Mechanisms of Aquatic Species Invasions Across the South Atlantic Landscape Conservation Cooperative Region
Cover: Photograph showing Falling Creek, Upper Suwannee subbasin, Columbia County, Florida (copyrighted by Alan Cressler, 2016; used with permission).
Mechanisms of Aquatic Species Invasions Across the South Atlantic Landscape Conservation Cooperative Region

By Amy J. Benson, Bradley M. Stith, and Victor C. Engel

Prepared in cooperation with the U.S. Fish and Wildlife Service and the South Atlantic Landscape Conservation Cooperative

Scientific Investigations Report 2016–5148

U.S. Department of the Interior
U.S. Geological Survey
Acknowledgments

We would like to acknowledge the effort of all those who have contributed to the U.S. Geological Survey Nonindigenous Aquatic Species Database Program. Thanks to P. Fuller and M. Bonneau for their reviews of this report and to M. Turtora for preliminary data analysis and creation of figure 12.
## Contents

Abstract ............................................................................................................................................... 1
Introduction ........................................................................................................................................ 1
Characterization of USGS NAS Database Data .................................................................................. 2
  Nonindigenous Aquatic Species in the SALCC Region ................................................................. 2
  Freshwater Species ......................................................................................................................... 5
  Marine Species ............................................................................................................................... 7
Patterns of Nonindigenous Aquatic Species Observations Across the SALCC Region ................. 9
Observations of Nonindigenous Aquatic Species in Relation to Land Use Patterns and Geographic Setting ................................................................................................................. 9
Impacted River Systems .................................................................................................................. 16
Potential Future Invasions .............................................................................................................. 19
Statistical Analyses and Point-Process Modeling ........................................................................... 21
  Methods for the Point-Process Model .......................................................................................... 21
Implications for Conservation Practices ......................................................................................... 22
Threats to Protected Areas .............................................................................................................. 23
  Methods for Protected Areas Analysis ......................................................................................... 23
  Results and Discussion of Protected Areas Analysis .................................................................. 25
Factors Associated With Nonindigenous Aquatic Species Richness ........................................... 29
  Methods for Species Richness Analysis ....................................................................................... 29
  Results and Discussion of Species Richness ................................................................................. 33
Network Analysis Tools for Modeling Diffusion Processes ............................................................ 39
  Methods for Modeling the Diffusion Process .............................................................................. 40
  Results and Discussion for the Diffusion Process Modeling ....................................................... 40
Discussion ....................................................................................................................................... 47
Conclusions ...................................................................................................................................... 49
References Cited ............................................................................................................................... 49
Appendices 1–4

1. Descriptions of Species With Possible Ecological Impacts That Have Been Introduced to the South Atlantic Landscape Conservation Cooperative Region in the Past Decade ................................................................................................................................. 56

2. Literature Reviewed for Planned Survey Data of Nonindigenous Aquatic Species ........................................................................................................................................................................ 58

3. Instructions for Importing NHDPlus Shapefile for Use in the R Library Called SSN ........................................................................................................................................................................ 64

4. R Code Developed To Analyze Nonindigenous Aquatic Species Point Data .......................... 66
Figures

1. Map showing study area consisting of the South Atlantic Landscape Conservation Cooperative region ..........................................................3

2. Graph showing cumulative number of nonindigenous aquatic species locations recorded in the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database for the South Atlantic Landscape Conservation Cooperative region by decade, 1850–2010 ........................................................................5

3. Map showing cumulative locations of species recorded in the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database from 1940 through 2013 for the South Atlantic Landscape Conservation Cooperative region ..............................................6

4. Graph showing number of nonindigenous aquatic species introduced by specific introduction methods in the South Atlantic Landscape Conservation Cooperative region as documented in the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database ........................................................................................................7

5. Map showing locations of *Pylodictis olivaris* (flathead catfish) documented in the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database for the South Atlantic Landscape Conservation Cooperative region, 1950–2013 ..................8

6. Map showing percent urban land cover of U.S. Geological Survey subbasins within the South Atlantic Landscape Conservation Cooperative region ........................................10

7. Map showing total number of nonindigenous aquatic species per subbasin recorded in the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database for the South Atlantic Landscape Conservation Cooperative region ..............11

8. Graph showing total number of distinct species per taxa across the 16 upland subbasins with 20 or more nonindigenous aquatic species recorded in the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database for the South Atlantic Landscape Conservation Cooperative region ........................................13

9. Graph showing number of new species introductions reported per decade, pre-1960s to 2000s, in the 16 upland subbasins with 20 or more nonindigenous aquatic species recorded in the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database for the South Atlantic Landscape Conservation Cooperative region ........................................14

10. Graph showing total numbers of distinct species reported in the 10 coastal subbasins with 20 or more nonindigenous aquatic species recorded in the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database for the South Atlantic Landscape Conservation Cooperative region ........................................................................................................15

11. Graph showing number of new species introductions reported by decade in the 10 coastal subbasins with 20 or more nonindigenous aquatic species recorded in the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database for the South Atlantic Landscape Conservation Cooperative region ........................................16

12. Matrix graph showing U.S. Geological Survey (USGS) subbasins with species recorded in the USGS Nonindigenous Aquatic Species (NAS) Database for the South Atlantic Landscape Conservation Cooperative region. Each point represents one or more occurrences of a single species in a specific subbasin ......17

13. Map showing total number of nonindigenous aquatic species recorded in the U.S. Geological Survey (USGS) Nonindigenous Aquatic Species (NAS) Database in USGS subregions in the South Atlantic Landscape Conservation Cooperative region ........................................................................................................18
<table>
<thead>
<tr>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.</td>
<td>Map showing observations recorded in the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database for three fish species with established populations in proximity to the South Atlantic Landscape Conservation Cooperative (SALCC) region</td>
</tr>
<tr>
<td>15.</td>
<td>Map showing protected areas within the South Atlantic Landscape Conservation Cooperative region from the Protected Areas Database of the United States (PAD-US, Conservation Biology Institute Edition)</td>
</tr>
<tr>
<td>16.</td>
<td>Graph showing the number of nonindigenous aquatic species reported from individual protected areas, from the Protected Areas Database of the United States (PAD-US, Conservation Biology Institute Edition) sorted by named areas with highest number of species, for the South Atlantic Landscape Conservation Cooperative region</td>
</tr>
<tr>
<td>17.</td>
<td>Graph showing the number of protected areas with reports of individual nonindigenous aquatic species, sorted by most commonly reported species, for the South Atlantic Landscape Conservation Cooperative region</td>
</tr>
<tr>
<td>18.</td>
<td>Graph showing the potential number of nonindigenous freshwater species reported outside of specific protected areas from the Protected Areas Database of the United States (PAD-US, Conservation Biology Institute Edition) but hydrologically connected by waterways at the subregion level in the South Atlantic Landscape Conservation Cooperative region</td>
</tr>
<tr>
<td>19.</td>
<td>Map showing number of nonindigenous aquatic species per subbasin within the South Atlantic Landscape Conservation Cooperative region</td>
</tr>
<tr>
<td>20.</td>
<td>Map showing housing density per subbasin within the South Atlantic Landscape Conservation Cooperative region</td>
</tr>
<tr>
<td>21.</td>
<td>Map showing elevation range per subbasin within the South Atlantic Landscape Conservation Cooperative region</td>
</tr>
<tr>
<td>22.</td>
<td>Map showing longitude value per subbasin within the South Atlantic Landscape Conservation Cooperative region</td>
</tr>
<tr>
<td>23.</td>
<td>Map showing road density per subbasin within the South Atlantic Landscape Conservation Cooperative region</td>
</tr>
<tr>
<td>24.</td>
<td>Map showing reservoir density per subbasin within the South Atlantic Landscape Conservation Cooperative region</td>
</tr>
<tr>
<td>25.</td>
<td>Map showing Suwannee River drainage (National Hydrography Dataset Plus [NHDPlus] shapefile) and data points representing observations for <em>Corbicula fluminea</em> (Asian clam)</td>
</tr>
<tr>
<td>26.</td>
<td>Time series showing distance from the earliest record to subsequent observations of <em>Corbicula fluminea</em> (Asian clam) in the Suwannee River Basin</td>
</tr>
<tr>
<td>27.</td>
<td>Time series showing distance between consecutive pairs of collection records of <em>Corbicula fluminea</em> (Asian clam) in the Suwannee River Basin</td>
</tr>
<tr>
<td>28.</td>
<td>Time series showing time interval between consecutive pairs of collection records of <em>Corbicula fluminea</em> (Asian clam) in the Suwannee River Basin, irrespective of consecutive pairs being upstream or downstream</td>
</tr>
<tr>
<td>29.</td>
<td>Map showing Suwannee River drainage exhibiting different network distances from the first record of <em>Corbicula fluminea</em> (Asian clam) to simulated points</td>
</tr>
</tbody>
</table>
Tables

1. Nonindigenous aquatic species data summaries for the South Atlantic Landscape Conservation Cooperative region .......................................................... 4
2. Results of Kruskal-Wallis H statistical tests comparing measurement variables of recorded nonindigenous aquatic species and nominal variables in subbasins in the South Atlantic Landscape Conservation Cooperative region ........................................ 12
3. Explanatory variables used in ordinary least squares (OLS) analysis of species richness for subbasins in the South Atlantic Landscape Conservation Cooperative region .......................................................................................................................... 33
4. Ordinary least squares (OLS) results for best models that pass all diagnostic tests with 3, 4, and 5 variables ........................................................................ 33
5. Protected area and distance to first recorded Corbicula fluminea (Asian clam) location in the Suwannee River Basin ................................................................................... 47

Conversion Factors

U.S. customary units to International System of Units

<table>
<thead>
<tr>
<th>Multiply</th>
<th>By</th>
<th>To obtain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td></td>
<td></td>
</tr>
<tr>
<td>inch (in.)</td>
<td>2.54</td>
<td>centimeter (cm)</td>
</tr>
<tr>
<td>inch (in.)</td>
<td>25.4</td>
<td>millimeter (mm)</td>
</tr>
<tr>
<td>mile (mi)</td>
<td>1.609</td>
<td>kilometer (km)</td>
</tr>
<tr>
<td>Area</td>
<td></td>
<td></td>
</tr>
<tr>
<td>square mile (mi²)</td>
<td>259.0</td>
<td>hectare (ha)</td>
</tr>
<tr>
<td>square mile (mi²)</td>
<td>2.590</td>
<td>square kilometer (km²)</td>
</tr>
<tr>
<td>acre-foot (acre-ft)</td>
<td>1,233</td>
<td>cubic meter (m³)</td>
</tr>
<tr>
<td>acre-foot (acre-ft)</td>
<td>0.001233</td>
<td>cubic hectometer (hm³)</td>
</tr>
</tbody>
</table>

International System of Units to U.S. customary units

<table>
<thead>
<tr>
<th>Multiply</th>
<th>By</th>
<th>To obtain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td></td>
<td></td>
</tr>
<tr>
<td>centimeter (cm)</td>
<td>0.3937</td>
<td>inch (in.)</td>
</tr>
<tr>
<td>millimeter (mm)</td>
<td>0.03937</td>
<td>inch (in.)</td>
</tr>
<tr>
<td>meter (m)</td>
<td>3.281</td>
<td>foot (ft)</td>
</tr>
<tr>
<td>kilometer (km)</td>
<td>0.6214</td>
<td>mile (mi)</td>
</tr>
<tr>
<td>kilometer (km)</td>
<td>0.5400</td>
<td>mile, nautical (nmi)</td>
</tr>
<tr>
<td>meter (m)</td>
<td>1.094</td>
<td>yard (yd)</td>
</tr>
<tr>
<td>Volume</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cubic meter (m³)</td>
<td>264.2</td>
<td>gallon (gal)</td>
</tr>
<tr>
<td>cubic meter (m³)</td>
<td>35.31</td>
<td>cubic foot (ft³)</td>
</tr>
<tr>
<td>cubic meter (m³)</td>
<td>1.308</td>
<td>cubic yard (yd³)</td>
</tr>
<tr>
<td>cubic meter (m³)</td>
<td>0.0008107</td>
<td>acre-foot (acre-ft)</td>
</tr>
</tbody>
</table>

Temperature in degrees Celsius (°C) may be converted to degrees Fahrenheit (°F) as °F = (1.8 × °C) + 32.
Mechanisms of Aquatic Species Invasions Across the South Atlantic Landscape Conservation Cooperative Region

By Amy J. Benson, Bradley M. Stith, and Victor C. Engel

Abstract

Invasive species are a global issue, and the southeastern United States is not immune to the problems they present. Therefore, various analyses using modeling and exploratory statistics were performed on the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database with the primary objective of determining the most appropriate use of presence-only data as related to invasive species in the South Atlantic Landscape Conservation Cooperative (SALCC) region. A hierarchical model approach showed that a relatively small amount of high-quality data from planned surveys can be used to leverage the information in presence-only observations, having a broad spatial coverage and high biases of observer detection and in site selection. Because a variety of sampling protocols can be used in planned surveys, this approach to the analysis of presence-only data is widely applicable. An important part of the management of natural landscapes is the preservation of designated protected areas. When the hydrologic connection was considered in this analysis, the number of potential invaders that could spread to each protected area within the SALCC region was greatly increased, with a mean exceeding 30 species and the maximum reaching 57 species. Nearly all protected areas are hydrologically connected to at least 20 nonindigenous aquatic species. To examine possible factors which may contribute to nonindigenous aquatic species richness in the SALCC region, a set of exploratory statistics was employed. The best statistical model that included a combination of three anthropogenic variables (densities of housing, roads, and reservoirs) and two environmental variables (elevation range and longitude) explained approximately 62 percent of the variation in introduced species richness. Highest nonindigenous aquatic species richness occurred in the more upland, mountainous regions, where elevation range favored reservoirs and attracted urban centers. Lastly, patterns seen in a diffusion model may reflect less about the diffusion process of the organism and more about the opportunistic nature of the data collection process. These results of the model are considered exploratory in nature.

Introduction

In the United States there is a long and well-documented history of anthropogenic introductions of aquatic organisms into environments where they were not historically known to exist (Lachner and others, 1970; U.S. Congress, Office of Technology Assessment, 1993; Simberloff, 1996; Benson, 2000). Only by virtue of their introduction into novel ecosystems are these organisms referred to by terms such as “alien,” “exotic,” “invasive,” “naturalized,” “nonindigenous,” and “nonnative” (Shafland and Lewis, 1984; Binggeli, 1994; Colautti and MacIsaac, 2004). Some scientists have made the case that as a result of these introductions we are losing biological diversity and becoming a biotically homogenized world (Vitousek and others, 1997; McKinney, 1998; Rahel, 2002; Smart and others, 2006; Olden and others, 2009). This homogenization is essentially an increasing similarity of species at a location through time (Olden and Poff, 2004); however, such a determination is dependent on spatial, temporal, and data resolution according to Olden and Poff (2003). There is a danger of this homogenization process occurring any time species are introduced, especially those which are considered broadly adapted generalists (McKinney and Lockwood, 1999). The cause of this homogenization may be, for example, simply the result of stocking of similar fish species across this country, including remote regions with sparse human populations (Rahel, 2002). Not only is there an ecological cost of these introductions, but there is also an economic cost of aquatic invasive species that has been estimated at more than $2 billion per year (Pimentel and others, 2005).

Another contributing factor that could cause a similar loss of biodiversity is climate change. It has been shown that plants known to be invasive in the United States respond positively
to increasing carbon dioxide levels in controlled monocultured environments, but are less responsive in a diverse community, therefore making predictions difficult (Dukes and Mooney, 1999). Changing climate can also come in the form of novel precipitation patterns where additional rainfall in arid climates could promote the growth of an invader (Dukes and Mooney, 1999). Global change of increasing temperature can shift an invasive species’ range northward, as well as promote native species to become invasive (Dukes and Mooney, 1999). For example, warmwater fish species that have been broadly stocked into novel habitats may expand their distributions as waters gradually warm (Eaton and Scheller, 1996).

