Spring 5-6-2013

Ground-Truthing Maxent in East Texas Rivers

David F. Ford

Recommended Citation

Ford, David F., "Ground-Truthing Maxent in East Texas Rivers" (2013). *Biology Theses.* Paper 12.
http://hdl.handle.net/10950/110

GROUND-TRUTHING MAXENT IN EAST TEXAS RIVERS

By

DAVID F. FORD

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science
Department of Biology

Lance Williams, Ph. D., Committee Chair

College of Arts and Sciences

The University of Texas at Tyler
May 2013

The University of Texas at Tyler
Tyler, Texas

This is to certify that the Master's Thesis of

DAVID F. FORD

has been approved for the thesis requirement on
April 9, 2013
For the Master of Science degree

Approvals:

Thesis Chair: Lance Williams, Ph. D.

Josh Banta

Member: Josh Banta, Ph. D.

Member: John Placyk, Ph. D.

Chair, Department of Biology

Martin

Dean, College of Arts and Sciences

Acknowledgments

For their continued patience and support I thank all of my friends. I know I have not always been the easiest to deal with when deadlines have approached, but you have all always been there for me when I needed you to be there. I thank my mother for all of the editing that she has been a part of. A special thanks goes to my father for insisting that mussels are “cool”, for teaching me my mussels, and for all of those early morning coffee trips we’ve had where we talked science. Without those, my thesis would not be nearly as good as it is.

I thank my committee members, Drs. John Placyk, and Josh Banta, for their support and guidance throughout this endeavor. A second thanks is extended to Dr. Banta, for all of his help with the ARCGIS and Maxent programs. I appreciate all the times you let me pop into your lab, and bug you with questions. I know we both wanted to pull out our hair a couple of times from them. I also thank my committee chair Dr. Lance Williams, for letting me into his lab, and for all of his help with my research.

My research would not have been possible without all of my field hands, most especially Joel Hunt, Judith Bilyea, and Kaitlyn Pettingill. These three people went out with me almost every day, and spent more hours in the field with me than they probably wanted to. I know there were several trips that we would all like to forget. I also would have no pictures without Judith, who was willing to take pictures on almost every trip I took.

Ashley Walters, I thank you for teaching me my mussels, and for giving me a starting point with my research. Without you, I would have no place from which my research could have started. Your help with Maxent, mussels, and my research in

general has been invaluable. I literally could not have finished my Master’s without you. Thank you as well for always letting me call you anytime I had questions, which we both know was a lot.

I thank Zara Environmental for the data that they let me incorporate into my thesis. Their data let my thesis extend into the Trinity River, and helped to make this project more valuable. Finally I thank the US Fish and Wildlife Service for their funding of this project.

Table of Contents

Appendices

List of Tables

List of Figures

Abstract

GROUND-TRUTHING MAXENT IN EAST TEXAS RIVERS

David F. Ford

Thesis Chair: Lance Williams, Ph. D.

The University of Texas at Tyler
May 2013

Unionid mussels are a guild of freshwater, sedentary filter-feeders, which play a critical role in freshwater systems. Mussels are currently experiencing a global decline in both species richness and abundance, due to invasive species, human alteration of water systems, and climate change. In North America, which is considered to have the highest global diversity of bivalve species, native mussels are currently declining rapidly with at least 37 species considered to already be extinct. If extant mussel species are to be preserved, then it is vital that conservation efforts be prioritized towards areas in which they are likely to be found. Often this is done through ecological niche models. Maxent for example uses the principle of maximum entropy on presence data and environmental variables to create a suitability score for a particular area, and is one of the most widely used of the ecological niche modeling programs. It has been used to make maps predicting the suitability scores for multiple species but very little ground-truthing to see if the maps are assigning the correct scores has been conducted. Ground-truthing was done by sampling at 138 sites throughout East Texas. These sites had been assigned suitability scores from a previous study, and the mussels found at the site were compared to those that were predicted to be there by Maxent. The new maps created by Maxent were compared to the original maps to see if new occurrence points added to the predictive ability of the maps by looking at the test AUC and test gain values. The influences of new data on Maxent's predictive ability for finding a particular mussel species at a site, and the number of mussels found at a site were also investigated by linear and logistic regression. Additional occurrence points were found to significantly improve the predictive maps for the Triangle Pigtoe, the Texas Heelsplitter, and the Southern Hickorynut, and all maps were found to accurately predict locations for mussels. Maxent's predictive ability via their suitability scores was improved for all species with additional occurrence points. However, for almost all of the species looked at, there was a data cap, which was a point at which additional data no longer improved the models. This suggests that the amount of data necessary to make accurate maps may not be as large as originally thought, and when trying to conserve an organism this could be important.

Chapter One: Life History and General Background

Extinctions are natural events that are caused by disturbances, and are counterbalanced by the evolution and migration of new species (Barnosky et al., 2011). The current wave of extinctions, however, is thought to have been brought about by human demands on the planet. Since the 1600's, there have been 490 animal and 580 plant extinctions worldwide. More than half of these have occurred in the last 100 years (May et al., 1995). Although the rate of the current extinction wave cannot be completely determined, it is clear that species are going extinct extremely rapidly, and that if trend continues, it will erode human well-being considerably, possibly leading to another mass extinction (Barnosk et. al, 2011; Jansson, 2009; Regan et al., 2001; Balmford et al., 2005; Lydeard et al., 2004).

Human modification of the land and climate are greatly influencing the habitat and biodiversity in ecosystems both worldwide and in North America. Some of the most vulnerable areas are freshwater communities such as those found in streams or rivers. The biodiversity in freshwater areas is decreasing by as much as 4% of their total biodiversity per decade worldwide (Poole & Downing, 2004; Vaughn, 2010). Freshwater communities depend upon the landscape around them, which constrain and control the local factors of the stream (Frissel et al., 1986; Smiley et al., 2008). In considerable danger in water systems are freshwater bivalves which are currently among the most endangered organisms in the world (Williams et al., 1992).

Worldwide, freshwater mussels are one of the most threatened groups of animals on the planet, and their decline is readily apparent in North American waters (Williams et al., 1992; Strayer et al., 2004; Galbraith & Vaughn, 2011). Freshwater mussels are divided into 19 families worldwide. Most of these families include only a few genera or species. The family Unionoida is one of the larger families, including six genera, and 850 species, and includes almost all freshwater bivalve families except the *Sphaeriidae* (Bogan, 2008; Haag, 2012). Unionids are found worldwide except for in Antarctica and the Pacific Islands (Bogan, 2008).

Freshwater mussels are a large guild of sedentary, burrowing, filter-feeding bivalves, which have a suite of life history traits that cause them to be vulnerable to changes in the aquatic ecosystem (Vaughn, 2010). Mussels are mostly sedentary and are unable to move away from areas which are no longer suitable for them. Mussels cannot escape from disturbances to their habitat such as impoundments (Galbraith, et al., 2010). Impoundments are features such as dams, which alter the flow of the water, and change the sediment flow and other aquatic processes that mussels need to survive. These changes can also alter the velocity of the water which changes the amount of sediment that settles in the river, and can suffocate mussels if they cannot get above this new sediment layer (Watters, 2000). An increase of just 2.5 cm of silt can lead to the mortality of about 90% of the mussels in that area. The amount of light entering the river is also altered by impoundments, which can change the amount of algae in an area. Algae can make up a large portion of a mussel’s diet so this can have a large impact on a mussel community (Ellis, 1936). Substrates in impounded area are often changed, which can lead to new mussel species entering an area, typically those that are more silt tolerant or those which can take harsh flood events (Watters, 1999). Dams have been found to impede fish migrations from one area to another, or to cause them to take another, longer route to an area, resulting in a necessary fish not being in an area during larval release by a mussel. Thus, environmental dynamics that can alter fish migrations can alter mussel recruitment as well (Watters, 2000).

Native freshwater bivalves tend to not reproduce until they are at least seven years old, leading to long generation times and an inability to rapidly recover from disturbances and invasive species (Kacar, 2011). Invasive species typically reproduce much faster than native species, and alter the physical, chemical and biological characteristics of an area (Strayer, 2006). One species in particular, *Dreissena polymorpha*, the zebra mussel, has had a major impact on mussels native to the United States. *Dreissena polymorpha* are epifaunal, and can attach themselves to almost any hard surface with byssal threads, allowing them to colonize new areas much faster than native freshwater mussels, where they quickly become the dominant mussel species. In any area in which they can survive they can reach densities of more than 100,000/m^2 (Strayer, 1999). Larval *D. polymorpha* will settle onto any surface, including other mussels, typically in densities of about 100-200 per individual mussel, but can exceed over 10,000 *D. polymorpha* on a single mussel (Hebert et al., 1991; Ricciardi et al., 1995; Martel et al., 2001). This increase in the mass of the host mussel, limits the mussel's movements along with its ability to open and close its shell, impacting its survivability (Vaughn & Heakenkamp, 2001). *D. polymorpha* waste increases toxic chemicals in the water, often in places where native mussels already congregate (Vaughn & Heakenkamp, 2001). The biggest impact that *D. polymorpha* have on native mussels, however, is through nutrient removal. In habitats where zebra mussels are found in high concentrations, there is a large decline in phytoplankton and zooplankton, sometimes as high as 90%. Plankton are eaten by native mussels, and native mussel body condition has been found to be significantly lower in areas where *D. polymorpha* occur, from a lack of nutrient uptake (Vaughn and Heakenkamp, 2001). Body condition is lower even for mussels who do not have zebra mussels directly on them but just inhabit the same areas (Haag, 1993; Strayer, 1999). In the Great Lakes area where *D. polymorpha* first appeared in the United States, they have virtually eliminated native mussel populations, and in areas where they have since spread to, *D. polymorpha* have led to greater than 90% declines in native mussel populations. This trend in mussel decline is expected to continue as zebra mussels move southwards (Ricciardi et al., 1998).

Mussels feed by pumping of large amounts of water through their body tissues. Consequently they are exposed to contaminants that are dissolved in the water, resulting in the bioaccumulation of toxic substances in their tissues (Naimo, 1995). Toxic metals and chemicals that enter an aquatic system can be absorbed by suspended particles, which then accumulate in the sediments (Salomons et al., 1987; Tessier & Campbell, 1987). These sediments are filtered by mussels, leading to the accumulation of these toxins in the mussels (Giesy & Hoke, 1989). Toxins can have a wide range of effects depending upon their concentration, as well as the species in question; however, all have at least some kind of a negative effect on mussel survivability. High concentrations of Cd, Cu, Hg, and Zn have been found to cause mortality, alterations in weight, changes in enzyme activity, and modifications in a mussel’s normal behavior, such as reduction in their time spent filtering water (Naimo, 1995). 3-Trifluoromethyl-4-nitrophenol (TFM), which is an endocrine disruptor used to kill larval sea lamprey, *Petromyzon marinus*, in the Great Lakes region, has a narcotic effect on mussels. When exposed to the chemical, mussels will relax most of their tissues, and extend their feet and other muscles (Waller et al., 1998). DNA damage occurs in the mussel *Unio pictorum,* when found downstream of paper mills and oil refineries (Štambuk et al., 2009). The abundance of mussels found both upstream and downstream of waste water treatment plants is significantly lower than what is historically known, and is thought to be a product of the ammonia and other chemicals that are pumped into the water (Goudreau et al., 1993). Chlorine, which is added to most drinking water, increases mortality in mussels, along with reductions in their growth and the detachment of settled larvae (Khalanski & Bordet, 1980).

Mussels are also thermo-conformers, whose physiological processes depend on the temperature of the surrounding water. Temperature controls when larval mussels complete their metamorphosis to adult mussels, and there is both a high and low temperature threshold that the water must reach. When the water is too cold, the larvae do not metamorphosis, and if it is too hot the larvae simply die (Watters, & O'Dee, 1999). Thus, changes in water temperature can lead to shifts in the rates and magnitudes of the biological processes that mussels employ for survival (Spooner & Vaughn, 2008; Vaughn, et al., 2008). These changes in temperature sometimes occur because of impoundments placed upon the river (Watters, 2000).

Direct harvest has also contributed to the endangerment of freshwater mussels, again because they lead a relatively sedentary lifestyle. Humans have used mussels for food, tools, and jewelry since prehistoric times (Strayer, 2006). In North America, humans have been harvesting mussels for about 10,000 years, mostly as a food source, although they have made use of their shells in other ways (Haag, 2012). Starting about 1,500 years ago, the fossil record indicates that mussel shells became a major element of tempering clay and in pottery production (Theler, 1990; Weinstein & Dumas, 2008). Shells were also used as implements such as spoons, dippers, scrapers and fishing lures (Parmalee & Bogan, 1998; Cartwright, 2003). Mussel harvesting has continued to modern day, and is still influencing mussel populations in North America (Haag, 2012).

Mussels can be an abundant and conspicuous component of freshwater ecosystems and often compose more than 50% of the total biomass in the community, therefore the loss of mussel biomass could have a serious impact on an ecosystem (Haag, 2012). Bivalves tend to be much larger in size compared to other benthic organisms and have numerous influences on the aquatic community. They can change the sediment and nutrient levels in the water by sequestering them in their tissues and shells as they feed (Vaughn & Hakenkamp, 2001). Bivalve filtration can lead to large decreases in plankton and other particles in the water column, especially when the biomass of the bivalves is very large (Kasprzak, 1986; Kryger & RiisgaÊrd, 1988; Welker & Walz, 1998; Strayer et al., 1999). The amount of water filtered by mussels in a dense bed can sometimes even exceed the daily stream discharge for an area which can lead to `biological oligotrophication' by decreasing phytoplankton biomass and total phosphorus, and increasing water clarity (Welker & Walz, 1998). The intensification in water clarity increases the food available for other organisms by increasing the algal content in the water (Haag, 2012). Mussels can also be important source of dissolved nutrients for the water column, since they are able to translocate (feed and pump back out) and transform (change the chemical form of) nutrients in the water (Kuenzler, 1961). Mussels are also important cyclers of nitrogen in water bodies, releasing ammonia and dissolved organic nitrogen that can be taken up directly by phytoplankton and benthic algal communities (Dame, 1996; Lauritsen & Mozley, 1989). These nutrients come from a hypo-osmotic urine that consists primarily of ammonia which is produced by freshwater mussels. The excretion rate and concentration of the urine varies between species, size classes, and seasons, as well as other variables (Burton, 1983; Vaughn & Hakenkamp, 2001).

Mussels can pull up nutrients that cannot readily be accessed by other organisms from the substrate, either through filtering water internally or by deposit feeding (McMahon, 1991). Mussels filter water internally by utilizing an elongated inhalant siphon to vacuum organic detritus and bacteria from the stream or lakebed surface (Way, 1989). Deposit feeding is a primitive bivalve function that has been found to be nearly universal in mussels, and contributes to about 50% of a mussels' diet (Vaughn & Hakenkamp, 2001). Mussels accomplish deposit feeding by using cilia on their foot to collect buried organic matter from the substrate, which they then ingest via those same cilia (Way, 1989).

In stream and lake communities the biodeposition of fecal matter and pseudofecal matter by bivalves can be an important sedimentation process that places high-quality pelagic resources back into the sediment. The resources that are put back into the water column by mussel’s can influence adjacent benthic communities, resulting in changes in species composition and abundance (Reusch, et al, 1994; Vaughn & Hakenkamp, 2001). In areas where mussels are in abundance there are significant increases in the amount and concentrations of organic matter in the substrate when compared to similar areas that lack large mussel populations (Prokopovich, 1969). The biodeposition can lead to an increase in the local abundance of macroinvertebrates, especially those that feed on decaying matter such as detritivores (Vaughn & Hakenkamp, 2001).

As bivalves locomote and burrow into the substrate they can release sediments into the water body for other organisms to access. Deposit feeding leads to an increase in the oxygen penetration into the sediments, and can stimulate microbial metabolisms (Dame, 1996; Levinton, 1995). As mussels burrow into the substrate they bioturbate the sediment, or remix the sediments, leading to changes in the chemical composition of the water body. These changes can increase the sediment levels in the water, increase sediment homogenization and increase the depth that oxygen can penetrate into the substrate (McCall et al., 1979). The rates at which these changes occur per mussel depend on the size of the mussel because larger mussels can mix sediments at greater rates (McCall et al., 1995). These changes also depend on the species of the mussel. Some species move and feed more than others (McCall et al., 1995). Bivalve activity also changes the flux rates of some solutes across the sediment-water gradient. Their movements help to enhance the release of nitrates and chloride and inhibit the release of calcium carbonate from the sediments (Matisoff et al., 1985).

The physical presence of mussels can influence the water body in several ways, in addition to the biological effects that bivalves can have. Both living and dead mussel shells act as suitable substratum for benthic algae, and as habitat for epizoic and epiphytic plants and algae to colonize (Beckett et al., 1996). Mussels can act as habitats for numerous types of invertebrates and small fish (Wooton, 1992; Navarrette, 1996). Shells act as refugia for small fish to avoid predators as they hunt for food (McCall et al., 1979; Strayer et al., 1994). Mussel shells can accumulate organic matter which can be a food source for other organisms, which can lead to increases in the abundance of chironomids and other detritivores (Gosselin & Chia, 1995). Mussels, both when alive and dead, alter the flow characteristics of a water body by helping to stabilize the substrate and limiting sheer stress on other organisms. Sheer stress is the stress on an organism from water flowing over it. Mussels simply by being present in an aquatic ecosystem can change the organism composition of an area by allowing creatures which are less tolerant of sheer stress to be able to live in an area (Zimmerman & de Szalay, 2007).

Bivalves are often very long lived, and store nutrients in their tissues that they have bioaccumulated during their lifetime (Vaughn, & Hakenkamp, 2001). When they die, they release these nutrients back into the water column, which can be an important source for calcium and other nutrients (Green, 1980). A bivalve’s ability to store nutrients can cause a mussel to serve as both nutrient sources and sinks depending on the circumstances. If the population of bivalves is declining and releasing more nutrients than they take in, it is a source, putting nutrients back into the water. If the mussel biomass is growing, or if mussels are being removed from the ecosystem permanently, then they are acting as a sink, removing nutrients from the water (Vaughn, &

Hakenkamp, 2001). Thus the health of the mussel population can influence the nutrient levels of the water body. Finally, mussels serve as food for numerous organisms, including humans, other mammals, fishes, reptiles, and birds (Haag, 2012).

The loss of diverse mussel assemblages can permanently alter the functioning of an aquatic ecosystem. The rates of ecological processes performed by bivalves are linearly related to the biomass of the organisms in the area, thus a significant decline in unionid biomass, regardless of species, could result in an alteration of the ecosystem. The degree and significance of these alterations will be context-dependent on system size, system stability and bivalve biomass (Vaughn & Hakenkamp, 2001). North America is seeing major declines in its freshwater mussel diversity (Howard & Cuffey 2003). Conservation of mussels, particularly in areas of high mussel diversity and density, has become a concern in recent years. Predictive models based upon various niche dimensions of mussels can be used to discriminate those areas with the highest potential to sustain healthy mussel communities and thus promote a more efficient allocation of resources to protect these areas.

A species' geographic distribution and its various niche dimensions are important factors in the effective study and conservation of freshwater mussels. Such data can indicate areas that may need conserving (Yom-Tov & Kadmon, 1998). However, creating individual conservation plans for every species, or even just the threatened species in an ecoregion, is impractical and nearly physically impossible (Margules & Pressey, 2000). Instead, a predictive model that combines information regarding landscape characteristics and occurrence data can be created that attempts to predict where a species should occur. These species distribution models can be used to improve ongoing and future conservation activities that are focused on species recovery and habitat restoration. From a more practical standpoint, these models can be used when planning surveys to economize limited resources such as time or money by prioritizing sites based on the probability of various impacts (Yom-Tov & Kadmon, 1998). In recent years, these models have been used to predict changes in factors such as global warming, invasive species, and species reintroductions (Esselman & Allan, 2011; Marmion et al., 2009; Lenoir et al., 2011).

Species distribution models (SDM) extrapolate species distribution data in space and time, usually using a statistical model (Franklin, 2009). These models are created by taking observations of species occurrences and matching those occurrences with environmental variables thought to influence habitat suitability and thus species distribution. SDMs produce outputs that can be interpreted in two different ways. First they give estimates of the probability that a species might occur at a given unrecorded location. The second are estimates of an area's suitability for a particular species. The interpretation depends on the assumptions that are made (Segurado & Araugo, 2004).

There are several elements that make up a useable SDM. There must be a theoretical or conceptual model of the biotic and abiotic factors that control a species' distribution through space and time, at different scales, and the organismal responses for these factors (Franklin, 2009). Data are needed on the occurrence of a species in geographical space, and can be measured in presence data, habitat use data, abundance data, or any other properties that can be used to indicate the occurrence of an organism (Franklin, 2009). Next there need to be maps of environmental variables representing those factors determining habitat quality, or that are correlated with those variables (Franklin, 2009). These types of data tend to be derived from remote sensing, spatial models of environmental processes, or other sources, and stored in a GIS software package. Then there needs to be a model linking habitat requirements or species occurrence data to the environmental variables. These models can be statistical, descriptive, logical or rule-based (Burgman et al., 2005). The model needs tools, such as thresholds, weights, coefficients and others, for applying values to the mapped environmental variables to produce a map of the organism’s occurrence or suitable habitats. Finally, there needs to be data and criteria that validate the predictions and a way to interpret error or uncertainty in the analysis (Franklin, 2009). Models for SDMs can be broken into two major groups, regression models and machine learning models (Franklin, 2009).

Regression models use relatively simple mathematical equations to explain the relationships between habitat variables and organism locations, and to predict species distributions (Austin, 2002). The other major type of modeling SDMs is machine learning models. Machine learning methods use various kinds of algorithms that are used to identify the mapping function or classification rules inductively, based on training data, that the algorithm uses to build the model (Franklin, 2009). Machine learning methods are relatively new and are a rapidly growing area of eco-informatics that is concerned with identifying structure in complex, large, and typically nonlinear datasets. Machine learning techniques have been promoted in ecology and SDMs as powerful alternatives to traditional statistical modeling approaches (Olden et al., 2008; Elith, 2002; Elith et al., 2006; Wisz et al., 2008).

Maximum entropy modeling (Maxent) is a general purpose machine learning method that is one of the most commonly used methods for inferring species distributions and environmental tolerances from occurrence only data (Warren & Seifert, 2011). Maxent has been used extensively with SDMs and outperforms other machine-learning and statistical methods even with small sample sizes (Elith et al., 2006). It was originally created for use with SDMs and has several advantages over other modeling systems (Franklin, 2009). It can predict species distributions based upon only species presence data and environmental information of the study area. Maxent has been used in studies of species richness, invasive species, climate change, endemism hotspots, and almost every other type of ecological study possible. However, maps have only been validated for areas that have already been sampled. We could find no studies that investigated new areas to validate the species distributions predicted by Maxent. Most studies simply made new maps, and accepted them as correct (Franklin, 2009). Only one other study was found that used some of its data to validate the model they created. However, their original sampling was not based on any habitat suitability maps (Rodriguez-Castaneda, et al., 2012). However, because it uses an exponential model for creating its species distribution probabilities, it can give very large predicted values for conditions that are outside the range of those found in the data used to develop the model (Franklin, 2009).

As of 2009, six mussels, the Texas Pigtoe (*Fusconaia askewi*), the Triangle Pigtoe (*Fusconaia lamanensis*), the Sandbank Pocketbook (*Lampsilis satura*), the Southern Hickorynut *(Obovaria jacksoniana*), the Louisiana Pigtoe (*Pleurobema riddelli*), and the Texas Heelsplitter (*Potamilus amphichaenus*) were all listed as threatened for the state of Texas. Conservation of these species has therefore become more important leading to the creation of habitat suitability models for these species with Maxent. Maxent suitability score maps have been created for each of these species, giving scores for all of Eastern Texas (Walters, 2013). These maps were based upon mussel and habitat data from the Sabine and Neches Rivers and predicted occurrence sites for all of East Texas.

My study had two objectives: 1) To determine if Maxent creates valid maps for threatened mussels in East Texas. Maps will be evaluated by the mussels found at the location and their suitability scores. The overall change in the maps will be looked at by their test AUCs and test gains. Comparisons will be done with each new set of data. As more data is added to the map the test AUCs and test gains should increase, indicating that the overall predictive power of the maps has increased. 2) To determine how much additional data affects the maps for mussels via their suitability scores. Suitability scores will be compared to the presence or absence of a species at a particular location, and the suitability scores will be compared to the number of individuals of a species at a particular location. Sites with higher scores will have more of a species at that site, and higher scored sites will be more likely to have a threatened species of mussel than a lower scored site. 3) To create new suitability maps for the six threatened mussel species. Maps were created using 20% of the new data each time, added to the original data. As new data is added to the maps, they should become better at predicting locations at which the mussels could be found. The scored areas in the maps are likely to shrink, as the scores will change as the maps become better at predicting. Many of the high areas will shrink and the lower scored areas will become more prominent, since many modeling systems over predict the high scores (Wisz et al., 2008).

Chapter Two: Methods

In 2012, mussels were collected in the Neches, Sabine, Angelina, and Trinity Rivers of East Texas between May and September of 2012 (Figure 2.1). Sampling locations were chosen based upon previously created habitat suitability maps for the six threatened species of mussels in East Texas (Walters, 2013). These maps were created using Maxent and were based upon previously collected mussel location and habitat data. Data for the original maps were from five sites on the Neches River and nine sites on the Sabine River (Walters, 2012). Suitability scores ranged from zero, which was a site which was predicted not to be suitable at all for the species, to one, which was a site that was predicted to be the best habitat for the species. Maps were divided into five different ranges of suitabilities based upon their suitability scores. Grids in the maps were scored as either high, low, mid high, or mid low. A uniform distribution of sampling sites were selected from the range of suitability scores for all six species. Sites were chosen via a stratified random sampling design to allow sites to be randomly chosen within the suitability scores found in the original maps created by Walters (Walters, 2013). That allowed sites to still be within a certain suitability score set, but to be randomly chosen within that set. Sites were chosen to allow at least five sampling efforts for each of the five score sections and for all six species, and to provide adequate coverage of all the major rivers in East Texas (Figure 2.1). Sites were sampled in a 50m reach containing as many geomorphic units as possible (i.e., riffle, pool, run) to be representative of the actual conditions in the area. Additional mussel data from the Sabine River were obtained from Neil Ford (Ford et al., 2009).

Each site was reached via kayak as sites tended to be some distance from bridge crossings, which had previously been sampled. At each site an initial reconnaissance of the shore was carried out for mussel shells. Only complete mussel shells were counted as dead mussels. Mussels were sampled using tactile and visual searches throughout the entire site, which provides the most accurate results for mussel species diversity, evenness, and richness (Hornbach & Deneka, 1996). Sampling continued until it was felt that the entire area had been sampled completely. Sites were thought to have been sampled completely when no new mussels were found for at least five minutes. All live unionids were collected, identified to species, enumerated, and replaced except for voucher specimens. Vouchers were retained in the University of Texas at Tyler collection. Recently dead mussels, indicated by the presence of tissue, shiny nacre, or uneroded shells were also enumerated at each site. Long deceased mussels were not included in samples because other forces, such as stream flow, can transport them from upstream locations that are outside of the sampling area, providing inaccurate evidence that the mussel was in that area.

Evaluation of the optimal sampling effort was done by breaking the sampling data into five sets with 0%, 20%, 40%, 60%, 80%, and 100% of the ground truthed data included (Appendix A). The sites in each set were randomly determined. The occurrence points from this first 20% were added to the original occurrence points used to create the first suitability maps. Occurrence points included both live and recently dead for a site. The new set of points was then run through Maxent to create new suitability maps with new values. These values were then assigned to the 80% of the data that remained. The process was repeated for each data set until all the data had been used and a final set of suitability maps had been created. Each time, the remaining data was considered to be the equivalent of sampling using these new maps since the suitability scores, AUC, and gain values were obtained from the new maps. Using the data in this way allowed the data not used to create the suitability maps to be a “new” sampling effort without actually having to obtain fresh data from the field. By breaking the data up into five different sets we obtained five different sampling efforts from only one summer of field research. The maps created with a certain amount of occurrence points could be compared to determine, to at least some extent, the amount of data needed to generate useful maps. To my knowledge this is the first study to use one data set to generate multiple sampling efforts to compare a model.

The software package Maxent was used for species distribution modeling (Dudik et al., 2010). The principal of maximum entropy states that a probability distribution with maximum entropy, i.e., the most spread out and closest to uniform, subject to known constraints, is the best approximation of an unknown distribution because it agrees with everything that is known and avoids assuming anything that is not known (Phillips et al., 2006). When this method is applied to presence-only SDMs, the pixels of the study area make up the space on which the Maxent probability distribution is defined. The information available about the target distribution of the species is presented as a set of real-valued variables, called features. These features are climatic variables, elevation, soil, and other environmental variables, and the model is constrained by these features. Species distribution is then directly modeled by estimating the density of environmental covariates conditional on species presence (Phillips et al., 2006). The output from Maxent assigns a probability to each site. These values are dependent on the number of background and occurrence sites used during the creation of the map. The values are not the probability of occurrence of a species in that particular site, but the suitability of that habitat for the species (Boyce et al., 2002).

Analysis was limited to locations falling within East Texas, including the Trinity, Sabine, Neches, and Angelina Rivers. Habitat suitability models were built separately for each species, based upon the 11 environmental layers found to influence the species by

Walters (Walters, 2012). Maxent's cross-validation option was used to assess the predictive ability and usefulness of the model (Pearson et al., 2007). The test gain and the test area under the operator receiving curve (AUC) that were produced by the software were used to determine model fit for each species. The test AUC measures the probability that a randomly chosen presence site will be scored above a randomly chosen pseudoabsence point (Fielding & Bell, 1997; Phillips & Dudik, 2008). AUCs > 0.75 are typically thought of as useful, and considered to indicate a well fit model (Elith, 2002). The test gain is the mean log probability of the occurrence samples after a constant has been removed that makes the uniform distribution have a zero gain. Suitability scores were obtained for each occurrence site based upon that species' suitability map.

