School of Graduate Studies
Bemidji State University
1500 Birchmont Dr NE, #48
Bemidji, MN 56601-2699
218-755-2027

Assessment of Walleye and Lake Whitefish Spawning Habitat Usage for Use in Predictive Model Validation

by

Benjamin D Erb

A Thesis Submitted to the Faculty of the
DEPARTMENT OF BIOLOGY

In Partial Fulfillment of the Requirements
For the Degree of

Master of Science in Biology

BEMIDJI STATE UNIVERSITY
Bemidji, Minnesota, USA

August 2026

STATEMENT BY THE AUTHOR

This thesis has been submitted in partial fulfillment of requirements for an advanced degree at Bemidji State University and is deposited in the University Library to be made available to borrowers under the rules of the library.

Brief quotations from this thesis are allowable without special permission, provided accurate acknowledgement of the source is indicated. Requests for permission to use extended quotations or to reproduce the manuscript in whole or in part may be granted by the Department of Biology or the Dean, College of Sciences and Health when the proposed purpose is in the interest of scholarship. In all other instances, however, permission must be obtained from the author.

Signed: Benjamin D Erb

APPROVAL BY THESIS ADVISOR

THIS THESIS HAS BEEN APPROVED ON THE DATE SHOWN BELOW:

Andrew W. Hafs, Ph.D. Committee Chair Professor of Biology	9 Sept 2026 Date
Marci Maple Dean, College of Sciences and Health	15 Sept 2026 Date

Assessment of Walleye and Lake Whitefish Spawning Habitat Usage for Use in Predictive Model Validation

Benjamin D Erb

This project exists to validate and enhance upon models provided by Environment and Climate Change Canada (ECCC) which predict spawning sites for Walleye and Lake Whitefish in Rainy Lake and Namakan Reservoir. Model validation was accomplished by sampling for fish eggs and comparing sample data to the predictive model. Walleye eggs were sampled on Rainy Lake only, and Whitefish data was collected from both Rainy Lake and Namakan Reservoir. Walleye eggs were sampled in <1 m of water using a scap net, and Lake Whitefish eggs were sampled in 1.5 – 5.3 m of water using egg-mats. Model performance differed largely between predictions of egg presence and absence. For both species, the models performed substantially better at predicting egg absence (63.5% for Walleye and 76.9% for Lake Whitefish) than egg presence (30.4% and 27.7%, respectively), indicating a greater ability to identify unsuitable spawning habitat. Field observations show that Walleye spawned on relatively shallow lake bottom slopes (3-15 °) and that depth and substrate were a key predictors for Walleye spawning sites (p=<0.01) Lake Whitefish were also selective of substrate type (p=0.04) but not selective of depth (p=0.62) Cobble was the most used substrate type of both species, with 97% of Walleye eggs and 80% of Lake Whitefish eggs being sampled over cobble substrates. The current ECCC model does not directly predict substrate, as a response, a ctree model was constructed using physical variables in GIS and through field-observed substrate presence to estimate locations of substrate on Rainy Lake and found depth to be the most useful predictor in estimating substrate types (p=<0.01).

Approved by:

9 Sept 2026

Committee Chair | Date

Ryan Maki | Digitally signed by Ryan Maki
Date: 2026.09.09 13:23:28 -05'00'

Committee Member

Bachand, Marianne | Signé numériquement par : Bachand, Marianne
Nom DN : CN = Bachand, Marianne C = CA O = GC
OU = EC-EC
Date : 2026.09.11 10:23:02 -04'00'

Committee Member

Carl Isaacson 9/11/26

Graduate Faculty Representative

ACKNOWLEDGMENTS

I would like to thank Bemidji State University and Dr. Andrew Hafs for the opportunity to perform this research, as well as my committee members Marianne Bachand and Ryan Maki for their excellent mentorship and project guidance. Additionally, I would like to thank the International Joint Commission, the Minnesota DNR, the Ontario MNR, and Environment and Climate Change Canada for their funding, time, and personnel. I want to specifically thank Audrey Moffet (ECCC), Olivier Champoux (ECCC), Melissa Moseley (Ontario MNR), Phil Talmage (MN DNR), and Ben Vondra (MN DNR) for their time and effort that they contributed to this project for data extractions, fieldwork logistics, and providing boats and equipment. This project would not have been possible without you.

Thank you as well to my undergraduate technicians Hunter Pawelski and Tegwen Romportl for their valuable time and assistance in my field work and data processing, and to my fellow graduate students, especially Collin Mitchell, Griffin Blegen, Baylor Short, Joey Heinen, and Amanda Schuermann for supporting me in my academic pursuits and providing a friendly and welcoming environment on campus.

I want to thank my parents and the countless friends and other family members of mine for having unwavering support and faith in me, even while I often did not.

Lastly, I would like to dedicate this thesis to the memory of my good friend Brian Mitchell, who passed away shortly after I began working on this project. If anyone reading this thesis is carrying weight on their shoulders, please speak about it with somebody, because the world is a better place with you here.

Table of Contents

List of Tables

List of Figures

Assessment of Walleye and Lake Whitefish Spawning Habitat Usage for Use in Predictive Model Validation

Benjamin D Erb

This project exists to validate and enhance upon models provided by Environment and Climate Change Canada (ECCC) which predict spawning sites for Walleye and Lake Whitefish in Rainy Lake and Namakan Reservoir. Model validation was accomplished by sampling for fish eggs and comparing sample data to the predictive model. Walleye eggs were sampled on Rainy Lake only, and Whitefish data was collected from both Rainy Lake and Namakan Reservoir. Walleye eggs were sampled in <1 m of water using a scap net, and Lake Whitefish eggs were sampled in 1.5 – 5.3 m of water using egg-mats. Model performance differed largely between predictions of egg presence and absence. For both species, the models performed substantially better at predicting egg absence (63.5% for Walleye and 76.9% for Lake Whitefish) than egg presence (30.4% and 27.7%, respectively), indicating a greater ability to identify unsuitable spawning habitat. Field observations show that Walleye spawned on relatively shallow lake bottom slopes (3-15 °) and that depth and substrate were a key predictors for Walleye spawning sites (p=<0.01) Lake Whitefish were also selective of substrate type (p=0.04) but not selective of depth (p=0.62) Cobble was the most used substrate type of both species, with 97% of Walleye eggs and 80% of Lake Whitefish eggs being sampled over cobble substrates. The current ECCC model does not directly predict substrate, as a response, a ctree model was constructed using physical variables in GIS and through field-observed substrate presence to estimate locations of substrate on Rainy Lake and found depth to be the most useful predictor in estimating substrate types (p=<0.01).

Introduction

Rainy Lake and Namakan Reservoir (including Lake Kabetogama, Sand Point, Crane, Little Vermillion and Namakan Lakes) are two large international boreal reservoirs located on the border of the United States (Minnesota) and Canada (Ontario) that support robust populations of Walleye *Sander vitreus* and Lake Whitefish *Coregonus clupeaformis*. The Walleye is a highly studied member of the Percidae family and is a species of ecologic, recreational, and socioeconomic significance (Scott and Crossman 1973; Colby et al. 1979; Barton 2011). Walleyes are a popular game fish and provide a popular sport fishery, and in turn provide an economic benefit to nearby communities (Vondra 2012). Approximately 60% of the fishing pressure on Rainy Lake and Namakan Reservoir is targeted at Walleye, which fiscally supports the many resorts and other tourism small businesses in the region (Kallemeyn et al. 2003). Walleyes are a spring-spawning species that broadcast spawn over hard substrates in shallow water shortly after ice-out in depths ranging from just a few centimeters to 1.5 m, and when water temperatures are between 2 and 12 °C (Papenfuss et al. 2017). Walleyes have a short incubation period, and eggs typically hatch about three weeks after they are laid (Morin et al. 2016). Lake spawning Walleye require coarse substrates such as pebble, cobble, and boulders, with many interstitial spaces for eggs to fall into (Zhao et al. 2009). Additionally, spawning typically occurs in areas that are frequently battered by waves, which cleans the substrate (Morin et al. 2016). These two factors ensure that freshly laid eggs are well oxygenated and have an optimal chance of survival. Due to the shallow water spawning habits of Walleye, a water level management concern exists to ensure that critical spawning habitat is available to Walleye and their eggs during the spring spawning and incubation period. Low water levels will result in lower success of spring Walleye eggs due to ideal shallow water habitat being unavailable (Morin et al. 2016). Increased spring water levels are not as detrimental to Walleye egg survival as low water levels, but can result in eggs laid in deeper, colder water which increases incubation periods and results in a decrease in hatching success rates (Bush et al. 1975; Jones et al. 2003).

There has been little research conducted on spawning habits of Lake Whitefish in Boreal Reservoirs such as those that make up the lakes in the Rainy-Namakan system, yet Lake Whitefish are still a culturally significant species to the communities surrounding the Rainy – Namakan System, supporting subsistence and sport netting fisheries, especially on Namakan Reservoir. Also, unlike Walleye, Lake Whitefish are a fall spawning species of fish that lay their eggs over hard substrates when autumnal water temperatures fall below 12 °C and can continue until after ice-on occurs (Lawler 1961). Lake Whitefish are broadcast spawners, and spawn on the surface, usually in pairs, but sometimes up to three fish, where their eggs then descend to the lake bottom, falling into the interstitial spaces of the substrate (Muir et al. 2025). The incubation period of Lake Whitefish eggs lasts throughout the winter, and the eggs hatch in the spring. Spawning Lake Whitefish have been netted on the Rainy-Namakan System by commercial fishermen over coarse substrates near shore and on main-lake reefs in depths of up to 6.5 m.

Both systems have their water levels controlled by dams, the Rainy Lake Dam was completed in 1909, and the Namakan Reservoir Dam (aka Kettle Falls Dam) was completed in 1914 (Morin et al. 2016). Several paper companies controlled the two dams from the dam's construction through 1949. During this time, water level management was based on industrial purposes and resulted in extremely low water that was often near natural water levels (Morin et al. 2016). Following 1949, control of the dams was given to the International Joint Commission (IJC), who sought to balance water levels to benefit not only hydropower production, but other stakeholders including the state of Minnesota, province of Ontario, First Nations communities (including the Naicatchwenin, Nigigoonsiminikaanmng, Couchiching and Mitaanjigamiing first nations of Ontario) as well as sportsmen, cabin owners, and riparian landowners (IJC 2001). Several water level management rule curves have been established since the transfer of power to the IJC. A "rule curve" represents target water levels to maintain throughout the year, generally higher water levels in the summer, and low water levels in the winter. The first IJC rule curve was established in 1949, and the most recent rule curve was established in 2018. As a result of high seasonal fluctuations in water levels on both Rainy Lake and Namakan Reservoir, careful water management decisions must be considered to help promote the survival of culturally and socioeconomically significant species such as Walleye and Lake Whitefish.

