Predicting *Dreissena* spp. presence and biomass as a function of Lake Huron environmental characteristics

Jennifer Wardell

A thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science (School for Environment and Sustainability) at the
University of Michigan
August 2021

Thesis Committee:

Catherine Riseng[1], Co-advisor

Peter Esselman[2], Co-advisor

Mark Rowe[3]

Ashley Elgin[3]

[1]University of Michigan School for Environment and Sustainability

[2]US Geological Survey Great Lakes Science Center

[3]National Oceanic Atmospheric Administration Great Lakes Environmental Research Laboratory

Table of Contents

List of Figures

List of Tables

Abstract

Invasive dreissenid mussels (*Dreissena polymorpha* and *Dreissena rostriformis bugensis*) have had multiple effects on Great Lakes ecosystems, changing benthic habitats and altering food web structure and nutrient cycling on lake-wide scales. The severity of these effects depends on dreissenid density and biomass distributions, therefore, dreissenid abundances must be well understood before their impacts can be accurately estimated and managed. This thesis explores whether environmental factors can quantify and predict dreissenid distributions and abundances on a lake-wide scale. Spatial trends in dreissenid distribution and abundance were examined using dreissenid mussel abundance data from the 2017 Lake Huron Coordinated Science and Monitoring Initiative (CSMI) survey. A suite of models was developed using ten environmental explanatory variables to interpolate and predict dreissenid presence and biomass. Specifically, I developed an empirical Bayesian kriging model ($r^2 = 0.61$), boosted regression tree models for abundance of quagga mussels ($r^2 = 0.27$) and both mussel species ($r^2 = 0.23$), and a boosted regression tree model of dreissenid presence (ROC score 0.735). The most important explanatory variables for dreissenid biomass were February maximum bottom temperature, substrate, bathymetry, tributary influence, and April colored dissolved organic matter. Dreissenid presence was most well described by distance from the Saginaw River, bathymetry, and February maximum bottom temperature. The inclusion of bathymetry and temperature in all the best performing models indicates the high importance of these variables for dreissenid distribution and abundance. While the models developed performed adequately, future distribution models could be improved through increases in survey effort and improved information about mussel habitat characteristics and environmental constraints. Advanced survey technologies such as autonomous underwater vehicles may be particularly helpful to quickly produce better species abundance and habitat data. Future predictive modeling efforts will be enhanced by incorporating data for other important explanatory variables, such as ionic calcium concentration, currents, and benthic food availability.

I. Introduction

A. Background

The Laurentian Great Lakes comprise the largest freshwater ecosystem in the world and the immense surface area and water volume support complex ecosystem functions and processes through physical, chemical, and biological components (Wang et al., 2015; Wehrly et al., 2013). Dreissenid mussels have increasingly infested the Great Lakes since the late 1980s (Baldwin et al., 2002). The first invasion was to Lake St. Clair in 1988 (spreading quickly into Lake Erie the same year) by the zebra mussels (*Dreissena polymorpha*), a well-known ecosystem engineer (O'Neill, 1994; Benson, 2014). Zebra mussels altered the entire lake bottom structure, providing more ideal attachment substrates in the benthos, contributing to the invasion success of the quagga mussels (*Dreissena rostriformis bugensis*) (Mayer et al., 2014; Orlova, 2014; Bootsma et al., 2015; Karatayev et al., 2018b). Dreissenids also have had large impacts on the food web, nutrient cycling, and benthic community by turning over large volumes of the water column as suspension-feeding planktivores (Conn et al., 2014). The magnitude of these impacts depends on mussel clearance rate, population density, and food availability (Ward & Ricciardi, 2014; Shen et al., 2020).

The Great Lakes encompass a wide range of climatic conditions, depths, geologies, and human influences that could be incorporated into analytical approaches to predict dreissenid mussel abundances. Importantly, the Great Lakes have been widely monitored for dreissenid presence and abundance for several decades, and basin-wide data exist summarizing habitat factors thought to limit the mussels. Therefore, the Great Lakes make an excellent test geography for combining available spatial datasets to explore models that predict mussel presence and biomass given a wide variety of lake conditions. Predictive habitat models have the potential to inform invasive species management, prevention, and control efforts in places that are in the early stages of invasion, or that may be invaded in the future.

Niche theory suggests that the local presence and abundance of a species will be determined by the interaction of the species' physiological and behavioral limits with the abiotic and biotic conditions present in the ecosystem (Hutchinson, 1957). Models of spatial

variation in species presence and abundance necessarily try to capture the spatial representations or proxies for the factors that limit the species' niche. Numerous studies of dreissenid ecology yield insights into the abiotic and biotic factors that correlate most strongly to their presence and abundance. These studies are summarized below.

B. Factors affecting Dreissena spp. distributions and biomass

The combination of physical, chemical and biological factors present at any specific geographic location affects the ability of dreissenid mussels to live, grow, and reproduce. If conditions are unsuitable for mussels, then they will need to move in order to survive, but their ability to move is limited by their morphology, life stage, and environment, and mussels may not have the energy necessary to move to a better location (Kobak, 2014). Therefore, adult mussels tend to stay in a location unless conditions are consistently unfavorable and will ultimately die if they cannot relocate. The factors described below have been found by previous research to effect dreissenid mussels, however none of these factors act alone and different combinations of conditions may lead to different population outcomes.

a. Water Chemistry

Some chemicals are toxic to dreissenids and mussels will not be found in areas where elevated concentrations of these chemicals are present. Chlorine kills mussels through asphyxiation and limited glycolysis if they are exposed over a prolonged period of time. Copper is also toxic to mussels and many other aquatic organisms, including algae and other invertebrates. High levels of potassium can kill mussels by interfering with their ability to transfer oxygen across gill tissue, resulting in asphyxia. Elevated salinity, above 4 ppt, kills mussels by causing an ionic tissue imbalance (Fernald & Watson, 2014).

In contrast, some chemicals are essential for the survival of dreissenids and need to be present in certain minimum concentrations for the success of mussel populations. The carrying capacity for dreissenid mussels is affected by many factors including nutrient loads, calcium concentrations, and total phosphorus (Vanderploeg et al., 2014). These essential chemicals are described in greater detail below.

i. Calcium

Dreissenid shells are made primarily out of aragonite (calcium carbonate) and therefore the stability of their shells is influence by both Ca^{2+} and carbonate ions (Phillips et al., 2015). Based on past studies, mussels will not colonize areas with less than 12-15 mg/L of Ca^{2+} (Marsden et al., 2014, Garton et al., 2014). Calcium is required by mussels for colonization, population expansion, and growth (Marsden et al., 2014). Mussels in areas with low calcium concentrations will increase their utilization efficiency and use less calcium for their shells (Marsden et al., 2014). Although the mussels can adapt to survive in these conditions, they tend not to thrive. Low calcium concentrations cause shells to lose their periostracum or outer shell layer, which causes them to be white, chalky, and more fragile (Marsden et al., 2014). In areas with high calcium, mussels will be able to make thicker shells, which provides them with more protection from predators. As a result, in low-calcium waters, mussels may be more vulnerable to predators and may need to alter their behavior more than mussels that are in high-calcium waters (Czarnoleski & Muller, 2014).

The correlation between mussel colonization and density with calcium concentrations is confounded by currents, temperature, and productivity gradients (Marsden et al., 2014). This along with the ability of dreissenid mussels to mobilize calcium from their shells to their hemolymph under low calcium conditions, in order to achieve the critical amount for survival, makes it very difficult to determine the minimum calcium concentration dreissenids require (Garton et al., 2014). The concentration of carbonate ions is also correlated with pH (Phillips et al., 2015).

ii. pH

Mussels typically are found where pH values are between 6.5 and 8.5 since outside of this range, mussels can be killed through shock or from the ionic loss of calcium, sodium, and potassium that they need to survive (Fernald & Watson, 2014). At lower pH values, the uptake rates of Na^+ and Ca^{2+} also decrease and impair the ability of mussels to regulate and maintain shell integrity, osmotically and ionically (Garton et al., 2014). In some cases, mussels are able to adapt to lower pH conditions (in the 5.5 to 6.0 range) if they are exposed for long periods of time (Fernald & Watson, 2014), however, it is not fully understood what conditions allow for mussels to adapt to low pH.

iii.   *Dissolved oxygen*

Mussels require oxygen to live and will die in areas where there are low dissolved oxygen concentrations (Schwalb et al., 2015). Fernald & Watson (2014) found that dissolved oxygen concentrations of less than 4.0 ppm will kill the entire population of dreissenid mussels. In addition, it is likely that dreissenid mussels have a low tolerance to hypoxia since mussels are poor regulators of oxygen uptake rate when oxygen concentrations decline (Garton et al., 2014). Zebra mussels may require higher oxygen concentrations than quagga mussels for survival and quagga mussels will survive longer under anoxic conditions than zebra mussels (Karatayev et al., 2014), especially in colder waters (less than 11°C) (Tyner et al., 2015; McMahon, 1996).

iv.   *Temperature*

Dreissenid mussels have an ideal temperature range just like most other species. Studies have identified different ideal temperature ranges for dreissenids, but between 15 and 28°C is generally accepted (Chakraborti et al., 2014; Kobak, 2014; Boeckman & Bidwell, 2014). Temperature extremes in either direction increase mussel mortality (Schwalb et al., 2015) and temperatures greater than 30°C can be lethal (Chakraborti et al., 2014; Boeckman & Bidwell, 2014). Zebra mussels are more tolerant to high temperatures than quagga mussels (maximum of 33°C versus 31°C) (Karatayev et al., 2014), but quagga mussels have a lower thermal range (2 to 3°C colder) than zebra mussels (Garton et al., 2014). However, both species have been found to have the same lower temperature limit of 0°C (Garton et al., 2014). The tolerance of quagga mussels to colder temperatures allows them to expand into the colder deep profundal zones of lakes where zebra mussels cannot survive (Williamson & Ozersky, 2019; Pilcher et al., 2015; Orlova, 2014; Higgins, 2014; Bootsma & Liao, 2014).

Higher temperatures have also been found to decrease veliger densities, such as associated with seasonal variation where veliger, the pelagic larval stage of *Dreissena,* densities will decrease in July and August and then increase again in September (Boeckman & Bidwell, 2014). The veliger life stage is critical to the dispersal of dreissenid mussels throughout the lake. Temperatures between 15°C and 20°C for adults appears to be best for byssal thread attachment strength (Kobak, 2014). Ideal temperature ranges vary with the size of the individual mussel since large mussels have been found to be less tolerant of higher

temperatures than smaller mussels (Boeckman & Bidwell, 2014). In fact, many dreissenid physiological processes are size-dependent, resulting in exponential changes in feeding and oxygen consumption based on size (Young et al., 1996; Tyner et al., 2015).

Variability in temporal temperature patterns is also an important factor in dreissenid biology. Earlier warming in the season results in higher fecundity and more rapid growth for adults (Marsden et al., 2014) due to a minimum temperature required for dreissenid spawning and larval development. Eggs are incapable of fertilization at less than 10°C and spawning does not typically occur until 15°C, although it has been found to occur in zebra mussels at as low as 11°C (Garton et al., 2014) and as low as 9°C for quagga mussels (Claxton & Mackie, 1998) in deeper waters.

Temperature also indirectly affects dreissenid populations through its effects on other chemical and physical factors. Variation in temperature affects the ability of dreissenid mussels to tolerate higher salinity levels (Wright et al., 1996). Temperature is directly related to relative humidity and dreissenids out of the water can withstand periods of desiccation for 3 to 5 days when conditions are cool and moist (Benson, 2014; Garton et al., 2014). In contrast, increased air temperature and decreased relative humidity cause a decrease in emersion and desiccation tolerance (Benson, 2014; Garton et al., 2014). Temperature also alters the effects of hypoxia on dreissenid mussels so that with the same low oxygen concentration, the time to death is significantly shorter in higher temperatures than in lower temperatures (Chakraborti et al., 2014).

Tolerance to temperatures outside of the ideal range differs based on several factors. Increased acclimation time, decreased latitude, increased nutritional condition, increased size, and the season of sampling all contribute to increased temperature tolerance (Garton et al., 2014). For example, Garton et al. (2014) found that individuals collected in the summer have higher temperature tolerances than individuals collected during the other seasons even after being exposed to lower acclimation temperatures.

b. Hydrodynamics

Hydrodynamics play a large role in the lives of dreissenid mussels since they are extremely important for several life history functions, including external fertilization of

gametes, dispersal of larvae, settlement of larvae and recruitment of juvenile mussels, suspension feeding, dislodgement of adult benthic stages, expansion, and growth (Ackerman, 2014; Marsden et al 2014). Horizontal and vertical water movement deliver food particles to mussel beds where they can be ingested while filter feeding. Ackerman (1999) found that the saturation of feeding for quagga mussels occurred between 10 to 20 cm/s. Tuchman et al. (2004) found that higher velocities (greater than 20 cm/s) delivered more algae to mussels, however, ingestion rates only increased in smaller individuals. Zebra mussels have also been found to outcompete quagga mussels under high water velocities (Peyer et al., 2009).

i. Currents

Currents, defined by both speed and direction, affect dreissenid abundance since some currents are necessary for food availability, but higher currents can be detrimental (Vaate et al., 2014; Ackerman, 2014). To maintain a sedentary position in areas with higher currents, dreissenids synthesize a strand of protein that is internally anchored and then attached to the substrate, a process known as byssal adhesion. In a study by Peyer et al. (2009), mussels were found to maintain byssal adhesion in current velocities up to 180 cm/s in the Lake Michigan nearshore. According to Chakraborti et al. (2014), mussels will settle, attach to, and grow on surfaces where water flows are less than 200 cm/s. However, Clarke & McMahon (1996) found that only 27 cm/s current velocity resulted in the decline of mussel populations in the lock walls of the Niagara River. Therefore, the literature asserts that the tolerable current velocity can vary greatly by population. The examples included here assume hard substrates are available for attachment; dreissenids are unable to attach with the same strength to soft substrates, limiting their ability to remain anchored to the benthos due to the movement and potential resuspension of soft substrates.

Currents are influenced by a variety of factors and can be difficult to predict. Some of these factors are solar energy, winds, hydraulic gradients, density gradients, and the earth's rotation (Ackerman, 2014). For example, stronger storm events lead to an increased flux of organic matter from the pelagic zone to the benthos, which provides mussels with more food than if the storm event was more mild or not present (Schwalb et al., 2015). Gradients in velocity determine the hydrodynamic forces that mussels experience (Ackerman, 2014). For example, shallower regions of the lake that have high current

velocities will have lower densities of mussels since mussels are not well adapted to stay attached to the substrate in these conditions (Garton et al., 2014). Currents disproportionately affect mussels of different sizes since larger mussels have a stronger attachment strength than smaller mussels (Kobak, 2014). This means that larger mussels will be found in areas with higher currents than smaller mussels since they are able to hold on to the substrates and not be washed away.

Currents are key to dreissenid veliger dispersal and survival. When mussels reproduce, the eggs and sperm are released simultaneously, and fertilization takes place externally (Kashian & Ram, 2014). Larvae will move away via active and passive transport from the site of fertilization, and these dispersal modes are sensitive to water velocity. Dreissenid larvae are capable of free-swimming, however, their swimming speeds have never been measured experimentally (Ackerman, 2014), which prevents us from determining their ultimate settling location. In general, larvae will settle near where the egg is fertilized and not move much further, except in areas of higher currents. Strong winds can cause currents that result in the movement of planktonic larvae to deeper in the water column or to windward shores (Lewandowski & Stanczykowska, 2014). Smaller larvae are generally thought to be too small and weak to overcome water movements and will settle wherever the currents take them (Kobak, 2014). Therefore, currents are primarily responsible for the spread of dreissenids throughout the lake and the expansion of their population range.

Currents may also result in larvae settling in areas of poorer condition. The future survival, growth, and reproductive success of an individual depends on where larvae end up settling (Kobak, 2014). Larvae have a preference of settling in the vicinity of adults, which is one explanation for the large and dense aggregations of mussels generally found (Kobak, 2014). Another possible explanation for dense aggregations is that adult mussels move to areas with other mussels due to better conditions and there are benefits to being in a group for predator avoidance (Kobak, 2014). Larval dispersal by currents could transport the fertilized egg far away from the aggregated adults and it could settle in a less than ideal area (Kashian & Ram, 2014).

ii. Turbidity

Currents also affect the turbidity of the lake through entraining fine sediments into the water column, which can inhibit feeding at high turbulence levels due to filtration instability, behavioral instability, and reduced optimal foraging (Ackerman, 2014). Turbidity greater than 20 NTU (Nephelometric Turbidity units) is generally the threshold for dreissenid survival since these higher levels can inhibit oxygen uptake and consumption rates (Garton et al., 2014).

iii. Vertical mixing

Mussels also rely on mixing of the water column to supply resources and remove wastes as well as to facilitate fertilization of gametes and to disperse larvae (Ackerman, 2014). Vertical mixing regulates the supply of seston from the water column to the mussels and affects the fate of dissolved phosphorus excreted by dreissenids (Bootsma & Liao, 2014).

c. Habitat conditions/properties

The overall productivity of dreissenid mussels is correlated with mussel colonization rates, population expansion, and growth (Marsden et al., 2014). As discussed above, larger mussels have stronger attachment strength than smaller mussels. This is advantageous for currents, but not good for other habitat conditions since this stronger attachment strength also means that more energy is required to move from an area that is not ideal (Kobak, 2014).

i. Depth

Mussel growth and recruitment rates are much higher at moderate depths than deeper depths, especially when comparing to areas below the thermocline (Karatayev et al., 2018b). This is likely due to limited food availability deeper in the lake, especially at depths greater than 100 m (Houghton et al., 2014). There is, however, a deep-water morph of the quagga mussel that spawns at a lower temperature and is more suited to the colder temperatures found in the deeper regions of the lake (Nalepa et al., 2014). This is important to consider, now that a large proportion of the mussels found in the Great Lakes, including Lake Huron, are quagga mussels. In general, with increased depth, turbulence is diminished, substrates become more available, and benthic fine organic material increases as material settles and accumulates (Nalepa et al., 2014). Because of this, if the deep-water areas were equal in all other respects, they would be better as dreissenid mussel habitat than the shallower areas.

However, this is not true since the lower temperature and lower food availability of deeper sites outweighs the benefits. Shallow areas that are near shorelines can support more mussel growth than deeper areas as a result of increased phytoplankton biomass (commonly measured using chlorophyll-a concentrations), substrate quality, temperature, and increased total suspended solids (TSS) (Chakraborti et al., 2014); however, dreissenid habitat in the nearshore is limited by higher currents that disrupt attachment. Therefore, biomass is highest at the middle depths of the lake, just below the thermocline. Supporting this, Nalepa et al. (2014) found that dreissenid biomass in southern Lake Michigan was the highest in the 31-50m depth range and concluded this was a result of the "best" combination of conditions for temperature, currents, food availability, substrate and more.

ii. *Substrates*

Dreissenid mussels can be found on a variety of substrate types but show preferences for different materials based on surface roughness, which facilitates adhesion and bacterial biofilms, which can either increase or decrease adhesion depending on the biofilm composition (Kobak, 2014). The type of substrate affects the attachment strength of mussel byssal threads. As stated previously, attachment strength is especially important in areas with faster and stronger water currents or in areas with high predator abundance. In most areas, it appears that hard substrates are preferred over soft substrates, but the ability to survive on soft substrates differs between quagga and zebra mussels. Zebra mussels will not be present on soft substrates and quagga mussels will likely be present in lower numbers on soft substrates, unless there are other factors at play (Kumar et al., 2016; Pilcher et al., 2017; Orlova, 2014; Higgins, 2014; Bootsma & Liao, 2014; Williamson and Ozerksy, 2019). This may be in part due to the lower oxygen concentrations found in finer substrates. The ability to settle on all lake bottom substrates is a competitive benefit to quagga mussels and is likely one of the main reasons for their rapid expansion to areas of the lake where zebra mussels are not found. Mussel density and biomass tends to rapidly decline with decreased substrate particle size from boulders through cobbles, gravel, sand, silt, and mud (Garton et al., 2014). Overall, hard substrates appear to be favored due to the higher habitat complexity providing more area to attach to, but also more protection from predators (Kumar et al., 2016, Pilcher et al., 2017). Settling on objects raised above the bottom, such as large rocks, also provides

benefits since mussels may have greater access to food delivered by currents (Kobak, 2014). Although mussels prefer large, raised objects, they appear to avoid inclined surfaces and will exert energy to detach and move to an area that is flatter so that they are not on a slope (Kobak, 2014). Dreissenids also prefer to settle on rocks that are rough, as opposed to smooth, due to the increased attachment surface (Kobak, 2014). Interestingly, concave oriented surfaces have been found to support greater settlement rates due to protection from water currents (Garton et al., 2014). Areas with large, rough substrate, such as boulders, will have even higher densities of dreissenids due to the ability of boulders to withstand the impacts of storm currents and provide protection to mussels so they are less likely to be involuntarily detached (Ackerman, 2014).

Dreissenid mussels prefer to settle on darker substrates instead of lighter substrates if given a choice due to predator avoidance benefits. Mussels are better able to blend in with darker substrates, making them harder to spot by predators, and more likely to open their valves to feed (Kobak, 2014). On lighter substrates with reduced feeding rates, mussels may not achieve the necessary energy for growth and survival and if they survive, they may be smaller in size (Kobak, 2014).

In some cases, mussels will attach to each other. In these cases, it is generally because there is no good substrate left to attach to and so they "double-up" to survive (Kobak, 2014). Dreissenid mussels have also been observed to attach to other species of mussels and other organisms, such as unionids and macrophytes (Kobak, 2014). Settling on other organisms is obviously more beneficial than settling on individuals of the same species, however, it is not necessarily better than settling on good quality substrates. Where there are rough substrates that are ideal for mussel attachment, there will be fewer mussels found attached to other organisms, however, when the only available substrate is sand or silt, then there will be more mussels found attached to other organisms (Lucy et al., 2014).

Substrates can also affect fertilization success in combination with water velocity. Bed roughness conditions in mussel beds generate recirculation zones that can enhance the likelihood of a sperm-egg encounter (Quinn & Ackerman, 2011). Understanding the

hydrodynamics of flow at the bed level can be useful for predictive models to identify areas that can be managed to reduce the probability of fertilization (Hasler et al., 2019).

d. Interactions with other organisms (interspecific interactions)

i. Food availability

Dreissenid mussels have varying distributions based on food availability. Food sources can come from a variety of places; however, it is agreed upon in the scientific community that phytoplankton and zooplankton are very important for the success of dreissenid mussels, both when adults and during their veliger (larval) stages (Claxton & Mackie, 1998; Wacker & von Elert, 2003; Wong et al., 2014; Pilcher et al., 2015; Boegehold et al., 2019). Mussels depend on water movements to deliver food particles which can arrive horizontally via advection, from above via particle settling and vertical mixing of the water column, or from below via resuspension of benthic materials (Ackerman, 2014). The best current speed for increased mussel density would be moderate to high in order to provide food, but not too high that it causes the mussels to get washed away or unable to feed due to the instability caused by currents (Ackerman, 2014).

Mussel diets shift in composition with the disappearance of phytoplankton blooms on a seasonal basis (Pilcher et al., 2015). When phytoplankton is not present or in low quantities, mussels will feed on dissolved organic matter, bacteria, and zooplankton in order to get the nutrients needed to survive (Wong et al., 2014). Because of this diet plasticity, mussels can survive and maintain growth even if there are low chlorophyll-a concentrations (Wong et al., 2014). Mussels prefer particulate food less than 50 μm, although they can utilize and filter particles anywhere from 0.4 to 450 μm (Alderstein et al., 2014). Generally, areas in the hypolimnion and the near-bottom boundary layer will have much lower mussel densities due to low food densities in these areas (Karatayev et al., 2018b). In addition, increased mussel longevity and larger mussel sizes are found in littoral habitats compared to profundal habitats (Karatayev et al., 2018b). Various factors affect the abundance of phytoplankton and other food sources for dreissenid mussels, but extreme eutrophication can affect the fatty acid content of food (Bril et al., 2017). If the food quality is lowered, even if food availability is high, metabolic processes of mussels will be affected and they will not

have as high of fecundity or growth compared to mussels found in areas with better food quality (Vanderploeg et al., 1996). Lastly, highly eutrophic lakes as well as oligotrophic lakes will have lower dreissenid body sizes and density compared to more moderate conditions (Lewandowski & Stanczykowska, 2014, Pollick, 2014).

ii. **Predators**

In the Great Lakes, round gobies appear to be the primary predator for adult mussels and alewife appear to be the main predator for dreissenid veligers (Lewandowski & Stanczykowska, 2014). Small and mid-sized mussels are significantly more susceptible to predation compared to larger mussels (Ghedotti et al., 1995). In general, predators focus on mussels 8-17 mm in size, however, this also varies based on the predator species (Czarnoleski & Muller, 2014). For example, the mussel size range for round gobies is 5-14 mm, with increases in the ability to consume larger dreissenids as round goby length increases (Houghton & Janssen, 2014). To feed on mussels, predators need to be able to crush mussel shells and the ability to do this requires specialized teeth and appropriately sized jaws. The round goby has pharyngeal teeth that can easily crush dreissenids, but there are still size limitations based on the size of the fish (Houghton & Janssen, 2014). The larger the mussel, the larger the fish needed to crush the shell, which is why these larger mussels tend to be preyed upon less due to the generally small size of round gobies. Smaller mussels, despite having greater resistance to being crushed, can instead be swallowed whole by predators, including those without crushing mouth parts (Houghton & Janssen, 2014).

In areas of strong predation pressure, dreissenids experience tradeoffs between diverting energy to predator resistance and growth. When danger is detected, mussels have been found to stop moving, stay near the bottom, and close their shell valves so they are less likely to be detected by predators (Kobak, 2014); however, this does not appear to hold true for larger dreissenids, who are less vulnerable to predators (Czarnoleski & Muller, 2014; Houghton & Janssen, 2014). Therefore, there is a balance between occupying sheltered habitats to hide from predators and occupying more exposed habitats, to have increased feeding opportunities (Houghton & Janssen, 2014). Dreissenids have been found clumped closer together in areas with high predator abundance and areas with "danger cues", such as crushed conspecifics (Kobak, 2014). When new mussels arrive in an area with crushed

conspecifics, they will attach less strongly and reduce their movements (Kobak, 2014). Attaching less strongly seems counterintuitive since stronger attachment makes it more difficult for individual mussels to be predated on, however, the process of creating and attaching byssal threads makes mussels more vulnerable and uses more energy, resulting in this necessary tradeoff (Kobak, 2014). The actual signaling pathway for how mussels detect crushed conspecifics is still unknown, however, it is likely a sort of kairomone that the fish predator exudes (Kobak, 2014). If this is the case, dreissenid mussels could have a different response based on different fish predators and if the predator kairomone is unknown it could be deleterious to mussel populations since they would not have an appropriate response to the eminent danger.

e. Potential adaptations/Physiological responses

The availability of food, food quality, competitors, predators, and pathogens all affect mussel physiological responses (Garton et al., 2014). To survive, dreissenids are sometimes forced to adapt to poor conditions. Below are a few examples of how dreissenids have been observed to adapt.

i. Size

Smaller mussels have been found to occur higher up on raised substrates than larger mussels, which are generally found right on the bottom (Kobak, 2014, Houghton & Janssen, 2014). There are multiple reasons why this occurs, including food availability, predation, and competition. These reasons could also be combined, such as smaller mussels could be higher up because there is more food present due to not having to compete with the larger mussels. It is also possible that mussels higher up do not live long enough to achieve larger sizes because they are less protected from predators, currents, and ice scour (Chase & Bailey, 1999).

ii. Food consumption

Dreissenids have been found to physiologically adapt to poor environmental conditions, which could confound distribution predictions since they may be in higher abundance than expected. For example, siphon length is known to be affected by several factors, including food availability, current velocities, temperature, and substrate type

14

(Nalepa et al., 2014). Dreissenid mussels in areas with soft substrates, low water velocities, and cooler temperatures have been found to have longer siphons (Nalepa et al., 2014). Longer siphon length could also be driven by food competition if the population density and biomass are higher than can be sustained by an ecosystem. Mussel grazing rate has also been found to differ with filtration rate, temperature, individual bivalve size, and water seston concentration (Schol et al., 2002; Schwalb et al., 2015; Shen et al., 2020).

iii.   Attachment strength

Attachment strength is negatively affected by many factors and is largely reduced in poor environmental conditions. Areas with hypoxia, increased light availability, effluent from crushed conspecifics, food depletion, and toxics, such as chlorine or copper, result in decreased byssal production and therefore decreased attachment strength for local dreissenids (Kobak, 2014). Areas with higher temperatures, currents for mixing and delivering food, healthy conspecifics, and predator kairomones all stimulate adhesion of dreissenids (Kobak, 2014). As mentioned above, there is an ideal range for currents, and shallower waters tend to have stronger currents and waves that can be detrimental to dreissenids since byssal threads may not be strong enough to resist washing away by these currents. Although mussels produce more byssal threads at higher water velocities, byssal thread synthesis is suppressed in areas with velocities above 180 cm/s (Hasler et al., 2019; Clarke & McMahon, 1996; Peyer et al., 2009).   Zebra mussels also have higher rates of byssal thread production and greater attachment strength relative to quagga mussels (Lucy et al., 2014; Marsden et al., 2014).

f. Notable differences between zebra and quagga mussels

Some differences between zebra and quagga mussels were mentioned in the sections above, but a few additional points are as follows. Quagga mussels devote less energy to respiration so they can focus their energy on assimilation and growth and therefore have a greater energetic efficiency compared to zebra mussels (Pilcher et al., 2015; Karatayev et al., 2014; Stoeckmann, 2003). Quagga mussels have a higher potential to survive during adverse conditions compared to zebra mussels due to their greater size, faster growth rates, and lower energetic costs (Garton et al., 2014). In addition, it is likely that quagga mussels also live

longer than zebra mussels which is inferred by their delayed reproduction, growth to a larger size, and enhanced survival (Garton et al., 2014).