Comparison of the models was done by graphing the test AUCs and test gains from each suitability map for each species, and looking for a general trend in the maps. If the models are improving with new data, then the test AUC and test gain should get larger with each data set. Determination of the suitability scores' overall ability to predict the number of individuals at a site was done via linear regression in Excel. Logistic regression was used to determine if sites with higher suitability scores were more likely to have a threatened mussel species than sites with lower scores. Percent contribution of each variable for each data set was examined to determine if there were any changes in the trends of the variables from the data found by Walters (Walters, 2012).

Chapter Three: Results

A total of 139 different sites were sampled throughout East Texas, from the Trinity, Sabine, Neches, and Angelina Rivers (Appendix A) (Figure 2.1). An additional 1,474 *F. askewi* mussels were collected thoughout the sample area, at 65 new occurrence sites. Adding this to the original 80 occurrence sites used to create the first habitat suitability map (Walters et al., 2012), resulted in a total of 145 occurrence sites for this species. We also found an additional 164 *F. lananensis* at 18 new occurrence sites, 118 *L. Satura* at 36 new occurrence sites, 14 *O. jacksoniana* at 5 new occurrence sites, 32 *P. amphichaenus* at 23 new occurrence sites, and 401 *P. riddelli* at 31 new occurrence sites (Table 3.1). Adding these new occurrence sites to their respective number of occurrence sites from the data used to create the original maps resulted in a total of 45 sites for *F. lananensis*, 79 sites for *L. Satura*, 17 sites for *O. jacksoniana*, 46 sites for *P. amphichaenus*, and 75 sites for *P. riddelli*. These totals were used to create the final suitability maps for each species (Walters et al., 2012) (Table 3.1).

The test gain was found to be significantly different with new data for only *F. lananensis* (p = 0.02) and the *O. jacksoniana* (p = 0.01) (Table 3.2) (Figure 3.1). The test gain increased from 1.21 to 1.95 for the *O. jacksoniana*, and decreased from 1.41 to 1.13 for *F. lananensis* (Appendix B). The test gains did not significantly change for any of the other mussel species. Test gains for *F. askewi* ranged from 1.03 at the beginning to 0.07 at the end (p = 0.07), test gains for *L. Satura* ranged from 1.25 to 1.81 (p = 0.10), test gains for *P. amphichaenus* ranged from 1.34 to 1.48 (p = 0.09), and test gains for *P. riddelli* ranged from 1.07 to 1.34 (p = 0.22) (Table 3.2). All of the test gains indicated that the models were potentially useful, although the drop in test gain for *F. lananensis* indicates that the new map is not as useful as the original map (Appendix B).

The test AUC values for both *P. amphichaenus* ($p = 4.38 \times 10^{-5}$) and *O. jacksoniana* (p = 0.01) were found to be significantly increased from the original test AUC values (Table 3.2) (Appendix B) (Figure 3.2). Test AUC values were 0.82 at the beginning for *P. amphichaenus*, and 0.90 at the end, and test AUC values for *O. jacksoniana* ranged from 0.78 to 0.83 (Table 3.2). All of the other species' test AUC scores were found to not be significantly different from the original map's test AUC scores. AUC scores for *F. askewi* ranged from 0.81 to 0.83 (p = 0.61), from 0.89 to 0.90 (p = 0.83) for *F. lananensis*, from 0.88 to 0.92 (p = 0.13) for *L. Satura*, and from 0.91 to 0.90 (p = 0.69) for *P. riddelli* (Appendix B) (Table 3.2). All of the test AUCs remained above 0.70 indicating that the maps were still accurately predicting habitat locations for the mussel species even though they did not become more accurate with additional data. The final maps for all of the species seemed to "tighten up". High suitability scored areas shrank and lower scored areas become more prominent (Figures 3.3-3.38).

Habitat variable importance remained almost identical to that reported by Walters (Walters et al., 2012). Soil remained the most important variable for all of the species, followed by vegetation and landcover usage. The other variables played very little role in the species modeling (Appendix B).

The suitability scores given by the model became better at predicting occurrence at a site of a mussel with the addition of new data in all species except for *O. jacksoniana* (p=0.04 at 0% of new data, p = 0.55 at 80% of new data). In this species the addition of data did not change how well the model predicted species occurrence at a site (Table 3.3). In all of the other species the addition of data significantly improved the model's ability to predict species occurrence. Model improvement eventually leveled out and stopped being significantly improved by the addition of new data in *L. Satura* (p =

4.89 X 10^{-5} at 0% of new data, p =0.06 at 80% of new data), *P. amphichaenus* (p = 0.08 at 0% of new data, p = 0.25 at 60% of new data), and *F. askewi* (p = 0.01 at 0% of new data, p = 0.31 at 80% of new data) (Table 3.3). Additional data continued to significantly improve the model for the Louisiana (p = 2.41 X 10^{-9} at 0% of new data, p = 9.28 X 10^{-3} at 80% of new data) and *F. lananensis* (p = 2.36 X 10^{-6} at 0% of data, p = 0.03 at 80% of data) though the significance of the new data became less with each new data set (Table 3.3).

In all species except *P. amphichaenus* (p = 0.14 at 0% of new data, p= 0.68), the addition of new data was found to significantly improve the model's prediction of high numbers of a species at sites with high suitability scores (Table 3.3). *F. askewi* improved with new data until an additional 84 sites had been added (p = 0.01 at 60% of new data), after those 84 sites new data no longer improved the model's predictive ability for (p = 0.32 at 80% of new data). *L. Satura* (p = 0.02 at 60% of new data, p = 0.12 at 80% of new data) and the *P. riddelli* (p = 0.001 at 60% of new data, p = 0.06 at 80% of new data) followed this same trend, and stopped improving after 84 sites had been reached (Table 3.3). New data continued to significantly improve the model's scores for *F. lananensis* (p = 0.39 at 0% of new data, p = 0.02 at 80% of new data) (Table 3.3).

Chapter Four: Discussion

A species can occur at a site that satisfies three sets of considerations. First, the physical characteristics of the environment must be suitable. Second, the correct suite of positive and negative interactor species must be present. Finally a site must be accessible for dispersal and colonization by the species in question (Owens, et al., 2011). Species distribution modeling systems typically use the first set of considerations in their predictions by using the abiotic variables for a species to predict new locations in which the species will be found based upon these variables. These results indicate that Maxent can accurately make these predictions.

The results from the field sampling indicate that my Maxent models accurately predict areas with similar habitats to known occurrence points for mussels in East Texas, which indicates that the models successfully predicted locations at which threatened mussels could be found. All models also improved in their predictive ability with additional data, either via an improvement in their total maps seen through an increase in their test gains, test AUCs, or the predictive ability of their suitability scores. Higher scored sites are more likely to have a threatened species than lower scored sites. More individuals of a species were also found at the higher scored sites. However no species was found to improve in all of my metrics (Table 3.2) (Table 3.3).

The test AUC only changed for *O. jacksoniana*, and the test gain only changed for *O. jacksoniana* and *F. lananensis*. *O. jacksoniana* had the least amount of data used to create the original map, almost 1/3 of the next lowest number of occurrence sites. Any new data will have an impact on this small sample size. *F. lananensis*, which had the second smallest number of occurrence sites in the original maps, had a decreasing test gain. *F. lananensis* are thought to occur only in the Angelina and lower Neches Rivers, and all new occurrence points for the species support that distribution. These new points would lead to a lowering of suitability scores in areas that are not in these two rivers, such as the Trinity River, and these changes result in the lowering of the test gain seen in this species (Appendix B). No overall change is seen in the test AUCs or test gains for the other species, but when you look at the maps, changes can be identified. In all six species the higher scored areas shrank in the amount of area they occupied. The lowest scored sites also shrank in the amount of area that it occupied. The middle scored areas increased in size (Figures 3.1-3.36). Shrinkage of the high and low scored areas indicates that Maxent over and underscores many areas. New data improves its ability to accurately predict new habitat locations.

In all species, more individuals were found at sites with higher suitability scores, and this trend continued as more data were added to the species, and scores become better at predicting the numbers of organisms of a particular species found at a site as they increase. However, there is a data plateau for each species, where new data no longer have an effect. The more data that were already used to create the original maps, the better those maps were at predicting new locations, and the less of an effect any new data will have, and the sooner the plateau appears. These plateaus are different for each species and depend on the number of occurrence sites that were used to create the original maps, the number of new occurrence points added with the new data, and the total number of new mussels that were found of that species.

No plateau in the effect of sampling effort on musseling maps was found for *O. jacksoniana* or *F. lananensis* which were the two species with the lowest original number of occurrence points, and the lowest number of new occurrence points. Suitability scores continued to improve in predictive power with new data for these species, and likely would have continued to improve if even more data had been available. *F. askewi*, *L.*

Satura, and *P. riddelli* all had suitability scores which increased in predictive ability with new data, until another 80 data points were added. These three species had fairly high numbers of occurrence sites constituting their original maps, and had the largest number of new occurrence points from the new data. Their maps were already better at predicting suitability scores than *O. jacksoniana* or *F. lananensis*, and so new data have less of a positive impact. *P. amphichaenus* had one of the lowest sets of starting occurrence points, and had very few new occurrence points found. Almost all of the original points were from the Sabine River, and this species is thought to be a possible endemic to that river body. After an additional 56 points were added, new data no longer affected the suitability scores. The original map was already a good predictor for the suitability scores for this species, and there were not enough new sites added to increase the effectiveness of the suitability scores. Additionally, the majority of the new sampling efforts were conducted outside of the Sabine River, and mostly included areas that did not have Heelsplitters (Figure 2.1). Had more sites been sampled in the Sabine River, or more Heelsplitters found in other areas, the suitability scores may have continued to increase in their predictive ability with more data for this species.

The suitability score was also a good predictor for whether or not a species would be present or absent at a site. Higher scored sites were more likely to have a threatened mussel species than lower scored sites. The predictive ability of the suitability scores tended to increase with more data, though eventually a plateau was reached, just as when the suitability score was used to predict the number of mussels at a site. The cap for a species depends on the number of occurrence sites used to make the original map and the number of new occurrence sites, compared to the total number of new sites. Maps that were made with larger numbers of occurrence sites have lower plateaus, and new data will not have as much of an effect on it. The suitability scores for *F. lananensis* and *P. riddelli* continued to improve in predictive ability as more data were added. The suitability scores for this species would have likely continued to increase had more data been available and plateaued eventually. Both of these species had large numbers of new location points discovered during this work, with the number of new *P. riddelli* occurrence sites almost twice the original amounts. The addition of new data continued to improve the suitability scores predictive ability for these two species. *F. askewi* and *L. Satura* both stopped improving in their scores after about 80 new sites had been added. The new data did not improve their scores very much because both species started with fairly large sample sizes, and so the new data had a smaller effect on their maps. *P. amphichaenus* score stopped improving after an additional 56 sites, likely because most of the sampling was done outside of the Sabine River. The small amount of sampling that was done in the Sabine River is likely what improved the model's scores (Table 3.1) *O. jacksoniana* suitability scores with regards towards presence and absence did not increase with new data, and no pattern was found for the species. Most of the sites that were sampled did not contain this species, including sites that had high suitability scores, and some of the sites that did have *O. jacksoniana* had low suitability scores. All locations that this species was found were in the Neches River, although it was random where it was found.

Many of the places that were sampled in this study are those that are thought to not contain mussels, such as tributaries or small channels of a river. Typically it is thought that mussels only occur in the main channels of a river (Ford et al., 2009). However, this study shows that mussels can be found in these smaller locations, and in deeper areas than was originally thought to contain mussels. Future mussel researchers may want to look in these smaller locations along with the more typical areas when sampling.

Finally it should be remembered that although species distribution modeling represents a scientifically important area in species conservation, it is still only a model of possible locations in which a species could live. Owing to the small sample sizes used to generate species distribution models, it would be unreasonable to expect these models to describe the complete realized niche of a mussel species (Owens et al., 2011), let alone the fundamental niche. Fish are important biotic components of mussel distributions. Unionids experience an obligate parasitic larval stage where they attach to a fish or salamander host after release from the adult mussel. Some species are able to parasitize a wide variety of fish species while others can use only a few closely related species (Trdan & Hoeh, 1982; Zale & Neves, 1982; Yeager & Saylor, 1995). Integrating information regarding fish host data through the identification of potential glochidia host relationships into my ecological niche models may provide a better understanding of the geographic distribution of these six East Texas mussels and improve the models. Future research should incorporate other abiotic and biotic factors that influence the fish hosts for these mussels into the ecological niche models.

In this study, I showed that Maxent creates accurate predictive maps for the six threatened species of mussels in East Texas and that the suitability scores from Maxent became more accurate at predicting with more data. Given the high degree of imperilment of freshwater mussel species and their ecosystems worldwide, Maxent species distribution maps could play a significant role in future conservation efforts. Protected areas have been noted as a critical strategy for conservation of organisms, and should be an important strategy to protect biodiversity (Groom, et al., 2006). The information provided from species distribution maps may aid in field surveys and allocation of conservation resources by providing valuable biogeographical information that will help in planning land use management around existing populations, locating new populations, identifying top-priority survey sites, or setting priorities to areas to restore to natural habitats (Kumar and Stohlgran, 2009; Raxworthy et al., 2003; Bourg et al., 2005).

Literature Cited

Anthony, J. L.; and Downing, J. A., 2001. Length-specific growth rates in freshwater mussels (*Bivalvia: Unionidae*): Extreme longevity or generalized growth cessation? Freshwater Biology, 46: 1349-1359.

Arbuckle, K. E.; and Downing, J. A., 2002. Freshwater mussel abundance and species richness: GIS relationships with watershed land use and geology. Canadian Journal of Fisheries and Aquatic Science, 59: 310-316.

Austin, M. P., 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modeling. Ecological Modeling, 157: 101-118.

Barnosky, A. D.; Matzke, N.; Tomiya, S.; Wogan, G. O. U.; Swartz, B.; Quental, T. B.; Marshall, C.; McGuire, J. L.; Lindsey, e. L.; Maguire, K. C.; Mersey, B.; and Farrer, E. A., 2011. Has the Earth's sixth mass extinction already arrived? Nature, 471: 51-57.

Beckett, D. C.; Green, B. W.; Thomas, S. A.; and Miller, A. C., 1996. Epizoic invertebrate communities on upper Mississippi River unionid bivalves. American Midland Naturalist, 135: 102-114.

Bogan, A. E., 2008. Global diversity of freshwater mussels (Mollusca, *Bivalvia*) in freshwater. Hydrobiologia, 595: 139–147.

Bourg, N. A.; McShea, W. J.; and D. E., 2005. Putting a CART before the search: successful habitat prediction for a rare forest herb. Ecology, 86: 2793-2804.

Boyce, M. S.; Vernier, P. R.; Nielsen, S. E.; and Schmiegelow, F. K. A., 2002. Evaluating resource selection functions. Ecological Modeling, 157: 281-300.

Burton, R. F., 1983. Ionic regulation and water balance. In: The Mollusca, Vol. 5, Physiology Part 2 (Eds A.S.M. Saleuddin & K.M. Wilbur): 291-352. Academic Press, New York.

Cartwright, C., 2003. Unusual use of freshwater mussel (Unio sp.) shells during the Early Bronze Age at Tell es-Sa’idiyeh, Jordan. Environmental Archaeology, 8: 85-89.

Cocker, R. E., 1919. Freshwater mussels and mussel industries of the U.S. Bulletin of the Bureau of Fisheres, 36: 13-89.

Crowell, E. F.; and Kinman, B. T., 1993. Musseling in Kentucky: The first 200 years. Conservation and management of freshwater mussels. Upper Mississippi River Conservation Committee, Rock Island, Illinois, USA.

Dame, R. F., 1996. Ecology of Marine Bivalves: an Ecosystem Approach. CRC Press, New York.

Dudik, M.; Phillips, S. J.; and Schapire, R. E., 2010. Maxent, version 3.3.3. http://www.cs.princeton.edu/~schapire/Maxent/.Accessed March 20, 2012.

Elith, J., 2002. Quantitative methods for modeling species habitat: comparative performance and an application to Australian plants. In: Quantitative methods for conservation biology (eds. Ferson S, Burgman M). Springer, New York.

Elith, J.; Graham, C.; Anderson, R.; Dudík, M.; Ferrier, S.; Guisan, A.; Hijmans, R.; Huettmann, F.; Leathwick, J.; Lehmann, A.; Li, J.; Lohmann, L.; Loiselle, B.; Manion, G.; Moritz, C.; Nakamura, M.; Nakazawa, Y.; Overton, J.; Peterson, A.; Phillips, S.; Richardson, K.; Scachetti-Pereira, R.; Schapire, R.; Soberón, J.; Williams, S.; Wisz, M.; and Zimmerman, N., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography, 29: 129-151.

Esselman, P. C.; and Allan, J. D., 2011. Application of species distribution models and conservation planning software to the design of a reserve network for the riverine fishes of northeastern Mesoamerica. Freshwater Biology, 56: 71-88.

Fielding, A. H.; and Bell, J. F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation, 24: 38-49.

Ford, N. B.; Gullett, J.; and May, M. E., 2009. Diversity and abundance of unionid mussels in three sanctuaries on the Sabine River in northeast Texas. The Texas Journal of Science, 61: 279-294.

Frissel, C. A.; Liss, W. J.; Warren, C. E.; and Hurley, M. D., 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. Environmental Management, 10: 199-214.

Galbraith, H. S.; Spooner, D. E.; and Vaughn, C. C., 2010. Synergistic effects of regional climate patterns and local water management on freshwater mussel communities. Biological Conservation, 143: 1175-1183.

Galbraith, H. S.; and Vaughn, C. C., 2011. Effects of reservoir management on abundance, condition, parasitism and reproductive traits of downstream mussels. River Research and Applications, 27: 193-201.

Giesy, J. P. and Hoke, R. A., 1989. Freshwater sediment toxicity bioassessment: rationale for species selection and test design. J. Great Lakes Res., 15: 539-569.

Goudreau, S. E.; Neves, R. J.; and Sheehan, R. J., 1993. Effects of wastewater treatment plant effluents on freshwater mollusks in the upper Clinch River, Virginia, USA. Hydrobiologia, 252: 211-230.

Green, R. H., 1980. Role of a unionid clam population in the calcium budget of a small Artic lake. Canadian Journal of Fisheries and Aquatic Sciences, 37: 219-224.

Groom, M. J.; Meffe, G. K.; and Carroll, C. R., 2006. Principles of Conservation Biology, 3rd edn. Sinauer Associates Inc., Sunderland, MA.

Haag, W. R., 2010. A hierarchical classification of freshwater mussel diversity in North America. Journal of Biogeography, 37: 12-26.

Haag, W. R., 2012. North American Freshwater Mussels: Natural History, Ecology, and Conservation. Cambridge University Press (1st Edition).

Haag, W. R.; Berg, D. J.; Garton, D. W.; and Farris, J. L., 1993. Reduced survival and fitness in native bivalves in response to fouling by the introduced zebra mussel (*Dreissena polymorpha*) in western Lake Erie. Canadian Journal of Fisheries and Aquatic Sciences, 50: 13-19.

Hebert, P. D. N.; Wilson, C. C.; Murdoch, M. H.; and Lazar, R., 1991. Demography and ecological impacts of the invading mollusc *Dreissena polymorpha.* Canadian Journal of Zoology, 69: 405-409.

Hornbach, D. J.; and Deneka, T., 1996. A comparison of a qualitative and quantitative collection method for examining freshwater mussel assemblages. Journal of North American Benthological Society, 15: 587-596.

Howard, J. K.; and Cuffey, K. M., 2003. Freshwater mussels in a California north coast range river: occurrence, distribution, and controls. Journal of the North American Benthological Society, 22: 63-77.

Kasprzak, K., 1986. Role of the Unionidae and Sphaeriidae (*Mollusca, Bivalvia*) in eutrophic Lake Zbechy and its outflow. International Revue der Gesamten Hydrobiologie, 71: 315-334.

Khalanski, M; and Bordet, F., 1980. Effects of chlorination on marine mussels. In: Jolly RL, Brungs WA, Cumming. Water chlorination: chemistry, environmental impact and health effects, 3: 557–567.

Kryger, J.; and RiisgaÊrd, H. U., 1988. Filtration rate capacities in six species of European freshwater bivalves. Oecologia, 77: 34-38.

Kuenzler, E. J., 1961. Phosphorus budget of a mussel population in a Georgia salt marsh. Limnology and Oceanography, 6: 400-415.

Kumar, S.; and Stohlgren, T. J., 2009. Maxent modeling for predicting suitable habitat for threatened and endangered tree *Canacomyrica monticola* in New Caledonia. Journal of Ecology and Natural Environment, 1: 94-98.

Lauritsen, D. D.; and Mozley, S. C., 1989. Nutrient excretion by the Asiatic clam *Corbicula fluminea*. Journal of the North American Benthological Society, 8: 134-139.

Lenoir, S.; Beaugrand, G.; and Lecuyer, E., 2011. Modeled spatial distribution of marine fish and projected modifications in the North Atlantic Ocean. Global Change Biology, 17: 115-129.

Levinton, J. S., 1995. Bioturbators as ecosystem engineers: control of the sediment fabric, inter-individual interactions, and material fluxes. In: Linking Species and Ecosystems (Eds C.G. Jones & J.H. Lawton): 29-38. Chapman & Hall, New York.

Marmion, M.; Parviainen, M.; Luoto, M.; Heikkinen, R. K.; and Thuiller, W., 2009. Evaluation of consensus methods in predictive species distribution modeling. Diversity and Distributions, 15: 59-69.

Martel, A. L.; Pathy, D. A.; Madill, J. B.; Renaud, C. B.; Dean, S. L.; and Kerr, S. J., 2001. Decline and regional extirpation of freshwater mussels (*unionidae*) in a small river system invaded by *Dreissena polymorpha*: The Rideau River, 1993-2000. Canadian Journal of Zoology, 79: 2181-2191.

Matisoff, G.; Fisher, J. B.; and Matis, S., 1985. Effect of microinvertebrates on the exchange of solutes between sediments and freshwater. Hydrobiologia, 122: 19-33.

McCall, P. I.; Tevesz, M. J. S.; and Schwelgien, S. F., 1979. Sediment mixing by *Lampsilis radiata siliquoidea* (*Mollusca*) from western Lake Erie. Journal of Great Lakes Research, 5: 105-111.

McMahon, R. F., 1991. *Mollusca: bivalvia*. In: Ecology and Classification of North American Freshwater Invertebrates (Eds J.H. Thorp & A.P. Covich): 315:390. Academic Press, New York.

Naimo, Teresa J., 1995. A review of the effects of heavy metals on freshwater mussels. Ecotoxicology, 4: 341-362.

Navarrette, A. A., 1996. Variable predation: effects of whelks on a mid-intertidal successional community. Ecological Monographs, 66: 301-321.

Neves, R. J., 1999. Conservation and commerce: Management of freshwater mussel (*Bivalvia: Unionidae*) resources in the United States. Malacologia, 41: 461-474.

Olden, J. D.; Lawler, J. J.; Poff, N. L.., 2008. Machine learning methods without tears: a primer for ecologists. The Quarterly Review of Biology, 83: 171-193.

Owens, H. L.; Bently, A. C.; and Peterson, A. T.., 2011. Predicting suitable environments and potential occurrences for coelacanths (*Latimeria* spp.). Biodiversity Conservation.

Parmalee, P. W.; and Bogan, A. E., 1998. The freshwater mussels of Tennessee. University of Tennessee Press, Knoxville, USA.

Pearson, R.G.; Raxworthy, C, J.; Nakamura, M.; and Peterson, A. T., 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. Journal of Biogeography, 34: 102-117.

Phillips, S. J.; and Dudík, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Echography, 31: 161-175.

Phillips, S. J.; Anderson, R. P.; and Schapire, R. E., 2006. Maximum entropy modeling of species geographic distributions. Ecological Modeling, 190: 231-259.

Poole, K. E.; and Downing, J. A., 2004. Relationship of declining mussel biodiversity to stream-reach and watershed characteristics in an agricultural landscape. Journal of the North American Benthological Society, 23: 114-125.

Prokopovich, N. P., 1969. Deposition of elastic sediments by clams. Journal of Sedimentary Petrology, 39: 891-901.

Raxworthy, C. J.; Martínez-Meyer, E.; Horning, N.; Nussbaum, R. A.; Schneider, G. E.; Ortega-Huerta, M. A.; and Peterson, A. T., 2003. Predicting distributions of known and unknown reptiles species in Madagascar. Nature, 426: 837-941.

Reusch, T. B. H.; Chapman, A. R. O.; and GroÈger, J. P., 1994. Blue mussels *Mytilus edulis* do not interfere with eelgrass *Zosteria marina* but fertilize shoot growth through bio-deposition. Marine Ecology Progress Series, 108: 265-282.

Ricciardi, A., et al., 1995. Predicting the intensity and impact of Dreissena infestation on native unionid bivalves from *Dreissena* field density. Canadian Journal of Fisheries and Aquatic Sciences, 52: 1449-1461.

Ricciardi A., et al., 1998. Impending extinctions of North American freshwater mussels (*Unionoida*) following the zebra mussel (*Dreissena polymorpha*) invasion. Journal of Animal Ecology, 67: 613-619.

Rodriguez-Castaneda, G.; Hof, A. R.; Jansson, R.; and Harding, L. E., 2012. Predicting the fate of biodiversity using species' distribution models: Enhancing model comparability and repeatability. PLoS ONE 7(9): e44402. doi:10.1371/journal.pone.0044402.

Salomons, W.; de Rooij, N. M.; Kerdijk, H.; and Bril, J., 1987. Sediments as a source for contaminants. Hydrobiologia, 149: 13-30.

Segurado, P.; and Araujo, M. B., 2004. An evaluation of methods for modeling species distributions. Journal of Biogeography, 31: 1555-1568.

Smiley, P. C.; and Dibble, E. D., 2008. Implications of a hierarchical relationship among channel form, instream habitat, and stream communities for restoration of channelized streams. Hydrobiologia, 548: 279-292.

Strayer, D. L., 1999. Effects of alien species on freshwater mollusks in North America. Journal of the North American Benthological Society, 18: 74-98.

Strayer, D. L., 1999. Use of flow refuges by unionid mussels in rivers. Journal of the North American Benthological Society, 18: 468–476.

Strayer, D. L., 2006. The effects of surface geology and stream size on freshwater mussel (*Bivalvia, Unionidae*) distribution in southeastern Michigan, U.S.A. Freshwater Biology, 13(3): 253-264.

Strayer, D. L.; Hunter, D. C.; Smith, L. C.; and Borg, C. K., 1994. Distribution, abundance, and roles of freshwater clams (*Bivalvia, Unionidae*) in the freshwater tidal Hudson River. Freshwater Biology, 31: 239-248.

Strayer, D. L.; Caraco, N. F.; Cole, J. F.; Findlay, S.; and Pace, M. L., 1999. Transformation of freshwater ecosystems by bivalves. Bioscience, 49: 19-27.

Strayer, D. L.; Downing, J. A.; Haag, W. R.; King, T. L.; Layzer, J. B.; Newton, T. J.; and Nichols, S. J., 2004. Changing perspectives on pearly mussels, North America's most imperiled animals. Bioscience, 54: 429-439.

Tessier, A.; and Campbell, P. G. C., 1987. Partitioning of trace metals in sediments: relationships with bioavailability. Hydrobiologia, 149: 43-52.

Theler, J. L., 1990. A possible Oneota freshwater mussel assemblages cache intended for ceramic tempering material. Journal of the Iowa Archeological Society, 37: 1-3.

Trdan, R. J.; and Hoeh, W. R., 1982. Eurytopic host use by two congeneric mussel *(Pelecypoda: Unionidae: Anodonta*). American Midland Naturalist, 108: 381-388.

Vaughn, C. C., 2010. Biodiversity losses and ecosystem function in freshwaters: Emerging conclusions and research directions. Bioscience, 60: 25-35.

Vaughn, C. C.; and Hakenkamp C. C., 2001. The functional role of burrowing bivalves in freshwater ecosystems. Freshwater Biology, 46: 1431-1446.

Waller, D. L.; Rach, J. J.; and Luoma, J. A., 1998. Acute toxicity and accumulation of the pesticide 3-trifluoromethyl-4nitrophenol (TFM) in freshwater mussels *(Bivalvia: Unionidae*). Ecotoxicology, 7: 113-121.

Walters, A.; Banta, J.; Williams, M.; and Williams, L., 2012. Using Maxent to predict the probability of occurrence of rare fish and mussel species in the Sabine River and Neches River in East Texas.

Warren, D. L.; and Seifert, S. N., 2011. Species distribution modeling in Maxent: the importance of model complexity and the performance of model selection criteria. Ecological Applications, 21(2): 335-342.

Watters, G. T.; and O'Dee, S. H., 1999. Glochidia of the freshwater mussel *Lampsilis* overwintering on fish hosts. Journal of Molluscan Studies, 65: 453-459.

Watters, G. T., 2000. Freshwater mussels and water quality: A review of the effects of hydrologic and instream habitat alterations. First Freshwater Mollusk Conservation Society Symposium., 261-274.

Way, C. M., 1989. Dynamics of filter feeding in *Musculium transversum* (*Bivalvia: Sphaeriidae*). Journal of the North American Benthological Society, 8: 243-249.

Weinstein, R. A.; and Dumas, A. A., 2008. The spread of shell-tempered ceramics along the northern coast of the Gulf of Mexico. Southeastern Archaeology, 27: 202-221.

