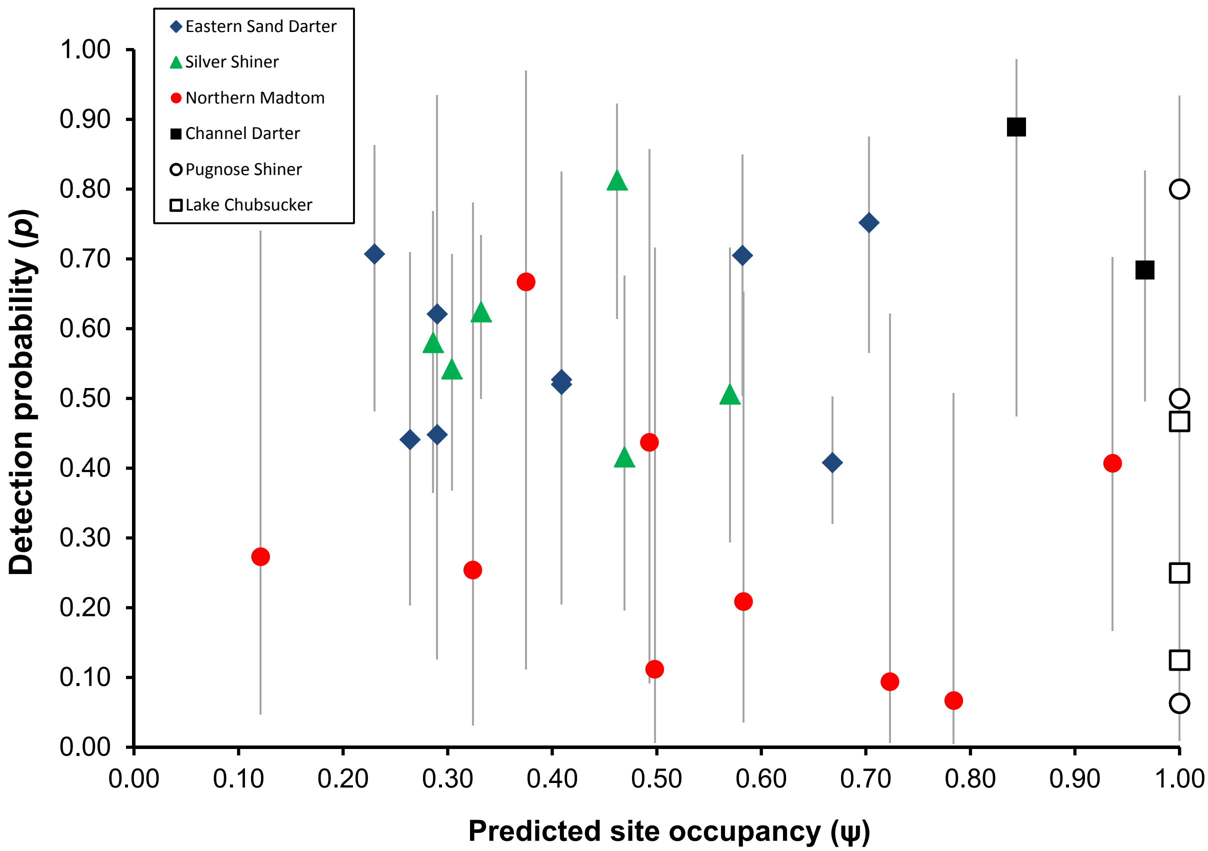


Fig. 9. Relationship between model-averaged predicted site occupancy and detection probability based on multinomial likelihood models (error bars represent 95% confidence limit of estimates).