It is important to note these differences since, although zebra mussels are one of the most studied freshwater invertebrates, much less is known about quagga mussels. In fact, since 1989 only 13% of all published papers relating to *Dreissena* focus on quagga mussels (Karatayev et al., 2014). Therefore, we know far less about quagga mussels than zebra mussels and until more research is conducted on quagga mussels individually, it is important to keep in mind key differences between these species.

C. Effects of **Dreissena** *on the lake ecosystem*

Dreissenids have had both positive and negative effects on the lake-wide ecosystem. This section will outline some of these effects to lay a foundation for understanding the importance of this research and to guide future studies. It is also important to keep these effects in mind when looking at dreissenids and the conditions they are found in within the lake.

Since the invasion of *Dreissena* spp. into the Great Lakes, there have been widespread changes to lake ecosystems. The direction and magnitude of the effects of dreissenids vary across different substrate types (Ward & Ricciardi, 2014) and the rate that mussels graze on phytoplankton and detritus is the product of individual mussel clearance rates, population density, and local availability of plankton and detritus (Shen et al., 2002).

a. Effects on nutrient cycling

Dreissenids have proven to be important for energy flow and nutrient cycling (Bootsma et al., 2015; Karatayev et al., 2018b). About 90% of the food taken in by dreissenid mussels is excreted, which has a large impact on the nutrient content in the ecosystem, especially nitrogen concentrations (Bril et al., 2017). Dreissenids also cause a consistent decline in pelagic chlorophyll (Kumar et al., 2016) and phosphorus (Marsden et al., 2014) and decrease the chlorophyll to total phosphorus ratio (Johengen et al., 2014). Dreissenids also contribute to eutrophication (Piacenza et al., 2015) by excreting bioavailable

phosphorus and through a direct transfer of dissolved phosphorus between the mussels and the benthic filamentous green algae *Cladophora* (Bootsma et al., 2015). In addition, dreissenids may facilitate phytoplankton growth and cyanobacterial blooms by modifying nitrogen and phosphorus forms to shorten turnover time (Boegman et al., 2008). The indirect effects of mussels, such as nutrient excretion or egestion, can stimulate phytoplankton growth through increased vertical mixing (Schwalb et al., 2015). Dreissenid filtration also causes increased water clarity, resulting in increased depth of light penetration which also has a positive effect on *Cladophora* growth rates (Bootsma et al., 2015; Marsden et al., 2014; Ward & Ricciardi, 2014).

Mussel excretion is an important benthic nutrient source (Shen et al., 2018). Dreissenid mussels affect nutrient cycling through nitrogen and phosphorus excretion (Johengen et al., 2014) and altering nutrient dynamics in general, resulting in elevated dissolved nutrient concentrations in upwelled water (Piacenza et al., 2015; Pothoven & Fahnenstiel, 2014). Mussels excrete bioavailable phosphorus when phytoplankton supply exceeds the needs for tissue growth (Schwalb et al., 2015) and mediate the transfer of phosphorus to the benthos through biodeposition of particulate matter (Schwalb et al., 2015). Strong filtration by mussels has decoupled lake production from phosphorus input (Jiang et al., 2015) and there have been reductions in dissolved phosphorus and phytoplankton biomass in the mid-depth lake area, however, phytoplankton biomass actually slightly increased in the epilimnion (Shen et al., 2020). Dreissenids also reduce the suspended solids concentration (Higgins, 2014) and because of increased mussel calcification, there has been a reduction in total alkalinity (Pilcher et al., 2015).

b. Effects on the lake bottom

Dreissenid mussels are known to accelerate the sedimentation process (Bootsma & Liao, 2014) and the rate of particle transfer from the water column to the lake bottom, which causes a reduction in the water column particle residence time (Shen et al., 2018). Furthermore, dreissenid mussels alter the vertical structure of benthic bed water flows when water velocity is weak (Troy et al., 2016).

Dreissenid mussels alter the entire lake bottom ecosystem by changing the available habitat. Presence of dreissenids can increase the resistance of silt and sand areas to disturbance from currents (Griffiths et al., 2014). In addition, interstitial spaces between mussel shells are used by smaller invertebrates to hide from environmental stressors, like currents, waves, and predators (Ward & Ricciardi, 2014). Dreissenids can also result in increased macroinvertebrate abundance on rocky substrates, potentially due to increased food availability or changes in habitat (Ward & Ricciardi, 2014). Nutrient rich particles that are egested by mussels may be stored in the sediment, further altering the nutrient cycle (Shen et al., 2018). A near-bed internal boundary layer is created by mussel filtration that may alter turbulence and mean flow characteristics (Troy et al., 2016). By altering the entire lake bottom substrate structure, dreissenids provide attachment substrate for invertebrates and benthic algae, like *Cladophora* (Mayer et al., 2014; Bootsma et al., 2015; Karatayev et al., 2018b), which were previous uninhabited by these organisms.

c. Effects on the food web

Dreissenids can alter the entire food web in addition to the impacts they have on more general nutrient cycling. Dreissenids affect both the water column food web (Pothoven & Fahnenstiel, 2014) and the water quality (Chakraborti et al., 2014). *Dreissena* spp. are capable of straining nearly all particulate matter from the water, given high enough densities, with the unwanted particles rejected and put back into the water (Pollick, 2014). The filtering of seston from the water column also leads to increased water clarity which prompts aquatic plants to grow and causes further community composition changes (Griffiths et al., 2014; Boegman et al., 2008; Kumar et al., 2016; Vanderploeg et al., 2015; Marsden et al., 2014). Dreissenids actively reject some toxic species of algae, like *Microcystis* spp., through selective filter feeding and thus contribute to harmful algal blooms (Chakraborti et al., 2014; Bril et al., 2017; Marsden et al., 2014). In addition, dreissenids remove phytoplankton and zooplankton that other organisms need to survive, resulting in declines in pelagic fish populations (Shen et al., 2020). Dreissenids can also alter the vertical migration of zooplankton and force them into unfavorable habitats (Vanderploeg et al., 2015), which in turn may further exacerbate declines to lake zooplankton populations.

There have been significant changes to fish behavior, growth rates, and reproduction following mussel invasion (Higgins, 2014). First, dreissenid filtration of the water column results in an increase in benthic and sight-feeding fish (Griffiths et al., 2014) as well as increases in benthic invertebrates as a result of increased benthic algae (Mayer et al., 2014) and deposition of wastes (Minchin & Zaiko, 2014), increasing the structural complexity of the benthos (DeVanna et al., 2014). Dreissenids create biodeposits which serve as a food source for other organisms, and can cause increased oligochaete densities (Nalepa et al., 2018). As a result of the nourishment from mussel biodeposits, dreissenids can cause an increase in the density and biomass of nearshore communities as well (Ward & Ricciardi, 2014).

Dreissenids tend to cause a decline in phytoplankton, specifically chlorophytes (Bootsma & Liao, 2014; Pothoven & Fahnenstiel, 2014), crustaceans, cyanophytes, rotifers, zooplankton (Alderstein et al., 2014), *Diporeia* (Nalepa et al., 2018; Marsden et al., 2014; Rennie, 2014), primary and secondary producers (Kumar et al., 2016), major fish species (i.e., lake whitefish) (Kumar et al., 2016; Rennie, 2014), and native bivalves and unionids (Ward & Ricciardi, 2014; Marsden et al., 2014; Strayer & Malcom, 2014). Dreissenids have also been documented to alter walleye depth distributions due to water clarity (Marsden et al., 2014) and deplete algae and plankton needed by perch for survival (Shen et al., 2018). Dreissenids have caused *Diporeia* to virtually disappear, which has resulted in declines to trout, whitefish, sculpin, and smelt (Marsden et al., 2014; Pollick, 2014; Bunnel et al., 2014). Due to the declines in prey, the diets of these lake fish must adjust as a result, causing a shift from pelagic primary production to nearshore, benthic primary production (Bootsma et al., 2015). In general, dreissenid mussels have caused declines in profundal benthic species abundance and diversity (Griffiths et al., 2014) and lake pelagic productivity (Marsden et al., 2014; Kumar et al., 2016).

Marsden et al. (2014) found that epilimnetic plankton does not decline with increasing mussel populations, however, overall literature supports that dreissenids reduce phytoplankton biomass (Boegman et al., 2008; Karatayev et al., 2018b; Kumar et al., 2016; Piacenza et al., 2015; Pilcher et al., 2015). Rowe et al. (2017) found that quagga mussel clearance rates are sufficient to exceed phytoplankton growth rates. Reductions in

phytoplankton negatively affect zooplankton and lead to further fish declines (Kumar et al., 2016). In addition to phytoplankton, dreissenids also cause declines in zooplankton (Griffiths et al., 2014) and these impacts are the most pronounced in microzooplankton (e.g., rotifers) (Marsden et al., 2014).

In contrast, dreissenids have been known to increase benthic prey consumption by fish (Mayer et al., 2014), increase the bioavailability of phosphorus, *Cladophora* (Marsden et al., 2014), cyanobacteria concentrations (Bootsma & Liao, 2014), dissolved silica (Higgins, 2014), water clarity (Bootsma & Liao, 2014), and macrophyte abundance (Bootsma & Liao, 2014). Dreissenids also have positive effects on densities of gammarid amphipods, leeches, flatworms, and certain insects (Ward & Ricciardi, 2014), oligochaetes (especially in areas less than 50m depth) (Nalepa et al., 2018), and littoral benthic species (Griffiths et al., 2014).

d. Other effects

Dreissenids can also cause economic and health costs to humans. For example, the high density of zebra and quagga mussels can clog water intakes and other pipes (Burrows, 2014). Dreissenids can also take up and retain parasites and pathogens that are harmful to other animals. They accumulate these during their filter-feeding, and while the parasites do not infect the mussels, they can be harbored in the tissues and spread to other organisms that consume them (Conn et al., 2014). For example, type E botulism, caused by *Clostridium botulinum* type E, affects aquatic vertebrates and waterfowl and has been found to spread through benthic organisms, including *Dreissen*a, due to their ability to carry spores and spread them to upper trophic levels through the food web (Perez-Fuentetaja et al., 2011). In addition to parasites and pathogens, mussels can also concentrate contaminants, such as PCBs (Pollick, 2014), which can then bioaccumulate through the food web into larger fish.

D. Recent Models

Given all the effects of dreissenids on the benthos, lake-wide ecosystem models that predict *Dreissena* spp. presence and biomass are important management tools. Despite this importance, few studies have attempted to quantity dreissenids, especially on a large scale.

Two examples of dreissenid models from recent years are by Rowe et al. (2015b) and Karatayev et al. (2021b).

Rowe et al. (2015b) created a geostatistical model to map *Dreissena* across four different years (1994/95, 2000, 2005, and 2010) using point observations from Ponar grabs across Lake Michigan. They predicted dreissenid biomass in grams of ash-free dry weight (AFDW) for both *Dreissena* species across space using bathymetry and geographic-coordinate covariates. The results showed that *Dreissena* biomass was highest at an intermediate-depth range. These results were presented as lake-wide distribution and standard deviation maps for each year to account for error in the spatial model. These errors were greater in areas with higher measured *Dreissena* biomass and lesser in areas with lower *Dreissena* biomass. Despite these errors, Rowe et al. (2015b) did show their model could predict *Dreissena* biomass in areas that were not sampled, including areas with hard substrates. The focus of this study was not to predict *Dreissena* biomass as a function of abiotic or biotic conditions, but to use lake-wide distributions to focus on the impacts dreissenids may have had on phytoplankton abundances over time.

Karatayev et al. (2021b) summarized dreissenid survey data across the Great Lakes over a span of almost 30 years (1990-2018). In order to visualize the survey data through space and time, Karatayev et al. (2021b) applied generalized additive models (GAMs). Through these GAMs, Karatayev et al. (2021b) found the 30-90 m depth zone to have the highest observed dreissenid biomass across all the Great Lakes, but the 71-90 m depth zone to have the highest biomass in Lake Huron. They hypothesized this could be due to decreasing respiration rates with lower temperature and therefore a lower food intake required to maintain biomass in these depth zones compared to shallower depths (Stoeckmann, 2003; Karatayev et al., 2021b). In addition, the shift from zebra to quagga mussels in these lakes resulted in greater biomass in deeper areas, due to the preference of quagga mussels for deeper waters and the authors hypothesized that the dreissenid populations will continue to move towards even deeper waters (Karatayev et al., 2021b). The authors also looked at other trends, for example, finding that on soft substrates, quagga mussel density was always higher than zebra mussel density and the percentage of zebra mussels increased with the availability of hard substrates. Although Karatayev et al. (2021b)

were able to put together datasets from almost 30 years, they were still limited to available long-term dreissenid sample sites that did not always adequately cover the wide variety of habitats present and did not always use the same sampling techniques.

E. Research Objective

The overarching objective of this thesis is to examine hypotheses about potential landscape-scale controls on dreissenid mussel presence and biomass using predictive modeling. Given the far-reaching effects of dreissenid mussels on ecosystems (Introduction Section C), having the ability to predict dreissenid presence and biomass on a lake-wide scale could inform management and science in important ways. Although it is important to quantify and predict *Dreissena* spp. extent in all the Great Lakes, this research focuses on Lake Huron. The benthic community of Lake Huron is considered to be the least studied of all the Great Lakes (Nalepa et al., 2007), justifying further focus on this lake specifically.

F. Hypotheses

Since the main objective focuses on testing hypotheses about landscape controls on dreissenid mussels, the published literature on factors that limit large scale dreissenid distributions was distilled into a set of potential limiting factors for dreissenid presence and abundance. These factors were further matched to available large-scale spatial datasets resulting in the hypotheses represented in Figure 1. Each hypothesis is summarized below.
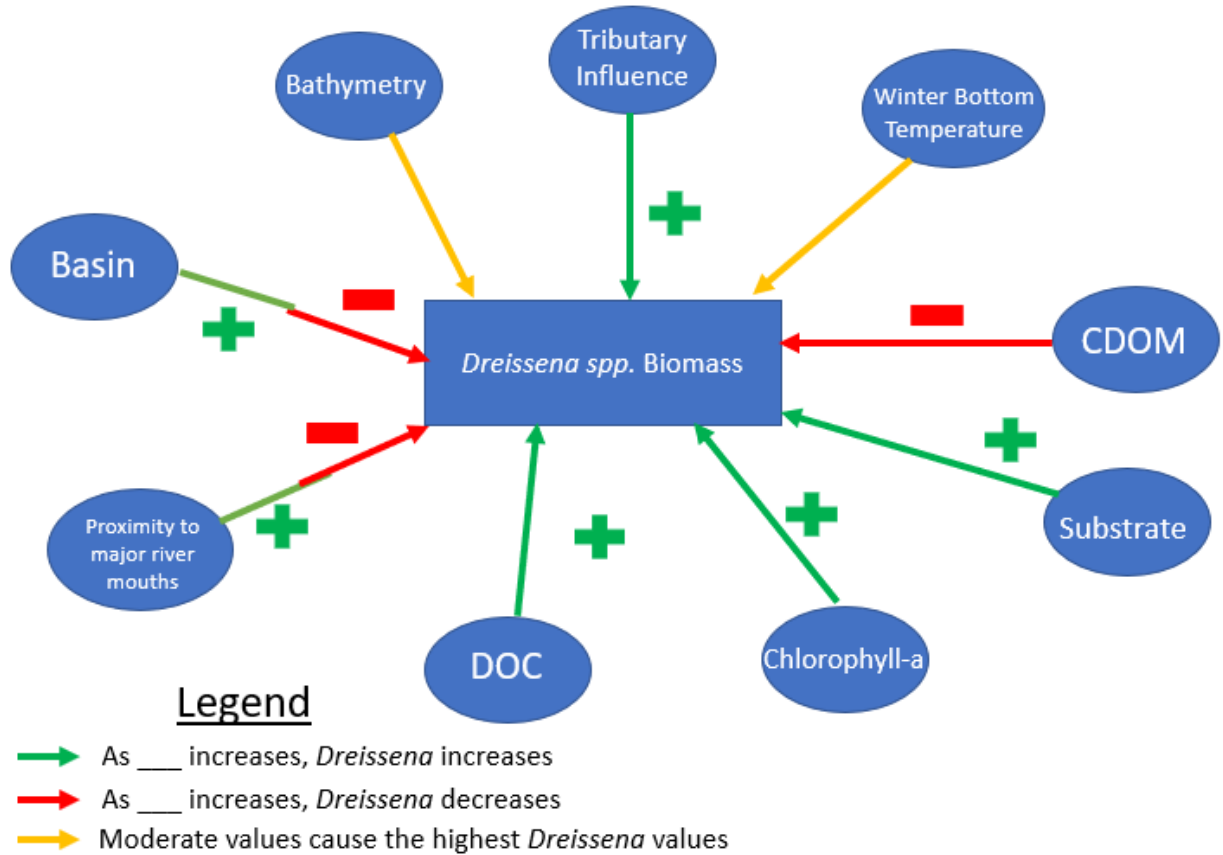
Figure 1: Conceptual Model showing the ten explanatory variables explored in this paper. Green arrows with plus sign syndicate positive effects, red arrows with negative sign indicators negative effects, and the yellow lines indicate positive effects at intermediate values of each variable. Basin and proximity to major river mouths are represented by a combined green and red arrow with both plus and negative sign indicators to show the positive and negative effects of different basins and rivers.

1) **Basin**: Basin was considered as a categorical variable summarizing four geographically distinct parts of Lake Huron with different water chemistry and habitat characteristics that may influence dreissenid mussels.

   a. Saginaw Bay is a shallow embayment of the lake with a large annual influx of nutrients from the Saginaw River and smaller tributary streams from agricultural activity upstream. I hypothesize that Saginaw Bay will have moderate dreissenid biomass because the positive effects of increased food availability (from the Saginaw River) will be offset by low winter water

temperatures, which tend to lower dreissenid biomass. Shallow areas of lakes also tend to have higher wind-driven current velocities which will result in lower dreissenid densities, especially on soft substrates to which they cannot attach well (Garton et al., 2014). This variable is similar to that of the Saginaw River influence variable, however, it encompasses differences in the Saginaw Bay area that may not be a result of the Saginaw River inflow, such as bathymetry and temperature differences from the rest of Lake Huron.

b. The influence of Lake Superior water delivered from the St. Mary's River to the North Channel is hypothesized to have low dreissenid biomass due to the influence of Lake Superior water quality (Barbiero et al., 2018). Lake Superior is inhospitable to dreissenid mussels because its low calcium concentrations inhibit shell formation. In waters with low Ca^{2+}, mussels are unable to make high quality shells and thus will feel stronger pressure from predators, either resulting in death or vacating the area (Czarnolesk & Muller, 2014; Marsden et al., 2014).

c. I hypothesize moderate dreissenid biomass across the Georgian Bay due to a combination of two factors. The watersheds along the eastern shore of Georgian Bay drain off the Canadian Shield basement rock leading to low calcium, tannin rich waters and rocky substrates. Low calcium concentrations will result in lower dreissenid biomass; however, rocky substrates may provide good habitat otherwise, resulting in moderate dreissenid biomass overall.

d. The main basin is large and heterogeneous, making it difficult to make generalized predictions. Nonetheless, I expect moderate to high dreissenid biomass overall due to the wide variety of habitats and substrates present, moderate to high calcium concentrations, deeper waters (especially important for quagga mussel success), and diversity of feeding conditions. The northern part of the main basin influenced by the St. Mary's River is hypothesized to have lower dreissenid biomass due to the influence of calcium-poor Lake Superior water.

2) **Bathymetry:** I hypothesize that dreissenid mussel biomass will be higher in areas of moderate depth, between 30-90 m. Karatayev et al. (2021b) found that this depth range has the highest observed dreissenid biomass values due to the balance between higher turbulence in shallow waters < 30m (disturbing the substrate and lowering attachment ability) (Nalepa et al., 2014), and poorer substrate quality and lower food availability (primarily phytoplankton) in deeper waters > 90 m (Chakraborti et al., 2014).

3) **Tributary influence:** I hypothesize that dreissenid mussel biomass will be highest in areas with higher tributary influence compared to areas of lower tributary influence due to the influx of nutrients from these tributaries. Nearshore waters, here defined as 0-30 m (Riseng et al. 2018), are known to often have higher nutrient and plankton concentrations, benthic biomass, and biological productivity than offshore waters (Bocaniov et al., 2013).

4) **Winter bottom temperature:** I hypothesize that higher winter maximum bottom temperatures, greater than 4°C, will result in increases in dreissenid biomass, due to the decreased chance of winter mortality due to freezing temperatures or ice scour. Winter bottom temperature is related to bathymetry; shallower areas of Lake Huron freeze in the winter, making these areas less hospitable to dreissenids (Thorp et al., 1998).

5) **Substrate:** Dreissenid mussels are known to have lower abundances on soft substrates than hard substrates, with zebra mussels having even lower biomass on soft substrates compared to quagga mussels (Kumar et al., 2016, Pilcher et al., 2017). However, the dreissenid sample sites used in this research are only on soft substrates due to sampling methods described below. Given this, I hypothesize that areas with sand will have higher dreissenid biomass than areas with silt or mud because of a higher abundance of attachment points with increasing grain size (Garton et al., 2014).

6) **Colored Dissolved Organic Matter (CDOM):** I hypothesize dreissenid biomass to be low in areas of higher CDOM concentrations due to decreased light availability

resulting in decreased phytoplankton concentrations (Tzortziou et al., 2015). Phytoplankton is the main source of food for dreissenid mussels and so any decrease in their abundances will negatively impact the biomass of dreissenids.

7) **Food availability:** I hypothesize that dreissenid biomass will increase with increased food availability. Food availability is commonly measured using chlorophyll-a concentration as a proxy for phytoplankton (Bold & Wynne, 1985; Bennion et al., 2019). Food availability is important since the success of the population relies on the ability of individual mussels to reproduce, spawn, and grow (Claxton & Mackie, 1998; Wacker & von Elert, 2003).

8) **Dissolved Organic Carbon (DOC):** I hypothesize increased dissolved organic carbon (DOC) will result in an increase in dreissenid mussels. Dissolved organic carbon is necessary for dreissenid biological functions and the success of other species dreissenids consider as prey, especially in areas where phytoplankton biomass is depleted as DOC serves as a nutritional supplement to dreissenids (Roditi et al., 2000).

9) **Proximity to major river mouths:** Two major rivers flow into Lake Huron—the St. Mary's River and the Saginaw River. The St. Mary's River connects Lake Superior and Lake Huron with water flowing from Lake Superior into Lake Huron's North Channel. I hypothesize that increased distance from the St. Mary's River mouth will result in higher dreissenid biomass, since the water chemistry in Lake Superior is inhospitable to dreissenid mussels due to low calcium concentrations (Grigorovich et al., 2003). The Saginaw River represents a large inflow of nutrients from a largely agricultural watershed that should contribute to the food availability for dreissenids in the Saginaw Bay and southern half of the main basin (Stow et al., 2014). Therefore, I hypothesize increased distance from the Saginaw River will result in decreasing dreissenid biomass with rapidly diminishing effects as the river water mixes with the low nutrient water of the main basin.

II.   Methods

A. Dreissenid sample data

The dreissenid sample site data used for this research is available through the Environmental Protection Agency (EPA) Great Lakes National Program Office (GLNPO) and was produced through efforts by EPA, Buffalo State College, and the National Oceanic Atmospheric Administration (NOAA) from surveys conducted during the summer of 2017 as a part of the Cooperative Science and Monitoring Initiative (CSMI). This survey is completed every 5 years and assesses dreissenid density and biomass across all of Lake Huron, alternating to the other lakes in the years in between. At each of the 119 sample sites spread throughout all four basins of Lake Huron, triplicate samples were taken with a 9-inch by 9-inch Ponar grab. Each sample was sorted by zebra mussels and quagga mussels in the lab, and samples summarized for numerical density, total wet weight (TWW), and ash-free dry weight (AFDW). Further information regarding the sampling methods and descriptions of these data can be found in a technical report by Karatayev et al. (2020).

Although the recorded values were split between zebra and quagga mussels, there were not enough non-zero values for zebra mussels to support model testing. Therefore, individual species trends could not be observed for zebra mussels, only quagga mussels. To have the most sample sites with non-zero values, zebra and quagga mussel data were summed to generate genus level models for both *Dreissena* species together. Although both TWW and AFDW were summarized, AFDW (grams carbon per meter squared) was used as the primary response variable due to its prevalence in literature as the least biased measure of dreissenid abundance (Karatayev et al. 2021b). Dreissenid numerical densities were used to create a binary response variable (presence versus absence), by assigning non-zero abundance values with a 1 and zero abundance as a 0. This was done both for quagga mussels alone and for the summed both *Dreissena* species variable.

B. Explanatory Variable Selection

Dreissenid mussels, like all organisms, depend on a wide variety of factors to live and grow. Candidate explanatory variables (Introduction Section B) were identified based on

their prevalence in the literature and only factors that were mentioned in multiple studies were considered for modeling. Lake-wide datasets were derived from the Great Lakes Aquatic Habitat Framework (GLAHF, Wang et al., 2015) and the National Oceanic Atmospheric Administration (NOAA) CoastWatch Program (Leshkevich et al. 1993). For data collected in multiple years, only those variables from the sampling year of the mussels (2017) were considered to ensure that mussels alive during the year of sampling experienced the conditions (i.e., temperature) during their lifespan. A total of 142 geospatial data layers were located and considered for inclusion in the model (Appendix 1).

Due to the high degree of redundancy among explanatory variables, a variable reduction process was undertaken. Initially, all variables that were not linearly correlated with the response variables ($r > 0.2$) were removed. I used Pearson correlation coefficients to examine explanatory variable pairs and identified collinear variables with $r^2 > 0.7$. One variable was selected for each explanatory category (i.e., only one temperature variable for each month) based on the published literature to ensure that plausible mechanistic relationships with dreissenid abundance were supported by prior research.

After variable reduction, ten explanatory variables remained: basin (categorical), bathymetry, tributary influence (categorical), February maximum bottom temperature, substrate (categorical), April colored dissolved organic matter (CDOM), July chlorophyll-a concentration, October dissolved organic carbon (DOC), distance from the St. Mary's River, and distance from the Saginaw River.

