Methods in Ecology and Evolution

MISS ALLISON L SUSSMAN (Orcid ID : 0000-0002-6996-9982)

DR BETH GARDNER (Orcid ID : 0000-0002-9624-2981)

A comparative analysis of common methods to identify waterbird hotspots

Allison L. Sussman[1,2*], Beth Gardner[3], Evan M. Adams[3,4], Leo Salas[5], Kevin P. Kenow[6],
David R. Luukkonen[7], Michael J. Monfils[8], William P. Mueller[9], Kathryn A. Williams[4],
Michele Leduc-Lapierre[10], Elise F. Zipkin[1,2]

[1]Department of Integrative Biology, Michigan State University, College of Natural Science,

288 Farm Lane RM 203, East Lansing, MI 48824.

[2]Ecology, Evolutionary Biology, and Behavior Program, Michigan State University, 293

Farm Lane RM 103, East Lansing, MI 48824.

[3]School of Environmental and Forest Science, University of Washington, Seattle, WA 98195.

[4]Biodiversity Research Institute, 276 Canco Road, Portland, ME 04103.

[5]Point Blue Conservation Science, 3820 Cypress Drive Suite 11, Petaluma, CA 94954.

[6]U.S. Geological Survey, Upper Midwest Environmental Sciences Center, 2630 Fanta Reed

Road, La Crosse WI, 54603.

[7]Department of Fisheries and Wildlife, Michigan State University, 5 Natural Resources Building, East Lansing, MI 48824.

[8]Michigan Natural Features Inventory, Michigan State University Extension, P.O. Box 13036, Lansing, MI 48901.

[9]Western Great Lakes Bird and Bat Observatory, 4970 Country Club Road, Port Washington, WI 53074.

[10]Great Lakes Commission, 2805 S. Industrial Highway Suite 100, Ann Arbor, MI 48104.

*CORRESPONDENCE AUTHOR. Email: sussman7@msu.edu Phone: 301-502-4244.

ABSTRACT

1. Hotspot analysis is a commonly used method in ecology and conservation to identify areas of high biodiversity or conservation concern. However, delineating and mapping hotspots is subjective and various approaches can lead to different conclusions with regard to the classification of particular areas as hotspots, complicating long-term conservation planning.

2. We present a comparative analysis of recent approaches for identifying waterbird hotspots, with the goal of developing insights about the appropriate use of these methods. We selected four commonly used measures to identify persistent areas of high use: kernel density estimation, Getis-Ord G_i*, hotspot persistence, and hotspots conditional on presence, which represent the range of quantitative hotspot estimation approaches used in waterbird analyses. We applied each of the methods to aerial survey waterbird count data collected in the Great Lakes from 2012-2014. For each approach, we identified areas of high use for seven species/species groups and then compared the results across all methods and with mean effort-corrected counts.

3. Our results indicate that formal hotspot analysis frameworks do not always lead to the same conclusions. The kernel density and Getis-Ord G_i^* methods yielded the most similar results across all species analyzed and were generally correlated with mean effort-corrected count data. We found that these two models can differ substantially from the hotspot persistence and hotspots conditional on presence estimation approaches, which were not consistently similar to one another. The hotspot persistence approach differed most significantly from the other methods but is the only method to explicitly account for temporal variation.

4. We recommend considering the ecological question and scale of conservation or management activities prior to designing survey methodologies. Deciding the appropriate definition and scale for analysis is critical for interpretation of hotspot analysis results as is inclusion of important covariates. Combining hotspot analysis methods using an integrative approach, either within a single analysis or post-hoc, could lead to greater consistency in the identification of waterbird hotspots.

KEYWORDS

Gamma distribution, Getis-Ord G_i^*, Great Lakes, Kernel density estimation, Lognormal distribution, Parametric and nonparametric models, Persistence, Spatial models, Spatial statistics

INTRODUCTION

Hotspots are defined as small geographic areas (within a predefined larger region) that exhibit persistent high concentrations of individuals or species (Harcourt 1999, Possingham and Wilson 2005). In the three decades since Myers (1988) first introduced the term, the definition of hotspots has expanded and adapted to reflect changes in conservation goals

(Briscoe et al. 2016). Animal hotspots are generally defined as areas with high levels of at least one of the following biological measures: species abundance, richness, or endemism; rare, threatened, or endangered species; and/or taxonomic distinctiveness (Possingham and Wilson 2005, Briscoe et al. 2016). Hotspots are typically designated on a case-by-case basis because patterns vary by species and location; thus, the threshold to differentiate between hotspots and other locations naturally varies (Nelson and Boots 2008). For example, common definitions of hotspots focus on determining areas with consistent high species abundance (Piatt et al. 2006, Davoren 2007), richness, or biological activity (Sydeman et al. 2006) or some combination of these (Nur et al. 2011). Hotspots have also been defined as locations where some metric exceeds a predefined threshold, such as the top five percent of the data (Harvey et al. 2017) or locations outside one (Suryan et al. 2012, Santora and Veit 2013) or three (Zipkin et al. 2015) standard deviations above the mean of a particular region or area sampled. Such definitions attempt to quantify hotspots (allowing for direct location comparison) as opposed to identifying hotspots using only qualitative criteria, which was common until recently (Mittermeier et al. 2011). The different approaches to identify hotspots have become increasingly complex and may lead to dissimilar or inconsistent results (Prendergast et al. 1993a, Araujo 2002, Orme et al. 2005, Possingham and Wilson 2005, Hobday and Pecl 2014, Daru et al. 2015, Harvey et al. 2017). The consequences of applying different metrics to define hotspots is a lack of congruence across measures and, thus, hotspot locations, culminating in controversy and conflict over long-term conservation efforts (Orme et al. 2005, Possingham and Wilson 2005, Marchese 2015). Such controversies could potentially be avoided through a transparent statement of objectives at the start of one's analysis accompanied by a thorough explanation for selecting a specific method or metric.

Waterbird species display extreme variability in habitat use over both space and time (Certain et al. 2007, Piatt et al. 2007, Votier et al. 2008). They often exhibit large, patchy aggregations offshore, making it difficult to measure their spatial distributions. As such, the foremost method to determine patterns of waterbird species is to identify locations of persistent aggregation or high use, such as hotspots. Hotspot identification is useful in studies of highly mobile organisms, such as waterbirds, because the likelihood that a survey event of any given location is representative of the true abundance at that location is low due to the extreme variability of their distributions (Santora and Veit 2013). There are many methods to examine the diversity and abundance patterns of open water populations, but locating persistent high-use areas is a frequent first step towards understanding the processes that generate spatial patterns of species distributions and informing effective conservation action (Nelson and Boots 2008).

For waterbird abundance data, hotspot analyses are typically conducted using one of the following approaches: 1) qualitative analyses (e.g., through mapping abundance); 2) spatial statistics; or 3) classic statistical modeling (parametric or non-parametric) techniques (Tremblay et al. 2009). Historically, areas of high density or concentration were displayed and compared visually, and mapping relative species abundances remains a prevalent conservation tool (Tremblay et al. 2009, Harvey et al. 2017). Yet, qualitative approaches are limited because they often do not reflect temporal changes (i.e., they are simply a snapshot in time), cannot adequately account for aggregations, and can be misleading based on how data are collected, classified, and presented (Marchese 2015). As a result, more rigorous quantitative approaches, typically in the form of spatial statistics or generalized linear models (GLMs), were developed. Spatial statistics use data collected in both focal and surrounding locations to identify areas of high use. As their name suggests, these techniques account for spatial patterns in the data (Harvey et al. 2017). In contrast, classic statistical modeling

techniques, which use GLM-based frameworks, do not typically consider explicit spatial correlation in hotspot identification and instead require the use of statistical distributions and a metric or threshold to account for variations in abundance patterns (Oppel et al. 2012, Santora and Veit 2013, Zipkin et al. 2015). Waterbirds are highly mobile and tend to aggregate in large groups, resulting in highly skewed data with many absences in space and over time. As such, selecting an appropriately skewed statistical distribution to model waterbird data is fundamental to accurately identifying hotspots using statistical modeling approaches (Zipkin et al. 2014).