Welker, M.; and Walz, N., 1998. Can mussels control the plankton in rivers? - a planktological approach applying a Lagrangian sampling strategy. Limnology and Oceanography, 43: 753-762.

Williams, J. D.; Fuller, S. L. H.; and Grace, R., 1992. Effects of impoundment on freshwater mussels (*Mollusca: Bivalvia: Unionidae*) in the main channel of the Black Warrior and Tombigbee rivers in western Alabama. Bulletin of the Alabama Museum of Natural History, No. 13:1-10.

Wisz, M.; Hijmans, R.; Li, J.; Peterson, A.; Graham, C.; and Guisan, A., 2008. Effects of sample size on the performance of species distribution models. Diversity and Distributions, 14: 763-773.

Wooton, J. T., 1992. Indirect effects, prey susceptibility, and habitat selection: impacts of birds on limpets and algae. Ecology, 73: 981-991.

Yeager, M. M.; and Saylor, C. F., 1995. Fish hosts for four species of freshwater mussels (*Pelecypoda: Unionidae*) in the upper Tennessee River drainage. American Midland Naturalist 133: 1-6.

Yom-Tov, Y.; and Kadmon, R., 1998. Analysis of the distribution of insectivorous bats in Israel. Diversity and Distributions, 4: 63-70.

Zale, A. V.; and Neves, R. J., 1982. Reproductive biology of four freshwater mussel species (*Mollusca: Unionidae*) in Virginia. Freshwater Invertebrate Biology, 1: 17-28.

Zimmerman, L. L.; and Neves, R. J., 2002. Effects of temperature on duration of viability for glochidia of freshwater mussels (*Bivalvia: Unionidae*). American Malacological Bulletin, 17: 31-35.

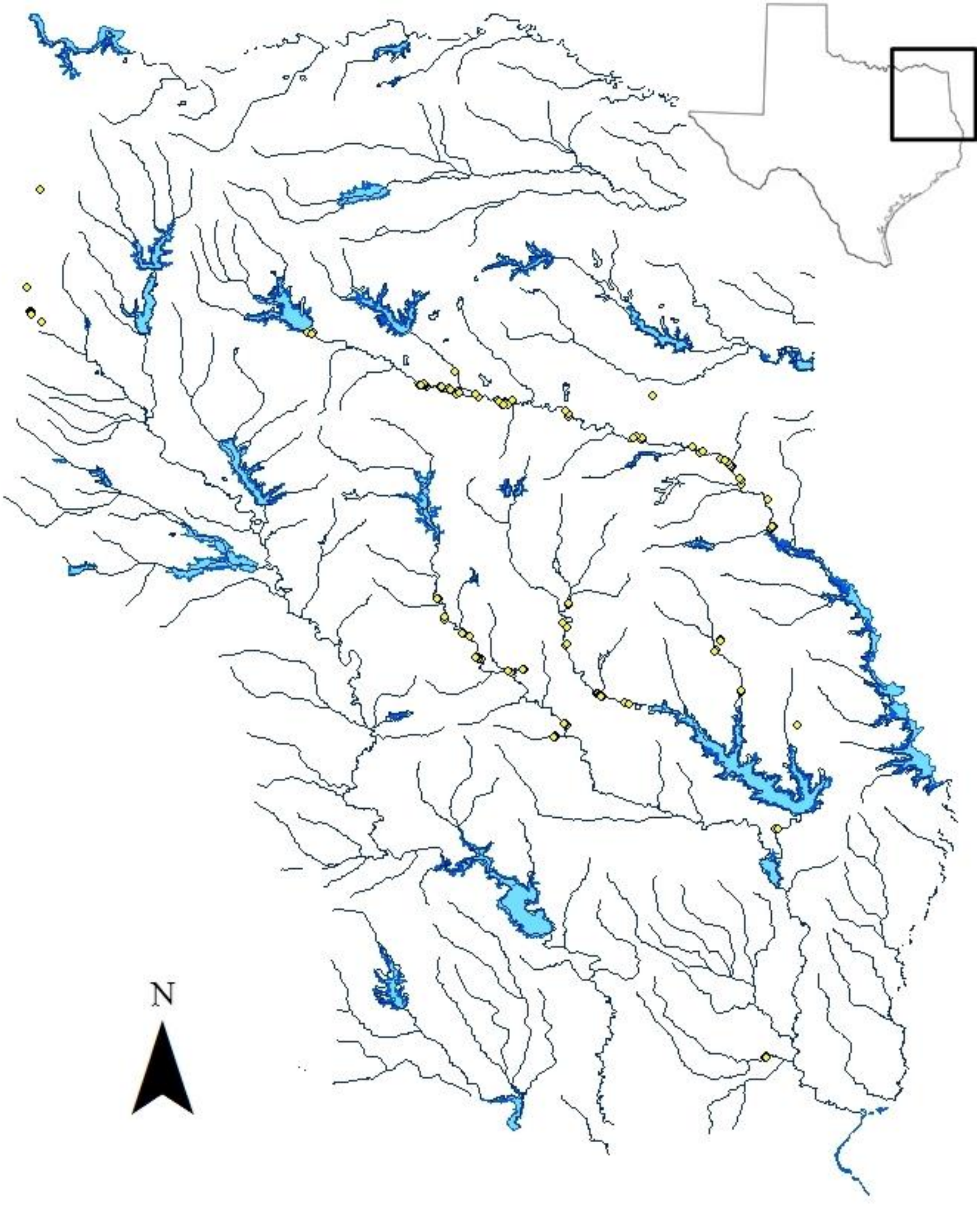

Figure 2.1. Map of the rivers of East Texas. Points indicate locations of sampling effort in the Sabine, Neches, Angelina, or Trinity Rivers.

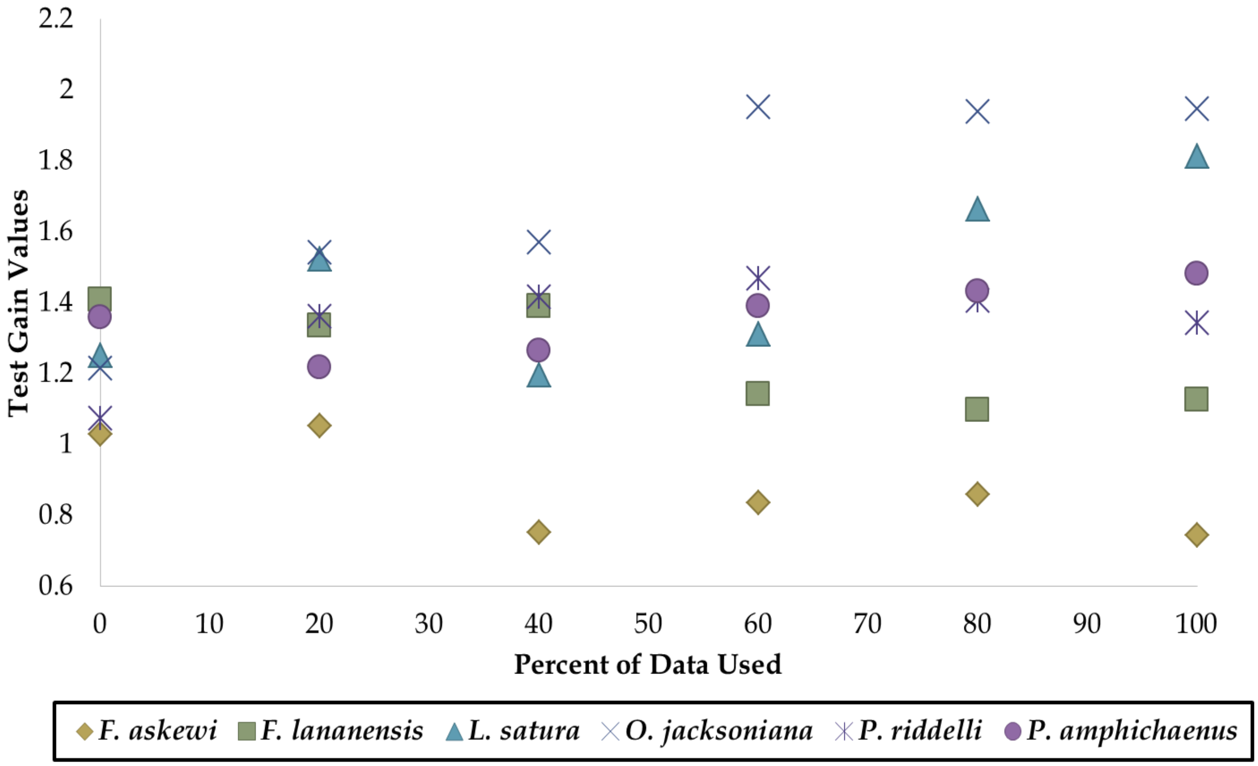

Figure 3.1. Test gain values for all six mussel species.

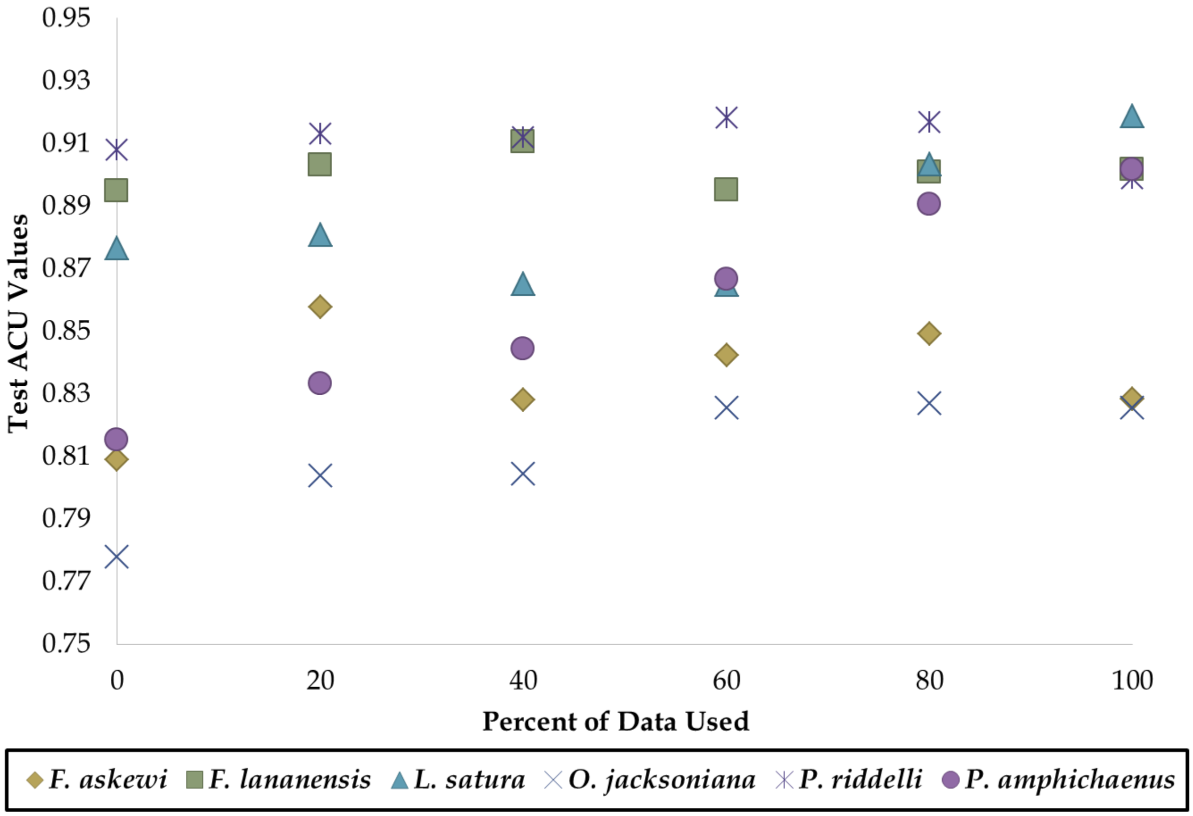

Figure 3.2. Test AUC values for all six mussel species.

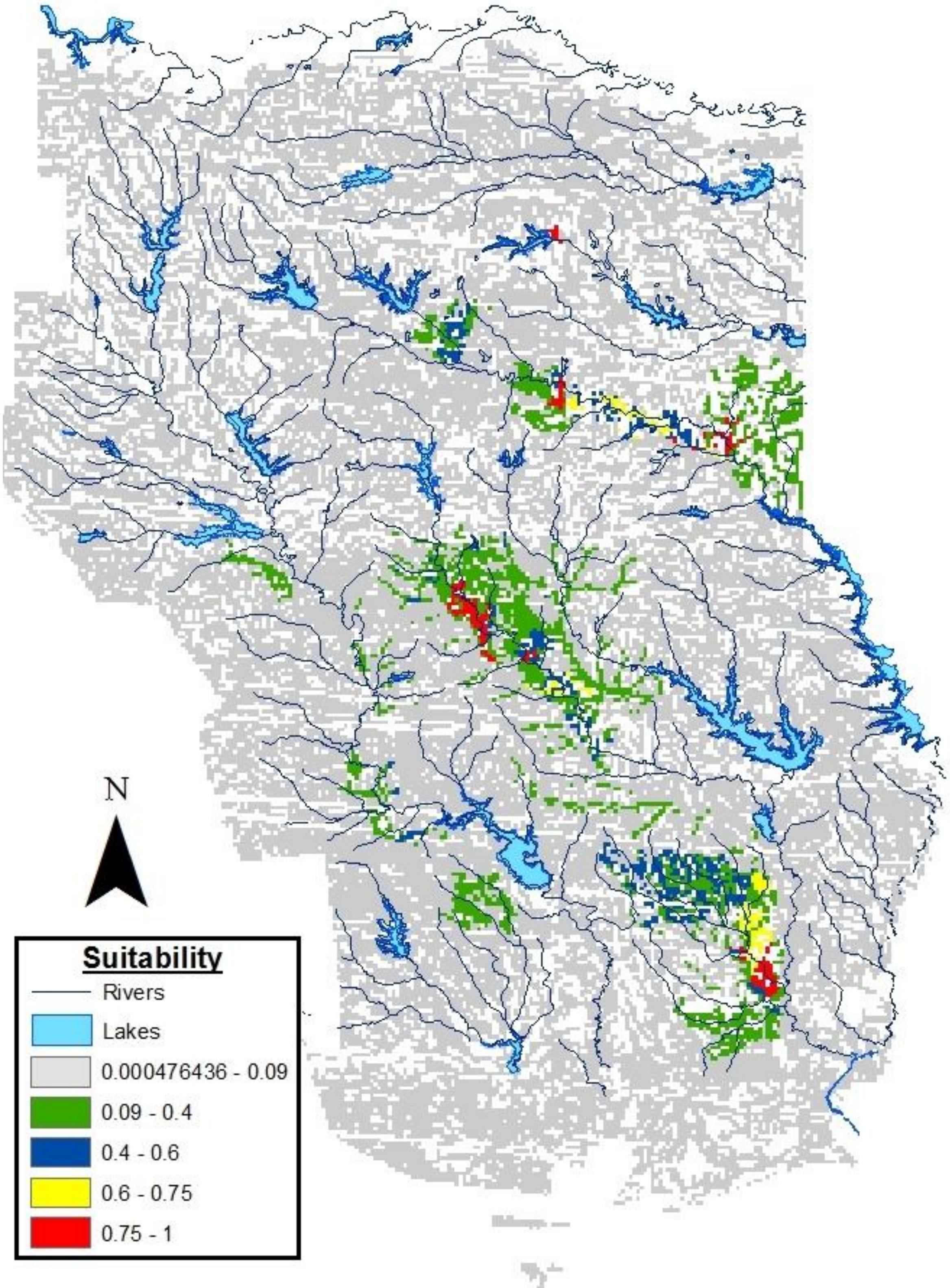

Figure 3.3. The predicted potential suitable habitat for *Fusconaia askewi* in East Texas with no new occurrence points added.

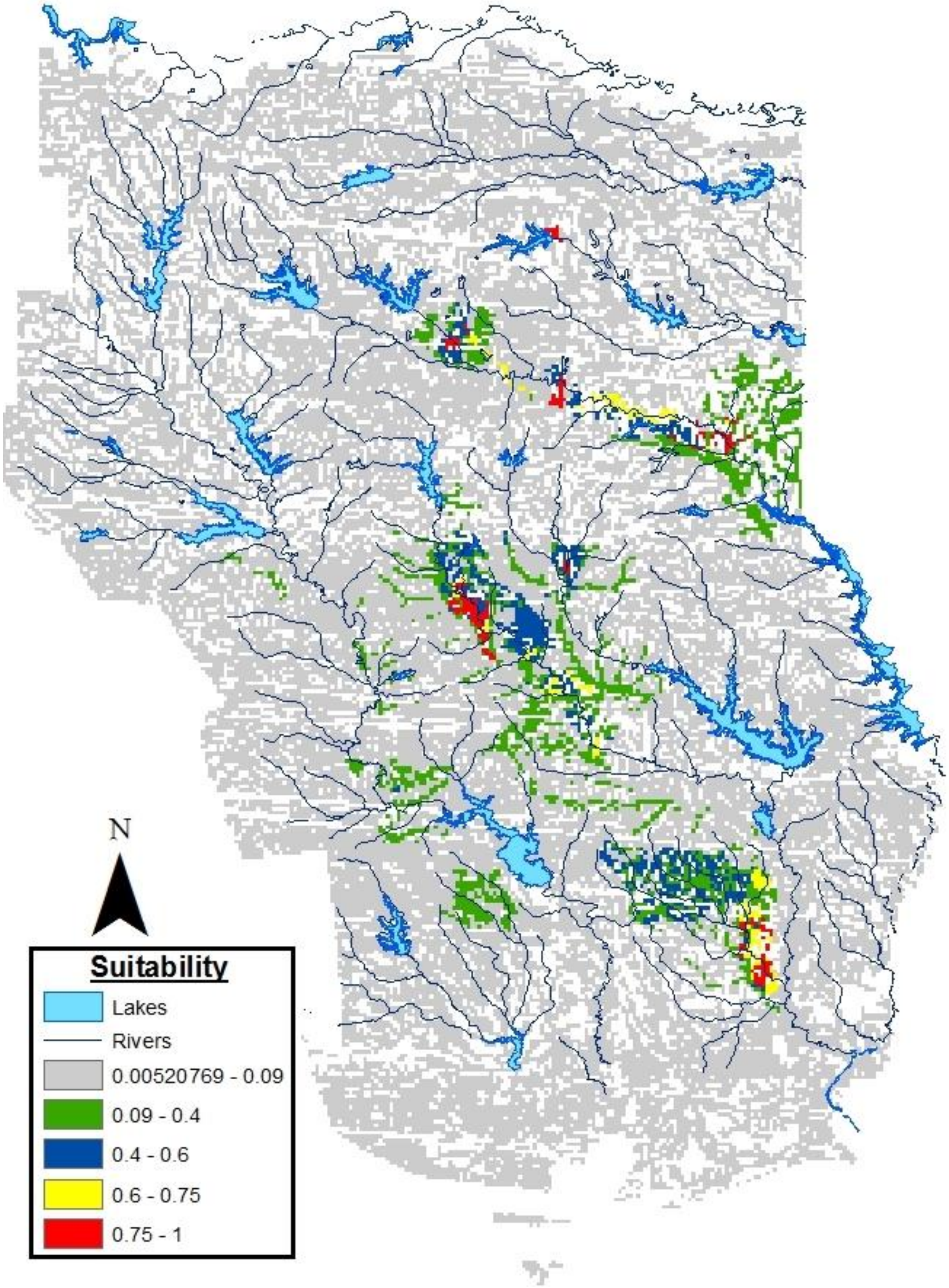

Figure 3.4. The predicted potential suitable habitat for *Fusconaia askewi* in East Texas with 20% of the new occurrence points added.

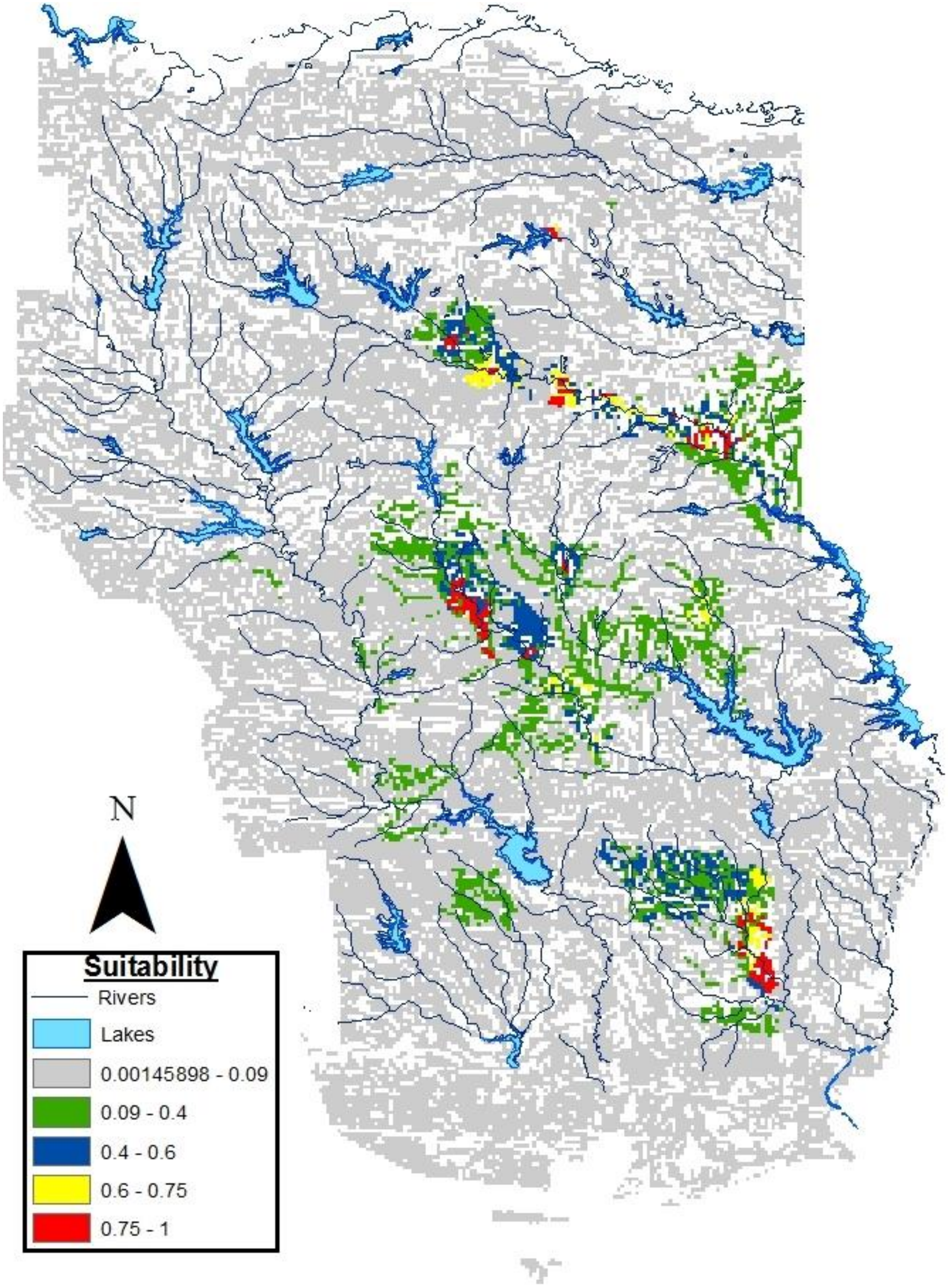

Figure 3.5. The predicted potential suitable habitat for *Fusconaia askewi* in East Texas with 40% of the new occurrence points added.

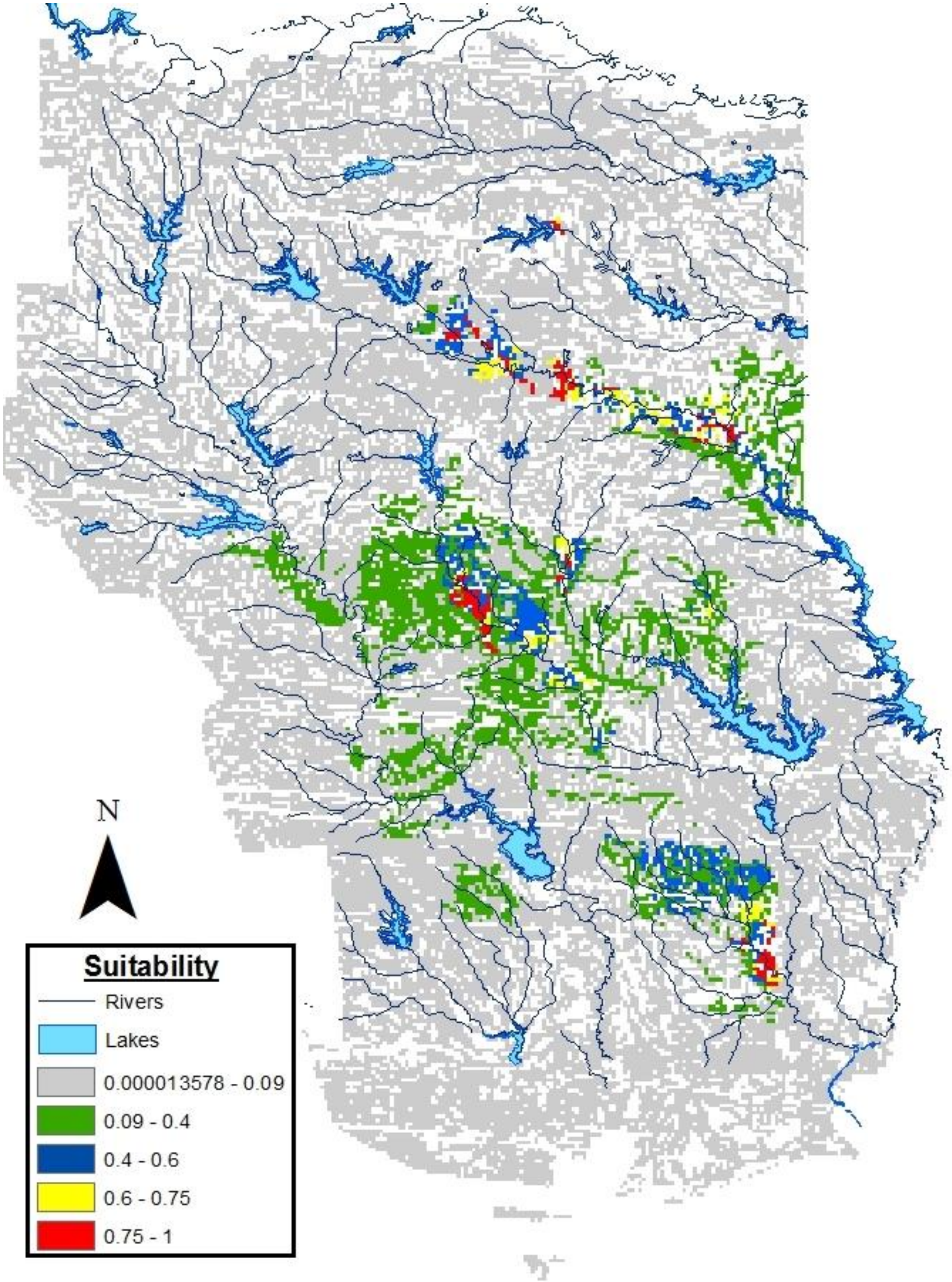

Figure 3.6. The predicted potential suitable habitat for *Fusconaia askewi* in East Texas with 60% of the new occurrence points added.

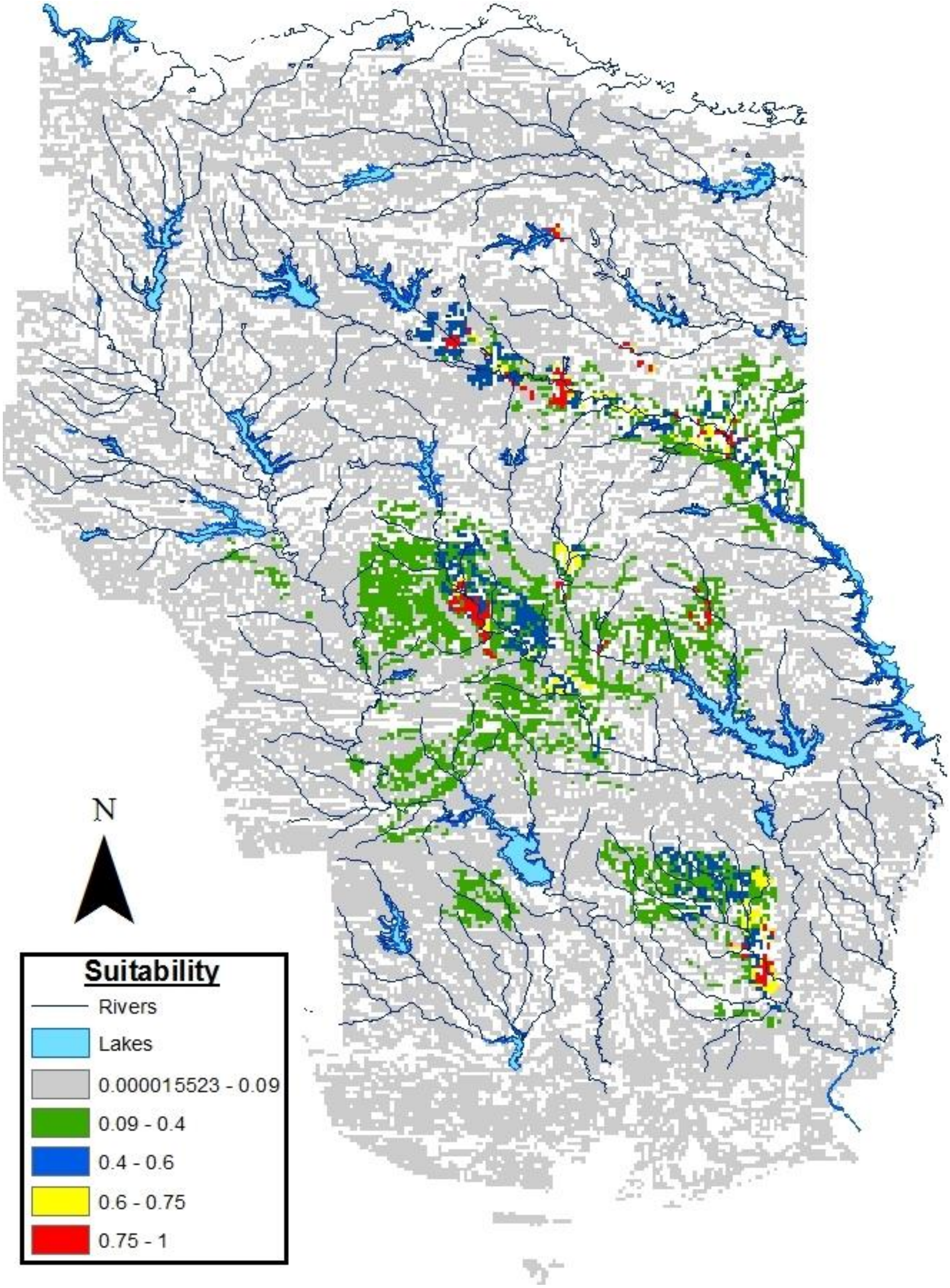

Figure 3.7. The predicted potential suitable habitat for *Fusconaia askewi* in East Texas with 80% of the new occurrence points added.