At the request of the International Joint Commission (IJC) predictive models were developed by ECCC to help the IJC prepare for water level management decisions in the Rainy Namakan System. This project examines ECCC model performance by sampling for Walleye and Lake Whitefish eggs during the spring and fall of 2024 and 2025 at sites predicted by the ECCC model and uses summer substrate sampling to develop a substrate proxy for Rainy Lake to be integrated with the ECCC model. The objectives of this study are to (1) assess the discriminatory capacity of the existing model to predict the presence and absence of eggs; (2) identify the habitat variables associated with egg observations; (3) develop an exploratory approach to predict substrate; and (4) formulate recommendations for a future update of the ECCC model.

Methods

Study Area:

This study took place on Rainy Lake and Namakan Reservoir (Figure 1). Rainy Lake is located just east of International Falls, Minnesota and the neighboring community of Fort Frances, Ontario. It is a 92,108-ha reservoir with over 400 islands and 3,501 km of shoreline, Namakan Reservoir is located approximately 22 km southeast of International Falls and Fort Frances, and is made up of Crane, Sand Point, Little Vermilion, and Kabetogama Lakes, as well as Namakan Lake itself. It measures 24,097-ha and has 1,114 km of shoreline.

ECCC Model:

For Walleye and Lake Whitefish, ECCC developed a model that included both 1D and 2D components. Their 1D model components estimate egg survival probabilities as a function of water level fluctuations during the spawning and egg incubation periods, and their 2D model components are spatially explicit and predict the extent of available spawning and egg incubation habitat (Morin et al. 2016). Using their physical variable inputs from 1D and 2D models, ECCC developed probability values ranging from 0 to 1, where a probability value coefficient of 0.62 was used to separate suitable and unsuitable spawning conditions. Inputs for Walleye and Lake Whitefish models include fetch, slope of the lake bottom, depth, and water temperature. For Walleye, wave action generated from winds of 10 to 45 km/h was added as an input, as weaker winds do not generate enough force to influence substrates, and stronger winds are rare and difficult to include in a model (Morin et al. 2016). The Lake Whitefish models incorporated sunlight intensity at the lake bottom as an input. The 1D model takes all available data and assigns each node a 1 or a 0, where 1 represents egg survival being likely, and 0 represents egg survival being unlikely. The 2D model uses spatial data to create a raster of values between 0 and 1, where sites with values <0.62 represent unsuitable spawning and incubation habitat, and sites with values >0.62 represent suitable spawning and incubation habitat. Data used in the models go as far back as the 1970’s and incorporate nodes spaced 10 m apart across the entire system (Morin et al. 2016). Previous scientific data for Walleye spawning sites on Namakan Reservoir had been collected (Papenfuss et al. 2017) but was absent from Rainy Lake. Thus, all the Walleye egg samples taken during this project were only taken on Rainy Lake. Whitefish data was provided by commercial fishermen who shared their fishing logs, but this study was the first scientific research conducted on Whitefish spawning in the Rainy - Namakan system, so Whitefish egg sampling for this project took place on both Rainy Lake and Namakan Reservoir.

One area where the ECCC model could be improved is identifying suitable spawning substrate for both Walleye and Lake Whitefish. Currently, the wind input used in the model is used in place of substrate, because there is a lack of physical substrate observation data for the system. Literature suggests that both species can be highly selective of the substrates they spawn over, with both species most frequently using coarse substrates such as cobble or boulder (Fischer et al. 2018; Papenfuss et al. 2018). Currently, the ECCC model does not use substrate as an input for their models, so integrating a near shore predictive substrate model could help improve ECCC model performance in identification of spawning sites for both Walleye and Lake Whitefish.

Site Selection:

ECCC created a heat map that represented their models for both Walleye and Lake Whitefish by creating a raster of 1's and 0's using all available model data. In this raster a 1 represents a suitable spawning condition for a model node on a given year, and a 0 represents an unsuitable spawning condition. This appears on their heat map as yellow-green shading for areas that frequently have suitable spawning conditions, dark blue for areas that frequently have unsuitable spawning conditions, and light blue for intermediate areas (Figure 2). Using the ECCC heat map, 120 total sites were selected (30 per spring Walleye egg sampling season of 2024 and 2025, and a different 30 per fall Whitefish sampling season of 2024 and 2025). For all sampling seasons, the selected 30 sites were distributed between the US and Canada to allow for each site to be visited once every 48 hours. Site selection for fall Lake Whitefish egg sampling was accomplished using the same methods used for spring Walleye egg sampling.

Walleye Egg Sampling:

Following the methods outlined by Papenfuss et al. (2017), a sample was taken from every site at 25, 50, 75, and 100 cm depths. The spring sampling period began in early spring shortly after ice-out and concluded when average water temps exceeded 12 °C, and a noticeable drop in egg counts per sample was observed. Before each sample was taken, a Secchi disc measurement was recorded just offshore of the sampling site in deep enough water to obtain a measurement to the nearest cm. Additionally, substrate was identified using the Wentworth scale (Wentworth 1922), and water temperature was recorded in °C before taking a sample. Samples were then captured via circular scap net (30-cm diameter, 250-micron mesh) by disturbing benthos with the heel of a wader boot, which would suspend eggs and other detritus into the water column, where they were then captured with a single swipe of the scap net. Every site was sampled every 48 hours, allowing sampling events to occur on the Minnesota and Ontario sides of Rainy Lake on alternating days. Captured eggs were preserved in 95% ethanol, a subset of which would later be sent to Dr. Loren Miller at the University of Minnesota Twin Cities for genetic identification. Site location was recorded using a Juniper Systems Geode paired with a Mesa 3 Tablet to take sub-meter GPS points which were imported to ARCpro GIS.

Walleye Egg Genetic Analysis:

Due to other species in Rainy Lake spawning during conditions that overlap with Walleye, it was important to be able to identify a proportion of spring egg samples. This was accomplished with the aid of Dr. Loren Miller and his lab at the University of Minnesota Twin Cities. DNA was extracted from 40 individual embryos crushed in 300 µl of a 5% (weight/volume) solution of a chelating resin (Chelex, Sigma Chemical, St. Louis, Missouri) and boiled at 56 °C for 8 minutes before use in PCR. PCR was performed using the COI-3 fish DNA barcoding cocktail (Ivanova et al. 2007). PCR reactions were screened using agarose gels and successful amplifications were submitted to a commercial DNA sequencing lab (Azenta) for sequencing with the M13F-21 primer. Sequences were viewed using Geneious Prime software and sequences were submitted to a BLAST search within the National Center for Biotechnology Information nucleotide database to identify the closest species match based on % pairwise identity. Additionally, genetic species identification was successful for 33 out of 40 fish embryos. Each individual had >500 bp of quality sequence and had >99% pairwise identity with the assigned species. The remaining 7 samples were unable to provide confident species identification. An assumption was made that enough Walleye are present in the system for their eggs to be detected by these methods.

Whitefish Egg Sampling:

Due to the depths at which Lake Whitefish spawn, the scap-netting method used for Walleye would prove to be ineffective. Instead, egg mats were constructed from flat half-cinder blocks with 20 x 40 cm hog hair furnace filter segments fastened to each block, the completed egg mat was then attached to a buoy. Sampling periods began when descending autumnal water temperatures reached 12 °C and concluded shortly before ice-on occurred. One sample was taken from each site with a targeted depth between 1.83 and 4.88 m. Egg mats were checked every 48 hours, alternating each day between Minnesota and Canada. Latitude and longitude were recorded on the initial set, and egg mats were redeployed in the same location until the end of the sample period. Additionally, water temperature was recorded before each trap retrieval, and substrate was recorded using a recreational-grade Marcum Pan-Cam and the Wentworth scale (Wentworth, 1920). Any captured eggs were extracted using tweezers and preserved in 95% ethanol. The only other species found in the Rainy-Namakan system that spawn under similar conditions is the Cisco *Coregonus arteidi,* the eggs of which can be identified using egg diameter (Pauve et al. 2020). Any egg over 3 mm was identified as a Lake Whitefish, and any egg under 2.5 mm was identified as a Cisco. If any eggs were measured between 2.5 and 3 mm, they would also be sent for genetic analysis, but none of the eggs sampled in this study fell within that range. As with spring Walleye, sampling site location was recorded using a Juniper Systems Geode GNS3 GNSS Receiver paired with a Juniper Systems Mesa 3 Tablet to take sub-meter GPS points which were imported to ARCpro GIS. An assumption was made that enough Lake Whitefish are present in the system for their eggs to be detected by these methods.

Model Validation:

All data processing and model validation techniques were performed using R statistical computing package (hereafter "R"; R Core Team 2020). Model performance can be interpreted as how well egg presence and absence were predicted by the ECCC model in both 2024 and 2025. Pie Charts were constructed to show model performance in predicting presence and absence. For both Walleye and Lake Whitefish, two pie charts each were made. The first of which showed model performance in predicting egg presence, and the other chart representing egg absence, with percentage slices showing the ratio of egg presence or absence represented by observations from the field.

A True Skill Statistics (TSS) analysis was used to compare each physical variables performance of egg presence estimation (Allouche et al. 2006). The two "readers" used in the TSS analysis were the ECCC model estimations, and presence/absence data collected during field sampling. Coefficient values provided in the ECCC model were used to determine yes/no values in the TSS analysis, where yes represents an estimated egg presence and no represents an estimated egg absence. If TSS values are close to 1, that suggests higher levels of agreement between readers, meaning that the model is performing well. TSS values close to 0 suggest that the model is not performing well and that agreement between readers is up to chance, meaning that the model is not performing well.

Two logistic regressions were performed on results from both years of Walleye and Lake Whitefish egg sampling. Binary data representing field observed egg presence and absence was plotted against probability values extracted from the ECCC model. A p-value of <0.05 and an upward sloping trendline would suggest that the model is performing well, whereas a p-value of >0.05 and a flat or downward sloping trendline would suggest that the model is not performing well at predicting egg presence/absence. In addition to logistic regressions, chi-squared tests were run on Walleye and Lake Whitefish selectivity of substrates and depth, vs available habitat visited during this study.

Additionally, bar charts were created and used to display quantitative data from Walleye and Lake Whitefish egg sampling. To display these results for Walleye, each sampling site was broken down into four subsites, one for each sample taken per location. This was not necessary for Whitefish because one sample was taken from each site instead of four.

Substrate Classification:

To develop a predictive model for substrate, physical substrate data was collected during the summer of 2025 using the ECCC model grid to randomly select 500 near-shore nodes from multiple regions of Rainy Lake to visit for substrate analysis. Each site was visited once (between 30-May-2025 and 10-September-2025), where date, depth, and substrate were recorded. The data collected from these nodes were used to build a conditional inference regression tree (cTree) model in program R. This substrate model was created using physical variables including fetch, bedrock outcroppings, depth, and distance from nearest tributary as input variables (Zentner et al. 2019). Physical variables that were not measured in the field including fetch and distance from nearest tributary were provided by ECCC and the Minnesota Geospatial Information Office. The classification tree method was selected for the following reasons; tree-based models have high accuracy rates while using GIS data (Pradhan 2013), they easily handle categorical and continuous predictors at multiple scales (i.e., whole lake and transect specific), and they hierarchically account for interactions between them. The classification tree was run under default settings in program R using the “Bonferroni” method.