Another form of global change impacting the distributions of invasive species is the growth of urban areas, or urbanization (McKinney, 2006). The world is quickly becoming more urbanized as an increasing number of people are moving to highly populated areas, thereby forcing cities to expand (Mcdonald and others, 2008). It has been found that along with urbanization and high human population densities of cities comes an increase in nonindigenous aquatic species (Spear and others, 2013). McKinney (2002) found that the number of nonnative species tends to increase whereas the number of native species tends to decrease as one moves towards an urban center, a trend which is often due to habitat disturbance, as well as propagule pressure, especially of cultivated plants. Because human population is estimated to reach 9 billion by 2044 (U.S. Census Bureau, International Programs, 2013), it would not be overly presumptive to assume there will continue to be biological introductions as a result of increased human activities. How will these natural resources be impacted as human populations and urban areas grow in the future? For many, if not most, conservationists, this trend of losing unique biological communities is not desirable.

This study represents the first attempt to utilize the U.S. Geological Survey (USGS) Nonindigenous Aquatic Species (NAS) Database (hereinafter referred to as the “USGS NAS Database”) to inform a Landscape Conservation Cooperative (a consortium of public and private agencies) conservation blueprint. Nonindigenous aquatic species data are generated from observations reported in the literature, from vouchered museum specimens, and by casual observations of individual people, including those from the general public. Many of the records are generated from localized biological surveys or natural resource assessments. These data are compiled in the USGS NAS Database (Williams and Jennings, 1991). Because of the diversity of data sources, the location and timing of observations recorded in the USGS NAS Database are not evenly distributed across the South Atlantic Landscape Conservation Cooperative (SALCC) region, which extends from southern Virginia to northern Florida (fig. 1). As a result, the inferences regarding the timing of nonindigenous aquatic species introductions and their dispersal across the SALCC should not be equated to inferences that might be drawn from dedicated monitoring programs. The USGS NAS Database indicates only where species have been observed (presence-only); other data sources are needed to determine where these species are not located (absence) in order to estimate population numbers. For some species, the numbers of observations vary over time, as does the spatial resolution of the location of the recorded observations. With these constraints in mind, we show how the USGS NAS Database can be uniquely applied towards an improved understanding of the relations among observations of nonindigenous aquatic species, cultural practices, geographic features, urbanization, and water resources.

The objectives of this study were fivefold: (1) characterize the USGS NAS Database as it relates to the SALCC region, (2) develop a new statistical modeling approach to analyze its presence-only data, (3) analyze threats to protected areas on the basis of known occurrences of nonindigenous aquatic species, (4) examine factors which may have contributed to new invasions and increases in species richness, and (5) model the diffusion of organisms in a river network.

Characterization of USGS NAS Database Data

Nonindigenous Aquatic Species in the SALCC Region

The southeastern United States is no exception to the invasion process and is in fact highly invaded with both extracontinental and transplanted North American species. The USGS NAS Database contains nearly 16,000 records for 279 nonindigenous aquatic species recorded in the SALCC region since 1850 (Benson, 2016). These species represent a broad range of taxa from amphibians and reptiles to fishes, mammals, invertebrates, and plants (table 1). Fishes are the most common type of nonindigenous aquatic species recorded in the SALCC region, followed by plants.

The total number of nonindigenous aquatic species reported in the SALCC region has steadily increased from 1850 to 2010 (fig. 2). On a decadal time scale, the rate of increase follows an exponential function ($R^2=0.95$) for the period from 1850 to 1980. From 1980 to 2010, the increase has been linear, with 50 new species reported for each decade in this interval. Although the observation effort is not constant over time, if this trend continues, as many as 46 new nonindigenous aquatic species introductions may be reported for the SALCC region by 2020.
Figure 1. Study area consisting of the South Atlantic Landscape Conservation Cooperative region.
Mechanisms of Aquatic Species Invasions Across the South Atlantic Landscape Conservation Cooperative Region

The earliest documented introductions in the SALCC region from the period 1850 to 1940 (fig. 3) include species such as *Ambloplites rupestris* (rock bass), *Micropterus salmoides* (largemouth bass), and *Pomoxis nigromaculatus* (black crappie), which were intentional introductions for sport fishing. During this same early period *Tinca tinca* (tench) and *Cyprinus carpio* (common carp) were introduced to supplement the food supply for human consumption. When viewed together, the distribution of observation records across the SALCC region for both plants and animals (fig. 3) reveals a few noteworthy patterns. First, until 1940 very few nonindigenous aquatic species were documented in the region. By 1960, relatively larger numbers of observations were reported in the Neuse River Basin of North Carolina. Closer examination of the data showed that these were observations of largemouth bass obtained from the North Carolina Museum of Natural Sciences indicating that perhaps the drainage was surveyed in the 1950s. Not known as native to this basin, largemouth bass were stocked for recreational fishing purposes. The number of observations reported from the SALCC region increased substantially from 1960 to 1980, as did the numbers reported along the Savannah River. The period from 1980 to 2000 was marked by a large, region-wide increase in the number of nonindigenous aquatic species observations primarily due to the increased availability of significant publications (fig. 3). As a result, more than one-half of the total number of observation records since 1850 is from this period.

Since 2000, there has been a marked increase in both the number of observations and the number of species reported in some coastal areas. Observations of nonindigenous aquatic species are now widespread at variable densities throughout the SALCC region, with the exception of some portions of central and southeast Georgia. The lack of observation records for these remote areas warrants direct investigation, as their absence may be due to the relatively low human population density (and correspondingly low probability of observing nonindigenous aquatic species) in addition to low numbers of nonindigenous aquatic species in these areas. A lack of planned surveys could also contribute to the low numbers of observations.

Table 1. Nonindigenous aquatic species data summaries for the South Atlantic Landscape Conservation Cooperative region.

<table>
<thead>
<tr>
<th>Taxa</th>
<th>Native transplant</th>
<th>Exotic**</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amphibians-frogs</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Annelids-polychaetes</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Bryozoans</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Coelenterates</td>
<td>0</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Crustaceans-amphipods</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Crustaceans-barnacles</td>
<td>1</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Crustaceans-cladocerans</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Crustaceans-copepods</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Crustaceans-crabs</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Crustaceans-crayfish</td>
<td>9</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Crustaceans-isopods</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Crustaceans-lobsters</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Crustaceans-shrimp</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Crustaceans-tanaids</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Fishes</td>
<td>107</td>
<td>38</td>
<td>145</td>
</tr>
<tr>
<td>Mammals</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Mollusks-bivalves</td>
<td>1</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Mollusks-gastropods</td>
<td>5</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>Mollusks-nudibranchs</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Plants</td>
<td>1</td>
<td>33</td>
<td>34</td>
</tr>
<tr>
<td>Platyhelminthes</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Reptiles-crocodilians</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Reptiles-lizards</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Reptiles-snakes</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Reptiles-turtles</td>
<td>8</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>Tunicates</td>
<td>0</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

*Native to the United States, but transplanted outside its native range.

**Not native to the United States.
Figure 2. Cumulative number of nonindigenous aquatic species locations recorded in the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database for the South Atlantic Landscape Conservation Cooperative region by decade, 1850–2010.

Information on how introductions occur is documented in the USGS NAS Database and is based on literature accounts or best professional judgement. Similar to those in most areas of the United States, the methods of introductions are derived from human activity and are diverse in the SALCC region (fig. 4). The dominant methods arise from the creation of recreational fishing opportunities through stocking programs, especially in man-made reservoirs.

**Freshwater Species**

There are 228 freshwater species recorded in the USGS NAS Database for the SALCC region including 142 fish species, 35 plants, 14 crustaceans, 13 reptiles, 12 mollusks, 3 amphibians, 2 mammals, and 7 miscellaneous invertebrates. The introduced fishes are dominated by members of the Cyprinidae family of minnows, both native and nonnative species, which include *Ctenopharyngodon idella* (grass carp), common carp, *Hypophthalmichthys molitrix* (silver carp), and *H. nobilis* (bighead carp); *Carassius auratus* (goldfish); and 15 members of the genus *Notropis* (eastern shiners). Other notable fishes that have been documented in the SALCC region include 19 members of the Centrarchidae family and members of the Cichlidae family such as *Astronotus ocellatus* (oscar), *Rocio octofasciata* (Jack Dempsey), *Oreochromis niloticus* (Nile tilapia), and *Tilapia zillii* (redbelly tilapia). The group of nonindigenous fishes found in the SALCC region also includes species that can tolerate a wide range of salinities for either short or long periods of time such as *Morone saxatilis* (striped bass) and *Osmerus mordax* (rainbow smelt).

An example of a nonindigenous freshwater fish species with known ecological impacts that is widespread throughout the SALCC region is *Pylodictis olivaris* (flathead catfish; fig. 5). Native to the Mississippi River Basin, this very large fish can reach 1.5 meters in length and is known to prey heavily on sunfishes of the genus *Lepomis* and on crayfishes (Quinn, 1988). The flathead catfish is believed to have been introduced into the Flint River by fishers in about 1950 (Quinn, 1988). In the early 1960s, they were introduced into J. Strom Thurmond Reservoir and Lake Marion in South Carolina. This species typically disperses throughout the river systems into which it is introduced, and this pattern is reflected with high fidelity in the USGS NAS Database. Records of this species from isolated water bodies are significantly less common primarily because they cannot reproduce in those environments. Descriptions of other freshwater species with possible ecological impacts that have recently been introduced to the SALCC region can be found in appendix 1.
Figure 3. Cumulative locations of species recorded in the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database from 1940 through 2013 for the South Atlantic Landscape Conservation Cooperative region.
Characterization of USGS NAS Database Data

Marine Species

There are 51 nonindigenous marine vertebrate and invertebrate species within the greater SALCC region that have been recorded in the USGS NAS Database. Only four nonindigenous marine fish species have been recorded in the SALCC region, all of which are also nonindigenous to waters of the United States. They include Pterois volitans (red lionfish), P. miles (devil firefish), Naso lituratus (orangespine unicornfish), and Gramma leto (fairy basslet). Of these four species, both the red lionfish and devil firefish are established in U.S. waters. The majority of the nonindigenous marine species are invertebrates, consisting of annelids, bryozoans, coelenterates, crustaceans, mollusks, and tunicates. These include organisms such as Phyllorhiza punctata (Australian spotted jellyfish) and Perna viridis (Asian green mussel). Similar to the fish species recorded in these marine areas, nearly all of these invertebrate species are nonnative not only to the SALCC region but also to all waters of the United States.

The most notorious marine invaders of recent years are red lionfish and devil firefish, from the Indo-Pacific region of the world. In little more than a decade, these two species together have established an extremely prolific population in the western Atlantic Ocean along the U.S. coast from North Carolina to Florida, as well as the entire Gulf of Mexico and Caribbean regions (Schofield, 2010). Direct predation of prey items by the red lionfish and devil firefish includes at least 41 fish species along with several species of crustaceans and mollusks (Morris and Akins, 2009). A relative newcomer with a similar nonindigenous range as the red lionfish and devil firefish in the western Atlantic is Penaeus monodon (Asian tiger shrimp) (Fuller and others, 2014). Although the Asian tiger shrimp spread quickly, it may be too early in this species’ invasion to realize any adverse effects on native shrimp populations. Another marine invader is Rapana venosa (veined rapa whelk). First collected in the region in 1988, it is a large, predatory aquatic snail from the western Pacific Ocean and is now found in southeastern Virginia near the mouth of the Chesapeake Bay. This species provides a new threat by preying directly upon native mussels and oysters (Harding and others, 2007).
Figure 5. Locations of *Pylodictis olivaris* (flathead catfish) documented in the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database for the South Atlantic Landscape Conservation Cooperative region, 1950–2013.
Patterns of Nonindigenous Aquatic Species Observations Across the SALCC Region

To maximize the information provided in the following spatial analyses we include both the number of USGS NAS Database records and the number of species associated with a location. Interpreting these results requires careful consideration of the nature of presence-only data. First, the number of opportunities to sight a nonindigenous species in a water body tends to increase with nearby human populations, depending in part on the activities in which they engage (for example, fishing) while on site. Therefore, we consider the number of records to be positively correlated to the number of opportunities to sight a nonindigenous aquatic species at any particular location. Second, individual species will vary in their detectability and observation frequency because of biological and behavioral differences: small, rare, or cryptic taxa are generally reported in lower numbers than are larger, more numerous, and conspicuous taxa. Last, the ability of nonindigenous aquatic species to persist in the environment after they have been reported varies considerably. For these reasons, the number of nonindigenous aquatic species observations reported over time should not be equated with their abundance at a location, although it is likely that these two measures are related for many species. Similarly, the number of species reported for a location does not necessarily represent the total number of nonindigenous aquatic species present. However, for areas with high numbers of opportunities to observe nonindigenous aquatic species (for example, near populated areas) our experience suggests that the two are closely related.

Data on land use categories (U.S. Geological Survey, National Gap Analysis Program, 2011) were merged with hydrography (U.S. Geological Survey, 2007–14), population densities (U.S. Census Bureau, 2000), and a subset of nonindigenous aquatic species location data for the SALCC region derived from the USGS NAS Database as a first step towards identifying the landscape characteristics that exhibit a strong correlation with the distribution of nonindigenous aquatic species. In this classification scheme, the United States is divided and then subdivided into successively smaller hydrologic units (HU) that are classified into four levels. The HUs are nested within each other from the largest geographic area (region) to the smallest geographic area (subbasin) (U.S. Geological Survey and U.S. Department of Agriculture, Natural Resources Conservation Service, 2013). Each HU is identified by a unique hydrologic unit code (HUC) consisting of 2, 4, 6, and 8 digits based on the four levels of classification (region, subregion, basin, and subbasin, respectively) in the HU system.

Using the 112 8-digit HUC (HUC8) hydrologic delineations (hereinafter referred to as “subbasins”) that were identified with a higher percentage of their area within the SALCC region (fig. 6), we quantified the percent of urban land cover in each subbasin and classified them according to their proximity to, or intersection with, urban areas containing populations greater than 50,000 people. Subbasins contacting the coastline were classified separately from upland subbasins, and subbasins containing large reservoirs (greater than 10 square miles) were distinguished from subbasins without reservoirs.

We identified 93 reservoirs in the SALCC region in this investigation. The number of reservoirs is significant because this type of water body is known to provide suitable habitat for many nonindigenous aquatic species (Johnson and others, 2008). Many reservoirs, including those in the SALCC region, are found near urban areas and managed according to multiuse criteria. Reservoirs are often intentionally stocked with nonindigenous recreational sport fish and are frequently the site of bait bucket and aquarium release methods of introduction. High numbers of introductions and frequent visitation generate relatively high opportunities for observations in reservoirs, and both the numbers of records and the numbers of species observed in the subbasins containing reservoirs are therefore expected to be larger than in other areas. Similarly, both the numbers of observations and the numbers of species in coastal zones are expected to be relatively high compared to other areas. Coastal subbasins in the SALCC region contain freshwater, brackish, and marine environments and therefore provide habitat conditions capable of supporting diverse and numerous taxa. Coastal zones also frequently exhibit high human population densities that lead, in turn, to increased opportunities for observations and numerous introduction methods. Since 2002, there have been approximately 450 new introductions of 123 unique aquatic species into any one of the 112 subbasins (including the coastal marine areas as one subbasin) in the SALCC region.