All grids and output maps were prepared using a combination of ESRI ArcGIS Pro (v. 2.7) and ESRI ArcMap (v. 10.8) and standardized to have the same spatial resolution (30m pixel size). All grids were resampled and put into the USA Contiguous Albers Equal Area Conic USGS version projection using ArcGIS Pro. Dreissenid abundance point localities were intersected with raster layers for all explanatory variables to generate the data matrix used in modeling.

a. Basin

The basin variable is a categorical variable that defined the four basins of Lake Huron with numerical values 1-4. These basins were Main basin, Georgian Bay, North Channel, and Saginaw Bay with values 1-4 in that order (Figure 2a). The majority of sample sites were located in the main basin with fewer sites in the other three basins (Main: 79, Georgian Bay: 11, North Channel: 16, Saginaw Bay: 12), however sampling density was highest in Saginaw Bay (Main: 26%, Georgian Bay: 8%, North Channel: 21%, Saginaw Bay: 45%).

b. Bathymetry

Bathymetry is represented as a continuous raster layer created by GLAHF using data from the NOAA National Centers for Environmental Information. This raster can be downloaded from the GLAHF website (https://www.glahf.org/data/; Wang et al., 2015). Water depths in Lake Huron ranged from 0-244 m (Figure 2b).

c. Tributary influence

Tributary influence is a categorical variable derived from the Great Lakes Hydrology Dataset (GLHD, Forsyth et al., 2016, Wang et al. 2015). These classes were defined by the values 1-3 in the nearshore zones and the value of 0 was given to offshore zones greater than 30 m depth (areas of no tributary influence). The numbers 1-3 correspond to low, moderate, and high tributary influence based on the catchment area; low is less than 30 km^2, moderate is 30-250 km^2, and high is greater than 250 km^2. A distance decay function weighted by depth was used to propagate the flow into the lake from each river mouth. The majority of sample sites were in areas of no tributary influence due to the large proportion of lake area beyond 30 m depth, typically considered to be offshore with minimal tributary influence (Figure 2c; Riseng et al. 2018).

d. February Maximum Bottom Temperature

February maximum bottom temperature was calculated from bottom temperature data from the Princeton Ocean Model (POM). These data are free to access from the National Oceanographic Atmospheric Administration (NOAA) CoastWatch Program Environmental Research Division Data Access Program (ERDDAP) server (https://coastwatch.glerl.noaa.gov/erddap/index.html; Leshkevich et al. 1993) and contain

bottom temperature values estimated eight times a day every day. This variable was calculated by finding the maximum value over all the time stamps across all of February for each pixel (30m spatial resolution) across the entire lake. February was selected to represent winter bottom temperatures due to higher correlation with dreissenid biomass in data exploration than other winter months. Higher February maximum bottom temperatures occurred in the deeper portions of the lake in 2017 (Figure 2d) and the coldest temperatures occurred in the nearshore, shallower zones.
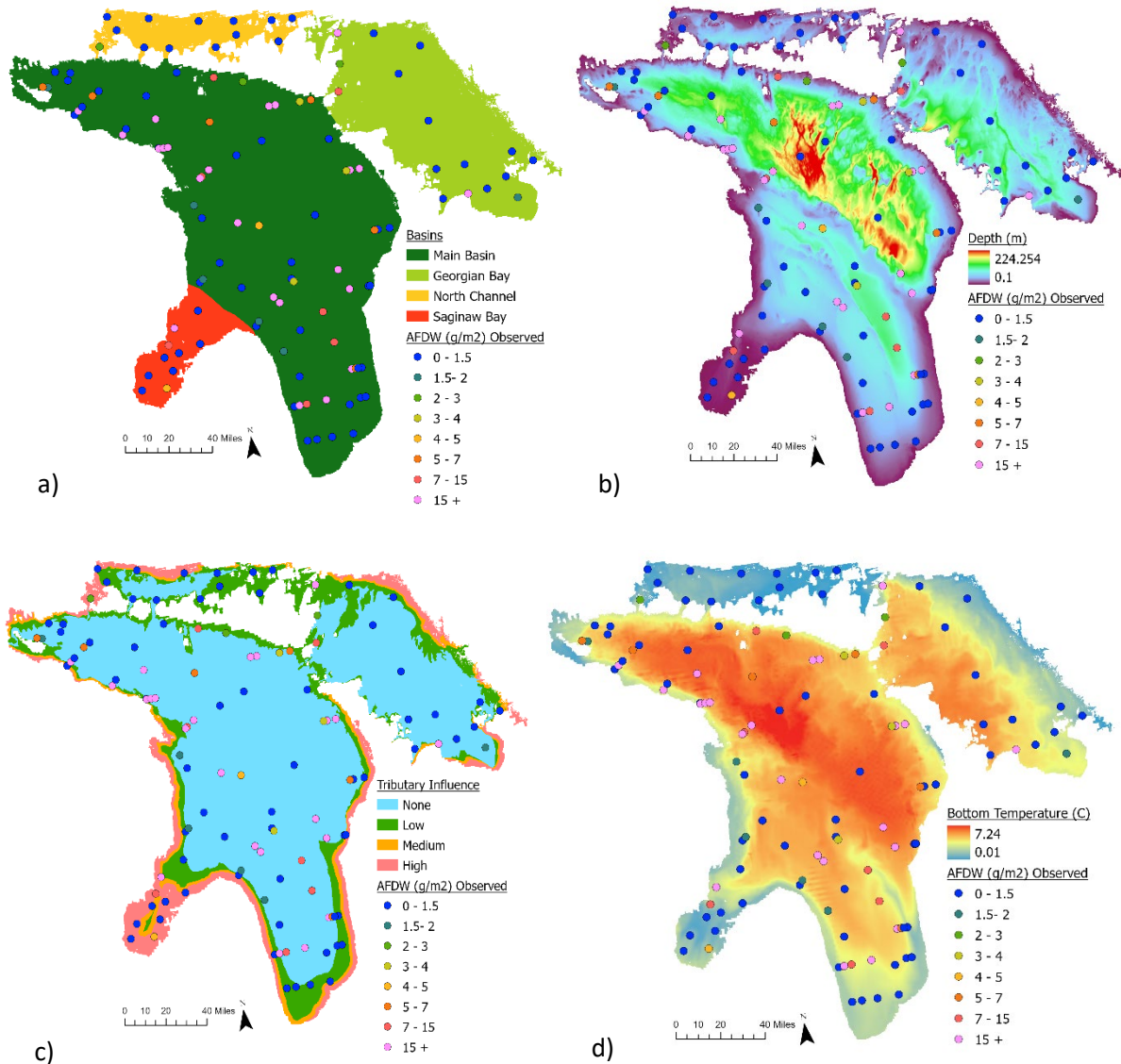
Figure 2: Predictor variables overlaid with values of Dreissena *spp. biomass (ash-free dry weight; AFDW)*
(points): (a) basin, (b) bathymetry, (c) tributary influence, and (d) February maximum bottom temperature.

e. Substrate

Substrate was the only predictor variable lacking a complete geospatial layer across
the lake, because it was drawn from the substrates sampled during Ponar surveys used to
collect mussels. Substrate samples were sorted into five categories: silt, sand, sandy
clay/sandy silt, silty clay/silty loam, and silty sand, which were assigned numbers from 1-5,
respectively, going from finer to coarser.

f. April Colored Dissolved Organic Matter (CDOM)

The CDOM dataset was downloaded from the NOAA ERDDAP server and was created using Moderate Resolution Imaging Spectroradiometer (MODIS) satellite imagery. The variable used in modeling was created by averaging the CDOM values in each 30-m pixel for all the timestamps available for April 2017. April was selected to represent CDOM due to higher correlation with dreissenid biomass in data exploration steps than other months. Higher April CDOM concentrations were observed in Saginaw Bay and along the eastern portion of the North Channel (Figure 3a).

g. July Chlorophyll-a

July chlorophyll-a data were downloaded from the NOAA ERDDAP server and were created using MODIS satellite imagery. This variable was calculated by averaging the chlorophyll-a values at each pixel (30m resolution) for all timestamps in July. July was selected to represent chlorophyll-a due to higher correlation with dreissenid biomass in data exploration steps compared to other months. Chlorophyll-a concentrations in July 2017 were generally low across the lake, except in Saginaw Bay (Figure 3b).

h. October Dissolved Organic Carbon (DOC)

October DOC data were downloaded from the NOAA ERDDAP server and created using MODIS satellite imagery. This variable was calculated by averaging the DOC value at each pixel for all timestamps in October 2017. October was selected to represent DOC due to higher correlation with dreissenid biomass in data exploration steps than other months. Lake Huron had fairly low levels of DOC in October 2017 with the highest values in the Saginaw Bay area (Figure 3c). The eastern portion of the North Channel and southern portion of the Georgian Bay had intermediate DOC concentrations.

i. Influence from Lake Superior (St. Mary's River)

Influence from Lake Superior water chemistry was assumed to be related to the outlet of the St. Mary's River and was calculated by finding the Euclidean distance from the St.

32

Mary's River mouth to every pixel in Lake Huron using the ArcGIS Pro Spatial Analyst toolbox (Figure 3d).
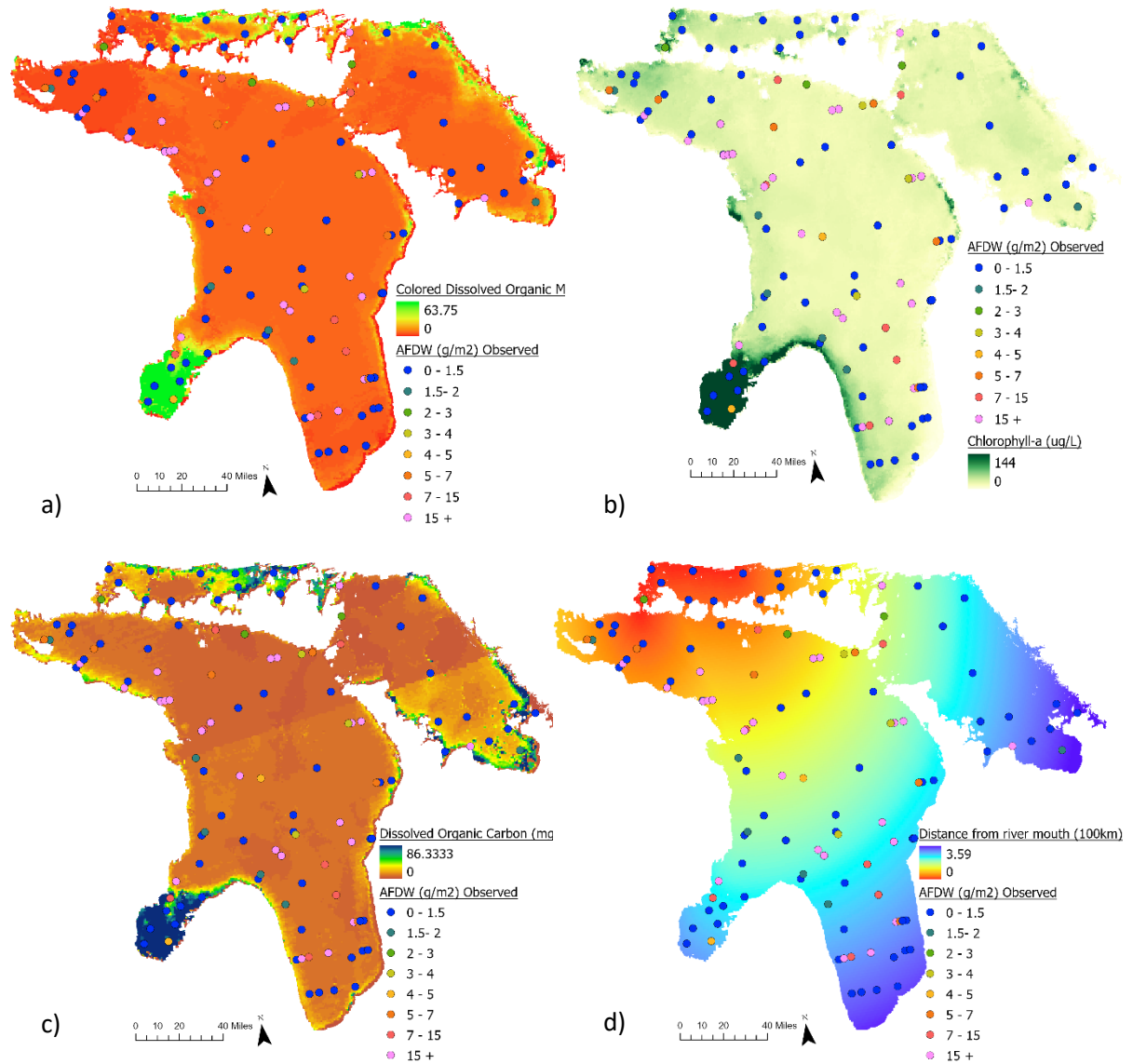

Figure 3: Predictor variables overlaid with values of Dreissena *spp. biomass (ash-free dry weight; AFDW) (points): (a) April CDOM, (b) July chlorophyll-a, (c) October DOC, and (d) Influence from Lake Superior.*

j. Influence of the Saginaw River

33

Influence of the Saginaw River was represented using Euclidean distance from the river mouth, calculated for every pixel in Lake Huron using the ArcGIS Pro Spatial Analyst toolbox. As with the distance to the St. Mary's River, this variable declined with distance (Figure 4).
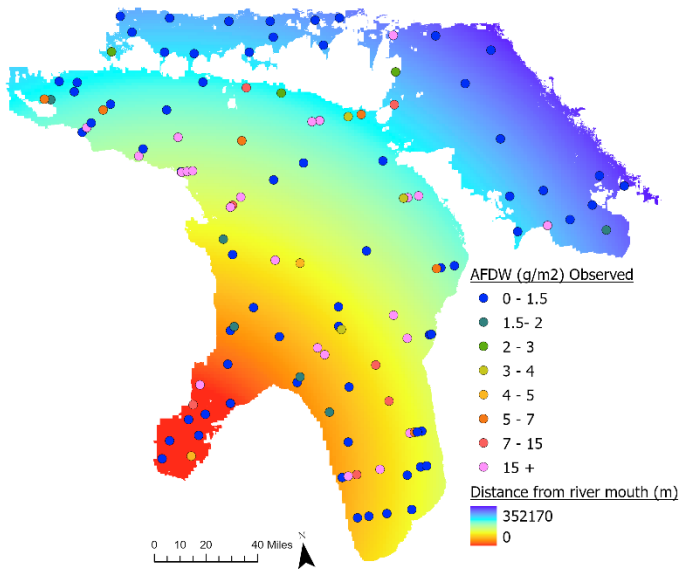
Figure 4: The influence of the Saginaw River variable overlaid with the observed values of Dreissena *spp. biomass (ash-free dry weight; AFDW).*

The influence of the Saginaw River and the influence from Lake Superior variables both overlap some with two other explanatory variables: tributary influence and basin. Both the Saginaw River and the St. Mary's River were included within the tributary influence variable. These two major rivers were isolated to focus on major water influx to Lake Huron that may be particularly influential due to water chemistry characteristics and could be partially obscured if only the tributary influence variable was used. Since the focus of separating these variables was to capture differences in these two basins, North Channel and Saginaw Bay, the basin variable also overlaps. However, the basin variable does not provide information on a pixel-by-pixel basis like that of the influence from Lake Superior and influence of the Saginaw River variables.

C. Modeling Methods

a. Empirical Bayesian Kriging

Kriging is a class of statistical techniques for optimal spatial prediction, originally developed for meteorological applications (Krivoruchko, 2012). Kriging is a probabilistic predictor and assumes a statistical model for the data, using a semivariogram to quantify the spatial dependence in the data and the standard errors quantify the uncertainty associated with the predicted values. A semivariogram is a function of the distance separating two locations used for measuring the degree of dissimilarity between observations, created by calculating half the average squared difference of all pairs of measurements and locations. When a semivariogram is plotted, the contribution of each observed data point to the prediction of new values at unsampled locations can be defined based on the nearby values.

Empirical Bayesian Kriging (EBK) differs from traditional kriging methods by accounting for error introduced while estimating the semivariogram model. Traditional kriging assumes the underlying random field is Gaussian and that the covariance function is known (Pilz & Spock, 2007), assumptions that are not met in practical applications. In EBK, many semivariograms are used rather than a single semivariogram. The EBK process estimates an initial semivariogram using the given data and then new values are simulated at each of the input data locations based on the semivariogram distribution. Then, using the simulated data, a new semivariogram model is created and given a weight, calculated using Bayes' rule; the probability of the observed data being generated from the semivariogram. In each of many repetitions a new semivariogram is created to estimate the values at the input locations. The result is a spectrum of semivariograms, each estimating the true semivariogram. These can all then be plotted, along with their weights, and used to predict the true values at the sample sites. Using a distribution of semivariogram models, instead of just one as in classical kriging models, tends to reduce errors due to the increased probability of correct value prediction.

EBK models were created using the Geostatistical Analyst tool in ArcGIS Pro (v. 2.7) to map the variation in dreissenid biomass across Lake Huron and predict distributions in unsampled areas of the lake. Model parameters such as number of simulations and subset

size were tested to find the strongest performing model across all the outputs and create the best fit semivariogram. Documentation regarding the meanings of these different parameters and how to fit the best semivariogram can be obtained from the ESRI ArcGIS Pro resources website (https://pro.arcgis.com/en/pro-app/latest/help/analysis/geostatistical-analyst/what-is-empirical-bayesian-kriging-.htm). This documentation was used to help pick the desired parameters for transformation and semivariogram type. The other parameter values were assigned through trial and error, based on performance metrics. This process created a continuous map for dreissenid biomass, expressed as ash free dry weight (AFDW), using leave-one-out cross validation. Following this, predicted point values were obtained from the predictive map at each sample site and compared to the measured values to calculate further performance metrics, including r-squared, standard error, and root mean squared error (RMSE). The measured values were then overlaid on the predictive map to qualitatively examine the spatial correspondence between predicted and observed values.

As an approach that relies on the semivariogram to make spatial predictions of dreissenid biomass, EBK does not allow for examination of relative effects of environmental variables on dreissenid response. Nonetheless, similar approaches to EBK have been used to good effect in the past with *Dreissena* in the Great Lakes (Rowe et al. 2015b), and a lake-wide predictive map, if accurate, may still have a benefit to managers.

b. Boosted Regression Trees

Boosted regression tree (BRT) models (Elith et al., 2008; Friedman, 2001) were used to evaluate the effects environmental variables had on dreissenid distributions and abundance. Boosted regression trees are a machine-learning approach that combines many simple regression trees in a stagewise process. Each tree trains on the residuals of the tree that precedes it until predictive deviance is minimized (see Elith et al., 2008 for overview). In the case of binary presence–absence data, the measure of predictive deviance used is the log-odds ratio (Friedman, 2001). In the case of a continuous response variable, deviance is measured as mean square error between predicted and observed biomass. In BRT models, three parameters must be considered: learning rate, tree complexity, and bag fraction

(described further in Elith et al., 2008). These parameters were altered to find the best performing models. Boosted regression trees (BRT) were created using the "gbm" package (Ridgeway 2020) in RStudio (v. 1.3.1056 "Water Lily"). BRTs were used to create predictive models for dreissenid biomass and presence throughout Lake Huron using the explanatory variables described above. Boosted regression trees allow for the exploration of different combinations of explanatory variables to produce both lake-wide continuous predictions of dreissenids, and to learn about the importance of each variable to the best-fit model projections.

c. Dreissenid Biomass (AFDW) BRT models

Both biomass response variables—quagga and quagga + zebra combined (both)—had distributions that were skewed right due to the number of zeros in the dataset (Figure 5), so a natural log + 0.0001 transformation was used to improve normality of the dataset.
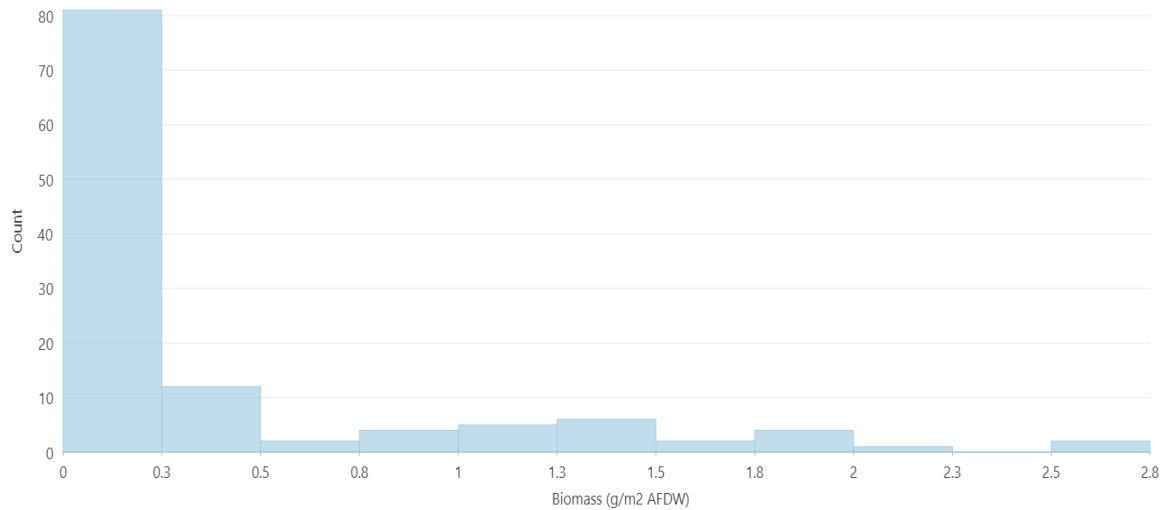
Figure 5: Distribution of biomass (ash-free dry weight; AFDW) sample values for both Dreissena species illustrating the large number of zero or near-zero values resulting in a right-skewed distribution.

An information theoretic approach was used to determine the most informative models for quagga mussels and both species biomass using all possible subsets of predictors. This resulted in 1,014 models with unique predictor subsets per group. Both models of dreissenid biomass were run in the R gbm library using the mean square error loss function, a tree depth

of 5, a learning rate of 0.0005, and a bag fraction of 0.50. Akaike's information criterion adjusted for small sample sizes (AICc) was used to identify the most informative model for each response variable (i.e., the "best" model with the lowest AICc). The most informative model was then hand-tuned by adjusting the learning rate and bag fraction to optimize the r-squared value of predicted versus observed values for held out data in 10-fold cross validation (see Elith et al., 2008 for details).

A non-parametric significance test was used to assess whether the mean cross validation deviance of each model was significantly lower than would be predicted at random. In this procedure, the cross-validation deviance for models trained on observed biomass samples was compared to a null distribution of deviance values produced from repeated model runs with the response values randomly reordered relative to rows of response data (following Raes & ter Steege, 2007). One thousand model iterations were run with randomly selected rows (with replacement) to establish the null distribution of deviance values. Frequency percentiles of the null distribution were used to judge the significance of mean cross validation deviance from the true model using a one-tailed test.

The relative importance of each predictor in the lowest AICc model was calculated using the summary.gbm procedure in the gbm library in RStudio. Relative importance is equal to the percentage of the total squared error improvement of the trained model that could be attributed to each variable (Friedman 2001).

d. Dreissenid presence-absence BRT models

A single presence-absence model was trained for both mussel species (zebra + quagga) using boosted regression trees in the "gbm" package of RStudio. The best model was selected using a backward elimination process wherein an initial model was trained using all ten predictors and then the least informative variables were dropped until the cross-validation deviance was maximized. This resulted in a model with three predictor variables. The final model was trained using the log odds ratio loss function appropriate to binary response variables, a tree complexity of 5, a learning rate of 0.0005, and a bag fraction of 0.5. This model predicted the probability of dreissenid presence and was evaluated using the area

under the curve (AUC) of the receiver operating characteristic (ROC) plot calculated using ten-fold cross validation (see Fielding & Bell, 1997).

III.   Results

Dreissenid biomass (AFDW) ranged from 0 to 52.6 g/m^2 with the highest values seen in the main basin on mixed substrates in temperatures greater than 3°C and depths between 30 and 50 m (Figures 2-4). Sampled biomass was lowest in the North Channel and moderate in the Georgian and Saginaw Bays. The southernmost part of the main basin also had sites with lower biomass. Higher values of biomass tended to be clustered in areas of the main basin that included both nearshore and deep-water environments. Several sites with higher biomass were also found in the nearshore areas of Saginaw Bay and Georgian Bay.

A. Dreissenid responses to basin, bathymetry, and substrate

Dreissenid mussel biomass varied significantly among major basins of Lake Huron, bathymetric zone, and substrate type. An ANOVA comparison of mean biomass between lake basins showed a significant difference between means by basin (F = 2.872, df = 3, P = 0.0394), and a Tukey multiple comparisons test showed that the only pairwise significant difference (P < 0.05) existed between the North Channel and the main basin (Figure 6).
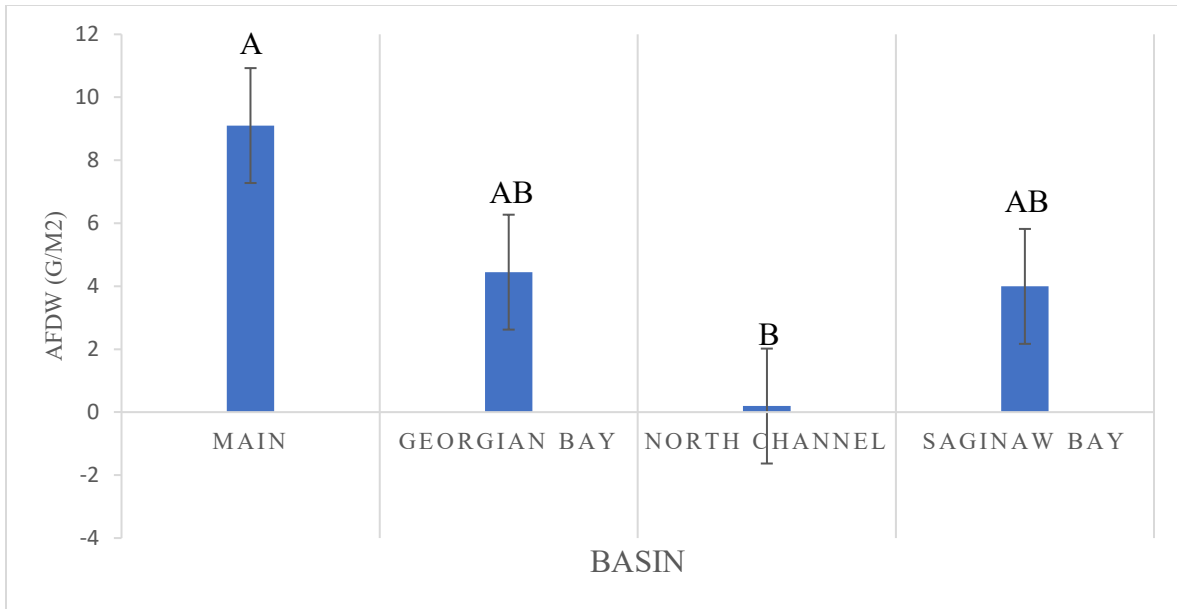
Figure 6: Bar chart showing average biomass (ash-free dry weight; AFDW) sampled in each basin with standard error bars and results of Tukey multiple comparisons of means above each bar, where significantly different groups are represented by different letters.

An ANOVA comparison of mean dreissenid biomass between depth zones showed a significant difference among means between 0-30 m, 30-60 m, 60-90 m, and 90+ m (F = 2.557, df = 3, P = 0.0586) at a 90% confidence level. The Tukey multiple comparisons of means showed that the 0-30 m zone had significantly lower biomass than the 30-60 m zone (Figure 7).
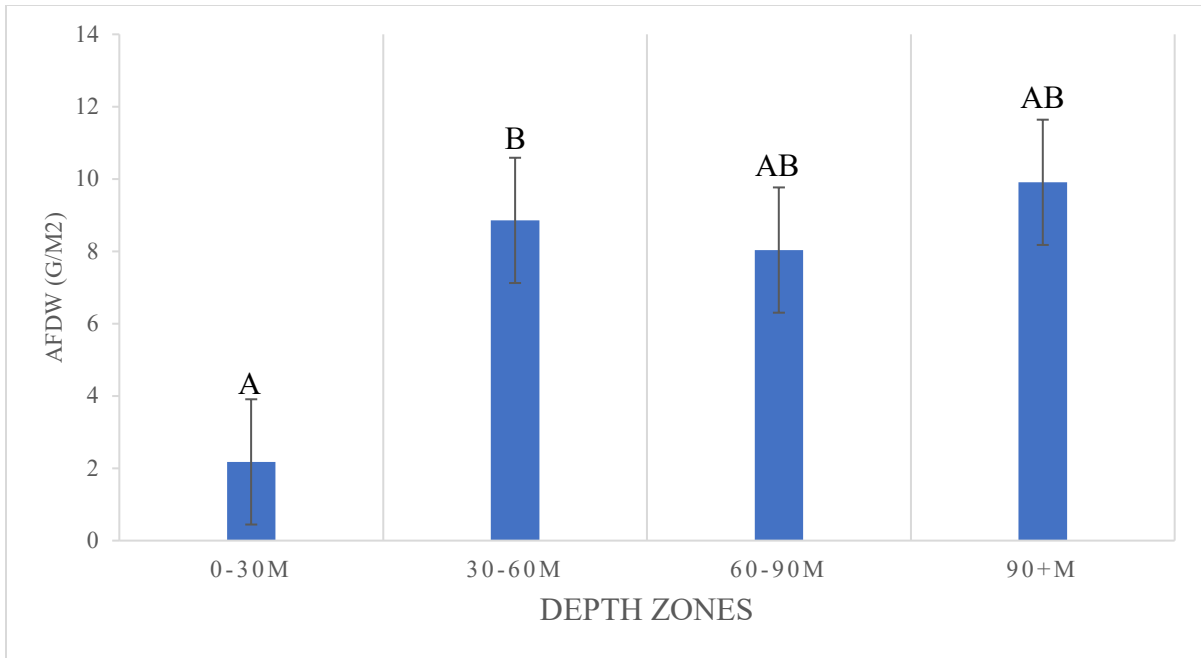
Figure 7: Bar chart showing average dreissenid biomass (ash-free dry weight; AFDW) for different depth zones and results of Tukey test. The modeled variable for bathymetry is continuous, however, it was split into categories for this chart for trend evaluation.

