Perception of Pattern and Process in Seabird Distributions: an Assessment of the Impact of Alternative Visualization Methods

David J. Lieske[1]

[1] Department of Geography and Environment, Mount Allison University, Sackville, NB, dlieske@mta.ca

Abstract

Advances in scientific visualization offer an ever-growing number of opportunities to depict and explore spatial information (DiBiase et al. 1994). However, parallel developments in different fields have limited communication about when and how different visualizations are effective (Gahegan 2008). As a member of an inter-disciplinary group interested in the spatial distribution of seabirds at-sea, the author sought to compare the relative effectiveness and interpretability of a range of visualization approaches. Through the use of an anonymous web-based survey, four different visualizations were evaluated by a small sample of study participants: 2D proportional-symbol maps, 2D interpolated-surface maps, and both static and dynamic 3D maps. There was evidence that the type of visualization impacted respondent's interpretations, with dynamic 3D visualizations stimulating conclusions not reported using any other method. Respondents commented on all four types of visualizations, but the provisioning of a 3D perspective was described as more "precise" and informative. The type of data was also important for determining the intelligibility of the 2D maps, with the more sparsely-distributed Thick-billed Murre more clearly showing hotspots of occurrence. User experience may have influenced the results and lead to more favourable evaluations of the role of interpolation surfaces and dynamic 3D visualizations.

Background and Relevance

Advances in scientific visualization offer an ever-growing number of opportunities to depict and explore spatial information (DiBiase et al. 1994). However, the historical development of visualization systems had proceeded across different disciplines in isolation from each other (Gahegan 2008), resulting in a lack of communication and inter-compatibility between different systems. This situation forces the spatial analyst to choose a system to study and utilize, with little guidance as to which visualization elements are going to be effective (Gahegan 2008). The question of "which visualization method will be most effective for exploring my particular phenomenon of interest?" remains unanswered.

As a member of an interdisciplinary group with different backgrounds and experiences with spatial analysis tools, but a shared interest and concern for the offshore distribution of seabirds inhabiting Canadian Atlantic waters, the author sought to reveal new "truths" about the spatial distribution of seabirds at-sea. But parallel to this main objective a secondary goal emerged: in what way does the type of visualization inform the "end user"? What impact does the visualization approach have on their ability to make inferences? And furthermore, what role do visualizations play in fostering collaboration and generating new research questions?

In this preliminary study the relative effectiveness and interpretability of a range of visualization approaches were compared. In a survey of the seabird literature there was a unanimous tendency for authors to fix the scale and extent of the visualization, and to rely on conventional 2D maps and symbology. Particular elements of symbolization were preferentially employed (e.g., proportional symbols: Certain *et al.* 2007, Serra-Sogas *et al.* 2008), as were uni-dimensional summaries of abundance as a function of distance along a survey track (Durazo *et al.* 1998, Skov and Durinck 1998). Occasionally, predictive models were employed to estimate abundances for particular points (Yen *et al.* 2004a) or entire surfaces (Fauchald *et al.* 2002, Clarke *et al.* 2003, Skov *et al.* 2008), but at 2D scales, extents and resolutions fixed by the authors. This is despite the evidence raised by Weimerskirch (2007) that seabird distributional patterns vary with scale as a result of changes in the nature of the species response. For example, at broad scales distributional patterns reflect prominent oceanographic features (such as the location of seamounts or shelf breaks; see Skov and Durinck 1998, Yen *et al.* 2004b) or constraints imposed by the need to be within tolerable distances to breeding colonies (Yen *et al.* 2004b, Weimerskirch 2007).

At finer spatial scales, distribution is much more sensitive to local combinations of prey abundance, and individual birds will show peculiarities in their daily movement patterns (Weimerskirch 2007), respond to the presence of other species (Durazo *et al.* 1998), etc. Seasonal variation in habitat usage also adds a source of temporal variation in seabird distribution that confounds simple depictions of species occurrence. Fixing the scale and extent of visualizations helps to make the task of interpreting distributions more manageable (Andrienko *et al.* 2006) but may also obscure important patterns. The problem of scale-dependent emergent behavior doesn't just complicate the understanding of animal distribution but represents a major research challenge facing any geovisualization study, methodology, or theoretical framework (MacEachren and Kraak 2001).

