Proceedings of the Fourth International
Partners in Flight Conference: Tundra to Tropics
292–302

MAPPING DATA-SPECIFIC AVIAN DISTRIBUTIONS UTILIZING CARTOGRAPHIC POINT AGGREGATING GENERALIZATION OPERATORS IN A MULTI-SCALE FRAMEWORK

M. Thomas Auer [1]

Department of Geography, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

Abstract. In mapping range distributions of avian species, there has been an historic trend to manually draw boundaries utilizing varied resources, including expert knowledge and anecdotal evidence. This older technique is in opposition to utilizing one unified dataset, which until recently has been largely unavailable for use. The efforts detailed here move known mapping methods of distribution ranges from arbitrary to actual, using only existing data sources. A related goal is to map the extent of migration routes, a facet of distribution mapping that has been scarcely attempted. The Rufous Hummingbird (*Selasphorus rufus*) is a suitable example for display at an exceptionally small scale, considering the extent of its autumnal migration to wintering grounds on both the North and South American continents. Public data, downloaded from the Avian Knowledge Network (AKN), provides a suitable sample for testing cartographic generalization operators in creating ranges of avian species for use in multi-scale databases. Within this project, thorough exploration and critique of point aggregation operators provides a basis for completing this work and yields suggestions regarding improved methods for generalizing points. The resulting distributions and framework are relevant for providing a multi-scale representation schema to conservation managers in need of contextualized, temporally relevant species occurrence information.

Key Words: avian distribution, Avian Knowledge Network (AKN), biogeography, bird ranges, generalization, multi-representation databases, point aggregation.

MAPEO DE DATOS ESPECÍFICOS DE DISTRIBUCION GEOGRAFICA DE AVES UTILIZANDO OPERADORES CAPAZES DE AGREGAR PUNTOS CARTOGRAFICOS EN UN ARMAZON CONCEPTUAL DE ESCALAS MÚLTIPLES

Resumen. Historicamente, la delineacion de la distribucion geografica de aves es un proceso que se realiza manualmente, basandose en la opinion de expertos y sujeto a evidencia anedoctal, entre otros factores. Esta forma de proceeder es contraria al proceso de utilizar datasets unificados, lo cual hasta recientemente, no era disponible. Los esfuerzos detallados aqui han logrado desarollar un metodo de delinacion que no es arbitrario y utiliza solamente information actual y objectiva. Enlazado con el objetivo anterior, era la delinacion de las rutas de migracion de aves, un aspecto de esta sciencia que todavia permanece pobremente explorado. El colibri rojizo (*Selasphorus rufus*) es un ejemplar particularmente ideal para demostrar la capacidad de este proceso novelo: sus requisitos ecologicos y ambientales son de scala diminutiva, aunque es capaz de migrar a ambos continentes Americanos para invernar. Informacion disponible a cualquiera, obtenida desde el Avian Knowledge Network (AKN), provee una muestra valida para comprobar el uso de operadores generales cartograficos para representar la distribucion geographica de aves en forma utilizable dentro de databases de multi-scala. En este proyecto se descubre que la exploracion y manipulacion de operadores capazes de agregar puntos cartograficos, es la base para realizar el objetivo deseado y resulta en un proceso con el potencial de identificar mejores metodos para la generacion de tales puntos. Las resultantes distribuciones geographicas y armazon conceptual son representaciones interpretables a varias escalas espaciales y son util a empresarios de conservacion que necesitan informacion sobre la distribucion de especies que es al corriente do lo que es conocido y en contexto specifico a la situacion o problema.

[1] E-mail: tomauer@psu.edu

INTRODUCTION

When designing maps, it is important to select and represent features such that map aesthetic and function is preserved across multiple scales, being careful not to overcrowd the map. The generalization of map features is needed to reduce the overall amount of information on a map between scales to preserve readability. Following Topfer's Radical Law (Topfer and Pillewizer 1966), which established levels of selection in scale change, that as the denominator of scale increases the amount of map objects visible should decrease, as "information on a map is reduced proportional to the square root of scale change." Working with existing generalization operator taxonomies, this paper explores the use of aggregation, a geometry generalization (Roth 2007), applied to point data. Aggregation, as used here, can be defined as the grouping of multiple map objects (often points, but also lines or polygons) present at a large-scale (small-extent) into a single map object (often a polygon). This is seen as a reduction of data density and map complexity.

Traditionally, generalization operations are performed on topographic or reference maps, in developing series of maps at different scales. In this scenario, the aggregation operator will be applied specifically to a thematic map (specifically, a distribution map), but the tenability of its use within other, non-thematic series remains. Critique, analysis, and development of aggregation operators for point objects is the focus of this project.

