EXAMINING THE SPREAD OF AQUATIC INVASIVE SPECIES, *Bythotrephes longimanus*, IN INLAND LAKES ACROSS ONTARIO

by

Joseph Arambulo

B.Sc. Ryerson University, 2016

A thesis

Presented to Ryerson University

in partial fulfillment of the

requirements for the degree of

Master of Applied Science

in the Program of

Environmental Applied Science and Management

Toronto, Ontario, Canada, 2019

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by the examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of the other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

ABSTRACT

EXAMINING THE SPREAD OF AQUATIC INVASIVE SPECIES, *Bythotrephes longimanus*, IN INLAND LAKES ACROSS ONTARIO

Joseph Arambulo

Master of Applied Science, 2019

Environmental Applied Science and Management

Ryerson University

The purpose of this study is to examine the secondary spread of *Bythotrephes longimanus*, commonly known as spiny water flea, across inland lakes in Ontario, and potentially determine predictors for the its invasion. Data for 190 inland lakes across 84 quaternary watersheds in Ontario were included in the database. Global Moran's I was used to analyze the spatial autocorrelation of the variables, and McFadden's Rho-Squared was used to determine if a variable was a predictor of invasion. Three independent variables, out of 28, were found to be good predictors of invasion: (1) mean temperature for watersheds during summer (MNTMPWSSU), (2) mean precipitation for watersheds during spring (MNPCPWSSP), and (3) mean precipitation for watersheds during summer (MNPCPWSSU). Of the three, mean precipitation for watersheds during summer was determined to be the best predictor.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my gratitude to Dr. Richard Ross Shaker for giving me this opportunity and guiding me throughout. As well, the faculty members of the Environmental Applied Science and Management program for giving me valuable knowledge during my time in the program.

Second, thank you to my committee members Dr. Wayne Forsythe and Dr. Eric Vaz for their critical observations and valuable input to my work.

Third, to my professors in the undergraduate department of Chemistry and Biology at Ryerson University. You have given me the inspiration to pursue this branch of science and have propelled me on a path that I am grateful for.

Lastly, to my family and friends who have given me support. Thank you for lifting me up and being there for me, especially during the roughest period of my journey.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0 INTRODUCTION

Invasive species are exotic or non-native species that have successfully managed to survive and thrive in a new environment (Kerr, Brousseau, & Muschett, 2005; Usher, 1988; Usher, Kruger, Macdonald, Loope, & Brockie, 1988). Successful invasions, given enough time, lead to significant changes in ecosystems such as a loss of biodiversity (Jarvis, 1979; Lachner, Robins, & Courtenay, 1970; Usher, 1986; Usher et al., 1988). Biodiversity is defined as the overall diversity of organisms in an ecosystem. Invasive species, alongside habitat destruction, are considered to be the leading cause of biodiversity loss (Allendorf & Lundquist, 2003; Havel et al., 2015; Pejchar & Mooney, 2009).

Freshwater aquatic ecosystems, in particular, are vulnerable to invasive species and their impacts (Carpenter, Lathrop, & Street, 1999; Havel et al., 2015; Pejchar & Mooney, 2009). This puts Ontario at risk to aquatic invasive species (AIS) because of its abundant freshwater resources. According to Environment and Climate Change Canada, Canada holds approximately 20% of the world's total freshwater; 15% of which is accessible to Ontario – encompassing approximately 250,000 inland lakes, and having access to four of the five Great Lakes (Kerr et al., 2005).

Although not solely responsible, human activity has greatly influenced the distribution and introduction of AIS. In its earliest form, human-aided distribution of exotic species were theorized to have began during migrations (Lachner et al., 1970; Usher, 1988). Early human civilizations would transport animals and plants that were deemed helpful for survival. During the late 1800s and early 1900s, it was common

1

practice to deliberately introduce exotic species, without the backing of proper scientific research, for recreational purposes and as counter measures to bothersome organisms (Jarvis, 1979; Lachner et al., 1970). Another example of human aided dispersal of species would be trading via ships, which carries and discharge ballast water – containing organisms – from one part of the world to another (Havel et al., 2015; Jarvis, 1979; Kerr et al., 2005; Lachner et al., 1970; Mills, Leach, Carlton, & Secor, 1993, 1994; Muirhead & MacIsaac, 2005). Human's advance in aviation technology, and the rise in its use as a form of transportation, also increased geographic connectivity at a greater rate than traditional seafaring. As a result, an increase in successful introductions were observed throughout the world. Specifically, Ontario and the Great Lakes have had several encounters with aquatic invasive species throughout the years.

1.1 Invasive Species in the Great Lakes

Mills et al. (1994) noted that the Great Lakes have struggled with invasive species since the 1800s. In 1831, a French immigrant brought along the Eurasian common carp – Asian carp that, for a long time, was bred and cultivated in Europe – to breed in his/her New York pond (Lachner et al., 1970; Mills et al., 1994). Around 1870, it was purposely stocked in lakes due to its perceived appeal as a food source, as well as for countermeasures to bothersome aquatic plants. It's appeal as a food source was not well received and its adverse effects became prominent, such as oophagous behaviours and habitat destruction of more commercially productive native fishes (Jarvis, 1979; Lachner et al., 1970; Mills et al., 1994). The sea lamprey is another example of an invasive species with devastating effects (Aron & Smith, 1971; Lachner

2

et al., 1970; Mills et al., 1994). Sea lampreys are believed to have been introduced to the Great Lakes through ballast exchange of ships. Their introduction was theorized to have begun since 1829 in the Welland Canal, a passage that connects St. Lawrence River and Lake Erie, when trading ships would need to change ballast water during passage (Aron & Smith, 1971; Lachner et al., 1970; Mills et al., 1994). However, sea lamprey's was not observable until a century after where its presence destabilized the fishing industry by causing a decline in the lake trout populations to unfishable levels (Aron & Smith, 1971; Lachner et al., 1970; Mills et al., 1994).

Around 1993, the Great Lakes were found to have contained 139 invasive species – a mix of fauna and flora (Mills et al., 1994). However, Ricciardi (2001; as noted by Kerr et al., 2005) noted that around 2000s the total number of AIS in the Great Lakes increased to 164 fauna and flora. As population increased in Ontario so did reliance on freshwater resources which ultimately lead to the further spread of AIS. Recreational activities – i.e. boating and fishing – increased the propagation of these organisms across the Great Lakes and inland lakes (Kerr et al., 2005; MacIsaac, Borbely, Muirhead, & Graniero, 2004; Muirhead & MacIsaac, 2005). Moreover, the utilization of inland lakes led to an increase in the distribution of AIS and has been described as hubs for invasion (Muirhead & MacIsaac, 2005). For example, Lake Muskoka in Ontario was documented to have propagated 39 further invasions of *Bythotrephes longimanus*, commonly known as spiny water flea (MacIsaac et al., 2004). Currently, the Great Lakes are recorded to have 187 AIS that are considered established (GLANSIS, 2017).

1.2 <u>Pathways to Introduction, the Process of Invasion, and Response to Invasion</u>

Kerr et al. (2005) reviewed the probable pathways of introduction of aquatic invasive species into Ontario. They enumerated 8 pathways; (1) shipping and ballast water; (2) canals and diversions; (3) fish stocking programs; (4) the aquarium and ornamental pond industry; (5) the bait industry; (6) live food fish industry; (7) recreational boating; and (8) private aquaculture. There are three steps involved in invasion. The first is initial dispersal or the displacement of an exotic species from its native area to a new area. Second is establishment, or the event when the invasive species' population growth becomes higher than its death. The third step is the spread/secondary spread (spread and secondary spread will henceforth be used interchangeably) of the invasive species. This step begins when the invasive species spread into nearby similar habitat in its new area (Hulme, 2006; Pimentel, Zuniga, & Morrison, 2005; Puth & Post, 2005).

Complete removal of an invasive species is almost impossible once established, so preventing rather than curing should be given focus (Carpenter et al., 1999; Kerr et al., 2005; Mack et al., 2000). Prevention is a key practice for initial dispersal and secondary spread. It includes keeping track of and predicting the advance of invasive species. In giving guidelines to predict which species can become invasive, Ricciardi & Rasmussen (1998) highlighted three methods. The first is using observed biological traits from a recorded invasive species to determine its "invasiveness"; traits such as resilience, rapid growth, and higher reproductive capability. The second is recognizing regions that have greatly contributed to recorded invasions and the likely ways of

introduction. Last is by reviewing invasion history of an invasive species; a species that is recorded to have successfully invaded a region can be considered a potential invader to another region. Historically, prevention is practiced to halter initial dispersal as this was found to be the most cost-effective way to deal with invasive species (Mills et al., 1993; Puth & Post, 2005). However, the species that managed to successfully disperse before preventative steps were even applied, along with the established and spreading invasive species, have shown to be problematic and costly. In a review of Ontario's struggle with AIS, Kerr et al. (2005) noted that Canada spends approximately $500 million per year just for controlling AIS. MacIsaac (2003; as noted by Kerr et al., 2005) estimated that Canada spends $750 million for damage mitigation of aquatic invasive species' impacts.

In a review of 873 scientific studies from the mid-1990s' to mid-2000's about invasive species, Puth & Post (2005) found that 73.1% of the published papers addressed the establishment of invasive species and its impact, 25.7% addressed the secondary spread of invasive species, and 11% addressed the initial dispersal of invasive species. Following their findings, the researchers concluded that each stage is important and decreasing efforts on one stage to increase the efforts of studying another would not do any good. However, they did point out the importance of the initial dispersal stage for cost-effective management strategies. Another point that was discussed is the similarities and differences between initial dispersal and secondary spread. Ultimately, the researchers settled on the fact that viewing an invasion on a temporal scale or considering the history of a species in an ecosystem will dictate its classification as either a native species or an invasive species.

1.3 <u>Invasive Zooplankton, *Bythotrephes longimanus*</u>

Invasive fish, plants, and molluscs has been the focus of many published articles throughout the years. Kerr et. al (2005), noted that in 2001 around 9% (14 out of 164) of the recorded aquatic invasive species in the Great Lakes and in inland lakes in Ontario are "miscellaneous invertebrates". This disparity in recorded invasions is reasonable because some invertebrates, particularly zooplanktons, are microscopic in size which draws attention away from them and add to that the shroud that being underwater provides. Nevertheless, zooplanktons play a big role in an aquatic ecosystem.

Figure 1. *Bythotrephes longimanus*, spiny water flea. Image duplicated, with permission, from Yan et al. (2011). (a) Mature spiny water flea, photographed by Bill O'Neill; (b) spiny water flea inside a lake herring's stomach, photographed by Bev Clarke; (c) sample of a spiny water flea retrieved from ballast water, photographed by Hugh MacIsaac; (d) a fishing cable covered by spiny water flea, photographed by A. Jaeger; (e) spiny water flea collected from a larval fish drift net, photographed by OMNRF.

Bythotrephes longimanus, commonly known as spiny water flea, is an invasive zooplankton that is well studied. Refer to Figure 1 for a depiction of spiny water flea. In 2002, a review of published literature on marine and freshwater invasive zooplanktons noted that 38 publications about *B. longimanus* were currently published (Bollens, Cordell, Avent, & Hooff, 2002); 38 out of 252 publications, with a total of 63 different species. In 2011, Yan et. al (2011) wrote a review of knowledge regarding spiny water flea. They found that ever since the 1950's approximately 250 articles were published about *B. longimanus*. However, of those publications only 130 journals are specifically about *B. longimanus* as an aquatic invasive species. They also noted that spiny water flea's extent in inland lakes across Ontario, as well as in inland lakes of neighbouring American states such as Wisconsin and Michigan, is progressing. In 2010, Yan et. al (2011) recorded that around 150 Ontario lakes were confirmed to be invaded. *B. longimanus* is known as a zooplanktivore (Bur, Klarer, & Krieger, 1986). As an aquatic invasive species, it has resulted in the decline of its prey's species richness (Yan, Girard, & Boudreau, 2002); ultimately affecting planktonic diversity negatively.

The planktonic community is crucial for the flow of energy and nutrients, and an overall healthy ecosystem. Planktons account for primary production in the pelagic zone of aquatic ecosystems (Yan, Leung, Lewis, & Peacor, 2011). Phytoplanktons gather energy from the sun, which is transferred, although not exclusively, to zooplanktons through the food chain. Research shows that plankton diversity in freshwater ecosystems is linked to efficient nutrient cycling (Cardinale, 2011; Yan et al., 2011). The greater the biodiversity, the more efficient nutrient cycling becomes. Diversity ensures that at least one species is present in every possible niche in an ecosystem, and that

each of these species are making use of every bioavailable resource (Cardinale, 2011). The success of invasive zooplanktons are increasing in occurrence (Yan et al., 2011), and it is threatening the diversity of planktons in aquatic ecosystems (Yan et al., 2002).

1.4 <u>Rationale for the Study, Objectives, and Goals</u>

Natural resources are abundant in Ontario, especially freshwater. The vast amount of inland lakes coupled with current, as well as future, environmental threats leave Ontario vulnerable to aquatic invasive species. Environmental threats such as mismanagement of natural resources, climate change and the shifting historical range of organisms, population growth of humans accompanied by increasing development, and chemical and physical alteration of inland lakes and other bodies of freshwater. As pointed out in the previous section, there is room for more studies on secondary spread of invasive species. Moreover, invasive species in freshwater ecosystems are studied less than invasive species in terrestrial ecosystems. Puth & Post (2005) found that 240 out of 873 (27.5%) published articles studied invasive species in freshwater ecosystems, whereas 440 out of 873 (50.4%) studied invasive species in terrestrial ecosystems; the rest of the published articles were of marine ecosystems and undefined ecosystems.

Reviewing current literature regarding the secondary spread of aquatic invasive species in inland lakes across Ontario highlighted a deficit. A few studies have been done on secondary invasion in Ontario with a big sample size of inland lakes (Gertzen & Leung, 2011; L. Wang & Jackson, 2011; Weisz & Yan, 2010), however all the lakes were within one watershed. Several studies discussed how the presence of certain lake

and landscape characteristics can create an accommodating environment that leads to the successful establishment of invasive species (Buchan & Padilla, 2000; Roley & Newman, 2008). Other studies were able to examine the correlation between certain characteristics and a lake's invasibility, and were able to develop models to map and predict the secondary spread of invasive species (MacIsaac et al., 2004; Muirhead & MacIsaac, 2005; Shaker, Rapp, & Yakubov, 2013; Shaker et al., 2017). But a question that remains to be addressed is whether the correlations still hold true if scrutinized across a bigger study area. *B. longimanus* was chosen as the invasive species to be examined because it is a known threat to freshwater ecosystems in Ontario due to its ability to spread rapidly, and its devastating impact on the native planktonic community. The goal of examining the spread of *B. longimanus* in inland lakes across Ontario was set and the following objectives were identified:

(1) to create a database of *B. longimanus* distribution across several watersheds in Ontario, and characteristics of several inland lakes and its landscape across Ontario;

(2) to analyze the correlation between the dependent variable (presence or absence of *B. longimanus*) and independent variable (lake and/or landscape characteristics); and

(3) to determine potential predictors of spiny water flea invasion.

The results of this research are also aimed at aiding natural resource management by making use of inexpensive, archived, and publicly available data. To fulfill the goals and

objectives a database will be built, and then analyzed using by spatial autocorrelation and logistic regression.

2.0 LITERATURE REVIEW

2.1 <u>The Journey of an Invasive Species</u>

Introduction of an invasive species is not simple, and it does not always lead to success. Several steps are involved, beginning with transportation from its native area to successfully establishing in a new area (Mack et al., 2000). The distribution or transportation of these organisms is mainly attributed to human activity (Allendorf & Lundquist, 2003; Havel et al., 2015; Mack et al., 2000; Pejchar & Mooney, 2009; Shaker et al., 2017; Usher, 1988); that is carrying and trading goods via ships, accidentally or deliberately carrying exotic species while traveling long-distance, and the deliberate introduction of exotic species as countermeasures. The earliest anthropogenic spread of organisms – domesticated species and their parasites – can be traced back to the migration and trade of ancient civilizations (Mack et al., 2000). European settlers in the 1500s brought with them crops and livestock that would assist with their living (Lachner et al., 1970; Mack et al., 2000; Usher, 1986, 1988). This is followed by a further rise in migration, trading and commerce due to innovations in shipbuilding (Mack et al., 2000). Ships carry exotic species purely by accident; i.e. rodents can hitch rides from one continent to another, and ballast water containing exotic species are released from one port to another (Kerr et al., 2005; Lachner et al., 1970; Mack et al., 2000; Mills et al., 1993, 1994). Barring human influence, natural processes would aid in accidental migration of organisms which include displacement via marine debris, predation (i.e. birds dropping prey accidentally), and hitchhiking (i.e. seeds binding to fur) (Jarvis, 1979; Lachner et al., 1970). Examples of deliberate introduction include Japan's Pacific

oyster (*Crassostrea gigas*), bumble bees in New Zealand, several plant species in the United States of America (Mack et al., 2000), and the Eurasian common carp in North America (Lachner et al., 1970; Mills et al., 1994).