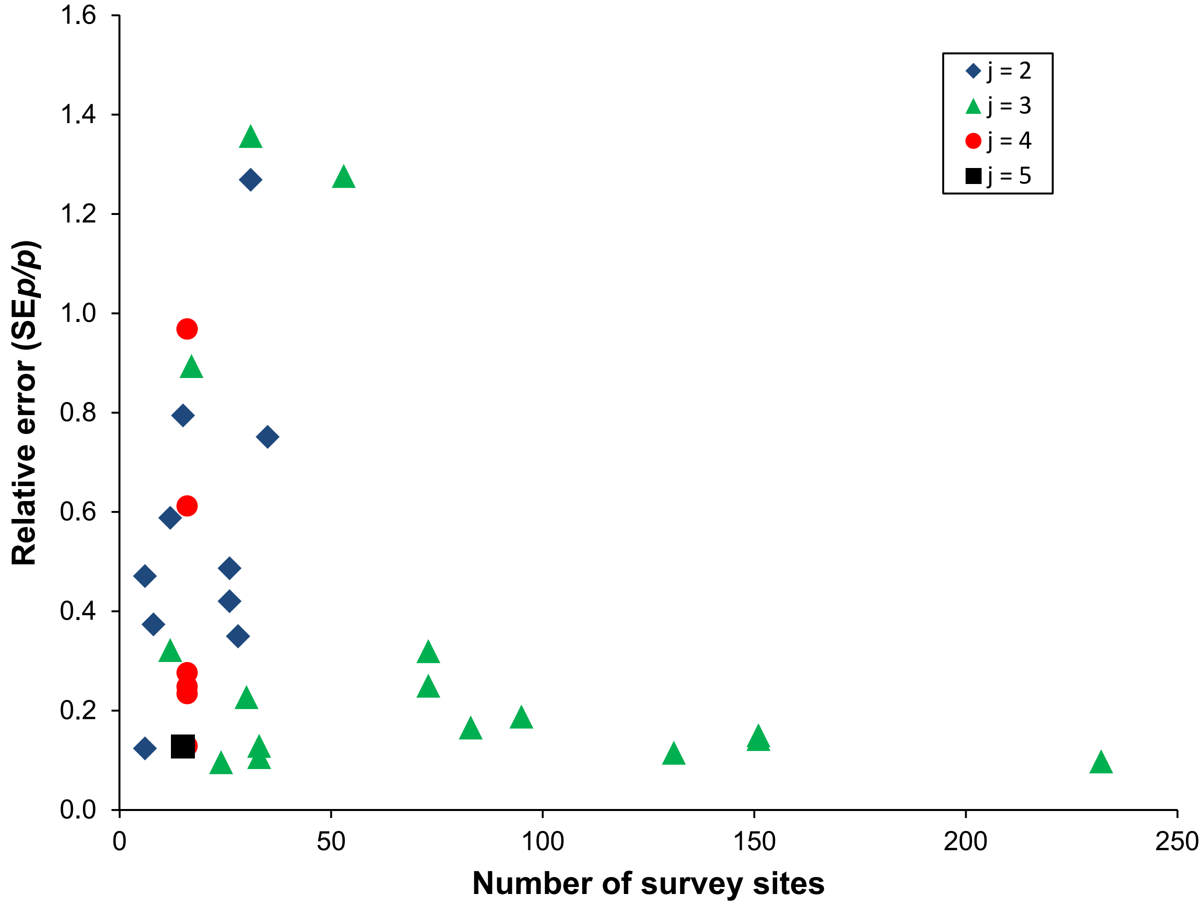


Fig. 10. Relationship between number of sites sampled to build site occupancy models and relative error of detection probability estimates (*j* = number of repeat surveys per site).

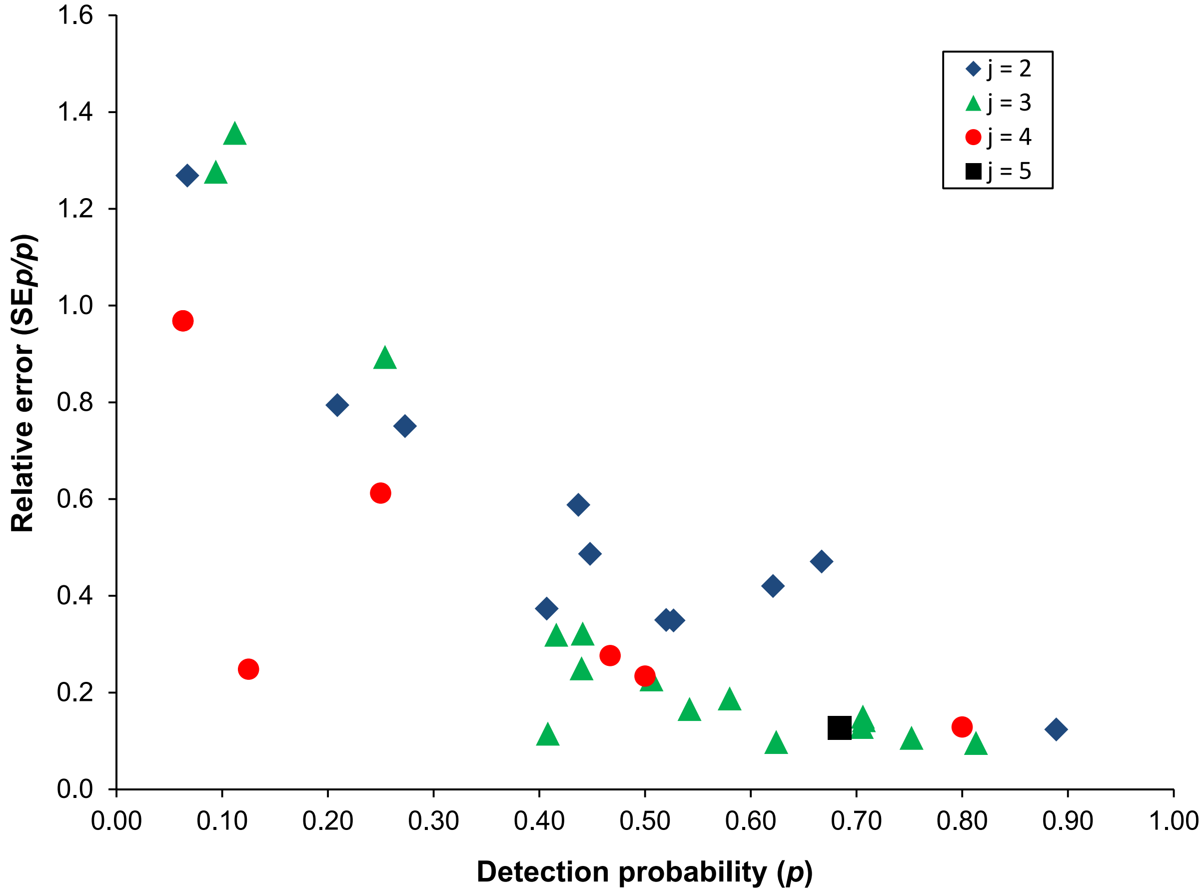


Fig. 11. Relationship between estimated detection probability and relative error of detection probability estimates (*j* = number of repeat surveys per site).