In this study, we evaluate four quantitative methods to identify waterbird hotspots using data collected in the Great Lakes: *kernel density estimation*, *Getis-Ord G_i**, *hotspot persistence*, and *hotspots conditional on presence*. We selected these four techniques because they are commonly used and represent the range of quantitative hotspot estimation approaches that have been employed in waterbird analyses, incorporating spatially-explicit processes to varying degrees. *Kernel density estimation* is perhaps the most well-known and widely used spatial method for identifying hotspots. *Kernel density estimation* is an interpolation technique that is used to estimate the probability density function of a variable of interest (e.g., abundance) to identify areas of high density (Wilson et al. 2009, O'Brien et al. 2012, Suryan et al. 2012, Wong et al. 2014). A less common spatial statistic for detecting hotspots is the *Getis-Ord G_i** statistic (G_i*), which allows for cluster evaluation within a specified distance of a single point but does not smooth over grid cells (Getis and Ord 1992, Santora et al. 2010, Kuletz et al. 2015). G_i* analysis is a spatial tool that identifies spatially explicit areas with values higher in magnitude than would be expected due to random chance, independent of the magnitude of abundance (Santora et al. 2010, Kuletz et al. 2015). For the other two models, we adapted GLM-based techniques which have been used to identify waterbird hotspot locations. *Hotspot persistence* defines hotspots for every unique sampling

event and calculates persistence over time (Suryan et al. 2012, Santora and Veit 2013, Johnson et al. 2015). *Hotspots conditional on presence* combines survey data from all sampling events and defines hotspots as locations with a long-term average abundance greater than three times the regional mean, conditional on species presence (Kinlan et al. 2012, Zipkin et al. 2015). Our objective was to compare consistency across methodologically different, yet commonly used, hotspot analysis techniques for several waterbird species and species groups. Other methods for identifying hotspots exist and may be useful, but we restricted our analysis to techniques that have been used for waterbird analyses. We applied the four hotspot methods to the species data and then performed pairwise correlations with each method and to the mean effort-corrected count data to measure the strength and association between the different approaches. This allowed us to quantify the degree to which the various estimators aligned in their assessments of waterbird hotspots. The results of our analyses can provide insights for more objective and goal-driven hotspot delineation to inform species conservation and research priorities.

METHODS

Study area & data description

　　　　We conducted systematic aerial transect surveys of waterbirds in portions of Lakes Erie, Huron, and Michigan, as well as Lake St. Clair during fall, winter, and spring seasons from late September 2012 through early June 2014 (Fig. 1a, Appendix S1). The data used in our analysis were collected as part of ongoing long-term survey efforts to monitor waterbirds in the Great Lakes. Although not collected for the explicit purpose of comparing hotspot identification techniques, these data provide an excellent case study because of the geographic scope and number of species observed. Transects ranged in length from 3 to 177 km covering approximately 8,000 km within the entire study area. Most of the transects

(97%) were surveyed repeatedly with an average number of 10.68 (SD: 3.85) sampling events per transect with approximately 83,000 km flown over the duration of the survey period. We defined a sampling event as a unique year-month-day (survey date) combination within each region of the Great Lakes. Transects were spaced 3.2 to 5 km apart and flight altitude ranged from 61 to 100 m above the lake's surface. Two observers, one on either side of the plane, recorded every waterbird flock that was detected in the observable portion of the transect (the area not obscured by the plane); observers were not permitted to use binoculars, and thus all detections were made using the naked eye. For each sighting, we recorded the species, flock size (i.e., number of individuals seen), and latitude and longitude (using onboard GPS) on the transect line. For large flocks that covered many square kilometers (i.e., up to 30 km), the location recorded was an approximate to the center of the flock. Birds were identified to the lowest taxonomic group possible when observers were unable to determine species.

We integrated the data into the open access Midwest Avian Data Center (MWADC), a regional node of the Avian Knowledge Network (AKN), hosted by Point Blue Conservation Science (http://data.pointblue.org/partners/mwadc/). In some cases, observational data (e.g., location, species, and flock size) were collected separately from effort data (e.g., survey date, location, transect flown, etc.) such that waterbird observations did not include the corresponding transect attribution (23.4% of records). For these records, we used a 1 m buffer to identify the closest transect, matched by survey date, to each observation record. We used data collected on the location of the transect line when available (42.26% of transect lines) and GIS to reconstruct the transect lines from observations in instances when that information was not recorded. Inclement weather and extensive ice coverage necessitated some surveys to be halted prematurely or conducted over a short period of time. We thus assumed that instances in which an area was surveyed over multiple consecutive days were a single

sampling event. During the survey period, 253 transects were surveyed resulting in 136 unique sampling events.

Species groups and composition

We recorded over two million individual birds representing 76 unique species in 41,803 observations. We focused our analysis on seven species/species groups: long-tailed duck *Clangula hyemalis*, common loon *Gavia immer*, gulls *Laridae sp.*, mergansers *Mergus & Lophodytes sp.*, scaup *Aythya affinis* and *A. marila*, loons *Gavia sp.*, and diving/sea ducks *Aythya sp., Bucephala sp., C. hyemlais, Melanitta sp., Mergus sp., Oxyura jamaicensis*, and *Somateria sp.* (Table 1). We chose these species and species groups (hereafter referred to as species groups) because they were fairly evenly distributed across the study area (i.e., occurred in most lakes with data), with observations in at least 200 grid cells (see Data Standardization in Methods), and were encountered at least 1,000 times during the survey period. The seven species groups used in our analysis comprised 33 species (Table 1) and nearly 90% of all observed birds, including some individuals that could not be identified to species (which were only used in multi-species group analyses). Canvasback was the most abundant bird species observed (i.e., most individuals), but long-tailed duck was encountered most often (Table 1).

We identified potential hotspots for the seven species groups, and then used the data from all species groups to analyze hotspots for an all-species-combined group. Some individual species appeared in multiple groups; for example, long-tailed duck was analyzed individually and in the diving/sea duck group. In such instances, an individual species was used only once in the all-species-combined group (i.e., not double counted). Additionally, to maximize data availability, we included some species that did not fall within an obvious or distinct species group. For example, ruddy duck is included in the diving/sea duck group

(although it is not a diving or sea duck) because it is often found with other more typical open water species which dominate our surveys on the Great Lakes.

Data standardization

We imposed a 5 km x 5 km grid (consisting of 17,746 cells) over the entire Great Lakes region and assigned transects to grid cells based on their spatial locations. We chose this scale because the maximum distance between survey transects was 5 km. Thus, a smaller grid would create a very patchy system of survey effort with many empty cells, whereas a larger size would lump together more data, obscuring fine scale aggregations of species. We segmented all transects using the grid system so that grid cells contained only the portion of the transect that occurred within the cell, such that a cell could contain anywhere from zero to many transect segments. Then, we calculated the total length of the transect segments(s) within a grid cell to determine the number of kilometers flown for all sampling events within each cell. We included in our analysis only those cells which contained a transect of at least 1 km in length. The sum of transect lengths within cells ranged from 1 to 16.02 km with an average of 4.69 km (SD: 2.16 km). A total of 1,699 of the 1,767 grid cells included in our analysis had bird observations on at least one sampling occasion.

We standardized the observation data within grid cells (Johnson et al. 2015) because the survey effort was unequal and highly variable across cells within the study area (Fig. 1a). We divided the number of observations of a species for the sampling event-grid cell combination by the summed transect length, resulting in a continuous effort-corrected count (Fig. 1b). We used data from all sampling events in our hotspot analysis; however, we limited the method comparisons, correlations, and hotspot maps to grid cells that contained at least four sampling events, for a total of 1,473 grid cells (83.4% of surveyed cells). Using grid cells with at least four sampling events allowed for analysis across the study region, while

minimizing the chances of false hotspot identification due to insufficient data (Fig. 1a; Kuletz et al. 2015, Zipkin et al. 2015). To calculate a mean effort-corrected count for each species group in surveyed grid cells, we divided the summed effort-corrected counts in a grid cell (across all sampling events) by the total number of sampling events for that grid cell.

