

# **Spatial Pattern Analysis of Chesapeake Bay Fish and Invertebrate Diversity, 2002 to 2012.**

## **INTRODUCTION**

Historically, the Chesapeake Bay has provided a legacy of commercial and recreational resources which are well represented by its fisheries. As the largest estuary in the United States it continues to provide great financial success while at the same time undergoing ecological stress in the form of eutrophication (Kemp et al., 2005), overfishing (Rothschild et al., 1994; Jackson et al., 2001), and habitat loss (Orth and Moore, 1983 and Kearney and Stevenson, 1991). Evidence suggests that the salinity gradient in Chesapeake Bay is a key factor responsible for changes in fish diversity (Jung and Houde, 2003). Biological estuarine salinity barriers occur at overlapping intervals of fully freshwater to 4 ppt, 2 to 14 ppt, 11 to 18 ppt, 16 to 27 ppt, and 24 ppt to fully marine (Bulger et al., 1993).

Diversity hotspot analysis using Geographic Information Systems (GIS) is a standard approach for evaluating geographic locations for potential conservation measures of threatened or endangered species (Mittermeier et al., 1998; Reid, 1998; and Myers et al., 2000). More recently, coldspot analysis has been advocated for similar purposes (Kareiva and Marvier, 2003; and Garofalo et al., 2007). While Chesapeake Bay has not been identified as a particular area of concern regarding threatened or endangered fisheries species, given its historical legacy of heavy resource use such diversity analyses could benefit future prospects for sustainable exploitation.

Typical tests for statistical significance assume the variables being tested are independent of one another; however, this is rarely the case in ecology, wherein phenomena are often structured with inherent patterns or patches (Legendre, 1993). When the null hypothesis of independence is rejected, the data are spatially autocorrelated as signified by the Moran's I statistic (Anselin, 1995). Because there are specific procedural measures to account for spatial autocorrelation in more robust statistical analyses, it is important to determine if the values are indeed autocorrelated.

## **METHODS**

The Chesapeake Bay Multispecies Monitoring and Assessment Program (ChesMMAP) is a long-term fishery-independent fish monitoring trawl survey. Five cruises per year (March, May, July, September, and November) employing a stratified random sampling design within the Chesapeake Bay main stem were completed

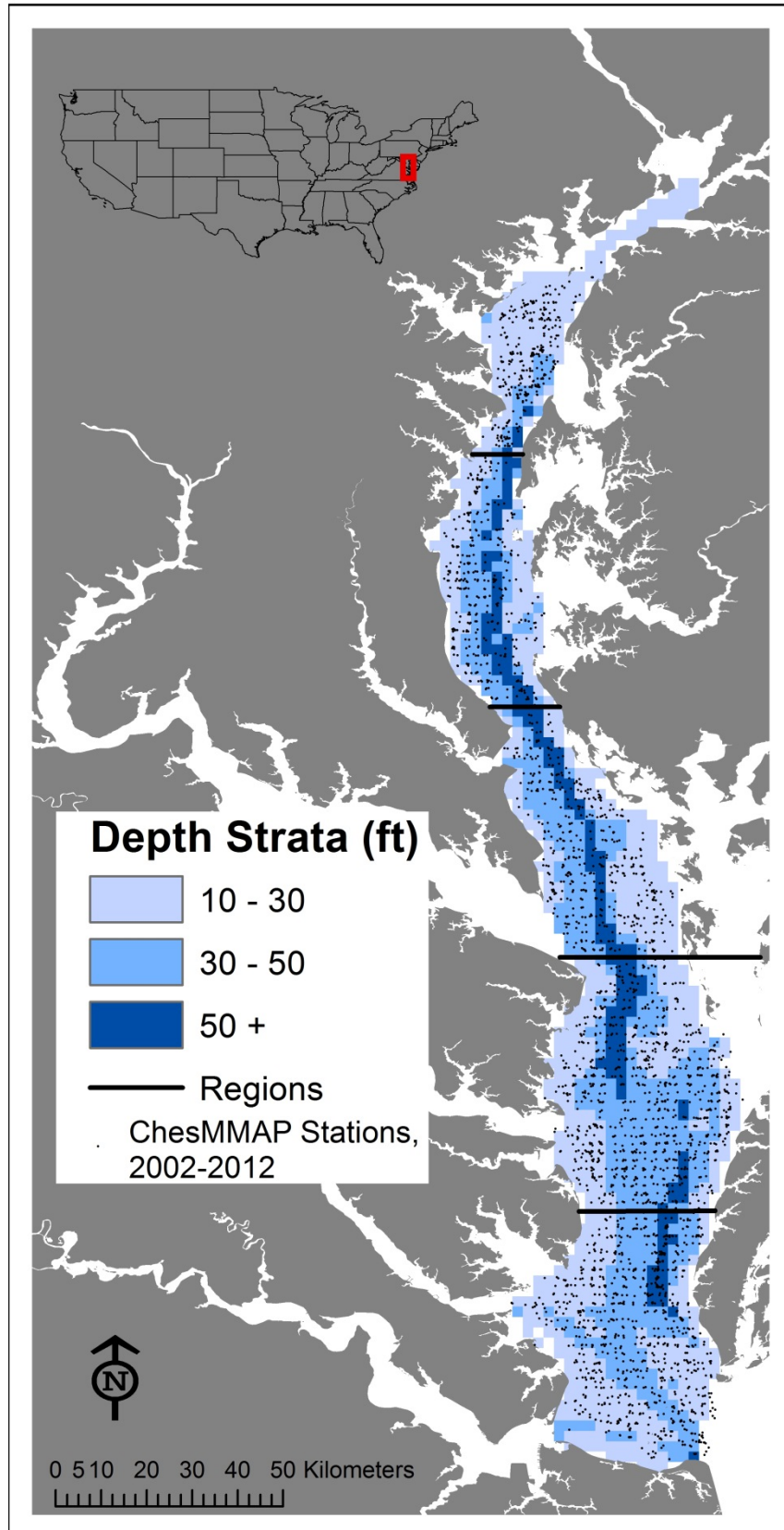
from 2002 to 2012. A total of 80 stations were sampled per cruise. The stratifications were defined by three depth strata and five roughly equal latitudinal regions (Figure 1, page 3). A georeferenced table of 3,052 station data points with total species count (species richness), total fish count, and salinity (ppt) attributes was analyzed. To account for the two sampling design stratification criteria, each depth strata within each of the five regions, a total of 14 strata (three strata in Regions 2, 3, 4, and 5 and two strata in Region 1) were analyzed separately.