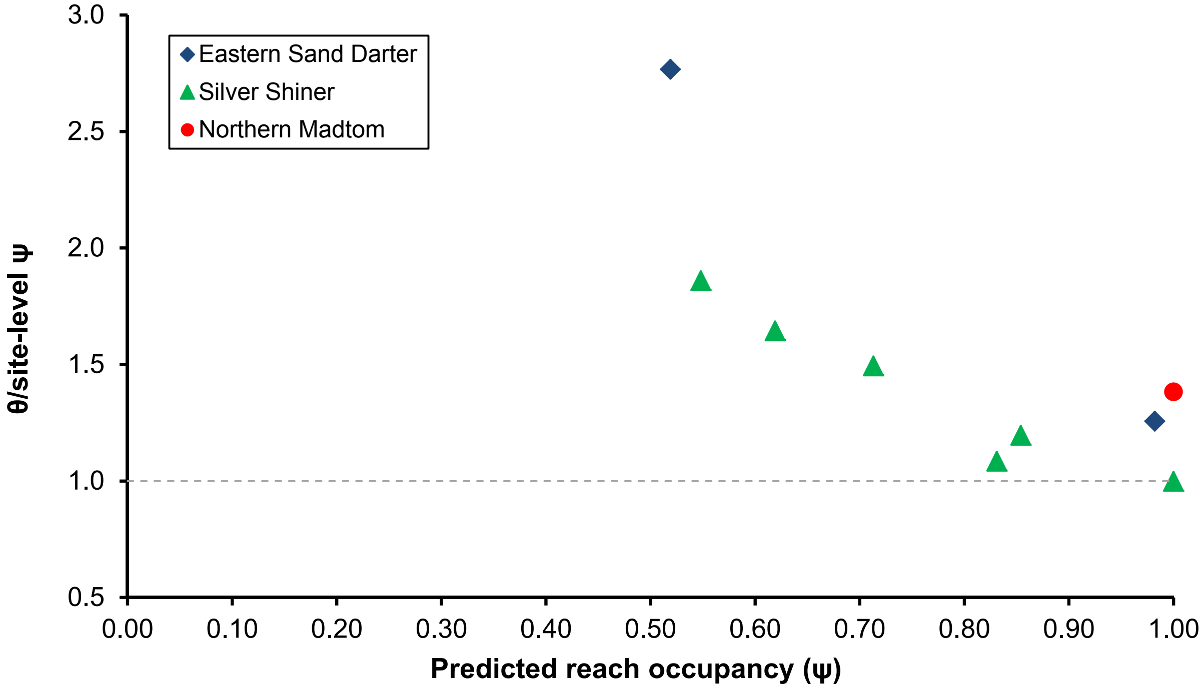


Fig. 12. Relationship between predicted reach occupancy from multi-scale occupancy models and the ratio of small-scale (site) occupancy (θ) from reach models to predicted site-level occupancy (ψ) from site occupancy models based on the same data (dashed line represents equivalency of θ and ψ).

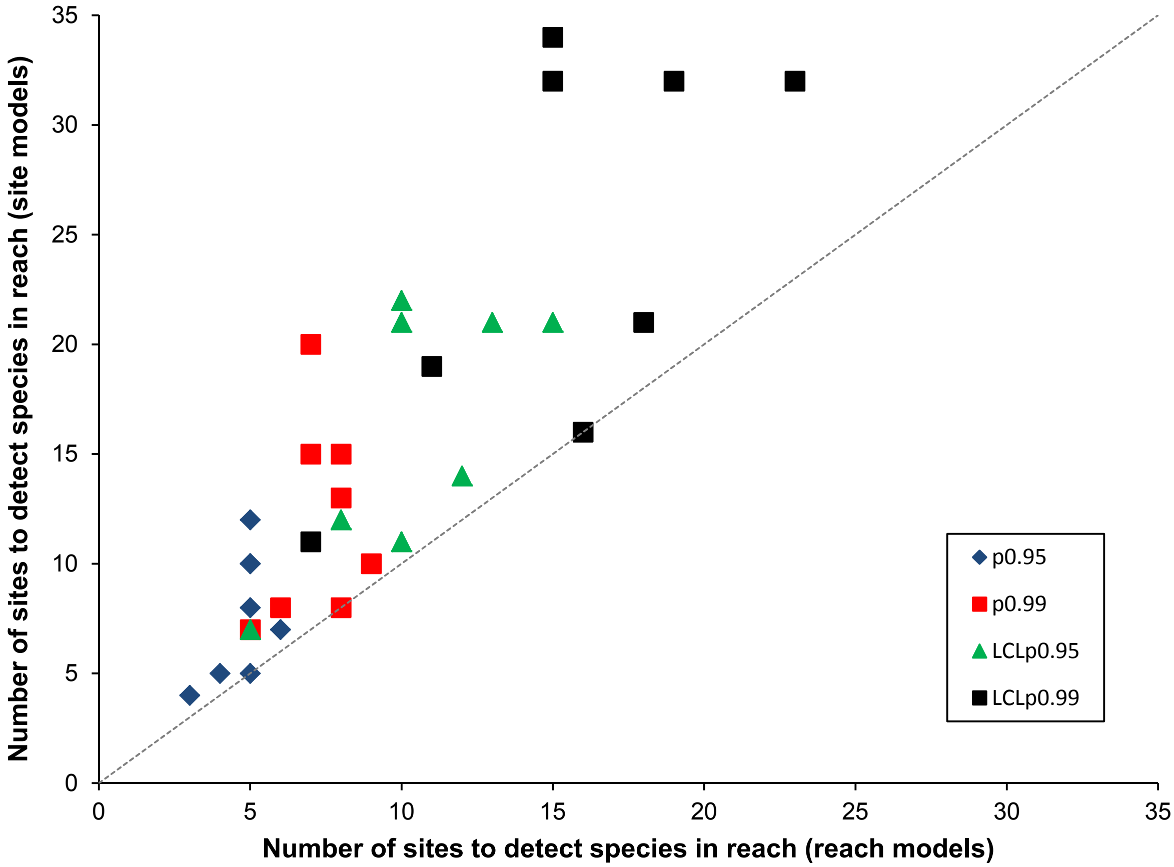


Fig. 13. Comparison of sampling effort (number of 92 m^2 sites with 3 seine hauls) required to detect Eastern Sand Darter and Silver Shiner at the reach level based on multi-scale reach occupancy models and site occupancy models based on the same data (for each model, survey effort was calculated for probabilities of detection of 0.95 and 0.99 based on the estimate of *p* and its lower 95% confidence limit; dashed line represents equivalency of number of sites based on reach and site models).