Hotspot analysis

Kernel density estimation

  Kernel density estimation is a common method for estimating relative density in animal populations that aggregate and has been used repeatedly to identify waterbird hotspot locations and marine areas in need of protection (Wilson et al. 2009, O'Brien et al. 2012, Suryan et al. 2012, and Wong et al. 2014). This modeling approach converts point data (i.e., effort-corrected counts) into a continuous surface grid reflecting relative densities across all grid cells, where the resulting density of each grid cell is weighted according to the distance from the focal location/grid cell (Wong et al. 2014).

  To implement the *kernel density* method, we calculated the midpoint of each grid cell and assigned all effort-corrected counts of the species group (across all sampling events) to the midpoint of the grid cell in which they occurred. We accounted for uneven sampling effort (grid cells were surveyed from one to thirty times, Fig. 1a) by dividing the summed counts in a grid cell by the total number of sampling events for that grid cell. We used the kernel density tool in the Spatial Analyst extension of ArcGIS 10.3.1 (ESRI 2015) to estimate bird density, inputting values for both bandwidth and cell size. The bandwidth, or size of the neighborhood over which the density is averaged, is the amount of smoothing applied to each kernel (Nelson and Boots 2008). Smoothing allows for abundance prediction in non- (or low-) sampled areas by assuming that neighbors behave more similarly to one another than to locations that are further away. Small bandwidth values may result in fragmented small-scale

kernels, leading to underestimation of hotspots, whereas large values result in oversmoothed general kernels, overestimating hotspots (Wong et al. 2014). We selected a 5 km bandwidth for kernel smoothing based on both the geographic extent of the data and the distance between survey transects (Suryan et al. 2012). *Kernel density estimation* results in a raster, which is a matrix of pixels where each pixel contains a value representing information (e.g., estimated bird abundance; ESRI 2015). The cell size for the output raster can also affect the interpretation of the kernel estimate: large cell sizes may result in a blocky raster that is a poor approximation of a continuous surface, and small cell sizes may result in a raster of many cells that is over-fit or takes an inordinate amount of time to calculate (Beyer 2014). We selected a cell size of 1 km x 1 km for the output raster. For each species group, we extracted the mean expected count from the resulting kernel density raster back to the 5 km^2 grid for comparison with the other methods. Each raster cell was assigned to a corresponding 5 km^2 grid cell based on its midpoint; we then averaged the density values from all raster cells for each grid cell, where the higher the value the more likely it is to be a hotspot.

Getis-Ord G_i*

The *Getis-Ord G_i** (G_i*) statistic detects hotspots while also indicating the statistical significance of those hotspots (Santora et al. 2010, Kuletz et al. 2015). The G_i* technique identifies grid cells whose data points cluster spatially by examining each grid cell within the context of the neighboring cells (Getis and Ord 1992). G_i* differs from *kernel density estimation* because it incorporates the value of each feature in the context of its neighbors, whereas *kernel density* estimates the neighbors based on the focal feature and then applies a smoothing over those neighbors.

To implement the G_i^* statistic, we again calculated the midpoint of each grid cell and assigned all effort-corrected counts of the species group across all sampling events to the midpoint of the cell into which they occurred. We accounted for uneven sampling effort by dividing the summed effort-corrected grid cell counts by the total number of sampling events for that cell. We built a neighbors list for all grid cells using Rook's case contiguity (i.e., grid cells that share a border), and then used the neighbors list to calculate a row-standardized spatial weights matrix (*spdep* package in R; Bivand et al. 2013, Bivand and Piras 2015, R Core Team 2015). The matrix informs every grid cell's relationship to all other cells in the neighborhood (Kuletz et al. 2015). We used the effort-corrected counts and the spatial weights matrix to calculate the G_i^* for each grid cell (*spdep* package in R; Bivand et al. 2013, Bivand and Piras 2015, R Core Team 2015). G_i^* produces a z-score for each grid cell, where high positive values are statistically significant and indicate the possibility of a local cluster of high species abundance (i.e., a hotspot) that is unlikely due to random chance.

Hotspot persistence

The *hotspot persistence* method quantifies the persistence of species counts within individual grid cells (Suryan et al. 2012, Santora and Veit 2013, Johnson et al. 2015). To implement this method, we first fit a gamma distribution to the effort-corrected continuous species group count data, summed within grid cells, for each unique sampling event (*fitdistrplus* package in R; Delignette-Muller and Dutang 2015, R Core Team 2015). We selected the two-parameter gamma distribution (shape and scale) because it can fit a variety of continuous right-skewed data (Dennis and Patil 1983, Bolker 2008) and because it has been used before with this hotspot analysis technique (Johnson et al. 2015). We then assigned a probability to each grid cell based on the fit of the data within the cumulative distribution curve for that sampling event. This allowed us to identify grid cells (for each unique sampling

event) with high abundance of the target species group relative to other cells. Within a unique sampling event, we identified grid cells as hotspots if the value of the cumulative distribution for the cell, based on the fit of the gamma distribution, was above the 75th percentile for that sampling event. After identifying which grid cells were categorized as hotspots for every unique sampling event, we calculated the proportion of sampling events in which a grid cell was identified as a hotspot to examine persistence. The final output was the proportion of sampling events, ranging from zero to one, in which a grid cell was considered a hotspot for the target species group. Values of zero indicate the grid cell was never a hotspot. The higher the proportion, the more frequently the grid cell was considered a hotspot, with a value of one indicating the grid cell was a hotspot for all sampling events.

Hotspots conditional on species presence

The *hotspots conditional on presence* method calculates the long-term probability that a grid cell is a hotspot for a particular species given observed abundances over time (Kinlan et al. 2012, Zipkin et al. 2015). To implement this method, we fit the effort-corrected count data using a lognormal distribution (*fitdistrplus* package in R; Delignette-Muller and Dutang 2015, R Core Team 2015). Because the lognormal distribution does not contain zero in its support, we used only the positive effort-corrected counts. The lognormal is a two parameter (mean and standard deviation), positive, continuous probability distribution characterized by a heavy tail and has been shown to fit waterbird data well because of its flexible shape and ability to fit heavily skewed data (Limpert et al. 2001, Zipkin et al. 2014). We then estimated prevalence in the reference region as the proportion of cells with occurrences for the target species group (at least one individual observed within the cell over all sampling events) relative to the total number of cells surveyed (Kinlan et al. 2012). We defined the reference region as the entire area sampled across the Great Lakes. We then simulated data with a two-

part Monte Carlo approach to calculate hotspot locations using the estimated mean and standard deviation from the lognormal distribution (the count component) and the prevalence estimate (the Bernoulli component; Kinlan et al. 2012, Zipkin et al. 2015). We defined a hotspot as a grid cell in which the long-term average effort-corrected count conditional on presence (with a $\alpha = 0.05$ threshold) was at least three times the mean of the reference region, also conditional on presence (Kinlan et al. 2012). The resulting values, ranging from zero to one, represent the proportion of simulated sample means that are greater than three times the average count. Values close to zero indicate the grid cell is not a hotspot. Values close to one indicate a high probability that the long-term average abundance in the grid cell is greater than three times the mean of the reference region.

Comparative analysis of the methods

Our objective was to determine the degree of congruence among the four methods across species groups and for all-species-combined. We compared methods using only grid cells that were surveyed four or more times. To quantify the consistency among the four approaches in their ability to detect hotspots, we performed a Pearson's product-moment correlation to evaluate the pairwise associations of the four approaches with a Bonferroni adjustment and an alpha level of 0.05 (*Hmisc* package in R; R Core Team 2015, Harrel1 Jr. 2019). In addition to comparing the four hotspot analysis methods, we also compared the methods with the mean effort-corrected count using the same correlation test. We analyzed the correlation coefficients (ranging from zero to one) to determine associations among the different approaches: the higher the value, the higher the correlation between two methods.

We produced maps for all species groups to visually compare the results of the four hotspot analysis approaches and mean effort-corrected counts (Appendices S2-S9). For the first set of maps, we plotted the values produced for each grid cell using each hotspot analysis technique. Direct visual comparison among the hotspot methods can be difficult because the scale of the results for each method varies (i.e., *kernel density* produces unbounded positive values, G_i* produces both positive and negative values, and *hotspot persistence* and *hotspots conditional on presence* range between zero to one). To resolve this issue, we created a second set of maps for each species group in which we considered a hotspot as any grid cell with a value above the 75th percentile (of all values for that method) and plotted those according to their percentiles.