Specific to the context of this project is the generalization of point data from a bird sighting database. This task arises from attempting to create bird distribution maps, emulating existing maps built from expert knowledge and multiple sources while improving on the ability of range maps to represent accurate and up-to-date information. The maps here are data-specific and are not manually adjusted to account for other sources or expert knowledge. This keeps the direction of this project in line with cartography's current demand for making generalization operations fully automated and not in need of manual adjustment. Finally, this project avoids techniques that apply raster-based interpolation or statistically-driven surface-building methods. That requires incorporating the value of the data at points, an issue of contention, as the data set on a raw level and for a continental scale is unlikely robust, homogenous, or uniform enough to trust this level of specificity. Statistical adjustments that account for first-order effects and underlying variability would need to be applied.

Following a review of the literature, this paper explores available methods for performing aggregations and yields a multi-scale framework for evaluating the aesthetic quality of their output. This framework and subsequent levels of map scale, from local to global, are established based on the respective management, conservation and eco-regional units and boundaries that relate to the data set at hand (avian species).

LITERATURE REVIEW

From the earliest mention of point aggregation (DeLucia and Black 1987), researchers have focused on two particular methods for describing, developing and detailing the operator-level work of point aggregation. The first is reducing map information by representing groups of points as polygonal, area features. The second identifies a need to indicate clustering as a way to identify when points should be aggregated (McMaster and Shea 1992, Regnauld and McMaster 2007). This is true from the first mention of point aggregation in DeLucia and Black (p. 181): "Points in space must form some association in order to determine their proximity with other points." The current state of point aggregation technique development needs to focus on this issue.

Despite DeLucia and Black (1987) acknowledging the need to find association amongst points to determine proximity, little mention of that issue follows the initial plea. Regnauld and McMaster (2007) address the issue, again lamenting (p. 42), "The critical problem in this operator is determining both the density of points needed to identify a cluster to be aggregated, and the boundary around the cluster." This comes despite the fact that McMaster had addressed the operator many times since 1987, yet failed to address methods for identifying clusters. Thus, the development of vector-based (as opposed to cell-based, raster format) strategies for cluster identification within the realm of point aggregating operators is still lacking.

Identifying clusters in a generalization framework proves to be a poorly developed aspect of aggregation. Aside from classic methods derived from computational geometry (described later as the feasible methods used in this study), there is little diversity in deriving clusters in geographic space from data points. However, two different approaches could potentially be extended to point aggregation; one statistical, the other visual.

While statistical methods for clustering points in geographic space through location and attribute value requires a unique attribute rarely found with topographic and non-thematic map data, it may be possible to apply certain point pattern analysis techniques for determining clusters. However, many available clustering analysis methods (such as Moran's I) merely acknowledge the presence or absence of clusters, but require more complex tools to their location and the subsequent significance. It is possible that more recent and efficient versions of the Geographic Analysis Machine (Openshaw et al. 1987) could provide cluster output guidance for aggregation, however, Sullivan and Unwin (2004) state that detecting clusters through underlying spatial variation present in almost all data is a difficult problem. It may be possible to extend hierarchical cluster analysis methods that work in attribute space to geographic space, as Grubesic and Murray (2001) suggest in detecting crime clusters. Unfortunately, this extension has been poorly studied or applied.

The other possibility for defining points for clustering is by modeling visual or perceptual clustering that map readers perform. Work by Slocum (1983) attempted to model users perceptions of clusters as a way of defining clusters for mapping. The first model was not entirely successful at predicting visually identified clusters, but suggested that with research a more developed model may succeed. Sadahiro (1997) continues this work by building and testing a similar model that proved to be valid, suggesting that it could be linked to GIS for choosing suitable symbol sizes. Further, Sadahiro suggests that it could be applied to map generalization as a way of selecting points for aggregation. Considering that generalization schemes work to present readable, efficient maps that are in part based on user perception, further exploration of this technique is highly suggested. However, due to its complexity and early stages of development, it was not within the scope of these efforts to pursue its application.

Aside from potentially applying Sadahiro's (1997) model, a scant few potential usable methods (all based in computational geometry) emerge from the literature. DeLucia and Black (1987) suggest a solution for point aggregation and cluster identity based on previously known computational geometry methods. Creating a Delaunay Triangle Network (a mesh of lines that connects each point in the network to its neighboring points) amongst all points and then using boundaries developed from Voronoi neighbor cells (artificial boundary cells derived from the mid-way distance between neighbors) as a means of segmenting clusters found some success in identifying groups of cells best represented as a single polygonal feature (DeLucia and Black 1987). Ware et al. (1995) continued with this method, applying it to aggregation of areal features. However, following these publications, this technique appears to be seldom used with points, departing into the realm of area aggregation. Development of this operator over the past decade remains static in the literature, with only simplistic description and little mention of technique.