The Shannon index is a widely used indicator of diversity in fisheries studies and shows a strong positive correlation with other indices of richness, evenness, and diversity (Garofalo et al., 2007). Based on catch number fields, the Shannon diversity index was calculated for each station trawled over the 11-year time span. The Shannon index is estimated as:

$$H' = - \sum_{i=1}^s p_i \ln p_i, \text{ where } p_i = \frac{n_i}{N},$$

where  $n_i$  is the number of individuals of species  $i$ ,  $N$  is the total number of individuals collected of all species, and  $s$  is the species richness, or total species count collected at that station (Pielou, 1966). The Shannon index increases as species richness increases (Levinton 1982), but is also influenced by rarely collected species (Peet 1974).

Figure 1. Location of ChesMMAP study area, depth strata, region lines, and sample points



To explore the evidence suggesting salinity plays a major role in determining fish diversity in Chesapeake Bay (Jung and Houde, 2003), spatial interpolation of the salinity attribute data was performed. Because the benthic habitat is the primary target for trawl gear in general, bottom salinity values were used. Ordinary Kriging with the linear semivariogram model produces the best fit for interpolation of salinity data sets from the Chesapeake Bay (Murphy et al., 2010). This method was used to produce a continuous surface of salinity with which to compare the diversity hotspots and coldspots. The interpolated salinity values were classified roughly according to the five overlapping salinity zone intervals determined by Bulger et al. (1993) into non-overlapping intervals as such: 1 to 3 ppt, 4 to 13 ppt, 14 to 17 ppt, 18 to 25 ppt, and 26 to 32 ppt.

To first test the null hypothesis that the point features were randomly distributed, the Average Nearest Neighbor (Spatial Statistics) analysis was run in ArcMAP 10.1. The degree of clustering was calculated using the High/Low Clustering (Getis-Ord General G) (Spatial Statistics) analysis in ArcMAP 10.1 via a spatial weights file based on the Shannon diversity index. To test the null hypothesis that the Shannon diversity values were independent of one another and to determine the degree, if any, of statistical autocorrelation, the Moran's I and probability statistics were calculated using the Univariate Local Indicators of Spatial Autocorrelation (LISA) procedure in GeoDa 0.9.9.14 (Anselin, 1995).