RESULTS

The highest correlation between methods for all species groups occurred between *kernel density* and G_i* estimation approaches with a correlation ≥0.80 for all species groups except mergansers (Table 2). For mergansers, the correlation between the two explicitly spatial methods was 0.67 and nearly identical to the correlation between G_i* and the *hotspots conditional on presence* method. For the other species groups, including all-species-combined, there was much higher congruency between *kernel density* and G_i* than any other combination of pairwise comparisons (excluding mean effort-corrected data, discussed below; Table 2; Appendices S2-S9). For example, for all-species-combined there was 94% overlap in identification of hot and non-hot grid cells between *kernel density estimation* and G_i* (Fig. 1c-d). In general, the two other methods were no more similar to one another than to either *kernel density* and G_i* with correlations between the models that varied by species (Table 2). *Kernel density estimation* and *hotspot persistence* showed the lowest correlations (0.03-0.56) for the species groups that we examined. For the all-species-combined group, we

found that the four methods identified the same 63% of grid cells as non-hot locations (below the 75th percentile), while approximately 8% of grid cells were identified as hotspot locations under all four methods (above the 75th percentile). The remaining 29% of the cells were identified as hotspots by one, two, or three of the methods. Loons and common loon were the only two species groups to have insignificant correlations ($\alpha = 0.05$, Table 2). The patterns observed with these two species groups may be due in part to insufficient data availability.

The *hotspot persistence* approach differed most significantly from the other three methods and had the lowest correlations overall with other methods (Table 2; Fig. 1-2). Unlike the other methods, the *hotspot persistence* approach calculates hotspots relative to the survey region, rather than the entire study area (i.e., Fig. 1a, Fig. 1e), and also explicitly incorporates temporal variability. For example, in the analysis of the scaup species group, we found that many grid cells in Lake St. Clair were identified as hotspots by all methods except for *hotspot persistence* (Fig. 2; Appendix S7). The counts for scaup were generally quite high in Lake St. Clair relative to other surveyed locations. However, the *hotspot persistence* method revealed that individual grid cells within Lake St. Clair did not often have high counts on repeated occasions (as evidenced with zeros and other low values in grid cells; Fig. 2e). The all-species-combined analysis produced similar results, with 9% of grid cells identified as hotspots by the persistence approach but not by the other three methods (Fig. 1b-f; Appendix S2).

Kernel density and G_i^*, which consider space in an explicit manner, are inherently different from *hotspot persistence* and *hotspots conditional on presence*, which rely on parametric distributions, as is evident in both the correlations and maps (Table 2; Appendices S2-S9). *Hotspot persistence* and *hotspots conditional on presence* tended to select single grid cells as hotspots, whereas *kernel density* and G_i^* selected small clusters of grid cells as hotspots (Fig. 3). In many cases, grid cells that were highly ranked with the *hotspots*

conditional on presence approach, were also highly ranked with *kernel density* and G_i^*. However, the surrounding grid cells tended to also be highly ranked with *kernel density* and G_i^* (Fig. 3).

Species-specific hotspots for common loon and long-tailed duck occurred in areas identified as hotspots for the corresponding species group (loons and diving/sea ducks, respectively). Common loon observations comprised 67% of the loon group observations (Table 1). A reanalysis of the loon group without common loons did not substantially alter correlations between the different methods (Appendix S10). Yet, when long-tailed duck data were removed from the diving/sea duck group (9% of the data), G_i^* showed a higher correlation than *kernel density estimation* with *hotspots conditional on presence* (although *kernel density* and G_i^* were still highly correlated with one another; Appendix S10), suggesting that long-tailed ducks may be disproportionately influencing the results of the hotspot analysis for the diving/sea duck group.

Correlations between the mean effort-corrected counts and each of the four methods were highly consistent among the species groups (Table 2). For most species groups, *kernel density estimation* had the highest correlation with mean effort-corrected counts, followed by G_i^* and *hotspots conditional on presence*, with *hotspot persistence* being least similar. Correlations between mean effort-corrected counts and *kernel density estimation* tended to be fairly high; though this was not true for the mergansers group, which had the highest correlation between mean effort-corrected counts and G_i^*. The only species group to demonstrate its highest correlation between mean effort-corrected count and a GLM-based technique (*hotspots conditional on presence*) was common loon. For seven of the eight species groups, mean effort-corrected count had the lowest correlation with the *hotspot persistence* approach (Table 2).

DISCUSSION

Despite the frequent use of hotspot analyses in management and conservation, we found that the various hotspot analysis approaches resulted in inconsistent identification of hotspots within the Great Lakes waterbird dataset, and that the degree of this inconsistency varied by species group. If conservation and management decisions are based on hotspot analyses, it is important to understand the advantages, limitations, and potential biases in the various approaches to facilitate selection of the most appropriate method to answer the question(s) of interest. Our analysis revealed that methods to estimate hotspots using explicitly spatial approaches, specifically *kernel density* and G_i*, produce highly correlated results and can likely be used as surrogates for one another. Not surprisingly, these spatial smoothing methods were often also highly correlated with mean effort-corrected counts. These approaches are likely to be most appropriate when the objective is to identify hotspot locations within specific temporal constraints (e.g., within a single season). Spatial hotspot techniques might also be useful in the creation and delineation of open water sanctuaries or marine preserves, as these areas often encompass large geographic ranges such that the information of spatial neighbors can be helpful. The other two modeling techniques that we examined produce less consistent results with each other, the spatial methods, and mean effort-corrected counts, the degree to which varied by species. The *hotspot persistence* approach differed the most from the other methods. *Hotspot persistence* estimates hotspots based on unique sampling events within survey regions and then identifies whether those areas persist as hotspots over time, while the other three approaches (and mean effort-corrected count data) focus on average abundance over the entire survey period. The *hotspot persistence* approach (perhaps more so than the other approaches) may thus perform best with a high number of sampling events rather than our minimum of four surveys (Kinlan et al. 2012). Studies constrained to a small geographic range would benefit from the *hotspot*

persistence approach because smaller areas are more likely to be surveyed more often resulting in a high sampling frequency. For example, we found that the Lake St. Clair location was a hotspot under most techniques for many species groups, but the *hotspot persistence* approach revealed that no grid cell within Lake St. Clair was more important than any other (Fig. 1e). The *hotspot persistence* approach could thus be useful in the assessment of development activities such as offshore wind energy. Analyzing data over time would allow regulators and energy developers to determine whether there are specific locations within a predefined spatial area with repeatedly and consistently high abundance, which can be used to minimize risk to wildlife populations.

Our analysis cannot determine which hotspot identification method is most appropriate generally. The best method for any given analysis will depend on the question and study region. *Kernel density* and G_i^* are useful in assessing large areas where waterbirds have been present (with high densities) on at least one occasion. Yet *kernel density* and G_i^* can have difficulty identifying individual grid cells as hotspots, especially at small spatial scales and with heavily skewed data (Songchitruska and Zeng 2010, Harris 2017). Both *kernel density estimation* and G_i^* demonstrated high similarity to the mean effort-corrected count data for almost all species groups. This suggests that effort-corrected count data may be a sufficient surrogate for hotspot techniques in some situations, such as coarse summaries, particularly if one does not plan on including covariate data. The other two methods, *hotspot persistence* and *hotspots conditional on presence*, provide fundamentally different hotspot information than a simple summary of mean counts. These techniques are likely to be relatively more conservative in identifying hotspots (e.g., Fig. 3), which may be desirable depending on conservation and management priorities.