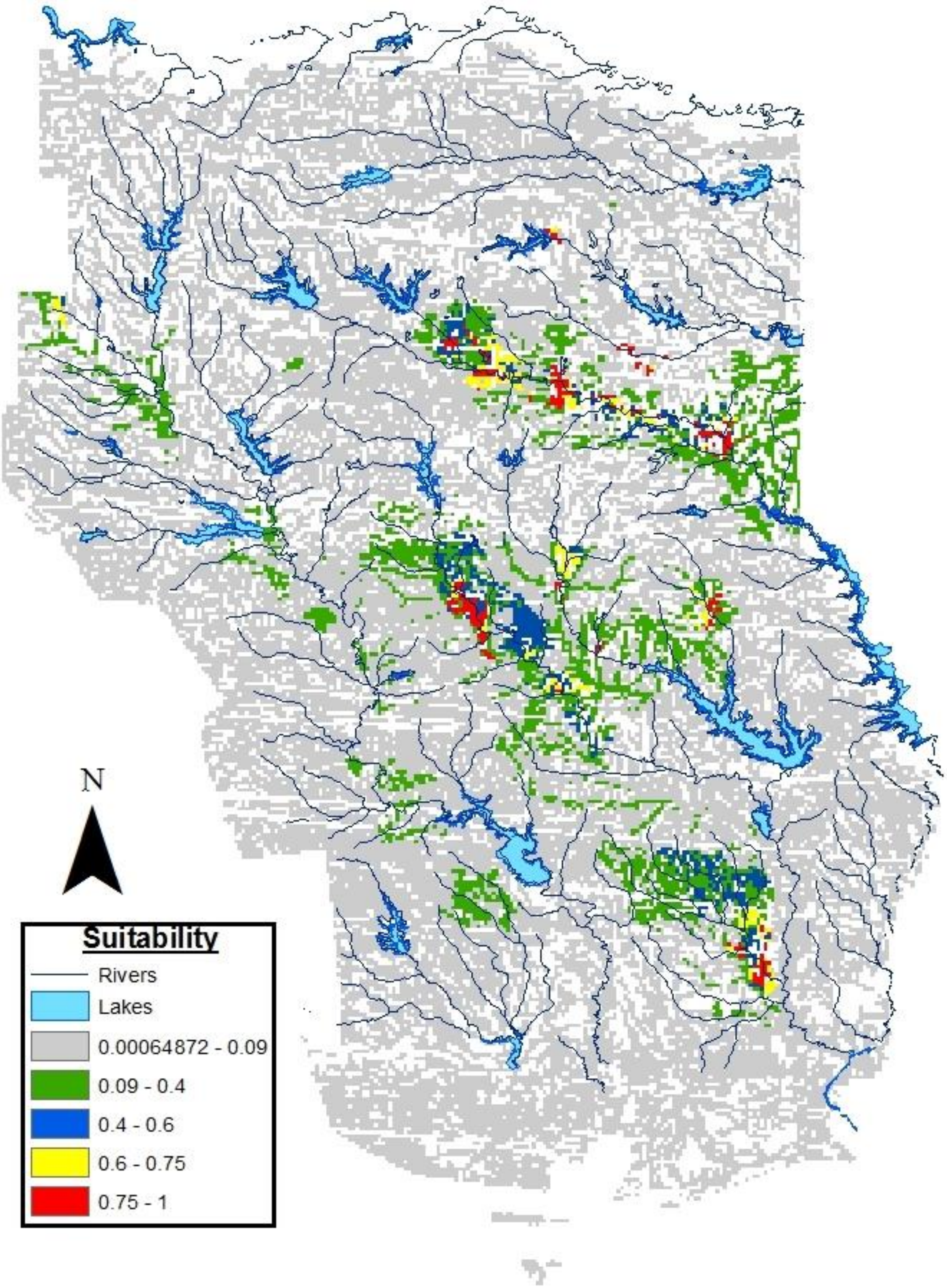

Figure 3.8. The predicted potential suitable habitat for *Fusconaia askewi* in East Texas with 100% of the new occurrence points added.

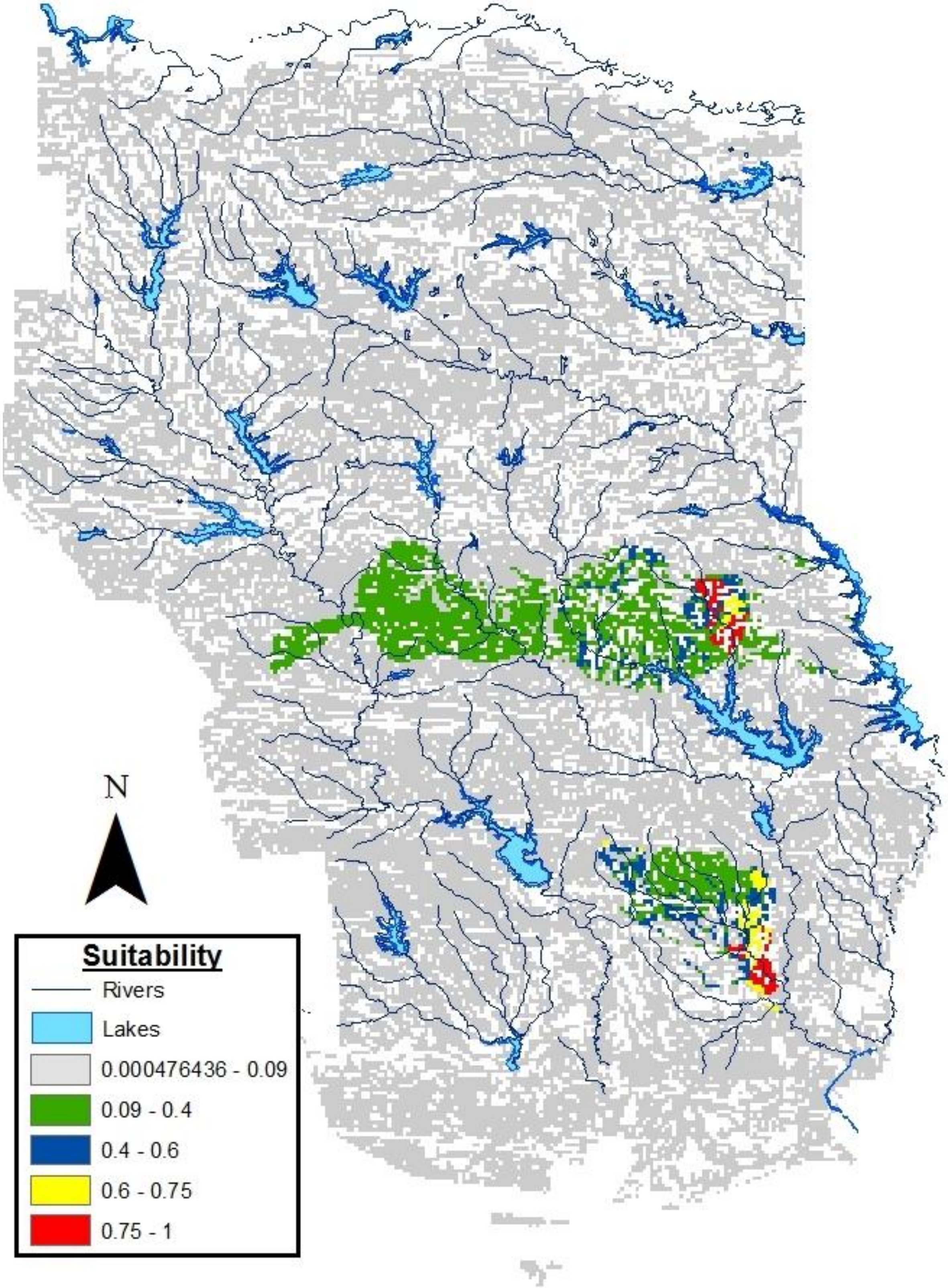

Figure 3.9. The predicted potential suitable habitat for *Fusconaia lananensis* in East Texas with no new occurrence points added.

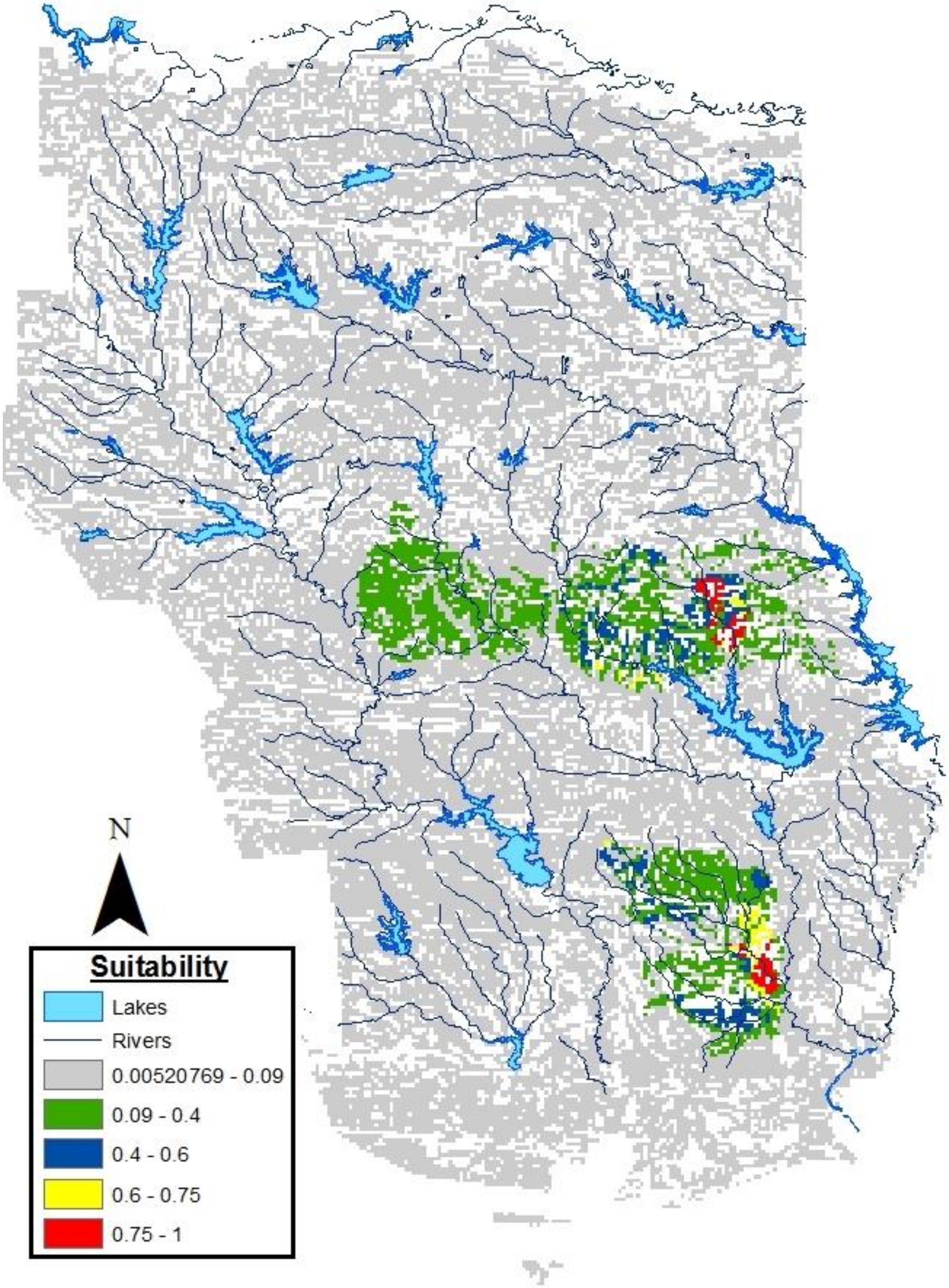

Figure 3.10. The predicted potential suitable habitat for *Fusconaia lananensis* in East Texas with 20% of the new occurrence points added.

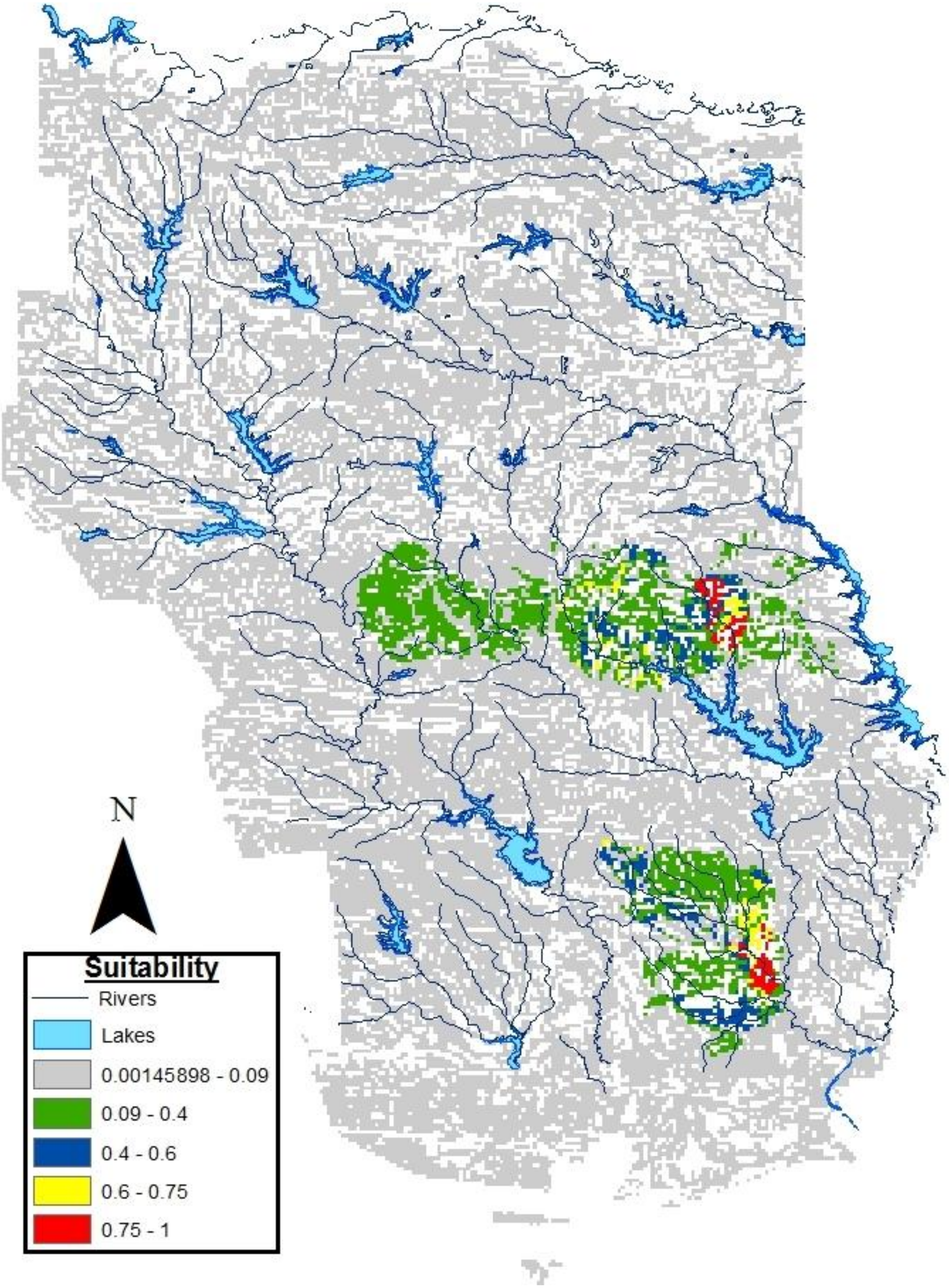

Figure 3.11. The predicted potential suitable habitat for *Fusconaia lananensis* in East Texas with 40% of the new occurrence points added.

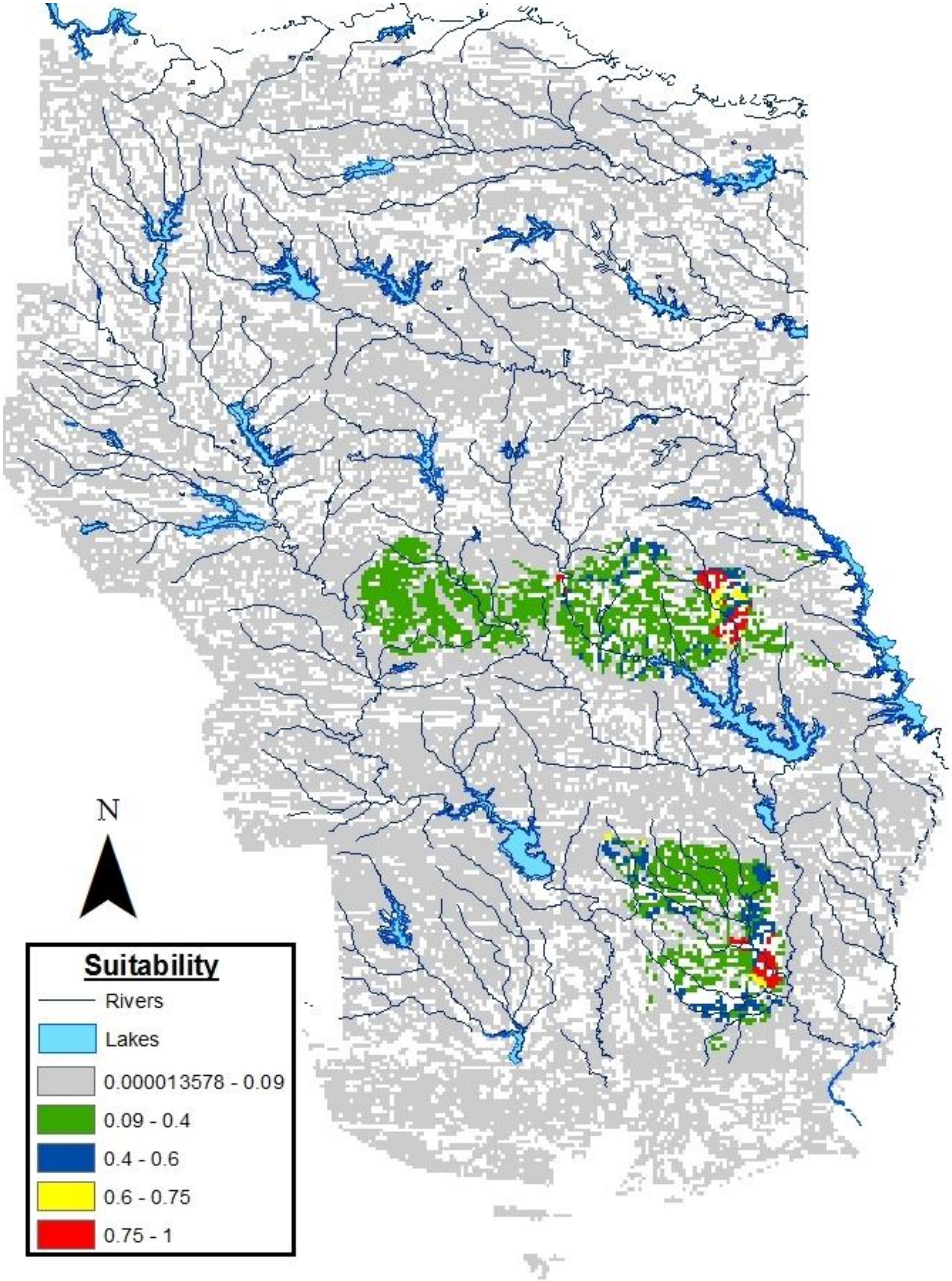

Figure 3.12. The predicted potential suitable habitat for *Fusconaia lananensis* in East Texas with 60% of the new occurrence points added.

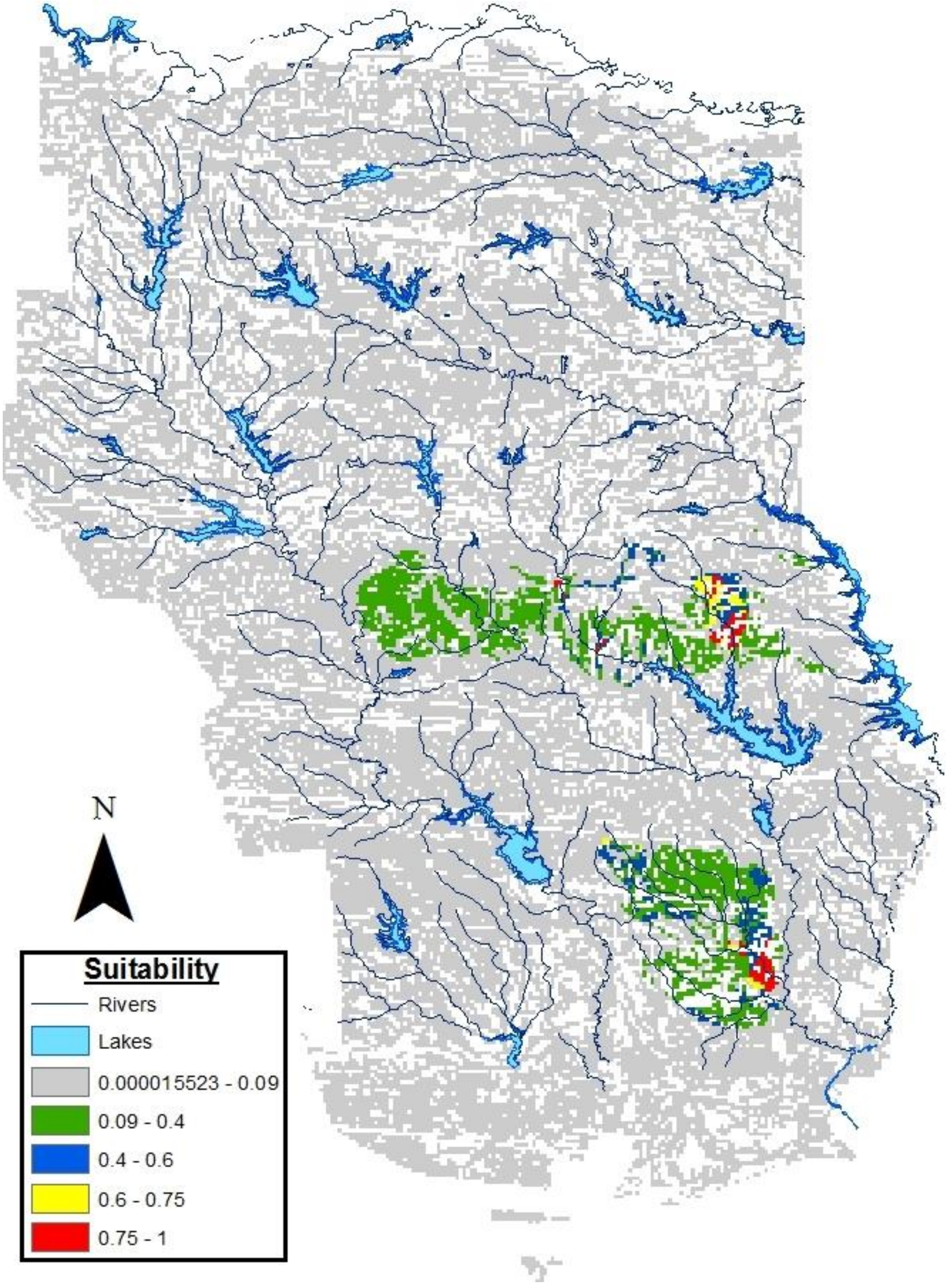

Figure 3.13. The predicted potential suitable habitat for *Fusconaia lananensis* in East Texas with 80% of the new occurrence points added.

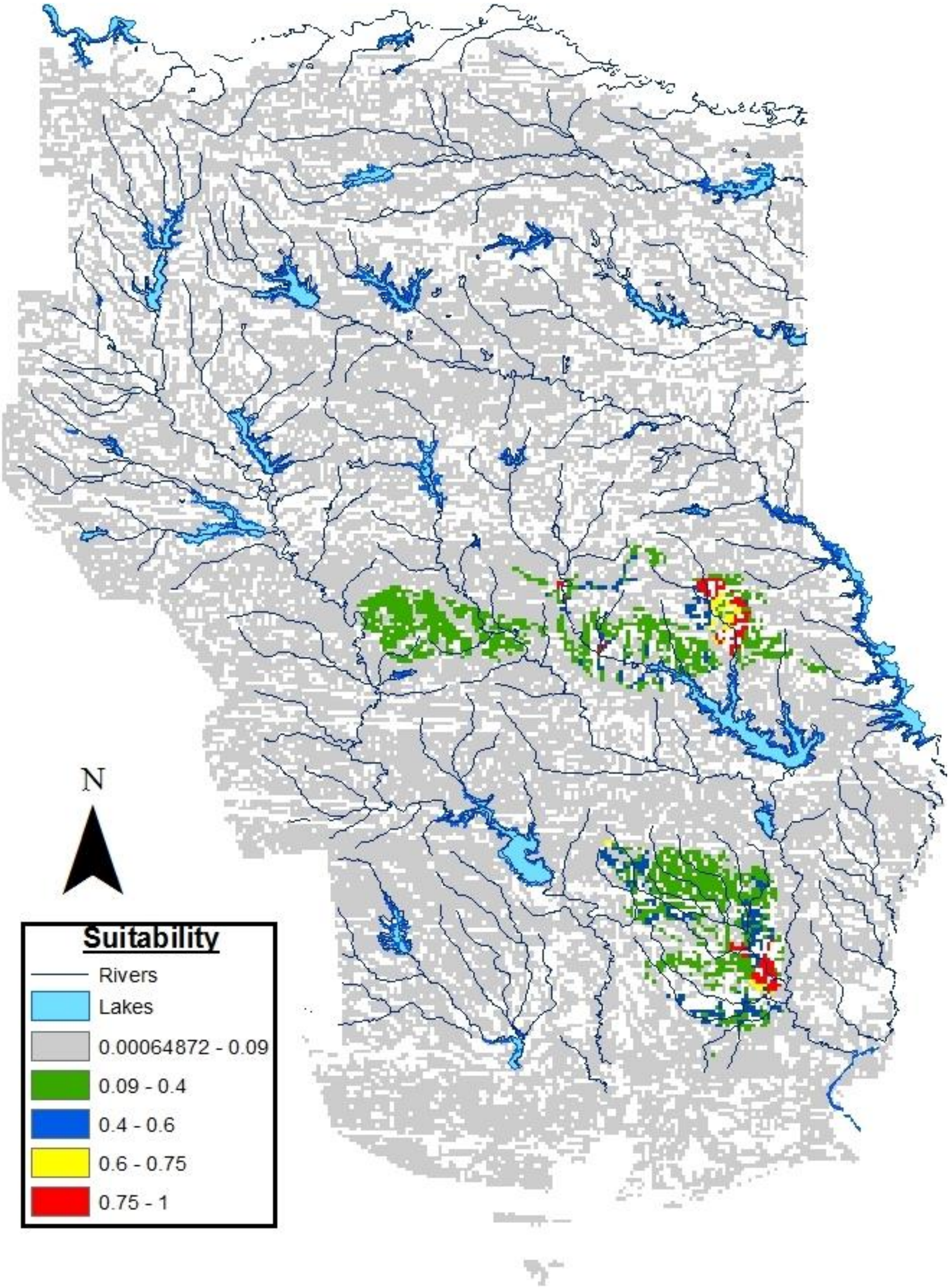

Figure 3.14. The predicted potential suitable habitat for *Fusconaia lananensis* in East Texas with 100% of the new occurrence points added.