Significant differences in mean dreissenid biomass also existed among substrate categories (F = 2.399, df = 4, P = 0.0541) at the P < 0.10 confidence level. The Tukey test revealed that mussels on silt had significantly lower biomass than those on silty sand (Figure 8). Sand also appeared to correlate with lower dreissenid biomass, though this tendency was not significantly different from the other classes. The three highest means were those with mixed substrate types.
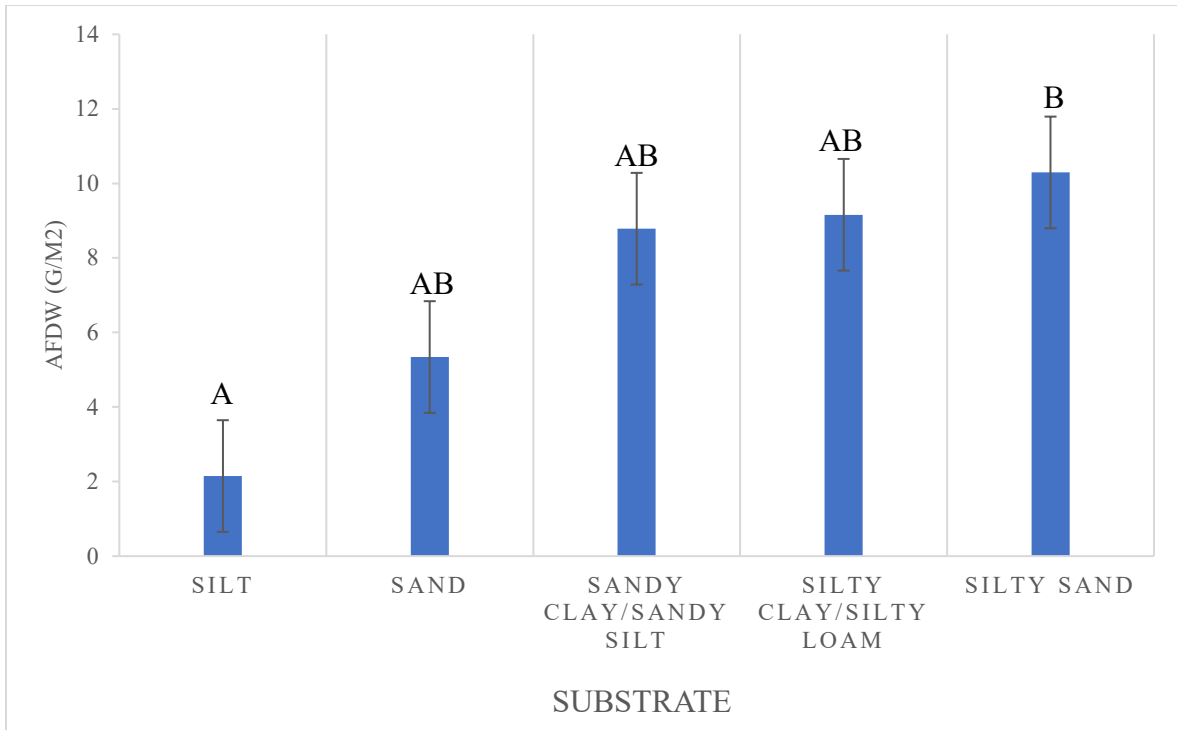
Figure 8: Bar chart showing the average Dreissena spp. biomass (ash-free dry weight; AFDW) for each substrate category with standard error bars and results of Tukey test over each bar.

B. Empirical Bayesian Kriging

The best EBK model used dreissenid biomass for both species of *Dreissena* across Lake Huron. A screenshot of the parameters used for the best model is in Figure 9 below. This figure also draws attention to the distribution of model semivariograms for the one point designated on the map.
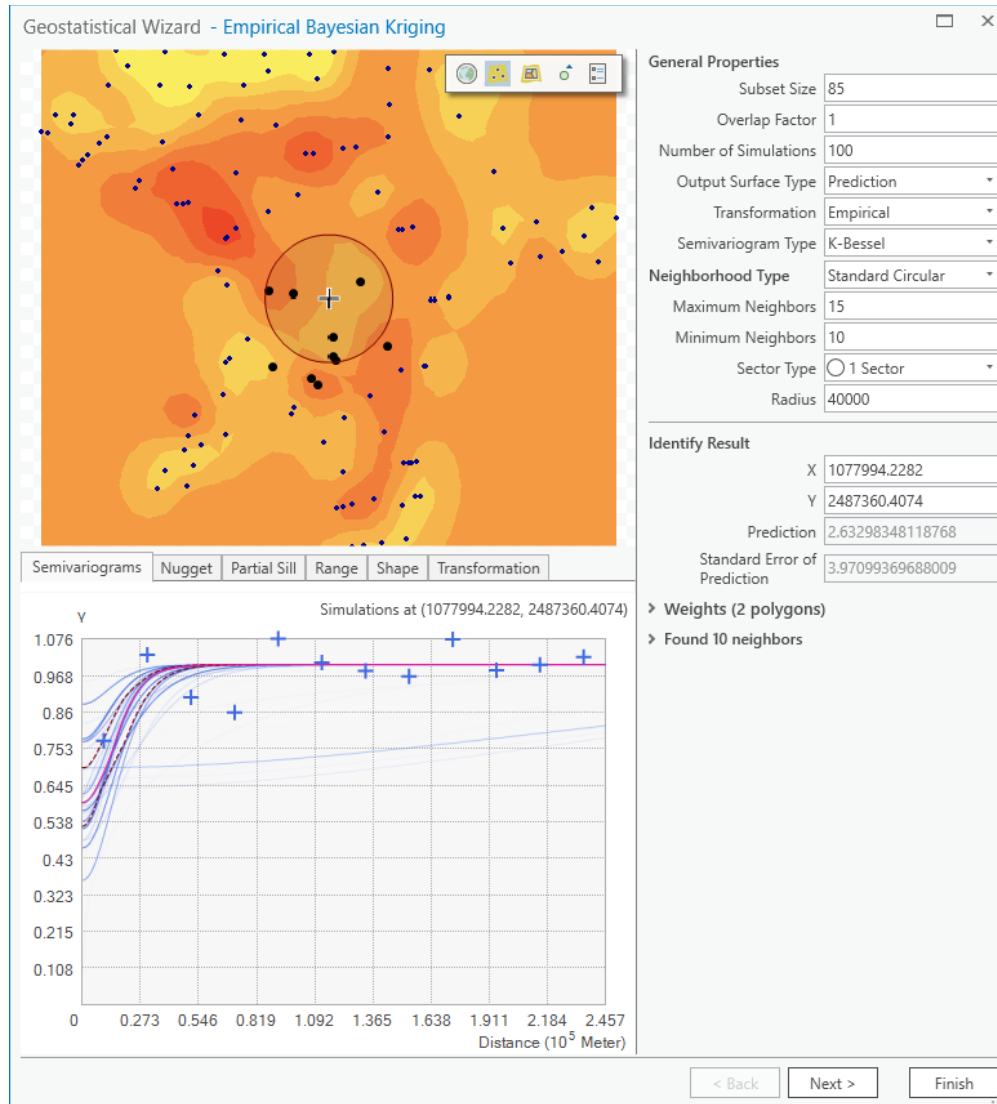
Figure 9: Screenshot of ArcGIS Pro Geostatistical Wizard showing the parameters and semivariogram of the best fit model for biomass (ash-free dry weight; AFDW) of both Dreissena species.

The model had reasonable precision but with a high standard error ($r^2 = 0.61$ +/-9.2), and a RMSE of 11.28. A scatter plot of predicted versus observed values suggested poor model performance with most values falling above the 1:1 line and no obvious trend between predicted and observed values (Figure 10).
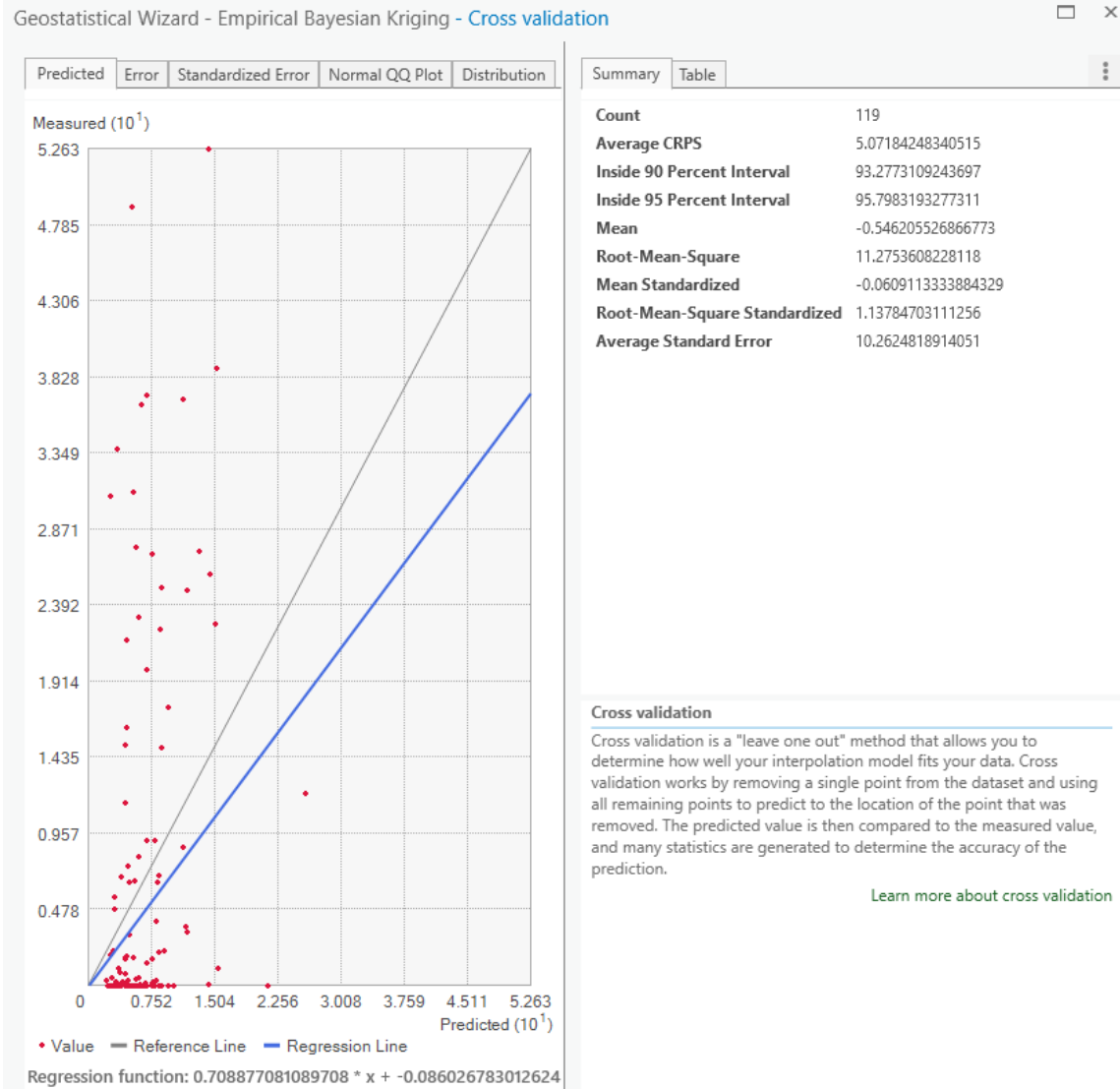
Figure 10: Empirical Bayesian Kriging leave one out cross validation statistics and graph of predicted versus measured biomass (ash-free dry weight; AFDW).

The whole lake distribution output map overlaid with sample values illustrates the poor model (Figure 11). Overall, the map shows a predicted "patchy" distribution as a result of the radius picked in the model parameters and the spread-out sample sites used to create the EBK model. Despite this, the predictive map appears to match fairly well with the observed values.

Dreissenid Ash Free Dry Weight (g/m2)
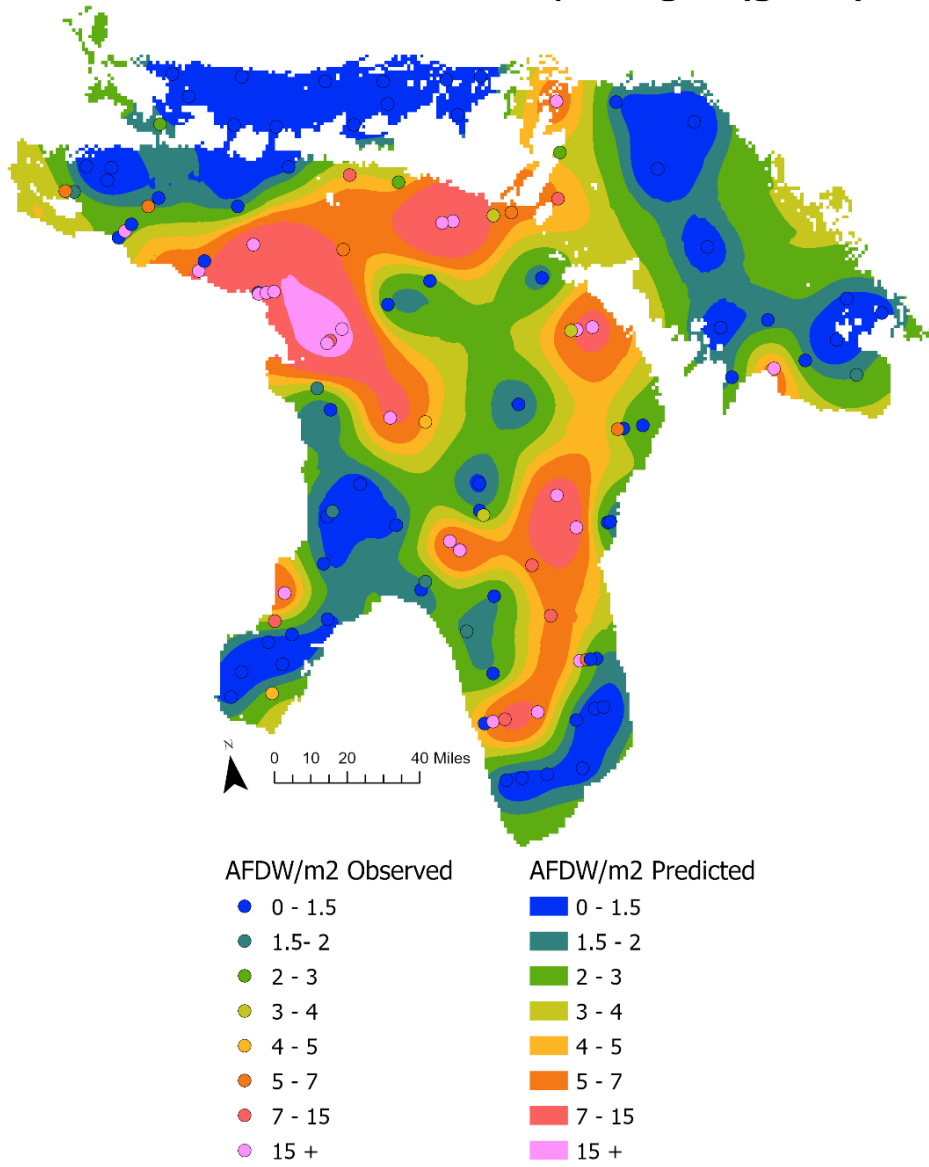
Figure 11: Map of Dreissena *spp. biomass (ash-free dry weight; AFDW) created using Empirical Bayesian Kriging. Overlaid on the predictive map are the actual sampled values of dreissenids using Ponar grabs. The Empirical Bayesian Kriging model uses only spatial coordinates of sampled values to create the predictive map, and no additional predictor variables.*

C. Boosted Regression Trees

The lowest AICc model (Table 1) of natural log transformed (+ 0.0001) biomass (AFDW) of quagga mussels contained four explanatory variables and had a R^2 of 0.27 (P < 0.0001). The lowest AICc model of natural log transformed biomass (AFDW) of both species of *Dreissena* (Table 2) had a R^2 of 0.23 (P < 0.0001). Each model used different combinations of explanatory variables with differing importance. February maximum bottom temperature, bathymetry, and substrate were the three most important variables selected in nine of the top ten most informative models (Tables 1 and 2). February maximum bottom temperature always accounted for more than 50% of the explained variation in each model.

Table 1: First ten rows of the all subsets procedure outputs for quagga mussels only. Only variables with greater than 1% of influence on dreissenids were included in the below table, however, the N_param column designates how many total variables (out of 10) were used for the model. The resid_dev column shows the residual deviance of the model using the predictor variables, train_r2 is the training data r-squared value, test_r2 is the test data r-squared value, and test_SE is the standard error of the model (using the test data). The other columns show the explanatory variable importance (%) to each model. The models in the table are ordered from lowest AICc (Akaike's information criterion adjusted for small sample sizes) to highest, showing the best model in the first row.

AICc	N_param	resid_dev	train_r2	test_r2	test_SE	Feb_max	DOC_10	bathy	SagRiver_influ	Basin_int	Substrate_int
2713.02	4	2634.79	0.43	0.27	0.05	61.50	0	17.94	0	0	20.35
2724.31	5	2635.05	0.43	0.25	0.09	61.64	0	17.90	0	0	20.21
2726.37	8	2631.67	0.47	0.23	0.07	56.19	3.29	12.91	6.96	1.18	19.21
2726.50	5	2634.86	0.43	0.22	0.07	61.57	0	17.91	0.00	0	20.27
2726.64	5	2630.81	0.47	0.23	0.07	57.89	0	14.58	8.29	0	19.22
2726.94	4	2630.78	0.47	0.23	0.04	57.87	0	14.59	8.34	0	19.20
2727.56	8	2631.96	0.47	0.26	0.04	56.41	3.48	12.77	7.76	0	19.30
2727.61	6	2634.22	0.44	0.20	0.09	59.94	0	16.70	0	2.96	20.16
2727.76	5	2631.49	0.47	0.26	0.05	56.68	3.24	12.91	7.76	0	19.41
2728.11	6	2641.43	0.44	0.24	0.06	60.73	4.70	0.00	12.94	0	21.17

Table 2: First ten rows of the all subsets procedure outputs for both mussel species. Only variables with greater than 1% of influence on dreissenids (in at least 2/10 of the top models) were included in the below table, however, the N_param column designates how many total variables (out of 10) were used for the model.

AICc	N_param	resid_dev	train_r2	test_r2	test_SE	Feb_max	st_mary	bathy	SagRiver_influ	Basin_int	Substrate_int
2704.34	5	2618.56	0.63	0.48	0.06	58.40	0	18.95	0	0	22.04
2706.25	5	2617.53	0.64	0.41	0.10	56.63	0	17.89	0	3.51	21.93
2709.33	4	2618.62	0.63	0.50	0.06	58.36	0	19.13	0	0	22.16
2710.13	7	2613.99	0.66	0.42	0.09	53.40	0	15.01	8.62	1.53	20.96
2710.32	6	2615.95	0.65	0.47	0.07	54.64	7.68	14.76	0	2.46	20.23
2715.32	4	2634.42	0.60	0.42	0.08	67.52	0	0	0	6.35	26.06
2715.47	10	2614.13	0.67	0.45	0.07	50.97	5.82	12.32	6.29	0.89	19.79
2715.67	4	2618.41	0.63	0.41	0.08	58.41	0	19.05	0	0	22.28
2717.08	6	2614.15	0.66	0.46	0.06	53.38	0	15.13	9.96	0	21.05
2718.68	5	2617.71	0.64	0.44	0.08	56.42	0	17.95	0	3.45	21.92

The most informative model of quagga mussel biomass used February maximum bottom temperature, bathymetry, substrate, and tributary influence. The partial dependence plots suggested a rapid increase in quagga mussel biomass as temperatures warmed from 1 to 3°C February maximum bottom temperature, after which the curve levels off and quagga mussel biomass remains high. In areas with less than 1°C maximum bottom temperature, quagga mussel biomass was predicted to be very low (Figure 12a). Silt substrate was predicted to support the lowest biomass of dreissenids, with the highest biomass predicted to be supported by silty sand. Mixed substrates including silt, clay, sand, and loam (categories 3, 4, and 5) were predicted to have higher dreissenid biomass (Figure 12b). Maximum quagga mussel biomass was predicted to occur at about 50 m depth and decline to a consistently low value at 75 m depth (Figure 12c). Tributary influence was predicted to support lower dreissenid biomass in areas with low to medium influence, and higher dreissenid biomass in areas with either no influence or high influence (Figure 12d).

Figure 12: Partial dependence plots from the best boosted regression tree model using the natural log of ash free dry weight (AFDW) as the response variable for quagga mussels. The plots show how each variable affects the predictions for dreissenid biomass. The x-axis shows the values for each explanatory variable on a continuous scale for plots a and c and categorical for plots b and d. The y-axis, the fitted function, is centered on the mean response value and represents the change in dreissenid biomass predicted by each explanatory variable. The red trendline is a smoothed spline line representing the best fit to the partial response predictions.

The predicted values of quagga biomass only ranged from 0 to close to 1 gram of AFDW/m^2 suggesting that the model explained little of the variation in the dataset. A map of model predictions (Figure 13) shows the highest predicted biomass in Saginaw Bay, the eastern half of the North Channel, and along the Georgian Bay shoreline. It should be noted that none of the areas with predicted high biomass had observed high biomass, confirming the generally low model accuracy.
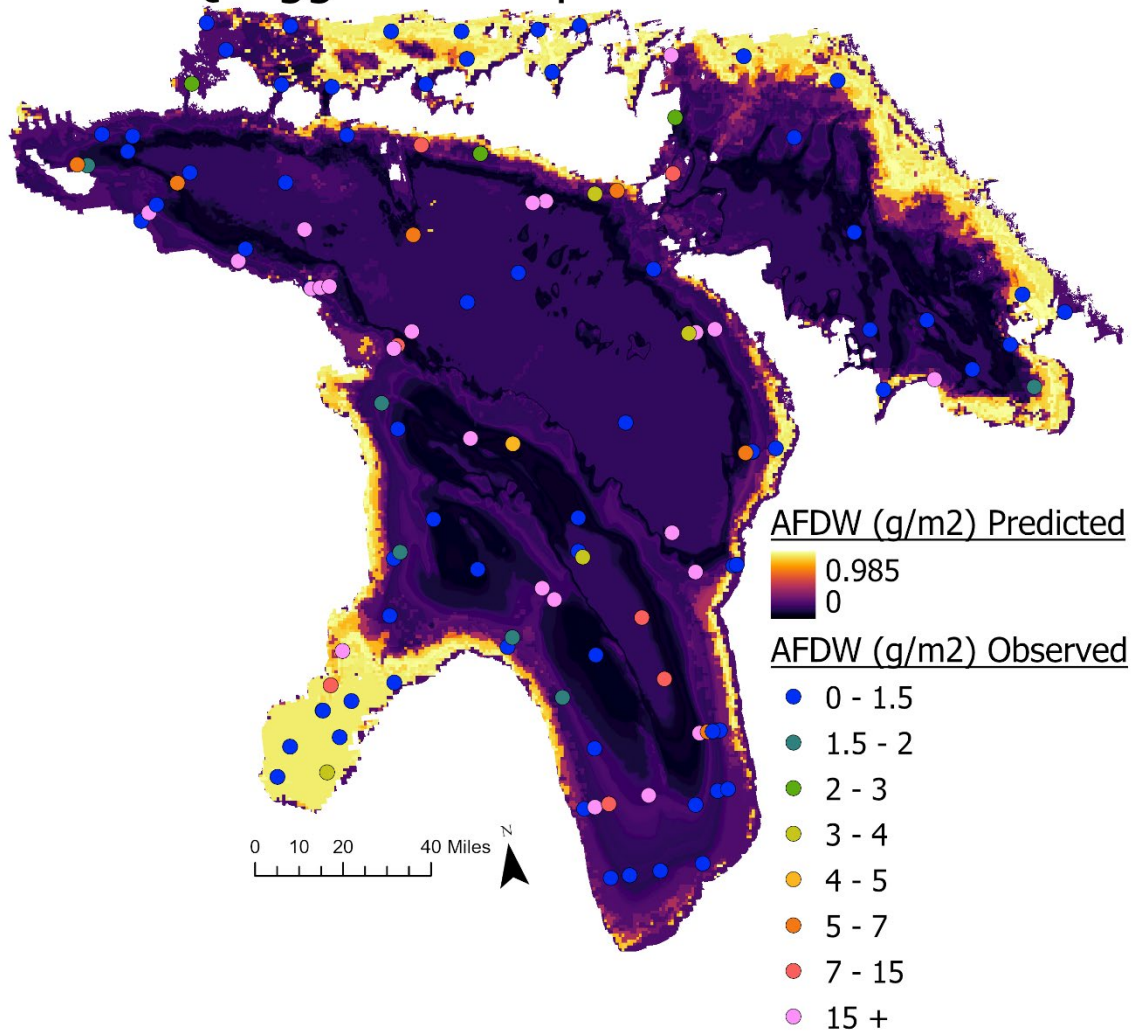
Figure 13: Map of predicted quagga mussel biomass (ash-free dry weight g/m²) created using a boosted regression tree model. This model included four explanatory variables: February maximum bottom temperature, substrate, bathymetry, and tributary influence.

The best model of combined *Dreissena* spp. biomass included five explanatory variables: February maximum bottom temperature, substrate, bathymetry, tributary influence, and April CDOM. The patterns in the partial dependence plots for the predictors in this model (Figure 14) were very similar to those described for quagga biomass (Figure 12). The only difference was the addition of CDOM (Figure 14e) which showed a minor effect on combined *Dreissena* density at very low values of CDOM.
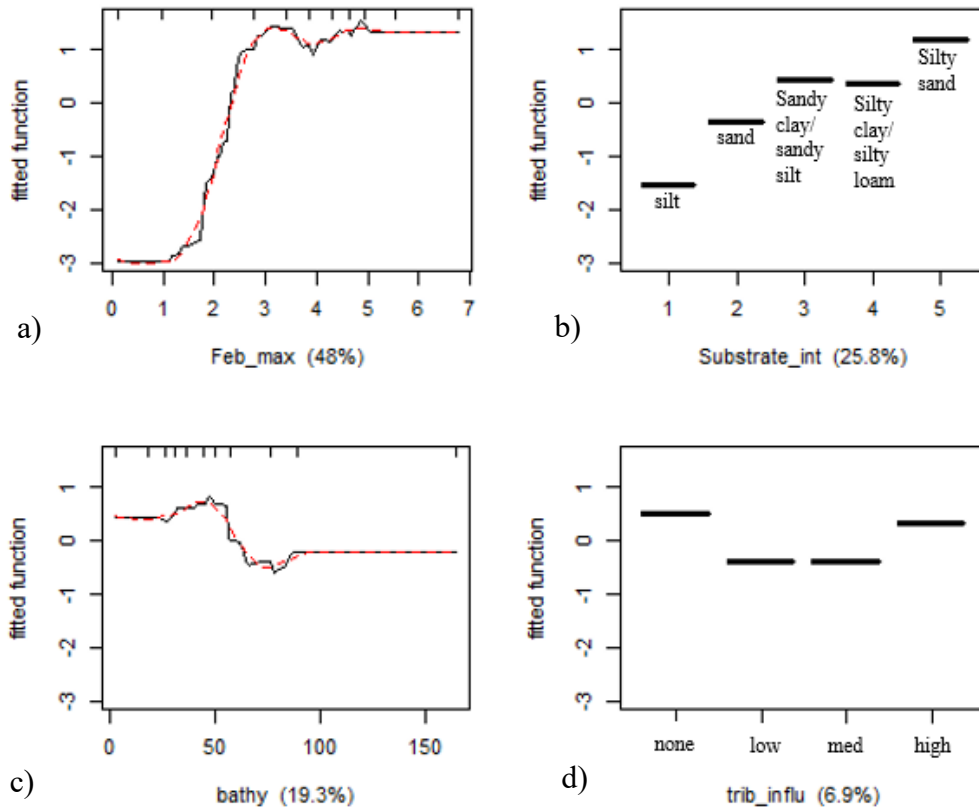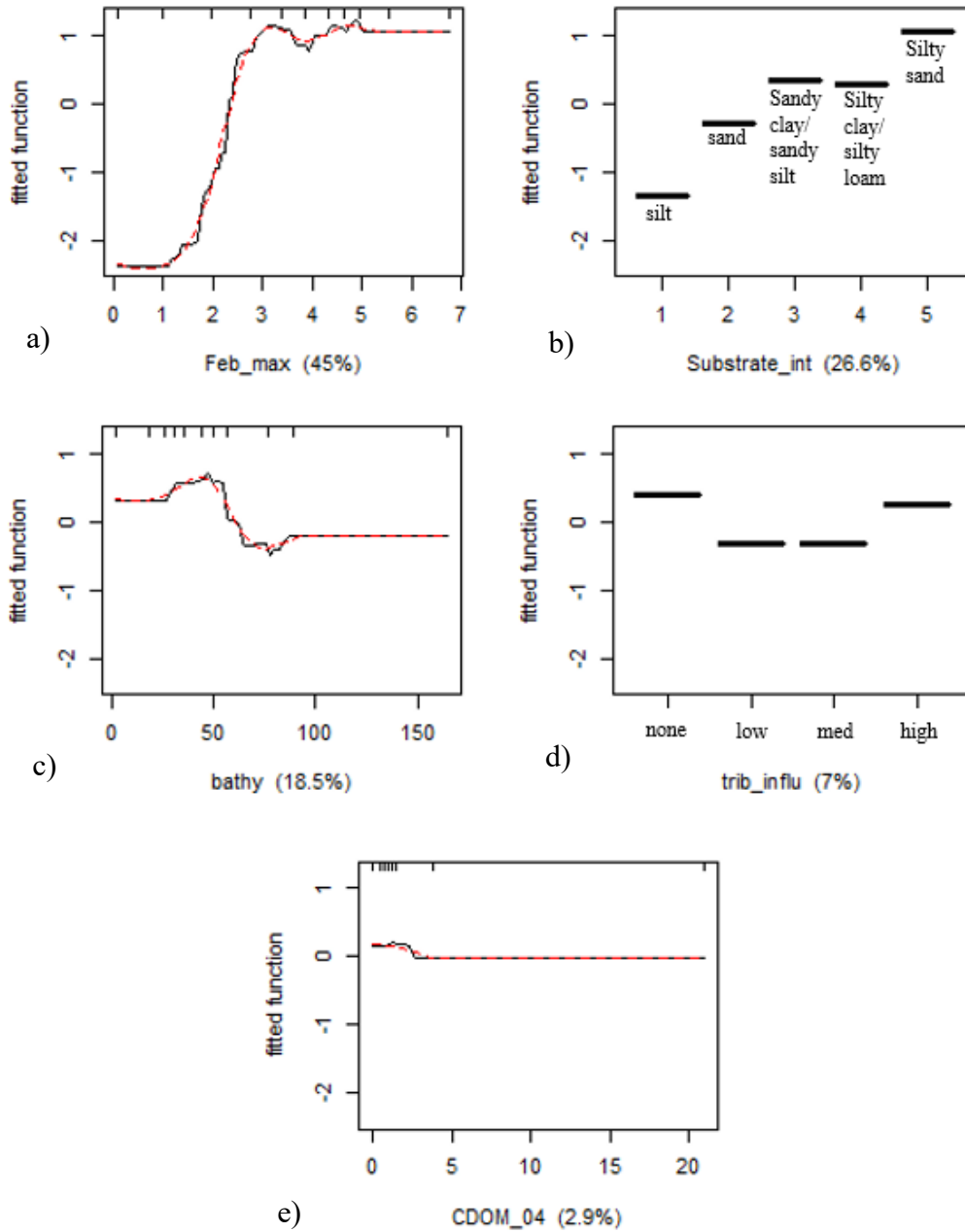
Figure 14: Partial dependence plots from the best boosted regression tree model using the natural log of ash free dry weight (AFDW) as the response variable for both Dreissena species. The plots show how each variable affects the predictions for dreissenid biomass. The x-axis shows the values for each explanatory variable on a continuous scale for plots a, c, and e and categorical for plots b and d. The y-axis, the fitted function, is centered on the mean response value and represents the change in dreissenid biomass predicted by each explanatory variable. The red trendline is a smoothed spline line representing the best fit to the partial response predictions.