There is subjectivity in all methods to identify hotspots because arbitrary thresholds are selected to delineate hotspot locations (Harvey et al. 2017). Thus, there is a natural tradeoff in the rate of Type I and Type II errors in selecting a threshold level (Kinlan et al. 2012). A high threshold will constrain the number of hotspots and the likelihood that an area is falsely categorized as a hotspot is comparatively low. A low threshold will increase the number of hotspots, possibly including locations that should not be considered hotspots. All four methods we selected incorporate at least one user-determined decision or metric, which can be changed to suit the user (Canadas et al. 2014, Marchese 2015) or may be different across various software implementations (Bivand and Wong 2018). The bandwidth and cell size must be specified in *kernel density estimation* and the neighborhood must be defined in the G_i^* method. For the other two models, the user must decide the level that constitutes a hotspot (i.e., top 25% of cells in the *hotspot persistence* approach; three times the mean in the *hotspots conditional on presence* method; Canadas et al. 2014, Zipkin et al. 2015). The production of hotspot maps is also subjective (Carolan 2009). We chose to classify and color hotspots as the top 25% of grid cell values for the four methods (Fig. 1c-f). However, a different threshold could produce maps that appear more or less similar (particularly when evaluating methods on their original scales; e.g., Fig. 2) and possibly lead to different conclusions on the congruency of the four approaches, illustrating the importance of examining quantitative correlations between the methods.

A combination of methods using an integrative approach that synthesizes results, either within a single analysis or post-hoc, could lead to greater consistency (Marchese 2015) and reduce the impacts of survey bias. Nur et al. (2011) found that the use of multiple criteria (when defining a hotspot within a single method) prevents misidentifying areas that may be overlooked when using only a single criterion. An integrated model combining a spatial method with a GLM-based technique could provide a desirable balance between the potential

for overestimation from spatial approaches and the more conservative non-spatial models. As waterbird data are often sparse or limited, an integrated approach will also help to identify hotspots in locations with possible data gaps and could be implemented by averaging standardized values from each model (e.g., Fig. 4). Such an integrated model would be beneficial in scenarios where both spatial coverage and temporal consistency are important.

Inconsistency in sampling required us to exclude data from areas that were sampled less than four times, although it is possible that even areas included in our analysis were not sampled enough for accurate hotspot detection (Hazen et al. 2014, Zipkin et al. 2015). Excluding grid cells with few visits may remove a substantial portion of available data, which is undesirable in any analysis. However, inclusion of low sampled grid cells could lead to false positives in the identification of hotspots, given the highly mobile and aggregated nature of waterbirds (Kuletz et al. 2015). For example, a specific grid cell that was surveyed only one time in our study had an effort-corrected count of almost 900 individuals (all-species-combined). Had we included cells that were surveyed less than four times, this cell would have been designated as a hotspot by all four methods. Yet, it is difficult to assess the validity of this designation without repeated sampling. For some rare species, though, it may be necessary to reduce the threshold of required survey events to maximize data use.

Variation in hotspot identification across the methods may be due in part to scale. Scale is important to several aspects of hotspot identification: 1) the scale at which the data are collected, 2) the spatial scale at which the data are analyzed, and 3) the scale at which management decisions will be made. It is important to account for potential discrepancies in the geographic scale of population-level processes and the resolution of different datasets when deciding the spatial scale for analysis. The 5 km^2 grid we used is a fine-scale resolution to identify and map hotspots and may result in less consistency across methods than a larger, coarser grid (Daru et al. 2015). The distance between and design of transect surveys can

directly affect the outcome of hotspot analyses, and we therefore recommend simultaneous and thorough consideration of survey design and analysis prior to data collection. Survey bias, the presence of spatial patterns in survey effort, is a ubiquitous concern when delineating hotspots (Prendergast et al. 1993b), and its effects may lead to different results depending on the analytical methods used.

The four approaches used in this study were selected because they have been previously used to identify areas of high-use for waterbird species in other studies. However, variations of these models (e.g., selecting a different threshold cutoff or using a metric other than the mean, such as standard deviation) as well as other potential hotspot methods could be used to model abundance data. For example, cluster analysis models that have been developed for monitoring crime rates or traffic accidents and patterns, which have not yet been applied to wildlife data, may provide more nuanced approaches to hotspot analyses (Tango 1995, Hengl 2009). We used R and ArcGIS to run our analysis; however, other software packages may also be useful for hotspot modeling. For instance, SaTScan is free software developed to detect disease clusters by analyzing spatial, temporal, and/or space-time data (Kulldorf 2018). The platform can run different types of models (such as normal, ordinal, and exponential models, as well as those we implement in this paper) and adjust for underlying spatial inhomogeneity as part of the default software features. Continued software and model development will allow for advances in hotspot analysis methods.

Hotspot analysis is a first step in understanding species distribution patterns, but it is often equally or more important to determine why certain areas contain persistent aggregations of waterbirds. The use of mechanistic or associative models with covariates that help explain and elucidate hotspots should be considered in those cases where knowledge of the system and adequate covariate data exist (e.g., Nur et al. 2011). We did not include environmental variables (e.g., bathymetry, surface temperature, ice coverage, etc.) or

seasonality in our hotspot analysis, although they most certainly play an important role in explaining the distribution and abundance patterns of waterbirds (Nur et al. 2011, Suryan et al. 2012). Although excluding covariates does not affect our model comparison results, it precludes understanding why certain locations are identified as hotspots. Some environmental variables, such as habitat suitability and food availability, are critical to discerning species behaviors and patterns and are occasionally used as proxies for hotspots when data are limited (Hyrenbach et al. 2000, Folmer et al. 2010, Shirkey 2012, Briscoe et al. 2016). However, challenges arise with incorporating environmental variables that are dynamic (e.g., ice cover or temperature), making identification of static hotspot locations relative to environmental variables difficult (Briscoe et al. 2016, Marchese 2015) and perhaps less useful for certain management-related questions. Waterbird hotspots may not be fixed locations, and may vary by season, annually, or on even longer time frames. Seasonal variability in waterbird species is an important factor that we did not consider in our analysis; abundances can fluctuate during migration or at overwintering locations and shifts in distributions may occur even within seasons (Suryan et al. 2016).

Survey methods and modeling techniques have improved over time, but waterbird species are highly mobile, making the identification of priority areas difficult (Arcos et al. 2012, Harvey et al. 2017, Marchese 2015). Through our study, we demonstrate that delineating hotspots is often subjective, as different analysis techniques and thresholds can produce varying results. Regardless of their drawbacks, hotspot analyses are likely to remain an important tool for conservation because of their relative ease to implement. Thus, the inconsistencies we found in our comparisons necessitates attention by conservation practitioners. Researchers should clearly identify conservation and management goals to select the most appropriate analysis method(s). This, along with incorporation of covariates, will allow hotspot analyses to be both useful and meaningful. Conservation and management

decisions on hotspots are often long-lasting and should be done carefully and with full understanding of model limitations.

AUTHOR'S CONTRIBUTIONS

ALS and EFZ conceived the ideas, analyzed the data, and led the writing of the manuscript; BG and EMA contributed to the development of the methodology; KPK, DRL, MJM, WPM, and KAW designed the surveys and collected the data; LS combined datasets and provided

data management & GIS support; MLL provided project oversight and management. All authors contributed to the drafts and gave final approval for publication.

DATA ACCESSBILITY

Data are available in the Dryad Digital Repository (doi:10.5061/dryad.rs776p3) and the Midwest Avian Data Center, a regional node of the Avian Knowledge Network, hosted by Point Blue Conservation Science (http://data.pointblue.org/partners/mwadc/). The data are also available on GitHub (https://zipkinlab.github.io/).

REFERENCES

Araujo M.B. 2002. Biodiversity hotspots and zones of ecological transition. Conservation Biology 16(6):1662-1663.

Arcos J.M., Becares J., Villero D., Brotons L., Rodriguez B., and Ruiz A. 2012. Assessing the location and stability of foraging hotspots for pelagic seabirds: An approach to identify marine Important Bird Areas (IBAs) in Spain. Biological Conservation 156:30-42.

Beyer H. 2014. Kernel density estimation *in* Introducing the Geospatial Modelling Environment. Spatial Ecology LLC. Version: 0.7.2 RC2. <www.spatialecology.com/gme>.

Bivand R. and Piras G. 2015. Comparing implementations of estimation methods for spatial econometrics. Journal of Statistical Software 63(18):1-36.

Bivand R.S. and Wong D.W.S. 2018. Comparing implementations of global and local indicators of spatial association. TEST 27:716-748.