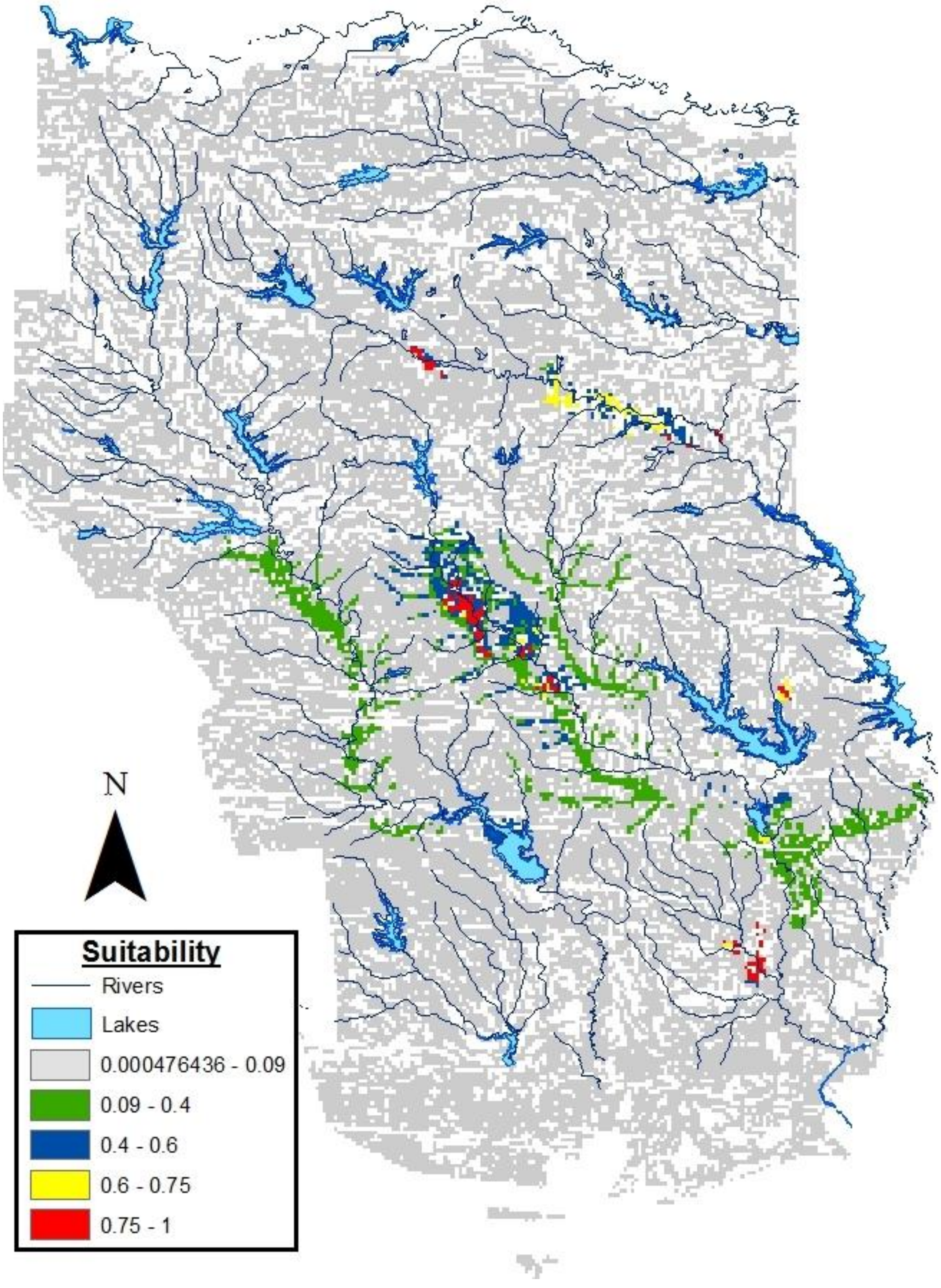

Figure 3.15. The predicted potential suitable habitat for *Lampsilis satura* in East Texas with no new data added.

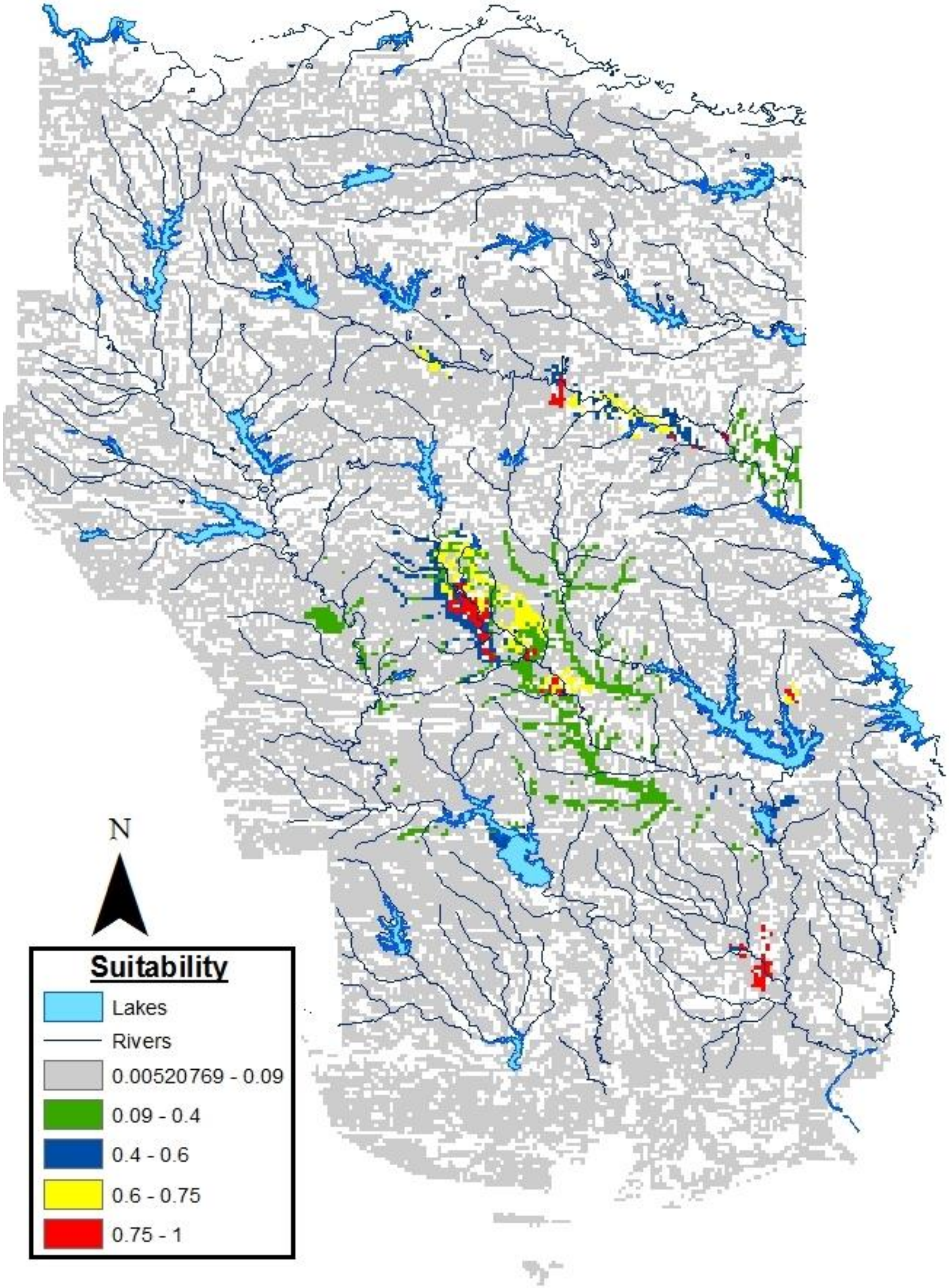

Figure 3.16. The predicted potential suitable habitat for *Lampsilis satura* in East Texas with 20% of the new occurrence points added.

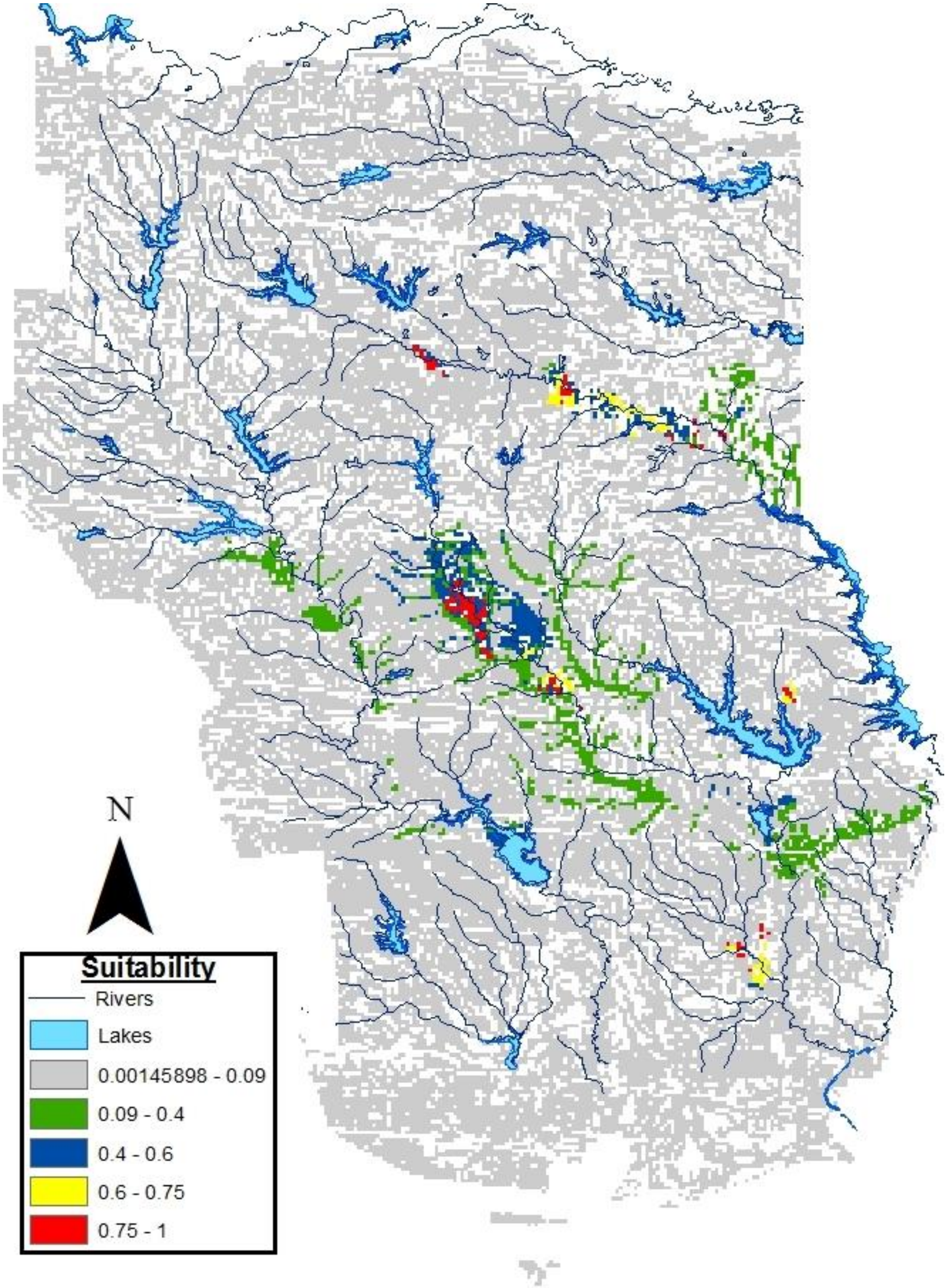

Figure 3.17. The predicted potential suitable habitat for *Lampsilis satura* in East Texas with 40% of the new occurrence points added.

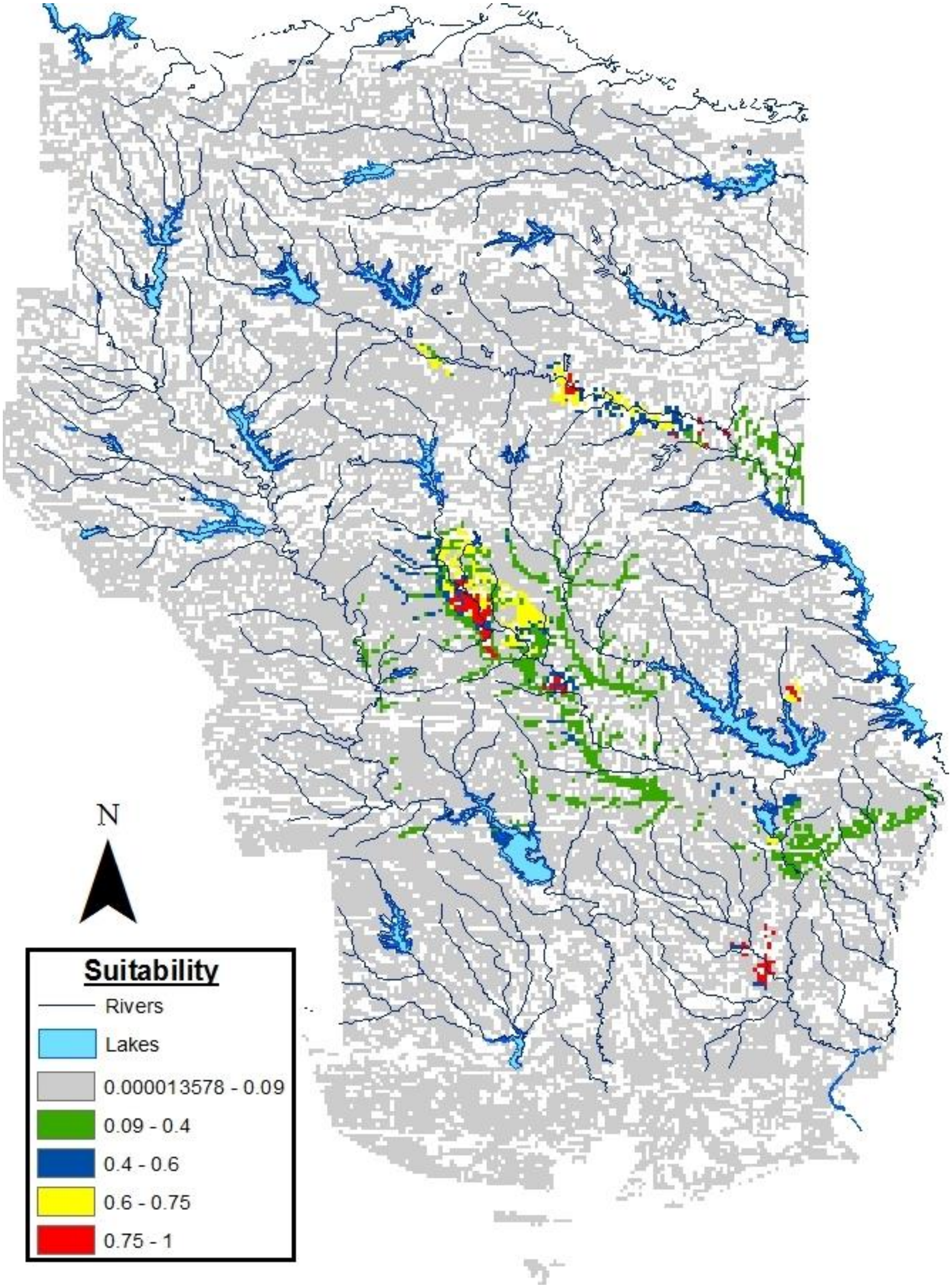

Figure 3.18. The predicted potential suitable habitat for *Lampsilis satura* in East Texas with 60% of the new occurrence points added.

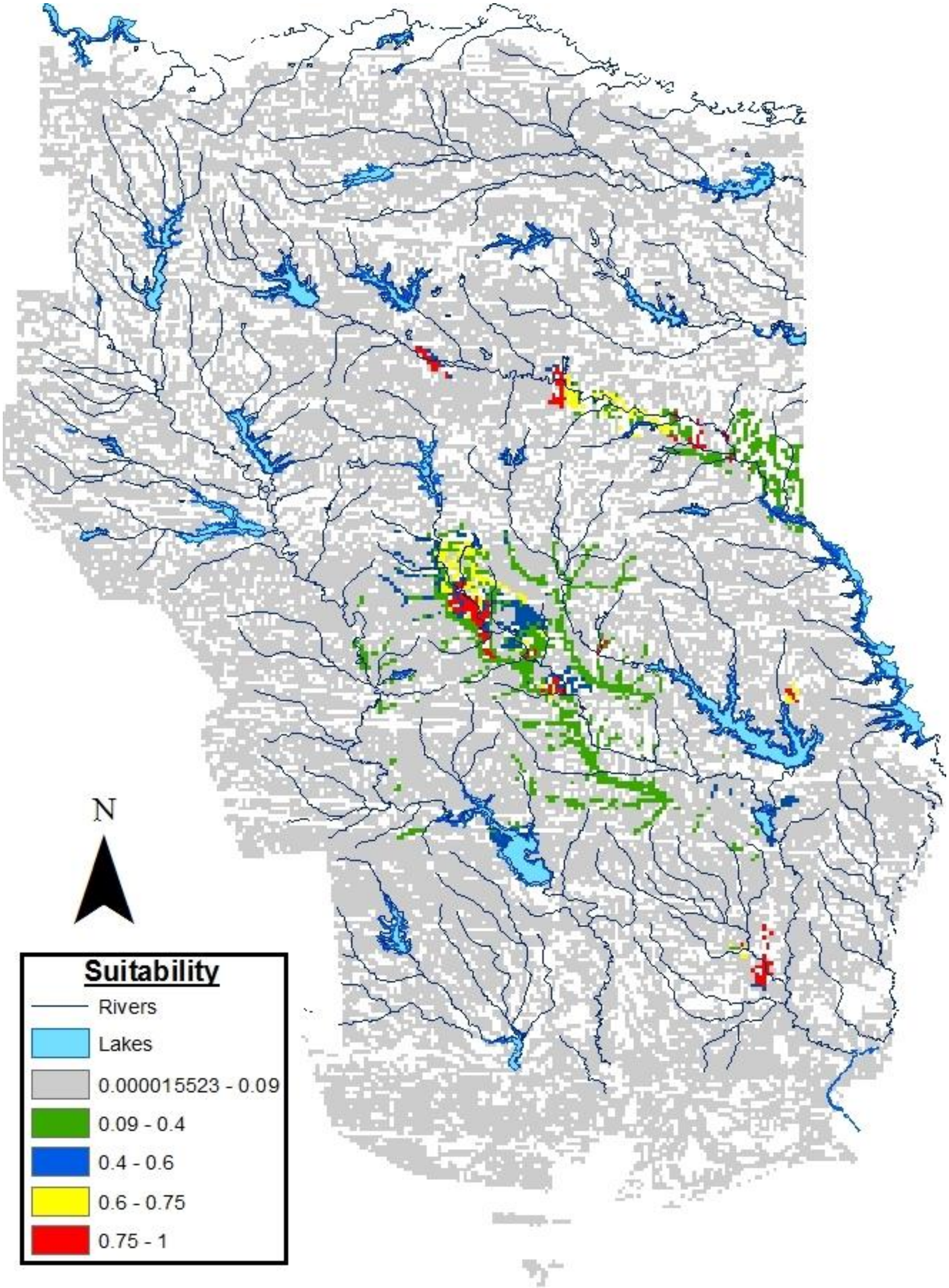

Figure 3.19. The predicted potential suitable habitat for *Lampsilis satura* in East Texas with 80% of the new occurrence points added.

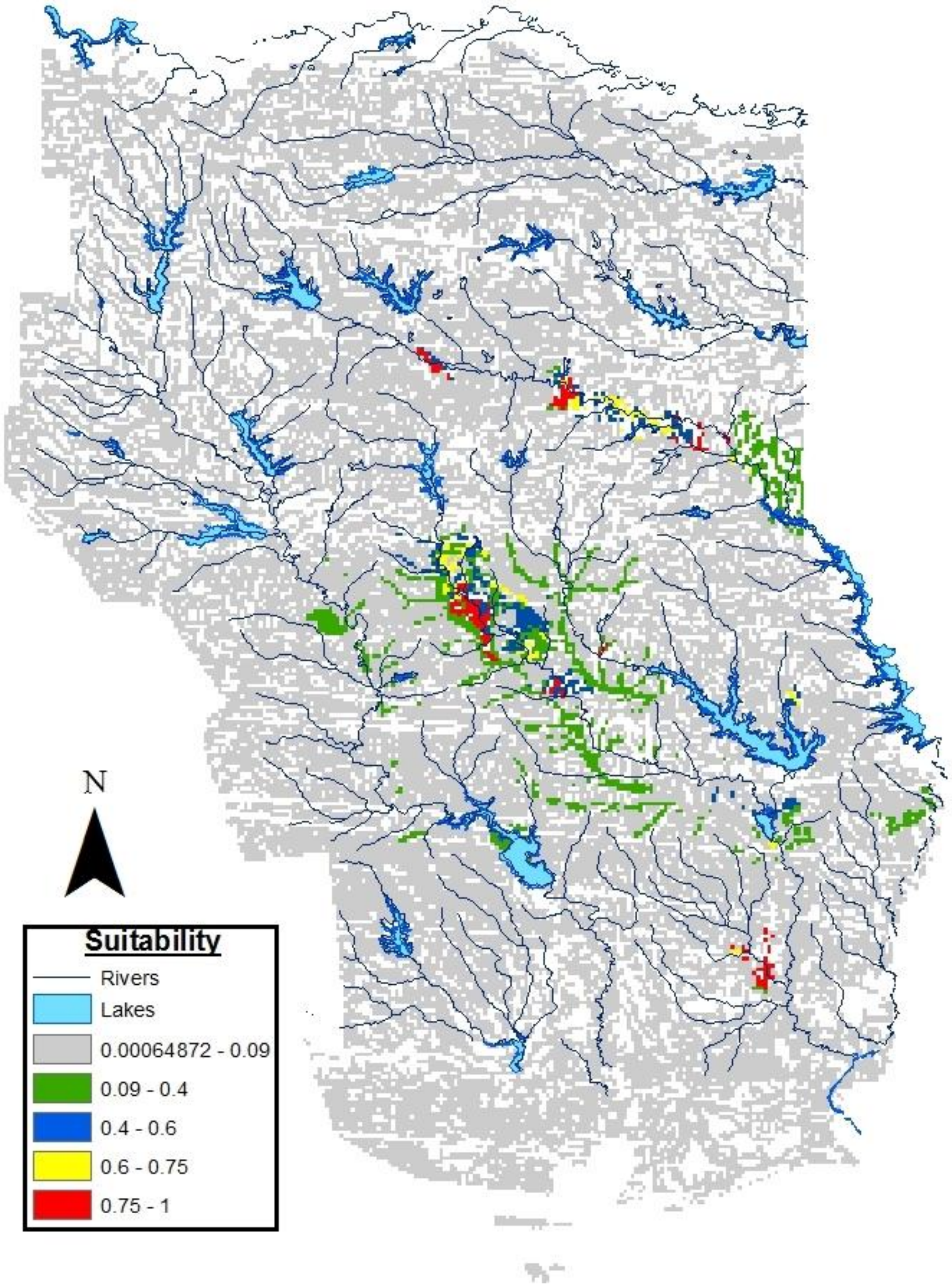

Figure 3.20. The predicted potential suitable habitat for *Lampsilis satura* in East Texas with all 100% of the new occurrence points added.

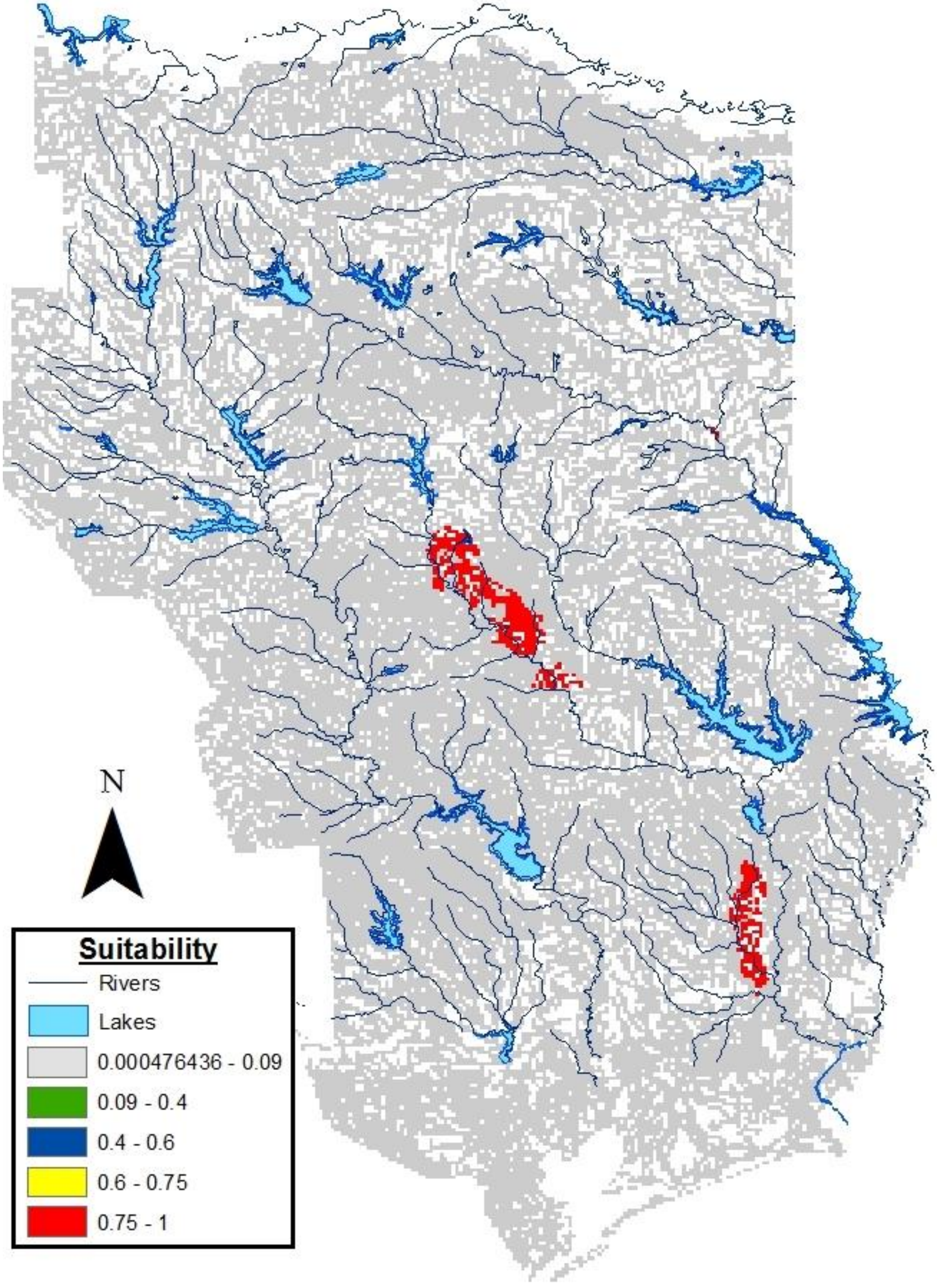

Figure 3.21. The predicted potential suitable habitat for *Obovaria jacksoniana* in East Texas with no new occurrence points added.

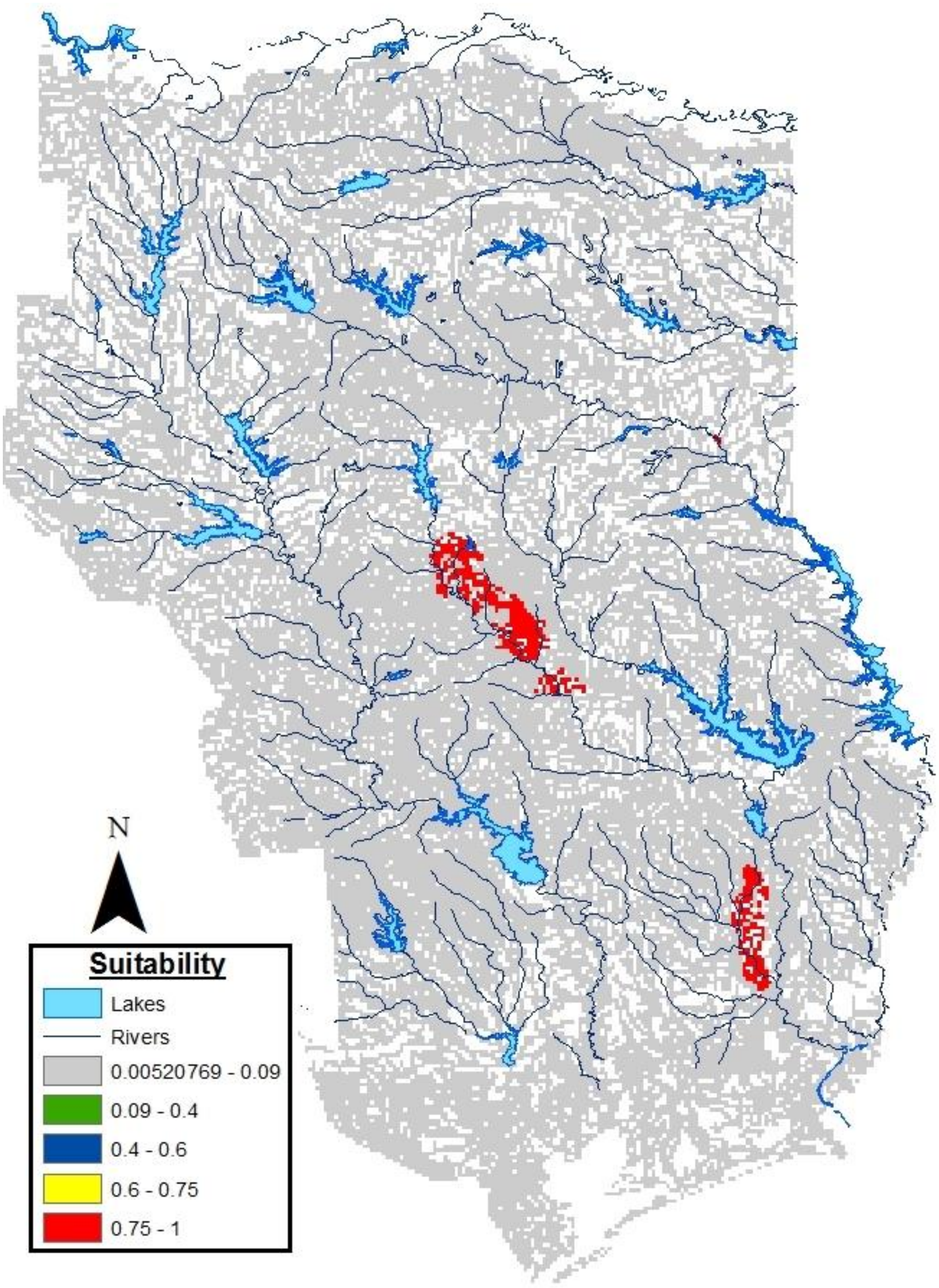

Figure 3.22. The predicted potential suitable habitat for *Obovaria jacksoniana* in East Texas with 20% of the new occurrence points added.