50

The range of predicted *Dreissena* spp. biomass, while slightly higher than predicted for quagga mussels, still substantially underestimated the sample values for biomass (Figure 15). However, the general patterns of high to low predictions were a more reasonable match to the sampled values of the variable. For example, lower dreissenid biomass in the North Channel was accurately depicted in this model.
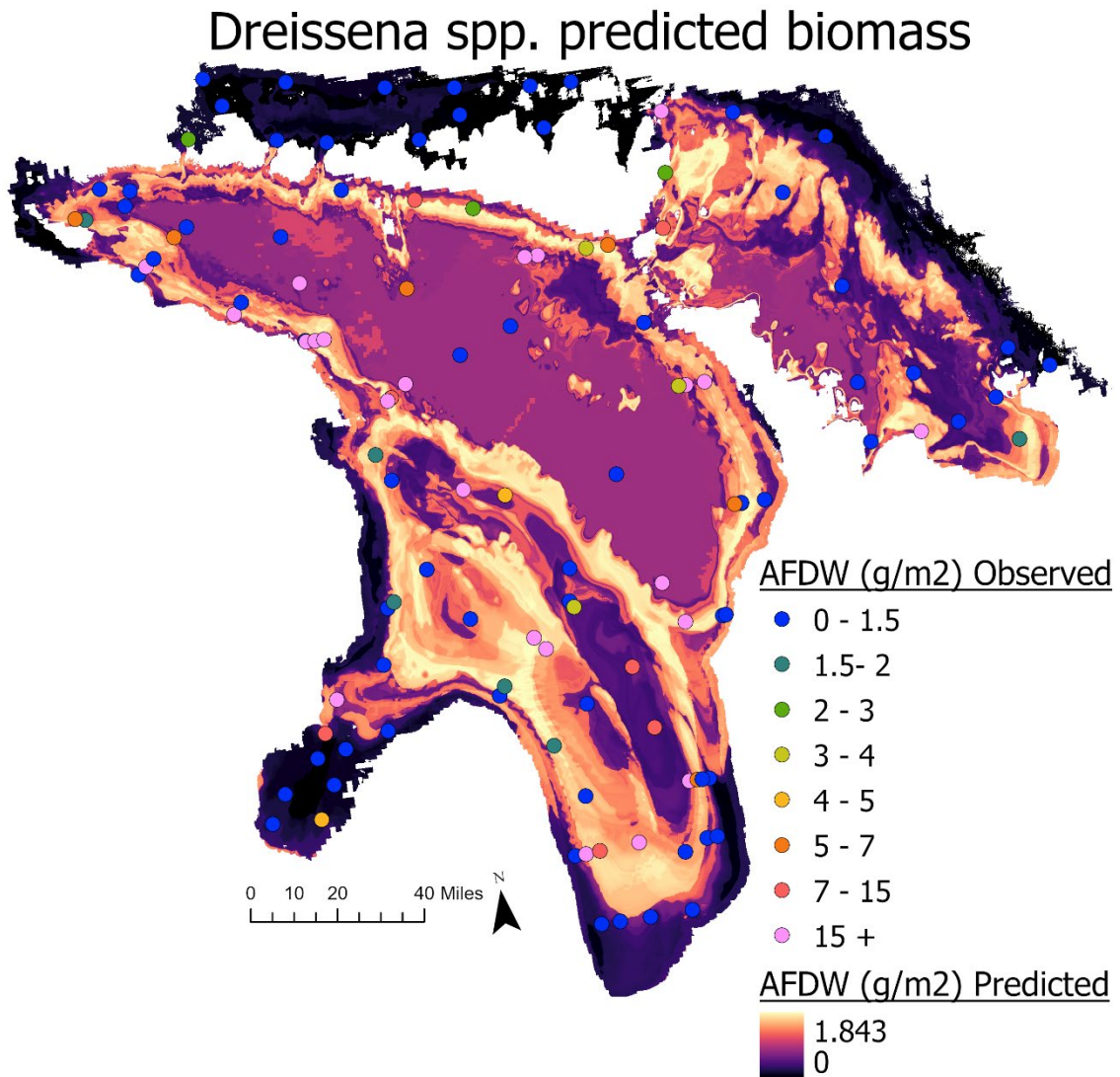
Figure 15: Map of predicted Dreissena *spp. (both zebra and quagga mussels) biomass (ash-free dry weight g/m². The map is the output of a boosted regression tree model using five explanatory variables: February maximum bottom temperature, substrate, bathymetry, tributary influence, and April colored dissolved organic matter (CDOM).*

a. Presence Models

The model for dreissenid mussel presence had a reasonable ROC score of 0.735 with a standard error of 0.051. In general, an ROC score of greater than 0.9 indicates high accuracy, 0.7 to 0.9 indicates moderate accuracy, and 0.5 to 0.7 indicates low accuracy (Akobeng, 2007; Manel et al., 2002). When compared to the null distribution, this model was highly significant ($P < 0.0001$). After backward elimination, three variables were retained in the model (Figure 16): distance from the Saginaw River, bathymetry, and February maximum bottom temperature. Each variable was approximately equal in terms of its relative importance in the model.

The model predicted a moderate probability of dreissenid presence in areas closest to the Saginaw River mouth and a low probability in areas greater than 280,000 m away in the North Channel and the Georgian Bay (Figure 16a). The peak in the middle of the plot suggests that dreissenid probability of presence was predicted to be highest between 150,000 and 250,000 m from the Saginaw River mouth, in the main basin of Lake Huron. While this variable accounted for a high proportion of deviance reduction in the model it is unlikely to represent a mechanistic relationship to the Saginaw River for reasons discussed below.

The predicted relationship of *Dreissena* spp. to bathymetry showed the highest probability of presence in areas with depths around 30-50 m with a very low probability of dreissenid presence in depths greater than 60 m (Figure 16b). The trend for February maximum bottom temperature suggested an optimal temperature range between 3 and 5°C with a peak at about 3.75°C (Figure 16c).

a) SagRiver_influ (38.8%)

b) bathy (31.5%)

c) Feb_max (29.7%)
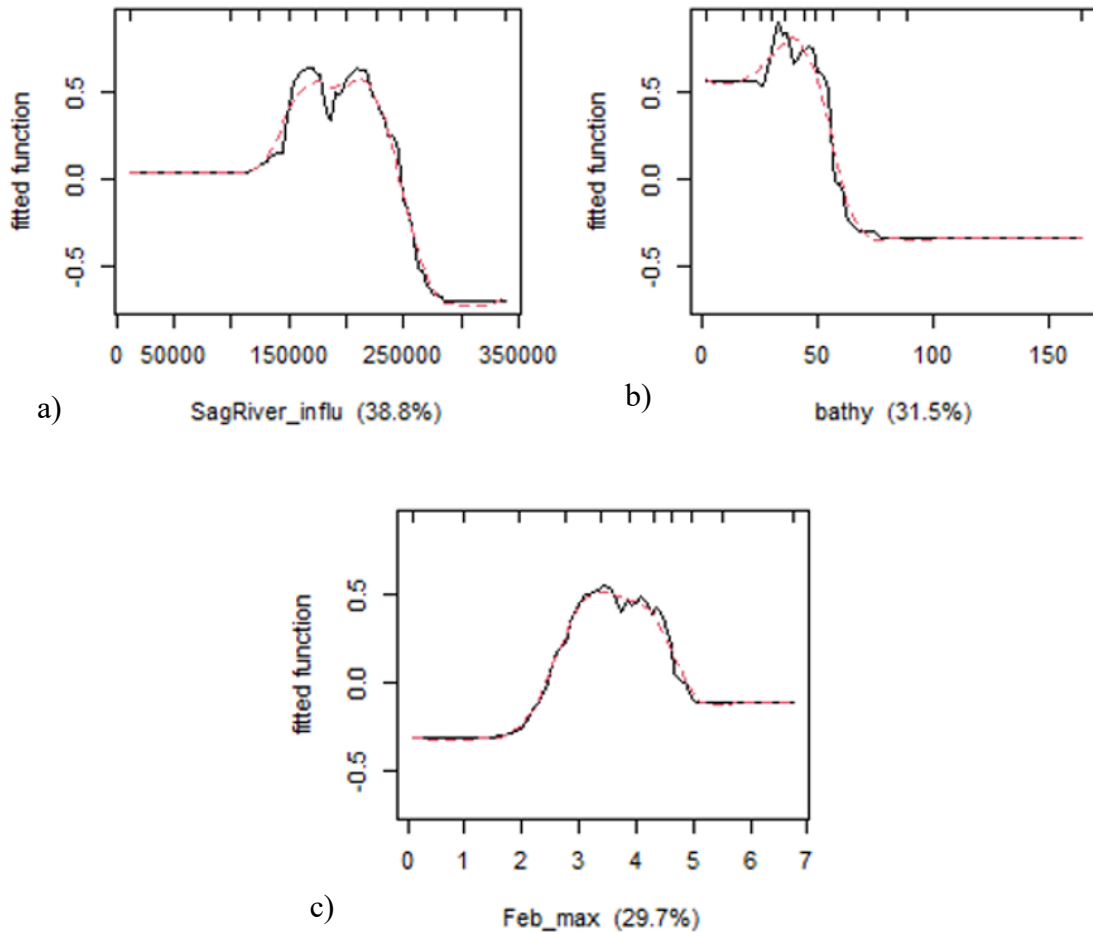
Figure 16: Partial dependence plots for the best explanatory variables for predicting Dreissena *spp. presence. The plots show how each variable affects the predictions for dreissenid presence. Y-axes are on the logit scale and centered to have zero mean over the data distribution. Ticks at the tops of the plots indicate deciles of the observed variable distribution. The solid line is the partial response curve of the taxon, and the dashed line is a LOESS curve (span 0.25) fitted to the partial response data.*

The geographic projection of the model (Figure 17) shows the importance of intermediate depths to mussel presence. The radial ring patterns in the main basin were likely caused by the influence of the distance from the Saginaw River and not likely to reflect a causal relationship to the river so far from its origin. The areas with the lowest probability of presence were found to be the southwestern portion of the Georgian Bay and the central portion of the North Channel.
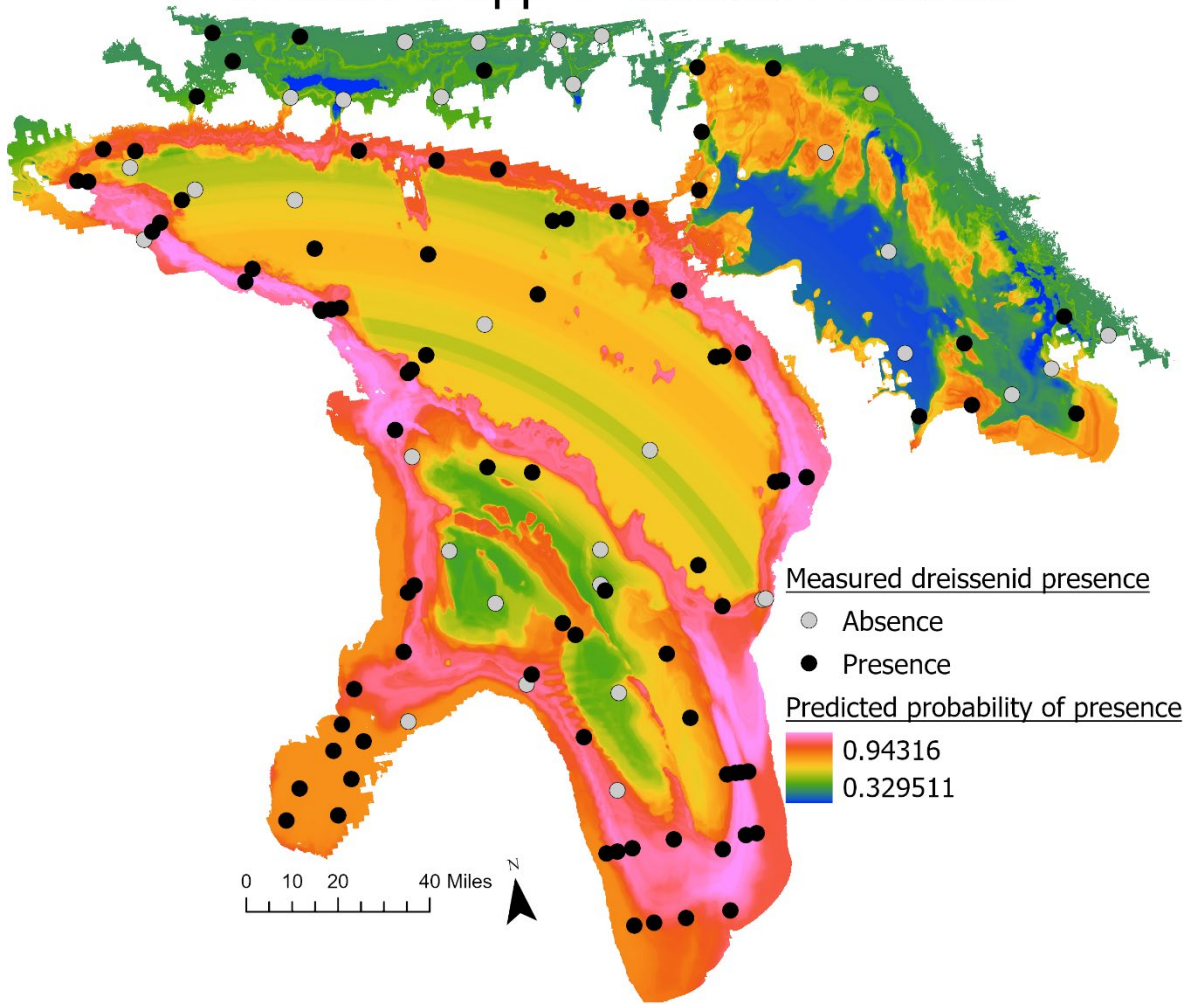
Figure 17: Map showing the probability of Dreissena *spp. presence across Lake Huron produced from a boosted regression tree model. This model used three explanatory variables: influence from the Saginaw River, bathymetry, and February maximum bottom temperature. This map is overlaid by measured values of presence or absence at each dreissenid sample site.*

IV. Discussion

A. Explanatory Variable Trends and Results

The objective of this research was to examine hypotheses about landscape controls on dreissenid presence and biomass and to use predictive models to map spatial patterns in *Dreissena* spp. across Lake Huron. Below, each hypothesis is examined relative to analysis results.

a. Basin

I hypothesized that the Saginaw Bay and Georgian Bay would have moderate dreissenid biomass, the North Channel low biomass, and the main basin higher biomass, which were supported by the comparison of means (Figure 6). Several factors may account for the observed differences in mean biomass. For instance, the North Channel is affected by calcium poor water from Lake Superior, which would tend to lower dreissenid biomass (Marsden et al., 2014). In contrast, the main basin has more favorable water chemistry and nutrients for mussels overall and a wide variety of substrates, resulting in good habitat conditions for dreissenids. Saginaw Bay and Georgian Bay had moderate values of dreissenid biomass on average. Georgian Bay has calcium poor waters in the northeast areas, due to the geology of the bay (Sly & Sandilands, 1988), is extremely deep in some areas, and oligotrophic primarily due to the lack of agriculture in its watersheds (Campbell & Chow-Fraser, 2018), which would tend to lower dreissenid biomass. Lastly, although mussel populations previously thrived in Saginaw Bay in the 1990s, following their discovery in the bay in 1991 (Francoeur et al., 2014), this is no longer the case. Given the apparent differences among basins, it is fair to wonder if different factors may be controlling dreissenid distributions and abundances in each. Were larger sample sizes available, within-basin analyses of controls on presence and abundance may yield deeper insights into the problem.

b. Bathymetry

I hypothesized that dreissenid mussels would prefer moderate depths within Lake Huron representing a balance between food availability and turbulence. This hypothesis was

supported by partial dependence plots showing predicted peak biomass between 30-50 m depth (Figures 12c & 14c), and peak probability of occurrence at similar depths (Figure 16b). Shallow areas generally appeared to have very low biomass. Further, bathymetry had a high relative importance to all models.

Findings by Rowe et al. (2015b) and Karatayev et al. (2021b) reinforce my findings, as both studies found dreissenid biomass to be the highest at moderate depths. Karatayev et al. (2021b) asserted that the ideal depth range was between 30-90 m, a slightly wider optimal depth range than observed here. Lastly, it is known that quagga mussels can colonize deeper and colder waters than zebra mussels (Pilcher et al., 2015; Orlova, 2014; Higgins, 2014; Bootsma & Liao, 2014), however, this was not observed in this analysis (Figures 12 and 14), potentially because zebra mussels have mostly been displaced by quagga mussels in offshore regions of the Great Lakes.

c. Tributary influence

Dreissenid biomass was hypothesized to increase with increasing tributary influence because of the available nutrients in these nearshore areas. My findings were largely inconclusive relative to this hypothesis, in part because of low sample sizes close to shore. Where sample sites did exist in the tributary influence zones, the measured values were mostly zero or low values in the Saginaw Bay and the North Channel. This result may be due to not enough sampling in nearshore sites, but also due to a different factor in the nearshore, such as shallower depths and greater turbulence, making conditions less favorable for dreissenids. The tributary influence dataset may also not be representative of actual tributary influence, since it does not fully account for all the mixing patterns within the lake. Physical conditions, such as temperature and wave disturbance, play a major role in influencing population dynamics at shallower depths (25 m) (Glyshaw et al., 2015; Karatayev et al., 2014; Nalepa et al., 2010). Glyshaw et al. (2015) found that although food supply may be greater at 25 m depth, biomass and average shell length were lower at 25 m than at 45 m, although the population densities were similar in the softer substrates sampled. Larger sample sizes across substrate types within the nearshore zone will be needed to test this hypothesis more robustly. Tributary influence was included in both BRT models of

biomass but had low relative importance values—3.8% and 3.7% for quagga mussel and combined dreissenid species biomass models respectively (Figures 12d & 14d). Although both these models performed significantly better than random (null distribution), neither model had sufficiently high predictive performance to draw strong conclusions about tributary influence.

d. February maximum bottom temperature

I hypothesized that there would be higher dreissenid biomass with higher February maximum bottom temperature due to the negative effects of winter freezing and ice scour on mussel populations. My hypothesis was supported by the model results. February maximum bottom temperature was the single most important predictor of dreissenid biomass in the quagga mussel and both species models (Figures 12a & 14a) with over 40% of the deviance reduction accounted for by this variable. This variable was also included in the presence-absence model, accounting for almost 30% of the explained deviance. The partial dependence plots for the biomass models showed increasing dreissenid biomass with increasing temperatures, but this increase leveled off around 4°C. In the presence-absence model, there was a predicted optimum temperature of 3°C and probability of presence decreased at higher temperatures. Karatayev et al. (2014) stated that dreissenids cannot survive in temperatures less than 0°C (but see Sprung, 1987), however, there was no mention of a decline in higher winter temperatures. In fact, the ideal temperature for zebra mussels is somewhere between 15 and 28°C (Chakraborti et al., 2014; Kobak, 2014; Boekman & Bidwell, 2014), but quagga mussels prefer temperatures 2-3°C colder (Pilcher et al., 2015; Orlova, 2014; Higgins, 2014; Bootsma & Liao, 2014). One explanation is quagga mussels have increased metabolic rates in higher winter temperatures than temperatures closer to 0°C (Vanderploeg et al., 2010). The higher presence probability at 3°C and higher dreissenid biomass in areas with greater than 4°C bottom temperature may correspond to areas that do not reach below 0°C, thereby reducing the physiological stress and mortality from lower temperature or ice scouring of the lakebed (i.e., mechanical stress) (Thorp et al., 1998; Hasler et al., 2019). Thus, my results are likely representing areas of the lake that do not freeze in the winter and not a temperature optimum. These temperatures also relate to bathymetry, with 4°C corresponding with the densest water and the deepest areas of the lake. A deeper

exploration of the effects of winter bottom temperatures (e.g., including other winter months, looking at minimum temperature more closely) is warranted.

e. Substrate

I hypothesized that areas with larger sediment grain sizes (i.e., sand) would have higher dreissenid biomass than areas with silt or mud because of a higher abundance of attachment points (Garton et al., 2014) and greater resistance to turbulence. Sediment grain size also appeared to be correlated with bathymetry, with deeper areas comprised of silt or mud. The ANOVA findings comparing biomass among fine sediment types (Figure 8) failed to support this hypothesis. Instead, sand and silt showed lower dreissenid biomass on average, while mixed substrates had higher biomass on average. Although this trend was observed, only the categories of silt and silty sand were found to have significantly different dreissenid biomass. Previous studies have tended to focus on abundance patterns relative to wide gradients in sediment particle size and hardness (i.e., cobble versus sand) (Garton et al., 2014). The finding that mixed substrate supported higher abundances may be due to the greater heterogeneity for mussel attachment provided by mixed sediments. This is possible but may not be the best explanation since dreissenids are likely still unable to attach to larger fine substrates and attach more to each other in areas with soft substrates (Rowe et al., 2015b). A more plausible explanation is mixed sediment may provide an advantage for predator avoidance as a result of more heterogeneous color schemes making it easier for dreissenids to blend into the background habitat (Kobak, 2014). Mixed substrates may also be indicative of a less turbulent benthic environment with greater mechanical stability and less local disturbance. In areas of turbulence, sandy substrate is easily mixed into the water column and thus is not an ideal substrate for mussels due to the high disturbance in these areas.

Expanding collection methods to sample on larger, coarse substrates, like gravel and cobbles, will lead to a more complete picture of lake-wide abundances. These larger substrates are better attachment surfaces and provide refugia from predators and currents (Kobak, 2014). Dreissenids generally prefer hard substrates to soft substrates (Kumar et al., 2016; Pilcher et al., 2017), a difference seen even more so with zebra mussels. In fact,

mussel density and biomass rapidly decline with decreased substrate particle size (Garton et al., 2014). Given the narrow set of non-preferred substrates sampled, lake-wide predictions should be interpreted with this caveat in mind.

f. April colored dissolved organic matter (CDOM)

Higher concentrations of CDOM were hypothesized to coincide with lower dreissenid biomass due to the decreased light availability resulting in decreased phytoplankton concentrations, and therefore, less food availability for dreissenids (Tzortziou et al., 2015). This hypothesis was not supported by the current analysis. While the biomass model for both *Dreissena* species included CDOM as a part of the best-performing predictive model (Figure 14e), the variable had little influence on the model, and the model itself had limited predictive power. CDOM is potentially a conflicting indicator of dreissenid mussels, since in addition to indicating lower food availability, CDOM may also indicate higher nutrient availability (Brighenti et al., 2018). These conflicting effects may explain the small influence CDOM had on the biomass models.

g. July Chlorophyll-a

I hypothesized that dreissenid biomass would increase with increasing concentrations of chlorophyll-a. This hypothesis was not supported in this study. Saginaw Bay was the only area of the lake found to have high concentrations of chlorophyll-a. Chlorophyll-a is a common proxy used for food availability in lakes since it directly correlates with phytoplankton abundance (Bold & Wynne, 1985; Bennion et al., 2019), an essential part of the lower food web and a primary food source for dreissenid mussels. The lack of influence of this variable may have been due to the fact that lake chlorophyll-a concentration data were only available from satellite data, which are only representative of about the top 10m of the water column (Bennion et al., 2019), rather than capturing information regarding the benthos where *Dreissena* reside.

A more accurate indicator of benthic food availability may be more predictive of dreissenid distributions and abundance. Remotely sensed spring chlorophyll-a (March-June) was used by Burlakova et al. (2018) since it is known that the spring maximum of

phytoplankton is used by zoobenthos as a source of energy for reproduction and growth (Jonasson, 2004). Chlorophyll-a from different months was evaluated in exploratory analyses and did not result in better performance than July chlorophyll-a. Chlorophyll-a has also been used as a proxy for cyanobacteria in past studies (see Boegehold et al., 2019), which is deleterious to mussels. Seston composition and quality influence oxygen consumption and should be considered when evaluating lake-wide metabolic responses to increased food availability (Tyner et al., 2015). Near-bottom turbidity has also been used in the past as a proxy for *Dreissena* food availability (Karatayev et al., 2018a). Obtaining data for these metrics may result in higher performing predictive models if these alterative proxies better capture the food availability for *Dreissena* in Lake Huron. Chlorophyll-a has also declined in the Great Lakes as a result of dreissenids, and so the low chlorophyll-a concentrations seen across Lake Huron may actually be the result of high *Dreissena* densities (Cha et al., 2012).

h. October dissolved organic carbon (DOC)

I hypothesized that an increase in dissolved organic carbon would result in an increase in dreissenid mussels due to the importance of carbon for biological functions of adult dreissenids, veligers, and their prey (Roditi et al., 2000; Barnard et al., 2006). However, DOC was not found to be predictive of either dreissenid presence or abundance, and therefore this hypothesis was not supported by my analyses. DOC should also be related to the CDOM and tributary influence variables, which does appear to be true across large sections of Lake Huron (Figures 2c, 3a, and 3c).

i. Lake Superior influence

Increased distance from the St. Mary's River mouth was hypothesized to increase dreissenid biomass due to the correlation between increased distance and less influence from Lake Superior conditions. Although this variable was not selected in the BRT models, the mapped results appeared to support my hypothesis (Figure 3d). Every sample site except for one in the North Channel was found to have either very low or zero dreissenid biomass. This variation may have been captured by other explanatory variables, such as February maximum bottom temperature, which could be why this variable was not selected in the BRT models.