Four types of visualization were developed, each of which could be naturally positioned along a continuum of visualization complexity (both in terms of the amount of information communicated as well as the level of interaction required to interpret them). The first two were strictly 2D and involved the use of either (1) proportional symbols, or (2) interpolated surfaces. These closely mimicked the types of visualizations currently employed in the seabird literature, and were of fixed scale and extent. The last two were 3D images, and were either (3) static 3D or (4) semi-interactive, animated 3D. These last two methods were totally novel approaches to representing seabird distributions, with the latter varying scale as well as orientation. All methods shared the same chief objectives, however: provide a synoptic view of distribution that would allow inferences to be made about the way birds were using the marine environment, and to illuminate how seasonality influenced that usage.

Through the use of a small, anonymous sample of colleagues from a seabird research project (the Atlantic Beached Bird Analysis), the utility of each visualization was assessed. Similar to Brunsdon et al. (2007), the following four questions were examined: (1) do the interpretations of spatial trends vary according to the types of visualization technique employed?; (2) is one technique clearly preferred, or is a combination the best way to enhance understanding?; (3) does the suitability of the visualization method depend on the type of data available?; and (4) does the effectiveness of the techniques vary with user experience?

Methods and Data

Observations of seabird occurrence and abundance were gathered during at-sea surveys conducted from March 2006 to March 2008 (Wilhelm et al. 2008). At regular positions along the survey track the presence or absence of different seabird species were noted, as were the numbers observed and the season in which observations were made. For the purposes of this study, "winter" was classified as observations gathered from November to January, "spring" from February to April, "summer" from May to July, and "fall" from August to October

The 2D proportional-symbol maps used symbol size to communicate relative differences in abundance, and their scaling was defined using the following formula:

$$scaling = \frac{k * CV * count}{\max(count)}$$

where *k* was a constant (defined as 5) and *CV* is the coefficient of variability of the counts (standard deviation / mean). The visualizations were conditioned on season, with a separate figure generated for each combination. These black-and-white figures are representative of many geographic visualizations that appear in the marine literature (e.g., Certain et al. 2007, Fig. 3a; Serra-Sogas et al. 2008, Fig.4). The figures were generated using the R Statistical Package (Ihaka and Gentleman 1996).

As with the 2D proportional-symbol visualizations, 2D interpolated-surface maps contained information about where (and in which season) seabirds were observed. Through the use of differences in hue and saturation, the four-panel display (one figure per season) was reduced to one single figure. Symbols were of uniform size – indicating at least one seabird observation – with color indicating the season in which the observation occurred. The locations of the survey tracks were added as uniformly-sized grey symbols, constituting a new datum missing from the previous visualization. Additionally, differences in relative abundance were symbolized using a color-graded interpolation surface (with blue = lowest abundance, red = highest abundance). The extent of the study area was also indicated using a dark-grey surface. This visualization was produced using ArcMap (Environmental Systems Research Institute 2006a).

The same information provided by the 2D interpolated-surface maps were also available in the static 3D visualizations, but rather than employ a colour-scaled interpolation surface, symbol height was used to represent relative differences in seabird abundance. Inevitably, the use of 3D perspective reduced the dependence on colour as an information medium and may have helped reduce the apparent complexity of the image.

Being a static display, the vantage point was defined by the map-maker (DJL), with users being presented with only one instance of orientation and scale. This visualization was produced using ArcMap (Environmental Systems Research Institute 2006a).

Lastly, dynamic 3D visualizations were produced to help alleviate some of the previously mentioned limitations. While respondents could control neither the scale nor the orientation of the visualizations, they were presented with a "fly by" panoramic view that took them through the landscape by rotating the surface 360^o counter-clockwise (thereby accommodating different orientations), and which varied the distance to the surface (thereby altering scale). The animations were produced using a series of still images gathered using ArcScene 9.2 (Environmental Systems Research Institute 2006b), and were distributed as large (~40Mb) avi animation files.