Most of these organisms do not survive en route to a new location. Upon arrival, they face physical barriers and/or biotic agents that inhibit invasion – i.e. temperature, light availability, presence of predators or pathogens – by causing immediate extirpation (Mack et al., 2000). The lucky few that do survive will have a chance to reproduce. However, the offsprings will still attempt to survive in the new area and will most likely die off immediately or after a few generations (Mack et al., 2000). Repeated introduction of these organisms is required to keep their population numbers stable in the new area (Mack et al., 2000; Mills et al., 1993, 1994). The moment that these organisms' population become self-sustaining – in the sense that repeated introduction is no longer necessary – solely by competing for resources then they have successfully become invasive species (Carpenter et al., 1999; Mack et al., 2000). Laurenson & Hocutt (1986) highlighted factors that could affect the success and survival of an invasive species, some are reproductive strategy of the organism (*r* or *K*-strategy), method and frequency of dispersal, fecundity, growth rate, the amount of time needed to become sexually mature, competition in the new environment, predation, parasitism and disease, physical and chemical features of the new environment, and availability of resources. Once a species establishes itself, and therefore successfully invading, it will then begin to spread further into nearby similar habitats in the new area and whole process will restart but will less barriers to slow it down.

2.2 <u>History of Aquatic Invasive Species (AIS) in the Great Lakes</u>

Species invasion in the Great Lakes, and inland lakes subsequently, is speculated to have begun right after the Wisconsin glacial ice stage receded, sometime between 14,000 and 4,000 years ago, and carved out basins that aided in organisms' establishment and propagation (Mills et al., 1993). These basins are what we recognize now as lakes. This event was followed by the transport of medicinal plants and domestic animals as First Nations searched for viable land to live in (Mills et al., 1993). However, European colonization and settlement quickly heightened species invasion. A common pattern for settlers throughout history is to aggregate and build near bodies of water. The French settlers resided near the Great Lakes to exploit their bountiful ecosystem during the 17th century (Mills et al., 1993). English settlers, after defeating the French, then took control over the Great Lakes. Both nations carried over technology that eventually lead to the opening of the St. Lawrence River canal system (antecedent to St. Lawrence Seaway), the Lake Erie Canal, and the Welland Canal around mid-1800s (Mills et al., 1993, 1994).

The Welland Canal connected Lake Erie to Lake Ontario, the Lake Erie Canal connected Hudson River to Lake Erie, and the St. Lawrence River canal system made the Great Lakes and distant North American communities accessible to European ships for trading (Aron & Smith, 1971; Mills et al., 1993). Both waterways increased the transferability of AIS from one body of water to another (Mills et al., 1993). Since ships around this time (mid-1800s) used solid ballast rather than water ballast, only fouling organisms were suspected to be introduced – i.e. sea lampreys (Mills et al., 1993,

13

1994). According to Mills et al. (1994), the following changes occurred around 1870 to 1930s. Technological advancement lead to ships using water ballast, which allowed the unintentional introduction of zoo- and phytoplankton, and pathogens. Fish introductions were implemented by government agencies – such as brown trout (*Salmo trutta*), rainbow trout (*Oncorynchus mykiss*), common carp (*Cyprinus carpio*), mosquitofish (*Gambusia affinis*), sunfish (*Centrarchus spp.*), and chinook salmon (*Oncorhynchus tshawytscha*). And minor improvements to the St. Lawrence River canal system which slightly increased the movement of trading vessels through the Great Lakes.

As noted by Mills et al. (1993, 1994), on June 26, 1959 the new and improved St. Lawrence River canal was opened for usage, and aptly given the new name St. Lawrence Seaway. Improvements to the canal focused on making it wider, which allows more trading vessels to pass, resulting to a boost in the economy for ports near the Great Lakes – such as Toronto, Duluth, and Chicago. This also meant that ballast water exchange was occurring at a higher rate than ever, and that invasive species establishment was inadvertently being promoted. As a result, around the 1970s, invasive phytoplanktons – i.e. diatoms – were introduced into the Great Lakes via ballast water exchange. Around the 1980s, accidental release of invertebrates and fish increased in occurrence because of ballast water exchange, and at the same time ship traffic in the St. Lawrence Seaway was its highest. This allowed the establishment of invasive species such as the Eurasian ruffe, gobies, spiny water flea, and zebra mussels. Around the same period, government agencies were still deliberately stocking fish into the Great Lakes and inland lakes, such as pink salmon's introduction into Lake

14

Superior, and the accidental introduction of exotic species via bait fishing started to occur, such as in the case of discovering ghost shiners in the Great Lakes.

Some of the worst cases of AIS and its impact on the Great Lakes include the alewife and the sea lamprey (Aron & Smith, 1971). Alewives were first observed in Lake Ontario around 1873. It was suspected to have been brought into Lake Ontario via the Lake Erie Canal because predatory fishes were noted to be abundant along the St. Lawrence River which made it difficult for alewives to migrate naturally. The presence of alewives was initially viewed positively as they were perceived to possibly become additional sustenance to valued native fishes. However, around 1870 up to 1890, the population of Atlantic salmon, lake trout, lake herring, whitefish, walleye, bass, and pike were rapidly declining while alewife population were steadily increasing. Alewives did not directly cause the population decline of the aforementioned native species as their population were already in its downward path even before alewife population increased, but the alewives did however compete with the native species for available resources (Aron & Smith, 1971).

Sea lampreys were observed in Lake Ontario at around 1880, although it is believed that they have been present in Lake Ontario a few years prior and managed to stay out of sight because of their long life cycle (Aron & Smith, 1971). At around 1885, there was a noticeable increase in sea lamprey population. It increased steadily up to the late-1890s. Alongside, native fish population were observed to be decreasing. Sea lamprey directly attributed to the decrease of native fish population – and ultimately destabilizing the fishing industry – by parasitizing or latching onto its prey and feeding

off of blood. Its effects were observed in other lakes such as Lake Huron, Michigan, and Superior. The reported lake trout catch in Lake Huron by the year 1950 was at its lowest at <1 (its highest was 3,812 caught lake trout in 1935), while the number of caught sea lamprey was recorded to be 210 during the same year. There were no reports of sea lamprey caught in fishing nets in Lake Huron until the year 1944, which was recorded to be 50 sea lampreys. In Lake Michigan, the lake trout catch was recorded to be <1 in 1953 (its highest was 6860 caught lake trout in 1943), while the number of caught sea lamprey during the same year was at 252. The earliest reported catch of sea lamprey in Lake Michigan was in the year 1945, which was reportedly three sea lampreys. In Lake Superior, the number of caught lake trout was reported to be 228 in 1966, while the number of caught sea lamprey during the same year was at 23. The year with the highest reported catch of lake trout in Lake Superior was in 1946 with 4,975 lake trouts, and the first year with reported catches of sea lamprey was in 1953 with 24 catches (Aron & Smith, 1971).

2.3 *Bythotrephes longimanus*, Spiny Water Flea

Bythotrephes longimanus is an endemic species to deep oligotrophic lakes of Eurasia, such as lakes in Sweden and water reservoirs near Moscow (Berg & Garton, 1988; Lange & Cap, 1986; Sprules, Riessen, & Jid, 1990; Yan, Dunlop, Pawson, & MacKay, 1992). It is theorized to have made the journey from Eurasia to North America through accidental dispersal on ballast water exchange. The earliest recorded sightings of Spiny water flea in Ontario dates back to 1984 in Lake Huron, and 1985 in Lake Erie (Bur et al., 1986). At first it was thought that the collected organism was of the genus

16

Bythotrephes, but of the species *cederstroemi*. The species *cederstroemi* and *longimanus* belong to the same cladoceran (minute crustaceans) family Cercopagidae and the same genus. It was originally thought that the two species were distinct due to their different morphology. Specifically, the caudal spine or the tail of *cederstroemi* is bent in an S-pattern and having 1-3 pairs of lateral spine near the bend, whereas *longimanus* has a somewhat straight caudal spine (Berg & Garton, 1988; Bur et al., 1986; Lange & Cap, 1986). However, genetic analysis show that the two species are the same and that *cederstroemi* is a polymorphic form of *longimanus* to adapt to warmer water temperatures (Berg & Garton, 1994; Bur et al., 1986; Lange & Cap, 1986; Yan et al., 2001). By 1989, this Eurasian invader was found to have spread to all of the Great Lakes through hitchhiking on fishing gear and watercrafts (Yan et al., 2001, 2002, 2011).

Bythotrephes longimanus are larger than any other cladoceran in North America. Samples collected near a cooling water intake in Somerset, New York were observed to have an average length of 11.9 mm (Lange & Cap, 1986); Berg & Garton (1988) noted that the samples they collected were all >10 mm in length. Its body comprises 1/3 of its length, and its caudal spine makes up the remaining 2/3 (Berg & Garton, 1988; Lange & Cap, 1986). The number of pairs of lateral spine present on its spine can be used to ascertain its age; neonates have one pair, and it gains a new pair per molt – it can molt up to two times in its life cycle (Berg & Garton, 1988). On its head is a small pair of antennae, and eyes that occupy most of its head. Located on the dorsal side of its body is a protruding brood sac (Lange & Cap, 1986). It exhibits sexual dimorphism in that the first limb of males would have a hook on the proximal end of its last endopodite

segment (Berg & Garton, 1988). Figure 2 illustrates photograph taken from by a microscope of *B. longimanus*.

Figure 2. Spiny water flea under a microscope. Image owned by J. Liebig from NOAA GLERL ("Spiny and Fishhook Waterfleas," 2001.).

It is a zooplanktivore with a particular taste for smaller cladocerans, copepods, and rotifers (Berg & Garton, 1988; Lange & Cap, 1986). It inhabits the water column, and it has a diurnal vertical migration feeding pattern wherein it vertically migrates closer to the surface at night to feed (Berg & Garton, 1988). There were opposing theories as to how *B. longimanus*' diet would impact native zooplankton community in the early 1990s. One is that it would negatively impact native zooplankton abundance as it does not have a known zooplankton predator in North America and as such will compete with planktivorous fish for resources (Lehman & Caceres, 1993); the other is

that *B. longimanus* will not be successful, that it will not reach a high enough population density to negatively impact the native zooplankton community (Sprules et al., 1990).

A laboratory experiment done in the late 1990s found that *B. longimanus*, in high densities, had a significant negative effect on both large and small zooplankton density; furthermore, the researchers also found that the presence of zooplanktivores in high population densities for a long period of time may negatively impact phytoplankton communities (Wahlström & Westman, 1999). Predation on spiny water flea was observed to be complex. Native fish (such as rainbow trout, deepwater sculpins, yellow perch) and invasive fish (such as alewife) were found to prey on spiny water flea, but fish size, developmental stage, and learned aversion seems to affect predation behaviour (Barnhisel & Harvey, 1995). Planktivorous fish that are <10 cm in length were observed to avoid spiny water flea due to the cladoceran's length and barbed caudal spine. The developmental stage of a fish determines its size, as well as the width of its jaw. As a result, older, bigger planktivorous fish are more indiscriminate to what they eat and are able to encounter spiny water flea more often due its better swimming capabilities. Learned aversion was noted to negatively impact preference for preying on spiny water flea as when a younger, smaller planktivorous fish attempts to eat the cladoceran it ends up getting hurt by the caudal spine and its barbs. Recorded reactions to failed attempts include whole body convulsions, and regurgitations. The fish then learns to avoid the cladoceran permanently or temporarily.

Bythotrephes longimanus can reproduce both asexually and sexually; females undergo parthenogenesis during times with low population density, and sexual

reproduction is preferred during times with higher population. The sexually produced eggs can undergo a resting egg state wherein the eggs are covered in a tough outer casing to lie dormant through unfavourable conditions (Berg & Garton, 1988; Lange & Cap, 1986; Sprules et al., 1990). Females are more predominant during summer, and the numbers of males increase as fall arrives due to fall conditions beings more favourable (Berg & Garton, 1988; Sprules et al., 1990). In a few studies to observe *B. longimanus* population and how it is affected by the changing seasons, it was found that population tends to be lower during summer and higher during fall (Berg & Garton, 1988; Garton, Berg, & Fletcher, 1990; Yan & Pawson, 1998). Parthenogenesis is the reproductive method used during summer, and sexual reproduction becomes more available during fall.

Water temperature has been observed to impact spiny water flea. Researchers caught 290 spiny water flea/m^2 in Lake Michigan while the surface water temperature was at 15˚C, while 80 spiny water flea/m^2 were caught in Rybinsk Reservoir near Moscow while the surface water temperature was at 21˚C (Berg & Garton, 1988). It reaches sexual maturity at a faster rate at temperatures around 15˚C where it becomes sexually mature in 13 days, but at a slower rate at temperatures around 8˚C where it becomes sexually mature in 26 days (Sprules et al., 1990). Egg development times are also affected similarly; from 8 days to 16 days at the same respective temperature (Sprules et al., 1990). Spiny water flea population was observed to decrease in temperatures above 25 ˚C (Yan et al., 2011).

2.4 <u>Impacts of Aquatic Invasive Species (AIS)</u>

Aquatic invasive species (AIS) are capable of changing the ecological, physical, and chemical conditions of a freshwater ecosystem. Successful invasion could lead to: ecosystem degradation; alteration of goods, through competition and alteration of trophic levels; alteration of ecosystem services, through negatively impacting drinking water quality and nutrient cycling; economic impacts; ultimately, a negative impact on human-wellbeing (Havel et al., 2015; Pejchar & Mooney, 2009; Pyšek & Richardson, 2010). As previously noted, once established an invasive species is almost impossible to eradicate (Carpenter et al., 1999; Kerr et al., 2005; Mack et al., 2000).

Any invasive species can negatively impact biodiversity as a result of the invading species outcompeting the native species and devastating the established trophic chain, altering the nutrient cycle, or even by preying on the native species (Havel et al., 2015; Paul & Kar, 2016; Pejchar & Mooney, 2009). Gutierrez et al. (2014) described the impacts of invasive species as direct or indirect. The authors defined direct impact as "changes in the stocks and transformations of energy and materials resulting solely from the presence and/or activities of the invasive species", or if the exotic species' diet (food translates to energy) is impacting the pre-established food web and directly causes an imbalance in the availability of food. For example, spiny water flea are voracious zooplankton eaters; altering the trophic pyramid and heightening the competition for other zooplankton with the same diet.

Gutierrez et al. (2014) defined indirect impacts as "when the invasive species alters the abundance and/or activity rates of one or more other species and, in so doing,

21

modulates their impacts on the stocks and transformations of energy and materials". In other words, indirect impact is observed when an exotic species' diet is inadvertently impacting the availability of food and energy to species higher in the food web. A good example of this, which Gutierrez et al. (2014) pointed out, is the effect of zebra mussels on the turbidity and light penetration in a lake; zebra mussel indirectly promote the growth of macrophytes by voraciously and indiscriminately feeding on any small particles – such as phytoplanktons and silt – within its surrounding, which greatly decreases turbidity and allowing more light to penetrate (more available energy for the macrophytes) in the littoral area.

Examples of impacts specific to the Great Lakes include (Mills et al., 1993, 1994): the introduction and stocking of salmonids for recreational activities and for the fishing industry which resulted in competition for resources and the native fish population struggling further. Fish stocking also introduced novel parasites and diseases to the native fish population. There is also the case of the European ruffe being introduced into the Great Lakes which competed for resources with yellow perch and walleye, the latter two species being important for the fishing industry. Invasive gastropods that caused biofouling of intake pipes, and invasive bivalves that out competed native species as well as biofouling any hard surface. Lastly, invasive aquatic flora that deterred recreational use of lakes by making the surface water or the nearby sediment unbearable and difficult to use.