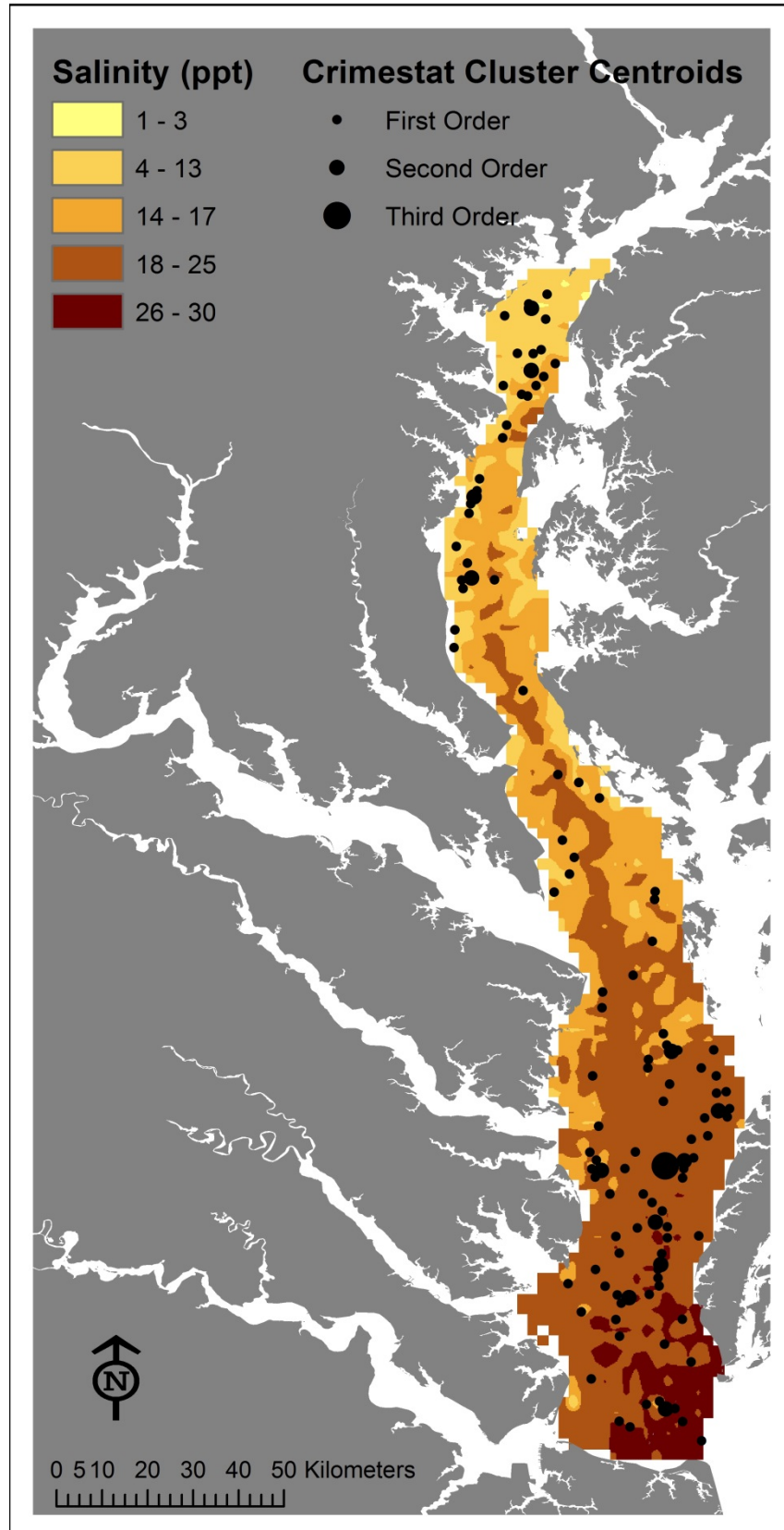
The Hotspot Analysis (Getis-Ord  $G_i^*$ ) (Spatial Statistics) was run in ArcMAP 10.1 to represent points where the diversity was significantly higher than expected ( $>+2.5$  and  $+3.5$  standard deviations) denoting hotspots, and significantly lower than expected ( $<-2.5$  and  $-3.5$  standard deviations) denoting coldspots. ArcMAP returns these standard deviation values as  $G_iZ$  Scores. Additionally, the Hotspot Analysis I was run in Crimestat III (Levine, 2007) using Nearest Neighbor Hierarchical Spatial Clustering (Nnh) with a minimum of 5 points per cluster. The results represent the standard deviational clusters as ellipses with centroids. CrimeStat calculates first-order clusters, then uses those clusters to determine second-order clusters (clusters of the first-order clusters) and so on (Smith and Bruce, 2008). These analyses enabled exploration of what species in particular make up the clusters and in what proportions as well as the value of the clusters to the total catch. The ellipses cluster centroids were superimposed on the interpolated salinity grid to visualize potential correlation.