A key criterion for the design of this study was the protection of respondents' anonymity. Furthermore, in order to maximize the likelihood of candidates choosing to participate in the study it was necessary to design an interface that would render it as simple as possible for them to observe the visualizations and provide their responses. For this reason, a web-based platform was programmed using a combination of the PHP 4 scripting language (PHP Group 2008) and a MySQL database backend (Sun Microsystems 2008). Users visiting the website were taken on a "guided tour" of each of the visualizations (first for the Northern Gannet, then for the Thick-billed Murre) and provided with an opportunity to submit comments for each one independently.

Results

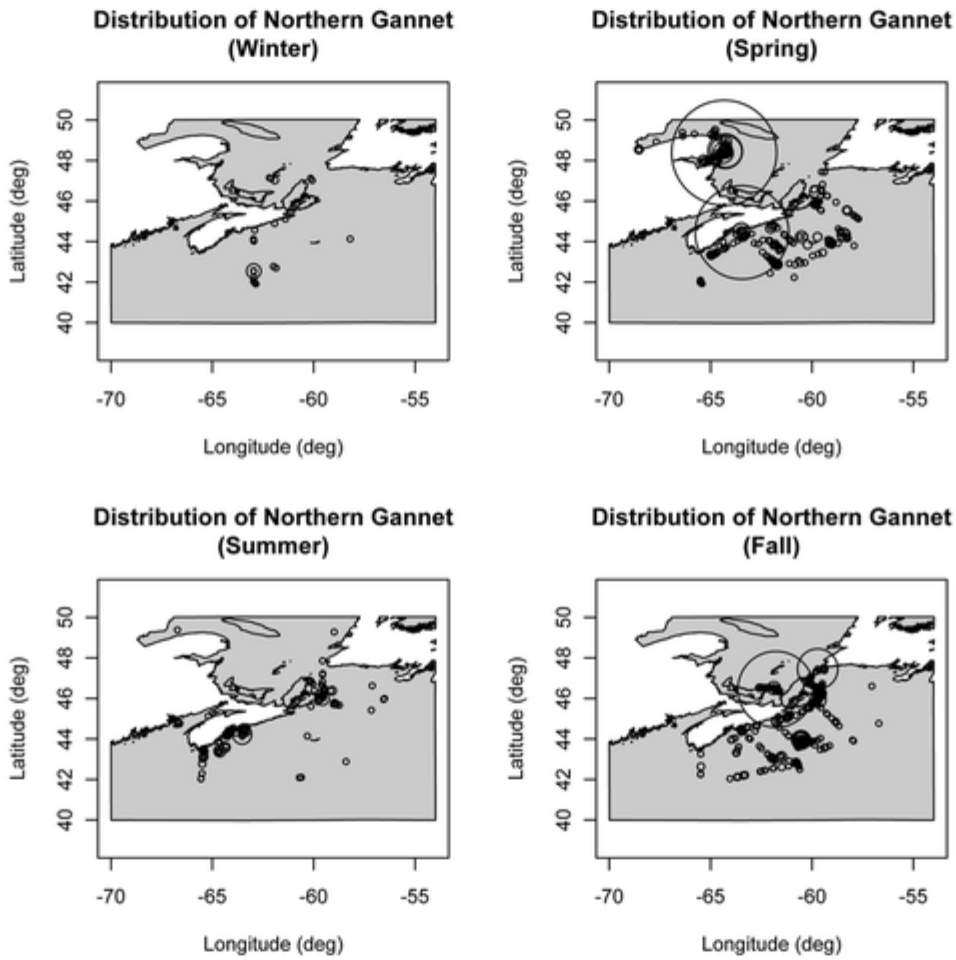
Participant's Prior Experience with GIS

The web interface received records from five participants. When posed the question "have you ever taken a course in cartography, surveying or GIS?", half of the participants who responded to this question reported that they had completed a course in GIS, while one declared that they used GIS but had never studied it formally. Another participant described themselves as having never used GIS.