Appendix 1. Specifications for netting gear used in field surveys.

Gear	Project	Net dimensions (m)	Mesh size (mm)
9.2 m bag seine	Eastern Sand Darter (Grand River, Thames River), Northern Madtom (Thames River)	wings – 1.8 x 3.7 bag – 1.8 x 1.8 x 1.8	wings – 6.4 bag – 3.2
	Eastern Sand Darter (Sydenham River)	wings – 1.8 x 3.7 bag – 1.8 x 1.8 x 1.8	3.2
8.2 m bag seine	Eastern Sand Darter (Sydenham 2002, 2003)	wings – 2.0 x 3.1 bag – 2.0 x 2.0 x 2.0	7.5
61 m boat seine	Pugnose Shiner, Lake Chubsucker (Old Ausable Channel)	1.8 x 61	12.5
Hoop net	Pugnose Shiner, Lake Chubsucker (Old Ausable Channel)	hoops – 0.9 wings – 0.9 x 3.0 lead – 0.9 x 4.5	6.4
Missouri trawl	Northern Madtom (Detroit River, St. Clair River 2010), Channel Darter (Detroit River, Ottawa River)	length – 4.0 throat – 2.5 head and foot ropes – 3.2 door width – 0.35 door length – 0.53	internal – 19 external – 3
Siamese trawl	Northern Madtom (St. Clair River 2012), Channel Darter (Lake St. Clair)	length – 5.0 throat – 3.0 head and foot ropes – 3.5 door width – 0.31 door length – 0.60	internal – 19 external – 3

Appendix 2. Ranking of hypothesized occupancy and detection covariates (mean rank in parenthesis, covariates in bold were included in candidate sets).

Species	Occupancy	Detection			
		Electrofishing	Trawling	Seining	Traps
Channel Darter	• **substrate** (1.6) • **water velocity** (1.8) • **depth** (3.4) • goby density (3.6) • water clarity (3.8) • water temp (3.8)		• **substrate** (1.3) • **depth** (3.3) • trawl speed (3.5) • **water velocity** (3.8)		
Eastern Sand Darter	• **substrate** (1.0) • **water velocity** (2.4) • **water clarity** (3.2) • goby density (3.6) • water temp (3.8)	• **water clarity** (2.0) • **conductivity** (2.4) • **depth** (3.4) • water velocity (3.4) • power (3.4) • substrate (3.6)	• **substrate** (1.0) • **depth** (3.3) • **water velocity** (3.3) • **trawl speed** (3.5)	• **substrate** (1.6) • **depth** (2.0) • **water velocity** (3.2) • **cover** (3.8)	
Lake Chubsucker	• water clarity (1.4) • macrophytes (1.6) • water velocity (3.6) • substrate (3.8) • dissolved oxygen (3.8)	• conductivity (2.8) • macrophytes (2.8) • depth (3.0) • water clarity (3.4) • power (3.4) • water temp (3.8)		• macrophytes (1.6) • depth (2.2) • substrate (3.0) • water clarity (3.4) • water temp (3.8)	• water clarity (2.4) • water temp (2.4) • macrophytes (3.4) • soak time (3.4) • water velocity (3.8)
Northern Madtom	• **water velocity** (1.6) • **substrate** (1.8) • **macrophytes** (3.2) • goby density (3.6) • depth (3.8)		• **substrate** (1.3) • **depth** (3.3) • **water velocity** (3.5) • **trawl speed** (3.5)	• **depth** (1.8) • **substrate** (2.0) • **water velocity** (3.2) • macrophytes (3.8) • cover (3.8)	
Pugnose Shiner	• macrophytes (1.4) • water clarity (1.6) • water velocity (3.6) • depth (3.6) • dissolved oxygen (3.6) • water temp (3.6)			• macrophytes (1.4) • depth (2.2) • substrate (3.2) • water clarity (3.4)	• water clarity (2.4) • water temp (2.6) • macrophytes (3.2) • mesh size (3.2) • soak time (3.6) • depth (3.8)
Silver Shiner	• **water velocity** (1.4) • **substrate** (2.6) • **depth** (2.6) • **water clarity** (3.2) • **macrophytes** (3.2)			• **depth** (1.6) • **substrate** (2.6) • **water velocity** (2.6)	

Appendix 3. Summary of occupancy modelling and detection probability analysis for Ontario fish species at risk.