Peng et al. (2004) present a short exploration of the topic and identify Convex Hull, adopted from computational geometry, as best suited for aggregation. Convex Hull (Fig. 1) is a geometric algorithm that involves drawing a line (or envelope) around the outer boundary of a group of points, creating a polygon or area feature that can represent distribution or range. Peng et al. (2004) go into detail on these triangulation methods, but finish saying (p. 2844), "the generalization methods of point clusters are still at the level of geometric operation and there is little breakthrough on geographic feature orientation," in that current methods for creating clusters rely on purely mathematical operations and do not take into account relevant geographic phenomenon that drive the patterns in the data (such as topography or habitat). This is a fair approximation of what the literature has to offer on algorithms and methods for performing clustering and grouping of geographic points for generalization.

An advancement of Convex Hull, the ShrinkWrap Hull (Revell 2004), which creates

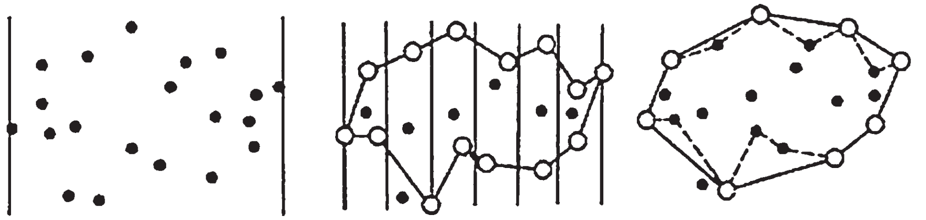

FIGURE 1. An example of a Convex Hull algorithm creating an envelope around points (Peng et al. 2004)

FIGURE 2. A ShrinkWrap Hull (the thick black line surrounding the buildings) (Revell. 2004)

a tighter wrap of points than a Convex Hull, also surfaced in the literature, but this author was unable to find specific algorithms for it. All mentions within the literature applied to generalization of building features on maps of residential areas, and ShrinkWrap Hull generally makes better use of orthogonal (right-angles most common in buildings) properties (Rainsford and Mackaness 2002, Revell 2004). It would appear that the algorithmic development of ShrinkWrap Hull, as shown wrapping a number of building groups in Figure 2, has become largely the domain of proprietary, non-academic groups; given that the references are from that realm.

Overall, there exists a general paucity of literature regarding point aggregating operators and it would seem that work in this field has either generated little interest in academia or has shifted into the realm of proprietary domains; the latter as suggested by Chaudhry and Mackaness (2004). Details on operator algorithms, methods, and potentials are scant in the literature. While there may not be much published concern for this subject, considering the rapid development of automated generalization schemes, developing techniques for identifying clusters and aggregating them in vector-based format is important.

Finally, in creating point aggregation operation outputs for multi-representation frameworks, it is appropriate to review the relevant literature that gives guidelines for such. Cecconi (2003) establishes a starting point for generalization models, levels of detail, operator selection, and database construction. Using his methods, the outputs of the aggregation operators created in this work can then be placed within a multi-scale framework, including the relevant management, conservation, and eco-regional layers. Brewer and Buttenfield (2007) explore multi-scale frameworks and develop context for scale selection and cross-scale operator function potential. These two pieces provide example frameworks and scale selection benchmarks for developing a multi-scale framework in evaluating the resulting distribution maps.

METHODOLOGY

The data for the project came from the Avian Knowledge Network (AKN 2007). From all of the North American bird species available to work with, Rufous Hummingbird (*Selasphorus rufus*) was selected because of its migratory nature and tendency for vagrancy at one particular season (winter). It provided a core range that was not too spatially extensive, in that it had well-defined patterns of occurrence associated with mountain ranges, making identifying accurate generalization obvious in most map reading tasks. Data were selected from 1970 through 2007. Initial exploration, however, shows that most of these data were entered within the last ten years. Of the twelve months available, the month of June was selected, being near the peak of the breeding season, providing a stable distribution, but also being near the beginning of migration for some individuals of the species and thus providing sightings beyond predicted occurrence locations (Healy and Calder 2006). The month of June yielded a total of 1263 points across the United States (including Alaska) and Canada. The scant data from Mexico were eliminated. The lack of sightings may have produced spatial anomalies. As well, most existing range maps are constructed for use within the US and Canada, not Mexico. However, it is possible that with new mapping methods and increasing amounts of data from previously less-studied regions that the inclusion of data from outside the US and Canada will be important and feasible. Existing range maps are overly simple and considering the rapid development of dynamic mapping technology, an upgrade is in order.

The data was brought into ArcMap 9.2 with the following attributes: Object ID, Latitude, Longitude, Year, Month, Day and Number. The number of sightings associated with each point was not used in the most aspects of this project as the goal was to primarily focus on extent as opposed to density, however those data were kept with this data set for further exploration. Personal Geodatabases and Feature Classes were constructed to organize data layers. At least some of these data were entered with handheld GPS units commonly using the WGS 1984 Datum and, thus, the data were assigned to this datum. While this assumption may be considered spurious, as spatial referencing uncertainty produced by this step may have produced error in the range of 15 meters, one datum had to be selected and the scales for analysis made this issue less important. Projection to North American Continental Albers Equal Area was used for final presentation. This projection was seen as best representing the entirety of the continent, ensuring equal area depiction for the aggregated polygons.