The Bonferroni test type was specified to correct for multiple testing that could lead to exaggerated P-values (Dalgaard 2008). The Bonferroni test type determined significant splits ($P \leq 0.05$) in the cTree and minimized error in variable selection (Dalgaard 2008). A total of 10,000 model simulations were ran using a train-test split of 400/100 observations (400 for training, 100 for testing).

Results

Spring Walleye

A total of 40 egg samples from 36 sites were sent to Dr. Loren Miller at the University of Minnesota Twin Cities for genetic analysis. Of those samples, 33 of 40 were able to be successfully identified and revealed 23 Walleye, 8 Longnose Suckers *Catastomus catastomus*, 2 Trout Perch *Percopsis omiscomaycus,* suggesting slight overlap of spawning habitat between species. All Walleye egg numbers have been adjusted to represent this proportion.

Spring egg sampling took place between 29-April and 9-May of 2024, and between 13-May and 24-May of 2025. A total of 1948 eggs were collected during spring egg sampling sessions (1662 in 2024, and 322 in 2025). Most of the Walleye spawning took place in shallow depths of 25 (84.6% of total eggs) and 50 cm (13.7% of total eggs), and very few eggs were sampled at depths of 75 or 100 cm depths (Figure 3). As such, there is significant evidence to suggest that Walleye are selecting shallow depths to spawn over (χ^2 =25.733, df = 3, p = <0.01). Substrates were an important physical variable for spawning Walleye, specifically the presence of cobble substrates. A total of 97% of all spring eggs sampled were over cobble substrates (Figure 4). There was enough evidence to suggest slope use was different than what was available (χ^2 = 32.996, df = 5, p = <0.01, Figure 4). During Walleye egg sampling, cobble was also the most sampled substrate class (Figure 4), with 41% of sites having cobble substrates. Of all the subsites that produced at least one egg, 91% had cobble substrates (Figure 4). Additionally, there was not enough evidence to suggest slope use was different than what was available (χ^2 = 2.257, df = 6, p = 0.93, Figure 5). Water temperatures ranged from 4 to 18 °C during sampling periods, with peak spawning occurring between 5 and 8 °C (Figure 6).

Fall Whitefish

Fall egg sampling took place between 17-October and 23-November of 2024, and 24-October and 14-November of 2025. Water temperatures ranged from 12 to 3 °C. There was not a clear peak temperature for Lake Whitefish spawning, eggs were first observed at temperatures just below 8 °C both years and were present throughout the rest of the sample periods.

A total of 25 Lake Whitefish eggs were collected during fall egg sampling sessions (17 in 2024 and 8 in 2025). As with Walleye, substrate was also an important predictor of Lake Whitefish egg presence. However, depth was not as valuable of a predictor with Lake Whitefish spawning sites as it was with Walleye. Eggs laid by Lake Whitefish during this study were most frequently observed on Cobble, with 80% of Lake Whitefish eggs sampled on cobble substrates (Figure 7). Results from substrate use by Lake Whitefish in this study suggest that cobble substrate use was higher than other substrates that were available ($\chi^2 = 9.547$, df = 4, $p = 0.04$, Figure 7). Depth ranges between 2.8 and 3.9 m were the most productive during the sampling sessions of both years, producing 19 of 25 eggs, however there is not enough evidence to suggest Lake Whitefish are selecting this depth range ($p=0.62$, Figure 8). A total of 14 Cisco *Coregonus spp.* eggs were also collected during fall Whitefish sampling. All eggs were identified using egg diameter (Pauve et al. 2020), with Whitefish eggs measuring 3.0 to 3.7 mm, and Cisco eggs measuring 1.7 to 2.5 mm (Figure 9).

Model Validation

The ECCC model showed higher success in predicting Walleye egg absence locations than egg presence locations. Of the 92 subsites that were estimated by the ECCC model to produce eggs, 64 of those subsites (69.6%) failed to produce eggs, and 28 subsites (30.4%) did produce at least one egg (Figure 10). Conversely, of the 156 subsites that were estimated by the ECCC model to not produce eggs, 99 of those subsites (63.5%) did not produce any eggs, whereas 57 (36.5%) of those subsites did produce at least one egg (Figure 10).

As with the results from Walleye validation, the Lake Whitefish model was also better at estimating egg absence locations than egg presence locations. Of the 47 sites that were estimated by the ECCC model to produce eggs, 13 (27.7%) sites produced at least one egg, and 34 (72.3%) sites failed to produce any eggs (Figure 10). Out of the 13 sites that were estimated to not produce eggs, 10 (76.9%) sites did not produce eggs, and 3 (23.1%) sites did produce eggs (Figure 10). There was poor agreement between ECCC model estimations and field observations. The TSS analysis ran on Walleye observations resulted in substantial disagreement between readers (T= -0.06, Table 1). Lake Whitefish TSS results were slightly better but still suggest poor agreement between field observations and model probabilities (T=0.04, Table 2).

Logistic Regressions

There was not a significant relationship between model estimated egg presence/absence and field-observed egg presence/absence. Both regressions resulted in downward-sloping regression lines and insignificant p-values of 0.397 for Walleye and 0.725 for Lake Whitefish (Figure 11), providing evidence to suggest that the current model showed limited ability to discriminate observed egg presence from non-detections under the sampling conditions of this study, particularly for predicted presence sites

ctree Substrate Model

Of the 500 sites visited in the summer of 2025, 10 were silt, 32 were sand, 9 were pebble, 133 were cobble, 107 were boulder, and 210 were bedrock substrates (Figure 12). The ctree was trained using 400 data points tested on the remaining 100. Overall model accuracy was 75.6% with a criterion of 0.95 to determine significant splits aka “branches” (Figure 13). Significant branches represented in this model were bedrock outcroppings (p=<0.001) and depth (p=<0.001). Depths of less than 31 cm were most likely to produce cobble, and depths over 31 cm were more likely to produce boulder substrates. As expected, sites that were on bedrock outcroppings were predicted by the ctree as bedrock outcroppings every time. Pebble and silt were rejected by the model due to lack of representation. The highest number of cobble sites, which was the most used substrate by spawning Walleye in this study, were found in nearshore habitats less than 31 cm deep.

Discussion

Spring egg sampling revealed that Walleye in the Rainy-Namakan system use highly specific habitats for spawning. The results from Papenfuss et al. (2017) on Namakan Reservoir led to expected results of shallow (<50 cm) samples to yield the highest quantity of eggs. Results from this project do reinforce those findings, however, Walleye in Rainy Lake showed higher selectivity of habitats than those represented in the Papenfuss study on Namakan Reservoir. In this study, Walleye eggs sampled in 25 and 50 cm depths accounted for 98.3% of total eggs, whereas in the Papenfuss study, 25 and 50 cm samples represented about 80% of total eggs sampled. However, 25 cm samples alone represented 84.6% of eggs sampled in Rainy Lake, compared to about 45% of eggs sampled in the Papenfuss study on Namakan Reservoir. These findings reinforce the concern for water level management on both Rainy Lake and Namakan Reservoir to ensure that these specific habitats are available to Walleye during their spawning periods. However, spawning habitat can occur at a variety of elevations, so minor water level fluctuations may not be cause for concern. While existing literature does suggest that spawning Walleye frequently use a variety of coarse substrates containing many interstitial spaces for eggs to be deposited (Zhao et al. 2009), Walleye in Rainy Lake also displayed high selectivity to the substrates used for spawning. Cobble was nearly the exclusive substrate type used by Walleye in Rainy Lake, suggesting that estimating the location of shallow cobble is a very important factor in spawning habitat modeling.

Lake Whitefish also displayed a preference for specific substrates, but not quite to the same degree that Walleye did. The most frequently used spawning habitat by Lake Whitefish in this study was 2.8 – 3.9 m deep cobble, accounting for 76% of total eggs, however chi-square testing did not suggest depth selectivity by Lake Whitefish. Results from the Lake Whitefish portion of this study reinforce the findings of Muir et al. (2025) that state that Whitefish spawn over coarse substrates with many interstitial spaces for eggs to fall into. The Muir paper suggests that substrates that would be classified as pebble, cobble, or boulder on the Wentworth scale (Wentworth, 1920) are suitable for Lake Whitefish spawning. This appears to hold true on the Rainy-Namakan system, although none of the sites sampled in this study happened to be on pebble substrate, larger substrates still produced all Lake Whitefish eggs. Due to a knowledge gap of Whitefish spawning habitat in the Rainy-Namakan system, a preliminary water level management concern existed to ensure that Whitefish eggs that were laid in shallow water were not destroyed by ice scour during harsh winters that can lead to ice thickness over 1.5 m. The Whitefish results from this study should not be used as reasoning to change water level management to prevent winter ice-scour of eggs. The spawning habits of Lake Whitefish in boreal reservoirs have been studied much less than those of Walleye. As such, the methods and data from this study were relatively new. Future research that replicates the fall Lake Whitefish egg sampling methods used in this study would be advised to push starting water temperatures back from 12 to 9 °C, as both sampling periods of this study did not produce any eggs until water temperatures fell below 8 °C, and there were many sampling attempts before then that produced 0 eggs. Lawler (1961) found that once Lake Whitefish spawning begins, it can continue all the way until ice forms and may continue under the ice as well. Results from this study support their findings, as once Lake Whitefish eggs appeared, at least one egg was found every day until the end of the sampling period. Future research would benefit from increasing Lake Whitefish effort if logistically reasonable due to low sample size in this study. Caution is advised for regulatory use of this study until the model is further refined.

Another suggestion for future research would be the use of additional egg mats if logistically sound, as this would increase spatial coverage for samples, and allow multiple depths per site to be sampled. Both fall sampling seasons for Lake Whitefish represented normal water levels within the current rule curves. If future Lake Whitefish research takes place during abnormally high or low water levels, an objective should be included to see if spawning sites change with water level. If spawning sites remain the same, that would imply that substrate type is more important than depth, and if spawning sites changed, that would imply that depth is at least as important as substrate type.

Model validation results suggest that both the Walleye and Lake Whitefish model are better at predicting egg absence than egg presence. Walleye and Lake Whitefish pie charts both showed above 60% success rates when predicting egg absence, but below 31% when predicting egg presence. Additionally, high p-value results from logistic regressions, coupled with TSS values near 0, imply that the predictive models are underperforming in estimating suitability of spawning sites for both species. These results suggest that the models should be re-evaluated. As previously stated, the data from this study and from Papenfuss et al. (2017) show that depth and substrate are vital predictors for spawning sites for both species. However, substrate is not a current input for the ECCC model, which highlights the value of building an effective substrate proxy to integrate with the ECCC model.

Substrate observations revealed that cobble is moderately abundant in Rainy Lake, with 41% of Walleye subsites being over cobble substrate, and random near shore model nodes that were visited during the summer of 2025 resulted in 26.6% of all substrate types being identified as cobble. This contrasts the data collected by Zentner et al. (2019) who found mostly fine substrates, however, this is explained by the lake used in the Zentner study which was much smaller and more eutrophic than Rainy Lake. Results from the ctree model suggest that nearshore substrate habitat on Rainy Lake is influenced the most by depth, which supports using supports findings from Zenter et al. (2019). Future research could benefit from the addition of more physical variables to the ctree model, however with Rainy Lake being on an international border, difficulties can be encountered in finding good GIS data from both countries.