Species	Location	Year	Scale	Spatial	Survey technique	n	naïve ψ	$\hat{\psi}$ (SE)	$\hat{p}$ (SE)	Number of repeat surveys (or sites) required to detect species at site (or reach): based on $\hat{p}$		based on LCL of $\hat{p}$	
										95%	99%	95%	99%
Notropis photogenis	Grand River	2007	Site	No	9.2 x 10 m seine haul	73	0.397	0.453 (0.103)	0.440 (0.110)	6	8	11	17
	Grand River	2007	Site	Yes	9.2 x 10 m seine haul	73	0.397	0.469 (0.103)	0.416 (0.133)	6	9	14	22
	Grand River	2007	Reach	No	1 seine hauls/site	14	0.714	0.831 (0.144)	0.438 (0.110)	12	19	36	54
	Grand River	2007	Reach	No	2 seine hauls/site	14	0.714	0.831 (0.144)	0.438 (0.110)	7	11	20	31
	Grand River	2007	Reach	No	3 seine hauls/site	14	0.714	0.831 (0.144)	0.438 (0.110)	6	9	15	23
	Four streams[1]	2011	Site	No	9.2 x 10 m seine haul	232	0.315	0.332 (0.041)	0.624 (0.061)	4	5	5	7
	Four streams	2011	Reach	No	1 seine haul/site	54	0.667	0.701 (0.080)	0.635 (0.063)	8	13	14	21
	Four streams	2011	Reach	No	2 seine hauls/site	54	0.667	0.701 (0.080)	0.635 (0.063)	6	9	9	13
	Four streams	2011	Reach	No	3 seine hauls/site	54	0.667	0.701 (0.080)	0.635 (0.063)	5	8	7	11
	Four streams	2011	Reach	Yes	1 seine haul/site	54	0.667	0.713 (0.096)	0.635 (0.063)	9	13	14	21
	Four streams	2011	Reach	Yes	2 seine hauls/site	54	0.667	0.713 (0.096)	0.635 (0.063)	6	9	9	13
	Four streams	2011	Reach	Yes	3 seine hauls/site	54	0.667	0.713 (0.096)	0.635 (0.063)	5	8	8	11
	Thames River	2011	Site	No	9.2 x 10 m seine haul	83	0.265	0.304 (0.071)	0.542 (0.090)	4	6	7	11
	Thames River	2011	Reach	No	1 seine haul/site	17	0.588	0.619 (0.147)	0.540 (0.101)	10	15	28	42
	Thames River	2011	Reach	No	2 seine hauls/site	17	0.588	0.619 (0.147)	0.540 (0.101)	6	10	16	25
	Thames River	2011	Reach	No	3 seine hauls/site	17	0.588	0.619 (0.147)	0.540 (0.101)	5	8	13	19
	Grand River	2011	Site	No	9.2 x 10 m seine haul	95	0.263	0.286 (0.061)	0.580 (0.109)	4	6	7	11
	Grand River	2011	Reach	No	1 seine haul/site	19	0.526	0.548 (0.113)	0.551 (0.090)	9	14	21	32
	Grand River	2011	Reach	No	2 seine hauls/site	19	0.526	0.548 (0.113)	0.551 (0.090)	6	9	13	19
	Grand River	2011	Reach	No	3 seine hauls/site	19	0.526	0.548 (0.113)	0.551 (0.090)	5	7	10	15
	Bronte Creek	2011	Site	No	9.2 x 10 m seine haul	30	0.500	0.570 (0.136)	0.506 (0.115)	5	7	9	14
	Bronte Creek	2011	Reach	No	1 seine haul/site	10	0.800	0.854(0.148)	0.484 (0.102)	8	12	27	42
	Bronte Creek	2011	Reach	No	2 seine hauls/site	10	0.800	0.854(0.148)	0.484 (0.102)	5	7	16	24
	Bronte Creek	2011	Reach	No	3 seine hauls/site	10	0.800	0.854(0.148)	0.484 (0.102)	4	6	12	18
	Sixteen Mile Cr.	2011	Site	No	9.2 x 10 m seine haul	24	0.458	0.462 (0.109)	0.813 (0.078)	2	3	4	5
	Sixteen Mile Cr.	2011	Reach	No	1 seine haul/site	8	1	1 (0)	0.813 (0.071)	7	10	16	24
	Sixteen Mile Cr.	2011	Reach	No	2 seine hauls/site	8	1	1 (0)	0.813 (0.071)	6	8	11	17
	Sixteen Mile Cr.	2011	Reach	No	3 seine hauls/site	8	1	1 (0)	0.813 (0.071)	5	8	10	16

Species	Location	Year	Scale	Spatial	Survey technique	*n*	naïve ψ	$\hat{\psi}$ (SE)	$\hat{p}$ (SE)	Number of repeat surveys (or sites) required to detect species at site (or reach): based on $\hat{p}$		based on LCL of $\hat{p}$	
										95%	99%	95%	99%
Notropis anogenus	Ausable River	2002	Site	No	200’ boat seine	16	0.813	1 (0)	0.500 (0.177)	5	7	14	21
	Ausable River	2002	Site	No	Hoop net	16	0.813	1 (0)	0.063 (0.061)	47	71	343	528
	Ausable River	2002	Site	No	Windermere/minnow trap	16	0.813	1 (0)	0.063 (0.061)	47	71	343	528
	Ausable River	2002	Site	No	Boat electrofisher (500 s)	16	0.813	1 (0)	0.800 (0.103)	2	3	4	7
Erimyzon sucetta	Ausable River	2002	Site	No	200’ boat seine	16	0.625	1 (0)	0.250 (0.153)	10	17	47	71
	Ausable River	2002	Site	No	Hoop net	16	0.625	1 (0)	0.125 (0.031)	23	35	94	145
	Ausable River	2002	Site	No	Windermere/minnow trap	16	0.625	1 (0)	0.000 (0.000)	—	—	—	—
	Ausable River	2002	Site	No	Boat electrofisher (500 s)	16	0.625	1 (0)	0.467 (0.129)	5	8	11	17
Noturus stigmosus	St. Clair River	2012	Site	No	50 m trawl set	31	0.065	0.784 (0.342)	0.067 (0.085)	44	67	597	918
	St. Clair River	2012	Site	No	3 x 50 m trawl sets	31	0.065	0.498 (0.309)	0.112 (0.152)	26	39	473	727
	St. Clair River	2010	Site	No	100 m adjacent trawl sets	17	0.176	0.324 (0.327)	0.254 (0.227)	11	16	94	144
	St. Clair River	2012	Site	No	10 x 100 m trawl sets	8	0.625	0.936 (0.174)	0.407 (0.152)	6	9	17	26
	Detroit River	2011	Site	No	2 x [5 x 100 m] trawl sets	6	0.333	0.375 (0.234)	0.667 (0.314)	3	5	26	40
	Detroit River	2003	Site	No	Minnow trap (3 traps/site)	12	0.250	0.493 (0.290)	0.437 (0.257)	6	9	32	49
	Detroit River	2008	Site	No	Minnow trap (2 traps/site)	25	0.080	lack of model fit		—	—	—	—
	Detroit River	2009	Site	No	Minnow trap (4 traps/site)	15	0.200	0.583 (0.407)	0.209 (0.166)	13	20	83	127
	Detroit River	2009	Site	No	Minnow trap (3 traps/site)	35	0.057	0.121 (0.106)	0.273 (0.205)	10	15	62	96
	Thames River	2006	Site	No	9.2 x 10 m seine haul	53	0.057	0.723 (0.713)	0.094 (0.120)	31	47	463	712
	Thames River	2006	Reach	No	1 seine haul/site	12	0.250	1 (0)	0.018 (0.010)	165	254	498	766
	Thames River	2006	Reach	No	2 seine hauls/site	12	0.250	1 (0)	0.018 (0.010)	83	127	249	383
	Thames River	2006	Reach	No	3 seine hauls/site	12	0.250	1 (0)	0.018 (0.010)	55	85	166	256
Ammocrypta pellucida	Grand River	2007	Site	No	9.2 x 10 m seine haul	151	0.225	0.234 (0.061)	0.706 (0.106)	3	4	5	8
	Grand River	2007	Site	Yes	9.2 x 10 m seine haul	151	0.225	0.230 (0.047)	0.707 (0.101)	3	4	5	8
	Thames River	2006	Site	No	9.2 x 10 m seine haul	131	0.504	0.668 (0.066)	0.408 (0.047)	6	9	8	12
	Sydenham River	2012	Site	No	9.2 x 10 m seine haul	12	0.250	0.264 (0.147)	0.441 (0.142)	6	8	14	21
	Grand River	2007	Reach	No	1 seine hauls/site	31	0.452	0.519 (0.116)	0.382 (0.052)	11	17	24	36
	Grand River	2007	Reach	No	2 seine hauls/site	31	0.452	0.519 (0.116)	0.382 (0.052)	6	10	13	20
	Grand River	2007	Reach	No	3 seine hauls/site	31	0.452	0.519 (0.116)	0.382 (0.052)	5	7	10	15
	Thames River	2006	Reach	No	1 seine hauls/site	30	0.867	0.982 (0.066)	0.398 (0.039)	8	12	12	19