In running operator functions on the data, the only layer necessary was the data itself. However, a background layer of North American political boundaries [1] provided a reference to ensure that, visually, procedures were running without error. For cartographic output within a multi-representation framework, a number of layers were used to provide base data reference that assisted in visual interpretation of scale levels. They are as follows: USGS National Elevation Dataset Shaded Relief [2], National Wildlife Refuge Boundaries [3], and Bird Conservation Regions [4]. These layers were then used without generalization across the following scales: 1:20 000; 1:250 000; 1:2 000 000; 1:4 000 000; 1:12 000 000, 1:40 000 000. One unique aspect of the project was the breadth of scales and the use of continental extent. While such a range of scales is not commonly used in generalization schemes, this dataset demanded consideration of all scales to appropriately depict the entirety of ranges. Construction of the multi-representation framework and display will be covered in the Discussion section.

Only the previously mentioned methods (Convex Hull and ShrinkWrap Hull), adapted from computational geometry, were used for aggregation. A simple Convex Hull for points is not readily available in the ArcToolbox. The "Bounding Containers" ArcScript downloaded from Environmental Systems Research Institute [5] provided a simple operator for wrapping all points in a data layer. Beyond this, efforts had to be made to emulate a ShrinkWrap Hull. This tool is not available within the ESRI suite, so I constructed a modeled process to create this operation. In the ArcToolbox, under Data Management Tools in the Generalization section, Aggregate Polygons essentially allows for

FIGURE 3. A polygon of Rufous Hummingbird sightings for June created with a Convex Hull algorithm allowing no holes.

a ShrinkWrap Hull. This tool works much in the way that a Convex Hull operates, except that it creates a tighter envelope of the perimeter through threshold distance selection.

The Aggregate Polygon tool introduces the first options to adjust the outcome. Two adjustable variables required evaluation. First, the threshold of aggregation had to be determined. The value of the aggregation threshold establishes that any points within X miles of each other will be aggregated. In other words, any point within the aggregated polygon is within at least X miles of at least one other point. Threshold values ranging from 1000 miles to 25 miles were tested and the outcomes will be presented in the Results and Discussion sections. The other option was to adjust for holes within the aggregation. It was possible to either allow holes of any size by setting the threshold for minimum allowable hole size (in square miles) or to not allow any holes, by setting the minimum allowable hole size to an impossibly large value. The differences, at a 50-mile threshold aggregation, can be seen in Figs. 3 and 4. In completing the simulated ShrinkWrap, each point was converted to a very small area (<0.01 miles) circular polygon through a simple buffer. These were then used as the input to the Aggregate Polygon tool itself. It was important to be careful in making the point buffers, because if their calculated area was too small the Aggregate Polygons tool failed to execute

FIGURE 4. A polygon of Rufous Hummingbird sightings for June created with a Convex Hull algorithm allowing holes to exist.

because of calculation errors. This constraint renders this modeled tool aesthetically unfeasible at the largest scales (1:250 000 or larger) because the buffers become obvious circles.

RESULTS

The Convex Hull operator (Fig. 5) was largely successful at only the smallest scale (1:40 000 000). This operator created a good result for the entire continental extent (seen here with original points drawn additionally). This operator provided the most simple, most general and easiest operation.

The simulated Shrinkwrap Hull created a generally more desirable result (Fig. 6) due to the variable thresholds of aggregation. This allowed for levels of simulated cluster identity (based on simple proximity). Selecting the threshold distance, however, became an arbitrary act and often left behind aesthetically unpleasing shapes, spiky artifacts and lines, or didn't aggregate enough points. Yet, this form gave the best sense of density distribution. A solution to some of these problems (Fig. 7) came from the Polygon Shrink Expand ArcScript, which smoothed borders and eliminated overly eccentric lines and spikes. This would have allowed for consistent creation of a desirable end product that could be used to create a range of threshold layers, had it not been so computationally intensive as to prohibit completing it for more than one example.