2.5 <u>Previous Studies with the Theme of Predicting or Examining Invasion of AIS</u>

Past studies that attempted to predict or examine secondary invasion can be divided into three categories depending on the method used. The first one is trait-based method wherein studies that fall within this category use the invasive species' known history and characteristics (Gertzen & Leung, 2011). Researchers would compile data on the species, such as temperature preference, prey, and predator, and then use those to predict if the species can become invasive in a new environment. The second method is niche-based method wherein studies that employ this method would look into the invasive species' native environment and its characteristics (Gertzen & Leung, 2011), such as climate, and available resources. These would then be cross referenced to potential new environment to predict if it is favourable for successful establishment of the invasive species. The last method is called propagule-pressure method wherein researchers focus on the dispersal of viable population of the invasive species to new locales (Gertzen & Leung, 2011). The movement of invasive species, both from an invaded habitat or into an uninvaded habitat, will be tracked and used to predict secondary spread of invasive species. The following summarized studies fall within one of the three categories, and some even contain elements of multiple methods.

MacIsaac et al. (2004) predicted future, as well as backtracked, invasions of Spiny water flea in inland lakes across Ontario by creating a gravity model that tracked the inflow and outflow of boats from invaded and uninvaded lakes. The researchers gathered the invaded and uninvaded lakes data from government agencies, non-government organizations, and unpublished data from another researcher. The boat

inflow and outflow data were gathered through surveying anglers, boaters, and other recreational users of lakes. Muirhead & MacIsaac (2005) followed up the aforementioned research by building up on the gathered data of invaded/uninvaded lakes and the outflow of boats from invaded lakes. Additional surveys in marinas and boat ramps about which lakes are frequented by boat owners were added to the data. As a result, the researchers were able to narrow down which lake will potentially become a source for the spread of spiny water flea.

Gertzen & Leung (2011) predicted the secondary spread of invasive species using the propagule pressure method, or by focusing on modes of dispersal, and chose the species *Bythotrephes longimanus* as the test subject. The sources of dispersal were recreational boat traffic and streams. They gathered data from government agencies, past research, the Canadian Aquatic Invasive Species Network, and surveys. They were able to gather spiny water flea presence/absence data for 336 lakes in south-central Ontario, 50 of the lakes were known to be invaded. All the lakes belonged in one watershed. They found that streams are unimportant to the secondary spread of spiny water flea, and boat traffic is highly significant. The model they developed also analyzed that the secondary spread of spiny water flea is slowing down because of possible maximum saturation of the lakes within the watershed.

In their research, Wang & Jackson (2011) considered the effects of biotic and abiotic conditions of lakes, and its resulting characteristics, in modelling the establishment of *Bythotrephes longimanus*. The biotic aspect involves recognizing the fish in the lake and its possible effect on the zooplankton population. Their study

focused on lakes in the same watershed from Gertzen & Leung's (2011) research which was discussed in the previous paragraph. The researchers found that Secchi depth, lake productivity, max depth, and lake surface area greatly influenced establishment. Furthermore, they found that the fish community in the lake, and lake water chemistry were significant to spiny water flea invasion.

In an attempt to predict the likelihood of Eurasian watermilfoil invasion in lakes throughout Wisconsin, Buchan & Padilla (2000) created logistic regression models using biological, chemical, anthropogenic, morphological, and physical variables of lakes as independent variables; and presence or absence of Eurasian watermilfoil data as the dependent variable. The data they were publicly accessible, and were gathered from government departments, government-associated organizations, and data gathered from the field. Examples include alkalinity, lake area, max depth, pH, Secchi depth, type of boat launch, and distance to highway. The researchers were able to determine the likelihood of a lake to support Eurasian watermilfoil growth. They went on to explain that the method they used to gather their data was inexpensive, and that this could help with better allocation of funds for the prevention of invasion. Furthermore, they detailed that variables that can affect the growth of Eurasian watermilfoil, such as forest cover, held greater weight as a predictor over the rest. Roley & Newman (2008) also tried to predict the secondary spread of Eurasian watermilfoil, but in Minnesota, using logistic regression models. The data used were also gathered from government and non-government organizations. During their analysis, they found that of all their gathered variables distance to nearest invaded lake, max depth, Secchi depth, duration

of invasion, alkalinity, and lake area lead to significant results for predicting Eurasian watermilfoil invasion.

Shaker, Rapp, & Yakubov (2013) assessed the correlation of lake and landscape data with the occurrence of 9 aquatic invasive species (including spiny water flea, and Eurasian watermilfoil) in 26 lakes within the Adirondack Park near the eastern end of New York state. Each lake was given an aquatic invasive species richness metric, which is the sum of aquatic invasive species in a lake. The rest of the data are lake area, elevation, mean depth, distance to highway (I-87) ramp, total boat ramp; landscape composition such as % private boat ramps, % public boat ramps, % mixed forest, % deciduous forest, % woody wetland, % developed open space, % evergreen forest; and landscape diversity metrics such as Simpson's evenness index, Simpson's diversity index, and relative patch richness. Through ordinary least squares regression analysis, the researchers found that lake area, total boat ramps, relative patch richness, and % developed open space were positively correlated to aquatic invasive species richness. On the other hand, they found that lake elevation, distance from highway (I-87) to boat ramp, Simpson's evenness index, Simpson's diversity index, and % evergreen forest was negatively correlated. Through logistic regression analysis, their findings confirm that the presence of boat ramps promotes invasions. Furthermore, although they found that public and private boat ramps promote invasion equally, public boat ramps were found to impact aquatic systems slightly higher than its counterpart by increasing human activity in and around the lake. The researchers concluded that, based on their results, landscape patterns such as relative patch richness have a higher value for predicting invasion over landscape composition such as % deciduous forest.

26

Shaker & Rapp (2013) published a follow-up article using the aforementioned article's data but with an added focus on determining the correlation between the presence of Eurasian water milfoil and the presence of curly-leaf pondweed, both being aquatic invasive plants, to lake and landscape metrics. The significant data for the presence of Eurasian water milfoil were lake area and lake elevation; lake area being positively correlated to the presence of Eurasian water milfoil, and lake elevation being negatively correlated. For the presence of curly-leaf pondweed, they found that relative patch richness, total number of boat ramps, mean lake depth, % developed open space, mean lake depth, and lake area were found to be positively correlated. Whereas elevation, distance to highway ramp (I-87), and Simpson's evenness index were found to be negatively correlated. They also found that the presence of a lake association significantly affected the presence of both invasive species.

Figure 3. Workflow of the process to fulfill the goals and objectives.

28

Figure 4. Map of Ontario. Source: ArcGIS Online, LIO.

3.0 METHODS

To carry out the goals and objectives, a database must first be built. Beginning with searching for data, followed by processing through ESRI's ArcGIS, and then compiling it into one dataset. The resulting dataset will then be analyzed through the software Statistical Package for the Social Sciences (SPSS) and Spatial Analysis in Macroecology (SAM). Figure 3 illustrates a summary of the Methods section.

3.1 <u>Study Area</u>

The study area was chosen based on the aim of examining inland lakes across Ontario. Figure 4 illustrates a map of Ontario. As the second largest province in Canada, Ontario spans over 1 million km^2 (Government of Ontario, 2017). Along with its large land mass, Ontario relies heavily on its abundant freshwater resources. There are over 250,000 lakes in Ontario which consist of varying temperature due to seasonal changes and varying landscape surrounding them such as urban, forest, and rocky terrain (Government of Ontario, 2017). The natural resources within Ontario drive the economy; the forestry and agriculture industry, both requiring access to good quality freshwater, are important players of Ontario's economy. Hydroelectricity is also a large contribution to the economy, supplying some of Ontario's most remote region electricity. Provincial parks provide a reprieve from city living, as well as recreational activities to approximately 10 million annual visitors (Government of Ontario, 2017).

3.2 Data Search

3.2.1 Invaded and Uninvaded Lakes by _Bythotrephes longimanus_

The spiny water flea sightings map was obtained from EDDMapSOntario. Originally, as an interactive map on a webpage, it illustrates if spiny water flea is present or not present in a lake. Figure 5 is a recreation of the map in EDDMapSOntario's website. It should be noted that the species' presence is based on submitted invasive species sighting reports. These reports are further examined to ensure that the reported description matches the description of a known invasive species. "Not present" depicts when the invasive species sighting report analysis finds that the reported species described is found to be a different species altogether. It is also important to note that after thorough examination of the unprocessed sightings map, one lake could contain more than one presence point, more than one absence point, or a mix of both presence and absence point.

3.2.2 Landscape Characteristics

3.2.2.1 Land Cover 2010

Land Cover 2010, the land cover dataset, was retrieved from the webpage Global Land Cover (GlobeLand30). This consisted of an interactive land cover map of the globe, as well showed a legend for each classification that was distinguishable on the map. In its original form, the dataset was available to download as rasters. A total of

12 rasters were downloaded to cover Ontario's boundaries (raster N15 45, 50, 55; N16 45, 50, 55; N17 40, 45, 50, 55; and N18 40, 45).

Figure 5. Distribution map of Spiny water flea (*Bythotrephes longimanus*) in inland lakes across Ontario. Red points indicate positive sighting, blue points indicate negative sighting.

3.2.2.2 Watershed-Quaternary

Watershed-Quaternary identifies Ontario's quaternary watershed division. Watersheds in Ontario are divided four ways and are categorized based on size. "Primary" identifies the watershed division with the biggest extent or the biggest drainage area, while "quaternary" identifies the watershed divisions with the smallest extent or smallest drainage area (Land Information Ontario, 2015b). Figure 6 illustrates Ontario's 1059 quaternary watersheds.

Figure 6. Quaternary watersheds of Ontario.

3.2.2.3 Precipitation and Mean Temperature

Both these datasets were gathered from and is displayed by Ontario Climate Change Data Portal. The foundation for the dataset was collected from 1986-2005 which was used to create models for that period, and to create future climate models (2006-2100). Seasonal variation are available for both dataset, and were downloaded for this project's use (Wang & Huang, 2015). Through initial inspection, the datasets were in point form; temperature and precipitation were displayed in points that were scattered throughout Ontario.

3.2.3 Lake Characteristics

3.2.3.1 Aquatic Resources Area

Aquatic Resource Area (ARA) was acquired from Land Information Ontario. It is available in survey point, line segment, and polygon segment form. ARA was created to support Ontario's fisheries management, forest management, municipal planning, and other land use planning by archiving spatial fisheries data. It tabulates the existing fish species in a waterbody, such as a stream, river, or lake, within Ontario and its physical characteristics. Data compiled in ARA include "WATERBODY_TYPE" which specifies if the waterbody is a pond, wetland, drain, stream, river, or lake, "SPECIES" which lists the known fish species present in the waterbody, and "REGIME" which details the surface water temperature of the waterbody during summer. The data that were of use for this project's database are "WATERBODY_TYPE", "OFFICIAL_WATERBODY_NAME" which is the government recognized name of the

waterbody, "MAX_DEPTH" which lists the maximum depth of the waterbody, and "MEAN_DEPTH" which lists the mean depth of the waterbody (Land Information Ontario, 2015a).

3.2.3.2 pH, Calcium, and Total Phosphorous

This dataset was acquired from Dorset Environmental Science Centre (DESC). It is a record of pH, calcium (mg/L), and total phosphorous (µg/L) for 835 inland lakes in Ontario from 2008-2012. This data is a part of Ontario's Broad-scale Monitoring program, which is spearheaded by the Ministry of Natural Resources and Forestry with the purpose of supporting fisheries management ("Broad-scale Monitoring Program," 2017).

3.2.3.3 Provincial Digital Elevation Model

Provincial Digital Elevation Model (PDEM) is a dataset that represents Ontario's true ground elevation. Although the elevation values in PDEM were gathered from several other datasets, it was not made for a specific application and has been used for various reasons (Land Information Ontario, 2018).

3.2.3.4 Bathymetry Line

This dataset is the culmination of up to date bathymetry data in Ontario. The data was gathered by the Ontario Ministry of Natural Resources and Forestry from 11,000 lakes between 1948 to 1995 with the intent to support fisheries management, water

resources management, and climate change modelling (Land Information Ontario, 2016).

3.2.3.5 Ontario Hydro Network-Waterbodies

Ontario Hydro Network-Waterbodies (OHN-W) is a dataset containing polygon features or information of waterbodies, such as stream, rivers, ponds, and lakes, throughout Ontario. The data that were of use to this project are "WATERBODY_TYPE" which specifies the surface water type/waterbody, "OFFICIAL_NAME_LABEL" which lists the government recognized name of the waterbody, "SYS_AREA" which lists the area of the waterbody, and "SYS_PERIM" which lists the perimeter of the waterbody (Land Information Ontario, 2010).

3.2.3.6 Cartographic Boundary Files

The Cartographic Boundary Files (CBF) are a part of the 2016 Census Boundary File compiled by Statistics Canada. By means of Canada's coastal islands and major land mass shorelines, the CBF identifies the nation's geographic areas for recognizing locations to which census data are distributed and collected from. The purpose of this dataset is to aid Geographic Information System (GIS) applications for land use management, market, social, and economic research, as well as demographic studies (Statistics Canada, 2017). Figure 7 illustrates geographic areas, in the form of metropolitan areas, designated places, and population centres, that are recorded by the CBF.

Figure 7. Artificial surfaces within the quaternary watersheds of Ontario.

3.2.3.7 Fishing Access Point

This data set was retrieved from Land Information Ontario (LIO). Mapping the distribution of boat access points in waterbodies across Ontario was mainly for the benefit for recreational fishers, experts, or visitors. To that end, this dataset was used to create a web application that can show which lake has an access point. Within the attribute table, the column "POINT_TYPE" represents the fishing access points as points. The categories in this column include "Shoreline Access", "Enhanced Shoreline Access", and "Boat Launch"; where shoreline access represents a patch of sandy shore that anglers can access, enhanced shoreline access describes a shoreline with a small

dock or small pier, and boat launch describes a full-on dock or pier (Land Information Ontario, 2012).

3.2.3.8 Road Network Files

The Road Network Files (RNF) was collected by Statistics Canada in 2016. The intent behind compiling this dataset is to support GIS applications by providing a framework for spatial analysis and geographic mapping which used in economic, social, and market research, as well as land use management and demographic studies (Statistics Canada, 2008). Refer to Figure 7 for an illustration of Ontario's roadways.

3.2.4 Data Sources

3.2.4.1 Early Detection and Distribution Mapping System – Ontario

Early Detection and Distribution Mapping System (EDDMapS) was founded in 2005 at University of Georgia's Center for Invasive Species and Ecosystem Health ("EDDMapSOntario," 2017). Its initial purpose was to better map the distribution of invasive plant species in each state of the U.S.A, but now it has become a user-friendly internet mapping tool that is accessible to the public and encourages public participation ("EDDMapSOntario," 2017). EDDMapSOntario is a sub-program of the aforementioned and is managed by the Invading Species Awareness Program (ISAP) – a program developed in partnership by the Ontario Ministry of Natural Resources and Forestry (OMNRF) and the Ontario Federation of Anglers and Hunters (OFAH) ("EDDMapSOntario," 2017; J. Birnsmead, personal communication, December 19,

2017). The Invasive Species Centre is also one of EDDMapSOntario's partners. At the time of EDDMapS Ontario's creation, all of the existing data on invasive species that were gathered by the ISAP and its partners were added onto EDDMapSOntario's distribution maps. New invasive species distribution data are constantly added (J. Birnsmead, personal communication, December 19, 2017).