While the ECCC model shows a need for re-evaluation, the results from this study show how model improvements can be made. The importance of depth and substrate as predictors were highlighted by field observations from both Walleye and Lake Whitefish spawning seasons, implying the value of this studies data for improvement of performance indicators. Additionally, the implementation of a predictive substrate model will aid in improving model performance, specifically for sites that predict egg presence. As stated above, substrate data provided by field observations serve to identify important habitat used by Walleye and Lake Whitefish for spawning and can be integrated into the ECCC model, and the substrate ctree produced by this study can be used to predict substrate by itself. This study demonstrates the necessity of testing predictive models, and the importance of making model improvements where needed.

References

Allouche, O., A. Tsoar, and R. Kadmon. 2006. Assessing the accuracy of species distribution models: prevalence, kappa, and the true skill statistic (TSS). Journal of Applied Ecology 43:1223–1232.

Barton, B. A. 2011. Biology, management, and culture of walleye and sauger. American Fisheries Society, Bethesda, Maryland.

Bush, W. D. N., R. L. Scholl, and W. L. Hartman. 1975. Environmental factors affecting the strength of walleye (*Stizostedion vitreum vitreum*), year-classes in western Lake Erie, 1960–70. Journal of the Fisheries Research Board of Canada 32:1733–1743.

Colby, P. J., R. E. McNicol, and R. Ryder. 1979. Synopsis of biological data on the walleye (*Stizostedion vitreum vitreum*). Ontario Ministry of Natural Resources, Goulais River, Ontario.

Congalton, R. G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37:35–46.

Dalgaard, P. 2008. Introductory statistics with R, 2nd edition. Springer, New York.

Graham, J. D. 2016. Assessment of American White Pelican predation and reproductive success of walleye in the Tamarac River, Minnesota. Master's thesis, Biology Graduate Theses 12.

Graham, J. D., A. W. Hafs, and A. J. Kennedy. 2017. Quantification of walleye spawning substrate in a northern Minnesota river using side-scan sonar. North American Journal of Fisheries Management 37(2):420–428.

IJC (International Joint Commission). 2001. In the matter of emergency regulation of the level of Rainy Lake and of other boundary waters in the Rainy Lake watershed. International Joint Commission:1–18.

Ivanova, N., T. Zemlak, R. Hanner, and P. Herbert. 2007. Universal primer cocktails for fish DNA barcoding. Molecular Ecology Notes 7(4):544–548.

Jones, M. L., J. K. Netto, J. D. Stockwell, and J. B. Mion. 2003. Does the value of newly accessible spawning habitat for walleye (*Stizostedion vitreum*) depend on its location relative to nursery habitats? Canadian Journal of Fisheries and Aquatic Sciences 60(12):1527–1538.

Kallemeyn, L. W., K. L. Holmberg, J. A. Perry, and B. Y. Odde. 2003. Aquatic synthesis for Voyageurs National Park. U.S. Geological Survey and Voyageurs National Park, International Falls, Minnesota.

Lawler, G. H. 1961. Egg counts of Lake Erie whitefish. Journal of the Fisheries Research Board of Canada 18:293–294.

Morin, J., M. Bachand, J. H. Richard, and S. Martin. 2016. Modeling the Rainy Lake and Namakan Reservoir ecosystem response to water level regulation. Scientific Report SR-110, Meteorological Service of Canada, Hydrology and Ecohydraulics Section, Québec. Prepared for the International Joint Commission study on Rainy Lake and Namakan Reservoir. 448 pages.

Muir, A. M., Y. Drebert, R. Lauzon, Z. Melnick, C. M. Ryther, E. S. Dunlop, and K. Skubik. 2025. A new observation of Lake Whitefish (*Coregonus clupeaformis*) spawning behaviour. Journal of Great Lakes Research. 51(2)

Papenfuss, J. T., T. Cross, and P. A. Venturelli. 2017. Comparing water-level policies in a boreal reservoir: how wave and ice energy can help to maintain walleye spawning habitat. Lake and Reservoir Management 33:249–259.

Papenfuss, J. T., T. Cross, and P. A. Venturelli. 2018. A comparison of the effects of water-level policies on the availability of walleye spawning habitat in a boreal reservoir. Lake and Reservoir Management 34(4):321–333.

Scott, W. B., and E. J. Crossman. 1973. Freshwater fishes of Canada. Bulletin of the Fisheries Research Board of Canada 184.

Vondra, B. A. 2012. Investigation of natural reproduction of walleye in Lake Kabetogama. Minnesota Department of Natural Resources, Division of Fish and Wildlife:24.

Zentner, D. L., T. K. Cross, J. K. Raabe, and P. C. Jacobson. 2019. Using GIS to predict habitat in lakes: an example using nearshore substrate categories. Limnology and Oceanography 17(1):1–16.

Zhao, Y. M., M. L. Jones, B. J. Shuter, and E. F. Roseman. 2009. A biophysical model of Lake Erie walleye (*Sander vitreus*) explains interannual variations in recruitment. Canadian Journal of Fisheries and Aquatic Sciences 66(1):114–125.

Tables

Table1: TSS analysis table of all Walleye sites from spring spawning seasons of 2024 and 2025 on Rainy Lake (T= -0.06)

	ECCC Presence	**ECCC Absence**	**Row Total**
Field WAE Presence	28	57	85
Field WAE Absence	64	99	163
Column Total	92	156	248

Table 2: TSS analysis table of all Lake Whitefish sites from fall spawning seasons of 2024 and 2025 on Rainy Lake and Namakan reservoir (T=0.04)

	ECCC Presence	**ECCC Absence**	**Row Total**
Field WHF Presence	13	3	16
Field WHF Absence	34	10	44
Column Total	47	13	60

Figures

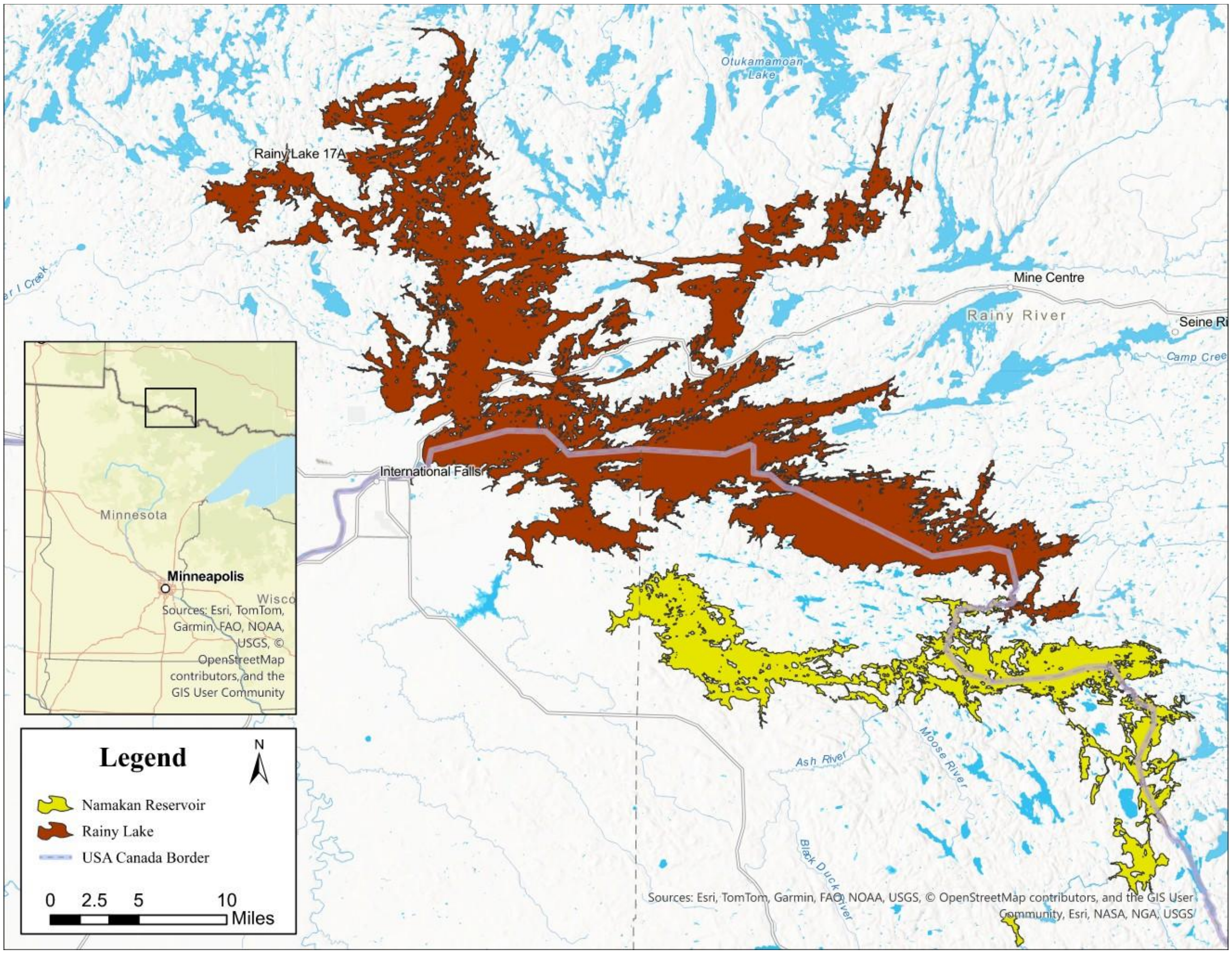

Figure 1: Location of the Rainy-Namakan System and its position on the Minnesota-Ontario Border. Rainy Lake (48°37'25.3"N 93°10'53.5"W) is directly east of the communities of International Falls, Minnesota, and Fort Frances, Ontario. Namakan Reservoir (48°26'57.3"N 92°41'48.0"W), is located south and east of Rainy Lake and contains Kabetogama, Namakan, Sand Point, Little Vermilion, and Crane Lakes.