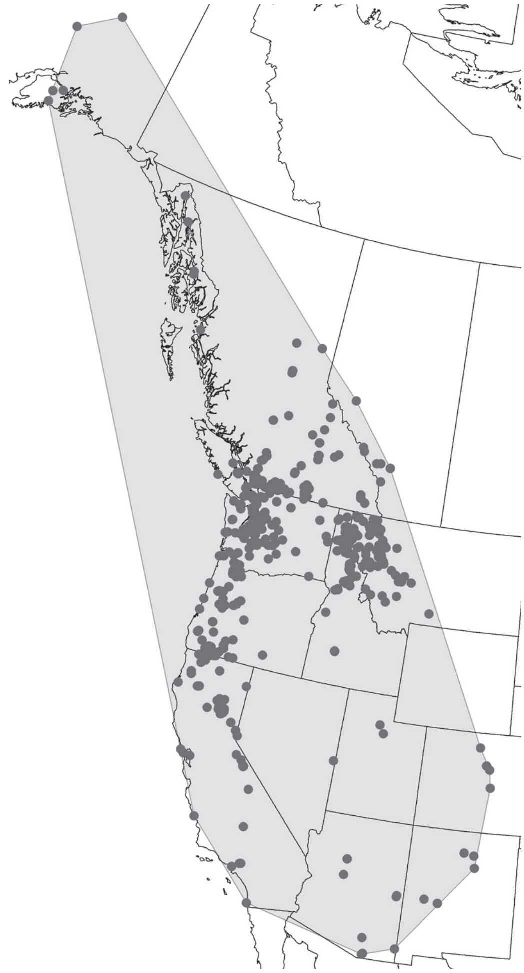

FIGURE 5. A polygon of Rufous Hummingbird sightings for June created with a Convex Hull algorithm, showing the hull and the original points.

Finally, combining multiple elements, including the ShrinkExpand tool, outputs for multiple months were merged together to create a composite map (Fig. 8: http://www.geovista.psu.edu/pif/extendedfigures.html) for the Summer, Fall and Winter seasons of the species' movement patterns. Extraneous points were those left over from threshold selection at 200 miles and were represented separately for each season. Intersections between the seasons were used to show overlap. The polygons were then smoothed (for aesthetic effect) using the Smooth Polygon operation available in ArcGiS 9.2. This output is seen as the most accurate recreation of existing range maps. It would also be possible, in an interactive environment, to give

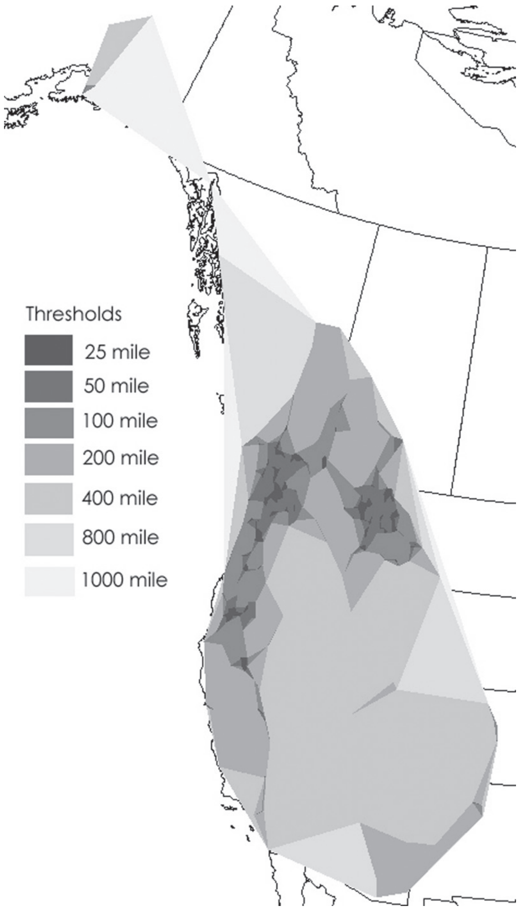

FIGURE 6. A layered, composite map using a simulated ShrinkWrap Hull method at different thresholds of wrapping distance.

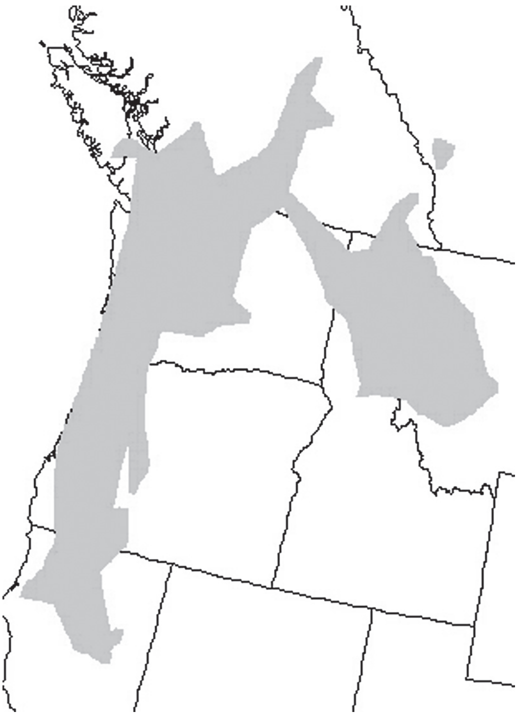

FIGURE 7. The boundaries here have been cleaned with the ShrinkExpand tool.