## RESULTS

A total of 169 taxa (fish and invertebrates) was captured and identified during the entire ChesMMAAP time series. A total of 3,052 stations were used to represent the sampling point data and all taxa were incorporated into the diversity index calculations. The species richness ranged from 1 to 21 with a mean of  $5.02 \pm 3.39$ . The Shannon index ranged from 0 to 2.52 with a mean of  $0.88 \pm 0.58$ . Six taxa (3.5% of all taxa) were captured at more than 25% of the sampling stations and included *Leiostomus xanthurus* (Sciaenidae, 52%), adult female *Callinectes sapidus* (Decapoda, 48%), *Paralichthys dentatus* (Bothidae, 47%), *Micropogonias undulatus* (Sciaenidae, 44%), *Cynoscion regalis* (Sciaenidae, 37%), male *Callinectes sapidus* (Decapoda, 32%), and *Morone saxatilis* (Moronidae, 32%). Eighty-eight percent of the taxa were captured at less than 10% of the sampling stations.

Spatial interpolation of the bottom salinity data revealed an expected pattern of higher salinity at the mouth of the bay while lower salinity values appear near the headwaters (Figure 2, page 6). Additionally, based on the kriging operation the bottom salinity appears higher in the main channel while values remain low near the shoreline.

Figure 2. Ordinary kriging with linear semivariogram model of ChesMMAp bottom salinity with Crimestat hotspot analysis cluster centroids



The Average Nearest Neighbor analysis in ArcMAP indicated a significantly clustered spatial pattern of the sampled points with a z-score of -52.73 ( $p=0.000$ ). The GeoDa LISA procedure to test for autocorrelation indicated there were significant spatial outliers (clustering) in the diversity index values (Figure 3, page 8). With a Moran's I of 0.2189 ( $p=0.001$ ), the Shannon diversity exhibited significant autocorrelation. The High/Low Clustering analysis based on the diversity index showed a significant predominance of high-value clusters with a G-score of 34.84 ( $p=0.000$ ).

The Hotspot Analysis in ArcMAP revealed a large number of points where the Shannon diversity was both significantly higher than expected (hotspots) and significantly lower than expected (coldspots). Hotspots and coldspots are indicated by points where the values were  $>2.5$  and  $<-2.5$  standard deviations from the expected value, respectively. To represent more extreme hotspots and coldspots,  $\pm 3.5$  standard deviations were used in addition to the  $\pm 2.5$  threshold (Figure 4, page 9). The GiZ Score represents standard deviations. The Hotspot Analysis in CrimeStat generated 3 orders of clusters. The cluster centroids from the CrimeStat analysis were superimposed on the interpolated bottom salinity grid (Figure 2, page 6) as well as the ArcMAP Hotspot Analysis results (Figure 4, page 9). Note that the CrimeStat results do not differentiate between hotspots and coldspots.

Diversity hotspots at the  $\pm 3.5$  threshold were comprised of 118 taxa and 4,472 individuals and were generally concentrated in the lower half of the bay while coldspots consisted of 60 taxa and 2,038 individuals occurring in the upper reaches. Ninety-five percent of the total catch was made up of 56 taxa at the hotspots and 22 taxa at the coldspots. Further, the most abundant species at the hotspots, *Paralichthys dentatus*, *Callinectes sapidus*, and *Leiostomus xanthurus* only made up 20% of the individuals while the dominant coldspot species, *Callinectes sapidus*, *Morone saxatilis*, and *Morone americana* made up 50% of the total (Figure 5, page 10).

Figure 3. GeoDa Univariate Local Indicators of Spatial Autocorrelation (LISA) results showing significant autocorrelation among Shannon diversity values