EDDMapSOntario greatly relies on the public for mapping invasive species distribution. Any citizen can submit a report, through EDDMapSOntario's website, that details the sighting/encounter of a supposed invasive species ("EDDMapSOntario," 2017). Before getting uploaded to an official distribution map, these reports undergo a confirmation process by an OFAH-ISAP staff ("EDDMapSOntario," 2017; J. Birnsmead, personal communication, December 19, 2017). The protocol for confirmation was developed by OMNRF and OFAH ("EDDMapSOntario," 2017; J. Birnsmead, personal communication, December 19, 2017). In occasions where a report is proving difficult to confirm (i.e. an unfamiliar species), the file is passed onto an OMNRF staff for further help (J. Birnsmead, personal communication, December 19, 2017). If difficulty is further encountered, then the report will be sent to local and overseas experts in taxonomy (J. Birnsmead, personal communication, December 19, 2017). Updating the official distribution map ensues confirmation ("EDDMapSOntario," 2017; J. Birnsmead, personal communication, December 19, 2017).

3.2.4.2 GlobeLand30

Global Land Cover (GlobeLand30) is a website that stores a 30 m resolution landsat dataset called GlobeLand30 (GLC30) ("Globe Land 30," 2017). Landsats are

land cover images taken by satellites. These are crucial for areas of study such as environmental science, resource management, and urban planning and sustainability ("Globe Land 30," 2017). The development of high resolution land cover imagery is necessary so as to get a better look at how land use has changed overtime; the higher the resolution, the higher the accuracy ("Globe Land 30," 2017). During the 2000s, land cover data with a resolution ranging from 300 m to 1 km were widely available ("Globe Land 30," 2017). In 2010, China responded by undertaking a project to map global land cover with high resolution which resulted to the development of GLC30 ("Globe Land 30," 2017). Some of the journals that have published research papers using GLC30 includes Nature, International Journal of Geo-Information, and Remote Sensing of Environment.

3.2.4.3 Land Information Ontario

Land Information Ontario (LIO) is a data repository managed by the Ontario Ministry of Natural Resources and Forestry (OMNRF) ("Land Information Ontario," 2017). Its purpose is to house geographic data pertaining to Ontario, as well as to become a hub for sharing the stored data. Communities and organizations involved in contributing data to LIO includes government agencies, municipalities, conservation authorities, non-profit organizations, Indigenous communities, public organizations, universities or colleges, public utility, and public health units ("Land Information Ontario," 2017). The types of data managed by LIO includes geographic information on waterbodies, laneways and pathways, geographic elevation, boundaries and official

names, and classification and management information ("Land Information Ontario," 2017).

3.2.4.4 Statistics Canada

Statistics Canada (Statcan) is the nation's premier agency for gathering and storing statistical information. The agency's undertaking is to provide "high-quality statistical information" to Canadians to assist in developing a better understanding of Canada's resources, culture, economy, society, and population. The organization believes that having access to objective statistical information creates a well-informed atmosphere amongst not-for-profit organizations, businesses, unions, elected representatives, and individual Canadians; ultimately, having reliable statistical information available is viewed as detrimental to achieving a society that is open and democratic. Statcan actively collects and archives hundreds of data; examples of data categories that are documented include "Aboriginal peoples", "Crime and justice", "Environment", "Government", "Population and demography", "Science and technology", and "Transportation" (Statistics Canada, 2019).

3.2.4.5 Ontario Climate Change Data Portal

Ontario Climate Change Data Portal (OCCDP) documents historical climate data,1986-2005, and future climate projections, 2006-2100. It was developed by University of Regina's Institute for Energy, Environment and Sustainable Communities, and was funded by the Ontario Ministry of the Environment and Climate Change. In the creation of the future models, the RCP 8.5 (representative concentration pathways) was

considered as the emissions scenario/climate forcing scenario. The 1986-2005 climate model was created using the Coupled Model Inter-comparison Project phase 5 (CMIP5), a guideline used to create the Intergovernmental Panel on Climate Change (IPCC)'s Fifth Assessment Report (AR5). The purpose is to make the most updated, and easy-to-use climate data available for inspection and use by professionals and non-professionals. Examples of climate data that are available to view and download include precipitation, temperature, wind speed, humidity, and air pressure (Wang & Huang, 2015).

3.2.4.6 <u>Dorset Environmental Science Centre</u>

Dorset Environmental Science Centre (DESC) is a facility dedicated to environmental science (The Dorset Environmental Science Centre, 2017). Established in the mid-1970s, DESC's initial purpose was to produce much needed scientific backed regulations for controlling cottage development (The Dorset Environmental Science Centre, 2017). Currently, it's known for its expertise on inland lakes and environmental issues that affect inland lakes such as nutrient enrichment related to lakeshore development, invasive species, metal contamination, atmospheric deposition, and climate change; the Ministry of Environment and Climate Change recognizes this and maintains a close partnership with DESC (The Dorset Environmental Science Centre, 2017). Aside from government agencies, DESC is in partnership with academic institutions and non-government organizations to facilitate programs such as Long-term Ecosystem Science, Lake Partner Program, and Ontario Benthos Biomonitoring

Network (The Dorset Environmental Science Centre, 2017). Table 1 provides a

summary of the collected data and its sources.

Table 1. Gathered data (Section 3.2) and the source (Section 3.3).

Data	*Source*
Invaded and uninvaded lakes	EDDMapSOntario
Land Cover 2010	GlobeLand30
Precipitation and Mean Temperature	OCCDP
Watershed-Quaternary Aquatic Resources Area pH, Calcium, and Total Phosphorous Provincial Digital Elevation Model Bathymetry Line Ontario Hydro Network-Waterbodies Fishing Access Point	LIO
Cartographic Boundary Files Road Network Files	StatCan

3.3 <u>Data processing</u>

3.3.1 Invaded and Uninvaded Lakes' Processing

3.3.1.1 <u>SPECIES_FO</u>

This variable was adapted (using ArcGIS's "Spatial join" tool) from the spiny water flea sighting dataset. As mentioned in Section 3.2.1, this variable indicates whether the sighting of an invasive species is "Positive" or "Negative". "Positive" refers to reported sightings of an invasive species and is confirmed to be the right invasive species by a specialist. "Negative" refers to reported sightings that were monitored and was found to be not present, or a case of species misidentification/mistaken identity. It is important to note that one lake polygon can have multiple "Positive" or "Negative" points, and that some lakes can have a mix of both.

3.3.1.2 <u>Prs_Abs</u>

The name stands for Presence or Absence. This variable was created using *SPECIES_FO* data, and was processed through ArcGIS. It contains the numerical equivalent of *SPECIES_FO*; "Positive" = 1 or present, and "Negative" = 0 or absent. To convert the data, a new column was added onto the attribute table with the name "*Prs_Abs*", the *SPECIES_FO* column was highlighted, and then the "Field calculator" tool was used (on the new *"Prs_Abs"* column) to convert the values "Positive" and "Negative" into "1" and "0" respectively. For further data processing, the "Positive"/"1" and "Negative"/"0" sightings were separated or exported into separate layers;

44

"Positive"/"1" sightings are in one layer, and the "Negative"/"0" sightings are in its own separate layer. Since *Prs_Abs* is based on *SPECIES_FO*, there were lakes with multiple points of 1s, 0s, and a mix of both. So a threshold is needed to "filter" the sample. The following variable, *CNFMDSIGHT*, was created for this purpose.

3.3.1.3 CNFMDSIGHT

It stands for Confirmed sighting. This column comprises the number of confirmed positive and negative sightings of the invasive species and acted as a threshold to narrow down the invaded and uninvaded lakes sample size. It was derived from the spiny water flea distribution dataset, specifically from the *Prs_Abs* data column and the separate sightings layers (1/positive layer, or 0/negative). The spiny water flea sightings layers were in points but needed to be represented as lake polygons; to do so, a "Spatial Join" was performed to transfer the attribute table from the sightings dataset onto a waterbody dataset that has been narrowed down to just lakes. After performing the join, ArcGIS automatically creates a column (JOIN_COUNT) that counts how many points ("Positive" or "Negative") fall within a lake polygon, which was used as the basis of a threshold.

A threshold was applied so that only lake polygons with a certain number of "1"/"Positive" or "0"/"Negative" sightings were chosen to be a part of the final sample size. The threshold for invaded lakes were at least 2 "Positive" sightings of spiny water flea, and at least 10 "Negative" sightings to be considered uninvaded lakes. To clarify, only lakes with at least 2 "Positive" sightings were considered invaded and were added into the sample size, while only lakes with at least 10 "Negative" sightings were

45

considered uninvaded. The following were used to fill out the "Spatial Join" interface: Target Features: lake polygon dataset; Join Features: invasive species distribution dataset; Join Operation: JOIN_ONE_TO_ONE; uncheck the Keep All Target Features option; Match Option: COMPLETELY_CONTAINED.

As mentioned, some lakes were classified as both positive or negative, or some lakes contained a mix of both. This became apparent once the 1/"Positive" layer and the "0/Negative" layer were joined. To resolve the latter situation, the "Editor" tool was used. While the tool was active, the lake names were sorted alphabetically. This resulted in the lakes that passed the threshold for "Positive" and "Negative" sightings to be listed twice/to be replicated. One of the replicates were deleted manually; the decision on which replicate to delete was based on whether the "JOIN_COUNT" value of a "Positive" or "Negative" sighting is higher/lower than the other. For example, if one of the replicates shows 9 "Positive" sightings and the other shows "11" Negative sightings, then the replicate with "9" Positive sightings gets deleted and the lake is considered uninvaded. Although solved, a lake having both positive and negative sightings is a limitation of the study. All in all, the final sample size of invaded and uninvaded lakes is 190; 76 invaded lakes and 114 uninvaded lakes. Figure 8 illustrates the lakes that were included in the sample size after applying threshold.
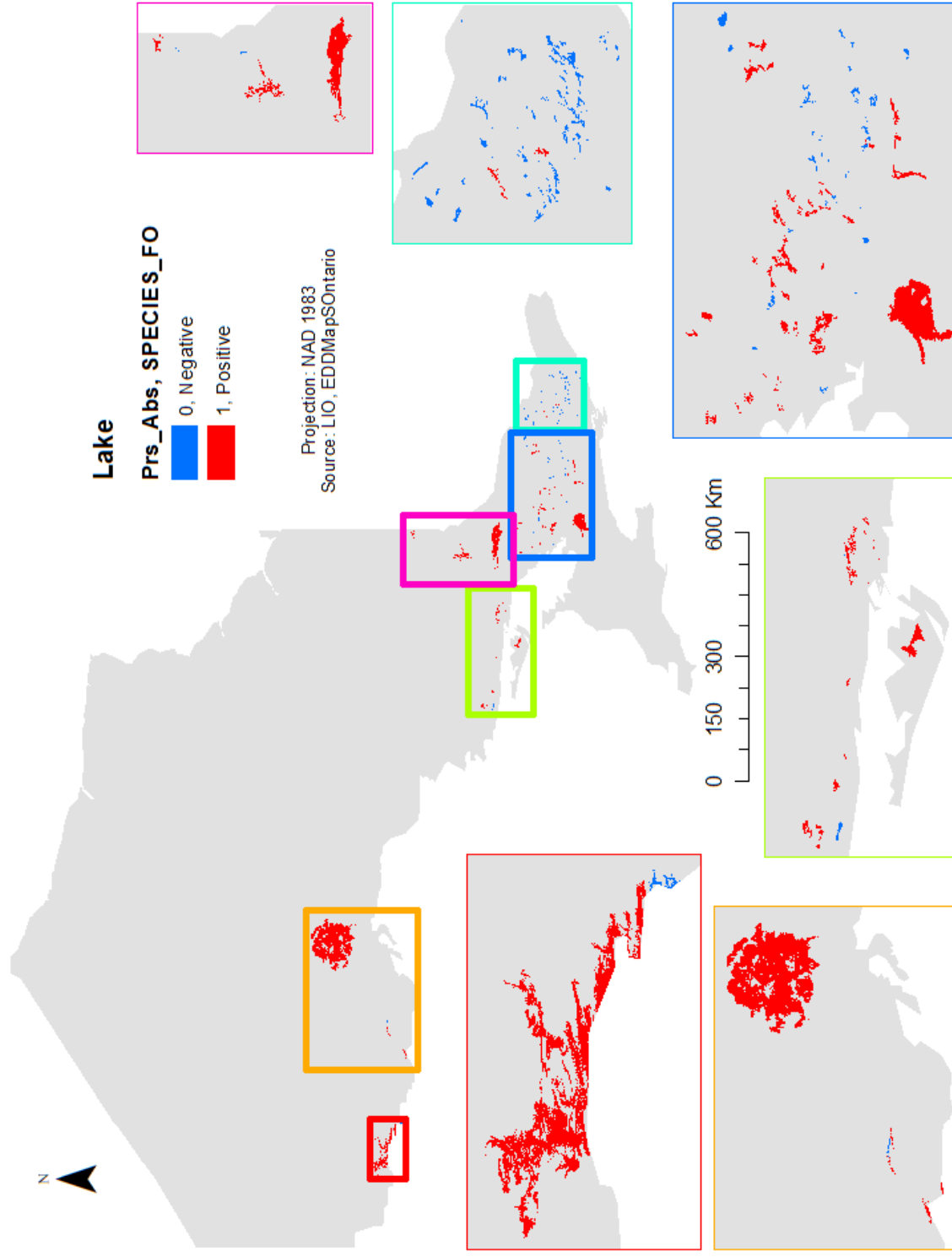
Lake

Prs_Abs, SPECIES_FO
- 0, Negative
- 1, Positive

Projection: NAD 1983
Source: LIO, EDDMapSOntario

0 150 300 600 Km

Figure 8. Invaded (1, Positive) and uninvaded (0, Negative) lakes.

47

3.3.2 Landscape Characteristics' Processing

3.3.2.1 <u>SDI, SEI, and %Landcover (FORST, WTRBODS, WETLND, GRASLND,</u>

<u>SHRBLND, BARELND, ARTFSURF, and CLTVLND)</u>

Patch Analyst was used to compute the landscape metrics (*%Landcover*) and

diversity metrics (*SDI or Shannon's Diversity Index, SEI or Shannon's Evenness Index*).

However, patch analyst only processes rasters with landscape classification, and the

unprocessed rasters from GlobeLand30 did not have any classification. The first task

was then to convert rasters into polygons, followed by classifying the landscape, and

then reverting the classified polygons to a raster that passed patch analyst's

requirement.

To start, the unprocessed rasters (from GlobeLand30) were mosaicked into one

polygon (refer to Figure 9a). Mosaicking all 12 rasters into one, and then converting that

one final raster into a polygon resulted into an error; the final raster was too big (over

four gb) to be converted into a polygon (ArcGIS has a 2.52 gb limit per conversion). To

bypass this error, the enormous raster was split into seven, an attribute table was added

to each, and were individually converted into a polygon. During conversion of raster to

polygon, great care was taken to make sure that the rasters were not simplified so that

its 30 m resolution gets preserved. The seven resulting polygons were then merged into

one final polygon (refer to Figure 9b).

To classify each unique cell value in the final polygon, the land cover

classification was assembled using a metadata provided by GLC30. The unique value in

the polygon were 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, and 255; each corresponding to Cultivated Land (*CLTVLND*), Forest (*FORST*), Grassland (*GRASLND*), Shrubland (*SHRBLND*), Wetland (*WETLND*), Waterbodies (*WTRBODS*), Tundra (not applicable to Ontario), Artificial Surfaces (*ARTFSRF*), Bareland (*BARELND*), Permanent Ice (not applicable to Ontario), and Ocean (not applicable to Ontario) respectively. The resulting land cover classification .csv file was then joined to the final polygon.