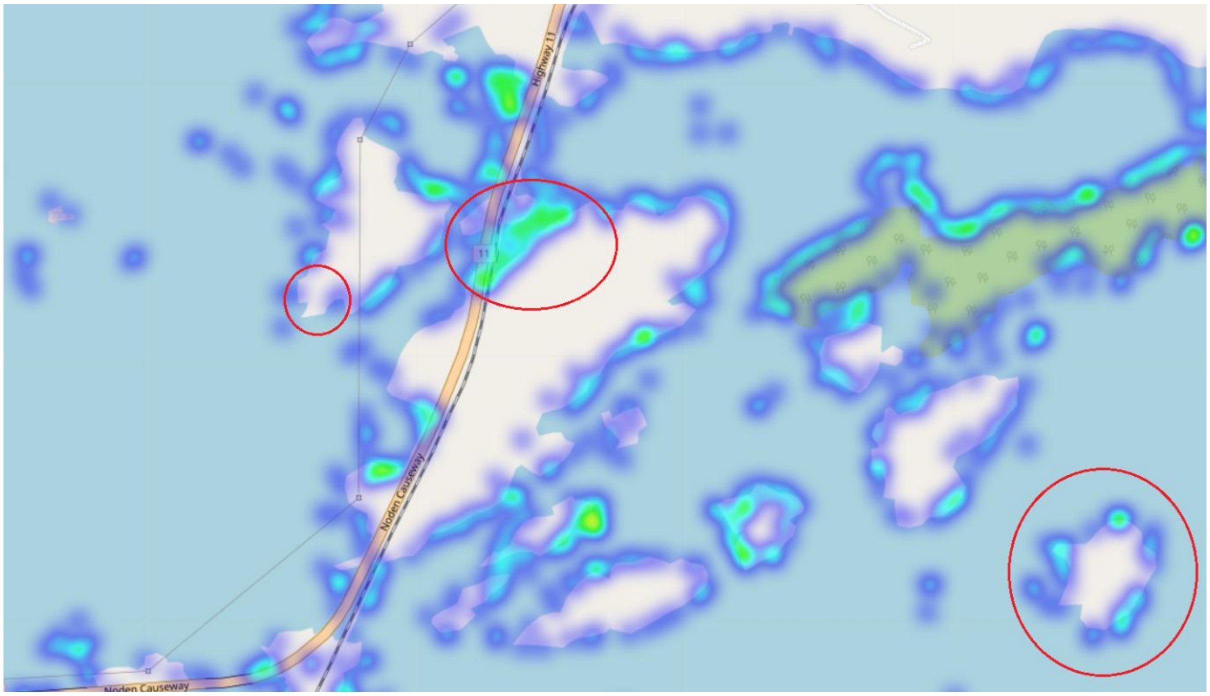

Figure 2: Example of the model-representative heat map provided by ECCC used for site selection. Areas of dark blue or no shading (left circle) are low-likelihood-of-spawning areas, yellow/green shading (middle circle) represents high-likelihood-of-spawning zones, and light blue shading (right circle) represents intermediate-likelihood-of-spawning zones.

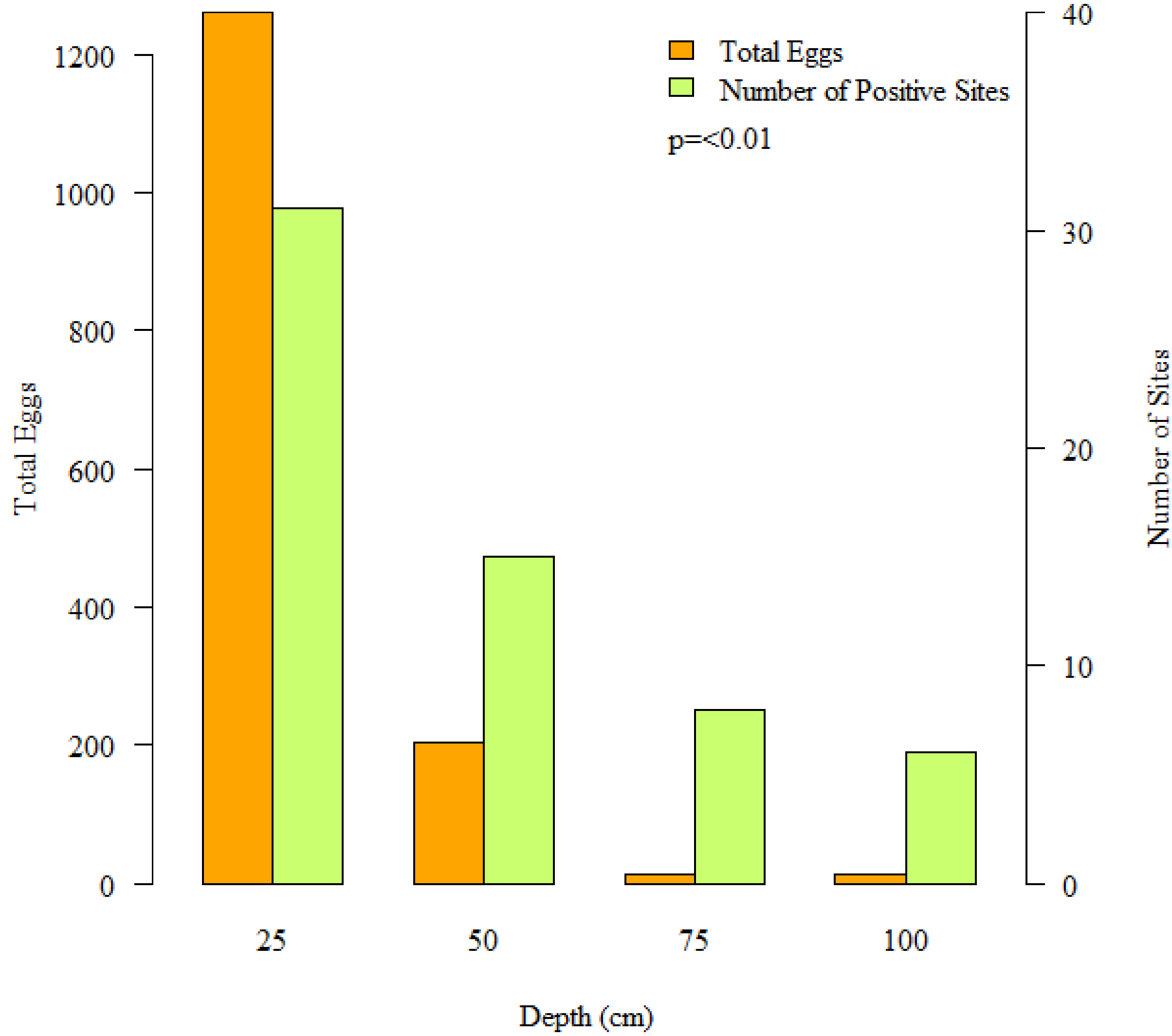

Figure 3: Cumulative total egg counts (y-axis) taken during spring Walleye egg sampling from depths of 25, 50, 75, and 100 cm of water (x-axis) the springs of 2024 and 2025 on Rainy Lake.

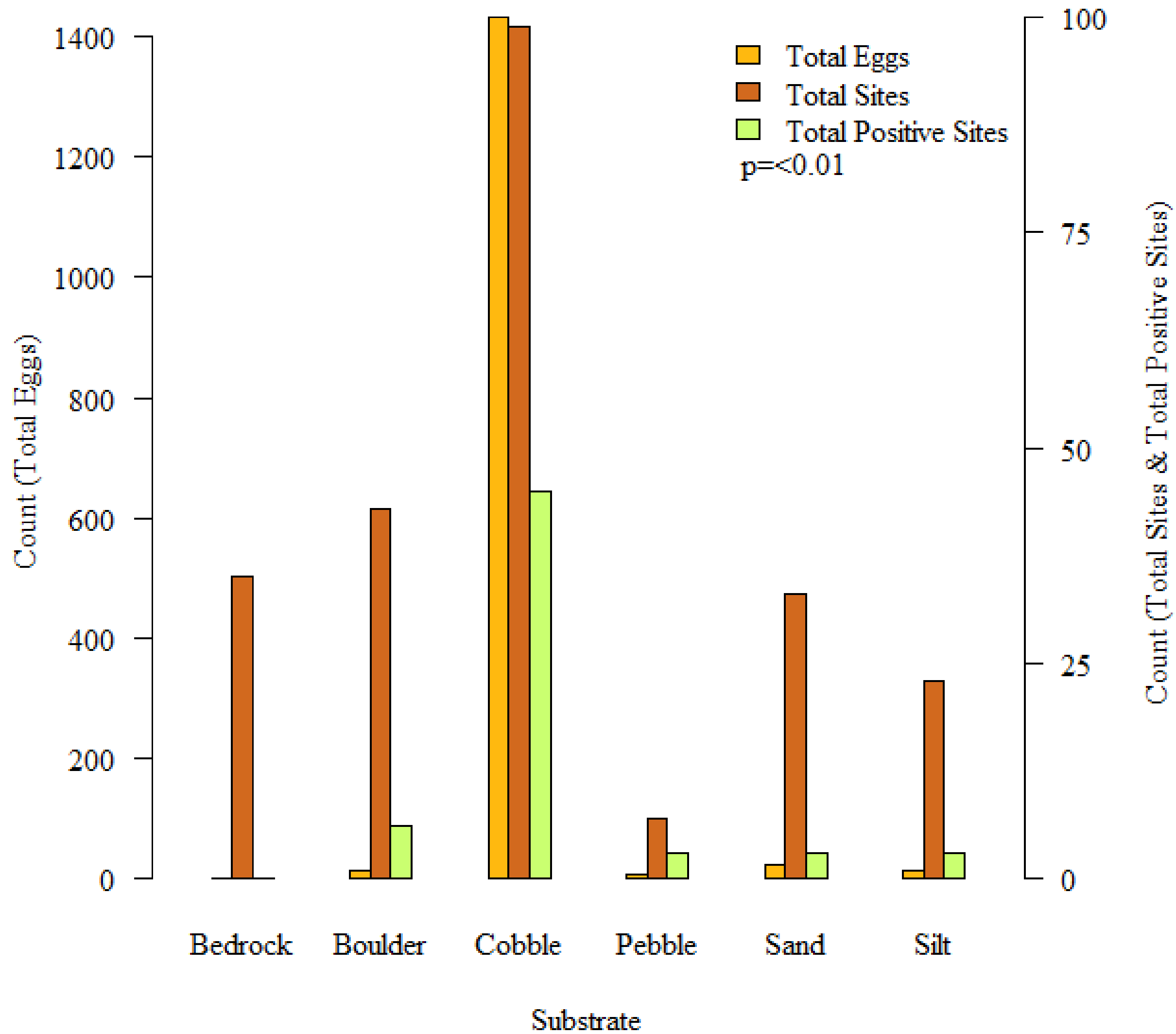

Figure 4: Cumulative total egg counts (y-axis, left), total sites, and positive sites (y-axis, right) taken during spring Walleye egg sampling over different substrates (x-axis) in the springs of 2024 and 2025 on Rainy Lake.