Observations of Nonindigenous Aquatic Species in Relation to Land Use Patterns and Geographic Setting

Distinct patterns across the SALCC region are apparent in the total number of species observed per subbasin (fig. 7). The subbasins with the highest reported diversity of nonindigenous aquatic species are found in the western and, in particular, the northwestern portions of the SALCC region. These portions are characterized by both the presence of reservoirs and the major metropolitan centers of the SALCC region. Relatively high species diversity is also apparent in the USGS NAS Database for several of the subbasins that border the Atlantic coast in the SALCC region. The number of species documented was greatest in subbasins adjacent to Jacksonville, Florida; Charleston, South Carolina; Myrtle Beach, South Carolina; and Virginia Beach, Virginia. In general, relatively few nonindigenous aquatic species were reported in subbasins characterized by low urban land cover and separated by at least 50 kilometers from major urban centers. For example, subbasins in the large, rural portions of southeastern Georgia were characterized by low nonindigenous aquatic species diversity in the USGS.
Figure 6. Percent urban land cover (U.S. Geological Survey, National Gap Analysis Program, 2011) of U.S. Geological Survey subbasins within the South Atlantic Landscape Conservation Cooperative region.
Figure 7. Total number of nonindigenous aquatic species per subbasin recorded in the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database for the South Atlantic Landscape Conservation Cooperative region.
NAS Database. These subbasins exhibit a comparatively low density of major water bodies, and this characteristic may result in a lack of available habitat for nonindigenous aquatic species relative to other subbasins. In addition to the low diversity of nonindigenous aquatic species reported in southeastern Georgia, the total numbers of observations in these subbasins were also relatively low. This low level of observations may reflect the low human population density and, when combined with a limited number of large water bodies, fewer opportunities relative to more populated areas to sight nonindigenous aquatic species. It is noted that, because of the limitations of presence-only data described above, the low diversity and low numbers of observations should not be regarded as conclusive evidence that nonindigenous aquatic species are uncommon in these areas.

A Kruskal-Wallis H nonparametric statistical test (equivalent to the parametric one-way analysis of variance test) was performed to further examine differences in the numbers of records, the number of species, and the numbers of records per species for subbasins classified as urban, nonurban, coastal, and upland. We performed the same tests on subbasins classified according to the presence or absence of reservoirs located anywhere within the subbasin or on its boundary. The results of these comparisons are summarized in table 2. These analyses show significant differences in the number of species reported in urban subbasins relative to those classified as rural and in subbasins with reservoir contact relative to those without reservoir contact. The total numbers of records were also found to differ significantly in urban versus rural subbasins and in subbasins with reservoir contact compared to those without reservoir contact. The results of these tests coincide with nonindigenous aquatic species data from other regions, though this is the first time that quantitative comparisons have been applied across multiple locations. Interestingly, the number of records per species in the USGS NAS Database did not differ significantly among any of the HUC classifications. This result suggests that the opportunities to observe nonindigenous aquatic species relative to the number of species present may not differ significantly across the SALCC. If this conclusion is valid, it would imply that the low numbers of species reported in the USGS NAS Database for some areas of the SALCC are a more accurate representation of actual nonindigenous populations than what might normally be expected on the basis of presence-only information. Normalizing the numbers of species in a subbasin by the subbasin area leads to a similar lack of significant differences in coastal versus upland or urban versus rural values. The implications of these findings warrant further exploration.

In the SALCC region, 20 or more nonindigenous aquatic species have been reported in each of 16 upland, freshwater subbasins, all of which contain, or intersect, large reservoirs. Also included in these subbasins are several large urbanized areas such as Augusta and Atlanta, Georgia; Charlotte, Greensboro, and Raleigh, North Carolina; and Greenville and Spartanburg, South Carolina. Fish compose a large majority of the nearly 140 unique species overall across these 16 subbasins (fig. 8). Some of the more notable species

<table>
<thead>
<tr>
<th>Measurement variable for each subbasin</th>
<th>Nominal variables for each subbasin</th>
<th>Coastal versus upland</th>
<th>Urban versus rural</th>
<th>Reservoirs present versus not present</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of species reported</td>
<td>H=0.011</td>
<td>H=14.39</td>
<td>H=33.81</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p=0.92</td>
<td>p=1.5E-4*</td>
<td>p=6.1E-9*</td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>H=0.02</td>
<td>H=8.87</td>
<td>H=19.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p=0.88</td>
<td>p=0.002*</td>
<td>p=1.2E-5*</td>
<td></td>
</tr>
<tr>
<td>Number of observations per species reported</td>
<td>H=1.21E-4</td>
<td>H=2.08</td>
<td>H=2.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p=0.99</td>
<td>p=0.15</td>
<td>p=0.11</td>
<td></td>
</tr>
</tbody>
</table>
Characterization of USGS NAS Database Data

Figure 8. Total number of distinct species per taxa across the 16 upland subbasins with 20 or more nonindigenous aquatic species recorded in the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database for the South Atlantic Landscape Conservation Cooperative region.

*Total number of distinct species in the 16 upland subbasins with 20 or more nonindigenous aquatic species in the South Atlantic Landscape Conservation Cooperative region.

documented in these subbasins with potentially significant ecological impacts are *Orconectes rusticus* (rusty crayfish), *Orconectes virilis* (virile crayfish), *Procambarus clarkii* (red swamp crayfish), *Channa argus* (northern snakehead), grass carp, common carp, *Monopterus albus* (Asian swamp eel), *Corbicula fluminea* (Asian clam), *Cipangopaludina chinensis* (Chinese mysterysnail), and *Trachemys scripta elegans* (red-eared slider). Nonindigenous aquatic plant species are also represented in the upland subbasins including *Eichhornia crassipes* (water hyacinth), *Hydrilla verticillata* (hydrilla), *Myriophyllum spicatum* (Eurasian watermilfoil), *Potamogeton crispus* (curlyleaf pondweed), and *Salvinia molesta* (giant salvinia). An examination of the periods during which species were recorded in these subbasins (fig. 9) shows that, with a few exceptions, most of the observations occurred during the 1990s.
Relatively high numbers of nonindigenous aquatic species have been reported in coastal subbasins in the SALCC region. The total number of nonindigenous aquatic species across these subbasins is 152, including 89 freshwater, 12 freshwater-marine (species that can use and/or tolerate the three habitat types), and 51 marine species (fig. 10). Ten coastal subbasins with 20 or more nonindigenous aquatic species have been documented. Notably, only 2 of these 10 subbasins are found to have a relatively high percentage of land classified as urban. The Cooper subbasin (13 percent urban land use) contains Charleston, South Carolina, and the Lower St. Johns subbasin (31 percent urban land use) contains Jacksonville, Florida. The Cooper subbasin has one major reservoir, Lake Moultrie, South Carolina, where as many as 10 nonindigenous aquatic plant species have been documented. Most of the plant species represent known ecological threats including water hyacinth (Sbardt, 1997), hydrilla, and Hygrophila polysperma (Indian hygrophila), the latter two of which are classified as noxious weeds in the United States (U.S. Department of Agriculture, 2015). Although the Lower St. Johns subbasin, which has no large reservoirs, was classified as coastal, most of the nonindigenous species...
aesthetic species, both plants and animals, in this subbasin are from freshwater habitats. An examination of the periods during which the species were recorded (fig. 11) shows that, for most of these coastal subbasins, relatively high numbers of new species were observed in the 1990s and 2000s. In the 1970s a large number of new observations were recorded in the Cooper subbasin, whereas in the Lower Ochlockonee and Lower St. Johns subbasins high numbers of new species were recorded in the 1980s.

Figure 12 illustrates in matrix form the individual subbasins in which each nonindigenous taxon has been observed in the SALCC region, along with the taxa observed within each subbasin. This visualization of the USGS NAS Database reveals several patterns. For example, 279 taxa have been recorded in the database at a spatial resolution sufficient to identify their presence in at least one of the subbasins. Of 279 taxa, 124 were recorded in only one subbasin, and 40 were recorded in only two subbasins. Asian clam was recorded in the most subbasins (n=88). Figure 12 also shows that those taxa that are rarely recorded are most likely to be recorded in subbasins where many other taxa have also been recorded. In addition, the concentration of points plotting on the left side of figure 12 suggests that taxa are distributed across subbasins more uniformly than subbasins are represented by taxa. As such, the right one-third of the plot represents subbasins with only one recorded taxon (fig. 12).
**EXPLANATION**

*Number of new species in the 10 coastal subbasins with 20 or more nonindigenous aquatic species in the South Atlantic Landscape Conservation Cooperative region

**Spalding and others, 2007**

**Figure 11.** Number of new species introductions reported by decade in the 10 coastal subbasins with 20 or more nonindigenous aquatic species recorded in the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database for the South Atlantic Landscape Conservation Cooperative region.

**Impacted River Systems**

Grouping nonindigenous aquatic species records according to the 4-digit HUC (HUC4) (hereinafter referred to as “subregion”) level provides a convenient and computationally tractable way of relating the spatial characteristics of these data to geographic and cultural features of the landscape. This approach should be complemented with additional analyses that recognize the importance of hydrologic connectivity as a critical factor in regulating the distribution of aquatic species across a region. With this in mind we examined the numbers of species reported in the major rivers and their tributaries in the SALCC region (fig. 13). This particular group of subregions is unique in that they each drain directly to either the Atlantic Ocean or the Gulf of Mexico and are therefore isolated hydrologically from one another. The two highest numbers of species were reported in the Edisto-Santee (n=118) and the Chowan-Roanoke (n=86) subregions. The results presented in figure 13 indicate that, while the species in these subregion river basins are reported more often in the urbanized upper reaches and along the coast, the degree of hydrologic connectivity within each basin should be taken into consideration when estimating the potential dispersion of nonindigenous populations.

The geomorphological characteristics of the southeast Atlantic Coastal Plain suggest that dispersion by nonindigenous aquatic species across the large river basins in the SALCC region will be limited to anthropogenic vectors and coastal “salt bridges.” Nonindigenous aquatic species with wide salinity tolerances are capable of moving alongshore in tidal waters and migrating upstream into previously unoccupied freshwater systems (Capps and others, 2011). New records of nonindigenous aquatic species with known ecological impacts and wide salinity tolerances should be monitored closely by the SALCC conservation community.
Species recorded in the U.S. Geological Survey (USGS) Nonindigenous Aquatic Species (NAS) Database observed in subbasins intersecting the South Atlantic Landscape Conservation Cooperative (SALCC) region.

**Figure 12.** U.S. Geological Survey (USGS) subbasins with species recorded in the USGS Nonindigenous Aquatic Species (NAS) Database for the South Atlantic Landscape Conservation Cooperative region. Each point represents one or more occurrences of a single species in a specific subbasin.
Figure 13. Total number of nonindigenous aquatic species recorded in the U.S. Geological Survey (USGS) Nonindigenous Aquatic Species (NAS) Database in USGS subregions in the South Atlantic Landscape Conservation Cooperative region.
Potential Future Invasions

Fish species with suspected negative ecological impacts have been recorded in areas close to the border of the SALCC region. Data from the USGS NAS Database for these species suggest that they are dispersing toward the SALCC region. These species include northern snakehead in Virginia, which is expanding its range in a southward direction; Hoplosternum littorale (brown hoplo) in Florida, which is expanding its range northward; and Nile tilapia in Alabama, Mississippi, and the Florida Panhandle, which is expanding its range eastward (fig. 14). Detailed below are these potential invaders.

Northern snakehead is native to China, Russia, and Korea and inhabits stagnant shallow ponds, swamps, and slow streams with mud or vegetated substrate, with water temperatures ranging from 0 degrees Celsius (°C) to about 30 °C (Courtenay and Williams, 2004). This broad temperature range leaves much of the United States vulnerable for the possible establishment of this fish (Herborg and others, 2007). Numerous established populations have been found primarily in the mid-Atlantic region of the United States, the first of which was discovered in a pond in Crofton, Maryland, in June 2002. This population, a deliberate introduction, was quickly eradicated by State biologists using a fish piscicide to poison the population (Orrell and Weigt, 2005). In August of that same year, two northern snakehead specimens were reportedly caught from Lake Wylie near Charlotte, North Carolina, by anglers. Five years later, in 2007, a large adult was caught by an angler in South Fork Catawba River, also in the Charlotte area of North Carolina. Since then, no additional fish have been reported in the State. Although there is no known established population of northern snakehead in North Carolina, in 2004 a population was discovered in the Potomac River of neighboring Virginia and several of its tributaries below the District of Columbia. The population has been predominantly moving downriver in the Potomac River where individuals have been reported as far south as the Rappahannock River in Virginia, more than 80 miles from the suspected point of introduction in the Potomac drainage. Northern snakehead specimens have also been found to the east of the Potomac drainage in Chesapeake Bay and to the north in Delaware River tributaries. It is suspected that northern snakehead can traverse moderately high salinities to disperse to these locations. Specimens have been collected in salinities as high as 7.6 parts per thousand (Starnes and others, 2011). Because of the proximity of northern snakehead populations to the SALCC region, this species may pose a threat, as it is a large piscivorous predator reaching 85 centimeters (cm) in length. Juvenile snakehead eat zooplankton, insect larvae, small crustaceans, and the fry of other fish. Adult snakehead feed almost exclusively on other fishes, with the remainder of their diet composed of crustaceans, frogs, small reptiles, and sometimes small birds and mammals. This diet of the adults shows a significant diet overlap with largemouth bass, with both species consuming a large proportion of Fundulus spp. (topminnow and killfish) and other centrarchids in the lower Potomac River (Saylor and others, 2012). Northern snakehead could eventually expand its range on its own, moving south along the Atlantic coast by using the salt bridges. However, the threat of illegal stocking also is present, of which the early captures in North Carolina likely represent, as northern snakehead does present a novel species for avid anglers.

Brown hoplo is indigenous to tropical South America. It is widespread east of the Andes Mountains and north of Buenos Aires and as far north as Trinidad in the Caribbean region (Reis, 1997). This armored catfish can reach 22 cm in length and is typically found in a variety of freshwater habitats, including slow-moving rivers, canals, streams, and pools (Nico and others, 1996). The first population in the United States was discovered in 1995 in the Indian River Lagoon system, Florida (Nico and others, 1996). Since then, the brown hoplo has been found in drainages both north and south of its first known collection site. At the present time, any negative impacts are unknown. However, this species is generally locally abundant where it occurs and feeds primarily on benthic invertebrates and detritus, possibly competing with native fishes if food is limited (Duxbury and others, 2010).

The brown hoplo is in the southernmost portion of the SALCC region, near Jacksonville, Florida, where it was collected in 2005. Subsequently, collections in the upper portion of the Ocklawaha River drainage in northern Florida beginning in 2006 (Shafland and others, 2008) may pose a threat to the adjacent Suwannee River Basin, designated as one of the “Outstanding Florida Waters” worthy of special protection.

The native range of Nile tilapia is tropical and subtropical Africa and Middle East. They inhabit lakes and rivers and can exceed 50 cm in length (de Moor and Bruton, 1988). This species is established in nearby Mississippi in freshwater locations along the Gulf of Mexico coast and in several locations of the upper Ocklawaha River drainage, and possibly established in Lake Seminole (Apalachicola-Chattahoochee Rivers drainage) on the Florida-Georgia border—all in proximity to the SALCC region. A recent and only collection to date in the SALCC region occurred in 2011 in a pond north of Wilmington, North Carolina, within the Northeast Cape Fear subbasin. Other collections just outside the SALCC region include those in Yates Reservoir and a small tributary of the Chattahoochee River, both in eastern Alabama. Although the diet of the Nile tilapia is primarily herbivorous, a small portion of their diet can consist of insects, crustaceans, and fish eggs (Khallaf and Al ne-na-ei, 1987). They are also known to prey on amphibians and juveniles of other fish species (Zambrano and others, 2006). In the southwestern United States, a population of Nile tilapia has led to the decline of two small native fish species, M o a p a co r i à c e a (Moapa dace) and Crenichthys b a il e y i mo a pae (Moapa White River springfish) (Canonico and others, 2005). For a description of additional species with possible ecological impacts that have been introduced to the SALCC region in the past decade, see appendix 1.
Figure 14. Observations recorded in the U.S. Geological Survey Nonindigenous Aquatic Species (NAS) Database for three fish species with established populations in proximity to the South Atlantic Landscape Conservation Cooperative (SALCC) region. Single specimens of these species have been collected within the SALCC region but are not established there.
Statistical Analyses and Point-Process Modeling

The motivation to model invasive species introductions is to gain an understanding of the process of how they spread in the environment with the ultimate goal of minimizing the threat they pose to novel ecosystems. As is the case for nearly all presence-only databases, the opportunistic observation records in the USGS NAS Database are contributed by (or are derived from) multiple types of observers across a wide range of locations and habitats. Corresponding absence data (that is, records of no occurrence) that are commonly reported as part of planned surveys are not typically contained in these databases. There are several other fundamental characteristics of presence-only databases that create challenges to their use in the development of models capable of simulating a dynamic biological process (Elith and Leathwick, 2007) such as the invasion and dispersal of a nonindigenous aquatic species.