Maps of spatial variation in calcium concentrations would create a more direct test of the underlying mechanism for Lake Superior influence but would require intensive field sampling. Further, accounting for the three-dimensional flow and mixing of water through Lake Huron could lead to a more direct proxy for St Mary's River influence that would be far more accurate than a simple Euclidean distance.

j. Saginaw River influence

Influence from the Saginaw River was hypothesized to result in increases in dreissenid biomass due to increased food availability and other nutrients that come from the outputs of such a large river (He et al., 2014). This variable was found to be the most important for predicting dreissenid presence in Lake Huron, accounting for 38.8% of the deviance reduction. The partial dependence plot for this variable (Figure 16a) showed intermediate probabilities of dreissenid presence in the vicinity of the Saginaw River mouth, increasing at around 150,000-225,000 m from the river mouth, well outside the likely influence of the river. Rather than capturing a causal relationship between the Saginaw River inflow and dreissenid presence, I believe that the correlation between this variable and dreissenid biomass may have been spurious. The radial distance from the Saginaw River generally matched the pattern of dreissenid abundance with peaks on the west and east coasts of the lake, and generally low abundances in the North Channel and Georgian Bay. A more robust way to test the influence of the Saginaw River would be to use modeled nutrient availability as a predictor variable. Such an approach could be particularly valuable for species like dreissenids that incorporate biological nutrients, such as total nitrogen or phosphorus. Modeling of the Saginaw Bay area of Lake Huron separately could also be useful for revealing the influence of the Saginaw River.

B. Model performance

Species distribution models (SDMs), also called ecological niche models, are important for quantifying species-habitat relationships and predicting species distributions in ecological research, conservation, and environmental management (Yu et al., 2020). A

variety of modeling techniques have been used in the past to develop SDMs, including generalized additive models (GAMs), generalized linear models (GLMs), boosted regression trees (BRTs), Random Forests (RFs), and maximum entropy (MaxEnt) models. Below I discuss the results of the three models presented here and some of the published literature as it relates to these results.

a. Empirical Bayesian Kriging

The EBK model performed well in terms of mapping the distribution of dreissenids using the sample sites, however, it did not perform well as a predictive model. Fair performance is seen both by considering the r^2 value of 0.61 and by looking at the geographically projected map (Figure 11) where predicted and sampled biomass corresponded fairly well. The map is "patchy" suggesting a high degree of spatial variation among the widely spaced sampling sites. Since kriging methods use semivariograms, which are functions of distance between sampling locations, greater geographic space between points makes accurate prediction more difficult, and greatly lowers the accuracy of the model.

EBK models are not usually used for SDMs, but more so for spatial mapping of substrates (i.e., Mulcan et al. 2015; Novaczek et al. 2019) or for visualization of other point data on a continuous scale. Despite this, EBK is still a part of some SDMs and has been used in past ecological studies. In one study, Leach & Leach (2020) used EBK to predict spotted lanternfly densities across vineyard blocks to analyze their spatial patterns. The use of EBK by Leach & Leach (2020) is similar to my own methods, since EBK was used to predict species distribution across geographic space, without using additional predictor variables.

Despite limitations, EBK is still a useful method for visualizing the spatial heterogeneity of dreissenid biomass across Lake Huron. This method is uncomplicated with explanatory variable data and thus helps us see patterns brought about by spatial variation alone.

b. Boosted Regression Tree Models of Abundance

BRT models were used to both examine which explanatory variables were most important for predicting dreissenid biomass, and to map predicted patterns. After exploring a high number of variable combinations, two models were determined as the best possible functions to describe dreissenid abundance, given the available data. Both models had relatively low precision with r-squared values of 0.27 and 0.23, for quagga mussel and both species models respectively. However, the non-parametric significance test showed that model deviance was lower than would be expected at random, suggesting that the results are informative, but should be interpreted conservatively. The all-subsets procedure suggested that bathymetry, winter temperature (specifically February maximum bottom temperature), substrate, and tributary influence may all be important for determining lake wide dreissenid distributions. Increased focus on the effects of substrate on dreissenid attachment and predator avoidance are suggested by these results (i.e., does substrate color effect mussel predation efficiency?). Increased sample collection in the nearshore zone (0-30 m) is also recommended.

c. Boosted Regression Tree Models of Presence

The presence absence model performed the best out of all the models presented here. This model had moderate performance with an ROC score of 0.735. This model used three explanatory variables with almost equal importance to the model: distance from the Saginaw River, bathymetry, and February maximum bottom temperature. Interestingly, bathymetry and February maximum bottom temperature were both included in this model, as well as in both biomass models. This solidifies the importance of these two variables to dreissenid mussels and guides the focus for future modeling efforts.

The success of this model compared to the biomass models may have been due to the simplicity of the response variable in a presence absence model. Compared to biomass models, presence-absence models reduce the variability of the response by using a binary scale. There are many instances of presence models being used to predict dreissenids in past literature, but few that used machine learning. Koutnik & Padilla (1994) used a geographical information system (GIS) to create a spatial model for predicting zebra mussel presence and abundance in Wisconsin lakes using limnological parameters (pH, ionic calcium, nitrate, and

phosphate) and landscape attributes around the lakes (surficial deposits, bedrock type, and USEPA ecoregions). This study was built on the methods of Ramcharan et al. (1992) who explored the same predictive variables, but in European lakes, where *Dreissena* have proliferated much longer. Ramcharan et al. (1992) found pH and calcium concentration to be the most important environmental factors in limiting the spread of zebra mussels, but nitrate and phosphate were also found to be important factors that could be used in modeling invaded systems. Koutnik & Padilla (1994) also used pH and calcium to attempt to quantify both zebra mussel presence/absence and density. Both models showed the same overall landscape-scale patterns for zebra mussels, although their models showed some inconsistencies between predicted and measured values. In a more recent study, Quinn et al. (2013) used maximum entropy (MaxEnt) methods to create SDMs for both zebra and quagga mussels in order to model the potential distributions (presence/absence) of both species to assess their ecological niche overlap. The most important variable contributors to both models were average annual temperature and altitude, and for quagga mussels, the maximum temperature was also important. They used air temperature and not benthic temperature, where the mussels reside, which may have influenced their conclusions. Quinn et al. (2013) found that 75% of the predicted quagga mussel distribution overlaps with the predicted zebra mussel distribution, and although the zebra mussel is close to filling its fundamental niche across both North America and Europe, the quagga mussel does not yet inhabit much of its predicted distribution. All of this suggests that the broad distributions of both species are limited by similar factors (explanatory variables) and therefore, both species of mussel can be combined within the same SDMs.

BRT models have been widely used in SDM modeling due to their usefulness at handling nonlinearity, selecting predictor variables, accounting for interactions among predictors, and quantifying predictor relative importance (Yu et al., 2020). Becker et al. (2020) used both BRTs and GAMs to create SDMs for cetacean species, both of which were successful at describing overall distribution patterns for the majority of species considered in their study. Young et al. (2020) conducted spatio-temporal modeling to look at the effects of climate change on abalone density in Australia. They used BRTs for their main modeling effort, to identify which factors best explain the distribution of productive abalone reefs and

used EBK to create some of their oceanographic parameters in the desired spatial resolution. Yu et al. (2020) developed SDMs for fluvial fish species in the northeastern US, comparing weighted and unweighted BRT approaches. The weighted BRT (WBRT) applies a weight to each species at each sampling site based on individual species abundance and overall species richness. Their results showed the WBRT models had a higher mean deviance explained than the BRT models for the majority of species tested. This could, therefore, be something interesting to explore for dreissenid data in the future. Hallstan et al., 2013 used SDMs to predict phytoplankton and invertebrate (both littoral and profundal) distributions in lakes across Sweden focusing on dispersal ability between lakes using six different modeling techniques: generalized linear models (GLM), multivariate adaptive regression splines (MARS), classification tree analysis (CTA), mixture discriminant analysis (MDA), generalized boosting models (GBM), and random forests (RF). Their study is one of the only studies of invertebrates in lake ecosystems, with the literature generally focusing on river or marine ecosystems.

One downside of presence-absence models is that they treat very small abundances as presence locations despite these habitats potentially being sub-optimal. For example, in the North Channel area of Lake Huron, sample site points show five sites with presence and eight sites with absence (Figure 17). This resulted in a relatively high portion of the North Channel having a probability of presence greater than 50%. Although this probability of presence is moderately high, dreissenid biomass in this area still remains low overall. Despite this limitation, there are still benefits to using presence absence models. For example, a small value for presence may be indicative of range expansion. This information is something that is lost when looking at the biomass models and further adds value to the presence/absence modeling method.

The presence model presented here has some limitations, especially with its reliance on the Saginaw River influence variable. Issues surrounding this variable were mentioned above and are important to keep in mind when looking at the model results. Future models should consider more direct proxies for the mechanisms hypothesized to influence dreissenid abundances—namely, food availability due to proximity to a high point source of nutrients.

C. Data Limitations

Below, I describe limitations in the data used to construct models and evaluate hypotheses, along with suggestions for potential improvements in future studies.

a. Explanatory variables

An important limitation of this study was the lack of direct measures of factors thought to be mechanistically related to dreissenid fitness and abundance, and the limited availability of proxies for these. Some variables did not have data that were spatially continuous across Lake Huron, and thus could not be used just due to the limited area covered by the sample sites. Increased availability of spatially extensive explanatory variable data could help improve model performance and help with hypothesis testing. For example, there is strong literature support for the importance of pH and calcium for dreissenid mussel shell development (Ramcharan et al., 1992; Koutnik & Padilla, 1994; Hincks & Mackie, 1997). In addition, nitrate, phosphate, turbidity, dissolved oxygen are all important water chemistry variables and should be explored in models (Ramcharan et al.,1992; Koutnik & Padilla,1994; Garton et al., 2014; Hincks & Mackie, 1997; Ackerman, 2014). Turbidity is also known to affect metabolic rate, along with size and temperature (Summers et al., 1996). Current velocity and turbulence are important predictors of microhabitat use (Ackerman, 2014; Garton et al., 2014). Circulation in Lake Huron shows more spatial variation than other large lakes due to the presence of many islands, bays, connecting channel flow, and other complex physical features, resulting in existing hydrodynamic models being unable to adequately resolve circulation across the entire lake, although models are more accurate at smaller spatial scales (Nguyen et al., 2014). Being able to accurately model bottom currents so they can be included as an explanatory variable in dreissenid predictive models could also be important.

Some of the variables that were included in the models could be improved further. For example, the chlorophyll-a, CDOM, and DOC variables all came from satellite data and are thus a more accurate representation of surface water conditions than conditions in the

benthos. These data were used since there was no spatially continuous data available for these variables in the benthos, just as was described by Burlakova et al. (2018). Turbulent mixing of the water column occurs between the top and bottom layers of the lake, transporting phytoplankton between the euphotic zone and the benthos, however, this mixing stops after stratification occurs (Rowe et al., 2015a). Therefore, chlorophyll-a concentrations, and potentially CDOM and DOC concentrations, could be similar at the surface and in the benthos, but only prior to lake stratification. However, there are only a few studies that mention using surface chlorophyll to represent benthic food availability (Jonasson, 2004; Sierszen et al., 2006; Burlakova et al., 2018) and this association is not studied for CDOM and DOC. Therefore, obtaining benthic data for chlorophyll-a, CDOM, and DOC is necessary to get a more accurate representation of benthic conditions than can be provided by satellite data.

In addition, distance from the Saginaw and St. Mary's River variables were created as proxies for flow of water with unique chemistries into Lake Huron, however, the explanatory variable datasets would be far more accurate if this influence could be measured and not approximated based on a simple Euclidean distance measure. In the case of the St. Mary's River influence, the intention was to capture the low calcium water flowing into Lake Huron from Lake Superior. If a spatially complete dataset for calcium was available, at least for the North Channel area and possibly the northern portions of the main basin, this would be able to capture the desired factor and this proxy variable would no longer be needed. The case of the Saginaw River variable is more complex. The intended use of the variable was to represent the increased flux of nutrients and food availability in the Saginaw Bay area as a result of the large inflow of nutrient rich water (He et al., 2014) from the Saginaw River. However, as described above, this was not what the resulting variable ended up capturing, at least in the context of the presence absence model. Therefore, creating a new explanatory variable to represent this trend is difficult since it is unknown what mechanism this variable actually captured that was important for determining dreissenid presence. This is a large limitation that can only be addressed by obtaining additional explanatory variable datasets and testing to see if any capture the variation in the dreissenid dataset as well as this proxy.

Substrate maps are a widely acknowledged data gap in the Great Lakes and affect the ability to make effective lake-wide predictions of ecological patterns (Riseng et al. 2018). Such was certainly the case in this study, despite the apparent importance of local scale measurements of substrate to the dreissenid models. Available lake-wide substrate maps fail to distinguish among soft substrate types, which were shown here to relate to dreissenid biomass. It is important to include multiple relevant fine sediment classes if future substrate maps are to be useful for predicting dreissenid presence and abundance. Although coarse substrate data would be useful, this should not be obtained at the expense of losing fine substrate variation for the most accurate representation of substrate types and their effects on dreissenid distributions.

Possible explanatory variables were only summarized from 2017, since that was when the dreissenid samples were taken. However, dreissenids integrate conditions over their entire 3 to 5-year life span (McMahon, 2012). Therefore, data from a single year may not be predictive of conditions that limited population success of dreissenid mussels. Summarizing variables over the typical lifespan of a mussel may be more appropriate. This was not done in this case to simplify the number of potential explanatory variables considered. Future work should consider expanding the considered timescale to the average lifespan of a mussel.

b. Dreissenid data

All the dreissenid data used for these models came from a single spatially intensive CSMI survey in the summer of 2017 and was collected using Ponar grabs at 119 sample sites spread across all four basins of Lake Huron. Since all these data points were in the summer of one year, dreissenid presence or biomass cannot necessarily be predicted for other seasons or other years. Including data from multiple years could lead to a more generalizable model (section D).

Although the 119 sample sites provided relatively good statistical power for this scale of analysis, there were still some areas where the point densities were low (e.g., Figure 11), thus the model may be less accurate in those locations. The CSMI surveys are some of the most spatially extensive surveys of the Great Lakes and have dramatically increased our ability to understand spatial distributions and mechanisms (Nalepa et al., 2018). However,

emerging technologies like autonomous underwater vehicles (AUVs) offer promise for gathering more well-resolved data over larger spatial domains. AUV surveys could simultaneously collect local scale water and habitat quality measures (calcium, substrate, etc.) and allow for interpolation at the lake-wide scale to create more informative predictor datasets. These technologies will be described further below (section D).

Ponar sampling is unable to collect organisms on hard substrates which are preferred by dreissenids (Kobak, 2014; Kumar et al., 2016; Pilcher et al., 2017), resulting in an incomplete picture of dreissenid mussel distributions across the lake. Obtaining additional dreissenid samples from coarse substrates could expand the response gradient by incorporating higher biomass values and more distinct substrate classes. The reliance on hard substrates is stronger for zebra mussels than for quagga mussels (Kumar et al., 2016; Pilcher et al., 2017), so the models may underestimate true zebra mussel biomass and presence more so than quagga mussels. In addition, a standard Ponar grab only samples a relatively small area of ~ 0.05 m^2, which cannot alone give insight into the broader benthic habitat or accurate mean dreissenid biomass, due to the patchy nature of dreissenid communities (Lietz et al., 2015; Karatayev et al., 2020). Benthic habitats, just like dreissenid populations, can vary greatly, even at small spatial scales. For example, Ponar grabs do not capture information about the surrounding substrates, such as if there is a large boulder a few meters away from the sample site. This would alter the conditions at that site, but the sampling technique would not reflect the reason why (the shelter from currents provided by a large boulder). The other issue is extremely heterogenous dreissenid distributions, even on homogeneous substrates, due to facilitation and competition (Karatayev et al., 2018b). Finally, lab protocols for estimating mussel abundance from Ponar samples are labor intensive and time consuming.

Sampling bias against zebra mussels effectively eliminated the possibility to evaluate their spatial variation across Lake Huron. Obtaining more zebra mussel samples in order to model the species separately would be helpful for modeling distributions more accurately, however, this information is likely difficult to obtain. Zebra mussels have continued to decline in Lake Huron, just as in the rest of the Great Lakes, as they are replaced by quagga mussel populations (Stewart, 2014; Xueying, 2016). Despite not having enough zebra

mussel samples to generate a model, there were enough quagga mussel samples to generate a separate model, and this can be used to look at the differences in prediction for various areas of Lake Huron when zebra mussels are left out of the model, versus when they are included (in the both species model).

D. Next steps

Below are some next steps laid out to address the data limitations above and how to integrate other studies with these models.

1. **Obtain more data to address the spatial gaps in the dreissenid sample sites and address limitations of current sampling methods**

Model results suggest a need for improved data density about *Dreissena* spp. at the scale of Lake Huron (especially for zebra mussels), and improved maps of key explanatory variables (see below). With additional samples, statistical power can be improved, and more accurate predictions may be obtained. For example, there is a lack of sampling in areas of Lake Huron greater than 100 m in depth. The ANOVA (Figure 7) and model results (Figures 12, 14, & 16) both support a decline in dreissenid biomass and presence in areas greater than 90 m, however, more data is necessary in order to see if this conclusion holds true or if it is just a result of sampling bias. Obtaining more data that represent dreissenid presence and biomass on a spatially continuous basis is essential for creating the most accurate models.

Techniques that can accurately sample dreissenid mussels, and the benthos in general, over larger spatial domains but with greater data density are needed to replace or supplement Ponar grabs. Since Ponar is only for soft substrates, traditionally, sampling on hard substrates has been done using SCUBA divers (Mehler et al., 2018). However, diver-based methods have many limitations, such as limits on diving time, weather conditions, total area able to be sampled, and depth range that can be covered (Lam et al., 2006), in addition to the costs of training specialized divers.

Photographic means of sampling and other advanced technology have shown promise in addressing sampling limitations. Mehler et al. (2018) used sonar technology coupled with

underwater imagery to model dreissenids in the Niagara River. Side scan sonar combined with underwater imagery analysis and Geographic Information Systems (GIS) has been used in many ecosystems for benthic mapping, including a few studies of *Dreissena* in Lake Erie (Coakley et al., 1997; Berkman et al., 1998; Haltuch & Berkman, 2000). While sonar works well in areas with soft substrates, *Dreissena* cannot be discriminated from coarse substrates in sonar images due to similar backscatter signals between mussels and large substrate particles, such as gravel (Mehler et al., 2018). Underwater imagery is an affordable and effective way to provide broader insight into habitat heterogeneity, including substrates, aquatic vegetation, and predator presence (Mehler et al., 2018). Despite this, Ponar sampling cannot be completely replaced because it allows for density and size-frequency estimation of dreissenids (Mehler et al., 2018).

Remote sensing methods for studying freshwater benthos are relatively new, although these methods have been explored in marine systems for decades (Karatayev et al., 2018b). Karatayev et al. (2018b) developed an innovative vessel-towed camera system containing a GoPro camera attached to a snowmobile chassis. This device can take photographs of the lakebed, which is made possible through the clearer waters caused by mussel filtration (Karatayev et al., 2018b). Photographic data may allow for much greater spatial coverage and less manual labor than using Ponar grabs. Although these methods can require labor-intensive photointerpretation to manually interpret, this is still less time intensive than traditional lab methods (Karatayev et al., 2018b). This new method can allow researchers to overcome issues with capturing the small-scale patchiness of dreissenid mussels to obtain population estimates that are more representative of the population mean mussel abundance. Following this study, Karatayev et al. (2021a) went on to develop a drop camera system to estimate *Dreissena* populations in almost real-time. This was also the first time they explored using this equipment outside of the nearshore zone and it showed great success in the areas of Lake Erie where this new equipment was used, except for the western basin where video image analysis greatly underestimated *Dreissena* densities (Karatayev et al., 2021a).

Underwater video methods using remotely operated vehicles (ROVs) and diver-operated video (DOV) are commonly used in benthic assessments of biodiversity (Boavida et

al., 2016), especially in marine systems. Technological advances have made video a promising tool for monitoring and sampling, including drop-down cameras, ROVs, AUVs, and DOVs (Boavida et al., 2016). These video surveys make collection of large amounts of data in a short amount of time possible (Boavida et al., 2016). Multiple studies have documented the power and precision of the human eye for in situ species identifications (Pelletier et al., 2011; Assis et al., 2013). This identification is more accurate with human identification versus remote sampling techniques such as ROV and AUV (Boavida et al., 2016).

Advanced technology research has rapidly developed in the past decade and has made autonomous robots possible. For instance, AUVs can acquire large amounts of data through sensors and cameras as they travel along the lake bottom. Therefore, AUVs would be able to provide not only information regarding dreissenid distributions, but also collect additional explanatory variable data at the same time. AUVs are becoming more widespread as a way to survey substantial spatial areas (Foster et al., 2014). Images from an AUV are typically higher quality than those from other image-based platforms due to their constant height above the seafloor, provided illumination for photography, and pre-planned routes that result in accurate geolocation (Foster et al., 2014). In addition, there is also development of AUV swarms that can work together, each with different data gathering capabilities (Dawson & Allison, 2021). Despite the advances in the hardware and software for these devices, the methods to automatically process these large amounts of data remain challenging and manual interpretation is impractical and time consuming. Automated approaches to dreissenid identification and enumeration would be particularly beneficial when coupled with robotic imaging systems. Such approaches, once mature, would make it possible to better quantify the current distributions of dreissenid mussels and make predictive biomass models possible.

Although these methods would still require physical sampling and lab analysis in order to calibrate and validate estimates, using photographic and other technologies could allow for greater aerial coverage of the lake bottom. Since Ponar surveys are a widely available sampling technique, this method should continue to be used, despite its limitations, in conjunction with advanced technologies. In addition, there are many other benefits to continuing physical sampling, such as assessing size of individual mussels and the

performance of advanced technology techniques, calibrating cameras/sensors, and aligning with past studies to maintain consistent data for analysis of temporal trends. Increased physical sampling density and additional habitat data (see below) will facilitate expanded analysis of the relationships between habitat and mussel distributions.

2. **Obtain more complete datasets for explanatory variables and explore trends seen in the models further**

The models presented here were limited in part by the lack of available lake-wide geospatial data about key explanatory variables. Imperfect proxies were used for some mechanisms. Improved data about key explanatory variables has a high potential to improve modeling performance. Specifically, the highest priority should be placed on obtaining data for calcium concentrations, food available to benthos, and complete substrate maps (i.e., including variation in both fine and coarse substrates).

One main trend seen in the models that should be explored further is winter bottom temperature. After exploring bottom temperatures across the entire year, February maximum bottom temperature was found to have the highest correlation with dreissenid biomass and produced models with greater accuracy. Due to the importance of February maximum bottom temperature to the BRT models (both biomass and presence), other winter bottom temperatures should be explored. The results showed a temperature optimum around 3-4°C, likely indicating areas that do not freeze in the winter. This may be further supported by looking at minimum winter temperatures or maximum temperatures during other winter months. In locations below the thermocline, the winter maximum temperature occurs in the fall when the water column first becomes isothermal (when the water column is vertically well mixed) (Beletsky & Schwab, 2001). These higher temperatures may be important for dreissenid growth, especially since they generally correspond to greater mixing and therefore greater access to phytoplankton (Fahnenstiel et al., 2010), and thus should be explored further. Studies on the importance of temperature are also relevant as climate change continues to increase the water temperature of the Great Lakes (Trumpickas et al., 2009;

Fichot et al., 2019); therefore, evaluating the effects of bottom temperature on dreissenids may help predict future trends.

3. Create individual models for each of the lake basins, instead of modeling all of Lake Huron together

A categorical variable that represents the different basins was used as an explanatory variable for these data, however, it was not a part of any of the best-fit models. Each of the basins of Lake Huron is bathymetrically and chemically distinct and may provide unique habitats for dreissenid mussels. Each of these basins has different nutrient concentrations due to the inflow from differing sources across geographic space, such as the Saginaw and St. Mary's Rivers. If sample sizes were sufficient, it could be beneficial to model the basins separately to account for more locally influential drivers. Based on the ANOVA (Figure 6) and the cartographic representation of this variable (Figure 2a), exploring each of the basins separately is warranted. Currently dreissenid sample size and available explanatory data limit this effort (next steps 1 & 2 above).

Based on the Tukey test, the focus should be on modeling the North Channel first since it is significantly different from the main basin. The St. Mary's River influence variable provides support for the effect of Lake Superior water conditions on the North Channel and thus part of the reason why it is significantly different from other areas. Despite not finding significant differences between the other basins, there is also precedent for analyzing Saginaw Bay separately from the rest of Lake Huron. There are many models that look at Saginaw Bay alone (Millie et al., 2008; Winslow et al., 2014; Lavrentyev et al., 2014) due to its differences from the rest of the lake. For example, only Saginaw Bay has reoccurring harmful algal blooms (Francoeur et al., 2015), which indicates nutrient differences between Saginaw Bay and the other basins. Cyanobacteria, part of harmful algal blooms, was found to inhibit reproductive success in quagga mussels by disrupting spawning and fertilization (Boegehold, 2018) and causing mortality to veligers (Boegehold et al., 2019). Filamentous green algae (nutritious algae), found in Saginaw Bay along with the blooms, provides a primary production base for food webs, habitat for epiphytes, and cover

74

for invertebrates, however, these large blooms are more problematic than helpful (Francoeur et al., 2015). Georgian Bay has also been the subject of individual modeling in the past, separate from the rest of Lake Huron due to the differing geology from the rest of the lake (Campbell & Chow-Fraser, 2018).

Therefore, it could be justified to explore models for different lake basins to account for more variability by including the explanatory variables most important to each unique area of the lake. Monitoring programs generally include sample sites from all basins of Lake Huron, however, frequently, management decisions and long-term trends are developed using samples primarily from the main basin (Paterson et al., 2020). Patterson et al. (2020) reported differences between the basins of Lake Huron in terms of bioaccumulation and found differences in the nitrogen stable isotope values in dreissenid mussels in the different basins, with the North Channel having higher values than the main basin and Georgian Bay. They posit the higher values in the North Channel reflect contributions to the waters from the cities surrounding the St. Mary's River (Patterson et al., 2020).

4. **Create a model that can predict the size distributions of dreissenid populations and not just presence or biomass**

The size frequency distribution of a dreissenid population can have large effects on the extent of the ecological impact since mussels of different size have different filtration rates (Bootsma and Liao, 2014; William and Ozersky, 2019; Barbiero et al., 2018). In addition, more productive lakes have been found to have mussels with heavier soft tissues and shells at any given length (Williamson and Ozersky, 2019). Larger mussels can also capture different kinds of particles from the water column than smaller mussels, which causes large mussels to compete more with native unionids than small mussels (Strayer & Malcom, 2014). Therefore, lakes and even potentially different areas of lakes need to be assessed separately to understand the full effects of different sized mussels. This size-related information is essential for better understanding the impacts of dreissenids on both the benthos and the whole lake-wide ecosystem, for both managing nutrient cycling and for economic reasons, such as being able to effectively manage and sustain fisheries. Obtaining

75

accurate size of individual dreissenids is another reason to continue physical sampling, through Ponar or other sampling techniques.

French et al. (2009) found that the size distribution of western Lake Huron quagga mussel populations differed temporally, between 2001 and 2007. They found an influx of newly recruited small quagga mussels at two shallower depths (27 m and 46 m) over those 6 years in addition to the larger mussels already present, but the size distribution of mussels at deeper sites (73 m) in 2007 was all small mussels, suggesting deeper sites were colonized via recruitment and larger individuals may not be able to persist on the soft substrates at these sites. There were also no quagga mussels at the deeper sites in 2001, indicating the continual spread of the population. These results suggest that the densities and size structure of quagga mussels on deep soft-substrate sites may be dynamic and follow recruitment patterns closely (French et al., 2009). If this study had just reported biomass at these different sites, this information regarding size distribution and what that means for the spread and success of the dreissenid populations would be lost.

5. **Address the larval stage of dreissenids (veligers) and attempt to model their distributions**

Dreissenid veligers are understudied despite their seasonal numerical dominance among plankton in the Great Lakes (Kirkendall et al., 2021). Recently, Kirkendall et al. (2021) used a GAM to model veliger densities and found that temperature, chlorophyll-a, and nitrates/nitrites were the most important explanatory variables. Further research on veligers is necessary to understand the invasion ecology of dreissenids, especially since veliger dispersal and survival plays a key role in the spread of dreissenid mussels (Kirkendall et al., 2021). Food availability may also affect gametogenetic development, the onset of spawning, determining the frequency of reproduction (Wacker & von Elert, 2003).