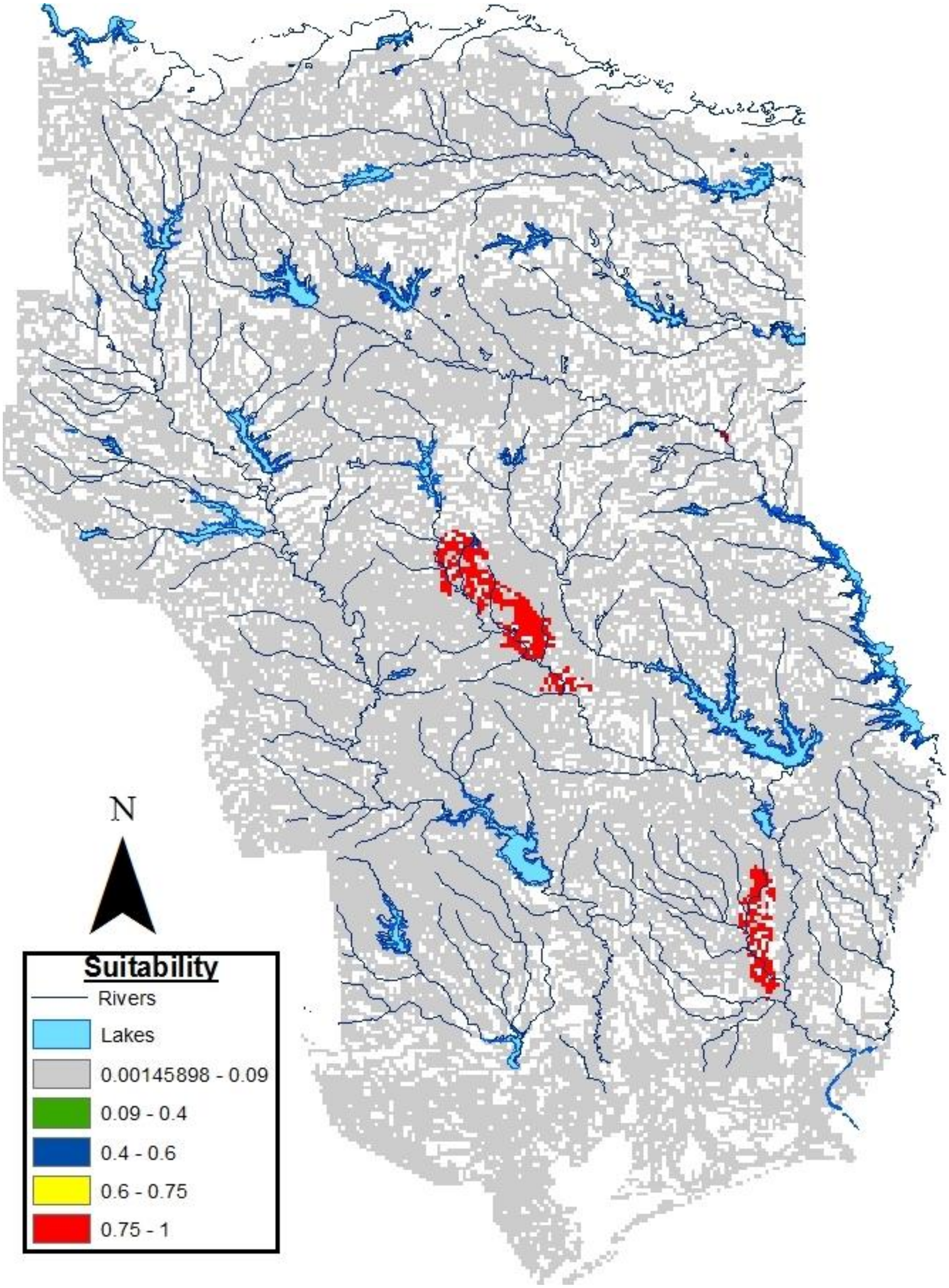

Figure 3.23. The predicted potential suitable habitat for *Obovaria jacksoniana* in East Texas with 40% of the new occurrence points added.

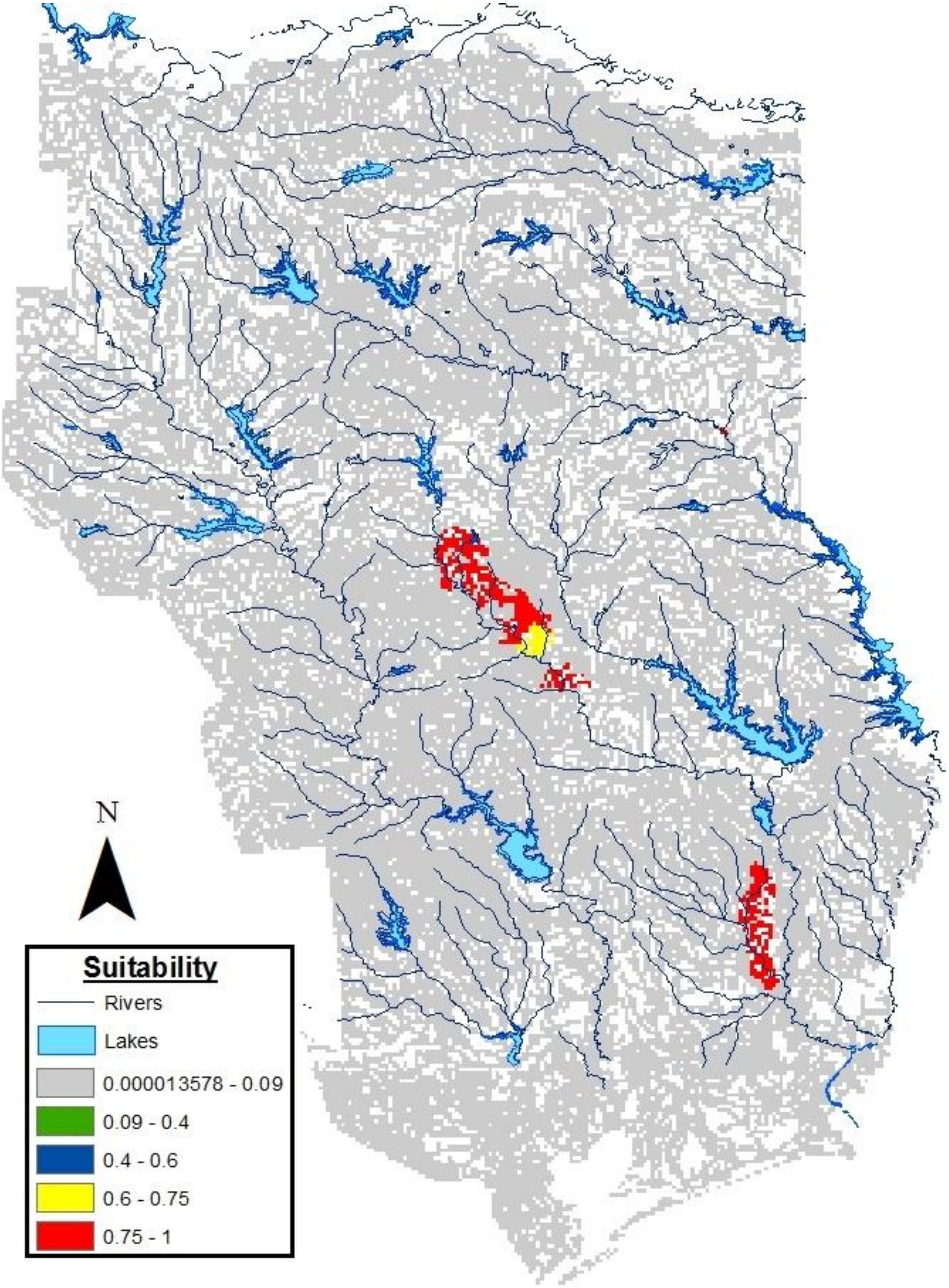

Figure 3.24. The predicted potential suitable habitat for *Obovaria jacksoniana* in East Texas with 60% of the new occurrence points added.

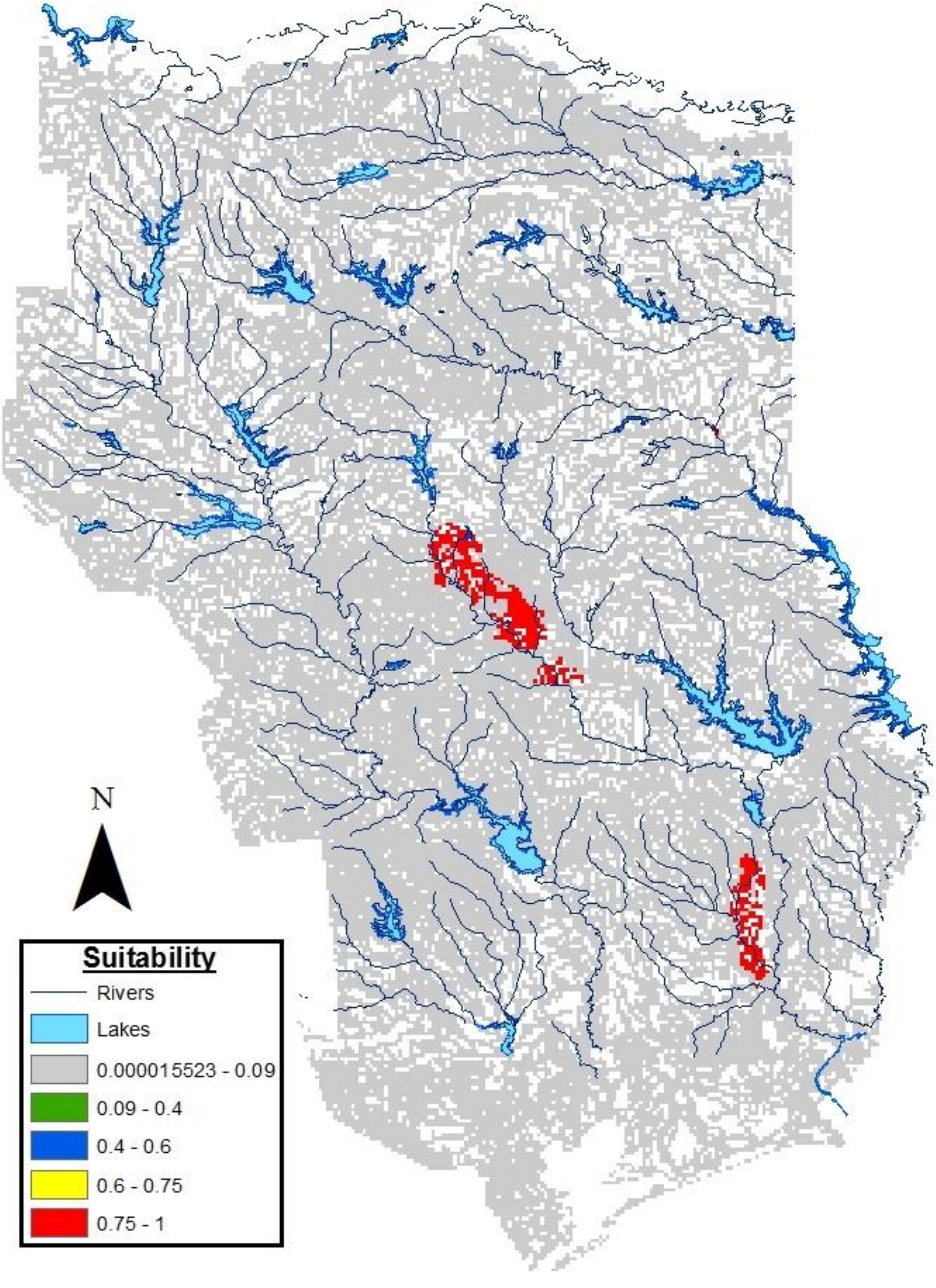

Figure 3.25. The predicted potential suitable habitat for *Obovaria jacksoniana* in East Texas with 80% of the new occurrence points added.

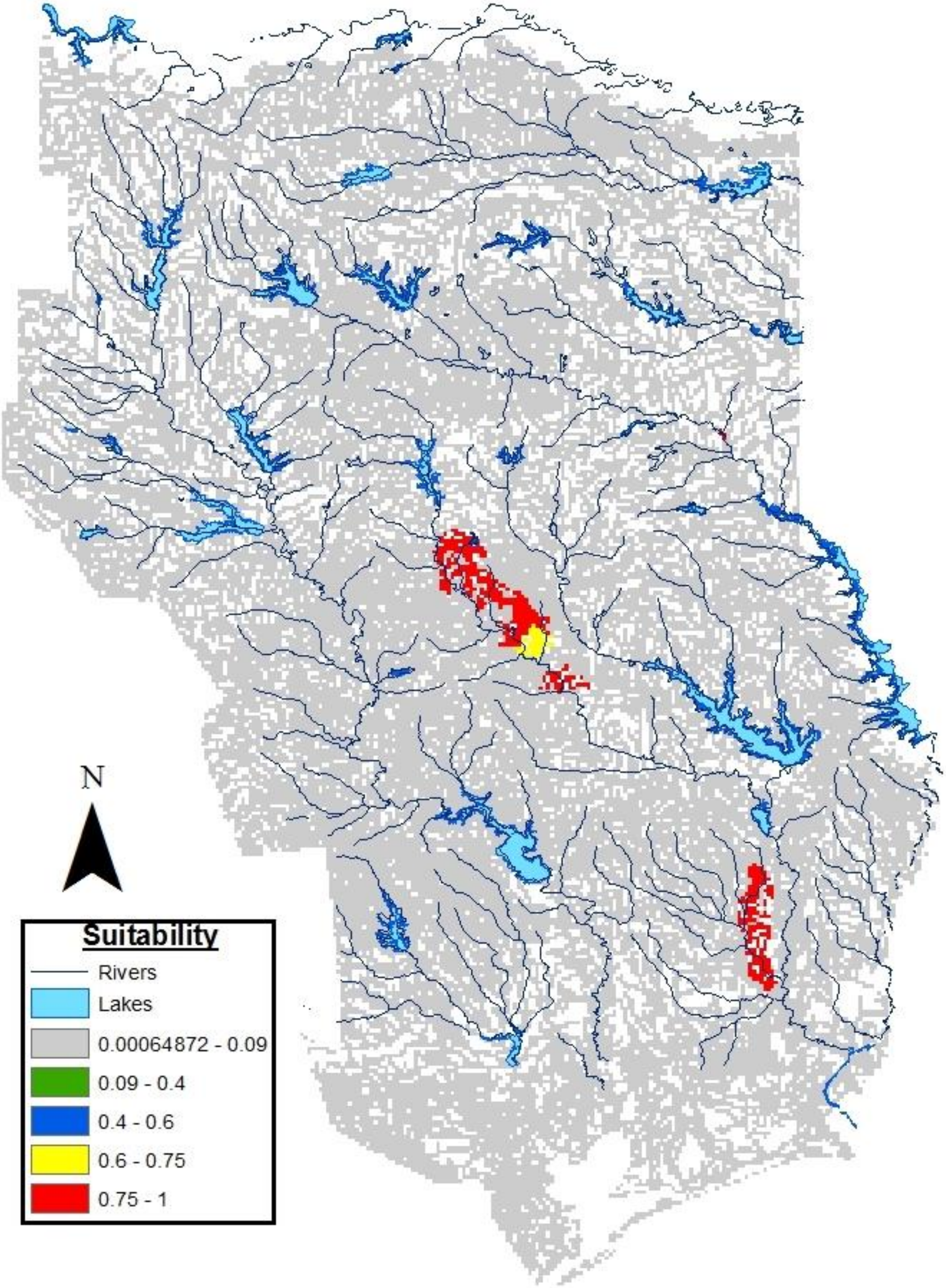

Figure 3.26. The predicted potential suitable habitat for *Obovaria jacksoniana* in East Texas with 100% of the new occurrence points added.

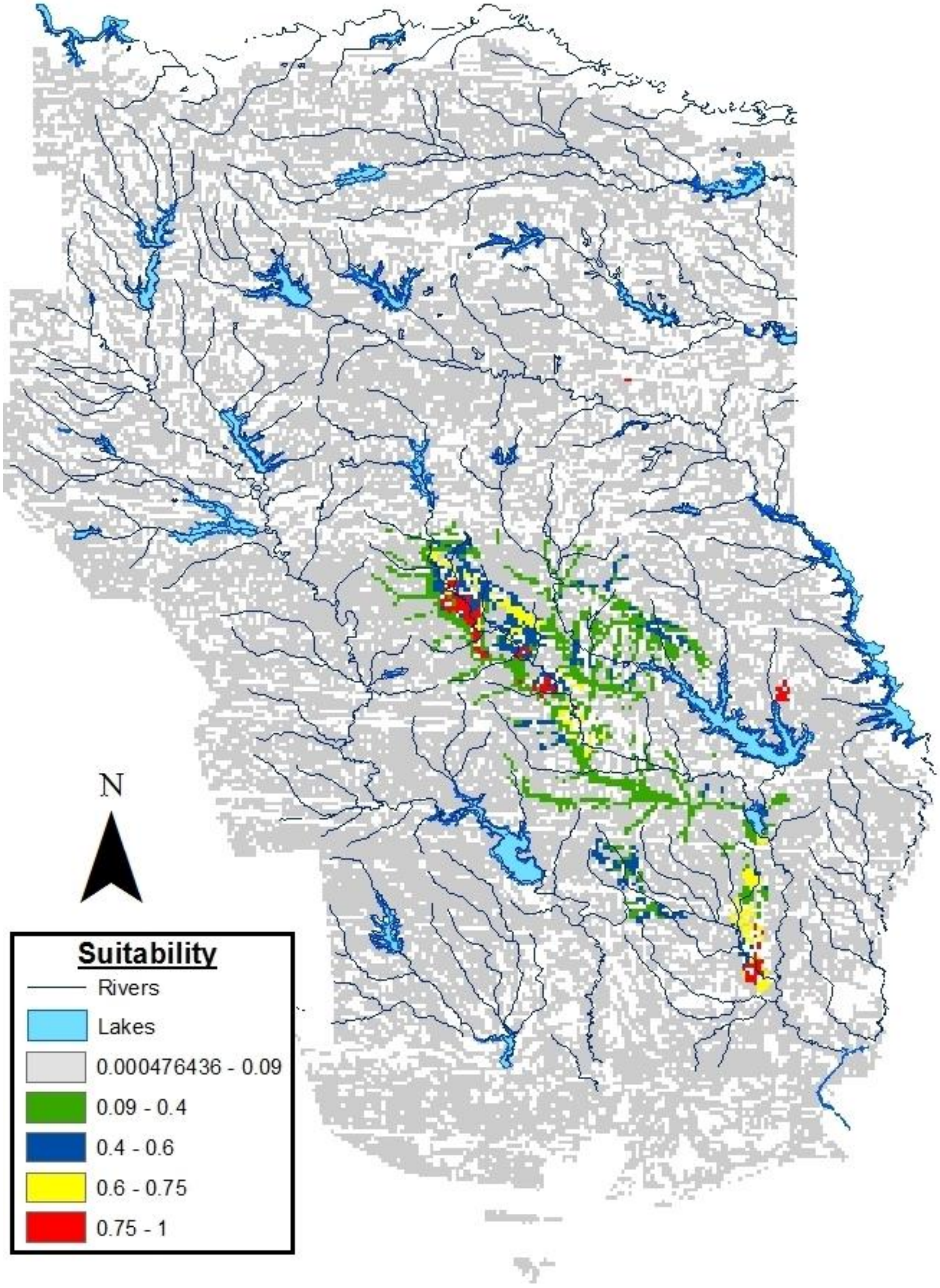

Figure 3.27. The predicted potential suitable habitat for *Pleuroblema riddellii* in East Texas with no new occurrence points added.

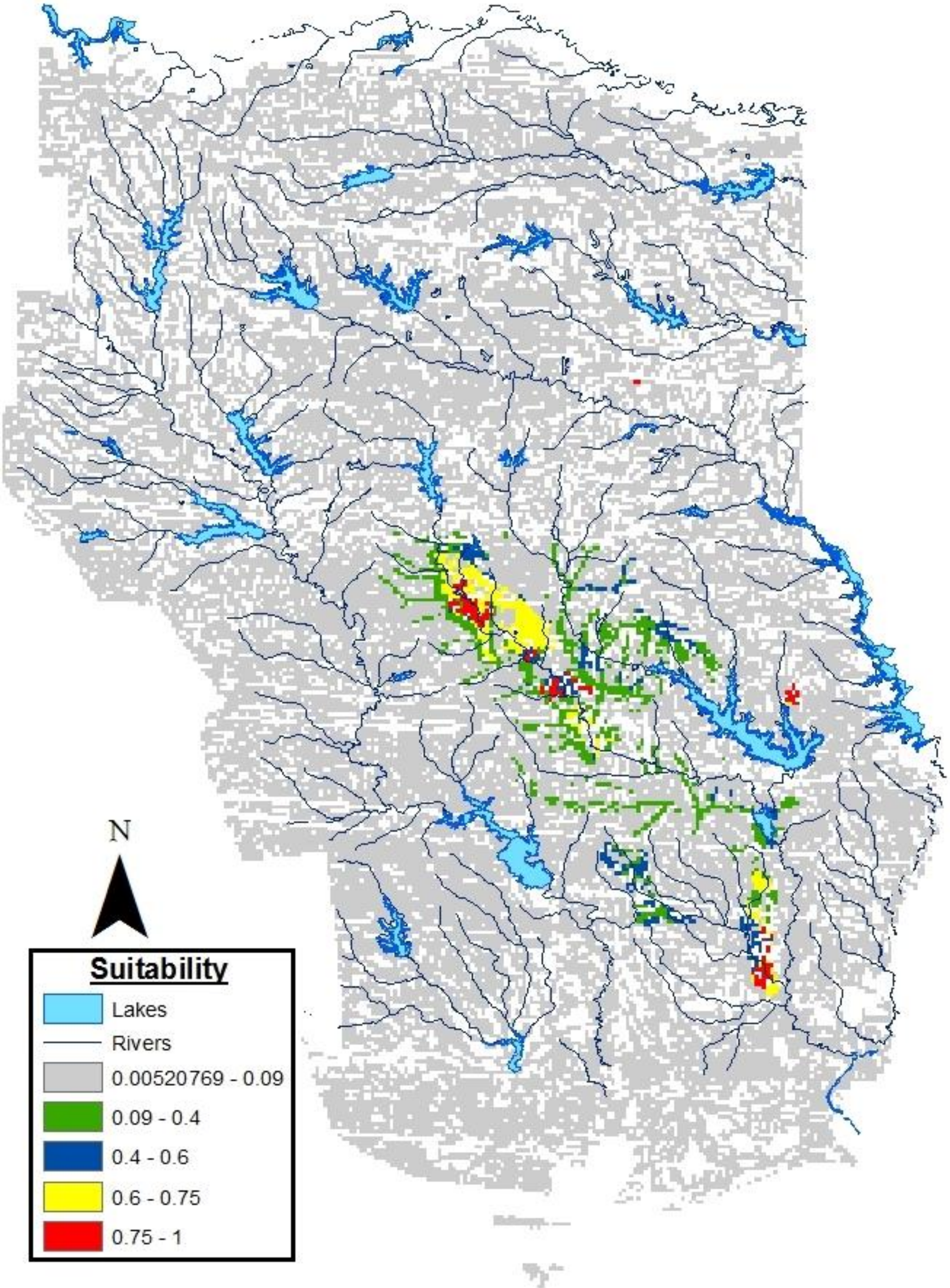

Figure 3.28. The predicted potential suitable habitat for *Pleuroblema riddellii* in East Texas with 20% of the new occurrence points added.

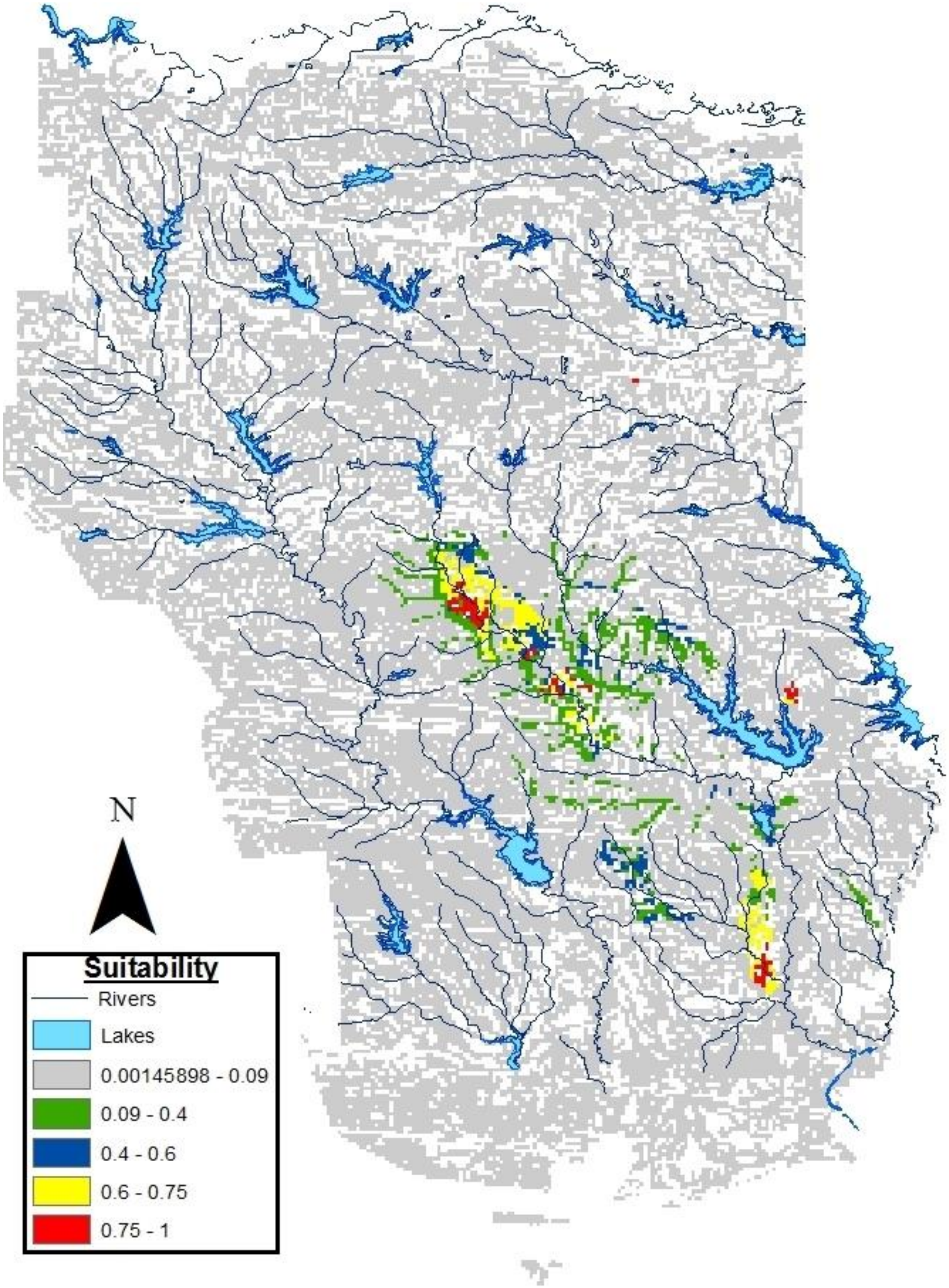

Figure 3.29. The predicted potential suitable habitat for *Pleuroblema riddellii* in East Texas with 40% of the new occurrence points added.