To revert the classified polygon to a raster, the Polygon to Raster tool was used. The following were typed onto the interface: Input Features: the desired polygon to be converted; Value field: the column that contained the landscape classification; Cell assignment type: unchanged; Cellsize: the same cellsize of the unprocessed raster, to preserve the resolution; and Priority field: unchanged. Converting the polygon, which was the province of Ontario and its landscape classification, resulted into an error because it was too big. To work around this, the polygon was first split based on the Quaternary-Watersheds of Ontario, and then individually converted to rasters. The watershed rasters were used for Patch Analyst's computation of the landscape and diversity metrics, and results for all metrics were tabulated into one excel file for ease with joining to other variables.
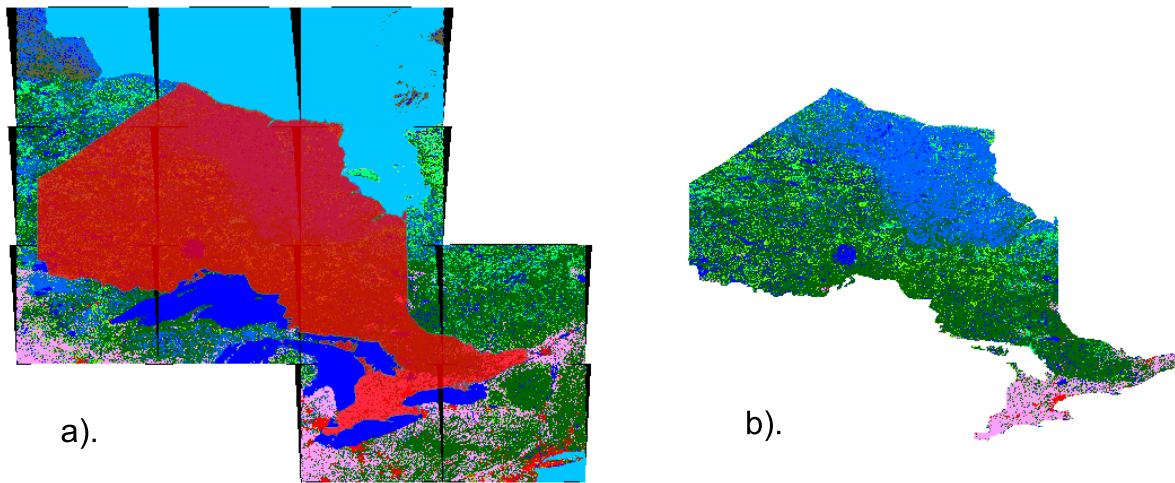
Figure 9. 30 m resolution land cover raster of Ontario. Quaternary watersheds of Ontario. (a) Pre-processing; (b) post-processing. Source: GlobeLand30; LIO.
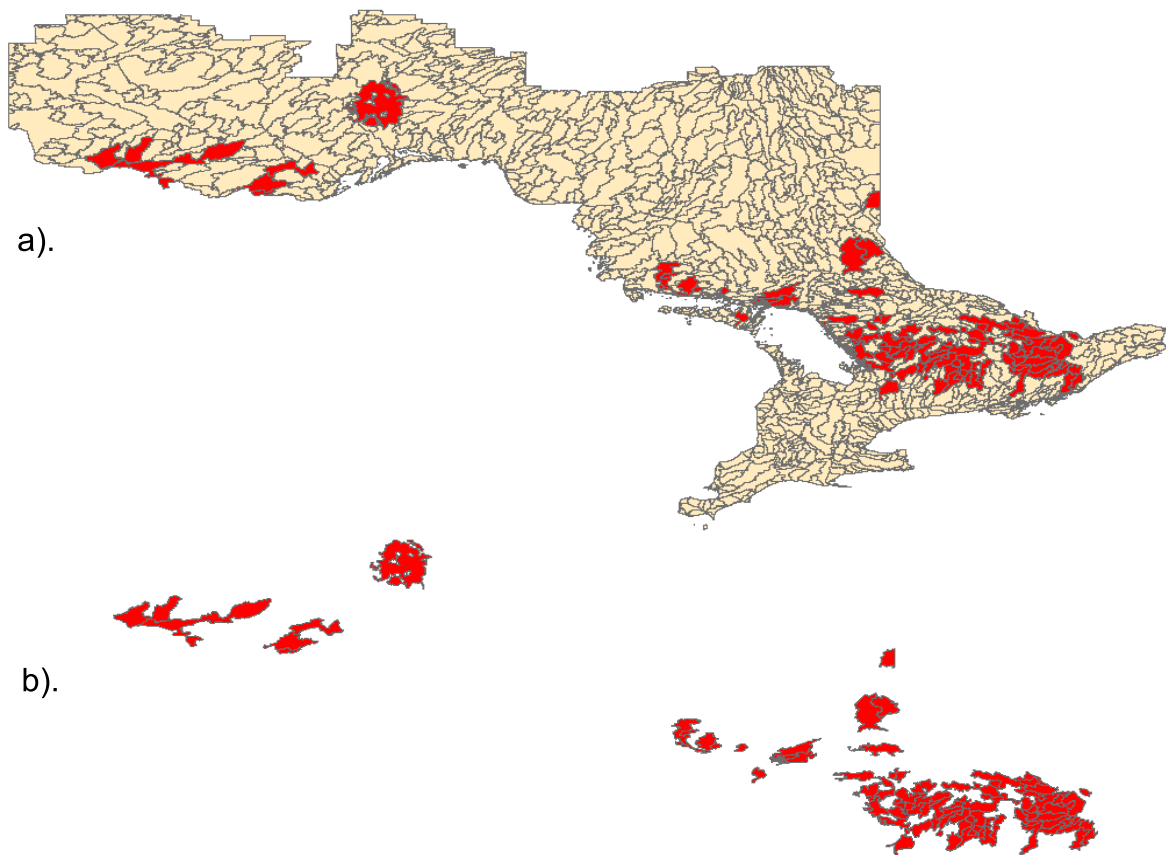

Figure 10. Quaternary watersheds of Ontario. (a) All watersheds pre-processing; (b) 84 watersheds that encompass the invaded and uninvaded lakes post-processing. Source: LIO.

3.3.2.2 <u>IDENT</u>

This column refers to a watershed's unique identifier and was derived from the Watershed-Quaternary dataset. The watershed was selected depending on the invaded and uninvaded lakes. A watershed was deemed important if an invaded or uninvaded lake is located within it, and it was then selected. In total, 84 watersheds were found to contain all 190 invaded and uninvaded lakes (refer to Figure 10b).

3.3.2.3 <u>MNPCPWS* and MNTMPWS*</u>

MNPCPWS stands for mean precipitation for watersheds, and *MNTMPWS* stands for mean temperature for watersheds (* stands for the season, refer to Table 1). Ultimately, "Zonal Statistics by Table" was used to calculate the mean precipitation and temperature of each watershed, but there several actions needed to be done before doing so. The unprocessed form of these datasets (in points, refer to Figure 11a) presented an issue as the points did not cover entire watersheds, and some watersheds were completely out of range from any single point.

To solve this issue, the "Buffer" tool was used to extend the range of the points by 15 km (refer to Figure 11b). A beneficial side effect of the solution was that the points got converted to polygons, the polygons retained the attributes of the points. The next step was to convert the new polygons to rasters using the "Convert to Raster" tool. Care was taken in making sure that the size/extent of the new polygons held in raster form. And finally, "Zonal Statistics as Table" tool was used to obtain the desired variables.

This step was repeated four times (for Winter, Summer, Spring, and Fall) for both temperature and precipitation dataset.

3.3.2.4 LKELV_MN, and MNELVWTRSH

These columns were derived from the Provincial Digital Elevation Model (PDEM) dataset. *LKELV_MN* stands for mean lake elevation (m), while *MNELVWTRSH* stands for mean watershed elevation (m); mean lake elevation should be sectioned within "3.3.3 Lake Characteristics' Processing", but since these two variables were derived similarly it was decided that they should be grouped together for ease. PDEM was in raster format, to derive elevation values for just the watersheds and invaded/uninvaded lakes polygon the "Zonal Statistics as Table" tool was used. To fill out the interface: Input raster or feature zone data: watersheds polygon; Zone field: IDENT (a unique identifier); Input value raster: PDEM; check the Ignore NoData in calculations box; Statistics type: All.

a).

b).

Figure 11. Mean precipitation (Spring) in Ontario, 1986-2005. (a) Spring precipitation points, pre-processing; (b) Spring precipitation polygon with a 15 km buffer. Source: OCCDP, post-processing.
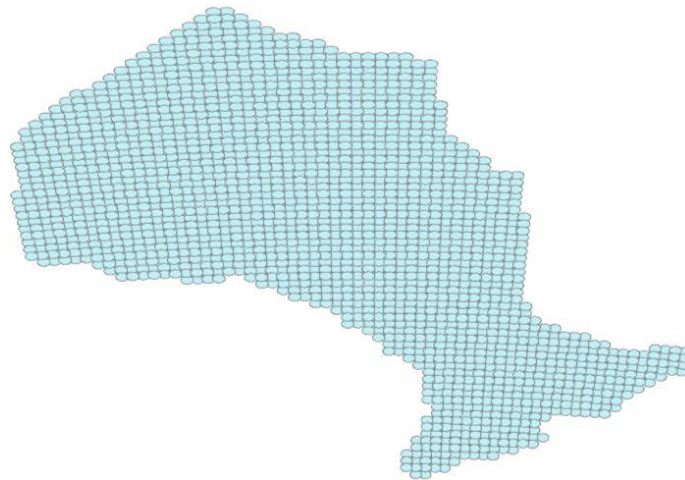
3.3.3 Lake Characteristics' Processing

3.3.3.1 OFF_NAME

It stands for official lake name. This variable lists the official name of the waterbody/lake. It was taken from the waterbody dataset and Aquatic Area Resources

(ARA) dataset. The two datasets were spatially joined on ArcGIS. Having taken the lake names from two datasets guaranteed the names were correct. Lake names were taken from both sources neither dataset had a complete list of every lake name, so it was thought that one would supply the missing name in the other. For the lake names that did match, it served to check if the name is correct or not.

3.3.3.2 pH, CALC, and TPHOS

CALC stands for calcium, and *TPHOS* stands for total phosphorous. Since the unprocessed dataset for these variables were in excel format, the first step was to save it as a .csv file to be usable in ArcGIS. The unprocessed dataset contained X and Y coordinates, which made it easier to project the dataset in point form. A "Spatial Join" was performed with the invaded and uninvaded lakes. The following were typed to fill out the tool interface: Target: the invaded and uninvaded lakes file; Join: the lake chemicals file; Operation: ONE_TO_ONE; "Keep all target" box was checked; Match Option: Intersect. Through examining the resulting file, it was found that only 68 out of 190 lakes had values for total phosphorous, 67 out of 190 lakes had values for pH, and 64 out of 190 lakes had values for calcium.

3.3.3.3 SYS_AREA, SYS_PERIM, and PERIM_AREA/AREA_PERIM

SYS_AREA lists the area (m^2) of the lakes and *SYS_PERIM* lists the perimeter (m), both were taken from the Aquatic Resources Area (ARA) dataset. *PERIM_AREA* and *AREA_PERIM* are ratios of perimeter and area, and vice versa. The "Field Calculator" tool was used to calculate the ratio.

3.3.3.4 <u>DIST2METRO, DIST2POPCN, DIST2DSPLC, DS2NRSHWYX, and</u>

<u>NRSTINVLK</u>

DIST2METRO refers to distance (km) to the nearest metropolitan area, which is an area with a population of at least 50,000. *DIST2POPCN* refers to distance (km) to the nearest population centre, which is an area with at least 1,000 people and a population density of at least 400 people/km^2. *DIST2DSPLC* refers to distance (km) to the nearest designated place which is a small community that does not have enough population to be considered a population centre. The three aforementioned data were derived using the 2016 Census Boundary File and the invaded and uninvaded lake polygons of each species. *DS2NRSHWYX* stands for distance (km) to nearest highway exit, which was derived using the 2016 Road Network File and the invaded and uninvaded lake polygons. *NRSTINVLK* stands for nearest invaded lake (km), which was derived using the invaded and uninvaded lake polygons. The "Measure" tool was used to determine the direct distance, or Euclidean distance, from point a to point b. To get an accurate measurement, centroids of each of the dataset were determined and used (i.e. centroid of a metropolitan area, point a, and centroid of an invaded/uninvaded lake, point b), except for *DS2NRSHWYX*. To measure *DS2NRSHWYX* values, the direct distance of the invaded and uninvaded lake centroid to the nearest highway exit was measured. The nearest highway exit was determined by overlaying highways and regular roads, both lines having different symbols to identify where highway exits are. A possible limitation to these specific variables would be the usage of direct distance, or a linear measurement of the distance between point a and b, because it does not take into account the meandering of roads.

3.3.3.5 MAX_DEPTH/MEAN_DEPTH

These variables were derived from the datasets Bathymetry line and Aquatic Resources Area, both had specific columns for mean and max depth. A "Spatial Join" was performed with the intent of adapting and combining the depth values in both datasets into one max and mean column, with the invaded and uninvaded lakes file as the target feature. The tool interface was not changed from the default aside from specifying the Target Features, Join Features, Output Feature Class, and the Field Map of Join Features. The resulting file showed that only 150 lakes out of 190 had values for mean and max depth.

3.3.3.6 BOATACC

This variable was derived from Fishing Access Point. The unprocessed dataset was overlayed with the invaded and uninvaded lakes. While using the "Editor" toolbar the lakes were individually checked and assigned a value. Lakes with shoreline access point were assigned the number "1", lakes with an enhanced access point were assigned the number "2", lakes with a boat launch the number "3" for lakes with a boat launch, and number "4" for lakes with different access types. The resulting file showed that only 80 out of 190 invaded and uninvaded lakes had records of access points.

3.4 <u>Joining Lake and Landscape Characteristics</u>

In prior sections, lake characteristics and landscape characteristics were processed separately. However, the desired outcome for building the database is to have every data and variable joined in one file. To do so, each variable was first joined to one representative file per main category (lake and landscape characteristics). For lake characteristics, lake variables (detailed in Section 3.3.3) were spatially joined to the invaded and uninvaded lakes polygon (detailed in Section 3.3.1). For landscape characteristics, landscape variables (detailed in Section 3.3.2) were spatially joined to watersheds polygon, specifically to the landscape variable *IDENT* (detailed in Section 3.3.2).

Joining the two representative files, lake polygon joined with every lake variable and watershed polygon joined with every landscape variable, was not as simple as using the spatial join tool. There are 190 lakes of interest, all of which are within 84 watersheds. The ideal outcome was for the final joined file to retain the number of lakes and lake variables, while each lake showing attributes of the watershed (along with all the landscape variables). Unfortunately, performing a spatial join resulted in the reduction of the number of lakes or the loss of variables. To work around this issue, several joining tools were tried and the tool that produced the result closest to the ideal outcome is the "Identity" tool (Analysis>Overlay>Identity). On the tool interface the lake polygon was assigned as the input feature, and the identity feature was the watershed polygon. Everything else were left unchanged.

The resulting layer exhibited all the desired lake and landscape variables while having polygons of all 190 lakes, but further inspection of the attribute table showed that the FID exceeded 190. In other words, there were 190 unique lake polygons being projected on ArcGIS, but there were more than 190 lake FIDs (or unique identification) shown in the attribute table. A closer look at the attribute table showed that, based on the unique lake names, some lakes were duplicated. The duplication was a result of lakes falling within two watersheds; bigger lakes and their basin cross watershed boundaries. ArcGIS partitioned the lakes that were situated on top of borders and it classified each part as a unique lake, thereby causing the duplication. To solve this issue, the same process done in Section 3.3.1 *CNFMDSIGHT* was performed; the "Editor" tool was used to alphabetically arrange the lakes, then one of the duplicates was deleted manually. The decision as to which one would be deleted was based the size of duplicated lakes; the bigger duplicate was kept, and the smaller duplicate was deleted. Fortunately, there never more than two duplicates.

3.5 <u>Statistical Analysis</u>

Analysis of the variables were done using the programs SPSS (Statistical Package for the Social Sciences by IBM), SAM (Spatial Analysis for Macroecology), and ArcGIS. SPSS was used to visually check the variables for Gaussian distribution by creating histograms (for variables pre- and post-transformation), as well as for transforming the variables with ratio data type. SAM was used for transforming the variables with count data types. It was also used for calculating bivariate logistic regression. ArcGIS was used to compute the variables' spatial autocorrelation.

Table 2. Variables derived from the processed data.