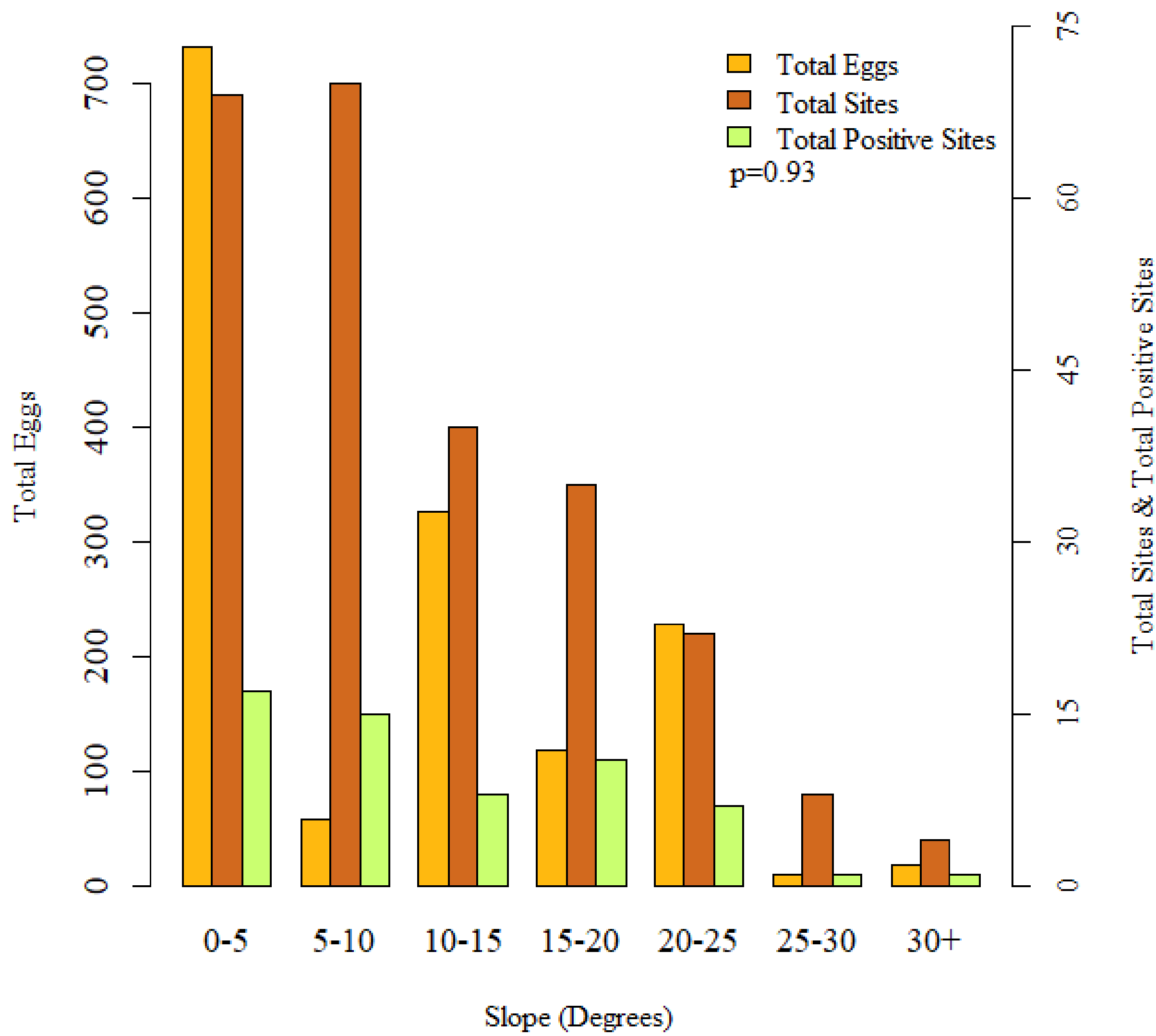

Figure 5: Cumulative total egg counts (y-axis, left), total sites, and positive sites (y-axis, right) taken during spring Walleye egg sampling from different lake bottom slopes (x-axis) the springs of 2024 and 2025 on Rainy Lake.

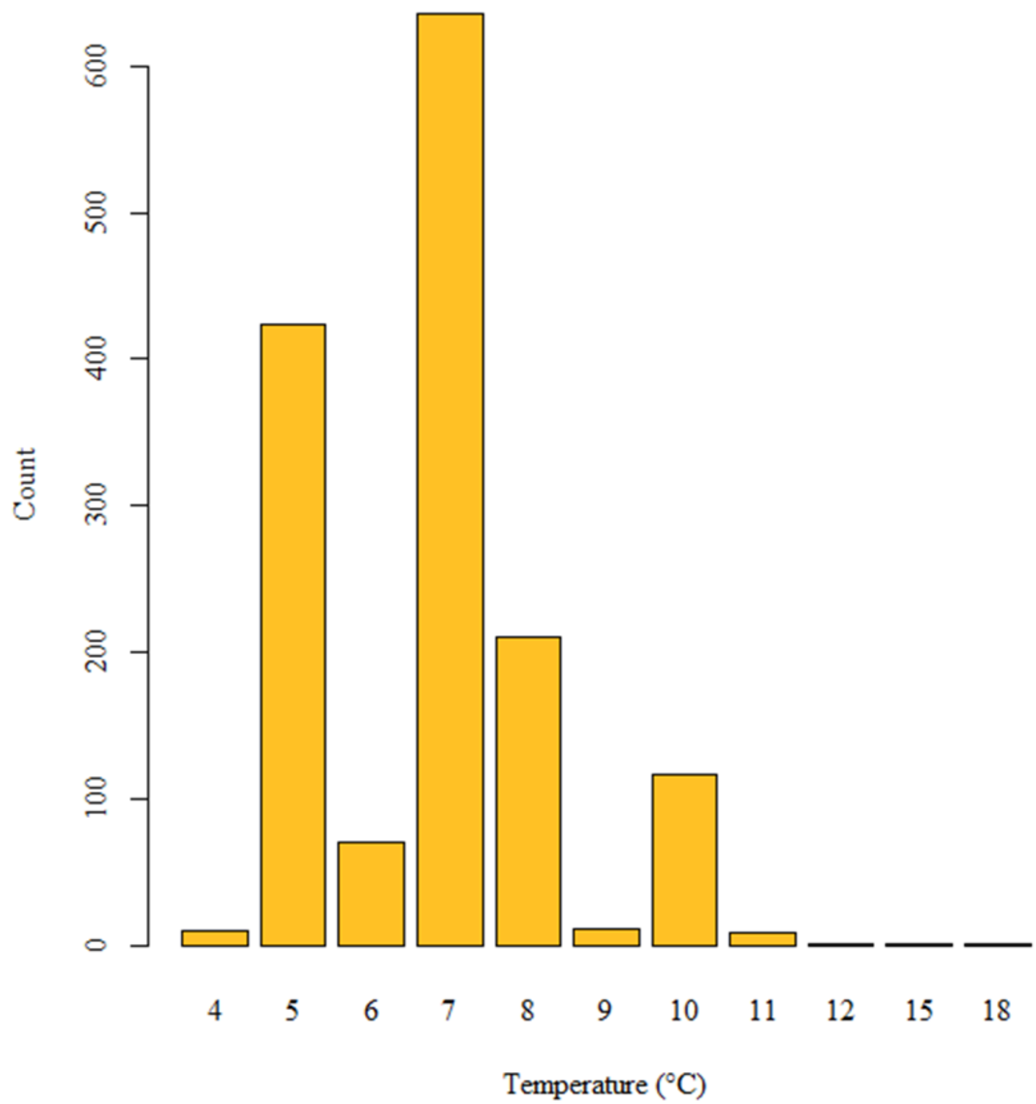

Figure 6: Total Walleye egg presence (y-axis) at temperatures in °C (x-axis) during the spring spawning seasons of 2024 and 2025 on Rainy Lake.

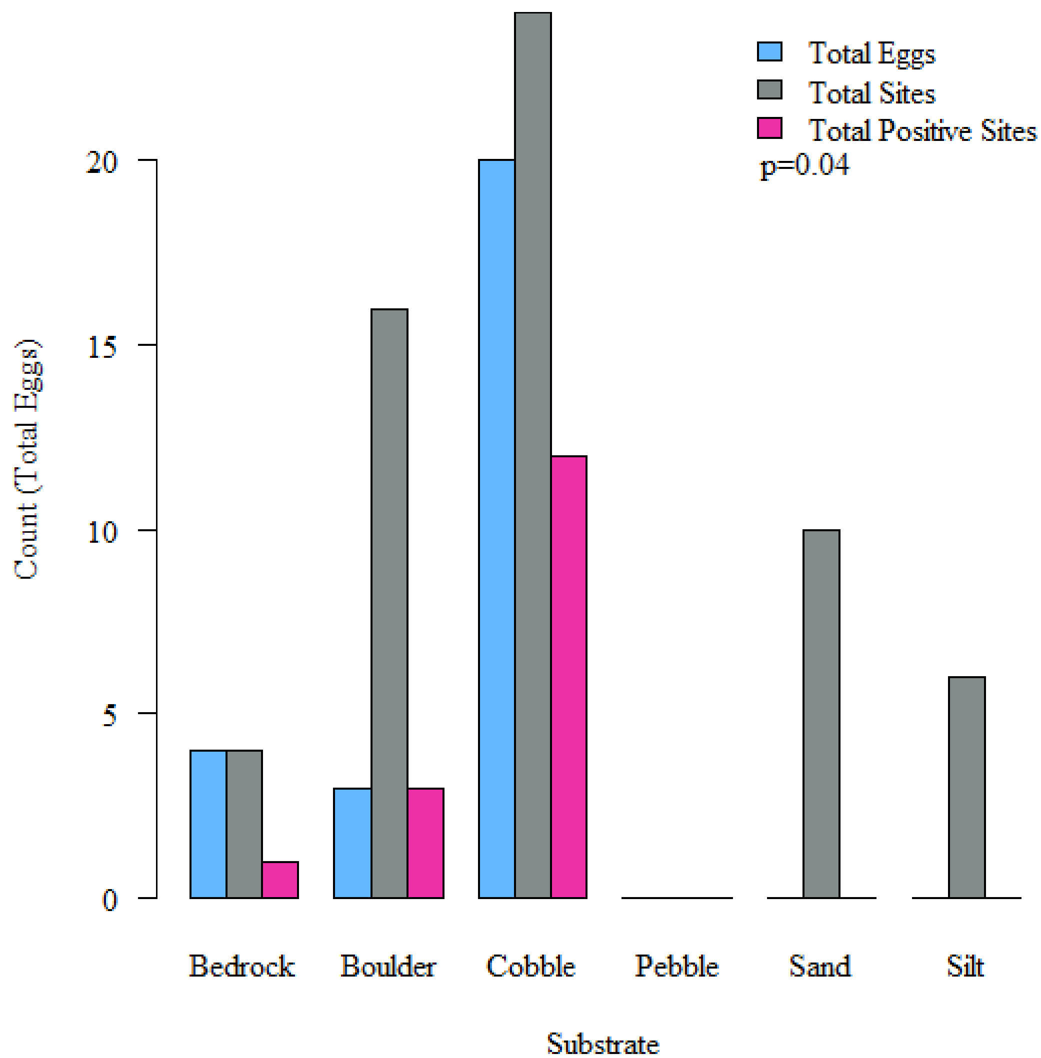

Figure 7: Counts of Lake Whitefish eggs and sites present at different substrates during the fall spawning periods of 2024 and 2025 on Rainy Lake and Namakan Reservoir.