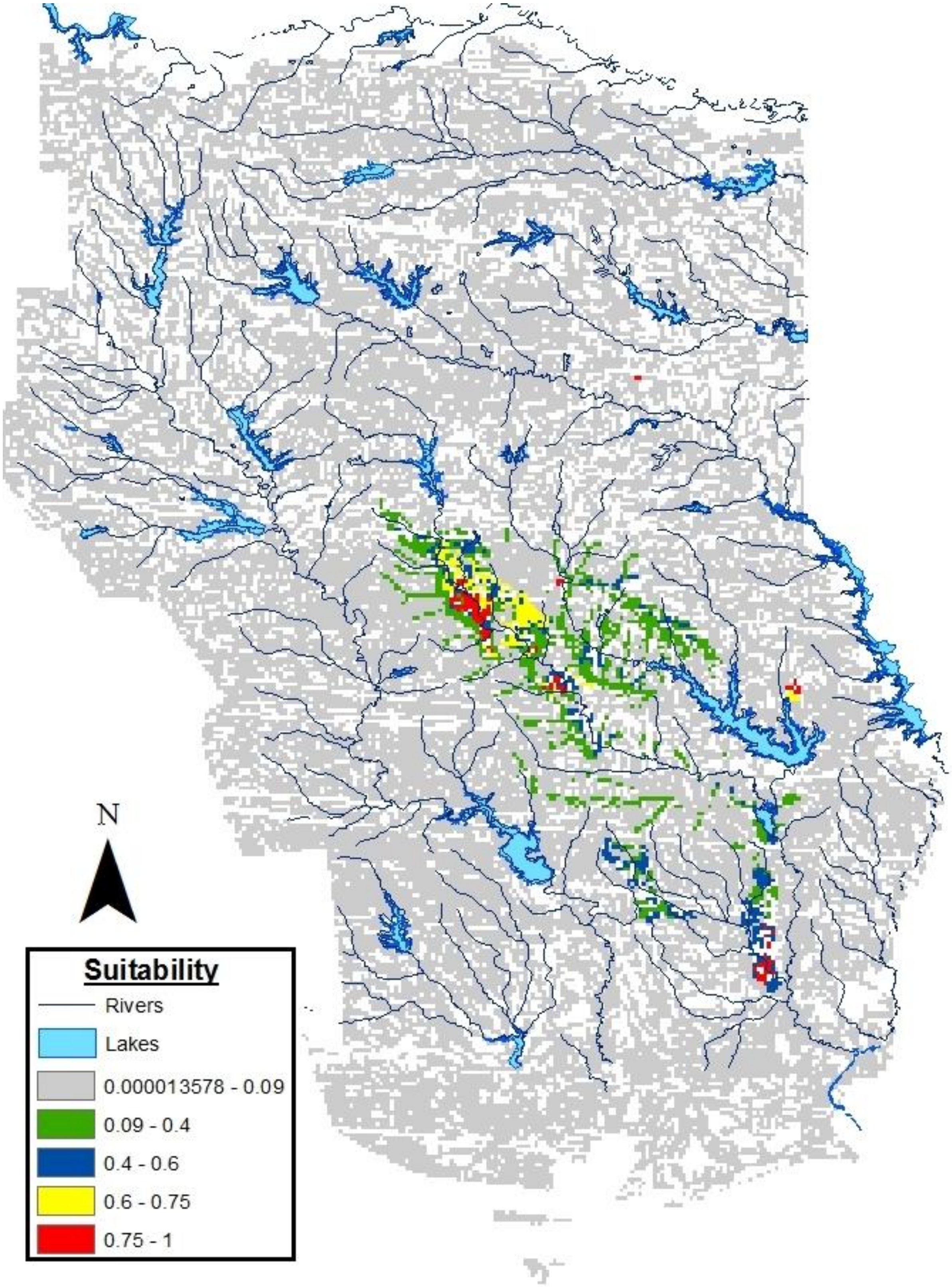

Figure 3.30. The predicted potential suitable habitat for *Pleuroblema riddellii* in East Texas with 60% of the new occurrence points added.

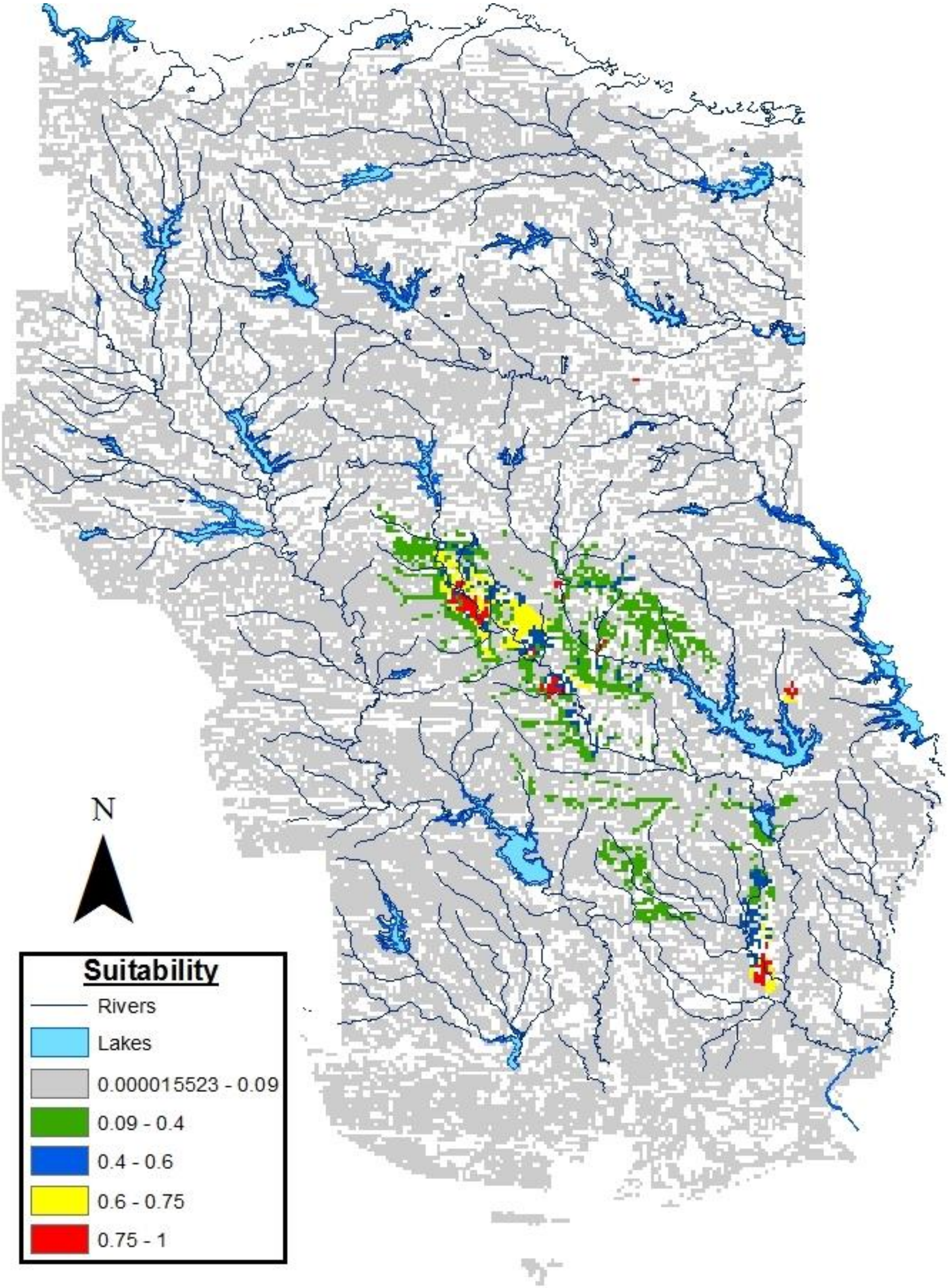

Figure 3.31. The predicted potential suitable habitat for *Pleuroblema riddellii* in East Texas with 80% of the new occurrence points added.

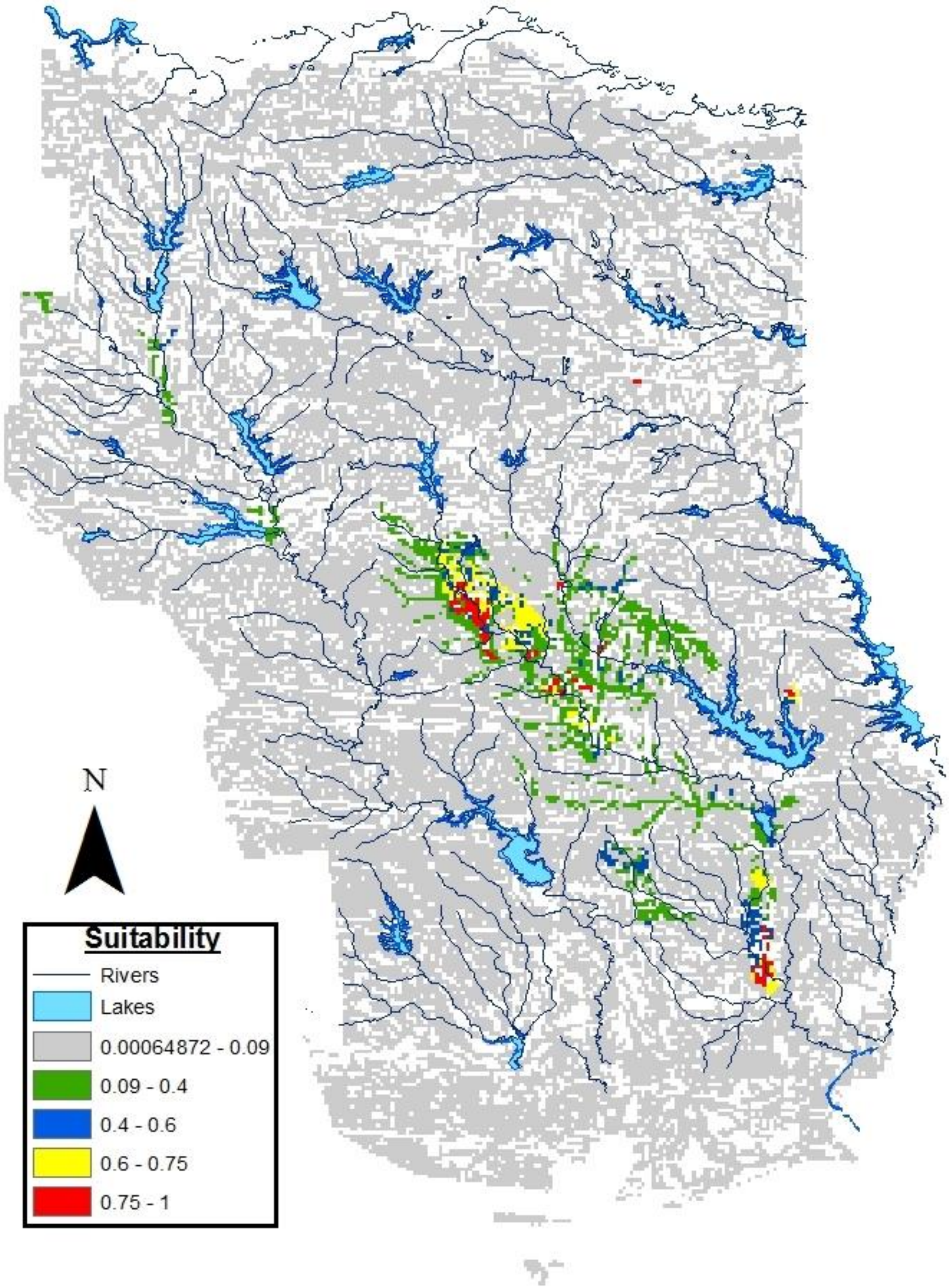

Figure 3.32. The predicted potential suitable habitat for *Pleuroblema riddellii* in East Texas with 100% of the new occurrence points added.

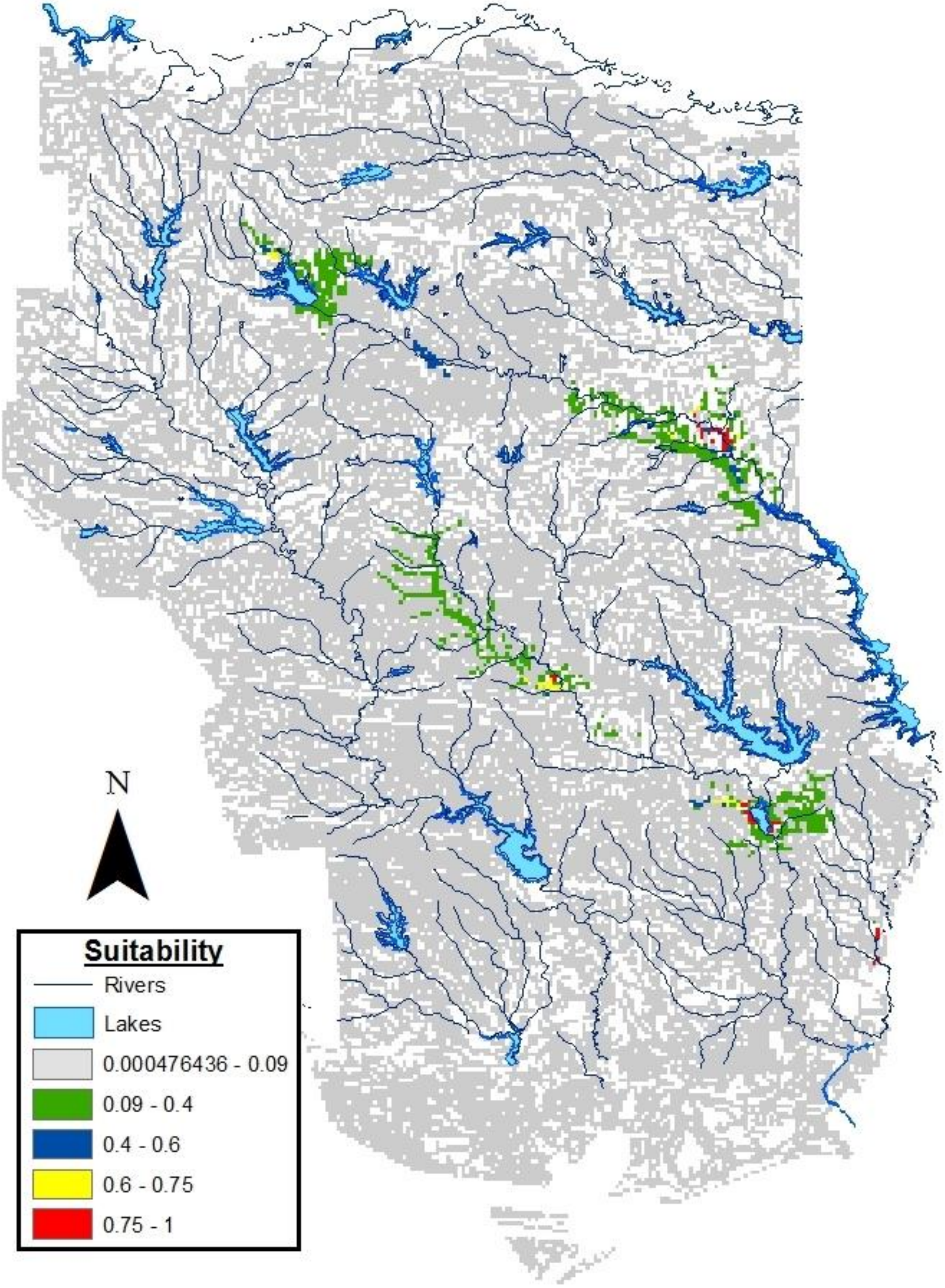

Figure 3.33. The predicted potential suitable habitat for *Potamilus amphichaenus* in East Texas with no new occurrence points added.

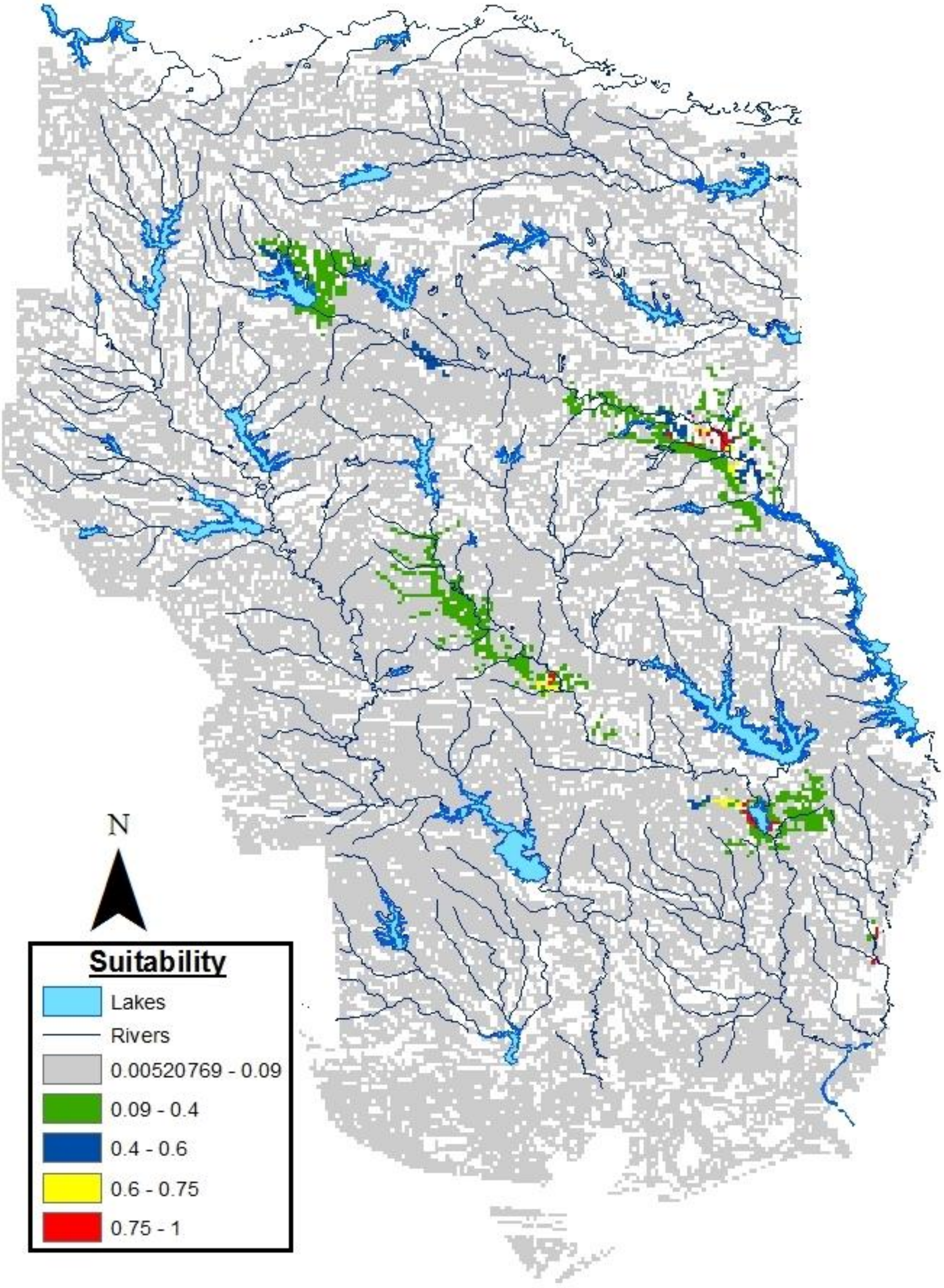

Figure 3.34. The predicted potential suitable habitat for *Potamilus amphichaenus* in East Texas with 20% of the new occurrence points added.

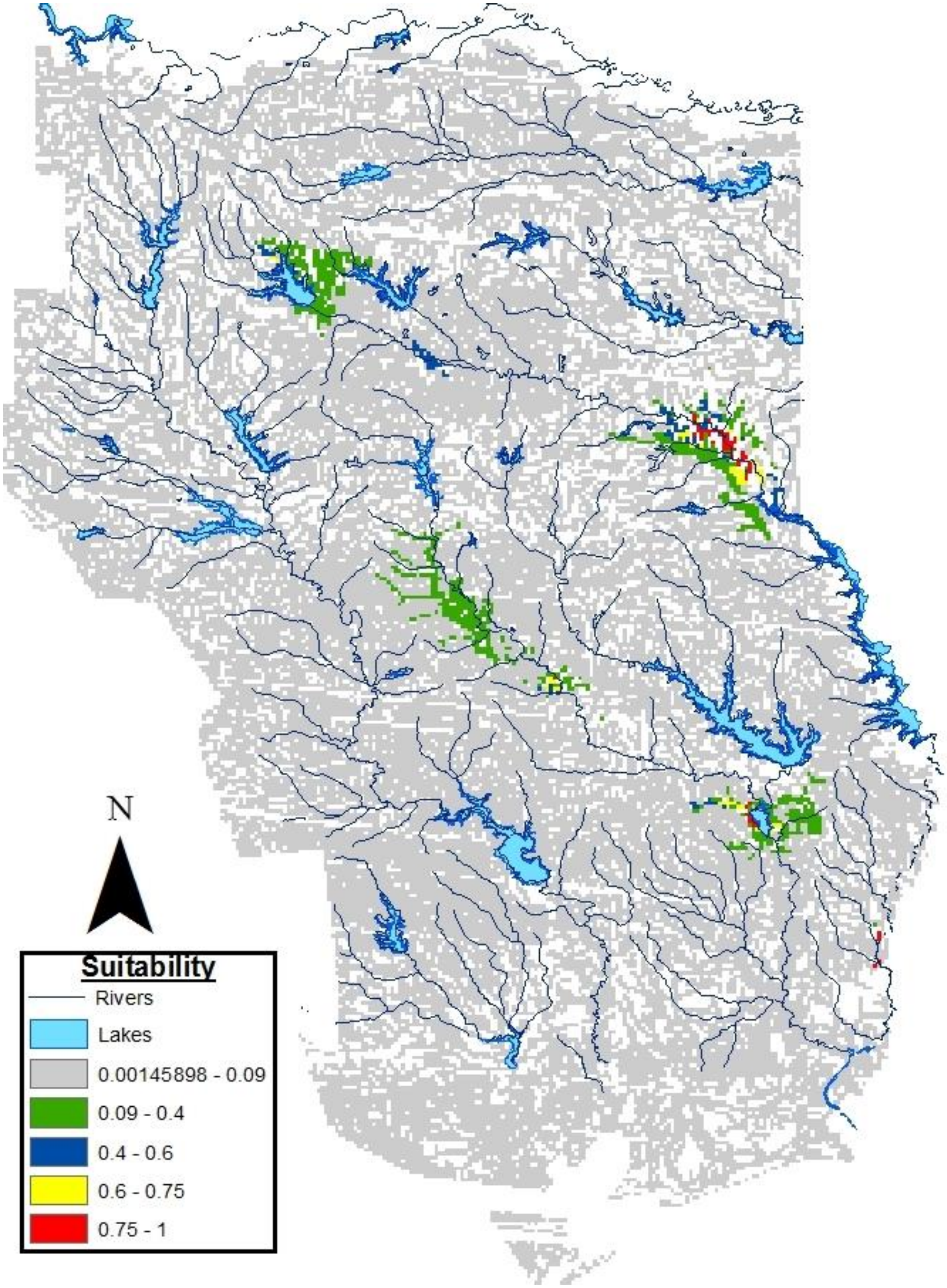

Figure 3.35. The predicted potential suitable habitat for *Potamilus amphichaenus* in East Texas with 40% of the new occurrence points added.

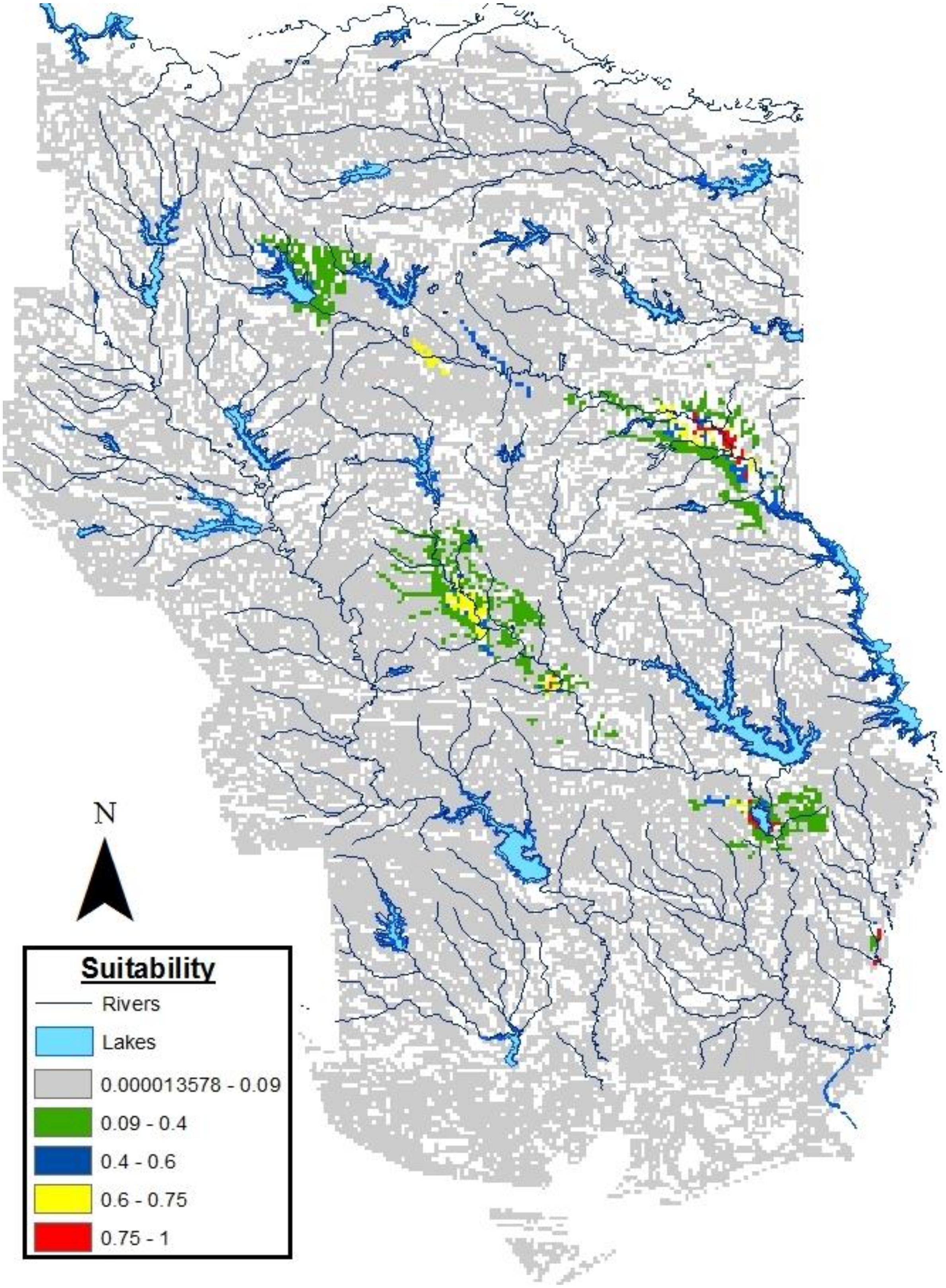

Figure 3.36. The predicted potential suitable habitat for *Potamilus amphichaenus* in East Texas with 60% of the new occurrence points added.

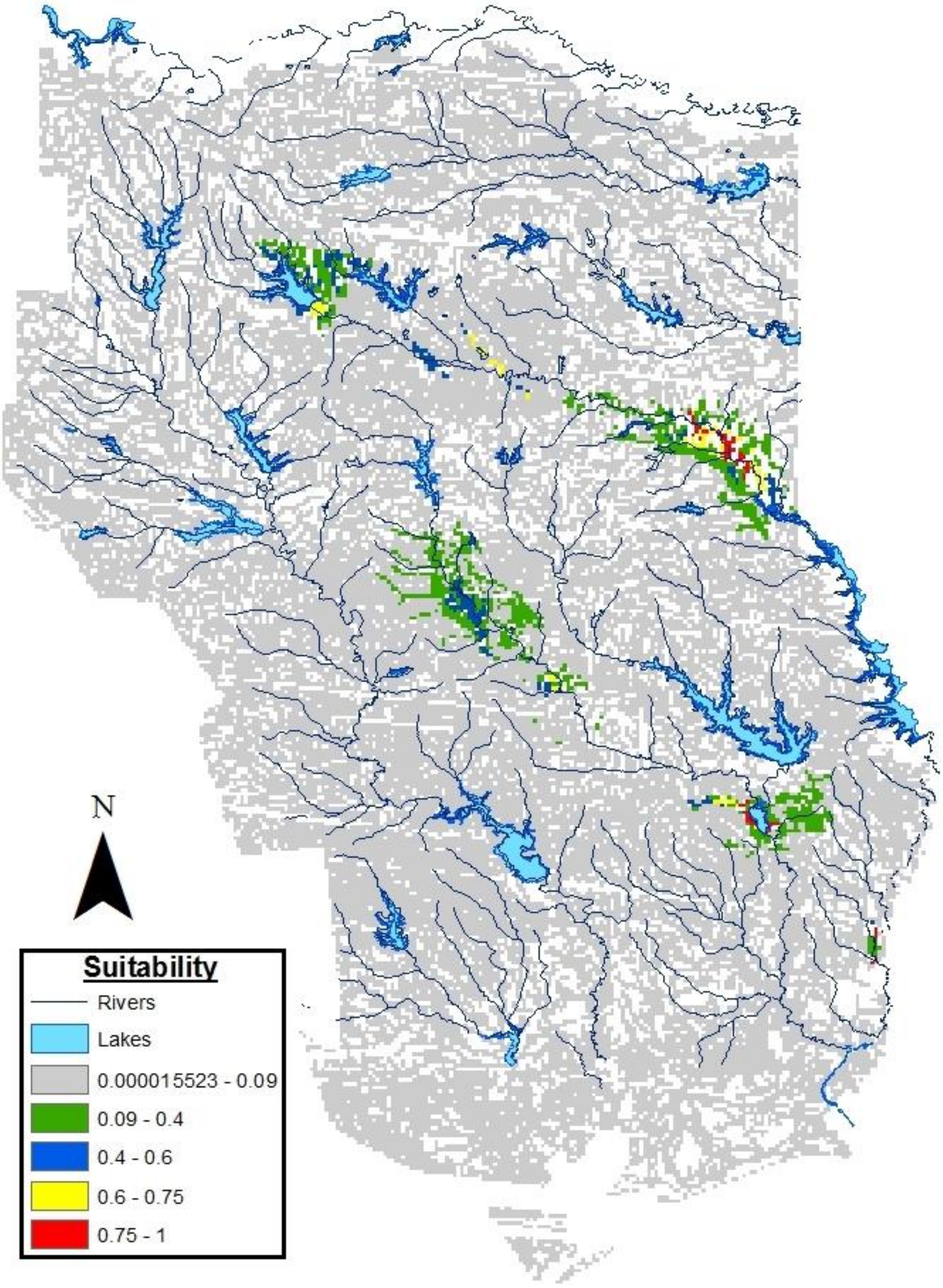

Figure 3.37. The predicted potential suitable habitat for *Potamilus amphichaenus* in East Texas with 80% of the new occurrence points added.