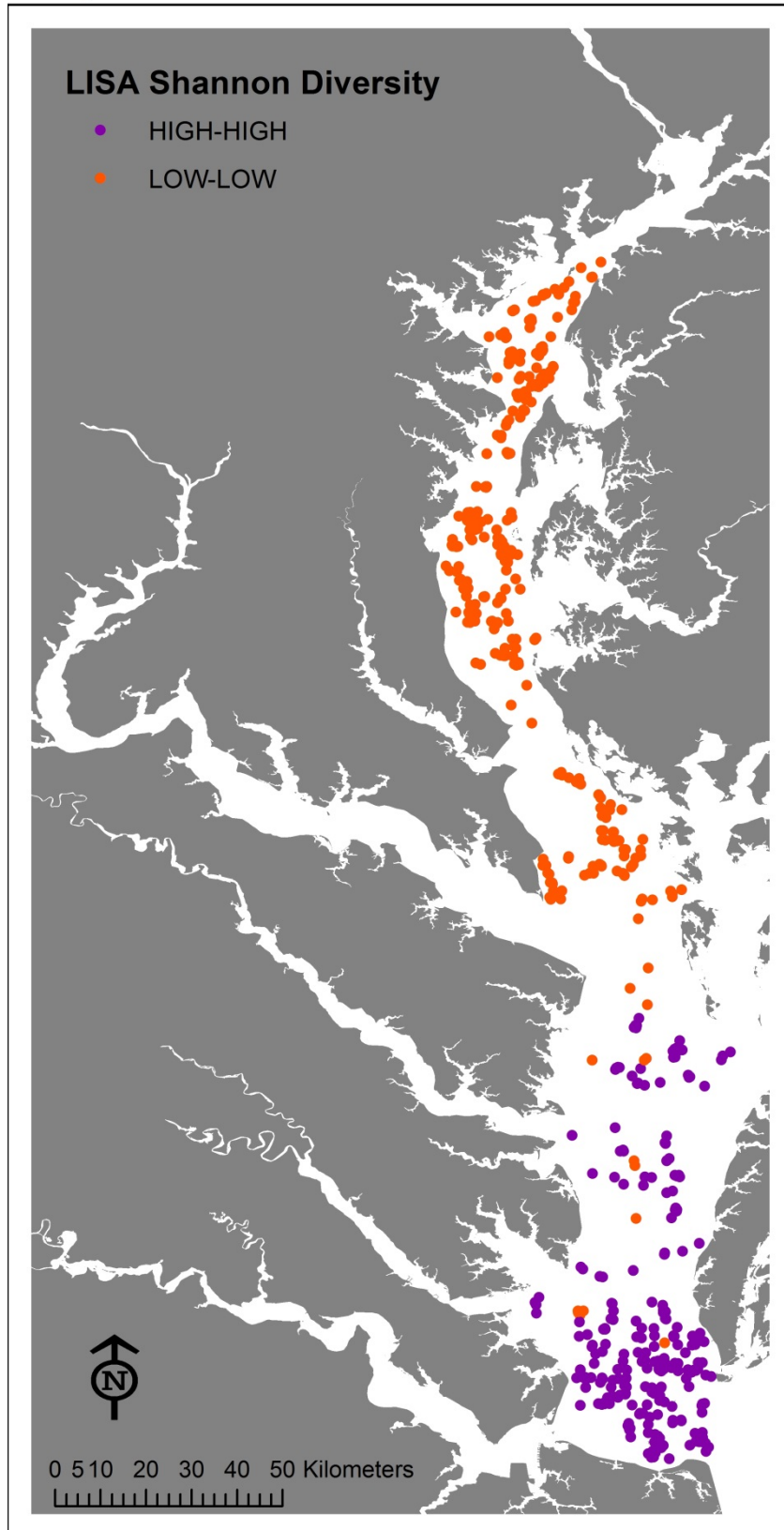




Figure 4. ArcMAP hotspots (red) and coldspots (blue) and Crimestat hotspot analysis cluster centroids