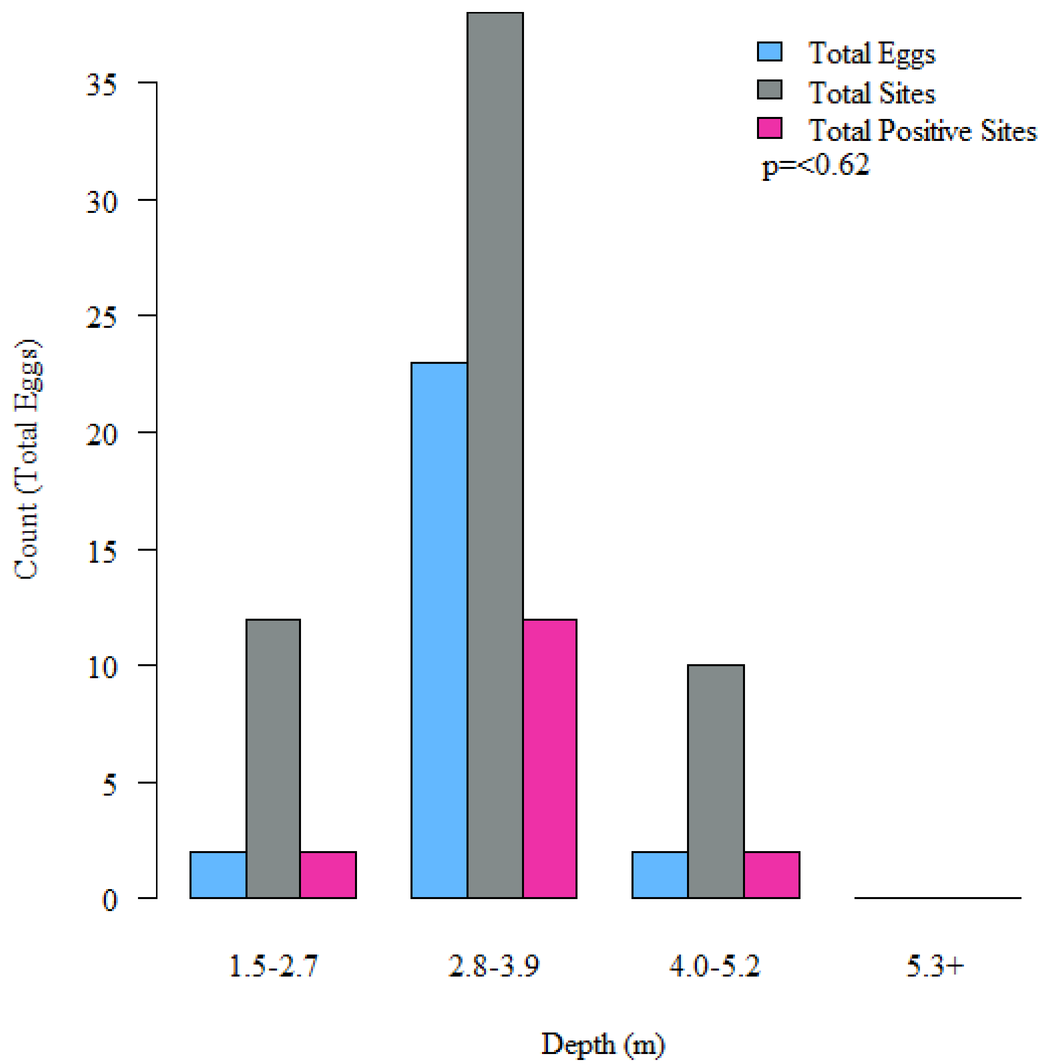

Figure 8: Counts of Lake Whitefish eggs and positive sites by depth range during the 2024 and 2025 fall spawning periods in Rainy Lake and Namakan Reservoir.

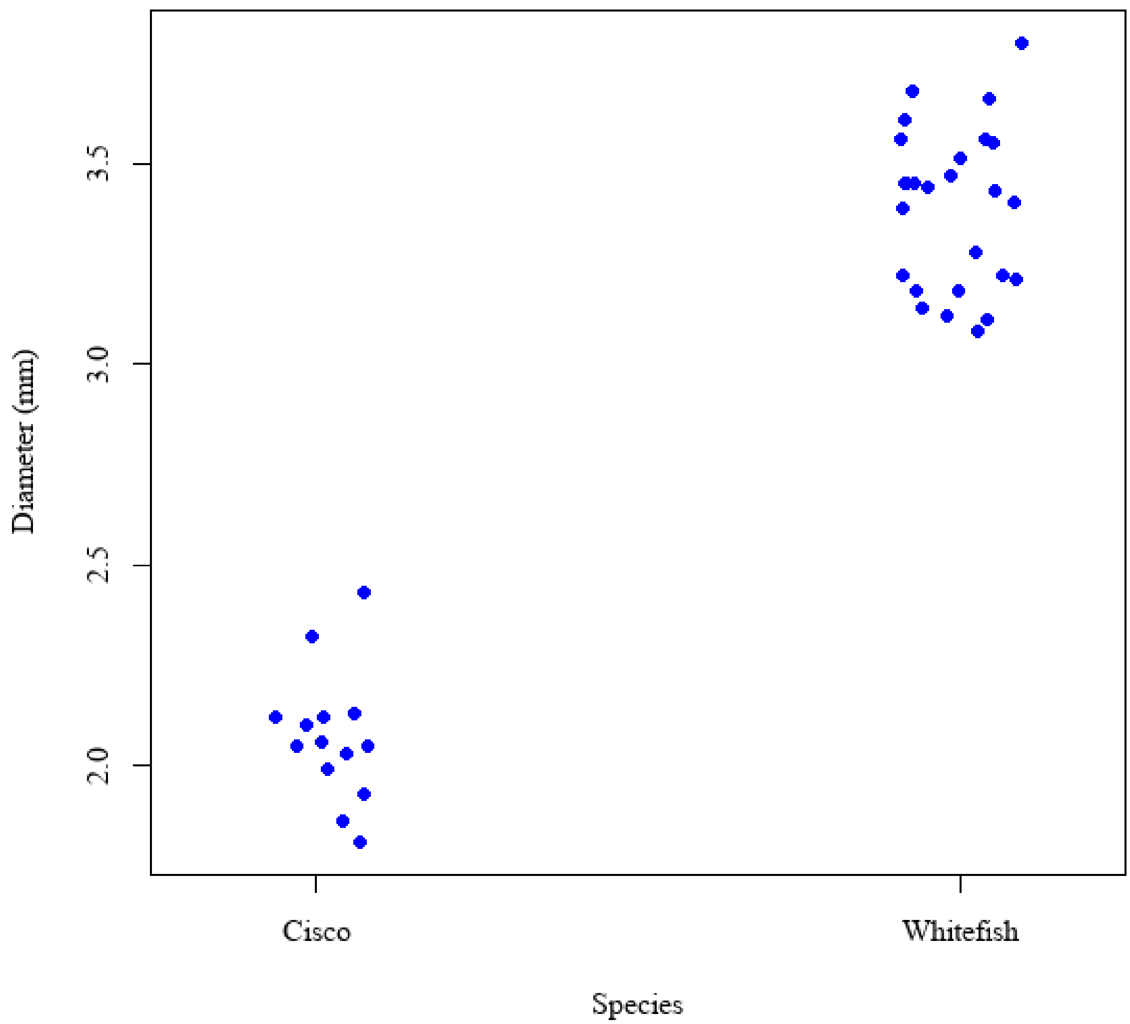

Figure 9: Egg diameters in mm (y-axis) of Cisco species (x-axis, left) and Lake Whitefish (right) taken from Rainy, Namakan, and Kabetogama Lakes during the fall of 2024 and 2025.

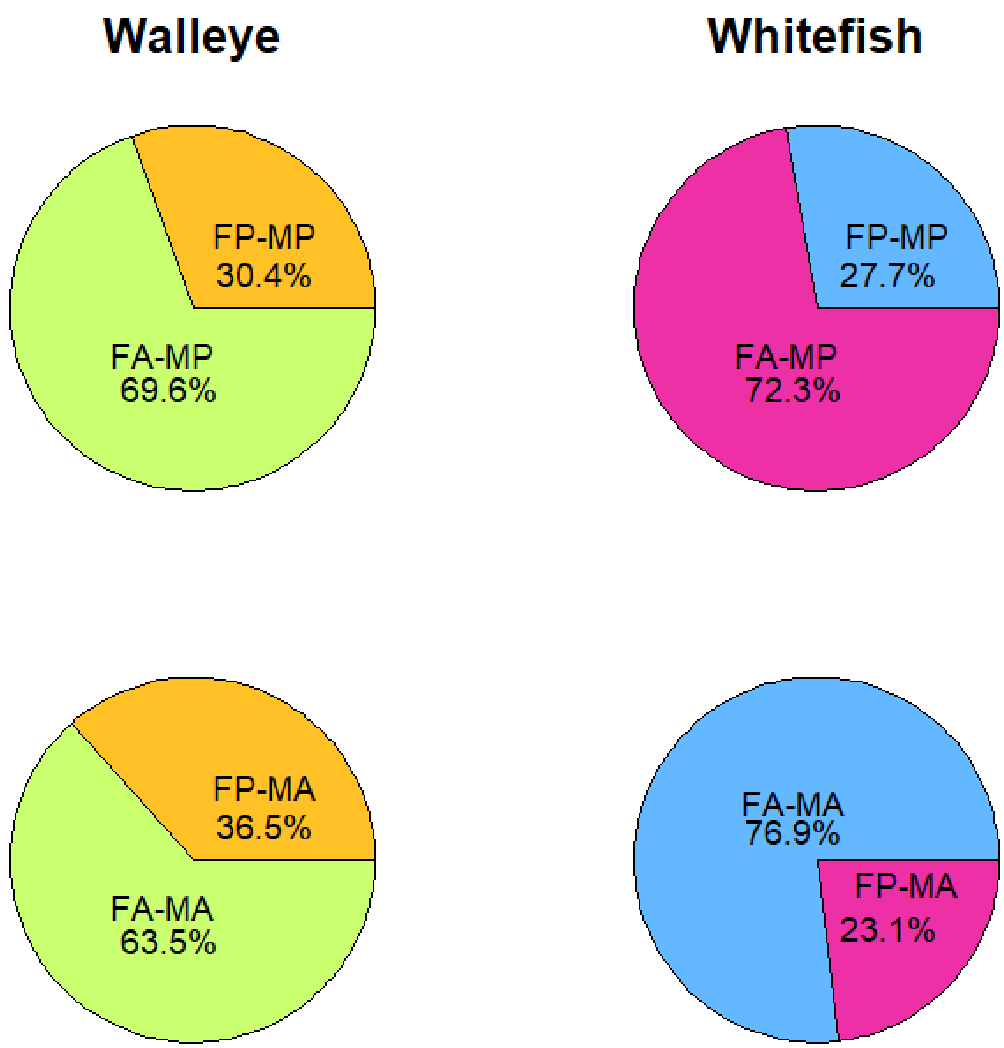

Figure 10: Pie charts showing predictive model success in predicting egg presence (top) and absence (bottom) for Walleye (left) and Whitefish (right) during the sampling seasons of 2024 and 2025 on the Rainy Namakan system. In this figure M = Model, F = Field, P = Presence, and A = Absence.