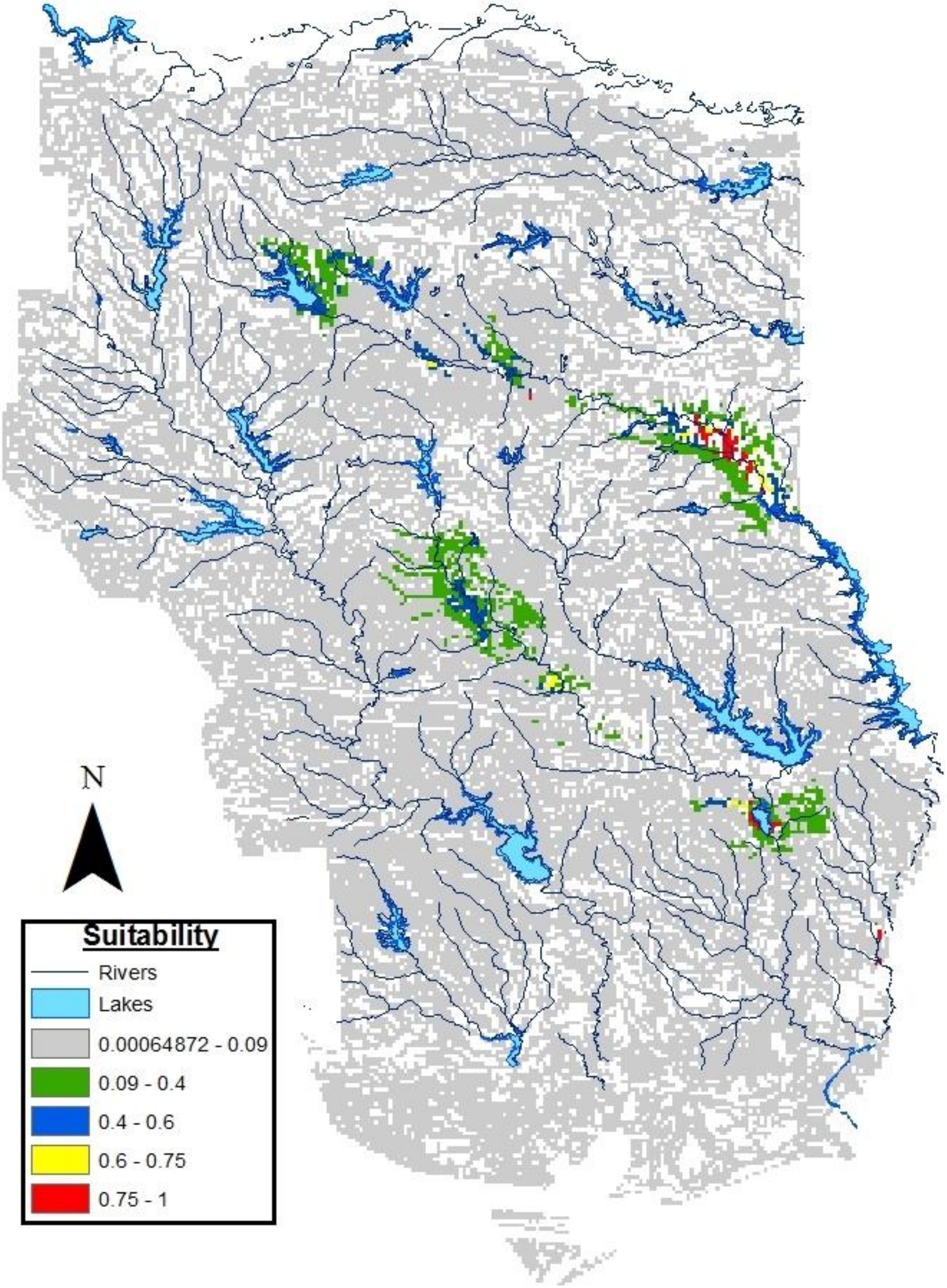

Figure 3.38. The predicted potential suitable habitat for *Potamilus amphichaenus* in East Texas with 100% of the new occurrence points added.

Table 3.1. Summary information for the number of sites a mussel species was found at before and after the new data was added.

Species	# of Original Occurrence Sites	# of New Occurrence Sites	Total Mussels	Total # of Occurrence Sites
Fusconaia askewi	80	65	1474	145
Fusconaia lananensis	27	18	164	45
Lampsilis satura	43	36	118	79
Obovaria jacksoniana	12	5	14	17
Potamilus amphichaenus	52	23	32	75
Pleuroblema riddelli	15	31	401	46

Table 3.2. The P values for the test AUC and test gain, and the r^2 values and equations for all six of the threatened species.

Fusconaia askewi			
	p	r^2	y=
Test AUC	0.61	0.07	y=0.0001x + 0.83
Test gain	0.07	0.60	y= 0.0028x + 1.02
Fusconaia lananensis			
	p	r^2	y=
Test AUC	0.83	0.01	y=2E-05x + 0.90
Test gain	0.02	0.79	y=0.0034x + 1.42
Lampsilis satura			
	p	r^2	y=
Test AUC	0.13	0.47	y=0.0004x + 0.87
Test gain	0.10	0.52	y=0.0048x + 1.22
Obovaria jacksoniana			
	p	r^2	y=
Test AUC	0.01	0.83	y=0.0005x + 0.79
Test gain	0.01	0.85	y=0.0075x + 1.32
Potamilus amphichaenus			
	p	r^2	y=
Test AUC	4.38E-05	0.99	y=0.0009x + 0.81
Test gain	0.09	0.55	y= 0.002x + 1.26
Pleuroblema riddelli			
	p	r^2	y=
Test AUC	0.69	0.05	y=-4E-05x + 0.91
Test gain	0.22	0.34	y= 0.0022x + 1.23

Table 3.3. The linear regression and logistic regression P values for each point at which new occurrence points are added to the original data.

Fusconaia askewi						
	0%	20%	40%	60%	80%	100%
Test	p	p	p	p	p	p
Linear Regression	0.65	0.00	0.00	0.01	0.32	8.49E-04
Logistic Regression	0.01	0.08	0.00	0.06	0.31	7.35E-07
Fusconaia lananensis						
	0%	20%	40%	60%	80%	100%
Test	p	p	p	p	p	p
Linear Regression	0.39	0.05	0.36	0.02	0.02	0.02
Logistic Regression	2.36E-06	5.36E-06	2.10E-04	1.33E-03	0.03	1.10E-07
Lampsilis satura						
	0%	20%	40%	60%	80%	100%
Test	p	p	p	p	p	p
Linear Regression	2.17E-03	3.28E-03	3.12E-04	0.02	0.12	1.71E-06
Logistic Regression	4.89E-05	3.60E-04	2.11E-04	0.02	0.06	3.68E-07
Obovaria jacksoniana						
	0%	20%	40%	60%	80%	100%
Test	p	p	p	p	p	p
Linear Regression	1.62E-05	1.47E-04	4.12E-04	1.06E-03	4.73E-03	1.13E-05
Logistic Regression	0.04	0.10	0.29	0.34	0.55	0.11
Potamilus amphichaenus						
	0%	20%	40%	60%	80%	100%
Test	p	p	p	p	p	p
Linear Regression	0.14	0.21	0.04	0.55	0.68	1.08E-07
Logistic Regression	0.08	0.05	0.02	0.25	0.29	7.35E-07
Pleuroblema riddellii						
	0%	20%	40%	60%	80%	100%
Test	p	p	p	p	p	p
Linear Regression	4.05E-08	6.62E-08	1.99E-06	1.48E-03	6.90E-02	1.52E-07
Logistic Regression	2.41E-09	1.92E-08	3.22E-07	7.84E-05	9.28E-03	8.77E-10

Appendix A: Site descriptions for the occurrence points for all species.

Site	Date	Lat	Long	Drainage	Set #	*Fusconaia askewi*	*Fusconaia lananensis*	*Lampsilis satura*	*Obovaria jacksoniana*	*Potamilus amphichaenus*	*Pleuroblema riddellii*
118	8/3/2006	32.806317	-95.906267	Sabine	1	0	0	0	0	3	0
135	9/27/2007	32.20129	-94.21	Sabine	1	47	0	9	0	4	0
68	7/28/2012	30.17874000	-94.21937000	Neches	1	0	2	0	0	0	0
66	7/26/2012	31.38270000	-94.09953000	Angelina	1	0	0	0	0	0	0
87	9/2/2012	32.61332000	-95.48736000	Sabine	1	0	0	0	0	0	0
20	6/4/2012	31.62297000	-95.26615000	Neches	1	25	0	4	0	0	0
31	6/15/2012	32.09328000	-94.19669000	Sabine	1	0	0	0	0	0	0
61	7/22/2012	31.68671000	-94.38370000	Angelina	1	0	1	0	0	0	0
13	6/1/2012	31.82275000	-94.94548000	Angelina	1	8	0	0	0	0	0
75	10/22/2011	32.66388611	-95.36239444	Neches	1	3	0	0	0	0	0
104	8/12/2005	32.5819	-95.35675	Sabine	1	0	0	0	0	0	0
94	9/11/2012	31.84060000	-95.42436000	Neches	1	101	0	5	3	0	48
77	8/3/2012	31.38442000	-94.95354000	Neches	1	13	0	2	0	0	8
22	6/7/2012	32.50129000	-94.94272000	Sabine	1	91	0	4	0	0	0
47	6/27/2012	31.50051000	-94.30700000	Angelina	1	0	0	0	0	0	0
136	7/27/2010	32.390183	-94.484717	Sabine	1	42	0	1	0	1	0
64	7/25/2012	31.00434000	-94.17068000	Angelina	1	0	0	0	0	0	0
78	8/3/2012	31.38765000	-94.96000000	Neches	1	56	0	5	0	0	40
101	9/17/2012	32.96588300	-96.94449500	Trinity	1	0	0	0	0	0	0
89	9/4/2012	31.70381000	-95.31088000	Neches	1	13	0	5	0	1	8
125	8/31/2006	32.371667	-94.449767	Sabine	1	19	0	0	0	0	0
124	9/1/2007	32.3469	-94.3835	Sabine	1	5	0	0	0	1	0
40	6/20/2012	31.46381000	-94.73476000	Angelina	1	0	0	0	0	0	0
41	6/20/2012	31.45617000	-94.72162000	Angelina	1	2	7	0	0	0	1

Continued on next page

Appendix A Continued

Site	Date	Lat	Long	Drainage	Set #	*Fusconaia askewi*	*Fusconaia lananensis*	*Lampsilis satura*	*Obovaria jacksoniana*	*Potamilus amphichaenus*	*Pleuroblema riddellii*
17	6/3/2012	31.58006700	-95.16773300	Neches	1	9	0	3	0	0	23
7	5/28/2012	32.32073000	-94.33807000	Sabine	1	21	0	1	0	0	0
99	9/24/2012	31.33876000	-94.99613000	Neches	1	0	0	0	0	0	0
133	10/6/2006	32.5443	-95.181483	Sabine	1	3	0	0	0	0	0
60	7/22/2012	31.68895000	-94.38317000	Angelina	2	0	1	0	0	0	0
12	6/1/2012	31.81929000	-94.94239000	Angelina	2	0	0	0	0	0	0
38	6/18/2012	32.27457000	-94.31361000	Sabine	2	0	0	0	0	1	0
14	6/1/2012	31.67040000	-94.95161000	Angelina	2	18	0	5	0	0	4
116	8/31/2006	32.542033	-95.189217	Sabine	2	1	0	0	0	0	0
23	6/7/2012	32.50691000	-94.94538000	Sabine	2	10	0	2	0	0	0
50	7/9/2012	32.87861100	-96.92944400	Trinity	2	0	0	0	0	0	0
10	5/28/2012	32.32699000	-94.34992000	Sabine	2	0	0	0	0	0	0
25	6/13/2012	32.41441	-94.6722	Sabine	2	0	0	0	0	0	0
35	6/18/2012	32.26912	-94.3068	Sabine	2	0	0	0	0	1	0
130	5/29/2006	32.37188	-94.4491	Sabine	2	42	0	3	0	0	0
132	9/26/2006	32.55857	-95.152	Sabine	2	2	0	0	0	0	0
112	9/1/2005	32.60058	-95.3859	Sabine	2	0	0	0	0	0	0
29	6/13/2012	32.42265	-94.7023	Sabine	2	0	0	0	0	0	0
134	9/27/2007	32.34125	-94.3639	Sabine	2	6	0	0	0	2	0
42	6/22/2012	31.49722	-94.8372	Angelina	2	5	2	0	0	0	0
105	8/12/2005	32.58563	-95.3467	Sabine	2	0	0	0	0	0	0
70	7/28/2012	30.17862	-94.2117	Neches	2	16	36	0	0	0	0
63	7/25/2012	31.00439	-94.1779	Angelina	2	0	0	0	0	0	0
103	10/23/2012	33.31903	-96.8922	Trinity	2	0	0	0	0	0	0

Continued on next page

Appendix A Continued

Site	Date	Lat	Long	Drainage	Set #	*Fusconaia askewi*	*Fusconaia lananensis*	*Lampsilis satura*	*Obovaria jacksoniana*	*Potamilus amphichaenus*	*Pleuroblema riddellii*
120	9/1/2007	32.36997	-94.4581	Sabine	2	149	0	8	0	0	0
58	7/20/2012	32.61879	-95.4799	Sabine	2	0	0	0	0	0	0
54	7/18/2012	31.64815	-94.4062	Angelina	2	5	0	0	0	0	0
131	9/26/2006	32.54395	-95.1663	Sabine	2	0	0	0	0	0	0
73	7/31/2012	31.48438	-94.82	Angelina	2	2	1	0	0	0	0
108	9/8/2005	32.57158	-95.2782	Sabine	2	0	0	0	0	0	0
83	8/28/2012	31.58583	-95.1076	Neches	2	0	0	0	0	0	0
86	9/2/2012	32.61444	-95.49	Sabine	2	0	0	0	0	0	0
106	8/12/2005	32.58518	-95.3467	Sabine	3	0	0	0	0	0	0
79	8/26/2012	31.62425	-95.2749	Neches	3	9	0	6	0	0	11
90	9/9/2012	31.71365	-95.3321	Neches	3	47	0	3	0	2	57
32	6/15/2012	32.10169	-94.1866	Sabine	3	0	0	0	0	0	0
36	6/18/2012	32.27501	-94.3101	Sabine	3	0	0	0	0	0	0
93	9/11/2012	31.83665	-95.4206	Neches	3	0	0	1	2	3	16
114	8/25/2005	32.60567	-95.416	Sabine	3	0	0	0	0	0	0
5	5/25/2012	31.7548	-94.9654	Angelina	3	51	12	0	0	0	1
8	5/28/2012	32.32148	-94.3429	Sabine	3	1	0	0	0	4	0
122	9/1/2007	32.34162	-94.3655	Sabine	3	13	0	0	0	1	0
115	8/25/2005	32.60658	-95.4143	Sabine	3	0	0	0	0	0	0
74	7/31/2012	31.48406	-94.8217	Angelina	3	0	0	0	0	0	0
126	8/1/2006	32.55542	-95.1753	Sabine	3	39	0	0	0	0	0
44	6/22/2012	31.49086	-94.8271	Angelina	3	0	0	0	0	0	0
82	8/26/2012	31.62829	-95.2839	Neches	3	9	0	1	0	0	2
107	9/1/2005	32.602	-95.3828	Sabine	3	0	0	0	0	0	0

Continued on next page

Appendix A Continued

Site	Date	Lat	Long	Drainage	Set #	*Fusconaia askewi*	*Fusconaia lananensis*	*Lampsilis satura*	*Obovaria jacksoniana*	*Potamilus amphichaenus*	*Pleuroblema riddellii*
65	7/26/2012	31.38222	-94.1001	Angelina	3	0	0	0	0	0	0
37	6/18/2012	32.27261	-94.31	Sabine	3	0	0	0	0	0	0
3	5/23/2012	31.77177	-95.3986	Neches	3	0	0	1	7	0	11
128	5/22/2006	32.55733	-95.2056	Sabine	3	32	0	0	0	0	0
137	7/27/2010	32.39068	-94.4856	Sabine	3	4	0	0	0	1	0
69	7/28/2012	30.18119	-94.2122	Neches	3	27	57	0	0	0	0
56	7/20/2012	32.61311	-95.4733	Sabine	3	0	0	0	0	0	0
46	6/27/2012	31.50051	-94.3061	Angelina	3	0	0	0	0	0	0
127	8/25/2006	32.54957	-95.1785	Sabine	3	17	0	0	0	0	0
76	8/3/2012	31.3805	-94.9473	Neches	3	20	0	3	0	0	7
19	6/4/2012	31.61998	-95.2626	Neches	3	6	0	0	0	0	0
91	9/9/2012	31.71217	-95.3339	Neches	3	17	0	6	0	1	25
24	6/7/2012	32.5215	-94.9534	Sabine	4	23	0	6	0	0	0
26	6/13/2012	32.4189	-94.6748	Sabine	4	0	0	1	0	0	0
95	9/11/2012	31.8419	-95.4269	Neches	4	48	0	0	0	0	20
92	9/9/2012	31.71444	-95.3363	Neches	4	7	0	0	0	0	5
43	6/22/2012	31.49334	-94.8334	Angelina	4	0	0	0	0	0	0
39	6/20/2012	31.46426	-94.7379	Angelina	4	1	0	0	0	0	5
30	6/15/2012	32.09131	-94.1985	Sabine	4	0	0	0	0	1	0
80	8/26/2012	31.6204	-95.2765	Neches	4	41	0	1	0	0	9
45	6/22/2012	31.49156	-94.8247	Angelina	4	34	10	4	0	0	7
109	9/8/2005	32.5802	-95.2886	Sabine	4	0	0	0	0	0	0
62	7/22/2012	31.68446	-94.3826	Angelina	4	0	1	0	0	0	0
18	6/3/2012	31.58007	-95.1677	Neches	4	12	0	7	0	0	12

Continued on next page

Appendix A Continued

Site	Date	Lat	Long	Drainage	Set #	*Fusconaia askewi*	*Fusconaia lananensis*	*Lampsilis satura*	*Obovaria jacksoniana*	*Potamilus amphichaenus*	*Pleuroblema riddellii*
100	9/24/2012	31.33863	-94.9992	Neches	4	0	0	0	0	0	0
51	7/9/2012	32.87722	-96.9294	Trinity	4	3	0	0	0	0	0
117	8/3/2006	32.80207	-95.8862	Sabine	4	0	0	0	0	1	0
85	9/2/2012	32.61512	-95.4913	Sabine	4	0	0	0	0	0	0
28	6/13/2012	32.42639	-94.6917	Sabine	4	0	0	0	0	0	0
123	9/1/2007	32.57333	-94.6321	Sabine	4	1	0	0	0	0	0
57	7/20/2012	32.61311	-95.4773	Sabine	4	0	0	0	0	0	0
111	8/25/2005	32.60603	-95.4124	Sabine	4	0	0	0	0	0	0
84	8/28/2012	31.58493	-95.1097	Neches	4	0	0	0	0	0	0
49	7/9/2012	32.88056	-96.9297	Trinity	4	0	0	1	0	0	0
21	6/4/2012	31.62597	-95.2749	Neches	4	28	0	2	0	0	2
59	7/22/2012	31.6903	-94.3842	Angelina	4	0	2	0	0	0	0
98	9/24/2012	31.33981	-94.9948	Neches	4	0	0	0	0	0	0
48	6/27/2012	31.508	-94.3061	Angelina	4	11	13	0	0	0	0
34	6/18/2012	32.252	-94.3075	Sabine	4	0	0	1	0	1	0
138	7/27/2010	32.3917	-94.4874	Sabine	5	0	0	0	0	1	0
52	7/10/2012	32.86683	-96.927	Trinity	5	24	0	0	0	0	7
88	9/4/2012	31.70523	-95.3111	Neches	5	19	0	3	0	0	12
81	8/26/2012	31.62514	-95.2807	Neches	5	2	0	0	0	0	1
96	9/23/2012	31.3409	-94.9943	Neches	5	0	0	0	0	0	0
67	7/28/2012	30.18405	-94.215	Neches	5	0	5	0	0	0	0
72	7/31/2012	31.48559	-94.8205	Angelina	5	2	2	0	0	0	0
11	5/28/2012	32.33353	-94.3558	Sabine	5	0	0	0	0	0	0
27	6/13/2012	32.42115	-94.6722	Sabine	5	0	0	0	0	0	0

Continued on next page

Appendix A Continued

Site	Date	Lat	Long	Drainage	Set #	*Fusconaia askewi*	*Fusconaia lananensis*	*Lampsilis satura*	*Obovaria jacksoniana*	*Potamilus amphichaenus*	*Pleuroblema riddellii*
71	7/31/2012	31.48633	-94.8192	Angelina	5	6	5	0	0	0	0
121	9/1/2007	32.54902	-95.1605	Sabine	5	39	0	0	0	0	0
110	8/25/2005	32.60478	-95.4113	Sabine	5	0	0	0	0	0	0
97	9/23/2012	31.34054	-94.9954	Neches	5	0	0	0	0	0	0
4	5/23/2012	31.77208	-95.3987	Neches	5	0	0	1	1	0	1
102	9/20/2012	32.84152	-96.8897	Trinity	5	4	0	0	0	0	0
16	6/3/2012	31.57078	-95.1562	Neches	5	60	0	6	0	0	46
2	5/23/2012	31.7641	-95.3997	Neches	5	0	0	3	0	0	9
129	5/22/2006	32.55308	-95.1997	Sabine	5	69	0	0	0	0	0
119	8/3/2006	32.79882	-95.891	Sabine	5	0	0	0	0	0	0
53	7/18/2012	31.65601	-94.3969	Angelina	5	1	0	0	0	0	0
15	6/1/2012	31.67668	-94.9512	Angelina	5	19	2	1	0	0	1
113	9/1/2005	32.59908	-95.3797	Sabine	5	0	0	0	0	0	0
1	5/23/2012	31.76901	-95.3989	Neches	5	0	0	2	1	0	1
33	6/15/2012	32.1036	-94.1899	Sabine	5	0	0	0	0	0	0
9	5/28/2012	32.33175	-94.3473	Sabine	5	11	0	0	0	2	0
55	7/20/2012	32.60967	-95.4681	Sabine	5	0	0	1	0	0	0
6	5/25/2012	31.7383	-94.9465	Angelina	5	3	5	0	0	0	1
98	9/24/2012	31.33981	-94.9948	Neches	4	0	0	0	0	0	0
48	6/27/2012	31.508	-94.3061	Angelina	4	11	13	0	0	0	0
34	6/18/2012	32.252	-94.3075	Sabine	4	0	0	1	0	1	0
138	7/27/2010	32.3917	-94.4874	Sabine	5	0	0	0	0	1	0
52	7/10/2012	32.86683	-96.927	Trinity	5	24	0	0	0	0	7
88	9/4/2012	31.70523	-95.3111	Neches	5	19	0	3	0	0	12

Appendix B. The test gain, test AUC, and habitat variables for all mussel species.

Species	% of Data	# of Sites	Test gain	Test AUC	Aquifers	Kernel reservoirs	Kernel roads	Kernel springs	Land form	N	Re charge	Soil	Sun light	Top model	Vege-tation
Fusconaia askewi	0	80	1.0289	0.8088	1.65	X	X	0.18	0.39	0.00	0.08	67.04	0.00	0.00	30.66
Fusconaia askewi	20	96	1.0534	0.8576	4.3793	X	X	2.2026	0.0398	0.087	0.0687	66.12	0	0	27.1036
Fusconaia askewi	40	107	0.7515	0.8281	6.9061	X	X	2.8007	0.7006	0.2882	0.2181	67.39	0.0016	0.1391	21.5551
Fusconaia askewi	60	121	0.8361	0.8424	7.8613	X	X	2.9237	0.9275	0.3936	0.2489	64.99	0.0081	0.1527	22.4992
Fusconaia askewi	80	132	0.8581	0.8491	8.6335	X	X	1.3214	0.3209	2.6406	0.1325	63.04	0.0291	0.1802	23.7053
Fusconaia askewi	100	145	0.7441	0.8284	9.0759	X	X	0.747	1.241	0.6117	0.2074	68.81	0	0.6701	18.6337
Fusconaia lananensis	0	27	1.4092	0.8948	3.6124	0.7817	X	X	18.2124	0	0	62.84	0.1348	0	14.4146
Fusconaia lananensis	20	30	1.3362	0.9031	3.3354	0.5836	X	X	15.8121	2.7838	0.3723	55.48	0.1744	0.0409	21.4141
Fusconaia lananensis	40	34	1.3899	0.9104	2.558	0.9027	X	X	14.9314	4.2096	0.4071	55.11	0.1297	0	21.7522
Fusconaia lananensis	60	36	1.1415	0.8952	2.2337	0.478	X	X	17.5409	2.4424	0.2649	59.6	0.1696	0.1248	17.148
Fusconaia lananensis	80	40	1.0978	0.9007	1.5648	0.3198	X	X	19.1465	4.1717	0.2557	60.33	0.1876	0	14.0243
Fusconaia lananensis	100	45	1.1268	0.9017	1.5031	0.4963	X	X	19.1202	3.6936	0.2945	56.08	0.0873	0.0666	18.6625
Lampsilis satura	0	43	1.2527	0.8766	5.7377	X	X	X	1.1982	0	0	70.89	0	0	22.1702
Lampsilis satura	20	53	1.5227	0.8808	6.7588	X	X	X	2.7475	0	0	65.22	0.0156	0.0069	25.2536
Lampsilis satura	40	57	1.1964	0.8652	8.3897	X	X	X	1.0964	0.0148	0.4077	68.11	0.0056	0	21.9775
Lampsilis satura	60	64	1.3124	0.8649	8.8575	X	X	X	2.2333	0.0633	0	66.34	0.0103	0.0019	22.4958
Lampsilis satura	80	72	1.664	0.9032	9.5476	X	X	X	1.8619	0.0619	0.0348	65.33	0.2072	0.0194	22.9342

Continued on next page

Appendix B continued

Species	% of Data	# of Sites	Test gain	Test AUC	Aquifers	Kernel reservoirs	Kernel roads	Kernel springs	Land form	N	Re charge	Soil	Sun light	Top model	Vege-tation
Lampsilis satura	100	79	1.8137	0.9186	10.2092	X	X	X	1.9882	0.0709	0.1338	69.24	0.6178	0.0105	17.7287
Obovaria jacksoniana	0	12	1.2142	0.7779	X	X	X	X	X	X	X	100.00	X	X	X
Obovaria jacksoniana	20	13	1.5419	0.8037	X	X	X	X	X	X	X	100.00	X	X	X
Obovaria jacksoniana	40	13	1.5709	0.8043	X	X	X	X	X	X	X	100.00	X	X	X
Obovaria jacksoniana	60	15	1.9518	0.8255	X	X	X	X	X	X	X	100.00	X	X	X
Obovaria jacksoniana	80	15	1.9388	0.827	X	X	X	X	X	X	X	100.00	X	X	X
Obovaria jacksoniana	100	17	1.9473	0.8254	X	X	X	X	X	X	X	100.00	X	X	X
Pleuroblema riddelli	0	44	1.0733	0.9078	1.7108	0.9718	X	X	7.1414	2.4725	0	67.01	0.0107	1.1651	19.5209
Pleuroblema riddelli	20	50	1.3605	0.913	1.3523	1.3553	X	X	9.7609	0.1266	0.2579	66.11	0.0118	2.4226	18.601
Pleuroblema riddelli	40	51	1.4154	0.9118	1.6542	1.4959	X	X	9.264	0	0.2282	65.47	0.006	2.3701	19.51
Pleuroblema riddelli	60	59	1.4689	0.9181	1.7678	2.0674	X	X	13.0017	0	0.2095	61.77	0.0004	2.4451	18.7373
Pleuroblema riddelli	80	66	1.4066	0.9166	1.6773	2.016	X	X	15.8401	0	2.3164	59.4	0.0179	0.8833	17.848
Pleuroblema riddelli	100	75	1.3419	0.8988	1.4823	1.7704	X	X	14.8559	0	0.1757	60.74	0.0001	0	20.9764
Potamilus amphichaenus	0	23	1.358	0.8153	0.7208	5.2384	X	4.6128	0.0707	0	X	51.61	X	0	37.747
Potamilus amphichaenus	20	28	1.2165	0.8331	0.6075	5.2932	X	6.1561	0.1586	0	X	48.95	X	0	38.8297
Potamilus amphichaenus	40	33	1.2656	0.8443	1.1257	7.1016	X	8.486	2.6909	0.0224	X	34.54	X	0	46.0338
Potamilus amphichaenus	60	40	1.3899	0.8664	0.3001	7.2979	X	9.1131	1.1376	0.0112	X	40.86	X	0.0059	41.274
Potamilus amphichaenus	80	43	1.4322	0.8904	0.6601	7.1347	X	9.7372	6.6756	0.436	X	41.02	X	0.0135	34.3227
Potamilus amphichaenus	100	46	1.4815	0.9017	0.7873	7.4604	X	9.2075	5.4075	0.3444	X	40.94	X	0.1879	35.6672