Presence-only data such as those in the USGS NAS Database that show a relatively unambiguous and largely predictable spatial evolution of observations over time are rare. Only a small percentage of the thousands of records of nonindigenous aquatic species in the entire United States display a pattern that is consistent with a dispersal pattern that would be expected on the basis of known biological processes and a species’ environmental requirements. Some examples of species whose records do display these patterns but are outside the SALCC region include Dreissena polymorpha (zebra mussel), bighead carp, silver carp, and Potamopyrgus antipodarum (New Zealand mudsnail). Simulation models of future distributions of these species calibrated by using presence-only data and habitat information can be generated by using MaxEnt or similar approaches (Poulos and others, 2012; Azzurro and others, 2013; Gallardo and others, 2013; McDowell and others, 2014). The potential high uncertainty in predicted distributions generated by these types of approaches motivated the effort to develop a new statistical approach.

Predicting the timing and location of nonindigenous aquatic species introductions and dispersal patterns in the SALCC region presents a challenging statistical problem. The most robust approaches to modeling population growth and dispersal require information on environmental covariates at the locations where a species is present and where it is absent. Depending on the modeling approach, other types of information may also be needed. In some cases, information on environmental covariates for presence-only records can be determined from other databases and monitoring networks. However, the location accuracy of the observation records often varies, thereby making it difficult to match appropriate datasets. Further, estimates of detection probability, also a requirement for many population models, cannot be readily derived from the opportunistic presence-only records such as those contained in the USGS NAS Database. These challenges to modeling nonindigenous aquatic species dispersal patterns in the environment provided additional motivation to develop a new approach.

Methods for the Point-Process Model

A significant effort was directed towards the development of a point-process species distribution model conditioned on nonindigenous aquatic species data from within the SALCC region. This effort began with a search for candidate species (or groups of species) in the USGS NAS Database whose records were characterized by (1) regular observations over time that displayed coherence with the spatial distribution of the species; (2) information on the locations of the observations, including habitat type; (3) a minimum number of qualified observers, some of whom reported observations as part of a planned survey; and (4) corresponding absence data.
Data meeting these criteria were considered to be optimal and would have enabled the exploration of novel point-process approaches to modeling species distributions. However, observation records that met these requirements were not found within the SALCC subset of the USGS NAS Database. Observation records of other nonindigenous aquatic species that may have been suitable were found in the database, but these were restricted to ranges outside of the SALCC boundaries and not pursued. Instead effort was devoted towards locating records of planned and, importantly, replicated surveys of nonindigenous aquatic species not yet represented in the USGS NAS Database that, when combined with existing records, would satisfy our criteria for a suitable dataset. A list of the grey and peer-reviewed literature reviewed is provided in appendix 2. We accessed three studies conducted in or just outside the SALCC boundary (Bart and others, 1994; Johnston and others, 1995; Dobbins and others, 2012). In these three studies, 12 species of fish were noted as nonindigenous to the waters from which they were collected. Remarkably, this search also did not yield any suitable data. We then turned to the Multistate Aquatic Resources Information System (MARIS) (http://www.marisdata.org/) for fish sampling data, this time in search of a dataset for either nonindigenous or native fish species that would meet our criteria. We focused on MARIS data from the State of Georgia. MARIS surveys cover exceptionally large spatial domains and for this reason alone are uniquely valuable for the scientific community. Our assessment of these surveys, however, also indicates that the relatively long intervals between sequential surveys in a given region do not allow an assumption of constant population size. This assumption is critical to the development of robust species distribution models. Our assessments also indicate that MARIS datasets do not include information on double-observer or similar sampling techniques that might also allow for the quantification of detection probability. Finally, the focus on sport and recreational species appears to have led to a lack of presence-absence records in the MARIS database for species not of interest to this program (including nonindigenous aquatic species).

We consider that the relative paucity of presence-absence data collected through repeated surveys that also quantify detection probability and habitat characteristics represents a significant barrier to the development of statistical models of populations that change over space and time, including populations of nonindigenous aquatic species. Because of the emergence of new and widespread environmental stressors such as a changing climate, sea level rise, and land use change, many native populations are likely to exhibit similarly dynamic characteristics as their ranges and abundance change in response to these stressors. Assessing how these populations are changing through a combination of monitoring and model development presents a unique challenge to the conservation community. More information than is routinely collected under many wildlife programs (for either native or nonindigenous aquatic species) may be needed to generate robust predictions of current and future species distributions. We consider this conclusion to be highly relevant to the effort to track biological indicators of the SALCC conservation blueprint.

The effort to locate an optimal set of nonindigenous aquatic species data delayed but did not prevent the development of a novel point-process species distribution model for this study. The model that has been developed relies on simulated point counts data generated from multiple observers. This foundational paper, funded by this study, is summarized by Dorazio (2014).

**Implications for Conservation Practices**

This point-process analysis shows that population distribution models for any species of wildlife, including nonindigenous aquatic species, that are based on presence-only records can be greatly enhanced by a relatively small amount of data collected as part of dedicated surveys. This finding is considered to be of potentially significant importance in ongoing efforts within the SALCC organization to develop biological indicators from existing monitoring programs and other historical datasets.

In this study, investigations have led to other theoretical and practical considerations for the development of population models of nonindigenous aquatic species. For example, while existing models, such as MaxEnt (Phillips and others, 2004), can be used to predict the spatial distribution of a nonindigenous aquatic species, the predictions are potentially biased because they do not account for the spatial variation in detectability of the species (Dorazio, 2012; Renner and Warton, 2013).
Our examinations of presence-only databases reveal a general lack of coherence between the timing of observations and changes in the spatial distribution of an invasive organism. Presence-only records also generally do not contain community-level information that would, in turn, enable inferences about the interactions of an invasive species with other organisms. As a result, observation data alone for many species may not be sufficient for representing or modeling a dynamic process such as the establishment, dispersal, and persistence of an introduced organism. As outlined above, statistical models representing these dynamics have not been developed in large part because of the overall lack of appropriate data. Developing these models in the future will require information on how habitats change over time, as well as how an organism’s behavior and life history characteristics respond to these changes. This consideration is relevant for all species whose native ranges may be changing as a result of new environmental stressors.

**Threats to Protected Areas**

Nonindigenous aquatic species constitute a significant threat to the integrity of protected areas. A primary goal of protected areas is to provide a natural refuge for native species. However, aquatic invaders may eliminate or greatly reduce populations of certain native species from protected areas through competition, predation, habitat modification, or other changes. Preventing the establishment of invasive species is widely recognized as more desirable and effective management strategy than eradication efforts (Leung and others, 2005). For many invasive species, eradication may only be effective if efforts are made early, before a species becomes well established in a protected area (Mehta and others, 2007). Therefore, identifying invasive species that occur near protected areas can provide managers with useful information for reducing the risk of invasion and inform managers on the need to monitor for specific invasive species.

In this section we describe a simple geographic information system approach that uses the USGS NAS Database to identify the risk of nonindigenous aquatic species invading or already occurring within individual protected areas of the SALCC region. We identify which nonindigenous aquatic species are already reported to occur within existing protected areas. Perhaps more importantly, we identify which nonindigenous aquatic species threaten to invade or may already exist within individual protected areas on the basis of the protected area having hydrologic connections to nonindigenous aquatic species observations occurring outside the protected areas.

Our first objective was to identify which nonindigenous aquatic species are known to occur within protected areas and provide basic statistics such as frequency of species occurrence. Our second objective was to identify which nonindigenous aquatic species are hydrologically connected to protected areas and therefore represent a threat of invasion or may already occur within specific protected areas.

**Methods for Protected Areas Analysis**

Nonindigenous aquatic species distribution data were obtained for the SALCC region from the USGS NAS Database (https://nas.er.usgs.gov/, accessed March 9, 2015). Slight modifications to the SALCC boundary were made to extend into adjacent coastal estuaries. Location data of marine species were not used, as this analysis only considered the potential threat of freshwater species in waters within protected areas on land. The nonindigenous aquatic species distribution data were then overlaid on the National Hydrography Dataset Plus (NHDPlus, version 2; http://www.horizon-systems.com/nhdplus/) (U.S. Environmental Protection Agency, 2012) flowline streams by using ArcGIS 10.2.2 (Esri, Inc., Redlands, California) tools; only nonindigenous aquatic species data with high spatial accuracy were used. Nonindigenous aquatic species data points that did not occur within 700 meters of any NHDPlus stream network were assumed to occur in isolated aquatic features and were not used in the analysis.

Next, the Protected Areas Database of the United States (PAD-US, Conservation Biology Institute Edition; http://consbio.org/products/projects/pad-us-cbi-edition) covering the SALCC region (fig. 15) was intersected with the NHDPlus...
Figure 15. Protected areas within the South Atlantic Landscape Conservation Cooperative region from the Protected Areas Database of the United States (PAD-US, Conservation Biology Institute Edition).
stream network; protected areas that did not intersect the stream network were deleted. Some protected areas occurred in more than one hydrologic unit (subregions of NHDPlus); these were split and analyzed as occurring in separate basins. Nonindigenous aquatic species data were also intersected with the protected area subbasins to determine whether each nonindigenous aquatic species data point occurred inside or outside of any protected area.

**Results and Discussion of Protected Areas Analysis**

A total of 345 protected areas (fig. 15) intersected the NHDPlus stream network in the SALCC region. The number of nonindigenous aquatic species occurring within these 345 protected areas varied considerably (fig. 16), with a maximum of 27 species and a mean/mode of 3.14/2 species per protected area. The species reported with the highest frequencies in 10 or more protected areas included Asian clam (125), 12 fish species introduced for sport fishing, 7 aquatic or emergent plant species, and 2 crustaceans (fig. 17). The number of potential invaders known to be hydrologically connected, but not yet reported from individual protected areas, varied considerably (fig. 18), with a maximum of 57 species and a mean/mode of 27/25 species per protected area (fig. 19).

The mean number of nonindigenous aquatic species reported for each protected area is surprisingly low (mean 3.14 species), but this statistic does not reflect the large number of potential invaders that are reported to occur within the same watershed for each protected area. When hydrologic connection is considered, the number of potential invaders that could spread to each protected area is greatly increased, with a mean of 27 species and the maximum rising to 57 species. Nearly all protected areas are hydrologically connected to at least 20 species.

Some protected areas within the SALCC region do not have direct hydrologic connection to stream networks identified within the NHD database. These protected areas may still be at risk of invasion from species that occur anywhere within the SALCC region by mechanisms that do not depend on hydrologic connection, such as deliberate introductions by humans through release of bait buckets or accidental movement by boat trailers or bilge water.
Figure 16. The number of nonindigenous aquatic species reported from individual protected areas, from the Protected Areas Database of the United States (PAD-US, Conservation Biology Institute Edition) sorted by named areas with highest number of species, for the South Atlantic Landscape Conservation Cooperative region.
<table>
<thead>
<tr>
<th>Number of species</th>
<th>Protected area</th>
<th>Number of species</th>
<th>Protected area</th>
<th>Number of species</th>
<th>Protected area</th>
<th>Number of species</th>
<th>Protected area</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Fort Sumter National Monument</td>
<td>2</td>
<td>Fort York SP</td>
<td>2</td>
<td>Griffins Run WMA</td>
<td>2</td>
<td>Cape Royal National Wildlife</td>
</tr>
<tr>
<td>4</td>
<td>Four Creek State Forest</td>
<td>2</td>
<td>Griffin Rock WMA</td>
<td>2</td>
<td>Guilford Courthouse</td>
<td>2</td>
<td>Capers Island Heritage Preserve</td>
</tr>
<tr>
<td>4</td>
<td>Great Dismal Swamp National</td>
<td>2</td>
<td>Honeycreek Creek</td>
<td>2</td>
<td>Indian Field</td>
<td>2</td>
<td>Carolina Sandhills NWR</td>
</tr>
<tr>
<td>4</td>
<td>Refuge</td>
<td>1</td>
<td>Ichetucknee Conservation Area</td>
<td>2</td>
<td>Kingfisher State</td>
<td>2</td>
<td>Carrs Island Preserve</td>
</tr>
<tr>
<td>1</td>
<td>Hyco State Game Land</td>
<td>1</td>
<td>Jordan Lake State Recreation</td>
<td>2</td>
<td>Kings Mountain SP</td>
<td>2</td>
<td>Castaway Island Preserve</td>
</tr>
<tr>
<td>4</td>
<td>Mecklenburg County Park</td>
<td>1</td>
<td>Lake Mountain State</td>
<td>2</td>
<td>Lake Stratford</td>
<td>2</td>
<td>Catawba Island Preserve</td>
</tr>
<tr>
<td>4</td>
<td>MCBB Camp Lejeune</td>
<td>1</td>
<td>Lower Haw River State</td>
<td>2</td>
<td>Manatee Springs State</td>
<td>2</td>
<td>Charlie Elliott Wildlife Center</td>
</tr>
<tr>
<td>4</td>
<td>Poosin Lakes NWR</td>
<td>1</td>
<td>Mitchell Mill SNA</td>
<td>2</td>
<td>Manatee State</td>
<td>2</td>
<td>CHIL-VOF-19</td>
</tr>
<tr>
<td>4</td>
<td>Raven Rock State Park</td>
<td>2</td>
<td>MCAS Cherry Point</td>
<td>2</td>
<td>Marion Swamp</td>
<td>2</td>
<td>Chowan Swamp SNA</td>
</tr>
<tr>
<td>4</td>
<td>Saint SP</td>
<td>2</td>
<td>MCAS New River</td>
<td>2</td>
<td>Claudenson Plantation</td>
<td>2</td>
<td>Colleton SP</td>
</tr>
<tr>
<td>4</td>
<td>Tanglewood State</td>
<td>2</td>
<td>Medoc Mountain Park</td>
<td>2</td>
<td>Coharie Creek</td>
<td>2</td>
<td>Colonial Dorchester SHS</td>
</tr>
<tr>
<td>4</td>
<td>Alligator River National</td>
<td>2</td>
<td>Mitchell River State</td>
<td>2</td>
<td>Moody Forest SNA</td>
<td>2</td>
<td>Congaree National Heritage</td>
</tr>
<tr>
<td>4</td>
<td>Wildlife Refuge</td>
<td>2</td>
<td>Neuse River State Game Land</td>
<td>2</td>
<td>Ninety Six National Historic Site</td>
<td>2</td>
<td>Congaree National Heritage</td>
</tr>
<tr>
<td>4</td>
<td>Cheaha-Chehaha</td>
<td>2</td>
<td>Old Santee</td>
<td>2</td>
<td>Ocmulgee West</td>
<td>2</td>
<td>Coastal National Wildlife</td>
</tr>
<tr>
<td>4</td>
<td>Plantation</td>
<td>2</td>
<td>Old Santee</td>
<td>2</td>
<td>Ocmulgee West</td>
<td>2</td>
<td>Cape Romain National Wildlife</td>
</tr>
<tr>
<td>4</td>
<td>George Plesants</td>
<td>2</td>
<td>Old Santee</td>
<td>2</td>
<td>Ocmulgee West</td>
<td>2</td>
<td>Cape Romain National Wildlife</td>
</tr>
<tr>
<td>4</td>
<td>Greenville Transmission Station</td>
<td>2</td>
<td>Old Santee</td>
<td>2</td>
<td>Ocmulgee West</td>
<td>2</td>
<td>Cape Romain National Wildlife</td>
</tr>
<tr>
<td>4</td>
<td>of Voice of America</td>
<td>2</td>
<td>Old Santee</td>
<td>2</td>
<td>Ocmulgee West</td>
<td>2</td>
<td>Cape Romain National Wildlife</td>
</tr>
</tbody>
</table>

**Figure 16.** The number of nonindigenous aquatic species reported from individual protected areas, from the Protected Areas Database of the United States (PAD-US, Conservation Biology Institute Edition) sorted by named areas with highest number of species, for the South Atlantic Landscape Conservation Cooperative region.—Continued
Figure 17. The number of protected areas with reports of individual nonindigenous aquatic species, sorted by most commonly reported species, for the South Atlantic Landscape Conservation Cooperative region.
Factors Associated With Nonindigenous Aquatic Species Richness

Spatial statistical models can provide a better understanding of factors associated with the occurrence of high numbers of nonindigenous aquatic species and may provide managers with information useful for preventing or controlling the introduction and spread of invasive species. These models can identify human or environmental factors that are associated with nonindigenous aquatic species richness, allowing managers to focus actions on those factors that are amenable to management actions. Additionally, the spatial models can be used in a predictive framework, providing spatially explicit forecasts of nonindigenous aquatic species richness increases based on forecasts of key independent variables. Our objective was to develop a spatially explicit regression model that identifies natural and human correlates of nonindigenous aquatic species richness in the SALCC region while examining spatial patterns of nonindigenous aquatic species richness and variables that might explain these patterns within a regression modeling framework. This aspect of the study represents an extension of and improvement over the initial exploratory analyses described in section “Characterization of USGS NAS Database Data.”