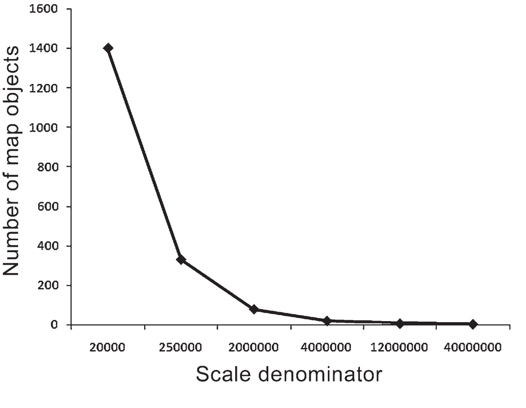

FIGURE 8. A predicted map object curve from Topfer and Pillewizer (1966).

the user the ability to institute threshold ranges for revealing density concentrations.

DISCUSSION

Applying these aggregation operator outputs in a multi-scale framework requires some theoretical consideration first. Using Topfer's Radical Law, the predicted trend of map object selection as applied to this particular framework is seen (Fig. 9) across the scales defined.

However, as depicted by the "Bird Points" curve (Fig. 10), moving from the largest scale to the smallest scale, an increasingly growing number of objects are encountered within the extent view (as points of bird sightings). When viewing the data at 1:20 000 only 25 points appear, yet scaling out to 1:40 000 000, all 1263 points are visible. To conform to the Radical Law, there should be significantly fewer objects at the smallest scale (1:40 000 000), not more objects. The "Bird Points and Polygons" curve in Fig. 9 shows that point aggregation needs to occur at a particular scale to reduce map information to the appropriate level. In this case, it is quite obviously at the scale of 1:4 000 000. Considering the large rate of increase in data points between 1:4 000 000 and 1:12 000 000, it then seems appropriate to substitute areal aggregations for points at this level and smaller scales.

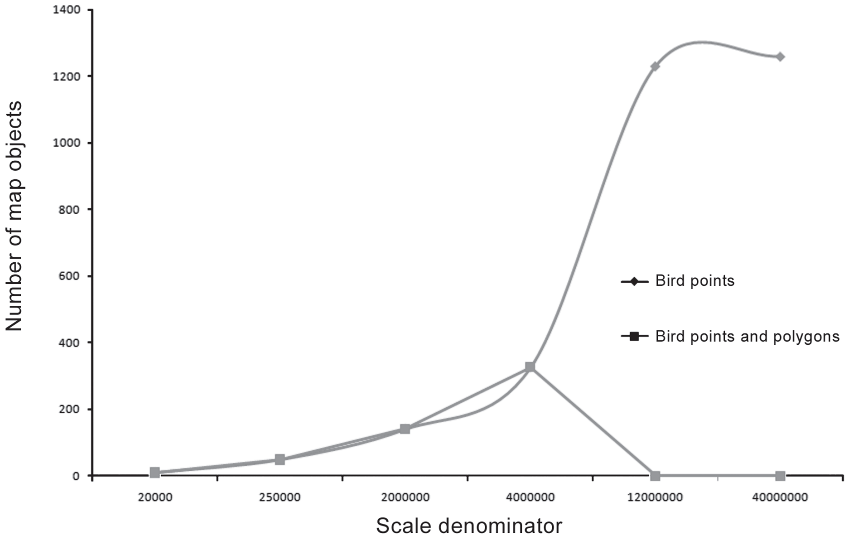

FIGURE 9. An actual observed map object curve.