Bivand R.S., Hauke J., and Kossowski T. 2013. Computing the Jacobian in Gaussian spatial autoregressive models: An illustrated comparison of available methods. Geographical Analysis 45(2):150-179.

Bolker B. 2008. Chapter 4: Probability distributions *in* Ecological Models and Data in R. Princeton University Press, Princeton, NJ. 508pp.

Briscoe D.K., Maxwell S.M., Kudela R., Crowder L.B., and Croll D. 2016. Are we missing important areas in pelagic conservation? Redefining conservation hotspots in the ocean. Endangered Species Research 29:229-237.

Canadas E.M., Fenu G., Penas J., Lorite J., Mattana E., and Bacchetta G. 2014. Hotspots within hotspots: Endemic plant richness, environmental drivers, and implications for conservation. Biological Conservation 170:282-291.

Carolan M.S. 2009. This is not a biodiversity hotspot: The power of maps and other images in the environmental sciences. Society and Natural Resources 22(3):278-286.

Certain G., Bellier E., Planque B., and Bretagnolle V. 2007. Characterising the temporal variability of the spatial distribution of animals: an application to seabirds at sea. Ecography 30:695-708.

Daru B.H., van der Bank M., and Davies T.J. 2015 Spatial incongruence among hotspots and complementary areas of tree diversity in southern Africa. Biodiversity Research 21:769-780.

Davoren G.K. 2007. Effects of gill-net fishing on marine birds in a biological hotspot in the northwest Atlantic. Conservation Biology 21(4):1032-1045.

Delignette-Muller M.L. and Dutang C. 2015. fitdistrplus: An R package for fitting distributions. Journal of Statistical Software 64(4):1-34. <http://www.jstatsoft.org/v64/i04/>.

Dennis B. and Patil G.P. 1983. The gamma distribution and weighted multimodal gamma distributions as models of population abundance. Mathematical Biosciences 68(2):187-212.

ESRI 2015. ArcGIS Desktop: Release 10.3.1. Redlands, CA: Environmental Systems Research Institute.

Folmer E.O., Olff H., and Piersma T. 2010. How well do food distributions predict spatial distributions of shorebirds with different degrees of self-organization? Journal of Animal Ecology 79:747-756.

Getis A. and Ord J.K. 1992. The analysis of spatial association by use of distance statistics. Geographical Analysis 24(3):189-206.

Great Lakes Fish and Wildlife Restoration Act (GLFWRA) of 2006. S. 2430 - 109th Congress. <https://www.govtrack.us/congress/bills/109/s2430>.

Harcourt A.H. 1999. Coincidence and mismatch of biodiversity hotspots: a global survey for the order, primates. Biological Conservation 93:163-175.

Harrell Jr. F.E, with contributions from Charles Dupont and many others. 2019. Hmisc: Harrell Miscellaneous. R package version 4.2-0. https://CRAN.R-project.org/package=Hmisc.

Harris N.L. 2017. Using spatial statistics to identify emerging hot spots of forest loss. Environmental Research Letters 12:1-13.

Harvey G.K.A., Nelson T.A., Fox C.H., and Paquet P.C. 2017. Quantifying marine mammal hotspots in British Columbia, Canada. Ecosphere 8(7):1-22.

Hazen E.L., Suryan R.M., Santora J.A., Bograd S.J., Watanuki Y., and Wilson R.P. 2013. Scales and mechanisms of marine hotspot formation. Marine Ecology Progress Series 487:177-183.

Hengl T. 2009. A Practical Guide to Geostatistical Mapping.

Hobday A.J. and Pecl G.T. 2014. Identification of global marine hotspots: sentinels for change and vanguards for adaptation action. Reviews in Fish Biology and Fisheries 24:415-425.

Hyrenbach K.D., Forney K.A., and Dayton P.K. 2000. Marine protected areas and ocean basin management. Aquatic Conservation: Marine and Freshwater Ecosystems 10:437-458.

Johnson S.M., Connelly E.E., Williams K.A., Adams E.M., Stenhouse I.J., and Gilbert A.T. 2015. Integrating data across survey methods to identify spatial and temporal patterns in wildlife distributions *in*: Wildlife densities and habitat use across temporal and spatial scales on the Mid- Atlantic Outer Continental Shelf: Final report to the Department of Energy Efficiency and Renewable Energy Wind & Water Power Technologies Office. Williams K.A., Connelly E.E., Johnson S.M., and Stenhouse I.J. (eds.) Award Number: DE-EE0005362. Report BRI 2015-11, Biodiversity Research Institute, Portland, Maine. 56 pp.

Kinlan B.P., Zipkin E.F., O'Connell A.F., and Caldow C. 2012. Statistical analyses to support guidelines for marine avian sampling: final report. U.S. Department of the Interior, Bureau of Ocean Energy Management, Office of Renewable Energy Programs, Herndon, VA. OCS Study BOEM 2012-101. NOAA Technical Memorandum NOS NCCOS 158. xiv+77 pp.

Kuletz K.J., Ferguson M.C., Hurley B., Gall A.E., Labunski E.A., and Morgan T.C. 2015. Seasonal spatial patterns in seabird and marine mammal distribution in the eastern Chukchi and western Beaufort seas: Identifying biologically important areas. Progress in Oceanography 136:175-200.

Kulldorf M. 2018. SaTScan User Guide for version 9.6.

Limpert E., Stahel W.A., and Abbt M. 2001. Log-normal distributions across the sciences: keys and clues. BioScience 51(5):341-352.

Marchese C. 2015. Biodiversity hotspots: A shortcut for a more complicated concept. Global Ecology and Conservation 3:297-309.

Mittermeier R.A., Turner W.R., Larsen F.W., Brooks T.M., and Gascon C. 2011. Global biodiversity conservation: The critical role of hotspots *in* Biodiversity Hotspots. pp 3-22.

Myers N. 1988. Threatened biotas: "Hot spots" in tropical forests. The Environmentalist 8(3):187-208.

Nelson T.A. and Boots B. 2008. Detecting spatial hot spots in landscape ecology. Ecography 31:556-566.

Nur N., Jahncke J., Herzog M.P., Howar J., Hyrenbach K.D., Zamon J.E., Ainley D.G., Wiens J.A., Morgan K., Balance L.T., and Stralberg D. 2011. Where the wild things are: predicting hotspots of seabird aggregations in the California Current System. Ecological Applications 21: 2241-2257.

O'Brien S.H., Webb A., Brewer M.J., and Reid J.B. 2012. Use of kernel density estimation and curvature to set Marine Protected Area boundaries: Identifying a Special Protection Area for wintering red-throated divers in the UK. Biological Conservation 156:15-21.

Oppel S., Meirinho A., Ramírez I., Gardner B., O'Connell A.F., Miller P.I., and Louzao M. 2012. Comparison of five modeling techniques to predict the spatial distribution of seabirds. Biological Conservation 156:94-104.

Orme C.D.L., Davies R.G., Burgess M., Eigenbrod F., Pickup N., Olson V.A., Webster A.J., Ding T., Rasmussen P.C., Ridgley R.S., Stattersfield A.J., Bennet P.M., Blackburn

T.M., Gaston K.J., and Owens I.P.F. 2005. Global hotspots of species richness are not congruent with endemism or threat. Nature 436(18):1016-1019.

Piatt J.F., Wetzel J., Bell K., DeGange A.R., Balogh G.R., Drew G.S., Geernaert T., Ladd C., and Byrd G.V. 2006. Predictable hotspots and foraging habitat of the endangered short-tailed albatross (*Phoebastria albatrus*) in the North Pacific: implications for conservation. Deep-Sea Research II 53:387-398.

Piatt J.F., Sydeman W.J., and Wiese F. 2007. Introduction: a modern role for seabirds as indicators. Marine Ecology Progress Series 352:199-204.

Possingham H.P. and Wilson K.A. 2005. Turning up the heat on hotspots. Nature 436(18):919-920.

Prendergast J.R., Quinn R.M., Lawton J.H., Eversham B.C., and Gibbons D.W. 1993a. Rare species, the coincidence of diversity hotspots and conservation strategies. Nature 365:335-337.