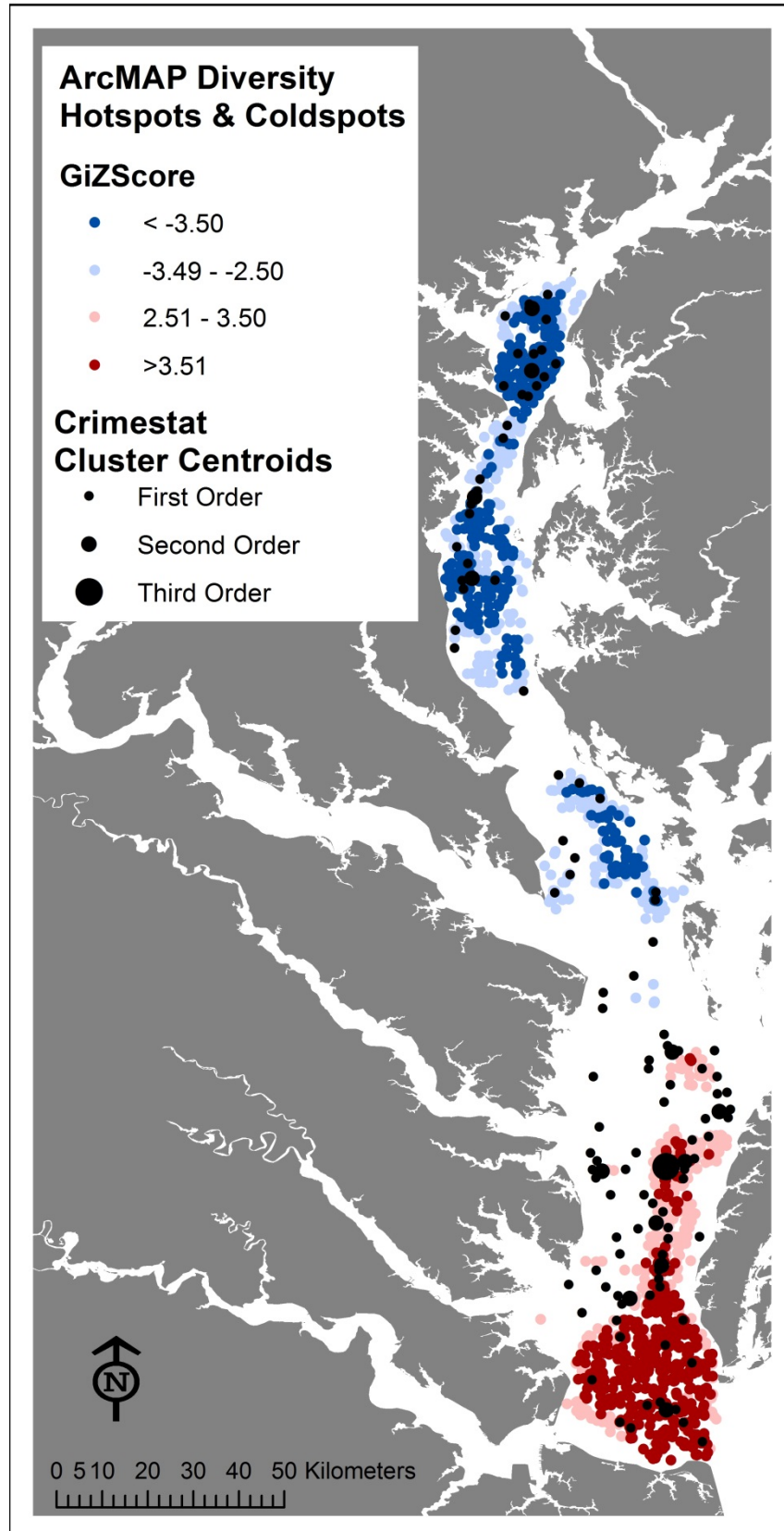
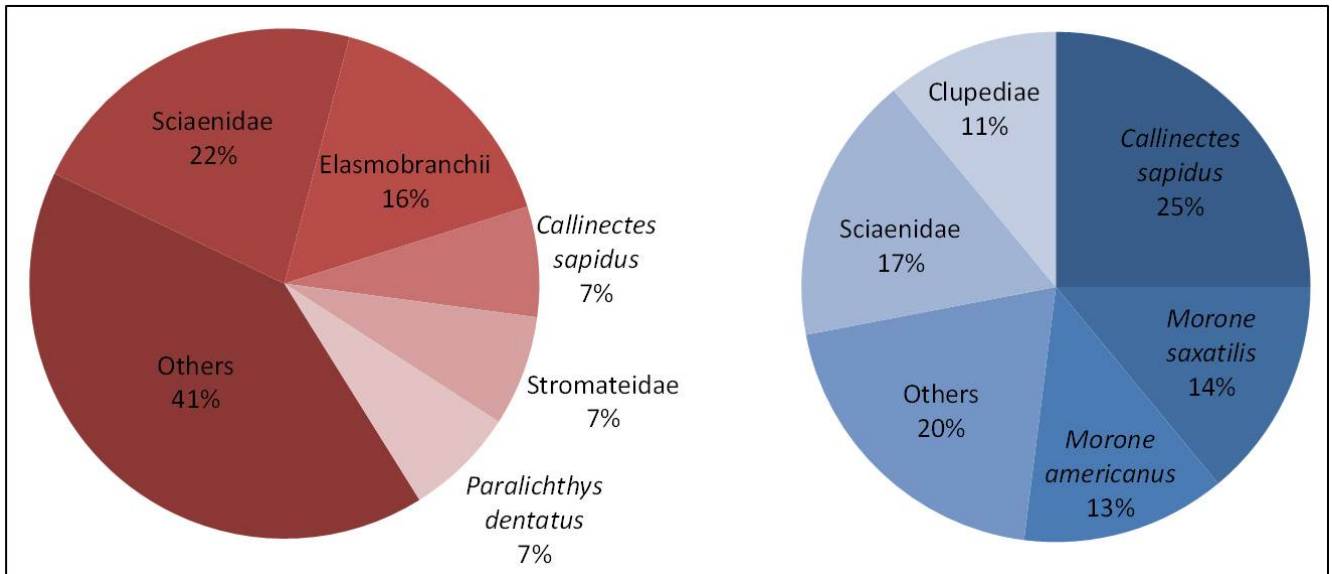