6. **Look at seasonal and long-term temporal variation in dreissenid populations**

Seasonal changes could be an especially helpful addition to the spatial distribution models presented here since many areas of the lake change widely with the seasons. *Dreissena* spp. are known to be influenced by seasonally varying food availability and abiotic conditions (Fanslow et al. 2001). Future incorporation of descriptors of seasonal variation may benefit modeling efforts, as temporal trends in niche limiting factors may affect survival and fitness. For example, seasonal variation in oxygen and food availability (partially due to algal blooms) could place a hard limit on the ability of mussels to survive the summer season (Karatyev et al., 2018a). In addition, changes to oxygen consumption relative to food conditions are greater at higher temperatures, specifically during the summer (Fanslow et al., 2001). Since temperature was found to be one of the main explanatory variables for representing the variation in dreissenid biomass values and presence across the lake, the importance of learning how temperature affects dreissenids in different seasons and over time is increased. For example, the duration of the spring and fall isothermal periods is important, as these are times when dreissenids have access to phytoplankton throughout the water column (Rowe et al., 2017; Karatayev et al., 2021b). Focusing on certain geographic areas of the lake may also be important, for example, Saginaw Bay is highly influenced by seasonal changes due to the presence of harmful algal blooms (Francoeur et al., 2015) and shallow waters that freeze in the winter.

There are also year-to-year differences in nutrient concentrations as a result of meteorological changes. For example, Nicholls (1998) found a strong relationship between phosphorus outflow from Lake Huron and percent ice cover and lake level, where during relatively ice-free winters, phosphorus is in suspension by circulation patterns driven by wind and results in substantially higher in-lake concentrations. The annual cycles of ice formation and loss affect the physical and ecosystem processes within the lake (Nguyen et al., 2017). Knowledge of water movement and temperature distribution during winter months is important for answering several questions related to the Great Lakes, especially for further understanding the hydrodynamic exchange between Saginaw Bay and Lake Huron (Nguyen et al., 2017). Food quality also varies with temperature, for example cyanobacteria are not a good food source for mussels, but they may be the only food available in certain seasons and water depths (Wacker & von Elert, 2003).

Assessing long-term and seasonal dreissenid temporal trends has been completed in past studies, both for Lake Huron and for other areas. Nalepa et al. (2007) assessed long-term changes in benthic macroinvertebrate populations in Lake Huron between the 1960s and 70s and the early 2000s. During the 2000s, quagga mussels were still increasing in density, although zebra mussels had stable populations. Nalepa et al. (2018) further expanded on Nalepa et al. (2007) to include data from surveys in Saginaw Bay from 2006-2009 and surveys of the rest of Lake Huron in 2007 and 2012. These surveys documented the shift from zebra to quagga mussels overtime, with quagga mussels increasing in all depth intervals from 2000 to 2012, however, zebra mussels were rarely found at any depth interval in the 2007 and 2012 surveys. The results stated in both reports were primarily focused on other benthic organisms and the impact dreissenid invasion had on the four major taxonomic groups (*Diporeia*, Oligochaeta, Sphaeriidae, and Chironomidae). Nalepa et al. (2018) mentioned the decline of *Diporeia* in Lake Huron, as has been seen in other studies, and noted an increase in oligochaete densities in areas of the main basin less than 50 m in depth. Burlakova et al. (2018) looked at temporal trends across all the Great Lakes, from 1998 to 2014. In lakes Michigan, Huron, and Ontario, *Diporeia* was replaced with *Dreissena* and Oligochaeta. They concluded that the invasion of dreissenids was among the most important drivers of benthic community change. Glyshaw et al. (2015) looked at southern Lake Michigan to examine temporal trends in the condition and spawning patterns of quagga mussels. They collected quagga mussels monthly in 2013 and concluded that long-term density trends indicate that quagga mussel populations were still expanding at deeper sites (93 m) but may be stabilizing in low and moderate depth sites (25 and 45 m, respectively). Strayer et al. (2019) conducted a large-scale long-term population dynamics study of both species of *Dreissena*. Interestingly, they found that there is not a strong stock-recruit relationship in dreissenids, since veliger density is not a good predictor of subsequent recruitment into juvenile or adult populations. Strayer et al. (2019) also found that populations at individual sites can have a great deal of interannual variation, which could be for a variety of reasons, including changes in nutrient inputs, intra- and interspecific competition, and physical disturbance (i.e., ice scour).

There are some limitations to identifying temporal trends mentioned in the various studies. One main limitation is not having sampling conducted using the same sampling methods over time, which makes it difficult to directly compare values from different studies. This is another reason for continuing Ponar sampling so that future studies can have comparable samples to those in the past and long-term dreissenid studies can be accurately produced. It is also important to use the same sampling units overtime, such as total wet weight (TWW) versus ash-free dry weight (AFDW). Karatayev et al. (2021b) mentioned this in their long-term study of dreissenids across the Great Lakes, stating that these unit differences resulted in lowered model precision since although calculations could be made to convert between these two measures of biomass, the accuracy of the values was still lowered. Second, most studies are of short duration and there are frequent temporal gaps, making temporal analysis difficult due to lengthy periods of unknown information. Third, few monitoring programs cover more than one life stage or population attribute, limiting the understanding of demographic processes and potential impacts of different life stages.

There are several reasons why temporal trends should continue to be explored, in addition to spatial trends. One of these reasons is for monitoring the continued shift in the Great Lakes from zebra to quagga mussels. Second, it is also important to know variations over time for understanding the effects of climate change (and other stressors) on Lake Huron, both in the benthos and in the lake overall (Thorp et al., 1998). Increased temperatures could allow *Dreissena* to colonize lakes and streams in higher elevations and latitudes and experience less ice scour and tissue freezing (Thorp et al., 1998). Responses to changes in average temperature are likely influenced by the genetics of the population, season of the year, and other environmental factors, such as sediment type and oxygen concentration (Thorp et al., 1998). Third, understanding dreissenid year-to-year variation and why populations differ could be the key to successful management. Dreissenids impact so much of the lake-wide nutrient cycle and so it is imperative to understand how they may change with varying conditions for appropriate management actions in the future.

E. Conclusions

To better understand the environmental and habitat constraints on the spatial distribution of dreissenid mussels, I explored three different modeling approaches and ten different environmental explanatory variables. The best performing model for these data was a model predicting probability of presence across the lake. Although this does not entirely satisfy the main objective of this research, it is still a meaningful result and poses new questions about the relationship between mussels and fine sediments with different compositions, and the influence of winter (February) maximum bottom temperature. Although EBK did not give any information about explanatory variable effects, this could still be a valuable model for visualizing dreissenid distributions across the lake, especially if more dreissenid sample sites were available. While the dreissenid biomass models did not perform well, these models still helped support several hypotheses regarding the importance of various explanatory variables for dreissenid biomass.

These models can be used to inform large-scale habitat management, linking environmental factors and dreissenid spatial distributions. These results showed the overall importance of many explanatory variables, but especially the importance of winter bottom temperature and bathymetry. These results also provided support for modeling the different basins of Lake Huron separately in the future. Following the six next steps above, these variables could be explored further, and more complete spatial datasets obtained. Being able to predict and map dreissenid abundance and distributions through both space and time will allow us to get a better idea of the big picture for how the Great Lakes and *Dreissena* spp. populations can best be managed now and in the future.

E. References

Ackerman, J. D. (2014). Role of fluid dynamics in dreissenid mussel biology. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.471-486). Boca Raton, FL: CRC Press.

Ackerman, J.D. (1999). Effect of velocity on the filter feeding of dreissenid mussels (*Dreissena polymorpha* and *Dreissena bugensis*): implications for trophic dynamics. *Canadian Journal of Fisheries and Aquatic Sciences*, 56: 1551-1561.

Akobeng, A.K. (2007). Understanding diagnostic tests 3: receiver operating characteristic curves. *Acta Paediatrica*, 96, 644-647.

Alderstein, S., Nalepa, T. F., Vanderploeg, H. A., & Fahnenstiel, G. L. (2014). Trends in phytoplankton, zooplankton, and macroinvertebrates in Saginaw Bay relative to zebra mussel (*Dreissena polymorpha*) colonization: a generalized linear model approach. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.525-544). Boca Raton, FL: CRC Press.

Assis, J.F., Claro, B., Ramos, A., Boavida, J., Serrao, E.A. (2013). Performing fish counts with a wide-angle camera, a promising approach reducing divers' limitations. *Journal of Experimental Marine Biology and Ecology*, 445, 93-98.

Baldwin, B. S., Mayer, M. S., Dayton, J., Pau, N., Mendilla, J., Sullivan, M., Moore, A., Ma, A., & Mills, E. L. (2002). Comparative growth and feeding in zebra and quagga mussels (*Dreissena polymorpha* and *Dreissena bugensis*): implications for North American lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, 59(4), 680-694. https://doi.org/10.1139/F02-043.

Barbiero, R.P., Lesht, B.M., Warren, G.J., Rudstam, L.G., Watkins, J.M., Reavie, E.D., Kovalenko, K.E., Karatyaev, A.Y. (2018). A comparative examination of recent changes in nutrients and lower food web structure in Lake Michigan and Lake Huron. *Journal of Great Lakes Research*, 44(4), 573-589.

Barnard, C., Martineau, C., Frenette, J., Dodson, J.J., Vincent, W.F. (2006). Trophic position of zebra mussel veligers and their use of dissolved organic carbon. *Limnology and Oceanography*, 51(3), 1473-1484.

Becker, E.A., Carretta, J.V., Forney, K.A., Barlow, J. Brodie, S., Hoopes, R…. Hazen, E.L. (2020). Performance evaluation of cetacean species distribution models developed used generalized additive models and boosted regression trees. *Ecology and Evoluion*, 10(12), 5759-5784.

Beletsky, D. & Schwab, D. (2001). Modeling circulation and thermal structure in Lake Michigan: annual cycle and interannual variability. *Journal of Geophysical Research*, 106, 19745-19771.

Bennion, D.H., Warner, D.M., Esselman, P.C., Hobson, B., Kieft, B. (2019). A comparison of chlorophyll a values obtained from an autonomous underwater vehicle to satellite-based measures for Lake Michigan. *Journal of Great Lakes Research*, 45(4), 726-734.

Benson, A. J. (2014). Chronological history of zebra and quagga mussels (Dreissenidae) in North America, 1988-2010. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.9-32). Boca Raton, FL: CRC Press.

Berkman, P.A., Haltuch, M.A., Tichich, E., Garton, D.W., Kennedy, G.W., Gannon, J.E., … Liebenthal, D.L. (1998). Zebra mussels invade Lake Erie muds. *Nature*, 393, 27-28.

bij de Vaate, A., van der Velde, G., Leuven, R., & Heiler, K. (2014). Spread of the quagga mussel (*Dreissena rostriformis bugensis*) in Western Europe. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.83-92). Boca Raton, FL: CRC Press.

Boavida, J., Assis, J., Reed, J., Serrao, E.A., Goncalves, J.M.S. (2015). Comparison of small remotely operated vehicles and diver-operated video of circalittoral benthos. *Hydrobiologia*, 766, 247-260.

Bocaniov, S.A., Barton, D.R., Schiff, S.L., Smith, R.E.H. (2013) Impact of tributary DOM and nutrient inputs on the nearshore ecology of a large, oligotrophic lake (Georgian Bay, Lake Huron, Canada). Aquatic Sciences, 75:321-332.

Boeckman, C. J. & Bidwell, J. R. (2014). Density, growth, and reproduction of zebra mussels (*Dreissena polymorpha*) in two Oklahoma reservoirs. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.369-382). Boca Raton, FL: CRC Press.

Boegehold, A.G. (2018). Sensitivity of quagga mussels (*Dreissena rostriformis bugensis*) to cyanobacteria at multiple life history stages. *Wayne State University. ProQuest Dissertations Publishing*. 108285512.

Boegehold, A.G., Johnson, N.S., Kashian, D.R. (2019). Dreissenid (quagga and zebra mussel) veligers are adversely affected by bloom forming cyanobacteria. *Ecotoxicology and Environmental Safety*, 182, 109426.

Boegman, L., Loewen, M. R., Culver, D. A., Hamblin, P. F., & Charlton, M. N. (2008). Spatial-Dynamic Modeling of Algal Biomass in Lake Erie: Relative Impacts of Dreissenid Mussels and Nutrient Loads. *Journal of Environmental Engineering*, *134*(6):456–468. https://doi.org/10.1061/(ASCE)0733-9372(2008)134:6(456)

Bold, H.C. & Wynne, M.J. (1985). Introduction to the algae. Prentice-Hall Inc., Englewood Cliffs, N.J.

Bootsma, H. A. & Liao, Q. (2014). Nutrient cycling by dreissenid mussels: controlling factors and ecosystem response. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.555-574). Boca Raton, FL: CRC Press.

Bootsma, H. A., Rowe, M. D., Brooks, C. N., & Vanderploeg, H. A. (2015). Commentary: The need for model development related to Cladophora and nutrient management in Lake Michigan. *Journal of Great Lakes Research*, 41:7–15. https://doi.org/10.1016/j.jglr.2015.03.023

Brighenti, L.S., Staehr, P.A., Brandao, L.P.M., Barbosa, F.A.R., Bezerra-Neto, J.F. (2018). Importance of nutrients, organic matter and light availability on epilimnetic metabolic rates in a mesotrophic tropical lake. *Freshwater Biology*, 63(9), 1143-1160.

Bril, J. S., Langenfeld, K., Just, C. L., Spak, S. N., & Newton, T. J. (2017). Simulated mussel mortality thresholds as a function of mussel biomass and nutrient loading. *PeerJ*, *5*, e2838. https://doi.org/10.7717/peerj.2838

Bunnell, D.B., Barbiero, R., Ludsin, S., Madenjian, C., Warren, G., Dolan, D., … Weidel, B. (2014). Changing ecosystem dynamics in the Laurentian Great Lakes: bottom-up and top-down regulation. *BioScience*, 64(1), 26-39.

Burlakova, L.E., Barbiero, R.P., Karatayev, A.Y., Daniel, S.E., Hinchey, E.K., Warren, G.J. (2018). The benthic community of the Laurentian Great Lakes: analysis of spatial gradients and temporal trends from 1998 to 2014. *Journal of Great Lakes Research*, 44(4), 600-617.

Burrows, M. J. (2014). Invading dreissenid mussels transform the 100-year-old international joint commission. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.187-194). Boca Raton, FL: CRC Press.

Campbell, S.D. & Chow-Fraser, P. (2018) Models to predict total phosphorus concentrations in coastal embayments of eastern Georgian Bay, Lake Huron. *Canadian Journal of Fisheries and Aquatic Sciences*, 75(11). http://dx.doi.org.proxy.lib.umich.edu/10.1139/cjfas-2017-0095

Cha, Y., Stow, C.A., Bernhardt, E.S. (2012). Impacts of dreissenid mussel invasions on chlorophyll and total phosphorus in 25 lakes in the USA. *Freshwater Biology*, 1-15. doi:10.1111/fwb.12050

Chakraborti, R. K., Madon, S., Kaur, J., & Gabel, D. (2014). Management and control of dreissenid mussels in water infrastructure facilities of the Southwestern United States. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.215-242). Boca Raton, FL: CRC Press.

Chase, M.E. & Bailey, R.C. (1999). The ecology of the zebra mussel (*Dreissena polymorpha*) in the lower Great Lakes of North America: I. population dynamics and growth. 25(1): 107-121.

Clarke, M. & McMahon, R.F. (1996). Effects of hypoxia and low-frequency agitation on byssogenesis in the freshwater mussel *Dreissena polymorpha* (Pallas). *Biological Bulletin*, 191: 413-420.

Claxton, W.T. & Mackie, G.L. (1998). Seasonal and depth variations in gametogenesis and spawning of *Dreissena polymorpha* and *Dreissena bugensis* in eastern Lake Erie. *Canadian Journal of Zoology*, 76, 2010-2019.

Coakley, T.P., Brown, G.L., Ioannou, S.E., Charlton, M.N. (1997). Colonization patterns and densities of zebra mussel *Dreissena* in muddy offshore sediments of western Lake Erie, Canada. *Water Air Soil Pollution*, 99, 623-632.

Conn, D. B., Lucy, F. E., & Graczyk, T. K. (2014). Dreissenid mussels as sentinel biomonitors for human and zoonotic pathogens. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.265-272). Boca Raton, FL: CRC Press.

Czarnoleski, M. & Muller, T. (2014). Antipredator strategy of zebra mussels (*Dreissena polymorpha*): from behavior to life history. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.345-358). Boca Raton, FL: CRC Press.

Dawson, H. A. & Allison, M. (2021) Requirements for autonomous underwater vehicles (AUVs) for scientific data collection in the Laurentian Great Lakes: a questionnaire survey. Journal of Great Lakes Research, 47(1):259-265.

DeVanna, K. M., Schloesser, D. W., Bossenbroek, J. M., & Mayer, C. M. (2014). Interactions between exotic ecosystem engineers (*Dreissena* spp.) and native burrowing mayflies (*Hexagenia* spp.) in soft sediments of Western Lake Erie. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.611-622). Boca Raton, FL: CRC Press.

Elith, J., Leathwick, J.R., Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77(4): 802-813.

Fahnenstiel, G.L., Pothoven, S., Vanderploeg, H., Klarer, D., Nalepa, T., Scavia, D. (2010b). Recent changes in primary production and phytoplankton in the offshore region of southeastern Lake Michigan. *Journal of Great Lakes Research*, 36, 20-29.

Fanslow, D.L., Nalepa, T.F., & Johengen, T.H. (2001). Seasonal changes in the respiratory electron transport system (ETS) and respiration of the zebra mussel, *Dreissena polymorpha* in Saginaw Bay, Lake Huron. *Hydrobiologia*, 448: 61-70.

Fernald, R. T. & Watson, B. T. (2014). Eradication of Zebra Mussels (*Dreissena polymorpha*) from Millbrook Quarry, Virginia: Rapid Response in the Real World. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.195-214). Boca Raton, FL: CRC Press.

Fichot, C.G., Matsumoto, K., Holt, B., Gierach, M.M., Tokos, K.S. (2019). Assessing change in the overturning behavior of the Laurentian Great Lakes using remotely sensed lake surface water temperatures. *Remote Sensing of Environment*, 235, 111427. Doi: 10.1016/j.rse.2019.111427

Fielding, A.H. & Bell, J.F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24(1), 38-49.

Forsyth, D., Riseng, C.M., Wehrly, K.E., Mason, L.A., Gaiot, J., Hollenhorst, T., … Whelan, G. (2016). The Great Lakes Hydrography Dataset: consistent, binational watershedsfor the Laurentian Great Lakes Basin. *Journal of the American Water Resources Association*, doi: 10.1111/1752-1688.12435.

Foster, S.D., Hosack, G.R., Hill, N.A., Barrett, N.S., Lucieer, V.L. (2014). Choosing between strategies for designing surveys: autonomous underwater vehicles. *Methods in Ecology and Evolution*, 5(3), 287-297.

Francoeur, S.N., Pillsbury, R.W., Lowe, R.L. (2015). Benthic algal response to invasive mussels in Saginaw Bay: a comparison of historical and recent data. *Journal of Freshwater Ecology*, 30(4), 463-477.

Francoeur, S.N., Winslow, K.P., Miller, D., Stow, C.A., Cha, Y., Peacor, S.D. (2014). Spatial and temporal patterns of macroscopic benthic primary producers in Saginaw Bay, Lake Huron. *Journal of Great Lakes Research*, 40(1): 53-63.

French, J.R.P., Schaeffer, J.S., Roseman, E.F., Kiley, C.S., Fouilleroux, A. (2009). Abundance and distribution of benthic macroinvertebrates in offshore soft sediments in Western Lake Huron, 2001-2007. *Journal of Great Lakes Research*, 35(1), 120-127.

Garton, D. W., McMahon, R., & Stoeckmann, A. M. (2014). Limiting environmental factors and competitive interactions between zebra and quagga mussels in North America. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.383-402). Boca Raton, FL: CRC Press.

Ghedotti, M.J., Smihula, J.C., Smith, G.R. (1995). Zebra mussel predation by round gobies in the laboratory. *Journal of Great Lakes Research*, 21: 665-669. doi:10.1016/S0380-1330(95)71076-0

Glyshaw, P.W., Riseng, C.M., Nalepa, R.F., Pothoven, S.A. (2015). Temporal trends in condition and reproduction of quagga mussels (*Dreissena rostriformis bugensis*) in southern Lake Michigan. Journal of Great Lakes Research, 41 (Supplement 3), 16–26

Griffiths, R. W., Schloesser, D. W., & Kovalak, W. P. (2014). Early responses to zebra mussels in the Great Lakes: a journey from information vacuum to policy and regulation. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.135-176). Boca Raton, FL: CRC Press

Grigorovich, I.A., Korniushin, A.V., Gray, D.K., Duggan, I.C., Colautti, R.I., MacIsaac, H.J. (2003). Lake Superior: an invasion coldspot? *Hydrobiologia*, 499, 191-210.

Gutiérrez, D., Harcourt, J., Díez, S.B., Illán, J.G., Wilson, R.J. (2013). Models of presence-absence estimate abundance as well as (or even better than) models of abundance: the case of the butterfly *Parnassius apollo*. *Landscape Ecology*, 28, 401-413.

Hallstan, S., Johnson, R.K., Sandin, L. (2013). Effects of dispersal-related factors on species distribution model accuracy for boreal lake ecosystems. *Diversity*, 5, 393-408. doi:10.3390/d5020393

Haltuch, M.A. & Berkman, P.A. (2000). Geographic information system (GIS) analysis of ecosystems invasion: exotic mussels in Lake Erie. *Limnology and Oceanography*, 45, 1778-1787.

Hasler, C.T., Leathers, J., Ducharme, A., Casson, N.J. (2019). Biological effects of water velocity and other hydrodynamic characteristics of flow on dreissenid mussels. *Hydrobiologia*, 837, 1-14.

He, C., Zhang, L., DeMarchi, C., Croley, T.E. II. (2014). Estimating point and non-point source nutrient loads in the Saginaw Bay watersheds. *Journal of Great Lakes Research*, 40, 11-17.

Higgins, S. N. (2014). Meta-analysis of dreissenid effects on freshwater ecosystems. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.487-494). Boca Raton, FL: CRC Press.

Hincks, S.S. & Mackie, G.L. (1997). Effects of pH, calcium, alkalinity, hardness and chlorophyll on the survival, growth, and reproductive success of zebra mussel (*Dreissena polymorpha*) in Ontario lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, 54, 2049-2057.

Houghton, C. J. & Janssen, J. (2014). Variation in predator-prey interactions between round gobies and dreissenid mussels. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.359-368). Boca Raton, FL: CRC Press.

Houghton, J. S., Paddock, R., & Janssen, J. (2014). Invasion of quagga mussels (*Dreissena rostriformis bugensis*) to the mid-lake reef complex. In T.F. Nalepa & D.W.

Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.65-70). Boca Raton, FL: CRC Press.

Hutchinson, G. E. 1957. Concluding remarks. Cold Spring Harbor Symp. Quant. Bioi. 22:415-27

Jiang, L., Xia, M., Ludsin, S. A., Rutherford, E. S., Mason, D. M., Marin Jarrin, J., & Pangle, K. L. (2015). Biophysical modeling assessment of the drivers for plankton dynamics in dreissenid-colonized western Lake Erie. *Ecological Modelling*, 308:18–33. https://doi.org/10.1016/j.ecolmodel.2015.04.004

Johengen, T. H., Vanderploeg, H. A., & Liebig, J. R. (2014). Effects of algal composition, seston stoichiometry, and feeding rate of zebra mussel (*Dreissena polymorpha*) nutrient excretion in two Laurentian Great Lakes. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.445-460). Boca Raton, FL: CRC Press.

Jonasson, P.M. (2004). Benthic Invertebrates. In P.E. O'Sullivan & C.S. Reynolds (Ed.), *The Lakes Handbook: Limnology and Limnetic Ecology*. https://doi.org/10.1002/9780470999271.ch12

Karatayev, A. Y., Burlakova, L.E., Mehler, K., Hinchey, E.K., Wick, M., Bakowska, M., Mrozinska, N. (2021b). Rapid assessment of *Dreissena* population in Lake Erie using underwater videography. *Hydrobiologia*. https://doi.org/10.1007/s10750-020-04481-x

Karatayev, A. Y., Karatayev, V. A., Burlakova, L.E., Mehler, K., Rowe, M.D., Elgin, A.K., Nalepa, T.F. (2021a). Lake morphometry determines *Dreissena* invasion dynamics. *Biological Invasions*. https://doi.org/10.1007/s10530-021-02518-3

Karatayev, A.Y., Karatayev, V.A., Burlakova, L.E., Rowe, M.D., Mehler, K., Clapsadl, M.D. (2018a). Food depletion regulates the demography of invasive dreissenid mussels in a stratified lake. *Limnology and Oceanography*, 63(5), 2065-2079.

Karatayev, A. Y., Mehler, K., Burlakova, L. E., Hinchey, E. K., & Warren, G. J. (2018b). Benthic video image analysis facilitates monitoring of Dreissena populations across

spatial scales. *Journal of Great Lakes Research*, 44(4):629–638. https://doi.org/10.1016/j.jglr.2018.05.003

Karatayev, A.Y., Burlakova, L.E., Padilla, D.K. (2015). Zebra versus quagga mussels: a review of their spread, population dynamics, and ecosystem impacts. *Hydrobiologia*, 746: 97-112.

Karatayev, A.Y., L. E. Burlakova, K. Mehler, S. E. Daniel, A. K. Elgin, and T. F. Nalepa. 2020. Lake Huron Benthos Survey Cooperative Science and Monitoring Initiative 2017. Technical Report. USEPA-GLRI GL00E02254. Great Lakes Center, SUNY Buffalo State, Buffalo, NY. Available at: http://greatlakescenter.buffalostate.edu/sites/greatlakescenter.buffalostate.edu/files/upl oads/Documents/Publications/LakeHuronBenthosSurveyCSMI2017FinalReport.pdf

Kashian, D. R. & Ram, J. L. (2014). Chemical regulation of dreissenid reproduction. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.461-470). Boca Raton, FL: CRC Press.

Kirkendall, D.S., Bunnell, D.B., Armenio, P.M., Eaton, L.A., Trebitz, A.A., Watson, N.M. (2021). Spatial and temporal distributions of *Dreissena* spp. veligers in Lake Huron: Does calcium limit settling success? *Journal of Great Lakes Research*. https://doi.org/10.1016/j.jglr.2021.04.001

Kobak, J. (2014). Behavior of juvenile and adult zebra mussels (*Dreissena polymorpha*). In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.331-344). Boca Raton, FL: CRC Press.

Koutnik, M.A. & Padilla, D.K. (1994). Predicting the spatial distribution of *Dreissena polymorpha* (zebra mussel) among inland lakes of Wisconsin: modeling with a GIS. *Canadian Journal of Fisheries and Aquatic Sciences*, 51, 1189-1196.

Krivoruchko, K. (2012). Empirical Bayesian kriging. *ArcUser Fall 2012.*

Kumar, R., Varkey, D., & Pitcher, T. (2016). Simulation of zebra mussels (Dreissena polymorpha) invasion and evaluation of impacts on Mille Lacs Lake,

Minnesota: An ecosystem model. *Ecological Modelling*, 331:68–76. https://doi.org/10.1016/j.ecolmodel.2016.01.019

Lam, K., Shin, P.K.S., Bradbeer, R., Randall, D., Ku, K.K.K., Hodgson, P., Cheung, S.G. (2006). A comparison of video and point intercept transect methods for monitoring subtropical coral communities.

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, 34(1), 1-17.

Lavrentyev, P.J., Vanderploeg, H.A., Franze, G., Chacin, D.H., Liebig, J.R., Johengen, T.H. (2014). Microzooplankton distribution, dynamics, and trophic interactions relative to phytoplankton and quagga mussels in Saginaw Bay, Lake Huron. *Journal of Great Lakes Research*, 40(1), 95-105.

Leach, A. & Leach, H. (2020). Characterizing the spatial distributions of spotted lanternfly (Hemiptera: Fulgoridae) in Pennsylvania vineyards. *Scientific Reports*, 10, 20588.

Leshkevich, G.A., Schwab, D.J., Muhr, G.C. (1993). Satellite environmental monitoring of the Great Lakes: a review of NOAA's Great Lakes CoastWatch program. *Photogrammetric Engineering and Remote Sensing*, 59(3), 371-379.

Lewandowski, K. & Stanczykowska, A. (2014). Summary of zebra mussel (*Dreissena polymorpha*) in polish lakes over the past 50 years with emphasis on the Masurian Lakes (Northeastern Poland). In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.103-116). Boca Raton, FL: CRC Press.

Lietz, J. E., J. R. Kelly, J. V. Scharold, P. M. Yurista. 2015. Can a rapid underwater video approach enhance the benthic assessment capability of the national coastal condition assessment in the Great Lakes? Environ. Manag. 55:1446–1456.

Lucy, F. E., Burlakova, L. E., Karatayev, A. Y., Mastitsky, S. E., & Zanatta, D. T. (2014). Zebra mussel impacts on unionids: a synthesis of trends in North America and

Europe. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.623-646). Boca Raton, FL: CRC Press.

Manel, S., Williams, H.C., Ormerod, S.J. (2002). Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology*, 38(5), 921-931.

Marsden, J. E., Stangel, P., & Shambaugh, A. D. (2014). Influence of environmental factors on zebra mussel population expansion in Lake Champlain, 1994-2010. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.33-54). Boca Raton, FL: CRC Press.