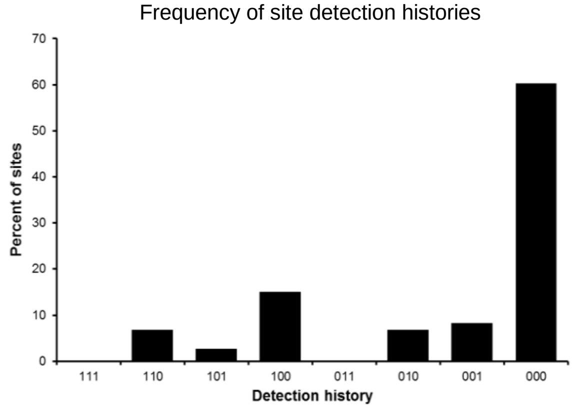
Species	Location	Year	Scale	Spatial	Survey technique	*n*	naïve ψ	$\hat{\psi}$ (SE)	$\hat{p}$ (SE)	Number of repeat surveys (or sites) required to detect species at site (or reach): based on $\hat{p}$ 95%	based on $\hat{p}$ 99%	based on LCL of $\hat{p}$ 95%	based on LCL of $\hat{p}$ 99%
Ammocrypta pellucida	Thames River	2006	Reach	No	2 seine hauls/site	30	0.867	0.982 (0.066)	0.398 (0.039)	4	7	7	10
	Thames River	2006	Reach	No	3 seine hauls/site	30	0.867	0.982 (0.066)	0.398 (0.039)	3	5	5	8
	Sydenham River	2002	Site	No	60 m - BP electrofisher	28	0.286	0.409 (0.161)	0.527 (0.184)	4	7	13	20
	Sydenham River	2002	Site	No	60 m - seine	28	0.286	0.409 (0.161)	0.520 (0.182)	5	7	14	21
	Sydenham River	2003	Site	No	60 m - BP electrofisher	26	0.231	0.290 (0.142)	0.621 (0.261)	4	5	18	28
	Sydenham River	2003	Site	No	60 m - seine	26	0.231	0.290 (0.142)	0.448 (0.218)	6	8	23	35
	Grand River	2010	Site	No	100 m trawl set	33	0.545	0.582 (0.115)	0.705 (0.091)	3	4	5	7
	Grand River	2010	Site	Yes	100 m trawl set	33	0.545	0.580 (0.134)	0.711 (0.086)	3	4	5	7
	Grand River	2011	Site	No	100 m trawl set	33	0.667	0.703 (0.101)	0.752 (0.080)	3	4	4	6
Percina copelandi	Ottawa River	2011	Site	No	100 m trawl set	15	0.933	0.967 (0.065)	0.684 (0.087)	3	4	5	7
	Lake St. Clair	2012	Site	No	100 m trawl set	36	0.028	did not reach convergence		—	—	—	—
	Detroit River	2011	Site	No	2 x [5 x 100 m] trawl sets	6	0.833	0.844 (0.156)	0.889 (0.110)	2	3	5	8

[1]Combined analysis for four streams sampled in 2011 (Bronte Creek, Grand River, Sixteen Mile Creek, Thames River).

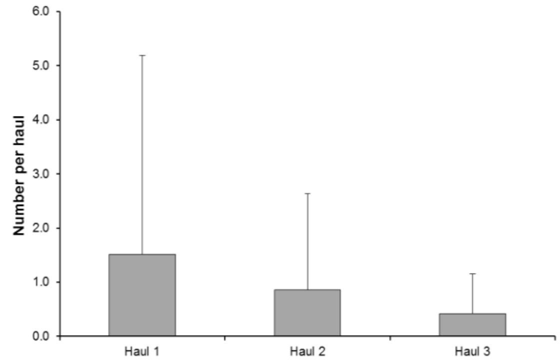
Appendix 4. Details of occupancy modelling and detection probability analysis for an individual project (Silver Shiner in lower Grand River).