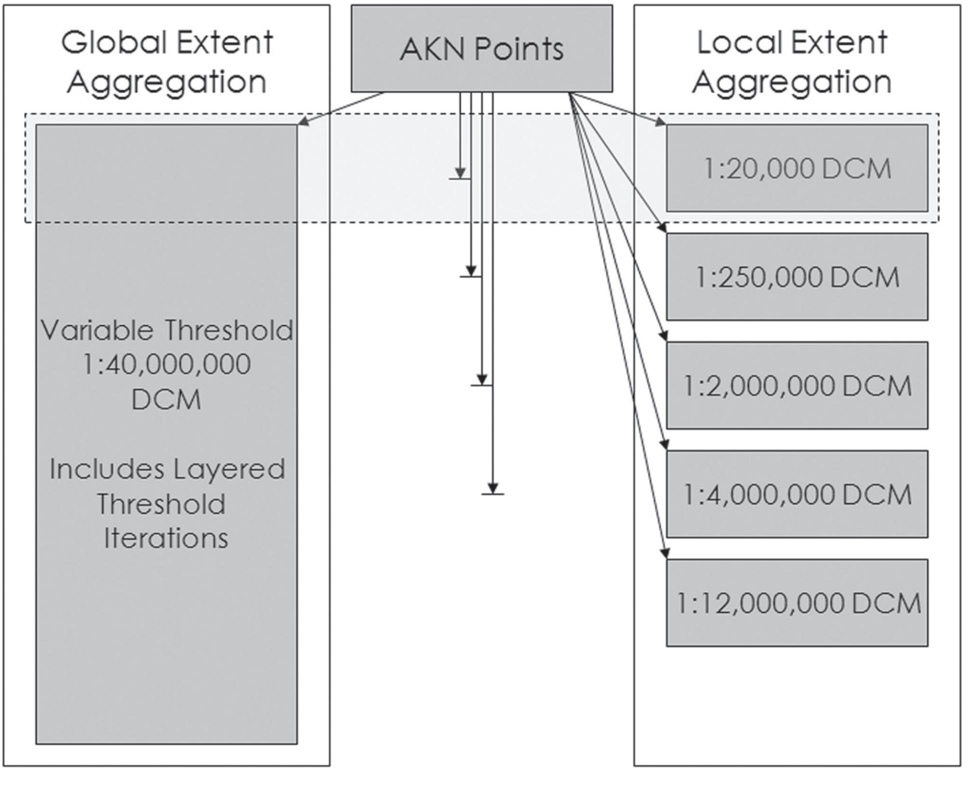

FIGURE 10. The generalization model used across multiple scales and showing the context of relevant layers.

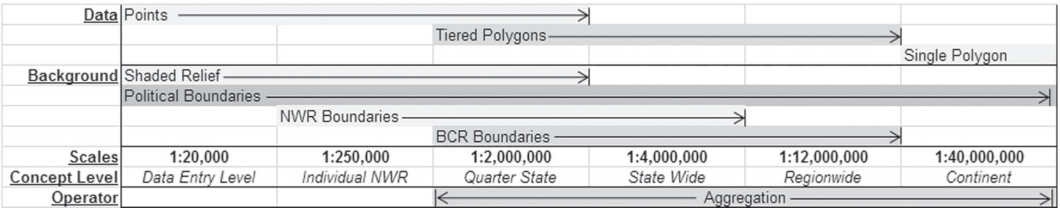

FIGURE 11. The composite multi-scale framework, including data, background layers, scales, conceptual levels, and operators.

Initially, following a Multi-Scale Framework Generalization Model (Fig. 11) for this work, it was possible to aggregate at both the local level and global level. Colleagues felt that being able to see both the global aggregation and local aggregation of points might be rewarding. With this in mind, a representative example was constructed, reviewed and deemed to generally be neither useful nor aesthetically desirable. One of the issues was that at the largest scales the simulated ShrinkWrap hull via the Aggregate Polygon tool no longer functioned properly, such that using Convex Hull for local aggregation was not particularly revealing. These graphical results have been omitted. From this, it was determined that combining a view of either points, points and polygons, or just a single or series of tiered polygons would prove to be most effective.

This composite framework (graphically depicted in Fig. 12), designed around the conceptual levels of scale and management units, is mapped in Figs. 13–18 (http://www.geovista.psu.edu/pif/extendedfigures.html). Despite having to utilize the unrefined simulation of the Shrinkwrap Hull (due to computational limitations), this was seen as an effective tool for exploring these ranges through point aggregation. The greatest amount of all represented data layers (not just the individual bird sighting points) happens at the scale of 1:2 000 000. While in some ways this runs counter to Topfer's Radical Law, there are multiple layers involved, each contextual and important to the overall appearance, such that a user's ability to read the map at this scale is not affected by the number of points represented. At 1:4 000 000, data reduction becomes dramatic and the representation rather simplified. In the eyes of this author the view at 1:2 000 000 provided the most useful representation. In future applications of this framework, this scale should be considered an important anchor scale.

In the end, most experts evaluating this series of procedures regarded the multi-scale framework as successful, however, they presented alternative routes via raster models for the actual generalization of these particular data. Most of these alternatives were particular to this project's context. Because of quantification limitations within the data set, exploring raster developments was not within the scope of this project and requires further study of adjustment and quantification methods to account for underlying variability. This author agrees that taking this particular project in the direction of raster would be most successful for interactive exploration, but still believes that there are those who may have lost sight of needing to accomplish non-statistical, vector-based strategies for aggregating points in topographic and navigational map series. Moreover, in creating a framework for producing range maps that can be updated based on current and recent sightings, this method has succeeded.

CONCLUSION

While the thematic context of this project suggests better execution based in raster format, there is still demand for solutions that aggregate points based on some level of clustering indication within the vector format. Instances might include cities and towns in metropolitan regions, any number of municipal features, buildings, and built up areas. This author suggests building on known Hull algorithms to incorporate cluster identification to best "shrink wrap" identified groups of points. A hybrid approach could incorporate cluster identification from hierarchical clustering methods or visual prediction models, passing identified clusters to geometrical algorithms for drawing. Proportions of regionalization can be based upon density measurements or clustering indices and then used to drive area symbolization.

For this particular dataset and context, the multi-scale framework provides decent guidance for representation selection across scale. The framework developed here is being applied to a web-tool designed to represent the entirety of this particular dataset. Better forums for exploratory analysis of this particular data could be found in raster interpolation techniques or within a specifically designed Space-Time GIS best suited to the strong temporal nature of the data. However, the best solutions for aggregating points in vector are currently held by non-geographic computational geometry. Some of the solutions presented here provide viable results to be placed into an automated framework for multi-scale representation. There is potential to extend this work for creating better solutions to aggregate points within any generalization framework.