Figure 5. Taxa comprising the hotspots (>3.5 stdev), in red, and coldspots (<-3.5 stdev), in blue, calculated using ArcMAP Hotpot Analysis



## DISCUSSION

While the range of Shannon diversity in this study (0 to 2.52) was relatively high, the average diversity (0.88) was comparable to that of similar studies in the Southeastern United States. The range of Shannon diversity was found to be 0.8 to 1.8 in a Georgia estuary (Dahlberg and Odum, 1970) and 0.36 to 1.47 in a North Carolina estuarine eelgrass habitat (Adams, 1976). The Chesapeake Bay mean diversity was lower compared to that of studies farther south: 0.92 to 2.09 in a Texas bay (Hoese et al., 1968) and 1.24 to 1.68 in a South American tropical estuary (Tito de Morais and Tito de Morais, 1994). However, the diversity was relatively high compared to that of a small California Lagoon where it ranged from 0.03 to 1.11 (Allen and Horn, 1975), and a polluted Boston, Massachusetts river estuary where it ranged from 0.33 to 1.03 (Haedrich and Haedrich, (1974).

The interpolated grid of bottom salinity values was as expected with higher values near the mouth and lower values near the headwaters. The density of water masses increases as their mineral content increases; therefore, higher salinity water sinks below lower salinity water. The higher values in the main channel versus lower values near the shoreline suggest an increase in salinity with water depth.

The significantly clustered pattern of the sampling points is most likely due to the stratified random design of the sampling protocol. According to the design, certain criteria must be met when selecting random points. All three depth strata and all five latitudinal regions must be adequately covered and, theoretically, all sampling points within those criteria are randomly chosen. In reality, in the event that the location is untrawable due to physical hazards or restrictions, an alternate point is used. There are one or more alternate points assigned to each primary point and the alternate points are subject to the same sampling design. Over time, it is evident that certain locations are more favorable than others. For this reason, along with the other two spatial criteria, it follows that the sampling points will appear clustered rather than dispersed at random.

The Shannon diversity expressed significant spatial autocorrelation (clusters), and the predominance of high-value clusters in the data set suggests there are areas where diversity is significantly higher. Further, hotspots and coldspots of diversity were so widespread that two threshold standard deviations (GiZ Scores) were used,  $\pm 2.5$  and  $\pm 3.5$ . Accordingly, the Crimestat analysis generated three orders of clustering, indicating smaller clusters of diversity significantly clustering with each other.

The hotspots of diversity represented roughly twice the number of both taxa and individuals compared to the coldspots. The diversity of the hotspots is characterized by a relatively low abundance of dominant species, many taxa present, and a high occurrence of rare taxa. Diversity coldspots are predictably dominated by vast numbers of only a few major taxa and two of the dominant species belong to the same genus. The locations of the hotspots and coldspots of diversity suggest a strong direct relationship with salinity, consistent with previous studies incorporating statistical significance evidence for such a relationship in the Chesapeake Bay (Jung and Houde, 2003) and elsewhere (Castillo-Rivera, 2002).

## REFERENCES

- Adams, S. M. 1976. The ecology of eelgrass, *Zostera marina* (L.), fish communities. I. Structural analysis. *Journal of experimental marine Biology and Ecology*. 22: 269-291.
- Allen, L. G. and M. H. Horn. 1975. Abundance, diversity and seasonality of fishes in Colorado Lagoon, Alamitos Bay, California. *Estuarine and Coastal Marine Science*. 3: 371-380.
- Anselin, L. 1995. Local Indicators of Spatial Association—LISA. *Geographical Analysis*. 27(2): 93-115.
- Bulger, A. J., B. R. Hayden, M. E. Monaco, D. M. Nelson, M. G. McCormick-Ray. 1993. Biologically-based estuarine salinity zones derived from a multivariate analysis. *Estuaries*. 16(2): 311-322.