Methods for Species Richness Analysis

The USGS NAS Database has a number of characteristics that pose challenges for statistical analysis. The database consists of presence-only data collected opportunistically from numerous observers over multiple decades. Consequently, the data are subject to a variety of biases, such as nonrandom or nonsystematic sampling, lack of independence of observations, and the inability to estimate detection probability. To reduce some of these biases, we aggregated the nonindigenous aquatic species point data spatially and temporally to a single index of nonindigenous aquatic species richness for each subbasin.

The SALCC region was divided into watersheds by using the NHD to delineate the subbasins. Because we were interested in only freshwater aquatic species, we omitted several small coastal subbasins that were strictly marine environments, resulting in a total of 109 subbasins. Nonindigenous aquatic species distribution data were obtained from the USGS NAS Database (https://nas.er.usgs.gov/, accessed June 11, 2015). Freshwater and diadromous aquatic species were included in the analysis, whereas marine species were omitted, resulting in a total of 10,786 individual records in the study area. All occurrence records irrespective of whether the species was known to be established were included. Nonindigenous aquatic species point data were aggregated at the subbasin level to obtain the recorded number of species occurring in each subbasin in order to reduce some of the biases inherent in these types of data (further discussion below).

An independent set of environmental and human variables (table 3) thought to be potentially important for explaining patterns in the nonindigenous aquatic species data was assembled from various sources. The variables included housing density, road density, reservoir density, river density, elevation (range, mean, and maximum), minimum temperatures (mean and minimum cold temperature for January 2010), and subbasin centroid (latitude and longitude) and area. These variables were summarized and aggregated for each subbasin. The variables were examined by using quantile-quantile and scatter plots, and nonnormal data were transformed by using a Box-Cox approach. The dependent variable (nonindigenous aquatic species richness) was square-root transformed. Several parameters were converted to densities based on subbasin area (density of roads, rivers, and reservoirs), and several were log transformed.

The data were examined for normality by using tests to help find a properly specified ordinary least squares (OLS) model that lacked multicollinearity and had residuals that showed normality and lack of spatial autocorrelation. OLS regression modeling tools in ArcGIS 10.2.2 were used to develop models potentially explaining nonindigenous aquatic species richness on the basis of the set of explanatory variables. The diagnostic tests in the OLS tool (Esri, 2011) were then used to screen out improperly specified models based on a suite of tests, including variable significance (p<0.05), variance inflation factor (VIF) test for collinearity (p<7.5), Moran’s I test for spatial autocorrelation of residuals (p>0.10), and Jarque-Bera test for nonnormal residuals (p>0.10). All subsets of models were evaluated with 1 to 5 dependent variables. Those that did not pass the suite of tests listed above were rejected (higher dimensioned models were excluded for computational reasons). Akaike information criterion comparisons (AICc), a measure of the relative goodness of fit and model complexity (Burnham and Anderson, 2002), were performed to compare models and identify the best passing OLS model.

As a final step, we tested the best OLS model for suitability for analysis by geographically weighted regression (GWR) analysis using the Breusch-Pagan & Koenker statistic. This statistic tested for nonstationarity or variation in parameterization of variables among subbasins versus variables remaining unchanged globally (Fotheringham and others, 2002). This result suggested that GWR was not warranted for these data.
Figure 18. The potential number of nonindigenous freshwater species reported outside of specific protected areas from the Protected Areas Database of the United States (PAD-US, Conservation Biology Institute Edition) but hydrologically connected by waterways at the subregion level in the South Atlantic Landscape Conservation Cooperative region.


*ACE is derived from Ashepoo, Combahee, and Edisto Rivers.

**Because of space limitations, not all protected areas are labeled on graph. See facing page.
<table>
<thead>
<tr>
<th>Number of species</th>
<th>Protected area</th>
<th>Number of species</th>
<th>Protected area</th>
<th>Number of species</th>
<th>Protected area</th>
<th>Number of species</th>
<th>Protected area</th>
</tr>
</thead>
<tbody>
<tr>
<td>51</td>
<td>Broad River State Natural Area</td>
<td>51</td>
<td>Crab Orchard Plantation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>Croft State Natural Area</td>
<td>51</td>
<td>Dilane Wildlife Management Area</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>Fort Goodwin</td>
<td>51</td>
<td>Freeman Tract</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>Springfield Plantation - Georgetown</td>
<td>51</td>
<td>Grasslands Reserve Program</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>Harris Neck National Wildlife Refuge</td>
<td>51</td>
<td>James W. Webb Wildlife Center</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>Kings Mountain State Park</td>
<td>51</td>
<td>King's Mountain Wildlife Management Area</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>Lake Warren State Park</td>
<td>51</td>
<td>Landscove Canal State Park</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>Mason Wildlife Preserve</td>
<td>51</td>
<td>McDuffie Public Fishing Area and Fish</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>Ninety Six</td>
<td>51</td>
<td>Pihag National Forest</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>Santee Coastal Reserve</td>
<td>51</td>
<td>Skidaway Island State Park</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>St. Catherine's Island</td>
<td>51</td>
<td>Torrence Creek Greenway</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>Cheechee-Combahkee Plantation</td>
<td>51</td>
<td>Fort Stewart</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>James Ross Wildlife Reservation</td>
<td>51</td>
<td>Lake Warren State Park</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>Mountain Island Educational State Forest</td>
<td>51</td>
<td>Pisgah Game Land</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>Pisgah State Game Land</td>
<td>51</td>
<td>Santee State Reserve</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>South Carolina Wildlife Refuge</td>
<td>51</td>
<td>Sork Island Wildlife Management Area</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>South Carolina State Game Land</td>
<td>51</td>
<td>Tuckahoe Wildlife Management Areas</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>ACE Basin National Wildlife Refuge</td>
<td>49</td>
<td>Ashepoo-Combahee-Edisto Basin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Basin National Estuarine Research Reserve</td>
<td>49</td>
<td>Magnolia Springs State Park</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Palacheo Wildlife Management Area</td>
<td>49</td>
<td>Palmetto State Park</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Santee State Park</td>
<td>49</td>
<td>Santee State Park</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Turkey River State Park</td>
<td>49</td>
<td>Cedar Creek State Park</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Catawba State Game Land</td>
<td>49</td>
<td>Colonial Dorchester Historic Site</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Colonial Dorchester Historic Site</td>
<td>49</td>
<td>Congaree Creek Heritage Preserve</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Congaree Creek Heritage Preserve</td>
<td>49</td>
<td>Crowder State Park</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Crowders Mountain State Park</td>
<td>49</td>
<td>Dawkas State Game Land</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Dawkas State Game Land</td>
<td>49</td>
<td>Kings Mountain</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Kintos Plantation</td>
<td>49</td>
<td>Lake Norman State Park</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Lake Norman State Park</td>
<td>49</td>
<td>Little Sugar Creek Greenway</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Mayo River State Park</td>
<td>49</td>
<td>Mc Alpine Creek Greenway</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Mclain Creek Greenway</td>
<td>49</td>
<td>Middleton Place Conservation Easement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Middleton Place Conservation Easement</td>
<td>49</td>
<td>Paris Mountain State Park</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Paris Mountain State Park</td>
<td>49</td>
<td>Richland Bridge State Park</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Sequoinsent State Park</td>
<td>49</td>
<td>Shuffletown Community Park</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Shuffletown Community Park</td>
<td>49</td>
<td>The Mulberry Tract</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Andrew Jackson State Park</td>
<td>49</td>
<td>Blackbeard Island National Wildlife Refuge</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Blackbeard Island National Wildlife Refuge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 18. The potential number of nonindigenous freshwater species reported outside of specific protected areas from the Protected Areas Database of the United States (PAD-US, Conservation Biology Institute Edition) but hydrologically connected by waterways at the subregion level in the South Atlantic Landscape Conservation Cooperative region.—Continued
Figure 19. Number of nonindigenous aquatic species per subbasin within the South Atlantic Landscape Conservation Cooperative region.
Factors Associated With Nonindigenous Aquatic Species Richness

Results and Discussion of Species Richness

Nonindigenous aquatic species richness among the 109 subbasins that compose the SALCC region’s watersheds ranged from 1 to 59 (mean 15.6, standard deviation 10.5) species (fig. 19). Many of the most species-rich subbasins were near the northwestern portion of the SALCC region or near major cities and towns, whereas the subbasins with the lowest numbers of species tended to be in the more rural areas of eastern Georgia.

Models that passed the diagnostic tests and had the lowest AICc scores all included five variables, all with statistically significant coefficients (table 4). The best OLS model (log-transformed housing density, reservoirs, and elevation range and nontransformed road density and longitude) produced an adjusted R² of 0.625 and an AICc of 272.601. At least three models with five variables had similar AICc scores (the lowest was 298.03), but these models failed the VIF test for collinearity. The VIF was less than 1.93 for all variables in the best model, indicating a strong lack of collinearity. Spatial autocorrelation (Moran’s I) was greater than 0.10 for all residuals, indicating lack of autocorrelation. The Jarque-Bera test was greater than 0.10 for all variables, indicating no evidence for nonnormality of residuals. The Breusch-Pagan & Koenker test was not statistically significant (p=0.18) for the best OLS model, indicating lack of evidence for spatial nonstationarity. The graphically weighted regression for the best OLS model did not improve the adjusted R², and the AICc was slightly higher than the OLS, suggesting that GWR was not an improvement over OLS. Figures 20–24 show maps of the input values for each of the five parameters that were included in the best OLS model.

Table 3. Explanatory variables used in ordinary least squares (OLS) analysis of species richness for subbasins in the South Atlantic Landscape Conservation Cooperative region.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Housing density</td>
<td>Average of housing density within each subbasin</td>
<td>GIS analysis of U.S. Census Bureau data (<a href="http://silvis.forest.wisc.edu/">http://silvis.forest.wisc.edu/</a>).</td>
</tr>
<tr>
<td>Road density</td>
<td>Total length of roads in a subbasin divided by subbasin area</td>
<td>USGS compilation of major roads.</td>
</tr>
<tr>
<td>Reservoir density</td>
<td>Total area of reservoirs in a subbasin divided by subbasin area</td>
<td>USGS compilation of major reservoirs.</td>
</tr>
<tr>
<td>River density</td>
<td>Total length of main rivers in a subbasin divided by subbasin area</td>
<td>USGS compilation of major rivers.</td>
</tr>
<tr>
<td>Elevation range</td>
<td>Difference between minimum and maximum elevation</td>
<td>USGS DEM 30-m resolution.</td>
</tr>
<tr>
<td>Elevation mean</td>
<td>Mean elevation within each subbasin</td>
<td>USGS DEM 30-m resolution.</td>
</tr>
<tr>
<td>Elevation maximum</td>
<td>Maximum elevation within each subbasin</td>
<td>USGS DEM 30-m resolution.</td>
</tr>
<tr>
<td>Latitude</td>
<td>Latitude of centroid of each subbasin</td>
<td>NHD dataset.</td>
</tr>
<tr>
<td>Longitude</td>
<td>Longitude of centroid of each subbasin</td>
<td>NHD dataset.</td>
</tr>
<tr>
<td>Subbasin area</td>
<td>Area in hectares</td>
<td>NHD dataset.</td>
</tr>
</tbody>
</table>

Table 4. Ordinary least squares (OLS) results for best models that pass all diagnostic tests with 3, 4, and 5 variables.

<table>
<thead>
<tr>
<th>AdjR²</th>
<th>AICc</th>
<th>JB</th>
<th>K(BP)</th>
<th>VIF</th>
<th>SA</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.504</td>
<td>300.620</td>
<td>0.609</td>
<td>0.585</td>
<td>1.730</td>
<td>0.105</td>
<td>+LOGHOUSDEN*** +ELEVMAX*** -DNROAD***</td>
</tr>
<tr>
<td>0.586</td>
<td>281.986</td>
<td>0.914</td>
<td>0.130</td>
<td>1.859</td>
<td>0.123</td>
<td>+LOGHOUSDEN*** +ELEVRNG*** -DNROAD*** +LGDNWATBOD***</td>
</tr>
<tr>
<td>0.625</td>
<td>272.601</td>
<td>0.501</td>
<td>0.047</td>
<td>1.928</td>
<td>0.837</td>
<td>+LOGHOUSDEN*** +LOGELEVRNG*** +LON*** -DNROAD*** +LGDNWATBOD***</td>
</tr>
</tbody>
</table>
Figure 20. Housing density per subbasin within the South Atlantic Landscape Conservation Cooperative region.
Factors Associated With Nonindigenous Aquatic Species Richness

Figure 21. Elevation range per subbasin within the South Atlantic Landscape Conservation Cooperative region.
Figure 22. Longitude value per subbasin within the South Atlantic Landscape Conservation Cooperative region.
Figure 23. Road density per subbasin within the South Atlantic Landscape Conservation Cooperative region.
Figure 24. Reservoir density per subbasin within the South Atlantic Landscape Conservation Cooperative region.
The best statistical model included a combination of five anthropogenic and environmental variables that explained approximately 62 percent of the variation in nonindigenous aquatic species richness. Three of these variables were anthropogenic (densities of housing, roads, and reservoirs), and two were environmental (elevation range and longitude). Highest species richness occurred in the more upland, mountainous regions, where elevation range favored reservoirs and attracted urban centers such as Atlanta, Charlotte, and Greensboro. The positive correlation of housing density with nonindigenous aquatic species richness is expected, as these urban centers with large populations represent increased observation opportunities and high availability of nonnative species such as those in retail aquarium stores, bait shops, and nurseries with aquatic plants. The positive correlation with large reservoirs may be associated with high visitation rates by boaters and anglers who release bait or move plants and other organisms via boat trailers or bilge water. Gido and Brown (1999) also found that the number of introduced species increased with number of reservoirs and drainage area, but decreased with native species richness. Elevation range also showed a positive correlation with nonindigenous aquatic species richness, which may be due to greater diversity of aquatic habitat and higher ranges of water temperatures that favor a greater range of aquatic species. A positive correlation with more easterly longitudes was found, which emphasized a declining trend in the number of species found in the coastal plains towards the southern portion of the SALCC region in southeastern Georgia. The negative correlation with road density is somewhat unexpected, but may reflect increasingly poor habitat conditions associated with extensive road networks in heavily urbanized areas.

The explanatory power of the best OLS model was greater than commonly used rejection thresholds (for example, 50 percent), and a next step may be to further examine the predictive accuracy of the model by using a cross-validation approach. However, because of the limitations of the input data discussed previously, including nonrandom or nonsystematic sampling, lack of independence of observations, and inability to estimate detection probability, this OLS model may not be appropriate for use as a predictive model. If the model and data were not subject to these statistical biases, the model could be used in a predictive manner by taking projections for the three anthropogenic variables (housing, road, and reservoir density; the two environmental variables would remain unchanged) and substituting them into the OLS model to obtain predictions of nonindigenous aquatic species richness for each subbasin.