Silver Shiner – seining at 92 m² sites in lower Grand River (May – July 2007), *n* = 73, *j* = 3

Frequency of site detection histories



Number of individuals detected in each haul at occupied sites



Summary of model selection procedure for Silver Shiner occupancy at sites in the lower Grand River based on non-spatial multinomial likelihood (n = 73, χ^2 = 12.8, p = 0.02, $\hat{c}$ = 2.57, naïve ψ = 0.397).

Model	QAICc	ΔQAICc	AIC weight	Number of parameters	-2[Log Likelihood]
ψ(macrophyte),p(removal)	82.80	0.00	0.1804	5	182.44
ψ(.),p(removal)[1]	83.51	0.71	0.1265	4	190.11
ψ(silt/clay+macrophyte),p(removal)	83.86	1.06	0.1062	6	179.11
ψ(silt/clay),p(removal)	84.53	1.73	0.0759	5	186.83
ψ(macrophyte+velocity),p(removal)	84.72	1.92	0.0691	6	181.29
ψ(velocity),p(removal)	85.01	2.21	0.0597	5	188.06
ψ(depth+macrophyte),p(removal)	85.10	2.30	0.0571	6	182.25
ψ(macrophyte+water clarity),p(removal)	85.17	2.37	0.0551	6	182.43
ψ(depth),p(removal)	85.38	2.58	0.0496	5	189.00
ψ(water clarity),p(removal)	85.80	3.00	0.0402	5	190.06
ψ(silt/clay+depth+macrophyte),p(removal)	86.27	3.47	0.0318	7	179.01
ψ(silt/clay+water clarity+macrophyte),p(removal)	86.31	3.51	0.0312	7	179.10
ψ(silt/clay+depth),p(removal)	86.53	3.73	0.0279	6	185.87
ψ(depth+velocity),p(removal)	86.65	3.85	0.0263	6	186.19
ψ(silt/clay+water clarity),p(removal)	86.83	4.03	0.0240	6	186.63
ψ(macrophyte+velocity+water clarity),p(removal)	87.17	4.37	0.0203	7	181.28
ψ(water clarity+velocity),p(removal)	87.35	4.55	0.0185	6	187.96

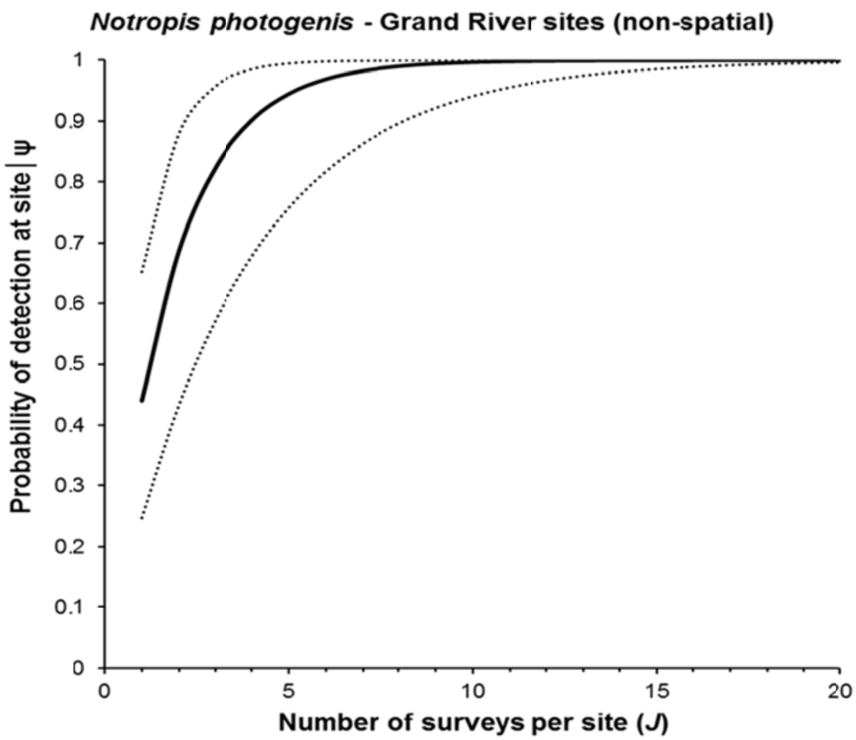
[1]Model for constant occupancy across sites.

Silver Shiner – seining at 92 m^2 sites in lower Grand River (May – July 2007), n = 73, j = 3

Model-averaged parameter estimates, standard errors (SE), 95% confidence limits (CL) and combined model weights of covariates for lower Grand River non-spatial multinomial likelihood Silver Shiner site occupancy models.