- Castillo-Rivera, M., J. A. Zavala-Hurtado, and R. Zárate. Exploration of spatial and temporal patterns of fish diversity and composition in a tropical estuarine system of Mexico. *Reviews in Fish Biology and Fisheries*. 12: 167-177.
- Dahlberg, M. D. and E. P. Odum. 1970. Annual Cycles of Species Occurrence, Abundance, and Diversity in Georgia Estuarine Fish Populations. *American Midland Naturalist*. 83(2): 382-392.
- Garofalo, G., F. Fiorentino, M. Gristina, S. Cusumano, & G. Sinacori. 2007. Stability of spatial pattern of fish species diversity in the Strait of Sicily (central Mediterranean). *Hydrobiologia*. 580: 117-124.
- Haedrich, R. L. and S. O. Haedrich. 1974. A seasonal survey of the fishes in the Mystic River, a polluted estuary in downtown Boston, Massachusetts. *Estuarine and Coastal Marine Science*. 2: 59-73.
- Hoese, H. D., B. J. Copeland, F. N. Moseley, and E. D. Lane. 1968. Fauna of the Aransas Pass Inlet, Texas. III. Diel and seasonal variations in trawlable organisms of the adjacent area. *Texas Journal of Science*. 20: 33-60.
- Jackson, J. B. C., M. X. Kirby, W. H. Berger, K. A. Bjorndal, L. W. Botsford, B. J. Bourque, R. H. Bradbury, R. Cooke, J. Erlanson, J. A. Estes, T. P. Hughes, S. Kidwell, C. B. Lange, H. S. Lenihan, J. M. Pandolfi, C. H. Peterson, R. S. Steneck, M. J. Tegner, R. R. Warner. 2001. Historical overfishing and the recent collapse of coastal ecosystems. *Science*. 293: 629-637.
- Jung, S. & E. D. Houde. 2003. Spatial and temporal variabilities of pelagic fish community structure and distribution in Chesapeake Bay, USA. *Estuarine Coastal and Shelf Science*. 58: 335-351.
- Kareiva, P. & M. Marvier. 2003. Conserving biodiversity coldspots. *American Scientist*. 91: 344-351.
- Kearney, M. S. & J. C. Stevenson. 1991. Island loss and marsh vertical accretion rate evidence for historical sea-level changes in Chesapeake Bay. *Journal of Coastal Research*. 7: 403-415.
- Kemp, W. M., W. R. Boynton, J. E. Adolf, D. F. Boesch, W. C. Botcourt, G. Brush, J. C. Cornwell, T. R. Fisher, P. M. Glibert, J. D. Hagy, L. W. Harding, E. D. Houde, D. G. Kimmel, W. D. Miller, R. I. E. Newell, M. R. Roman, E. M. Smith, & J. C. Stevenson. 2005. Eutrophication of Chesapeake Bay: historical trends and ecological interactions. *Marine Ecology Progress Series*. 303: 1-29.
- Legendre, P. 1993. Spatial Autocorrelation: Trouble or New Paradigm? *Ecology*. 74(6): 1659-1673.
- Levine, N. 2007. CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations (v 3.1). Ned Levine & Associates, Houston, TX, and the National Institute of Justice, Washington, DC. March.
- Levinton, J.S. 1982. *Marine ecology*. Prentice-Hall Inc., Englewood Cliffs, New Jersey.
- Mittermeier, R. A., N. Myers, & J. B. Thomsen. 1998. Biodiversity hotspots and major tropical wilderness areas: approaches to setting conservation priorities. *Conservation Biology*. 12: 516-520.
- Murphy, R. R., F. C. Curriero, W. P. Ball. 2010. Comparison of spatial interpolation methods for water quality evaluation in Chesapeake Bay. *Journal of Environmental Engineering*. 136: 160-171.
- Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B da Fonseca, & J. Kent. 2000. Biodiversity hotspots for conservation priorities. *Nature*. 403: 853-925.
- Orth, R. J., & K. A. Moore. 1983. Chesapeake Bay: An unprecedented decline in submerged aquatic vegetation. *Science*. 222: 51-53.
- Peet, R.K. 1974. The measurement of species diversity. *Annual Review of Ecology and Systematics* 5: 285-307.
- Pielou, E.C. 1966. The measurement of diversity in different types of biological collections. *Journal of Theoretical Biology* 13:131-144.

- Reid, W. V. 1998. Biodiversity hotspots. *Trends in Ecology and Evolution*. 13: 275-280.
- Rothschild, B. J., J. S. Ault, P. Gouletquer, & M. Héral. 1994. Decline of the Chesapeake Bay oyster population: a century of habitat destruction and overfishing. *Marine Ecology Progress Series*. 111: 29-39.
- Smith, S.C. and C.W. Bruce. 2008. *Crimestat III User Workbook*. National Institute of Justice, Washington, DC. 113 pages.
- Tito de Morais, A. and L. Tito de Morais. 1994. The abundance and diversity of larval and juvenile fish in a tropical estuary. *Estuaries*. 17(1): 216-225.

#### ACKNOWLEDGEMENTS

The author wishes to thank Dr. Hua Liu for her expert knowledge, guidance, and instruction throughout the Spatial Analysis of Coastal Environments class at Old Dominion University. The data was collected by the Virginia Institute of Marine Science Multispecies Research Group led by James Gartland, Christopher Bonzek, and Dr. Robert Latour.