Network Analysis Tools for Modeling Diffusion Processes

The diffusion of organisms across a landscape has been modeled by using a variety of methods, including analytical models such as Fisher’s diffusion equation (Skellam, 1951; Shigesada and Kawasaki, 1997), gravity models, and statistical models (Frieden and others, 2014). Diffusion of terrestrial organisms often assumes homogenous landscapes and random movement, allowing the simplification of uniform range expansion in all directions (Bartak and others, 2013). However, many aquatic organisms within the USGS NAS Database are restricted in their natural movements to expanding within the existing stream networks that are hydrologically connected. These dendritic patterns may drastically alter movement patterns, and movement may be very different in the upstream versus downstream direction (Peterson and others, 2013). In this part of the study we investigated the use of network analysis tools to look at the nonindigenous aquatic species observations from a time series perspective. Network structures provide an efficient means of measuring accurate hydrologic distances between points, as opposed to using straight-line Euclidean distances. A primary objective of this study was to look at basic patterns in the spread of nonindigenous aquatic species observations for individual species within a single, hydrologically connected network.

In addition to developing tools to examine empirical patterns in the USGS NAS Database, we wanted to develop a simple approach to model the potential spread of an organism throughout a stream network from a single starting point by using network data structures to represent hydrologic connectivity. Although fully developing a diffusion model was beyond the scope of this project, an important component of such models is a measure of the rate of spread, which in the case of aquatic organisms in stream systems requires realistic measurements of distance along the stream network. Therefore, we have described some of the capabilities of network tools to provide these measurements and suggest how they can be useful to managers.

We again note that there are characteristics of the USGS NAS Database that limit its suitability for modeling diffusion processes. These are the same limitations described earlier and include the unavailability of absence data and reliance on opportunistic rather than systematic sampling. As such, any patterns in the data may reflect more about the data collection process than actual diffusion of an organism. We also note that deliberate or accidental movement of organisms by human activities often results in the expansion of nonindigenous aquatic species between hydrologically unconnected stream networks or in long distance dispersal in the same hydrologic unit. For this diffusion process model, anthropogenic or other effects that allowed range expansion between unconnected water bodies were not addressed.
Methods for Modeling the Diffusion Process

The NHDPlus (version 2) was used for the diffusion process model. We also examined the NHD (http://nhd.usgs.gov/), which is available at higher resolution and was more up to date, but we determined that the NHDPlus resolution (1:100,000 scale) had adequate detail and greatly reduced computer and disk space demands compared to the NHD dataset. Additionally, the NHDPlus datasets combine additional features from the National Elevation Dataset (NED) (http://nationalmap.gov/elevation.html), the National Land Cover Dataset (NLCD) (http://www.mrlc.gov/nlcd2011.php), and the Watershed Boundary Dataset (WBD) (http://nhd.usgs.gov/wbd.html).

Initially, the use of the ArcGIS tool Utility Network Analyst, which provides interactive methods of analyzing data along a network, was explored. NHDPlus was imported into ArcGIS as a “geometric network” that consisted of a flowline and node file used to create a network file. The “geometric network” format differs from the ArcGIS “network” format used in the Network Analyst Extension, which is not available with the standard ArcGIS installation and is mostly used for transportation networks. Instead, the Utility Network Analyst toolbar was used to explore the possibility of analyzing the nonindigenous aquatic species data and modeling diffusion within the standard ArcGIS environment. It was then possible to use the Network Analyst tool to interactively find hydrologically connected reaches from one or more starting points. This capability is potentially useful for investigating the threat to protected areas and reaches that are hydrologically connected to point observations of aquatic species. However, further investigation of the ArcGIS programming environment needed to perform these analyses without user interaction suggested that there were alternative programming tools in the R programming environment (R Core Team, 2015) that better served the needs of this study.

The ArcGIS Python language tool Spatial Tools for the Analysis of River Systems (STARS) (Peterson and Ver Hoef, 2014) was used to help generate spatial networks that can be used by the R software library called Spatial Stream Network (SSN) (Ver Hoef and others, 2014). STARS is able to import a polyline shapefile from NHDPlus and generate a stream network object that can be analyzed by SSN and accessed by other R libraries. STARS also can snap points (for example, observation locations from the USGS NAS Database) to the stream network and calculate network distances between points, treating upstream and downstream distances differently. The STARS tool is well documented and has an active user development group involving numerous Federal agencies including the USGS (see http://www.fs.fed.us/rm/boise/AWAE/projects/SpatialStreamNetworks.shtml).

The STARS procedures required extensive and complex ArcGIS editing of the NHDPlus shapefiles, most notably the requirement that braided streams must be converted to single stream segments and that multiple confluences converging on a single point (so-called complex confluence) must be modified so that no more than two stream segments converge on a single node. Appendix 3 provides a simplified user manual for the STARS procedures that was followed for this study. The basic procedures were run within ArcGIS ArcMap by using the STARS toolbar and included the following steps: (1) extract the flowlines and node shapefile for the study area from NHDPlus, (2) project these files to a Cartesian coordinate system, (3) check and correct the shapefile topology by using the STARS tools, (4) create a Landscape Network (LSN) file by using STARS, (5) snap the nonindigenous aquatic species data points to the network by using STARS, and (6) export the LSN format to the SSN object by using STARS.

By using STARS, a stream network object was developed for the Suwannee River Basin within the SALCC region of northern Florida, which spatially snapped the nonindigenous aquatic species data points for a single species, Asian clam, to the Suwannee River Basin network and calculated network distances among points. This species and the Suwannee River Basin were chosen because of the large number of recorded observations in this well-studied river system. Next, the R library SSN was used to develop an R script (see app. 4) to read the stream network objects, plot maps of the nonindigenous aquatic species data points on the stream network, create all-pairs distance matrices for nonindigenous aquatic species data points, and generate graphs showing the expansion of network distances from the original observation over time and the network distances between consecutive observations.

In addition to developing an R script for generating spatiotemporal graphs for existing nonindigenous aquatic species data points, a simple method was also illustrated to portray a key underlying variable needed to simulate the spread of a diffusion front across a network, namely, the distance from a single origin point to any other location across the stream network. The SSN was used to attach the distance from the origin point to a set of simulated points spread evenly across the network, which allows for easy display of network distances in a geographic information system context. The simulated network points were intersected with the protected areas data layer and used to tabulate the smallest distance from the earliest recorded observation for Asian clam to each protected area.

Results and Discussion for the Diffusion Process Modeling

We successfully converted the NHDPlus shapefile for the Suwannee River Basin into a stream network object by using the ArcGIS tool STARS and developed further analyses for the largest dataset of a single species in the USGS NAS Database for the SALCC region, Asian clam. Figure 25 provides a synoptic view of Asian clam locations along the Suwannee River Basin network. The spatiotemporal progression of distances from the earliest recorded observation (1967) to subsequent observations (fig. 26) and distances
Figure 25. Suwannee River drainage (National Hydrography Dataset Plus [NHDPlus] shapefile) and data points representing observations for *Corbicula fluminea* (Asian clam). The first recorded observation is shown as occurring near the mouth of the Suwannee River.
Figure 26. Distance from the earliest record to subsequent observations of *Corbicula fluminea* (Asian clam) in the Suwannee River Basin. Note that the earliest record was near the mouth of the Suwannee River; therefore, most of the subsequent observations necessarily were upstream rather than downstream.
between temporally consecutive pairs of Asian clam data points (indicating whether the more recent record was upstream or downstream from the previous point) (fig. 27) were determined. The large increase in data after 2000 was due to intense sampling efforts, above casual reporting. Because the earliest recorded observation was near the mouth of the Suwannee River, most of the subsequent observations necessarily were upstream rather than downstream. The time interval between consecutive pairs of records for Asian clam shows that multiyear gaps in records are not uncommon (fig. 28). The R code used to produce figures 26–28 is provided in appendix 4.

We mapped the first recorded observation location for Asian clam and the simulated points across the stream network, representing different distances from that first location (fig. 29). Categories of equidistant points are given the same color, illustrating clusters of points in different sections of the network that are the same distance from the first location. The colors also reflect whether a point is upstream or downstream from the origin. The protected areas are delineated, showing the large number of areas potentially impacted by the first location of Asian clam and the distances from this point to the protected areas. Table 5 tabulates the minimum distance from the Asian clam origin point to each protected area. Although this information is simulated, except for the first location, the distances could be used in a diffusion model to provide more accurate distance measurements rather than the more commonly used straight line distances. Additionally, if rough estimates are available on dispersal rates for an organism, the distances in table 5 could be used as a rough indicator of how long it may take for an organism to begin invading a given protected area.

The network analysis tools provided by STARS and SSN allowed for the creation of useful maps and graphs by using standard R scripts. Importing the NHDPlus hydrology data and merging the nonindigenous aquatic species data into standard R objects that can be further analyzed by the numerous R libraries mean that this approach may be useful for other research. Other researchers have found the ability to measure actual hydrologic distances and conduct analyses or simulations along a network to provide useful advances (Peterson and others, 2013). However, we found that the procedures required to import the NHDPlus data into STARS were time intensive, requiring many hours of manual editing to prepare a single subregion, and suggest that this factor should be considered when evaluating this approach. The editing required was not due to errors in the NHDPlus dataset, but was due to topological simplifications required by the STARS format, most notably the prohibition against braided streams and complex confluences. Because the interest in this study was focused on exploring network analysis tools from a methodological perspective, the analysis was limited to the Suwannee River Basin and a single species, Asian clam.

The time series graph of the range expansion of Asian clam recorded observations (fig. 26) shows a pattern of rapid expansion in 1970, followed by multiple decades of no expansion until 2002, when a large collection of observations greatly expanded the documented range because of sampling efforts of the USGS and the Georgia Department of Natural Resources. Figure 26 illustrates some of the limitations of the data, such as the impact of uneven spatial and temporal sampling effort on the apparent maximum range or invasive front. Even though the Asian clam dataset had the largest number of observations for any species in the Suwannee River Basin (and the SALCC region), limitations of the data are apparent from the resulting analysis. This issue is readily seen by the large increases in the range (fig. 26) and multiyear gaps between consecutive pairs of data points (fig. 28), suggesting that the sample interval can strongly influence patterns in the apparent range expansion. The patterns seen in these analyses may reflect less about the diffusion process of the organism and more about the opportunistic nature of the data collection process.

The STARS and SSN tools provided a programming environment that was useful for making basic distance calculations that would be needed for simulating network processes. In particular, once the network data were imported into the R environment of the SSN library, calculating network distances between any two points was easily programmed. The signage of the distance measurements also indicated upstream versus downstream direction, thus allowing consideration of potential asymmetries in movement, which may be especially important for simulating passive dispersal. The illustrative map for Asian clam (fig. 29) and the table of distances to each protected area along the stream network (table 5) provide potentially useful information for managers that supplements information provided in the earlier protected areas analyses.
Figure 27. Distance between consecutive pairs of collection records of *Corbicula fluminea* (Asian clam) in the Suwannee River Basin.
Figure 28. Time interval between consecutive pairs of collection records of *Corbicula fluminea* (Asian clam) in the Suwannee River Basin, irrespective of consecutive pairs being upstream or downstream.
Figure 29. Suwannee River drainage exhibiting different network distances from the first record of *Corbicula fluminea* (Asian clam) to simulated points. See table 5 for distance to specific protected areas tabulated from this information.
Discussion

One widely used method to preserve biodiversity has been to set aside lands and designate them as “protected areas” (Mcdonald and others, 2008). Yet in as much as we would like to believe that these areas are indeed protected, they are still susceptible to change by human intervention (Spear and others, 2013). Urbanization encroaches on these protected areas and impacts them in multiple ways including pollution, global warming, noise, garbage, disease, and the introduction of invasive species (Mcdonald and others, 2009). The proximity of humans to protected areas has been shown to be a driver of initial invasions, whereas the subsequent spread tends to be driven by environmental factors, as for native species (Stohlgren and others, 2006). Iacona and others (2014) found that the number of households near protected areas was positively correlated with invasion as a result of an increase in both propagule pressure and disturbance. The size of those protected areas was also an important predictor of “invadedness” (a relative proportion of an area covered by an

Table 5. Protected area and distance to first recorded Corbicula fluminea (Asian clam) location in the Suwannee River Basin (see fig. 29 for illustration).

<table>
<thead>
<tr>
<th>Protected area</th>
<th>Distance (km)</th>
<th>Protected area</th>
<th>Distance (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alapaha Research Farm</td>
<td>322.3</td>
<td>Lower Suwannee National Wildlife Refuge</td>
<td>-40.9</td>
</tr>
<tr>
<td>Alligator Lake Park and Recreation Area</td>
<td>105.8</td>
<td>Mallory Swamp Restoration Area</td>
<td>75.0</td>
</tr>
<tr>
<td>Andrews Wildlife Management Area</td>
<td>0.0</td>
<td>Manatee Springs State Park</td>
<td>-9.8</td>
</tr>
<tr>
<td>Banks Lake National Wildlife Refuge</td>
<td>258.3</td>
<td>Moody Air Force Base</td>
<td>285.3</td>
</tr>
<tr>
<td>Bay Creek Conservation Area</td>
<td>254.1</td>
<td>Moody Air Force Base Annex</td>
<td>199.5</td>
</tr>
<tr>
<td>Beadles Lumber Co.</td>
<td>327.1</td>
<td>O’Leno State Park</td>
<td>110.2</td>
</tr>
<tr>
<td>Big Shoals Conservation Area</td>
<td>230.6</td>
<td>Olustee Creek Conservation Area</td>
<td>114.3</td>
</tr>
<tr>
<td>Big Shoals State Forest</td>
<td>233.2</td>
<td>Olustee Experimental Forest</td>
<td>164.2</td>
</tr>
<tr>
<td>Big Shoals State Park</td>
<td>229.3</td>
<td>Osceola National Forest</td>
<td>247.3</td>
</tr>
<tr>
<td>Camp Branch Conservation Area</td>
<td>190.8</td>
<td>Paradise Public Fishing Area</td>
<td>345.6</td>
</tr>
<tr>
<td>Deep Creek Conservation Area (SRWMD)</td>
<td>250.0</td>
<td>Parcners Branch Conservation Area</td>
<td>130.0</td>
</tr>
<tr>
<td>Deep Creek Plantation Conservation Easement</td>
<td>240.6</td>
<td>Peacock Springs Conservation Area</td>
<td>106.3</td>
</tr>
<tr>
<td>Florida Youth Ranches Conservation Easement</td>
<td>181.5</td>
<td>Raiford Wildlife Management Area</td>
<td>166.4</td>
</tr>
<tr>
<td>Fort White Mitigation Park Wildlife and Environmental Area</td>
<td>75.6</td>
<td>River Rise Preserve State Park</td>
<td>105.0</td>
</tr>
<tr>
<td>G. L. Drummond Conservation Easement</td>
<td>-17.4</td>
<td>Santa Fe River Ranch</td>
<td>124.2</td>
</tr>
<tr>
<td>Gaskins</td>
<td>364.5</td>
<td>Santa Fe Swamp Conservation Area</td>
<td>183.7</td>
</tr>
<tr>
<td>Giddens Pond</td>
<td>325.0</td>
<td>Stephen Foster Folk Culture Center State Park</td>
<td>215.8</td>
</tr>
<tr>
<td>Grady Conservation Area</td>
<td>66.2</td>
<td>Suwannee River State Park</td>
<td>155.2</td>
</tr>
<tr>
<td>Graham Conservation Area</td>
<td>153.9</td>
<td>Suwannee Valley Conservation Area</td>
<td>212.9</td>
</tr>
<tr>
<td>Grand Bay Wildlife Management Area</td>
<td>250.4</td>
<td>Swift Creek Conservation Area</td>
<td>203.3</td>
</tr>
<tr>
<td>Hatchbend Conservation Area</td>
<td>55.7</td>
<td>Troy Spring Conservation Area</td>
<td>80.2</td>
</tr>
<tr>
<td>Ichetucknee Conservation Area</td>
<td>67.0</td>
<td>Twin Rivers State Forest</td>
<td>136.9</td>
</tr>
<tr>
<td>Ichetucknee Springs State Park</td>
<td>74.6</td>
<td>Upper Alapaha Conservation Area</td>
<td>195.7</td>
</tr>
<tr>
<td>Lafayette Blue Springs State Park</td>
<td>121.9</td>
<td>Upper Alapaha Conservation Easement</td>
<td>191.3</td>
</tr>
<tr>
<td>Little River Conservation Area</td>
<td>76.1</td>
<td>Warner/Harrell Conservation Easement</td>
<td>162.8</td>
</tr>
<tr>
<td>Log Landing Conservation Area</td>
<td>29.3</td>
<td>Wetlands Reserve Program</td>
<td>314.5</td>
</tr>
<tr>
<td>Loncala, Inc. Conservation Easement #1</td>
<td>77.5</td>
<td>Woods Ferry Conservation Area</td>
<td>194.8</td>
</tr>
</tbody>
</table>
invasive plant) where larger areas were less invaded (Iacona and others, 2014). Human population density surrounding parks was the most consistently significant predictor of numbers of alien (nonnative) and invasive (nonnative and harmful) species of both plants and animals across diverse environments in South Africa’s national park system (Spear and others, 2013). Nonnative species for which the numbers are low or there is a lag period where the numbers are initially low and go undetected may be harder to detect than truly invasive species (Spear and others, 2013). Higher human density and wealth supported more invasions in some taxa (plants, aquatic invertebrates, reptiles, birds, and mammals), but not in others, such as fish, which have been so widely stocked that human density and wealth most likely have no relation to invasions (Pyšek and others, 2010). A review of studies on general species richness of invertebrates and nonavian vertebrates showed a consistent decline with urbanization, whereas plant species richness tended to increase with moderate urbanization; however, extreme urbanization nearly always reduced plant species richness (McKinney, 2008). The caveats of these types of urban gradient studies such as spatial scales, locations of cities, habitat quality in cities, and the fact that species richness does not take into account species abundance are not considered (McKinney, 2008). In practice, conservation planners could plan new protected areas that may resist invasions (Byers and others, 2002).