Mayer, C. M., Burlakova, L. E., Elkov, P., Fitzgerald, D., Karatayev, A. Y., Ludsin, S. A., Millard, S., Mills, E. L., Ostapenya, A. P., Rudstam, L. G., Zhu, B., & Zhukova, T. V. (2014). Benthification of freshwater lakes: exotic mussels turning ecosystems upside down. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.575-586). Boca Raton, FL: CRC Press.

McMahon, R.F. (1996). The physiological ecology of the zebra mussel, *Dreissena polymorpha*, in North America and Europe. *American Zoology*, 36: 339-363.

McMahon, R.F. (2012). Quagga/zebra mussel life history. [PowerPoint Slides] https://seagrant.oregonstate.edu/sites/seagrant.oregonstate.edu/files/invasive-species/mussel-workshop-2012/ag_ais_workshop_quagga_zebra_life_history_mcmahon_082212.pdf

Mehler, K., Burlakova, L.E., Karatayev, A.Y., Biesinger, Z., Valle-Levinson, A., Castiglione, C., Gorsky, D. (2018). Sonar technology and underwater imagery analysis can enhance invasive *Dreissena* distribution assessment in large rivers. *Hydrobiologia*, 810, 119-131.

Millie, D.F., Fahnenstiel, G.L, Dyble, J., Pigg, R., Rediske, R., Klarer, D.M., Litaker, R.W., Tester, P.A. (2008). Influence of environmental conditions on late-summer cyanobacterial abundance in Saginaw Bay, Lake Huron. *Aquatic Ecosystem Healther & Management*, 2, 196-205. https://doi-org.proxy.lib.umich.edu/10.1080/14634980802099604

Minchin, D. & Zaiko, A. (2014). Variability of zebra mussel (*Dreissena polymorpha*) impacts in the Shannon River system, Ireland. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.587-598). Boca Raton, FL: CRC Press.

Mulcan, A., Mitsova, D., Hindle, T., Hanson, H., Coley, C. (2015). Marine benthic habitats and seabed suitability mapping for potential ocean current energy sitting offshore southeast Florida. *Journal of Marine Science and Engineering*, 3, 276-298.

Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics*, 33, 341-365.

Nalepa, T. F., Riseng, C. M., Elgin, A. K., Lang, G. A. (2018). Abundance and distribution of benthic macroinvertebrates in the Lake Huron system: Saginaw Bay, 2006-2009, and Lake Huron, including Georgian Bay and North Channel, 2007 and 2012. NOAA Technical Memorandum GLERL-172. NOAA, Great Lakes Environmental Research Laboratory, Ann Arbor, MI, 54 pp. (DOI:10.25923/aqe2-ma69)

Nalepa, T. F., Fanslow, D. L., Pothoven, S. A., Foley, A. J., & Lang, G. A. (2007). Long-term trends in benthic macroinvertebrate populations in Lake Huron over the past four decades. *Journal of Great Lakes Research*, 33:421-436.

Nalepa, T. F., Pavlova, V., Wong, W. H., Janssen, H., Houghton, J. S., & Mabrey, K. (2014). Variation in the quagga mussel (*Dreissena rostriformis bugensis*) with emphasis on the deepwater morphotype. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.315-330). Boca Raton, FL: CRC Press.

Nalepa, T.F., Quigley, M.A., Ziegler, R.W. (1988). Sampling efficiency of the ponar grab in two different benthic environments. *Journal of Great Lakes Research*, 14(1):89-93.

Nguyen, T. D., Hawley, N., & Phanikumar, M. S. (2017). Ice cover, winter circulation, and exchange in Saginaw Bay and Lake Huron: Ice cover, winter circulation, and exchange. *Limnology and Oceanography*, *62*(1):376–393. https://doi.org/10.1002/lno.10431

Nguyen, T.D., Thupaki, P., Anderson, E.J., Phanikumar, M.S. (2014). Summer circulation and exchange in the Saginaw Bay-Lake Huron system. *Journal of Geophysical Research: Oceans*, 119(4), 2713-2734.

Nicholls, K.H. (1998). El Nino, ice cover, and Great Lakes phosphorus: implications for climate warming. *Limnological Oceanography*, 43, 715-719.

Novaczek, E., Devillers, R., & Edinger, E. (2019). Generating higher resolution regional seafloor maps from crowd-sourced bathymetry. *PLoS ONE*, 14(6): e0216792. https://doi.org/10.1371/journal.pone.0216792

O'Neill, C.R. Jr. (1994). The introduction and spread of the zebra mussel in North America. Proceedings of the Fourth International Zebra Mussel Conference, Madison Wisconsin, March 1994.

Onikura, N., Inui, R., Oikawa, S., Kawamoto, T. (2014). Performance of a species distribution model based on presence/absence for various population sizes in the amphidromous fish species *Sicyopterus japonicus*. *Ichthyological Research*, 61, 90-96.

Orlova, M. I. (2014). Origin and spread of quagga mussels (*Dreissena rostriformis bugensis*) in Eastern Europe with notes on size. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.93-102). Boca Raton, FL: CRC Press.

Paterson, G., Di Pierdomenico, L.L., Haffner, G.D. (2020). Basin-specific pollutant bioaccumulation patterns define Lake Huron forage fish. *Environmental toxicology and chemistry*, 39(9), 1712-1723.

Pavlova, V. & Izyumov, Y. (2014). Morphological variability of *Dreissena polymorpha* and *Dreissena rostriformis bugensis* (Mollusca: Bivalvia). In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.287-314). Boca Raton, FL: CRC Press.

Pelletier, D., Leleu, K., Mou-Tham, G., Guillemot, N., Chabanet, P. (2011). Comparison of visual census and high definition video transects for monitoring coral reef fish assemblages. *Fisheries Research*, 107, 84-93.

Perez-Fuentetaja, A., Clapsadl, M.D., Getchell, R.G., Bowser, P.R., Lee, W.T. (2011). *Clostridium botulinum* type E in Lake Erie: inter-annual differences and role of benthic invertebrates. *Journal of Great Lakes Research*, 37(2), 238-244. doi:10.1016/j.jglr.2011.03.013

Peyer, S.M., McCarthy, A.J., Lee, C.E. (2009). Zebra mussels anchor byssal threads faster and tighter than quagga mussels in flow. *Journal of Experimental Biology*, 212: 2027-2036.

Phillips, J.C., McKinley, G.A., Bennington, V., Bootsma, H.A., Pilcher, D.J. (2015). The potential for CO $<$inf$>$2-induced acidification in freshwater: a great lakes case study. *Oceanography*, 28(2), 136-145. http://doi.org/10.5670/oceanog.2015.37

Piacenza, S. E., Barner, A. K., Benkwitt, C. E., Boersma, K. S., Cerny-Chipman, E. B., Ingeman, K. E., Kindinger, T. L., Lee, J. D., Lindsley, A. J., Reimer, J. N., Rowe, J. C., Shen, C., Thompson, K. A., Thurman, L. L., & Heppell, S. S. (2015). Patterns and Variation in Benthic Biodiversity in a Large Marine Ecosystem. *PLOS ONE*, *10*(8): e0135135. https://doi.org/10.1371/journal.pone.0135135

Pilcher, D. J., McKinley, G. A., Bootsma, H. A., & Bennington, V. (2015). Physical and biogeochemical mechanisms of internal carbon cycling in Lake Michigan. *Journal of Geophysical Research: Oceans*, *120*(3):2112–2128. https://doi.org/10.1002/2014JC010594

Pilz, J. & Spock, G. (2007). Why do we need and how should we implement Bayesian kriging methods. *Stochastic Environmental Research and Risk Assessment*, 22, 621-532.

Pilcher, D. J., McKinley, G. A., Kralj, J., Bootsma, H. A., & Reavie, E. D. (2017). Modeled sensitivity of Lake Michigan productivity and zooplankton to changing nutrient concentrations and quagga mussels: Lake Michigan Mussels and Nutrients. *Journal of*

Geophysical Research: Biogeosciences, *122*(8):2017–2032. https://doi.org/10.1002/2017JG003818

Pollick, S. (2014). One reporter's perspective on the invasion of dreissenid mussels in North America: reflections on the news media. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.117-134). Boca Raton, FL: CRC Press.

Pothoven, S. A. & Fahnenstiel, G. L. (2014). Lake Michigan after dreissenid mussel invasion: water quality and food web changes during the late winter/spring isothermal period. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.545-554). Boca Raton, FL: CRC Press.

Quinn, A., Gallardo, B., Aldridge, D.C. (2013). Quantifying the ecological niche overlap between two interacting invasive species: the zebra mussel (*Dreissena polymorpha*) and the quagga mussel (*Dreissena rostriformis bugensis*). *Aquatic Conservation: Marine and Freshwater Ecosystems*, 24(3), 324-337.

Quinn, N.P. & Ackerman, J.D. (2011). The effect of near-bed turbulence on sperm dilution and fertilization success of broadcast-spawning bivalves. *Limnology and Oceanography Fluids and Environments*, 1(1): 176-193.

Raes, N. & ter Steege, H. (2007). A null-model for significance testing of presence-only species distribution models. *Ecography*, 30, 727-736. https://doi.org/10.1111/j.2007.0906-7590.05041.x

Ramcharan, C.W., Padilla, D.K., & Dodson, S.I. (1992). Models to predict potential occurrence and density of the zebra mussel, *Dreissena polymorpha*. *Canadian Journal of Fisheries and Aquatic Science*, 49, 2611-2620.

Rennie, M. D. (2014). Context-dependent changes in Lake Whitefish populations associated with dreissenid invasion. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.661-680). Boca Raton, FL: CRC Press.

Ridgeway, G. (2020) Generalized boosted models: a guide to the gbm package. https://cran.r-project.org/web/packages/gbm/vignettes/gbm.pdf

Riseng, C.M., Wehrly, K.E., Wang, L., Rutherford, E.S., McKenna, Jr., J.E., Johnson, L.B., Mason, L.A., Castiglione, C., Hollenhorst, T.P., Sparks-Jackson, B.L., Sowa, S.P. (2018). Ecosystem classification and mapping of the Laurentian Great Lakes. *Canadian Journal of Fisheries and Aquatic Sciences 75:1693-1712*, https://doi.org/10.1139/cjfas-2017-0242

Roditi, H.A., Fisher, N.S., Sanudo-Wilhelmy, S.A. (2000) Uptake of dissolved organic carbon and trace elements by zebra mussels. Nature, 407:78-80.

Rowe, M. D., Anderson, E. J., Wang, J., & Vanderploeg, H. A. (2015a). Modeling the effect of invasive quagga mussels on the spring phytoplankton bloom in Lake Michigan. *Journal of Great Lakes Research*, *41*:49-65. https://doi.org/10.1016/j.jglr.2014.12.018

Rowe, Mark D., Anderson, E. J., Vanderploeg, H. A., Pothoven, S. A., Elgin, A. K., Wang, J., & Yousef, F. (2017). Influence of invasive quagga mussels, phosphorus loads, and climate on spatial and temporal patterns of productivity in Lake Michigan: A biophysical modeling study: Lake Michigan productivity model. *Limnology and Oceanography*, *62*(6): 2629–2649. https://doi.org/10.1002/lno.10595

Rowe, Mark D., Obenour, D. R., Nalepa, T. F., Vanderploeg, H. A., Yousef, F., & Kerfoot, W. C. (2015b). Mapping the spatial distribution of the biomass and filter-feeding effect of invasive dreissenid mussels on the winter-spring phytoplankton bloom in Lake Michigan. *Freshwater Biology*, *60*(11): 2270–2285. https://doi.org/10.1111/fwb.12653

Schol, A., Kirchesch, V., Bergfeld, T., Scholl, F., Borcherding, J., Muller, D. (2002). Modelling the chlorophyll a concent of the River Rhine—Interrelation between riverine algal production and population biomass of grazers, rotifers and the zebra mussel, *Dreissena polymorpha*. *International Review of Hydrobiology*, 87(2-3): 295-317.

Schwalb, A. N., Bouffard, D., Boegman, L., Leon, L., Winter, J. G., Molot, L. A., & Smith, R. E. H. (2015). 3D modelling of dreissenid mussel impacts on phytoplankton in a large lake supports the nearshore shunt hypothesis and the importance of wind-driven

hydrodynamics. *Aquatic Sciences*, *77*(1), 95–114. https://doi.org/10.1007/s00027-014-0369-0

Shen, C., Liao, Q., & Bootsma, H. A. (2020). Modelling the influence of invasive mussels on phosphorus cycling in Lake Michigan. *Ecological Modelling*, *416*, 108920. https://doi.org/10.1016/j.ecolmodel.2019.108920

Shen, C., Liao, Q., Bootsma, H. A., Troy, C. D., & Cannon, D. (2018). Regulation of plankton and nutrient dynamics by profundal quagga mussels in Lake Michigan: A one-dimensional model. *Hydrobiologia*, *815*(1):47–63. https://doi.org/10.1007/s10750-018-3547-6

Sierszen, M.E., Peterson, G.S., Scharold, J.V. (2006). Depth-specific patterns in benthic-planktonic food web relationships in Lake Superior. *Canadian Journal of Fisheries and Aquatic Sciences*, 63(7), 1496-1503.

Sly, P.G. & Sandilands, R.G. (1988). Geology and environmental significance of sediment distributions in an area of the submerged Niagara escarpment, Georgian Bay. *Hydrobiologia*, 163, 47-76.

Stewart, K. M. (2014). Replacement of zebra mussels by quagga mussels in the Erie Canal, New York, USA. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.55-64). Boca Raton, FL: CRC Press.

Stoeckmann, A. (2003) Physiological energetics of Lake Erie dreissenid mussels: a basis for the displacement of *Dreissena polymorpha* by *Dreissena bugensis*. Canadian Journal of Fisheries and Aquatic Sciences, 60:126-134.

Stow, C.A., Dyble, J., Kashian, D.R., Johengen, T.H., Winslow, K.P., Peacor, S.D. … Miller, D. (2014) Phosphorus targets and eutrophication objectives in Saginaw Bay: a 35 year assessment. Journal of Great Lakes Research, 40(1):4-10.

Strayer, D. L. & Malcom, H. M. (2014). Long-term change in the Hudson River's bivalve populations: a history of multiple invasions (and recovery?). In T.F. Nalepa & D.W.

Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.71-82). Boca Raton, FL: CRC Press.

Strayer, D.L., Adamovich, B.V., Adrian, R., Aldridge, D.C., Balogh, C., Burlakova, L.E., … Jeschke, J.M. (2019). Long-term population dynamics of dreissenid mussels (*Dreissena polymorpha* and *D. rostriformis*): a cross-system analysis. *Ecosphere*, 10(4), e02701.

Summers, R.B., Thorp, J.H., Alexander, J.E. Jr., Fell, R.D. (1996). Respiratory adjustment of dreissenid mussels (*Dreissena polymorpha* and *Dreissena bugensis*) in response to chronic turbidity. *Canadian Journal of Fisheries and Aquatic Sciences*, 53, 1626-1631.

Thorp, J.H., Alexander, J.E. Jr., Bukaveckas, B.L, Cobbs, G.A., Bresko, K.L. (1998). Responses of Ohio River and Lake Erie dreissenid molluscs to changes in temperature and turbidity. *Canadian Journal of Fisheries and Aquatic Sciences*, 55, 220-229.

Todem, D., Kim, K., Hsu, W. (2016). Marginal mean models for zero-inflated count data. *Biometrics*, 72(3), 986-994.

Troy, C., Cannon, D., Liao, Q., & Bootsma, H. (2016). Logarithmic velocity structure in the deep hypolimnetic waters of Lake Michigan. *Journal of Geophysical Research: Oceans*, *121*(1): 949-965. https://doi.org/10.1002/2014JC010506

Trumpickas, J., Shuter, B.J., Minns, C.K. (2009). Forecasting impacts of climate change on Great Lakes surface water temperatures. *Journal of Great Lakes Research*, 35(2009), 454-463. doi: 10.1016/j.jglr.2009.04.005

Tuchman, N.C., Burks, R.L, Call, C.A., Smarrelli, J. (2004). Flow rate and vertical position influence ingestion rates of colonial zebra mussels (*Dreissena polymorpha*). *Freshwater Biology*, 49: 191-198.

Tyner, E.H., Bootsma, H.A., Lafrancois, B.M. (2015). Dreissenid metabolism and ecosystem scale effects as revealed by oxygen consumption. *Journal of Great Lakes Research*, 41(5), 27-37.

Tzortziou, M., Zeri, C., Dimitriou, E., Ding, Y., Jaffe, R., Anagnostou, E., Pitta, E., Mentzafou, A. (2015) Colored dissolved organic matter dynamics and anthropogenic influences in a major transboundary river and its coastal wetland. Limnology and Oceanography, 60(4):1222-1240.

Vanderploeg, H.A., Bunnell, D. B., Carrick, H. J., & Höök, T. O. (2015). Complex interactions in Lake Michigan's rapidly changing ecosystem. *Journal of Great Lakes Research*, 41:1–6. https://doi.org/10.1016/j.jglr.2015.11.001

Vanderploeg, H.A., Liebig, J.R., Nalepa, T.F., Fahnenstiel, G.L, Pothoven, S.A. (2010). *Dreissena* and the disappearance of the spring phytoplankton bloom in Lake Michigan. *Journal of Great Lakes Research*, 36, 50-59.

Vanderploeg, H.A., Wilson, A. E., Johengen, T. H., Bressie, J. D., Sarnelle, O., Liebig, J. R., Robinson, S. D., & Horst, G. P. (2014). Role of selective grazing by dreissenid mussels in promoting toxic *microcystis* blooms and other changes in phytoplankton composition in the Great Lakes. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.509-524). Boca Raton, FL: CRC Press.

Vanderploeg, H.A., Liebig, J.R., Gluck, A.A. (1996). Evaluation of different phytoplankton for supporting development of zebra mussels larvae (*Dreissena polymorpha*): the importance of size and polyunsaturated fatty acid content. *Journal of Great Lakes Research*, 22(1), 36-45.

Wacker, A. & von Elert, E. (2003). Food quality controls reproduction of the zebra mussel (*Dreissena polymorpha*). *Oecologia*, 135, 332-338.

Walz, N. (1978). The energy balance of the freshwater mussel *Dreissena polymorpha* Pallas in laboratory experiments and in Lake Constance. 3. Growth under standard conditions. *Arch Hydrobiologia Supplement*, 55, 121-141.

Wang, L., Riseng, C.M., Mason, L. A., Wehrly, K. E., Rutherford, E. S., McKenna, J. E., Catiglione, C., Johnson, L. B., Infante, D. M., Sowa, S., Robertson, M., Schaeffer, J., Khoury, M., Gaiot, J., Hollenhorst, T., Brooks, C., & Coscarelli, M. (2015). A spatial

classification and database for management, research, and policy making: the Great Lakes aquatic habitat framework. *Journal of Great Lakes Research*, 1-13.

Ward, J. M. & Ricciardi, A. (2014). Impacts of *Dreissena* on benthic macroinvertebrate communities: predictable patterns revealed by invasion history. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.599-610). Boca Raton, FL: CRC Press.

Wehrly, K.E., Wang, L., Infante, D., Cooper, A., Geddes, C., Stanfield, L., Rutherford, E. (2013). Landscape change and its influences on aquatic habitat and fisheries in the Great Lakes Basin. In: Taylor, W., Leonard, N., Lynch, A. (Eds.), Great Lakes Fisheries Policy and Management: A Binational Perspective, 2nd edition Michigan State University Press, East Lansing.

Williamson, F. & Ozersky, T. (2019). Lake characteristics, population properties and invasion history determine impact of invasive bivalves on lake nutrient dynamics. *Ecosystems*, 22:1721-1735.

Winslow, K.P., Francoeur, S.N., Peacor, S.D. (2014). The influence of light and nutrients on benthic filamentous algal growth: a case study of Saginaw Bay, Lake Huron. *Journal of Great Lakes Research*, 40(1), 64-74. https://doi.org/10.1016/j.jglr.2014.01.012

Wong, W. H., Holdre, G. C., Tietjen, T., Gerstenberger, S., Moore, B., Turner, K., & Wilson, D. C. (2014). Effects of invasive quagga mussels (*Dreissena rostriformis bugensis*) on chlorophyll and water clarity in Lakes Mead and Havasu of the Lower Colorado River Basin, 2007-2009. In T.F. Nalepa & D.W. Schloesser (Ed.), *Quagga and Zebra Mussels* (2nd ed., pp.495-508). Boca Raton, FL: CRC Press.

Wright, D. A., Setzler-Hamilton, E. M., Magee, J. A., Kennedy, V. S., & McIninch, S. P. (1996). Effect of Salinity and Temperature on Survival and Development of Young Zebra (Dreissena polymorpha) and Quagga (Dreissena bugensis) Mussels. *Estuaries*, *19*(3): 619. https://doi.org/10.2307/1352522

Xueying, M., Xiufeng, Z., Kassam, S., Rudstam, L.G. (2016). Will the displacement of zebra mussels by quagga mussels increase water clarity in shallow lakes during summer? Results from a mesocosm experiment. *PLoS One*, 11(12). e0168494. https://doi.org/10.1371/journal.pone.0168494

Young, M.A., Treml, E.A., Beher, J., Fedle, M., Gorfine, H., Miller, A.D., … Ierodiaconou, D. (2020). Using species distribution models to assess the long-term impacts of changing oceanographic conditions on abalone density in southeast Australia. *Ecography*, 43(7), 1052-1064.

Yu, H., Cooper, A.R., & Infante, D.M. (2020). Improving species distribution model predictive accuracy using species abundance: application with boosted regression trees. *Ecological Modelling*, 432, 109202.

F. Appendix 1

explanatory variable	source
January vertical water temperature 0-20m depth	GLAHF
January vertical water temperature 20-40m depth	GLAHF
January vertical water temperature 40-214m depth	GLAHF
February vertical water temperature 0-20m depth	GLAHF
February vertical water temperature 20-40m depth	GLAHF
February vertical water temperature 40-214m depth	GLAHF
March vertical water temperature 0-20m depth	GLAHF
March vertical water temperature 20-40m depth	GLAHF
March vertical water temperature 40-214m depth	GLAHF
April vertical water temperature 0-20m depth	GLAHF
April vertical water temperature 20-40m depth	GLAHF
April vertical water temperature 40-214m depth	GLAHF
May vertical water temperature 0-20m depth	GLAHF
May vertical water temperature 20-40m depth	GLAHF
May vertical water temperature 40-214m depth	GLAHF
June vertical water temperature 0-20m depth	GLAHF
June vertical water temperature 20-40m depth	GLAHF
June vertical water temperature 40-214m depth	GLAHF
July vertical water temperature 0-20m depth	GLAHF
July vertical water temperature 20-40m depth	GLAHF

July vertical water temperature 40-214m depth	GLAHF
August vertical water temperature 0-20m depth	GLAHF
August vertical water temperature 20-40m depth	GLAHF
August vertical water temperature 40-214m depth	GLAHF
September vertical water temperature 0-20m depth	GLAHF
September vertical water temperature 20-40m depth	GLAHF
September vertical water temperature 40-214m depth	GLAHF
October vertical water temperature 0-20m depth	GLAHF
October vertical water temperature 20-40m depth	GLAHF
October vertical water temperature 40-214m depth	GLAHF
November vertical water temperature 0-20m depth	GLAHF
November vertical water temperature 20-40m depth	GLAHF
November vertical water temperature 40-214m depth	GLAHF
December vertical water temperature 0-20m depth	GLAHF
December vertical water temperature 20-40m depth	GLAHF
December vertical water temperature 40-214m depth	GLAHF
Annual vertical water temperature 0-20m depth	GLAHF
Annual vertical water temperature 20-40m depth	GLAHF
Annual vertical water temperature 40-214m depth	GLAHF
Vertical water temperature cumulative degree days	GLAHF
Surface water temperature cumulative degree days	GLAHF
January surface water temperature	GLAHF
February surface water temperature	GLAHF

March surface water temperature	GLAHF
April surface water temperature	GLAHF
May surface water temperature	GLAHF
June surface water temperature	GLAHF
July surface water temperature	GLAHF
August surface water temperature	GLAHF
September surface water temperature	GLAHF
October surface water temperature	GLAHF
November surface water temperature	GLAHF
December surface water temperature	GLAHF
Annual surface water temperature	GLAHF
bathymetry	GLAHF
tributary influence	GLAHF
January maximum bottom temperature	NOAA-GLCFS
January average bottom temperature	NOAA-GLCFS
January minimum bottom temperature	NOAA-GLCFS
January standard deviation bottom temperature	NOAA-GLCFS
February maximum bottom temperature	NOAA-GLCFS
February average bottom temperature	NOAA-GLCFS
February minimum bottom temperature	NOAA-GLCFS
February standard deviation bottom temperature	NOAA-GLCFS
March maximum bottom temperature	NOAA-GLCFS
March average bottom temperature	NOAA-GLCFS

March minimum bottom temperature	NOAA-GLCFS
March standard deviation bottom temperature	NOAA-GLCFS
April maximum bottom temperature	NOAA-GLCFS
April average bottom temperature	NOAA-GLCFS
April minimum bottom temperature	NOAA-GLCFS
April standard deviation bottom temperature	NOAA-GLCFS
May maximum bottom temperature	NOAA-GLCFS
May average bottom temperature	NOAA-GLCFS
May minimum bottom temperature	NOAA-GLCFS
May standard deviation bottom temperature	NOAA-GLCFS
June maximum bottom temperature	NOAA-GLCFS
June average bottom temperature	NOAA-GLCFS
June minimum bottom temperature	NOAA-GLCFS
June standard deviation bottom temperature	NOAA-GLCFS
July maximum bottom temperature	NOAA-GLCFS
July average bottom temperature	NOAA-GLCFS
July minimum bottom temperature	NOAA-GLCFS
July standard deviation bottom temperature	NOAA-GLCFS
August maximum bottom temperature	NOAA-GLCFS
August average bottom temperature	NOAA-GLCFS
August minimum bottom temperature	NOAA-GLCFS
August standard deviation bottom temperature	NOAA-GLCFS
September maximum bottom temperature	NOAA-GLCFS

September average bottom temperature	NOAA-GLCFS
September minimum bottom temperature	NOAA-GLCFS
September standard deviation bottom temperature	NOAA-GLCFS
October maximum bottom temperature	NOAA-GLCFS
October average bottom temperature	NOAA-GLCFS
October minimum bottom temperature	NOAA-GLCFS
October standard deviation bottom temperature	NOAA-GLCFS
November maximum bottom temperature	NOAA-GLCFS
November average bottom temperature	NOAA-GLCFS
November minimum bottom temperature	NOAA-GLCFS
November standard deviation bottom temperature	NOAA-GLCFS
December maximum bottom temperature	NOAA-GLCFS
December average bottom temperature	NOAA-GLCFS
December minimum bottom temperature	NOAA-GLCFS
December standard deviation bottom temperature	NOAA-GLCFS
winter maximum bottom temperature	NOAA-GLCFS
winter average bottom temperature	NOAA-GLCFS
winter minimum bottom temperature	NOAA-GLCFS
winter standard deviation bottom temperature	NOAA-GLCFS
spring maximum bottom temperature	NOAA-GLCFS
spring average bottom temperature	NOAA-GLCFS
spring minimum bottom temperature	NOAA-GLCFS
spring standard deviation bottom temperature	NOAA-GLCFS

summer maximum bottom temperature	NOAA-GLCFS
summer average bottom temperature	NOAA-GLCFS
summer minimum bottom temperature	NOAA-GLCFS
summer standard deviation bottom temperature	NOAA-GLCFS
fall maximum bottom temperature	NOAA-GLCFS
fall average bottom temperature	NOAA-GLCFS
fall minimum bottom temperature	NOAA-GLCFS
fall standard deviation bottom temperature	NOAA-GLCFS
April CDOM	NOAA-ERDDAP
May CDOM	NOAA-ERDDAP
June CDOM	NOAA-ERDDAP
July CDOM	NOAA-ERDDAP
August CDOM	NOAA-ERDDAP
September CDOM	NOAA-ERDDAP
October CDOM	NOAA-ERDDAP
spring and summer CDOM	NOAA-ERDDAP
April chlorophyll-a	NOAA-ERDDAP
May chlorophyll-a	NOAA-ERDDAP
June chlorophyll-a	NOAA-ERDDAP
July chlorophyll-a	NOAA-ERDDAP
August chlorophyll-a	NOAA-ERDDAP
September chlorophyll-a	NOAA-ERDDAP
October chlorophyll-a	NOAA-ERDDAP

spring and summer chlorophyll-a	NOAA-ERDDAP
July dissolved organic carbon	NOAA-ERDDAP
August dissolved organic carbon	NOAA-ERDDAP
September dissolved organic carbon	NOAA-ERDDAP
October dissolved organic carbon	NOAA-ERDDAP
distance from the St. Mary's River	created
distance from the Saginaw River	created