Prendergast J.R., Wood S.N., Lawton J.H., and Eversham B.C. 1993b. Correcting for variation in recording effort in analyses of diversity hotspots. Biodiversity Letters 1(2):39-53.

R Core Team 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

Santora J.A., Reiss C.S., Loeb V.J., and Veit R.R. 2010. Spatial association between hotspots of baleen whales and demographic patterns of Antarctic krill *Euphasia superba* suggests size-dependent predation. Marine Ecology Progress Series 405:255-269.

Santora J.A. and Veit R.R. 2013. Spatio-temporal persistence of top predator hotspots near the Antarctic Peninsula. Marine Ecology Progress Series 487:287-304.

Shirkey B.T. 2012. Diving duck abundance and distribution on Lake St. Clair and western Lake Erie. M.S. Thesis, Michigan State University, East Lansing.

Songchitruska P. and Zeng X. 2010. Getis-Ord spatial statistics to identify hot spots by using incident management data. Journal of the Transportation Research Board 2165:42-51.

Suryan R.M., Santora J.A., and Veit R.R. 2012. New approach for using remotely sensed chlorophyll a to identify seabird hotspots. Marine Ecology Progress Series 451:213-225.

Suryan R.M., Kuletz K.J., Parker-Stetter S.L., Ressler P.H., Renner M., Horne J.K., Farley E.V., and Labunski E.A. 2016. Temporal shifts in seabird populations and spatial coherence with prey in the southeastern Bering Sea. Marine Ecology Progress Series 549:199-215.

Sydeman W.J., Brodeur R.D., Grimes C.B., Bychkov A.S., and McKinnell S. 2006. Marine habitat 'hotspots' and their use by migratory species and top predators in the North Pacific Ocean: introduction. Deep Sea Research II 53:247-249.

Tango T. 1995. A class of tests for detecting 'general' and 'focused' clustering of rare diseases. Statistics in Medicine 14(21-22):2323-2334.

Tremblay Y., Bertrand S., Henry R.W., Kappes M.A., Costa D.P., and Shaffer S.A. 2009. Analytical approaches to investigating seabird-environment interactions: a review. Marine Ecology Progress Series 391:153-164.

Votier S.C., Birkhead T.R., Oro D., Trinder M., Granthan M.J., Clark J.A., McCleery R.H., and Hatchwell B.J. 2008. Recruitment and survival of immature seabirds in relation to oil spills and climate variability. Journal of Animal Ecology 77:974-983.

Ward E.J., Marshall K.N., Ross T., Sedgley A., Hass T., Pearson S.F., Joyce G., Hamel N.J., Hodum P.J., and Faucett R. 2015. Using citizen-science data to identify local hotspots of seabird occurrence. PeerJ 3:e704.

Wilson L.J., McSorley C.A., Gray C.M., Dean B.J., Dunn T.E., Webb A., and Reid J.B.

2009. Radio-telemetry as a tool to define protected areas for seabirds in the marine

environment. Biological Conservation 142:1808-1817.

Wong S.N.P., Gjerdrum C., Morgan K.H., and Mallory M.L. 2014. Hotspots in cold seas:

The composition, distribution, and abundance of marine birds in the North American

Arctic. Journal of Geophysical Research: Oceans 119:1691-1705.

Zipkin E.F., Kinlan B.P., Sussman A., Rypkema D., Wimer M., and O'Connell A.F. 2015.

Statistical guidelines for assessing marine avian hotspots and coldspots: A case study

on wind energy development in the U.S. Atlantic Ocean. Biological Conservation

191: 216-223.

Zipkin E.F., Leirness J.B., Kinlan B.P., O'Connell A.F., and Silverman E.D. 2014. Fitting
statistical distributions to sea duck count data: implications for survey design and abundance
estimation. Statistical Methodology 17:67-81.

Table 1. List of species and species groups used in the analysis, including the total number of encounters and total number observed.

Species/Species Group	Number of Encounters	Number Observed	Species Included
Diving/sea ducks	19,183	1,700,311	Bufflehead *Bucephala albeola*, canvasback *Aythya valisineria*, common eider *Somateria mollissima*, long-tailed duck *Clangula hyemalis*, redhead *Aythya americana*, ring-necked duck *Aythya collaris*, ruddy duck *Oxyura jamaicensis*, all eiders *Somateria sp*, all goldeneye *Bucephala sp*, all mergansers *Mergus sp*, all scaup *Aythya sp*, all scoters *Melanitta sp*, and all unidentified diving ducks *Aythya sp*.
Gulls	12,233	81,399	Bonaparte's gull *Chroicocephalus philadelphia*, glaucous gull *Larus hyperboreus*, great black-backed gull *Larus marinus*, herring gull *Larus smithsonianus*, Iceland gull *Larus glaucoides*, mew gull *Larus canus*, ring-billed gull *Larus delawarensis*, and all unidentified gulls *Laridae sp*.
Long-tailed duck	6,011	149,542	Long-tailed duck *Clangula hyemalis*
Mergansers	4,865	95,702	Common merganser *Mergus merganser*, hooded merganser *Lophodytes cucullatus*, red-breasted merganser *Mergus serrator*, all unidentified mergansers *Mergus sp*, and all unidentified merganser/goldeneye *Mergus/Bucephala sp*.
Scaup	3,431	383,495	Greater scaup *Aythya marila*, lesser scaup *Aythya affinis*, and all unidentified scaup *Aythya sp*.
Loons	2,111	4,364	Common loon *Gavia immer*, red-throated loon *Gavia stellata*, and all unidentified loons *Gavia sp*.
Common loon	1,688	2,922	Common loon *Gavia immer*

Table 2. Pearson correlation matrix of pairwise comparisons between the four hotspot analysis methods (*kernel density estimation*, *Getis-Ord G_i**, *hotspot persistence*, and *hotspots conditional on presence*) and mean effort-corrected counts, with a Bonferroni adjustment. Correlations are significant unless otherwise denoted (**) at an alpha level of 0.05. Values above and below of the diagonals are mirror images (gray values are duplicates).

	Kernel density estimation	Getis-Ord G_i*	Hotspot persistence	Hotspots conditional on presence	Mean effort-corrected count
All species/groups					
Kernel density estimation	1.000	0.870	0.121	0.441	0.928
Getis-Ord G_i*	0.870	1.000	0.125	0.435	0.650
Hotspot persistence	0.121	0.125	1.000	0.332	0.093
Hotspots conditional on presence	0.441	0.435	0.332	1.000	0.375
Mean effort-corrected Count	0.928	0.650	0.093	0.375	1.000
Diving/sea ducks					
Kernel density estimation	1.000	0.874	0.160	0.467	0.929
Getis-Ord G_i*	0.874	1.000	0.167	0.458	0.655
Hotspot persistence	0.160	0.167	1.000	0.364	0.123
Hotspots conditional on presence	0.467	0.458	0.364	1.000	0.395
Mean effort-corrected Count	0.929	0.655	0.123	0.395	1.000
Gulls					
Kernel density estimation	1.000	0.808	0.318	0.457	0.865
Getis-Ord G_i*	0.808	1.000	0.325	0.459	0.715
Hotspot persistence	0.318	0.325	1.000	0.446	0.274
Hotspots conditional on presence	0.457	0.459	0.446	1.000	0.443
Mean effort-corrected Count	0.865	0.715	0.274	0.443	1.000
Long-tailed duck (LTDU)					
Kernel density estimation	1.000	0.804	0.407	0.475	0.725
Getis-Ord G_i*	0.804	1.000	0.451	0.432	0.469
Hotspot persistence	0.407	0.451	1.000	0.583	0.350