Human population distribution may change with climate change primarily as it relates to the possibility of water redistribution. In response, the pattern of introduced aquatic species will most likely follow the changes in human populations. Even without the human factor, invasive species distributions have been predicted to change with global climate change (Rahel and Olden, 2008; Carey and others, 2012; Bellard and others, 2013; Fobert and others, 2013). There may be a need to plan control methods for invasive species in anticipation of changes in water regimes due to climate change (Rahel and Olden, 2008). Bellard and others (2013) used six climatic and 12 land use variables in their distribution models and found that some invasive species ranges increased for aquatic and terrestrial invertebrates while some actually decreased for amphibians and birds. Increased impacts of introduced species due to climate change and globalization are expected to also include higher elevations that have been up to now relatively less impacted (Seimon and others, 2007; Pauchard and others, 2009).

The intention behind forecasting invasions is to provide an early warning system with the goal of preventing or delaying invasions (Byers and others, 2002; Welk, 2004). Iacona and others (2014) stated that the predictions need to be based on readily available data that do not require intensive survey work in order to keep costs considerably low. Forecasts or predictions can help prioritize monitoring efforts, thereby allowing for cost-effective measures to be put in place, which may minimize negative impacts. Caution should be taken, as the accuracy of predictions can be lessened if the ecosystem is undergoing increasingly rapid changes (Byers and others, 2002). Further complications of the accuracy of predictions may include evolution of invasive species. Phillips (2015) modeled invasion speed incorporating evolution; his models exhibited more than four times the variance in the spread rate versus not including evolution scenarios, therefore making spreading rates very hard to predict and demonstrating that evolutionary processes make invasions faster and more unpredictable. Once invasive species become established, the management, control, and eradication are expensive (Pimentel and others, 2005). Control of invasions will sometimes demand difficult decisions about land use and management priorities (Hobbs and Humphries, 1995). Many natural areas contain innumerable nonindigenous aquatic species, likely far more than managers can control or manage (Byers and others, 2002). However, accurate predictions of their future distribution may help to prioritize future budgeting (Buchan and Padilla, 2000; Keller and others, 2008).

Many studies have used a variety of methods to analyze presence-only data for predicting invasions such as regression analyses (Stohlgren and others, 2006), ecological niche modeling employing maximum entropy (MaxEnt) (Rödder and Lötters, 2010; Gallardo and others, 2013; McDowell and others, 2013; Lübcker and others, 2014; Quinn and others, 2014), and genetic algorithm rule-set production (GARP) models (Drake and Bossenbroek, 2004; Drake and Lodge, 2006; Chen and others, 2007), thereby identifying risk. Johnson and others (2006) used presence-only data to determine invasion rates, and Mingyang and others (2008) modeled potential habitats for nonindigenous aquatic species. A nonmodeling approach was used for a pathway risk analysis for ornamental marine nonindigenous aquatic species by using presence-only data (Zajicek and others, 2009). Guo and Olden (2014) used presence-only data in a quantile regression analysis to describe invasion patterns at multiple spatial scales. They found that the relation between natives and exotics varied across the spatial scales (2-, 4-, 6-, 8-digit HUCs).

The variables we selected for the nonindigenous aquatic species richness OLS model explained some of the variance, but overall the model was relatively weak. A similar correlate analysis by Lapointe and others (2012) on riverine invasion drivers found that colonization pressure, research effort, and range in elevation were most predictive of nonindigenous aquatic species richness. Using presence-only data, Johnston and Purkis (2011, 2012) used a cellular automata model to determine potential parameters for red lionfish transport in the western Atlantic Ocean. Modeling the diffusion process in our analysis by using the STARS tool was probably the least informative, as the results conveyed more about the nature of how the data were compiled and less about the true invasion process. Just as in the point-process modeling, small amounts of planned survey data would likely improve the OLS and STARS models as well.
Conclusions

This project was the first attempt to utilize the USGS NAS Database to inform a Landscape Conservation Cooperative conservation blueprint. A significant amount of effort was dedicated to determining the most appropriate use of these data for the purposes of identifying the mechanisms and patterns of aquatic species invasions. These investigations were regarded as a fundamental step required to undertake predictive modeling of future invasion pathways.

This investigation has resulted in substantial progress in overcoming some of the challenges associated with using presence-only data for predictive modeling. Importantly, we showed that only a relatively small amount of high quality data from planned surveys will greatly enhance the usefulness of presence-only databases such as the USGS NAS Database. This information will be useful in the development of future sampling protocols undertaken by SALCC collaborators that are designed to improve the understanding of the location and dispersal rates of new nonindigenous aquatic species observed in this region. The model can be applied to datasets for any species, native or nonnative, and it represents a potentially valuable tool for leveraging other presence-only information being gathered on SALCC conservation indicators.

In this study, the most direct use of presence-only data is the risk analysis of threats to protected areas. The USGS NAS Database can provide useful information for managers of protected areas by identifying nonindigenous aquatic species that are hydrologically connected to individual protected areas and therefore represent a threat of invasion or may indicate an undetected species that is already present. By identifying these threats as early as possible, management actions can be expected to be more effective to prevent invasion or implement eradication efforts at the early stages of establishment.

These analyses have shown that the SALCC region is characterized by a growing number of nonindigenous aquatic species, some of which are expected to have significant ecological impacts over time. We have also generated a modeling framework and strategy for greatly improving predictions of future distribution of nonindigenous aquatic species from presence-only records, such as those represented in the USGS NAS Database. Despite these efforts, we advise treating these methodologies as exploratory in nature.

References Cited


Mechanisms of Aquatic Species Invasions Across the South Atlantic Landscape Conservation Cooperative Region


Gallardo, B., zu Ermgassen, P.S.E., and Aldridge, D.C., 2013, Invasion ratcheting in the zebra mussel (Dreissena polymorpha) and the ability of native and invaded ranges to predict its global distribution: Journal of Biogeography, v. 40, p. 2274–2284.


Herborg, L., Mandrak, N.E., Cudmore, B.C., and MacIsaac, H.J., 2007, Comparative distribution and invasion risk of snakehead (Channidae) and Asian carp (Cyprinidae) species in North America: Canadian Journal of Fisheries and Aquatic Sciences, v. 64, p. 1723–1735.

Iacona, G.D., Price, F.D., and Armsworth, P.R., 2014, Predicting the invadedsness of protect areas: Diversity and Distributions, v. 20, no. 4, p. 430–439.


Quinn, A., Gallardo, B., and Aldridge, D.C., 2014, Quantifying the ecological niche overlap between two interacting invasive species: The zebra mussel (Dreissena polymorpha) and the quagga mussel (Dreissena rostriformis bugensis): Aquatic Conservation: Marine and Freshwater Ecosystems, v. 24, p. 324–337.


References Cited


Appendixes 1–4
Appendix 1. Descriptions of Species With Possible Ecological Impacts That Have Been Introduced to the South Atlantic Landscape Conservation Cooperative Region in the Past Decade

Osteopilus septentrionalis (Cuban treefrog).—Although the Cuban treefrog is native to Cuba, the Bahamas, and the Cayman Islands, it has adapted to a cooler climate in the United States. This species has been found within the South Atlantic Landscape Conservation Cooperative (SALCC) region in Florida and Georgia. In Florida, the earliest appearance in the Lower St. Johns subbasin occurred sometime in the early 1990s. Since then, several collections have been made in the drainage from 2002 to 2008. In northern peninsular Florida, the Cuban treefrog is established, considered invasive, and expanding its range northward. There have been four observations of Cuban treefrog in Georgia, the first of which was a collection near Savannah in 2004. The other three observations were documented as recently as 2013 from central Georgia and Jekyll Island. These animals are considered waifs (lone individuals); established populations are yet to be determined in Georgia. It has been predicted that the species could disperse from peninsular Florida to the panhandle region of the State and westward to Mexico, as well as additional locations in Georgia. The greatest threat the species poses may be to native frogs. In Florida they are known to prey on native hylid frogs.

Pomacea maculata (giant applesnail).—This very large freshwater snail (up to 155-millimeter shell height) is native throughout much of tropical and subtropical South America. The earliest documented introduction in the United States was in southern Florida in the late 1980s; the earliest documentation of giant applesnail in the SALCC region was in Lake Munson, near Tallahassee, Florida, in 2006. Since then, this applesnail has been found in numerous locations in southern Georgia, north along the Savannah River and to several coastal areas of South Carolina. Whether by natural dispersal or separate anthropogenic introductions outside the SALCC region, giant applesnail has been documented in Alabama, Mississippi, Louisiana, and Texas. It is known to be an agricultural crop pest, especially of rice paddies, and there has been very little success in controlling them.

Leptodiaptomus tyrrelli (northern snakehead).—This small, brown ectoparasitic copepod ranges from 4 to 9 millimeters long. It is native to Asia, where its common hosts, Carassius auratus (goldfish) and Cyprinus carpio (common carp), are also native species. It is present in Florida and Georgia, as well as the Great Lakes region and the West Coast. Japanese fishhouse was very likely introduced to North America with its host goldfish through the aquarium industry. This parasite can cause severe damage to the integument of its hosts, sometimes resulting in death by affecting the appetite of a fish, and subsequently affecting growth rates. This parasite attaches to the skin, fins, or gills by way of suckers and then feeds. It also may aid in the transmission of Rhabdovirus carpio (spring viraemia rhabdovirus), larval nematodes, and the fungal disease caused by Saprolegnia species in some parts of the world.

Skistodiaptomus pallidus (a calanoid copepod).—This small freshwater copepod is native to the north-central Plains States, the southern part of the Mississippi River Basin, Texas, and west to Colorado. It has expanded its range into Georgia, North Carolina, and South Carolina. Skistodiaptomus pallidus could have been introduced accidentally in bait buckets, on fishing equipment, by recreational boaters, or with hatchery stockings from the Mississippi River Basin. This species is an efficient omnivorous predator, with the ability to prey on preferred rotifers and microzooplankton. It also has been known to attain very high densities in suitable habitats, reaching 10,000 individuals per cubic meter. Skistodiaptomus pallidus became the primary calanoid copepod in a particularly eutrophic part of Lake Tahoe along the California-Nevada border after its introduction there, dominating two previously common species, Leptodiaptomus tyrrelli and Epischura nevadensis. Additionally, based on evidence it has been suggested that S. pallidus is an intermediate host for the parasitic worm Tanaorhamphus longirostris, although study of this occurrence has been limited.

Procambarus hayi (straightedge crayfish).—This crayfish is native to western drainages of the Tombigbee River from the Noxubee River in Mississippi and northward in the State into drainages of the Yalobusha, Skuna, Yocona, Tallahatchie, and Coldwater Rivers. Straightedge crayfish was documented in two streams in north-central Georgia in 2009. The origin of its introduction is unknown, but many crayfish have been introduced as the result of discarded bait. It may too early to see any impacts from its introduction; however, there is the potential to threaten native crayfish.

Channa maculata (blotched snakehead).—Like Channa argus (northern snakehead), the blotched snakehead is also native to shallow, vegetated areas of streams and ponds in eastern Asia (Japan, southern China, Vietnam, Taiwan, and the Philippines). It is smaller than northern snakehead, reaching approximately 33 centimeters in total length, and was most likely introduced as a food source for human consumption. However, a larger 31-inch specimen was collected in 2009 from the North Carolina portion of Lake Wylie, a reservoir on the Catawba River. Blotched snakehead is not established in the State. Because of its subtropical native distribution, it is not expected to survive the temperate climate of the southeastern United States except perhaps for southern Florida. But as with any large predator, it has the potential to impact native fauna if it becomes established.
Atractosteus spatula (alligator gar).—The alligator gar inhabits fresh, brackish, and occasionally saltwater coastal estuaries and major, slow-moving rivers. It is native to the Mississippi River Basin from southwestern Ohio and southern Illinois south to the Gulf of Mexico and the Gulf Coastal Plain from Econfina River, Florida, to Veracruz, Mexico. This range places the SALCC region well within the thermal tolerances of this very large piscivorous predator. This fish can reach 3 meters in length and can live for decades. They are voracious predators feeding on fish, birds, small mammals, and turtles. A single alligator gar was collected in July 2010 from Lake Wateree, a reservoir on the Wateree River, South Carolina. It is not known how this fish was introduced. Because this species has become popular recently in the aquarium trade, it may have been intentionally released after it had outgrown its aquarium.

Pterygoplichthys anisitsi (southern sailfin catfish).—This species is native to South America, where it can be found in a wide variety of habitats, ranging from relatively cool, fast-flowing, and oxygen-rich highland streams to warm, slow-flowing, lowland rivers and stagnant pools poor in oxygen. They are tropical fish, and populations typically are limited only by their lower lethal temperature, which has been found to be about 8.8–11 degrees Celsius (°C) in some species of this genus. Pterygoplichthys spp. are highly tolerant to poor water quality and commonly are found in polluted waters. They are known to use outflow from sewage treatment plants as thermal refugia and can readily adapt to changing water quality. Some species are salt tolerant. The southern sailfin catfish has been collected in the St. Johns River Basin, Florida, the presence of the specimen most likely being the result of an aquarium release. Pterygoplichthys spp. may compete with native fish and are also believed to have displaced several species of native minnows in Texas, including the federally threatened Dionda diaboli (Devils River minnow). Pterygoplichthys spp. are capable of surviving mesohaline conditions (up to 10 parts per thousand) for extended periods of time, thereby allowing for the use of estuarine and coastal areas for dispersal.
Appendix 2. Literature Reviewed for Planned Survey Data of Nonindigenous Aquatic Species


Appendix 2


Appendix 3. Instructions for Importing NHDPlus Shapefile for Use in the R Library Called SSN

The goal of this procedure is to create one or more stream network files from the National Hydrography Dataset Plus (NHDPlus) that can be used in the R library called Spatial Stream Network (SSN). To create the SSN network file(s), we used an ArcGIS tool called Spatial Tools for the Analysis of River Systems (STARS) (v. 2.0.1), which is written in Python. STARS was loaded into the ArcGIS toolbox as described in the STARS manual.