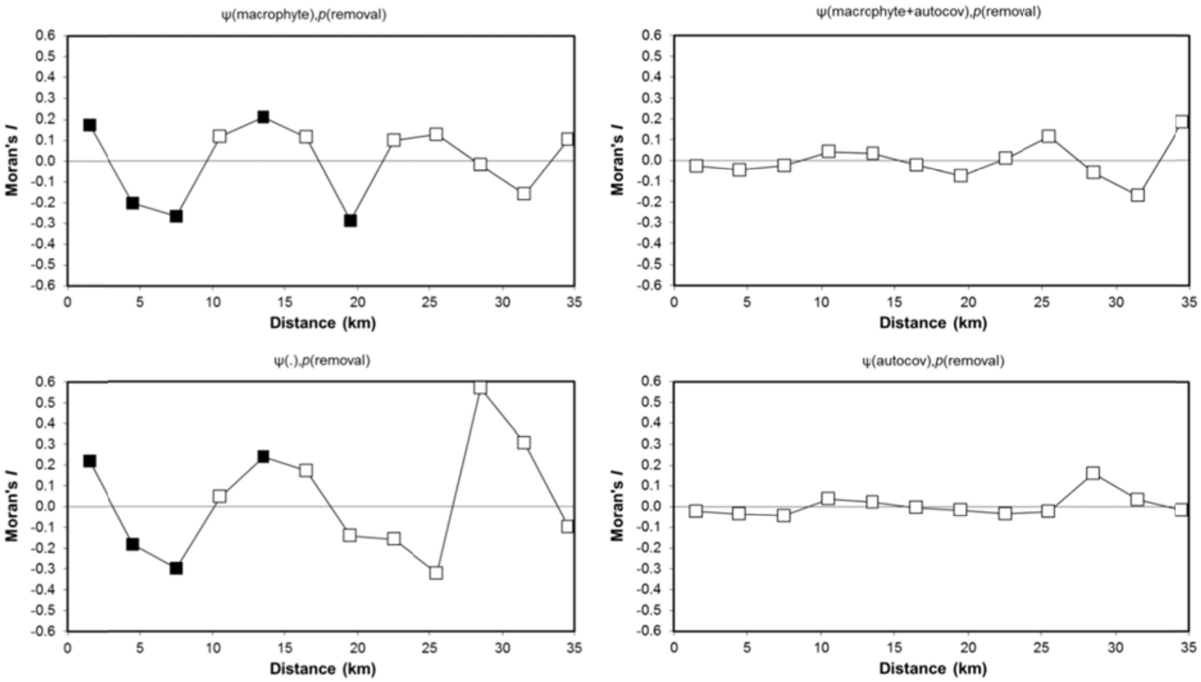
Parameter	Estimate (SE)	95% CL_{lower}	95% CL_{upper}	Combined model weight
Occupancy				
ψ	0.453 (0.103)	0.268	0.651	—
Silt/clay	-0.32 (0.69)	-1.66	1.03	0.27
Macrophyte cover	-2.09 (2.54)	-7.06	2.89	0.46
Water velocity	0.08 (0.21)	-0.34	0.49	0.19
Depth	0.04 (0.16)	-0.28	0.37	0.19
Water clarity	-4.0E-03 (0.13)	-0.25	0.25	0.15
Detection				
p	0.440 (0.110)	0.247	0.652	—
Removal	-1.52 (0.58)	-2.70	-0.33	1.00

Number of repeat surveys per site required to detect Silver Shiner based on estimated detection probability (p) and the upper and lower 95% confidence limits of p.



Silver Shiner – seining at 92 m^2 sites in lower Grand River (May – July 2007), n = 73, j = 3

Correlograms of Residual SAC for top-ranked site occupancy model (plots on right side are equivalent to models on left, but include a spatial autocovariate).



Summary of model selection procedure for Silver Shiner occupancy at sites in the lower Grand River based on spatial multinomial likelihood (n = 73, χ^2 = 10.3, p = 0.07, $\hat{c}$ = 1.98, naïve ψ = 0.397).

Model	QAICc	ΔQAICc	AIC weight	Number of parameters	-2[Log Likelihood]
ψ(macrophyte+auto),*p*(removal)	102.65	0	0.1092	6	176.65
ψ(auto),*p*(removal)	103.41	0.76	0.0747	5	182.86
ψ(silt+macrophyte+auto),*p*(removal)	103.96	1.31	0.0567	7	174.41
ψ(silt+auto),*p*(removal)	104.56	1.91	0.0420	6	180.42
ψ(macrophyte+velocity+auto),*p*(removal)	104.68	2.03	0.0396	7	175.82
ψ(velocity+auto),*p*(removal)	104.95	2.30	0.0346	6	181.20
ψ(macrophyte+wc+auto),*p*(removal)	105.03	2.38	0.0332	7	176.51
ψ(depth+macrophyte+auto),*p*(removal)	105.09	2.44	0.0322	7	176.64
ψ(depth+auto),*p*(removal)	105.60	2.95	0.0250	6	182.49
ψ(wc+auto),*p*(removal)	105.77	3.12	0.0229	6	182.83
ψ(silt+wc+macrophyte+auto),*p*(removal)	106.43	3.78	0.0165	8	174.28
ψ(silt+depth+macrophyte+auto),*p*(removal)	106.49	3.84	0.0160	8	174.40
ψ(silt+depth+auto),*p*(removal)	106.88	4.23	0.0132	7	180.17
ψ(depth+velocity+auto),*p*(removal)	106.89	4.24	0.0131	7	180.19
ψ(silt+wc+auto),*p*(removal)	107.01	4.36	0.0123	7	180.42
ψ(velocity+wc+macrophyte+auto),*p*(removal)	107.17	4.52	0.0114	8	175.75
ψ(velocity+wc+auto),*p*(removal)	107.40	4.75	0.0102	7	181.20

Silver Shiner – seining at 92 m^2 sites in lower Grand River (May – July 2007), *n* = 73, *j* = 3

Model-averaged parameter estimates, standard errors (SE), 95% confidence limits (CL) and combined model weights of covariates for lower Grand River spatial multinomial likelihood Silver Shiner site occupancy models.

Parameter	Estimate (SE)	95% CL_{lower}	95% CL_{upper}	Combined model weight
Occupancy				
ψ	0.469 (0.103)	0.228	0.725	—
Autocovariate	2.40 (1.34)	-0.23	5.03	1.00
Silt/clay	-0.25 (0.64)	-1.50	1.00	0.16
Macrophyte cover	-1.94 (2.54)	-6.93	3.04	0.31
Water velocity	0.07 (0.22)	-0.35	0.50	0.11
Depth	0.02 (0.16)	-0.29	0.33	0.10
Water clarity	0.02 (0.14)	-0.26	0.29	0.11
Detection				
p	0.416 (0.133)	0.197	0.675	—
Removal	-1.41 (0.68)	-2.75	-0.07	1.00

Number of repeat surveys per site required to detect Silver Shiner based on estimated detection probability (*p*) and the upper and lower 95% confidence limits of *p*.