Category	Variable	Full name
Invaded and uninvaded lakes (dependent variable)	CNFMDSIGHT *	Confirmed sight
Lake characteristics (independent variables)	SYS_AREA **	Lake area
	SYS_PERIM **	Lake perimeter
	PERIM_AREA ***	Perimeter/area ratio
	DIST2METRO **	Distance to nearest metropolitan area
	DIST2POPCN **	Distance to nearest population centre
	DIST2DSPLC **	Distance to nearest designated place
	NRSTINVLK **	Distance to nearest invaded lake
	DS2NRSTHWY **	Distance to nearest highway exit
	LKELV_MN **	Mean lake elevation
Landscape characteristics (independent variables)	MNELVWTSHD **	Mean watershed elevation
	MNTMPWSF **	Mean watershed temperature, Fall
	MNTMPWSSP **	Mean watershed temperature, Spring
	MNTMPWSSU **	Mean watershed temperature, Summer
	MNTMPWSW **	Mean watershed temperature, Winter
	MNPCPWSF **	Mean watershed precipitation, Fall
	MNPCPWSSP **	Mean watershed precipitation, Spring
	MNPCPWSSU **	Mean watershed precipitation, Summer
	MNPCPWSW **	Mean watershed precipitation, Winter
	SDI ***	Shannon's Diversity Index
	SEI ***	Shannon's Evenness Index
	FOREST ***	% Forest cover
	WATERBODS ***	% Waterbody cover
	WETLAND ***	% Wetland cover
	GRASSLAND ***	% Grassland cover
	SHRUBLAND ***	% Shrubland cover
	BARELAND ***	% Bareland cover
	ARTSURFCS ***	% Artificial surface cover
	CULTVLAND ***	% Cultivated land cover

3.5.1 Histograms and Normalization

Histograms are done to visualize the distribution of data. What is assessed during this process is whether the data follows Gaussian distribution/normal distribution or if the distribution of the data falls within a bell-shaped curve (Kandane-Rathnayake, Enticott, & Phillips, 2013; Lyon, 2014). SPSS was used to create histograms for the variables. After creating histrograms, the variables were then subjected to normalization or transformation. The ratio data type variables were transformed using SPSS's arcsine transformation formula; SPSS>Transformation>Compute Variable>Arithmetic>Arsin. The count data type variables were transformed using SAM's logarithmic transformation formula; SAM>Data>Data handling>Transformations>Log10(x+1). Histograms were also created for the transformed variables.

3.5.2 Spatial Autocorrelation

Spatial autocorrelation analyzes the connection between a variable found in a spatial unit, and the same variable but found within a neighbouring spatial unit (Getis, 2007, 2008; Legendre, 1993; Valcu & Kempenaers, 2010). In more technical terms, as quoted from Getis (2008), "spatial autocorrelation shows the correlation within variables across georeferenced space". Legendre (1993) loosely defined it as "the property of random variables taking values, at pairs of locations a certain distance apart, that are more similar (positive autocorrelation) or less similar (negative autocorrelation) than expected for randomly associated pairs of observations". There are multiple methods to determine spatial autocorrelation, as detailed in studies by Getis (2007) and Legendre (1993), but Global Moran's I was used for this research.

60

Global Moran's I (index) analyzes spatial autocorrelation by creating a parameter that ranges from -1 to 1. A positive index represents clustered variables; high values are neighbouring other high values, or low values are neighbouring other low values. A negative index represents variables that repel each other (imagine a checkerboard and its pattern); high values are far away from other high values, or low values are far away from other low values. A zero represents random spatial distribution of the variables (Environmental Research Systems Institute (ESRI), 2019; Fu, Zhao, Zhang, & Tunney, 2011) . For the purposes of this paper, the GIS software ArcMAP 10.6.1 was used to calculate Global Moran's I, specifically the tool Spatial Autocorrelation (Morans I) (Spatial Statistics Tools>Analyzing Patterns>Spatial Autocorrelation). The tool interface was kept unchanged except for specifying the distance band or threshold distance. To measure the appropriate distance band so that every polygon would have at least one neighbouring sample, the tool Calculate Distance Band from Neighbor Count was used (Spatial Statistics Tools>Utilities>Calculate Distance Band from Neighbor Count). The result showed that a distance band of three meters would guarantee that every sample would have at least one neighbour.

3.5.3 Logistic Regression

Logistic regression is a statistical method that can predict the outcome of an event by scrutinizing the correlation between two variables. In this case, it was performed to determine if there is a correlation between the presence or absence of *B. longimanus* (dependent variable) and lake or landscape characteristics (independent variables) in 190 lakes across several watersheds. It is a good statistical tool to use

when dealing with a bivariate/dichotomous/categorical dependent variable; or if the dependent variable could be interpreted as "yes" or "no" or "1" or "0" (Menard, 2013; Osborne, 2017; Rupert, Cannon, Gartner, Michael, & Helsel, 2008). It has been used in several branches of science; Menard (2013) cited its use in several "hard" and "soft" sciences. Logistic regression produces the parameter *McFadden's Rho-Square (ρ^2)* with a value ranging from 0 to 1; wherein values closer to 1 corresponds to more significant results (Rupert et al., 2008; Shaker & Rapp, 2013). This is similar to linear regression's R^2 in that its value has a small range. However, smaller values are not indicative of poor results; *McFadden's Rho-Square* with values between 0.2-0.4 are considered excellent results (McFadden, 1977). SAM (Spatial Analysis for Macroecology, version 4.0) was used to compute logistic regression (Rangel et al., 2010).

3.0 RESULTS

Figure 12. Example histograms for the variables listed on Table 1. (a) SYS_AREA (pre-transformation) and LGAREA (post-transformation) histograms; (b) LKELV_MN (pre-transformation) and LGMNLKELV (post-transformation) histograms.

The pre-transformation and post-transformation histograms generated by SPSS showed various results. Histograms, such as those in Figure 12a and b, showed a conclusive difference between the distribution of data. Some histograms were less

conclusive, such as the histograms of the variable artificial surfaces in Figure 13. Table 3 provides a list of the variables after transformation, as well as the variables that were chosen for statistical analysis.

Figure 13. Histogram of the variable ARTSURFCS pre- (left) and post-transformation (right).

Table 4. Variables pre- and post-transformation.

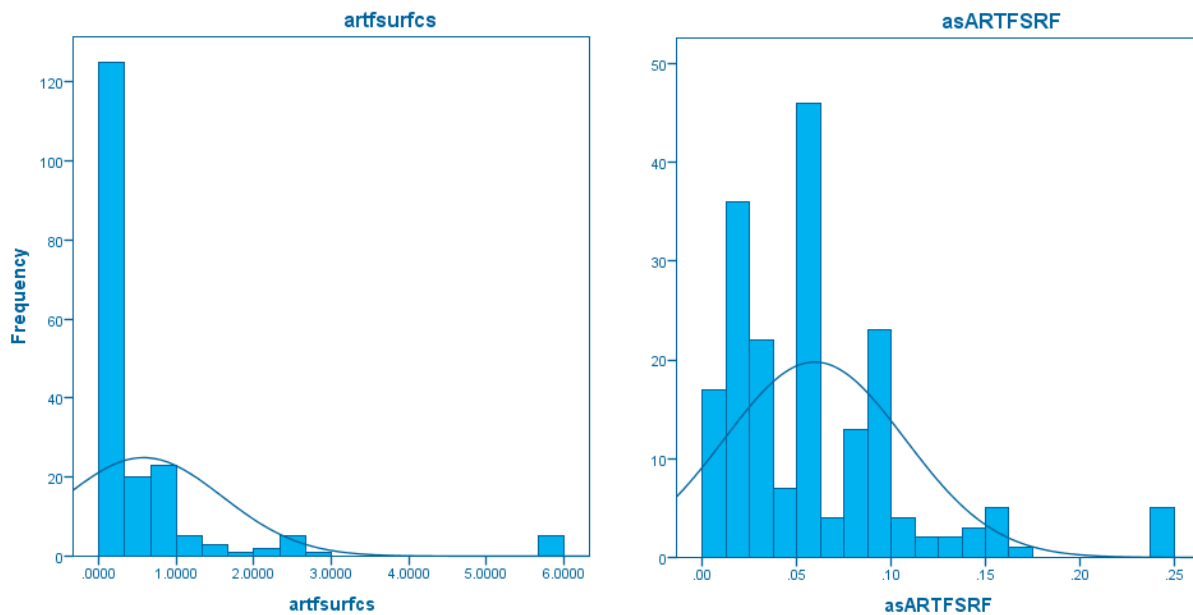| Category	Variable	Transformed
Invaded and uninvaded lakes (dependent variable)	CNFMDSIGHT	*n/a*
Lake characteristics (independent variables)	SYS_AREA	LGAREA*
	SYS_PERIM	LGPERIM*
	PERIM_AREA	ASPERIAREA*
	DIST2METRO	LGDSMETRO*
	DIST2POPCN	LGDSPOPCN*
	DIST2DSPLC	LGDSDGPLC*
	NRSTINVLK	LGNRSINVLK*
	DS2NRSTHWY	LGNRSHWY*
	LKELV_MN*	LGMNLKELV
Landscape characteristics (independent variables)	MNELVWTSHD*	LGMNWSELV
	MNTMPWSF*	LGMNTMPF
	MNTMPWSSP	LGMNTMPSP*
	MNTMPWSSU*	LGMNTMPSU
	MNTMPWSW*	LGMNTMPW
	MNPCPWSF*	LGMNPCPF
	MNPCPWSSP*	LGMNPCPSP
	MNPCPWSSU*	LGMNPCPSU
	MNPCPWSW*	LGMNPCPW
	SDI*	ASSDI
	SEI	ASSEI*
	FOREST	ASFORST*
	WATERBODS	ASWTRBODS*
	WETLAND	ASWETLND*
	GRASSLAND	ASGRASLND*
	SHRUBLAND	ASSHRBLND*
	BARELAND	ASBARELND*
	ARTSURFCS	ASARTFSRF*
	CULTVLAND	ASCLTVLND*

* variables that were used for spatial analysis and logistic regression.

Most of the Global Moran's I value, illustrated in Table 4, are positive. The variable ASARTSRF, however, returned a negative value of -0.00043. The corresponding z-score for ASARTSRF is also on the lower end of the spectrum at 1.11. Whereas the corresponding z-scores for the rest of the variables range between 2.62 for ASPERIAREA – 72.7 for MNPVPWSSP. The resulting p-value for spatial autocorrelation of the variables are either >0.05, or 0.268. For logistic regression, McFadden's Rho-Squared resulted in varying numbers. The lowest being 0 for ASARTSRF, and the highest being 0.303 for MNTMPWSSU. The corresponding p-values range from <0.0001 – 0.995. The True Skills Statistic either produced results or resulted in errors which are represented by asterisks on Table 4.

Table 6. Spatial autocorrelation and logistic regression results.

Variable	Spatial Autocorrelation			Logistic Regression		
	Global Moran's I	z-score	p-value	McFadden's Rho-Squared	p-value	TSS
LGAREA	0.057	14.2	*	0.142	<0.0001	0.158
LGPERIM	0.053	13.3	*	0.119	<0.0001	0.132
LGDSMETRO	0.100	23.9	*	0.109	<0.0001	0.044
LGDSPOPCN	0.019	5.46	*	0.001	0.561	*
LGDSDGPLC	0.048	12.2	*	0.016	0.046	*
LGNRSINVLK	0.132	31.2	*	0.140	<0.0001	0.123
LGNRSHWY	0.026	7.01	*	0.003	0.408	*
LKELV_MN	0.162	38.1	*	0.123	<0.0001	0.035
MNELVWTSHD	0.175	41.1	*	0.174	<0.0001	0.118
MNTMPWSF	0.183	42.9	*	0.193	<0.0001	0.136
LGMNTMPSP	0.208	48.6	*	0.194	<0.0001	0.162
MNTMPWSSU	0.299	69.2	*	0.303	<0.0001	0.351
MNTMPWSW	0.098	23.7	*	0.068	<0.0001	0.066
MNPCPWSF	0.145	34.2	*	0.029	0.007	*
MNPCPWSSP	0.314	72.7	*	0.223	<0.0001	0.333
MNPCPWSSU	0.263	60.9	*	0.211	<0.0001	0.263
MNPCPWSW	0.145	34.2	*	0.029	0.007	*
ASPERIAREA	0.006	2.62	*	0.029	0.006	**
SDI	0.083	20.1	*	0.037	0.002	*
ASSEI	0.971	23.3	*	0.030	0.006	*
ASFORST	0.035	9.10	*	0.005	0.249	*
ASWTRBODS	0.010	3.43	*	0.029	0.007	0.039
ASWETLND	0.134	31.7	*	0.043	0.001	*
ASGRASLND	0.052	12.9	*	0.012	0.076	*
ASSHRBLND	0.060	14.7	*	0.040	0.001	0.004
ASBARELND	0.110	26.2	*	0.039	0.002	0.092
ASARTFSRF	-0.0004	1.11	0.268	0.000	0.995	**
ASCLTVLND	0.162	38.0	*	0.155	<0.0001	*

Spatial autocorrelation, p-value * = <0.05.
Logistic regression, True Skills Statistics (TSS) * = error due to constant values among results;
** = cannot process fitting.

4.0 DISCUSSION

5.1 <u>Building the Database</u>

Since building the database was a big part of this project, the decisions that were made while going through the process will discussed. Starting with the invaded and uninvaded lakes. As mentioned in Section 3.2.1, the raw data did not specify if a lake is invaded or uninvaded only that if there were correct sightings of the species in the lake or if the reported sighting is a mistaken identity. Translating this information to invaded/uninvaded or presence/absence of spiny water flea is detailed in Section 3.3.1.1 and 3.3.1.2. Section 3.3.1.3, CONFMDSIGHT, discusses in detail the application of a threshold to filter and improve the sample size of invaded and uninvaded lakes. In that section, uninvaded lakes are held at a higher standard with the threshold being ≥10 negative sightings to be considered uninvaded. The rationale is that having a greater amount of negative or mistaken sightings would strengthen the uninvaded status of a lake. In the case of invaded lakes and the threshold of ≥ 2 positive sightings, having less than the threshold could mean that spiny water flea is in the process of establishing its population in the lake.

In regard to the precipitation and mean temperature data, detailed in Section 3.3.2.3, the rationale behind creating a buffer is that climatic factors are not point specific. Temperature and precipitation are measured over a big scale, a neighbourhood for example, and is reported at the same scale. However, it is acknowledged that testing four seasons of precipitation and temperature data against invaded and uninvaded lakes is questionable. Specifically, because the latter variable

will not change along with the former. Total population of spiny water flea changes along with the season (Berg & Garton, 1988; Garton et al., 1990; Yan & Pawson, 1998), but a lake would still be invaded or uninvaded regardless of the season.

There were potential variables that were not included for the statistical analysis portion (refer to Table 2 and 3 to review the variables); (1) pH, CALC, and TPHOS, (2) MAX_DEPTH/MEAN_DEPTH, and (3) BOATACC. The three variables were excluded because they were not as extensive as the other rest of the variables. Out of 190 invaded and uninvaded lakes, only 68 had recorded values of total phosphorous, only 64 had recorded values for pH, and only 64 had recorded values for calcium. Only 150 lakes had recorded values of maximum and mean depth, and 80 lakes were recorded to have boat access.

Previous studies have found these variables to be indicative and predictive of invasion. Lake chemistry can determine the favourability of a lake. Phosphorous levels, for example, in inland lakes is directly linked to its trophic state. Eutrophic lakes, or lakes with high levels of phosphorous among other nutrients, tend to have higher levels of primary producers and consumers which means more resources for spiny water flea. Especially since spiny water flea is a voracious zooplanktivore (Berg & Garton, 1988; Bur et al., 1986; Lange & Cap, 1986). Spiny water flea establishment is positively correlated to deeper lakes (L. Wang & Jackson, 2011), and it is also known to capitalize on a lake's depth by going deeper into the water column during the day as a reprieve from the warmer surface water temperature (Berg & Garton, 1988; Garton et al., 1990). The presence of boat ramp or boat launch increases a lake's desirability for human

69

recreational activity, and is directly linked to the spread of spiny water flea and other

aquatic invasive species (Gertzen & Leung, 2011; MacIsaac et al., 2004; Muirhead &

MacIsaac, 2005; Shaker et al., 2013)

5.2 <u>Statistical Analysis</u>

The histograms in Figure 12a, 12b, and 13 is a good representation of every

histogram generated by the 28 independent variables. In cases as observed in Figure

12a where there is a clear contrast, the version of the variable that produced the better

histogram was used for statistical analysis. For Figure 12a, the transformed variable,

LGAREA, was chosen. In cases that are similar to what Figure 12b illustrates, where

both version's histogram is visually represent Gaussian distribution, the version with a

bell curve that's closer to the centre was chosen. For Figure 12b, the pre-transformed

variable, LKELV_MN, was chosen. In cases similar to what is represented by Figure 13,

the version that produced a curve that is closer to a bell shape was chosen. For Figure

13, the transformed variable, ASARTFSRF was chosen.