Hotspots conditional on presence	0.475	0.432	0.583	1.000	0.495
Mean effort-corrected Count	0.725	0.469	0.350	0.495	1.000
Mergansers					
Kernel density estimation	1.000	0.672	0.433	0.600	0.739
Getis-Ord G_i*	0.672	1.000	0.547	0.669	0.812
Hotspot persistence	0.433	0.547	1.000	0.575	0.539
Hotspots conditional on presence	0.600	0.669	0.575	1.000	0.778
Mean effort-corrected Count	0.739	0.812	0.539	0.778	1.000
Scaup					
Kernel density estimation	1.000	0.878	0.562	0.661	0.950
Getis-Ord G_i*	0.878	1.000	0.586	0.623	0.707
Hotspot persistence	0.562	0.586	1.000	0.686	0.469
Hotspots conditional on presence	0.661	0.623	0.686	1.000	0.611
Mean effort-corrected Count	0.950	0.707	0.469	0.611	1.000
Loons					
Kernel density estimation	1.000	0.808	0.047**	0.210	0.910
Getis-Ord G_i*	0.808	1.000	0.063	0.258	0.517
Hotspot persistence	0.047**	0.063	1.000	0.454	0.026**
Hotspots conditional on presence	0.210	0.258	0.454	1.000	0.136
Mean effort-corrected count	0.910	0.517	0.026**	0.136	1.000
Common loon (COLO)					
Kernel density estimation	1.000	0.800	0.027**	0.049**	0.060
Getis-Ord G_i*	0.800	1.000	0.032**	0.075	0.087
Hotspot persistence	0.027**	0.032**	1.000	0.606	0.536
Hotspots conditional on presence	0.049**	0.075	0.606	1.000	0.718
Mean effort-corrected count	0.060	0.087	0.536	0.718	1.000

Figure 1. (a) Map of the study area showing the number of sampling events per 5 km x 5 km grid cell during the entire survey period. (b) Mean effort-corrected count data for the all-species-combined species group (species list found in Table 1). Potential hotspots (values above the 75% percentile) across all sampled locations for the all-species-combined group as estimated with each of the four hotspot analysis approaches: (c) *kernel density estimation*, (d) *Getis-Ord G_i**, (e) *hotspot persistence*, and (f) *hotspots conditional on presence*. The 'Hotspots' legend on panel (c) is used for all four hotspot analysis approach maps (c-f). Grid cells sampled less than four times were excluded from the analysis and are shaded in gray. Note the survey regions are delineated for the *hotspot persistence* approach (e) because hotspots in this method are calculated relative to other grid cells within these specific regions.

Figure 2. Hotspot maps for the scaup species group (greater scaup *Aythya marila* pictured in top left, panel a) in western Lake Erie, including (b) mean effort-corrected counts. The hotspot values are shown on the raw scales for each of the four methods: (c) *kernel density estimation*, (d) *Getis-Ord G_i**, (e) *hotspot persistence*, and (f) *hotspots conditional on presence*.

Figure 3. Potential hotspots (values above the 75% percentile) in a portion of Lake Michigan for the diving/sea duck species group, highlighting (a) single cell selection in *hotspots conditional on presence* method (a GLM-based technique) and (b) clustered cell selection in *kernel density estimation* (a spatial statistic approach).

Figure 4. Potential hotspots from an integrated hotspot modeling approach across all sampled locations for the all-species-combined species group (species list found in Table 1). We combined two hotspot analysis techniques, one spatial statistic approach (*Getis-Ord G_i**) and one GLM-based technique (*hotspots conditional on presence*). The results shown in this map were produced by calculating the average percentile of the two selected methods for all grid cells, resulting in a value from zero to one where high values closer to one are more likely to be hotspots than low values closer to zero. Grid cells sampled less than four times were excluded from the analysis and are shaded in gray. We binned the percentiles for mapping purposes and consider a species hotspot as any grid cell with an average value above the 75th percentile.
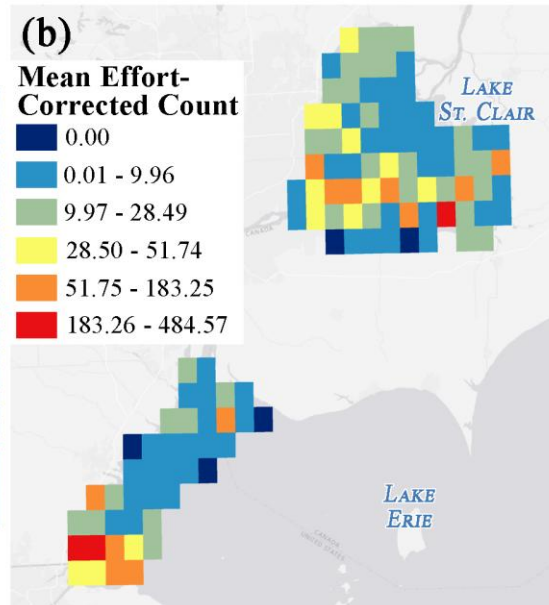
(a) **Number of Sampling Events**
- 1 - 3
- 4 - 6
- 7 - 10
- 11 - 16
- 17 - 20
- 21 - 30

Survey Effort

(b) **Mean Effort-Corrected Count**
- 0.00
- 0.01 - 29.83
- 29.84 - 136.90
- 136.91 - 399.06
- 399.07 - 1002.27
- 1002.28 - 3917.29
- sampled less than 4 times

Mean Effort-Corrected Count

(c) **Hotspots**
- < 75th
- 75 - 90th
- 90 - 95th
- 95 - 99th
- > 99th
- sampled less than 4 times

Kernel Density Estimation

(d) *Getis-Ord G_i^**

(e) **Research Entity**
- Biodiversity Research Institute
- Michigan Department of Natural Resources
- Michigan Natural Features Inventory
- U.S. Geological Survey
- Western Great Lakes Bird & Bat Observatory
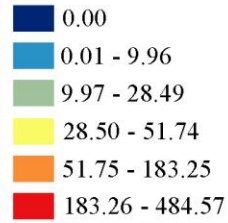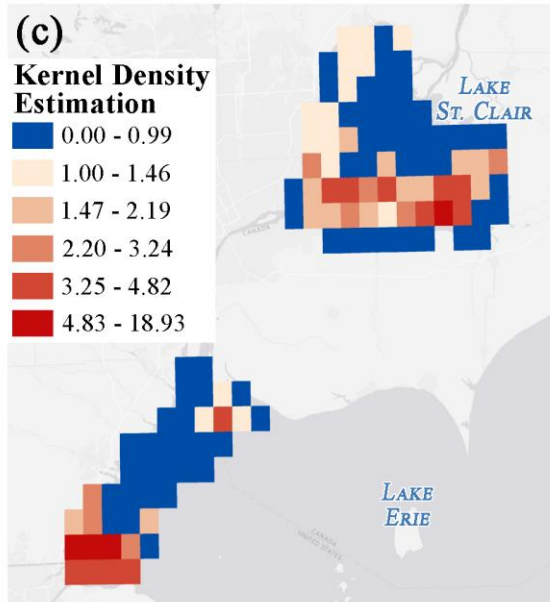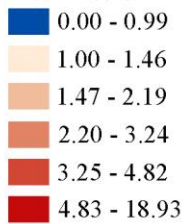
Hotspot Persistence

(f) *Hotspots Conditional on Presence*

(a)

(b) Mean Effort-Corrected Count
- 0.00
- 0.01 - 9.96
- 9.97 - 28.49
- 28.50 - 51.74
- 51.75 - 183.25
- 183.26 - 484.57

LAKE ST. CLAIR

LAKE ERIE

(c) Kernel Density Estimation
- 0.00 - 0.99
- 1.00 - 1.46
- 1.47 - 2.19
- 2.20 - 3.24
- 3.25 - 4.82
- 4.83 - 18.93

LAKE ST. CLAIR

LAKE ERIE

(d) Getis-Ord G_i^*
- -1.65 - 1.65
- 1.66 - 1.96
- 1.97 - 2.56
- > 2.56

LAKE ST. CLAIR

LAKE ERIE

(e) Hotspot Persistence
- 0.00 - 0.19
- 0.20 - 0.59
- 0.61 - 0.70
- 0.71 - 0.80
- 0.81 - 0.90
- 0.91 - 1.00

LAKE ST. CLAIR

LAKE ERIE

(f) Conditional on Presence
- 0.00 - 0.19
- 0.20 - 0.59
- 0.61 - 0.70
- 0.71 - 0.80
- 0.81 - 0.90
- 0.91 - 1.00

LAKE ST. CLAIR

LAKE ERIE

N

0 5 10   20 Kilometers