The basic STARS procedure was to start with an NHDPlus flowline shapefile that represents the stream network of interest. STARS was used to convert this shapefile to an intermediate network format called LSN, which is a personal geodatabase format. The LSN file was checked for various errors by using STARS preprocessing steps. The LSN file cannot be edited, so any problems found must be fixed in the original NHDPlus shapefile first, and then the LSN file must be recreated and checked again for errors. After the LSN file was correct and finalized, STARS was used to export it to the desired SSN format.

DETAILED USER INSTRUCTIONS DEVELOPED FOR THIS PROJECT

Page numbers and sections below refer to STARS 2.0.0 tutorial manual. The STARS toolbox must be loaded into ArcToolbox (see sections 2 and 5). Before installing STARS, you should ensure that PythonWin version 2.7.3 is also installed for use with ArcInfo. (NOTE: Use the 219 version of 2.7 called pywin32-219.win32-py2.7.exe. There are newer versions of PythonWin, but ArcInfo does not work with them.)

Load an NHDPlus flowline shapefile into ArcMap (make sure it is projected to, for example, Albers). The STARS manual recommends deleting any Objectid fields in the shapefile table (p. 45), but I failed to do this, and all was well.

Create or empty an LSN directory to house the new LSN files (do not put other files in this directory; p. 12, 45–56).

1. Run the STARS “Polyline to LSN” preprocessor option (p. 12; proper file extensions are needed for the names: be sure to put lsn.mdb for output and .shp for input). This should only take a few minutes for a Suwannee River Basin size hydrologic unit code (HUC).

(Optionally load and inspect the LSN geodatabase in ArcMap. Per p. 12, everything should overlay the original NHDPlus shapefile.)

2. Run the STARS “Check Network Topology” option (p. 13–16; file extensions for nodes and edges do not matter at this point). This will report any errors in the LSN file and may take at least 20 minutes to run. Use “node_cat” as the field name in the STARS input dialogue (p. 14). Because the NHDPlus data are already clean, you can use a very small tolerance (for example, 1 meter) for the snap distance. If you are lucky, the NHDPlus data will have no errors. You should inspect the “node_errors” shapefile (p. 15). This file does not tell you what the error is, so you will need to look for several possible errors. A common problem mentioned on p. 14 of the STARS manual shows a converging stream snapped to another stream but with no corresponding vertex (you can tell by selecting the flowlines, which should terminate at a shared vertex—the fix is to split the flowline at the vertex). This is called an “end point snapped to edge” on p. 23 of FLoWS v1 manual.

You will also need to manually check for “converging stream” nodes or “downstream divergence” nodes in the “node_cat” field of the “nodes” feature in the “lsn.mdb” geodatabase. These two types of nodes are not permitted in SSN, so you will need to zoom to each of these nodes and edit the original NHDPlus flowline shapefile to fix these problems.

The “converging stream” error is caused by two streams that join together at a single node, but that have no outlet downstream (p. 15). A typical fix would be to add a segment to form an outlet flowline below the shared vertex.

The “downstream divergence” error is caused by streams that diverge and rejoin farther downstream, as in a braided channel. SSN cannot handle these, so one side of the braided flowline should be deleted (p. 16).

3. After you fix the errors identified by the Check Network Topology option, run the STARS Identify Complex Confluences tool (p. 17; file extensions matter in this dialogue: include .txt for the output and lsn.mdb for the input). You will probably get a few complex confluences errors showing up in the output .txt file showing node ids with this confluence problem. These must be fixed by manually editing the original NHDPlus flowline .shp file by using ArcMap.

ArcMap Editing Procedure

- Open original flowline shapefile in ArcMap by using Editor
- Open “nodes” table in /LSN/lsn
• Select pointid to be edited

• Zoom to selected

• Open Topology toolbar, click on topology button, and select flowline for temporary topology

• Click on topology select button, click on polylines to decide what node to move where, possibly delete a flowline, or add a vertex to move one of the polylines to

• On Editor bar, choose arrow select

• Click on polyline segment to move to, right-click, and insert vertex

• On Topology toolbar, choose arrow select

• Double-click on node and move to vertex elsewhere on flowline

4. Once edits are made to fix any errors in the original NHDPlus shapefile, quit ArcMap and delete all files in /lsn (note that ArcMap will create a file lock which requires closing ArcMap to delete lsn files). Rerun the STARS options (steps 1–3 above): Polyline to LSN, Check Network Topology, and Identify Complex Confluences.

5. Now you are ready to overlay the NIS data for a particular species onto the stream network. Run the STARS Snap Point to Landscape Network (p. 30–32) for a particular species (for example, Corbicula) for which you wish to calculate network distances. The species file should be a point shapefile in the same projection as the flowline shapefile. Note that the first field in the NIS point shapefile (that is, “SpecimenID”) will be reset to zero, so before running the STARS Snap option, add another field to the shapefile table and copy over the information from the first field to the new field. You will have to provide a reasonable value to use for the search radius for the snap distance, which can be derived by running the ArcGIS Near tool and using as input the NIS versus flowline shapefile to see what distance will snap most of the points to a nearby flowline. You might need to use a fairly “large” distance, such as 700 meters.

6. Run the STARS options to Calculate Upstream Distance-Edges (p. 35) and Calculate Upstream Distance (p. 35–36—Sites) (Note: you should run the Edges option before the Sites option). Make sure nothing is selected in Edges or Sites shapefiles; otherwise data will not populate in the “upDist” field.

(Additional functions could be run if RCAs are available, but in our case we are not using all of the options in STARS, so you can ignore anything in the manual having to do with RCAs or Watershed attributes (that is, ignore sections 9, 10, 11, 12, 14, 16, and 17).

7. Run the STARS option to convert the LSN to an SSN (p. 40–44). This will create the final output file needed for the R scripts to calculate distances among nonindigenous aquatic species observations. If you are lucky, the conversion will finish without error. However, you might get some new errors here that will need to be fixed. The error message will provide information about the offending nodes. The only error I have seen so far, a digitizing direction error, required looking up the duplicate rid in the “fromfeat” field in the LSN relationship table and then finding/deleting this duplicate in the original NHDPlus flowline shapefile. Once this correction is made, you will have to go back through all of the original STARS options, including deleting files in the LSN directory and starting from the polyline to LSN option (that is, steps 1–6).
Appendix 4. R Code Developed To Analyze Nonindigenous Aquatic Species Point Data

The process described in this appendix was used to produce figures 26–28 in the section “Network Analysis Tools for Modeling Diffusion Processes.” By convention, flow-connected locations lie on the same stream network (share a common downstream junction), and water flows from one location to the other. Flow-unconnected locations also lie on the same stream network, but do not share flow; that is, they flow down to a common junction, which is the primary distance provided in the distance matrix. Users may need to close ArcGIS if file locks exist for any files referenced here.

```r
library(SSN) #for help manual -- help(SSN)
#setwd("…/Suwannee/lsn_suwannee_nodes/lsn.ssn")
#list.files() #display files in directory - edges.shp, sites.shp, distance table netID1.dat, and binaryID.db
#setwd("…/Suwannee/lsn_suwannee_nodes") #includes Corbicula localities as “sites”; that is, burned in as “observations”; no “prediction” sites added

setwd("…/Suwannee/lsn_ok/lsn.ssn")
list.files() #display files in directory - edges.shp, sites.shp, distance table netID1.dat, and binaryID.db
setwd("…/Suwannee/lsn_ok") #includes equi-distant nodes as “sites”; that is, burned in as “observations”; no “prediction” sites added

#load ssn conversion from lsn personal geodatabase for Suwannee River.
Suw=importSSN("lsn.ssn", o.write = TRUE) #can include “prediction” file here, but not included here. Prediction file should be a shapefile in /lsn.ssn created by STARS
#Suw=importSSN("lsn.ssn", predpts="suwannee_nodes", o.write = TRUE) #can include “prediction” file here, but not included here. Prediction file should be a shapefile in /lsn.ssn created by STARS
names(Suw) #shows list of 31 variables across 111 sites
plot(Suw) #plot map using SNN defaults – figure 25
obsDF=getSSNdata.frame(Suw) #pull out the observation points into a dataframe
dim(obsDF) # dim() gives number of sites as rows [1], and fields as cols [2]
head(obsDF)
obsDSort=obsDF[order(obsDF$Date_),] #sort by date

# create downstream distance matrix among observed data points - distances are from one point to common junction of both points (asymmetric),
# NOT total hydrologic distance -distinguish between downstream (flow-connected) and total hydrologic distance (ignores flow direction=flow-Unconnected)
# note that the points must be on the same network to be flow connected or unconnected
# total distances can be created from this matrix by adding transpose, etc. (see below)
# two points on different forks will give positive distances to downstream junction when from-to matrix access is reversed, hence sum is total hydrologic distance
createDistMat(Suw, o.write = TRUE)
distObs=getStreamDistMat(Suw)
str(distObs)
names(distObs) # returns single field called “dist.net1”

#list out some distances
distObs$dist.net1[1:5,1:5]
distObs$dist.net1[39,9] #distance of 0 means this from-to pairing goes downstream
distObs$dist.net1[9,39] #distance should be positive if going upstream

#to get total in-stream distances between all point pairs (add asymmetric matrix plus its transpose)
# this is distance irrespective of flow direction; that is, total hydrologic distance
strDistNet1=distObs$dist.net1 + t(distObs$dist.net1)
strDistNet1[1:5,1:5]
strDistNet1[39,9] #find total in-stream distance between pid = 39 & 9, irrespective of flow direction
```
#setup to write to output text file
#write(array,fileconn,ncols,append=TRUE,sep=""," set filecon="" to write to screen
#filenm="corbicula_out9.csv"
#write header in first line
#write(paste("Date_", "pidfrom","pidto","dist_pr","dist_up","dist_down","dist_org","dist_org_up","dist_org_down," dist_cum_up","dist_cum_down"), file=filenm, append=TRUE)

#Upstream point measurements returned by SSN as positive distance, downstream as 0 distance (output as negative distance in the script).
dist_org=0
dtime=0
speed=0

#need to initialize these to coerce them to the correct vector data type
date_to=as.Date(as.character(obsDFsort$Date_[2]),"%Y%m%d")
date_from=as.Date(as.character(obsDFsort$Date_[1]),"%Y%m%d")
dist_cum_down[1:2]=0 #largest cumulative downstream distance from origin (NEGATIVE NUMBER)
dist_cum_up[1:2]=0 #largest cumulative upstream distance from origin
pid_org=obsDFsort$pid[1]
nrecs=dim(obsDFsort)[1] - 1
i=2

while(i <= nrecs){
  date_to[i-1] = as.Date(as.character(obsDFsort$Date_[i-1]),"%Y%m%d")
  date_from[i-1] = as.Date(as.character(obsDFsort$Date_[i-1]),"%Y%m%d")
  pid_to[i-1]=obsDFsort$pid[i]
  pid_from[i-1]=obsDFsort$pid[i-1]
  dist_pr1 = distObs$dist.net1[pid_from[i-1], pid_to[i-1]] #will return 0 if pid_to point is downstream
  dist_pr2 = distObs$dist.net1[pid_to[i-1], pid_from[i-1]] #use this to get downstream distance if any
  dist_up[i-1] = NA
  dist_down[i-1] = NA
  if(dist_pr1 > 0 && dist_pr2 > 0){
    dist_pr[i-1] = dist_pr1 + dist_pr2 #both points upstream from common junction -> total in-stream distance
    dist_down[i-1] = - dist_pr2
    dist_up[i-1] = dist_pr1
  }else if(dist_pr1 == 0){
    dist_pr[i-1] = - dist_pr2 #this point is downstream, so output as negative number...
  }else {
    dist_pr[i-1] = dist_pr1 #this point is upstream, so output as positive number...
  }
  dtime[i-1] = abs(difftime(date_to[i-1],date_from[i-1],"days"))
  if(dtime[i-1] > 0){ speed[i-1] = abs(dist_pr[i-1])/dtime[i-1]} else {speed[i-1] = NA}

  #track how far each point is back to origin
  dist_org1 = distObs$dist.net1[pid_org, pid_to[i-1]] #will return 0 if pid_to point is downstream
  dist_org2 = distObs$dist.net1[pid_to[i-1], pid_org]
  dist_org_up[i-1] = NA
  dist_org_down[i-1] = NA
  if(dist_org1 > 0 && dist_org2 > 0){
    dist_org[i-1] = dist_org1 + dist_org2 #both points upstream from common junction -> total in-stream distance
    dist_org_down[i-1] = - dist_org2
    dist_org_up[i-1] = dist_org1
  }else if(dist_org1 == 0){
dist_org[i-1] = - dist_org2 #this point is downstream of origin, so output as negative number…
dist_org_down[i-1] = dist_org[i-1]
}
else {
    dist_org[i-1] = dist_org1 #this point is upstream, so output as positive number…
    dist_org_up[i-1] = dist_org[i-1]
}
if(i > 2){
    if(dist_org[i-1] < dist_cum_down[i-2]) dist_cum_down[i-1] = dist_org[i-1] else dist_cum_down[i-1] = dist_cum_down[i-2]
    if(dist_org[i-1] > dist_cum_up[i-2])   dist_cum_up[i-1] = dist_org[i-1] else dist_cum_up[i-1] = dist_cum_up[i-2]
}
#write(paste(obsDFsort$Date_[i],“,",pid_from,“,",pid_to,“,",round(dist_pr,1),“,",round(dist_up,1),“,",round(dist_down,1),“,",round(dist_org_up,1),“,",round(dist_org_down,1),“,",round(dist_cum_up,1),“,",round(dist_cum_down,1)),file=filenm, append=TRUE)

i=i+1
}

spread_frm=data.frame(date_from,date_to,dtime,pid_from,pid_to,dist_pr,speed,dist_up,dist_down,dist_org,dist_or g_up,dist_org_down,dist_cum_up,dist_cum_down)

#spread_frm=data.frame(date_from,date_to,pid_from,pid_to,dist_pr,dist_up,dist_down,dist_org,dist_org_up,dist_org_ down,dist_cum_up,dist_cum_down)

spread=read.csv(filenm, header=T,sep=“,”,as.is=T)
mydt=as.Date(as.character(spread$Date_),”%Y%m%d”)
spread_frm=data.frame(mydt,spread)

#plot distances from first report, distinguish upstream from downstream – figure 26
plot.new()
yrange=range(spread_frm$dist_org_up,spread_frm$dist_org_down,finite=T)
#points 21=circle, 22=square, 23=diamond,24=triangle point up, 25=triangle point down
plot(spread_frm$date_to,spread_frm$dist_org_up, ylim=yrange,ylab=”Distance from First Report (m)”, xlab=”Date”, pch=24, col=”blue”)
points(spread_frm$date_to,spread_frm$dist_org_down,pch=25,col=”green”) abline(h=0)
lines(spread_frm$date_to,spread_frm$dist_cum_up,col=”blue”) lines(spread_frm$date_to,spread_frm$dist_cum_down,col=”green”) 

#plot distances between consecutive points, distinguish upstream from downstream – figure 27
plot.new()
yrange=range(spread_frm$dist_up,spread_frm$dist_down,finite=T)
#points 21=circle, 22=square, 23=diamond, 24=triangle point up, 25=triangle point down
plot(spread_frm$date_to,spread_frm$dist_up, ylim=yrange,ylab=”Distance between consecutive pairs (m)”, xlab=”Date”, pch=24, col=”blue”)
points(spread_frm$date_to,spread_frm$dist_down,pch=25,col=”green”) abline(h=0)
#plot time between consecutive points – figure 28
plot.new()
yrange=range(spread_frm$dtime,finite=T)
#points 21=circle, 22=square, 23=diamond,24=triangle point up, 25=triangle point down
plot(spread_frm$date_to,spread_frm$dtime, ylim=yrange,ylab=”Time between consecutive pairs (days)”, xlab=”Date”, pch=24, col=”blue”)