Spatial autocorrelation, as mentioned in Section 3.5.2, investigates the

correlation between nearby objects. In this case the nearby objects would be a lake and

its characteristic or its watershed's characteristics. Most of the variables resulted in

positive Moran's I, therefore they exhibit positive spatial autocorrelation and that most of

the lake and landscape characteristics are spatially clustered. The corresponding z-

score values also support the aforementioned, that the observed spatial patterns are

correlated and that they are not caused by random events. This is reasonable as the

glaciers that receded to the poles, which carved out the land and created basins that

are the foundation of lakes, are humongous enough to create a uniform effect on a massive mass of land such as Ontario. The p-values for this analysis are mostly highly significant (p-value>0.01), which strengthens the corresponding z-scores by suggesting that there is a 99% probability that the spatial patterns observed are not born from random events.

Further examination of the spatial autocorrelation parameters in Table 4 shows that ASARTFSRF (artificial surfaces) had the opposite result. Artificial surfaces, or man-made surfaces, show negative spatial autocorrelation (-0.00043). It means that there is a pattern to how artificial surfaces are laid out, and that they are somewhat far apart from each other. As can be observed in Figure 14, most artificial surfaces are concentrated in the south eastern watersheds (light green extent frame) and are far apart from each other. A possible explanation for the pattern could be that humans historically developing close to resources, as well developing latitudinally lower to avoid colder temperatures. ASARTFSRF's z-score, however, shows that it is roughly one standard deviation away from the mean, and that the pattern observed is caused by random events. Its p-value, 0.268, is not significant (p-value<0.05) and reinforces its z-scores interpretation that the observed spatial pattern is likely caused by random events.
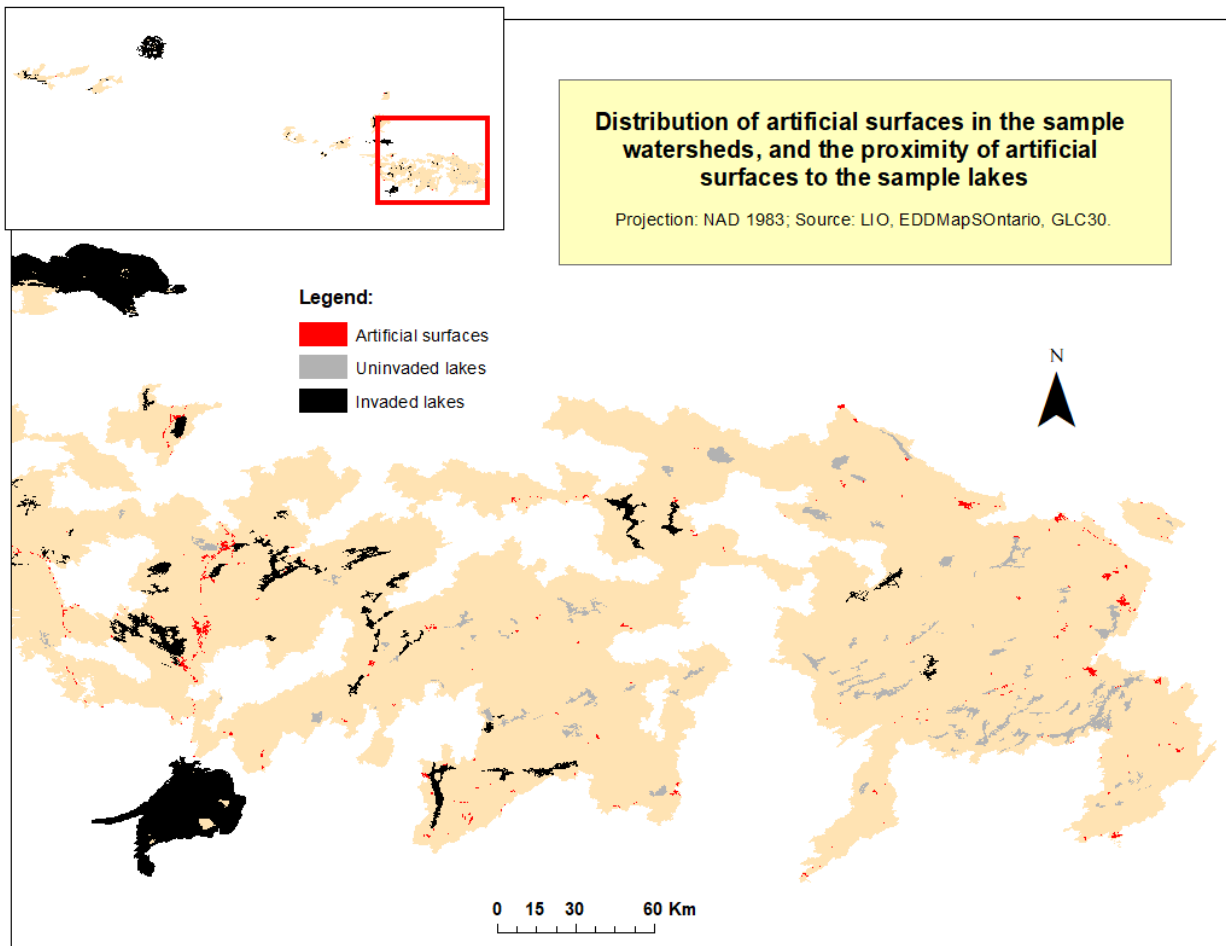
Figure 14. Distribution of artificial surfaces in the sample watersheds. This figure depicts the distribution of artificial surfaces across the 84 watersheds, as well as artificial surfaces' proximity to invaded and uninvaded lakes.

In regards to logistic regression, the results presented in Table 4 examines the correlation between the presence or absence of spiny water flea with lake and landscape characteristics. As previously mentioned in Section 3.5.3, logistic regression is a tool that can predict the occurrence of events. It creates a parameter, McFadden's ρ^2, that has a value ranging between 0 to 1; 1 being the best possible fit, but values

between 0.2 to 0.4 is still considered good (McFadden, 1977). Most variables have a ρ^2 >0, only 3 variables are within the 0.2 to 0.4 range, none attained a $\rho^2 > 0.4$, and 1 variable had a $\rho^2 = 0$. The variables that resulted to a good fitting ρ^2 are MNTMPWSSU (mean temperature for watersheds during summer), MNPCPWSSP (mean precipitation for watersheds during spring, $\rho^2 = 0.223$), and MNPCPWSSP (mean precipitation for watersheds during summer, $\rho^2 = 0.211$).

MNTMPWSSU had the highest ρ^2 with 0.303, suggesting that warmer temperatures brought about by summer is predictive of the presence of spiny water flea. This is supported by the variable's highly significant p-value suggesting that the pattern, warmer temperatures in a watershed can be predictive of invasion, is not a result of random chance. However, it is acknowledged that MNTMPWSSU is measured on a watershed scale. Even though climatic temperature affects lake temperature, it is not a direct measure of surface water temperature. This interpretation is in accord with results from previous studies that found that spiny water flea thrives in lakes with warmer surface water (between 8°C to 25°C) which are common in lakes during spring, summer, and fall (Berg & Garton, 1988; Garton et al., 1990; Sprules et al., 1990; Yan & Pawson, 1998). A few studies also mentioned that fall is when spiny water flea's population density reaches its peak (Berg & Garton, 1988; Garton et al., 1990; Sprules et al., 1990); however, fall watershed temperature (MNTMPWSF) resulted only in an adequate ρ^2 of 0.193, paired with a highly significant p-value.

MNPCPWSSP (mean precipitation for watershed during spring) and MNPCPWSSU (mean precipitation for watershed during summer) resulted in a ρ^2 value

73

of 0.223 and 0.211 respectively, and both variables have a highly significant p-value. Precipitation in a watershed could point towards the amount nutrients that flows into a lake along with the water it receives. This would provide a possible explanation for precipitation during spring, but not for summer. The former season is known to receive more nutrients from the winter thaw, while the latter is known to have lower precipitation. The nutrient level, or trophic state, of a lake affects its ability to support organisms living in it. Eutrophic lakes can support more prey for spiny water flea. This interpretation is both in accordance and conflicting with previous research; spiny water flea is known to thrive in oligotrophic lakes of Eurasia (Berg & Garton, 1988; Lange & Cap, 1986), but it has also managed to successfully invade Lake Erie (Berg & Garton, 1988; Bur et al., 1986; Garton et al., 1990) which is a eutrophic lake.

A surprising result is ASARTFSRF (artificial surfaces) which generated the lowest ρ^2 of 0, with a corresponding p-value that is >0.05. Past studies have shown that human development within 200 m of a lake is a good predictor of the presence of invasive species (Shaker & Rapp, 2013; Shaker et al., 2013). In this case however, a possible explanation for the conflicting results is that artificial surfaces were measured on a watershed scale and, as illustrated by Figure 14, most lakes do not have any artificial surface in its vicinity. This could be a reflection of the land cover data that was used. Other unexpected results are the poor fitting ρ^2 of LGDSMETRO, LGNRSHWY, SDI, and SEI as these variables were considered to be indicative of lake accessibility and human impact.

5.0 CONCLUSION

This study was successfully able to build a database that was used for examining the secondary spread of spiny water flea in 190 inland lakes across 84 watersheds in Ontario. This study can also be of aid to natural resource management by how it repurposed archived, inexpensive, and publicly available data. The following variables were produced for statistical analysis:

(1) presence or absence of spiny water flea/invaded or uninvaded lakes as the dependent variable;

(2) and 28 lake and landscape characteristics as the independent variables.

The variables were then analyzed using spatial autocorrelation and logistic regression. In terms of spatial autocorrelation, most variables were found to be close in proximity or are spatially clustered. The exception was artificial surfaces (ASARTFSRF) which was found to be far apart from each other when viewed on a watershed scale.

In regard to logistic regression, although most variables had inadequate results, three variables were found be good predictors if a lake is invaded or uninvaded by spiny water flea. The best predictor was the variable mean temperature for watersheds during summer (MNTMPWSSU). This finding is in accord with past research which showed a consensus with spiny water flea's preference for water temperatures between 8-25°C. The other two variables were mean precipitation for watersheds during spring (MNPCPWSSP) and during summer (MNPCPWSSU). It is speculated that precipitation in a watershed is tied to the amount of nutrients that a lake receives; ultimately, it could

affect a lake's favourability for aiding the establishment of invasive species. Other notable results were mean temperature for watersheds during fall (MNTMPWSF), and artificial surfaces (ASARTFSRF). The former variable was notable because it fell short of getting categorized as a good predictor, whereas past studies have discussed the spiny water flea's fondness of fall and of water temperatures during that season. The latter variable was notable not only because it resulted to a McFadden's $\rho^2 = 0$, but also because it was regarded by previous studies as a good predictor of invasion. In conclusion, although there were variables that were found to be good predictors, secondary invasion of spiny water flea appears to be more complex when scrutinized on a bigger scale rather than on a smaller scale. The following section elaborates further.

6.1 <u>Limitations of the study</u>

It has been expressed that some variables did not produce their expected result. It is recognized that the scale at which these variables were examined could have restricted their potential result. In previous studies, the inland lakes included in the sample size were within a watershed or within a national park; conditions in one watershed or national park will most likely be different compared to another. Ultimately, most of the independent variables used for this study did not reflect the scale of the study area.

Furthermore, the use of watersheds to analyze certain landscape characteristics could have negatively affected some of the results. As in the case of artificial surfaces, refer to Figure 14, where it was found to be insignificant as a predictor for secondary invasion when analyzed on a watershed scale; whereas previous studies have shown

otherwise when analyzed within 200 m of the lake. Also, the True Skill Statistics (TSS) error observed in Table 4 seems to be a result of having values that are too similar or the data not having enough variation. This is because 1 watershed can contain multiple lakes, which lead to multiple lakes having the same data.

Difficulties were encountered while gathering and processing data. Lake chemistry data were gathered during the early stages of this study. However, after being processed the records were not extensive enough to have value for all the lakes included in the sample. The same problem was observed for lake depth and boat access. Other data that were reported by previous studies to be indicative of a lake's habitability to invasive species, such as surface water temperature, Secchi depth, and chlorophyll levels were difficult to track down, and some were not yet available for download from the data sources that are listed in Section 3.2.4.

6.2 Direction for future studies

For future studies, coming up with variables that better reflect the scale of the study area is recommended. Water flow direction, for example, could point towards the amount of invasive species that are getting dispersed. As well, it is a variable that crosses boundaries. Another example would be soil composition of the lake's immediate surrounding. Moreover, developing models that are made of multiple independent variables should be considered to account for the complexity of examining secondary invasion on a large scale. Rather than having one independent variable be analyzed against the dependent variable, testing a complex model containing several independent variables against the dependent variable should be given consideration.

Certain landscape characteristics should also be analyzed in a scale that is closer in proximity to the lake to achieve better results. For example, the forest cover of a lake's vicinity would directly impact its water chemistry; whereas forest cover past a certain distance would most likely have little to no impact on the lake's water chemistry. Another example is the distance of cultivated land to a lake; cultivated land that are closer to a lake would have a greater impact on the lake rather than the distant ones.

Establishing a working relationship with government agencies can make data gathering easier. Government agencies and its affiliated organizations, such as OMNRF and DESC, have archived data that could be of use to a study with a similar theme as this one. Especially since some of their data are still undergoing the process of being published to the public. Gathering field data could also strengthen future studies. If going to the field is unrealistic, developing a survey that can be sent to park rangers or lake managers could be an option. Creating and using biological indexes could also address the complexities of a bigger study area. For example, biological indexes such as an index of invasive species in the lake, an index of potential predator or prey in the lake, an index of native or game fish in the lake, and an index of migratory birds that frequent the lake.

REFERENCES

Allendorf, F. W., & Lundquist, L. L. (2003). Introduction: Population biology, evolution, and control of invasive species. *Conservation Biology*, *17*(1), 24–30. https://doi.org/10.1046/j.1523-1739.2003.02365.x

Aron, W. I., & Smith, S. H. (1971). Ship Canals and Aquatic Ecosystems. *Science*, *174*(4004), 13–20.

Barnhisel, D. R., & Harvey, H. A. (1995). Size-specific fish avoidance of the spined crustacean *Bythotrephes* : field support for laboratory predictions. *Canadian Journal of Fisheries and Aquatic Sciences*, *52*(4), 768–775. https://doi.org/10.1139/f95-076

Berg, D. J., & Garton, D. W. (1988). Seasonal Abundance of the Exotic Predatory Cladoceran, Bythotrephes Cederstroemi, in Western Lake Erie. *Journal of Great Lakes Research*, *14*(4), 479–488. https://doi.org/10.1016/S0380-1330(88)71580-4

Berg, D. J., & Garton, D. W. (1994). Genetic differentiation in North American and European populations of the cladoceran *Bythotrephes*. *Limnology and Oceanography*, *39*(7), 1503–1516. https://doi.org/10.4319/lo.1994.39.7.1503

Bollens, S. M., Cordell, J. R., Avent, S., & Hooff, R. (2002). Zooplankton invasions: A brief review, plus two case studies from the northeast Pacific Ocean. *Hydrobiologia*, *480*, 87–110. https://doi.org/10.1023/A:1021233018533

Broad-scale Monitoring Program. (2017). Retrieved from https://www.ontario.ca/page/broad-scale-monitoring-program

Buchan, L. A. J., & Padilla, D. K. (2000). Predicting the Likelihood of Eurasian Watermilfoil Presence in Lakes, a Macrophyte Monitoring Tool. *Ecological Applications*, *10*(5), 1442–1455. https://doi.org/10.1890/1051-0761(2000)010[1442:PTLOEW]2.0.CO;2

Bur, M. T., Klarer, D. M., & Krieger, K. A. (1986). First Records of a European Cladoceran, Bythotrephes Cederstroemi, in Lakes Erie and Huron. *Journal of Great Lakes Research*, *12*(2), 144–146. https://doi.org/10.1016/S0380-1330(86)71711-5

Cardinale, B. J. (2011). Biodiversity improves water quality through niche partitioning. *Nature*, *472*(7341), 86–89. https://doi.org/10.1038/nature09904

Carpenter, S. R., Lathrop, R. C., & Street, N. P. (1999). Lake restoration : capabilities and needs, 19–28.