LITERATURE CITED

Brewer, C., and B. Buttenfield. 2007. Framing guidelines for multi-scale map design. Cartography and Geographic Information Science 34:3–15.

Cecconi, A. 2003. Integration of Cartographic Generalization and Multi-Scale Databases for Enhanced Web Mapping. Ph.D. dissertation, Geographisches Institut, Universität Zürich.

Chaudhry, O., and W. Mackaness. 2005. Modelling Geographic Phenomena At Multiple Levels of Detail. Institute of Geography, University of Edinburgh.

DeLucia, A., and R. Black. 1987. A Comprehensive Approach to Automatic Feature Generalization. Proc. 13th Conference of the International Cartographic Association: Morelia, Mexico 4:173–191.

Download Data. 2007. The Avian Knowledge Network (AKN). Cornell Lab of Ornithology. [Online.] <http://www.avianknowledge.net/content/> (accessed 20 September 2007).

Grubesic, T., and A. Murray. 2001. Detecting Hot Spots Using Cluster Analysis and GIS. Fifth Annual International Crime Mapping Research Conference Dallas, TX.

Healy, S., and W. A. Calder. 2006. Rufous Hummingbird (*Selasphorus rufus*). *In* A. Poole (ed.), The Birds of North America Online Cornell Lab of Ornithology, Ithaca, NY. [Online.] <http://bna.birds.cornell.edu/bna/species/053> (Retrieved from the Birds of North America Online).

Li, Z. 2007. Algorithmic Foundation of Multi-Scale Spatial Representation. CRC Press, Boca Raton, FL.

McMaster, R., and K. S. Shea. 1992. Generalization in Digital Cartography. Association of American Geographer, Washington, D.C.

O'Sullivan, D., and D. J. Unwin. 2003. Geographic Information Analysis. John Wiley & Sons, Inc. Hoboken, NJ.

Openshaw, S., and M. E. Charlton. 1987. OA Mark 1 Geographical Analysis Machine for the Automated Analysis of Point Data Sets. International Journal of Geographical Information Science and Systems 1:335–358.

Peng, H., Q. Qingwen, and L. Zhaoli. 2004 Progress in studies on automated generalization of spatial point cluster. Geoscience and Remote Sensing Symposium (IGARSS '04) Proceedings 2004 IEEE International (5) pp. 2841–2844.

Rainsford, D., and W. Mackaness. 2002. Template Matching in Support of Generalisation of Rural Buildings, pp. 137–151. *In* D. Richardson D. and P. van Oosterom [eds.], Advances in Spatial Data Handling. Proceedings 10th International Symposium on Spatial Data Handling, Springer, Berlin; Heidelberg.

Regnauld, N., and R. McMaster. 2007. A Synoptic View of Generalisation Operators, pp. 37–66. *In* W. A. Mackaness, A. Ruas, and L. T. Sarjakoski [eds.], Generalisation of Geographic Information: Cartographic Modelling and Applications. Elsevier.

Revell P. 2004. Building on Past Achievements: Generalizing OS MasterMap Rural Buildings to 1:50,000. ICA Workshop on Generalization and Multiple Representation, 20–21 August 2004, Leicester.

Roth, R. 2007. ScaleMaster.Org Theoretical Background: Identifying and Defining Geometry, Symbology, and Filtering Generalization Choices for Legible Representation of Geospatial Datasets At Multiple Scales. Seminar in Cartography—Mini-Conference. Department of Geography, University Park, PA. 12 Dec. 2007.

Sadahiro, Y. 1997. Cluster Perception in the Distribution of Point Objects. Cartographica: The International Journal for Geographic Information and Geovisualization 34:49–62.

Slocum, T., R. McMaster, F. Keller, and H. Howard. 2005. Thematic Cartography and Geographic Visualization. 2nd ed. Prentice Hall, Upper Saddle River, NJ.

Topfer, F., and W. Pillewizer. 1966. The Principles of Selection, a Means of Cartographic Generalization. Cartographic Journal 3:10–16.

Ware, J., C. Jones, and G. Bundy. 1995. A triangulated spatial model for cartographic generalization of areal objects. Springer–Verlag Lecture Notes in Computer Science 988:173–193.

DATA SOURCES

[1] North American Political Boundaries (http://www.nationalatlas.gov/atlasftp.html)

[2] USGS NED Shaded Relief (http://www.esri.com/data/download/usgs_ned/index.html)
[3] National Wildlife Refuge Boundaries (http://www.fws.gov/data/datafws.html)
[4] Bird Conservation Regions (http://mbirdims.fws.gov/nbii/)
[5] ESRI "Bounding Containers" ArcScript (http://arcscripts.esri.com/details.asp?dbid=14535)
[6] ESRI "Polygon Shrink Expand" ArcScript (http://arcscripts.esri.com/details.asp?dbid=15019)