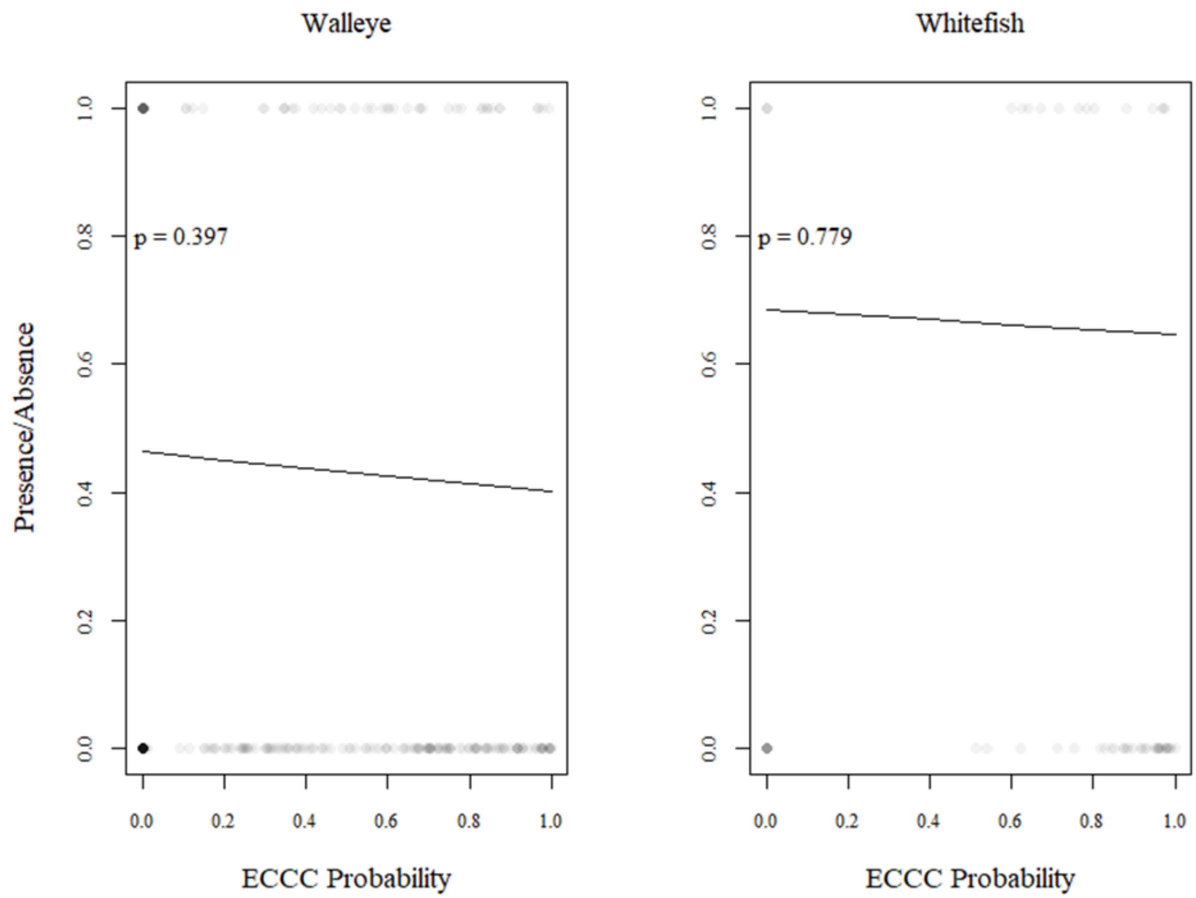

Figure 11: Logistic regression of spring Rainy Lake Walleye (left) and fall Rainy Lake and Namakan Reservoir Lake Whitefish (right) field observed presence (1) and absence (0) represented on the y-axis, and ECCC subsite probabilities (0 = least probable, 1 = most probable) displayed on the x-axis

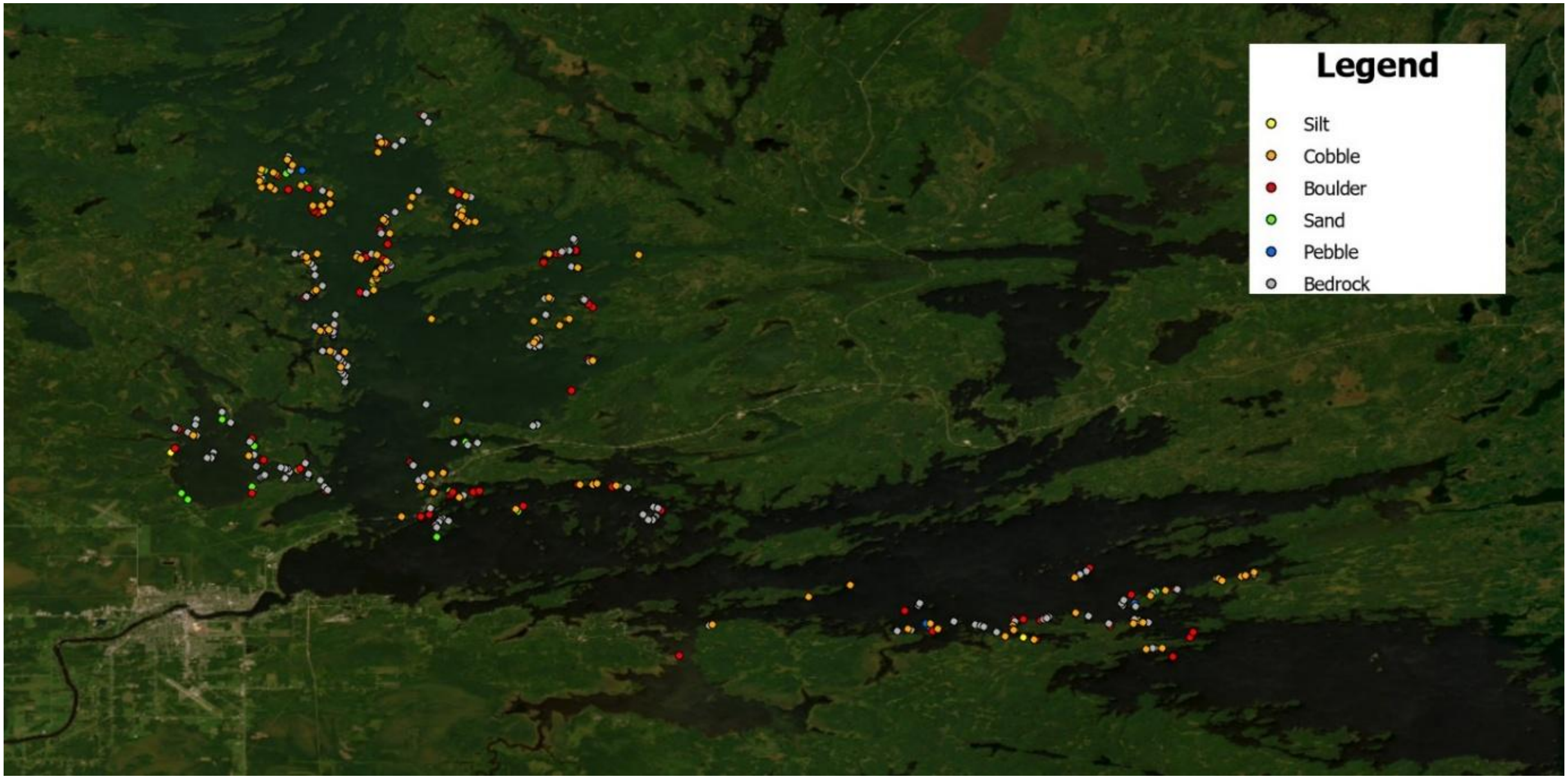

Figure 12: Total substrate sampling coverage of Rainy Lake during the summer of 2025.

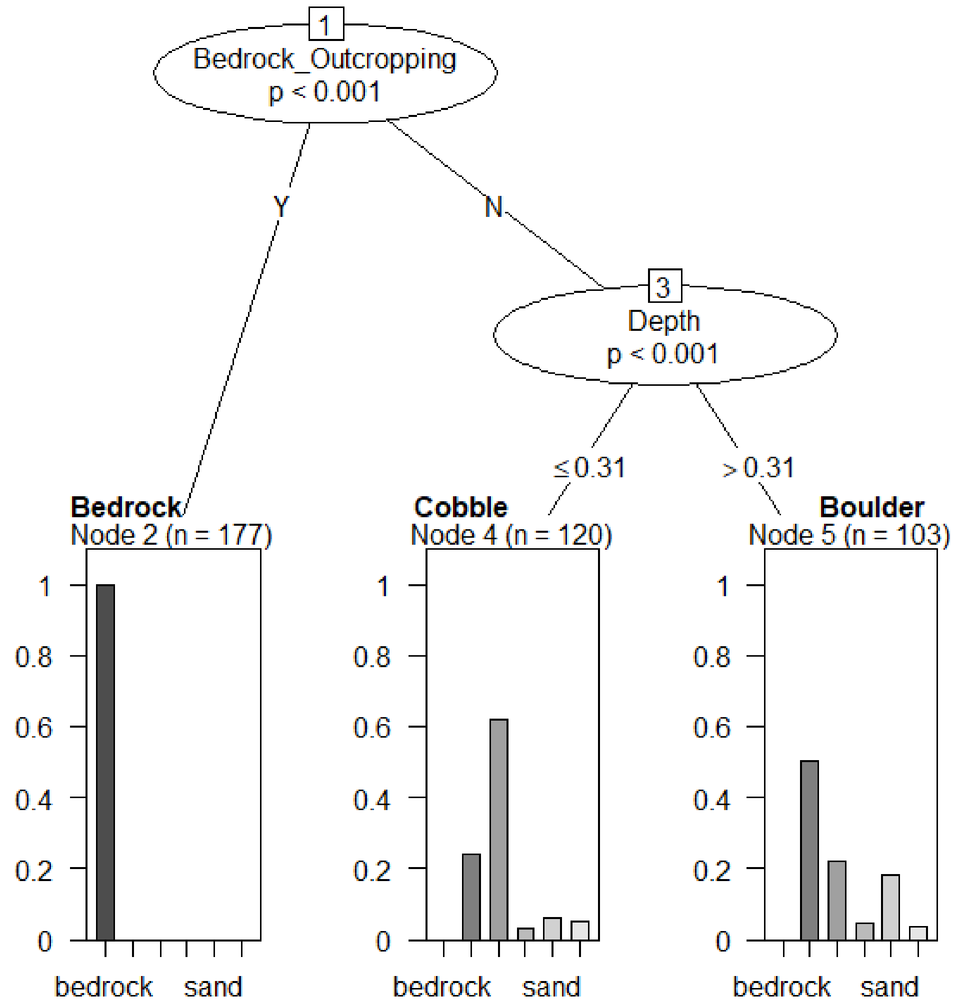

Figure 13: Final substrate ctree model results. P-values are indicated at each split, and bar charts representing substrate percentages are at the end of each node. Overall model accuracy was 75.6%. A total of 10,000 model simulations were ran using a train-test split of 400/100 observations (400 for training, 100 for testing).