EDD MapS Ontario. (2017). Retrieved December 17, 2017, from https://www.eddmaps.org/ontario/

Environmental Research Systems Institute (ESRI). (2019). ArcGIS 10.6 Desktop Web

Help. Retrieved March 15, 2019, from
http://desktop.arcgis.com/en/arcmap/latest/tools/spatial-statistics-toolbox/h-how-
spatial-autocorrelation-moran-s-i-spatial-st.htm

Fu, W., Zhao, K., Zhang, C., & Tunney, H. (2011). Using Moran's I and geostatistics to
identify spatial patterns of soil nutrients in two different long-term phosphorus-
application plots. *J. Plant Nutr. Soil Sci*, *174*, 785–798.
https://doi.org/10.1002/jpln.201000422

Garton, D. W., Berg, D. J., & Fletcher, R. J. (1990). Thermal Tolerances of the
Predatory Cladocerans Bythotrephes cederstroemi and Leptodora kindti :
Relationship to Seasonal Abundance in Western Lake Erie. *Canadian Journal of
Fisheries and Aquatic Sciences*, *47*, 731–738. https://doi.org/10.1139/f90-083

Gertzen, E. L., & Leung, B. (2011). Predicting the spread of invasive species in an
uncertain world: Accommodating multiple vectors and gaps in temporal and spatial
data for Bythotrephes longimanus. *Biological Invasions*, *13*(11), 2433–2444.
https://doi.org/10.1007/s10530-011-0082-z

Getis, A. (2007). Reflections on spatial autocorrelation. *Regional Science and Urban
Economics*, *37*(4), 491–496. https://doi.org/10.1016/j.regsciurbeco.2007.04.005

Getis, A. (2008). A History of the Concept of Spatial Autocorrelation: A Geographer's
Perspective. *Geographical Analysis*, *40*(3), 297–309. Retrieved from
http://dx.doi.org/10.1111/j.1538-4632.2008.00727.x

GLANSIS. (2017). Great Lakes Aquatic Nonindigenous Species Information System.
Retrieved March 25, 2019, from
https://www.glerl.noaa.gov/glansis/about.html#progHist

Globe Land 30. (2017). Retrieved December 19, 2017, from
http://www.globallandcover.com/GLC30Download/index.aspx

Government of Ontario. (2017). About Ontario | Ontario.ca. Retrieved February 25,
2019, from https://www.ontario.ca/page/about-ontario#section-4

Havel, J. E., Kovalenko, K. E., Thomaz, S. M., Amalfitano, S., & Kats, L. B. (2015).
Aquatic invasive species: challenges for the future. *Hydrobiologia*, *750*(1), 147–
170. https://doi.org/10.1007/s10750-014-2166-0

Hulme, P. E. (2006). Beyond control: Wider implications for the management of
biological invasions. *Journal of Applied Ecology*, *43*(5), 835–847.
https://doi.org/10.1111/j.1365-2664.2006.01227.x

Jarvis, P. J. (1979). The ecology of plant and animal introductions. *Progress in Physical
Geography*, *3*(2), 187–214. https://doi.org/10.1016/B978-0-08-047514-1.50014-7

Kandane-Rathnayake, R. K., Enticott, J. C., & Phillips, L. E. (2013). Data distribution:

normal or abnormal? *TRANSFUSION*, *53*, 257–259. https://doi.org/10.1111/j.1537-2995.2012.03814.x

Kerr, S. J., Brousseau, C. S., & Muschett, M. (2005). Invasive Aquatic Spcies in Ontario: A Review and Analysis of Potential Pathways for Introduction. *Fisheries*, *30*(1), 10–17. https://doi.org/10.1577/1548-8446(2005)30

Lachner, E. a., Robins, C. R., & Courtenay, W. R. (1970). Exotic fishes and other aquatic organisms introduced into North America. *Smithsonian Contributions to Zoology*, (59), 1–29. https://doi.org/10.5479/si.00810282.59

Land Information Ontario. (2010). *Ontario Hydro Network (OHN) – Waterbody*.

Land Information Ontario. (2012). *Fishing Access Point*. Retrieved from https://www.sse.gov.on.ca/sites/MNR-PublicDocs/EN/CMID/Fishing Access Point - Data Description.pdf

Land Information Ontario. (2015a). *Aquatic Resource Area Summary*.

Land Information Ontario. (2015b). User Guide for Watershed.

Land Information Ontario. (2016). *Bathymetry*.

Land Information Ontario. (2017). Retrieved December 18, 2017, from https://www.ontario.ca/page/land-information-ontario

Land Information Ontario. (2018). *Provincial Digital Elevation Model User Guide*. Retrieved from https://www.sse.gov.on.ca/sites/MNR-PublicDocs/EN/CMID/PDEM_UserGuide.pdf

Lange, C., & Cap, R. (1986). Bythotrephes Cederstroemi (Schödler). (Cercopagidae: Cladocera): A New Record for Lake Ontario. *Journal of Great Lakes Research*, *12*(2), 142–143. https://doi.org/10.1016/S0380-1330(86)71710-3

Laurenson, L. B. J., & Hocutt, C. H. (1986). COLONISATION THEORY AND INVASIVE BIOTA: THE GREAT FISH RIVER, A CASE HISTORY. *Environmental Monitoring and Assessment*, *6*(1985), 71–90.

Legendre, P. (1993). *SPATIAL AUTOCORRELATION: TROUBLE OR NEW PARADIGM? Ecology* (Vol. 74). Retrieved from https://about.jstor.org/terms

Lehman, J. T., & Caceres, C. E. (1993). Food-Web Responses to Species Invasion by a Predatory Invertebrate: Bythotrephes in Lake Michigan. *Limnology and Oceanography*, *38*(4), 879–891. Retrieved from https://www-jstor-org.ezproxy.lib.ryerson.ca/stable/pdf/2837929.pdf?refreqid=excelsior%3A79b598006bc36e1a418158d2b9251a26

Lyon, A. (2014). Why are Normal Distributions Normal? *British Journal for the*

Philosophy of Science, 65(3), 621–649. https://doi.org/10.1093/bjps/axs046

MacIsaac, H. J., Borbely, J. V. M., Muirhead, J. R., & Graniero, P. A. (2004). Backcasting and forecasting biological invasions of inland lakes. *Ecological Applications*, *14*(3), 773–783. https://doi.org/10.1890/02-5377

Mack, R. N., Simberloff, D., Lonsdale, W. M., Evans, H., Clout, M., & Bazzaz, F. A. (2000). Biotic invasions: Causes, epidemiology, global consequences, and control. *Ecological Applications*, *10*(3), 689–710. https://doi.org/10.1890/0012-9623(2005)86[249b:IIE]2.0.CO;2

McFadden, D. (1977). *QUANTITATIVE METHODS FOR ANALYZING TRAVEL BEHAVIOUR OF INDIVIDUALS: SOME RECENT DEVELOPMENTS. Cowles Foundation Discussion Paper.* New Haven, Connecticut: Cowles Foundation for Research in Economics at Yale University. Retrieved from http://cowles.yale.edu/sites/default/files/files/pub/d04/d0474.pdf

Menard, S. (2013). *Introduction: Linear Regression and Logistic Regression In: Logistic Regression: From Introductory to Advanced Concepts and Applications Introduction: Linear Regression and Logistic Regression*. https://doi.org/10.4135/9781483348964

Mills, E. L., Leach, J. H., Carlton, J. T., & Secor, C. L. (1993). Exotic Species in the Great Lakes: A History of Biotic Crises and Anthropogenic Introductions. *Journal of Great Lakes Research*, *19*(1), 1–54. https://doi.org/10.1016/S0380-1330(93)71197-1

Mills, E. L., Leach, J. H., Carlton, J. T., & Secor, C. L. (1994). Exotic Species and the Integrity of the Great Lakes. *BioScience*, *44*(10), 666–676. https://doi.org/10.2307/1312510

Muirhead, J. R., & MacIsaac, H. J. (2005). Development of inland lakes as hubs in an invasion network. *Journal of Applied Ecology*, *42*(1), 80–90. https://doi.org/10.1111/j.1365-2664.2004.00988.x

Osborne, J. W. (2017). *A Conceptual Introduction to Bivariate Logistic Regression In: Best Practices in Logistic Regression*. https://doi.org/10.4135/9781483399041

Paul, P., & Kar, T. K. (2016). Impacts of invasive species on the sustainable use of native exploited species. *Ecological Modelling*, *340*, 106–115. https://doi.org/10.1016/j.ecolmodel.2016.09.002

Pejchar, L., & Mooney, H. A. (2009). Invasive species, ecosystem services and human well-being. *Trends in Ecology and Evolution*, *24*(9), 497–504. https://doi.org/10.1016/j.tree.2009.03.016

Pimentel, D., Zuniga, R., & Morrison, D. (2005). Update on the environmental and economic costs associated with alien-invasive species in the United States.

Ecological Economics, *52*(3 SPEC. ISS.), 273–288. https://doi.org/10.1016/j.ecolecon.2004.10.002

Puth, L. M., & Post, D. M. (2005). Studying invasion: Have we missed the boat? *Ecology Letters*, *8*(7), 715–721. https://doi.org/10.1111/j.1461-0248.2005.00774.x

Pyšek, P., & Richardson, D. M. (2010). Invasive Species, Environmental Change and Management, and Health. *Annual Review of Environment and Resources*, *35*(1), 25–55. https://doi.org/10.1146/annurev-environ-033009-095548

Rangel, T. F., Alexandre, J., Diniz-Filho, F., Bini, L. M., Rangel, T. F., Diniz-Filho, J. A. F., & Bini, L. M. (2010). SAM: a comprehensive application for Spatial Analysis in Macroecology. *Ecography*, *33*. https://doi.org/10.1111/j.1600-0587.2009.06299.x

Ricciardi, A., & Rasmussen, J. B. (1998). Predicting invasions: Propagule pressure and the gravity of allee effects. *Canadian Journal of Fisheries and Aquatic Sciences*, *55*, 1759–1765. https://doi.org/10.1890/02-0571

Roley, S. S., & Newman, R. M. (2008). Predicting Eurasian watermilfoil invasions in Minnesota. *Lake and Reservoir Management*, *24*(4), 361–369. https://doi.org/10.1080/07438140809354846

Rupert, M. G., Cannon, S. H., Gartner, J. E., Michael, J. A., & Helsel, D. R. (2008). *Using Logistic Regression to Predict the Probability of Debris Flows in Area Burned by Wildfires, Southern California, 2003-2006*. Reston, Virginia. Retrieved from https://pubs.usgs.gov/of/2008/1370/pdf/OF08-1370_508.pdf

Shaker, R. R., & Rapp, C. J. (2013). Investigating Aquatic Invasive Species Propagation within the Adirondack Region of New York: A Lake and Landscape Approach. *Papers in Applied Geography*, *36*, 200–209.

Shaker, R. R., Rapp, C. J., & Yakubov, A. D. (2013). Examining Patterns of Aquatic Invasion Within the Adirondacks: an Ols and Glm Approach. *Middle States Geographer*, *46*(1979), 1–11.

Shaker, R. R., Yakubov, A. D., Nick, S. M., Vennie-Vollrath, E., Ehlinger, T. J., & Wayne Forsythe, K. (2017). Predicting aquatic invasion in Adirondack lakes: a spatial analysis of lake and landscape characteristics. *Ecosphere*, *8*(3), e01723. https://doi.org/10.1002/ecs2.1723

Spiny and Fishhook Waterfleas. (n.d.). Ontario's Invading Species Awareness Program. Retrieved from http://www.invadingspecies.com/invaders/invertebrates/spiny-and-fishhook-waterflea/

Sprules, W. G., Riessen, H. P., & Jid, E. H. (1990). *DYNAMICS OF THE BYTHOTREPHES INVASION OF THE ST. LAWRENCE GREAT LAKES*. *Journal of Great Lakes Research* (Vol. 16). https://doi.org/10.1016/S0380-1330(90)71429-3

Statistics Canada. (2008). *Road Network File, Reference Guide 92-500-GWE. Statistics Canada, Ottawa*.

Statistics Canada. (2017). *Boundary files, reference guide census year 2016.* Retrieved from http://www.statcan.gc.ca/pub/92-160-g/92-160-g2016002-eng.pdf%0Ahttp://epe.lac-bac.gc.ca/100/200/301/statcan/boundary_files_guide_2011-ef/92-160-g2011001-eng.pdf

Statistics Canada. (2019). About StatCan. Retrieved from https://www.statcan.gc.ca/eng/about/statcan?MM=1

The Dorset Environmental Science Centre. (2017). Retrieved December 19, 2017, from http://desc.ca/

Usher, M. B. (1986). Invasibility and Wildlife Conservation : Invasive Species on Nature Reserves. *Philosophical Transcations of the Royal Society of London. Series B, Biological Sciences*, *314*(1167), 695–710.

Usher, M. B. (1988). Biological invasions of nature reserves: A search for generalisations. *Biological Conservation*, *44*(1–2), 119–135. https://doi.org/10.1016/0006-3207(88)90007-9

Usher, M. B., Kruger, F. J., Macdonald, I. A. W., Loope, L. L., & Brockie, R. E. (1988). The ecology of biological invasions into nature reserves: An introduction. *Biological Conservation*, *44*(1–2), 1–8. https://doi.org/10.1016/0006-3207(88)90002-X

Valcu, M., & Kempenaers, B. (2010). Spatial autocorrelation: an overlooked concept in behavioral ecology. *Behavioral Ecology*, *21*(5), 902–905. https://doi.org/10.1093/beheco/arq107

Wahlström, E., & Westman, E. (1999). Planktivory by the predacious cladoceran <i>Bythotrephes longimanus</i>: effects on zooplankton size structure and abundance. *Canadian Journal of Fisheries and Aquatic Sciences*, *56*(10), 1865–1872. https://doi.org/10.1139/cjfas-56-10-1865

Wang, L., & Jackson, D. A. (2011). Modeling the establishment of invasive species: habitat and biotic interactions influencing the establishment of Bythotrephes longimanus. *Biological Invasions*, *13*, 2499–2512. https://doi.org/10.1007/s10530-011-0071-2

Wang, X., & Huang, G. (2015). Technical Report: Development of High-Resolution Climate Change Projections under RCP 8.5 Emissions Scenario for the Province of Ontario, 1–90.

Weisz, E. J., & Yan, N. D. (2010). Relative value of limnological, geographic, and human use variables as predictors of the presence of Bythotrephes longimanus in Canadian Shield lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, *67*(3),

462–472. https://doi.org/10.1139/F09-197

Yan, N. D., Blukacz, A., Sprules, W. G., Kindy, P. K., Hackett, D., Girard, R. E., & Clark, B. J. (2001). Changes in zooplankton and the phenology of the spiny water flea, *Bythotrephes*, following its invasion of Harp Lake, Ontario, Canada. *Canadian Journal of Fisheries and Aquatic Sciences*, *58*(12), 2341–2350. https://doi.org/10.1139/cjfas-58-12-2341

Yan, N. D., Dunlop, W. I., Pawson, T. W., & MacKay, L. E. (1992). *Bythotrephes cederstroemi* (Schoedler) in Muskoka Lakes: First Records of the European Invader in Inland Lakes in Canada. *Canadian Journal of Fisheries and Aquatic Sciences*, *49*(2), 422–426. https://doi.org/10.1139/f92-048

Yan, N. D., Girard, R., & Boudreau, S. (2002). An introduced invertebrate predator (*Bythotrephes*) reduces zooplankton species richness. *Ecology Letters*, *5*(4), 481–485. https://doi.org/10.1046/j.1461-0248.2002.00348.x

Yan, N. D., Leung, B., Lewis, M. A., & Peacor, S. D. (2011). The spread, establishment and impacts of the spiny water flea, *Bythotrephes longimanus*, in temperate North America: A synopsis of the special issue. *Biological Invasions*, *13*(11), 2423–2432. https://doi.org/10.1007/s10530-011-0069-9

Yan, N. D., & Pawson, T. W. (1998). Seasonal variation in the size and abundance of the invading Bythotrephes in Harp Lake, Ontario, Canada. *Hydrobiologia*, *361*, 157–168.