In response to the following question "if you are familiar with, and use GIS on a regular basis, please indicate how long you have used these software tools", two of the five reported using it for more than 2 years, while one individual reported using if for more than 6 months but less than 2 years. Two of the five reported that they had never used GIS.

All respondents responded positively to the question "do you enjoy studying and using maps?".

The final question was one of self-assessment: "how would you characterize your GIS and mapping skills?'. Four of the five respondents (80%) described themselves as "informed" users, while one described themselves as a "novice".



Distribution of Northern Gannet (Winter)
Distribution of Northern Gannet (Spring)
Distribution of Northern Gannet (Summer)
Distribution of Northern Gannet (Fall)
Latitude (deg)
Longitude (deg)
40
42
44
46
48
50
-70
-65
-60
-55

Fig 1a.

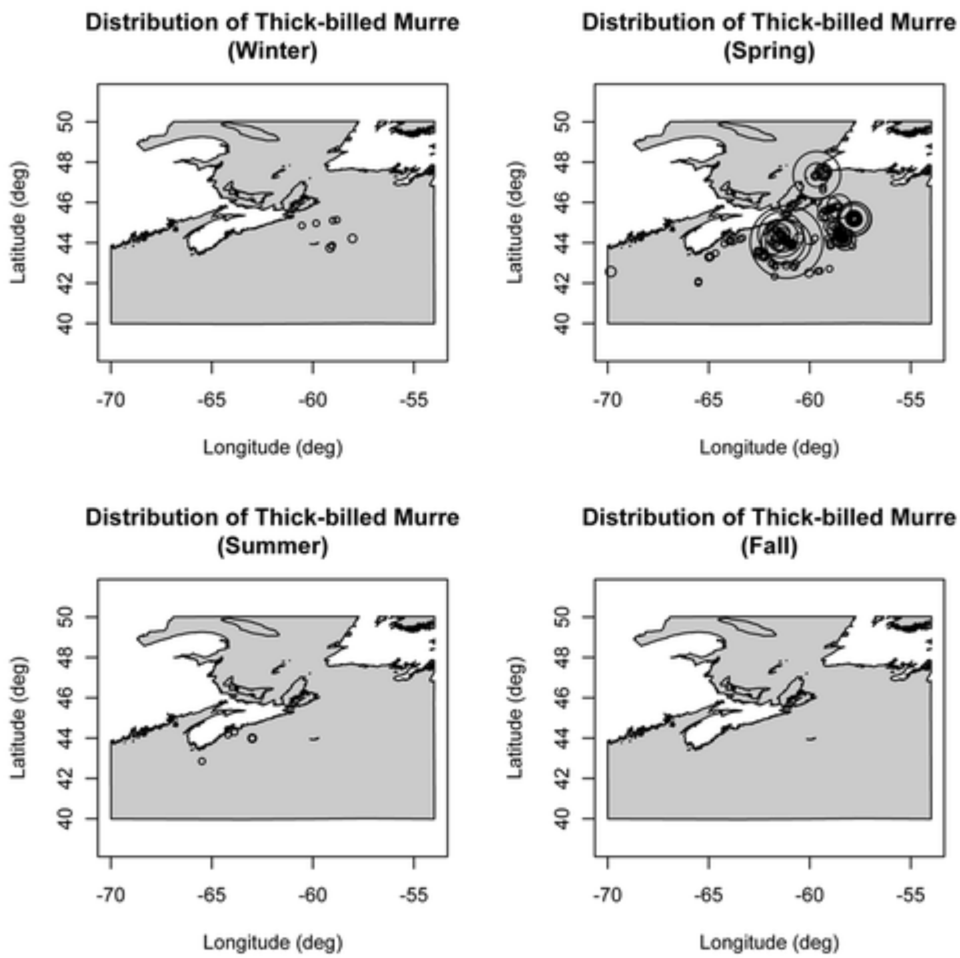


Fig. 1b.

Receptivity to 2D, Proportional Symbols Maps

Participants were first presented with two-dimensional proportional symbol maps for each species (Fig. 1a,b). In these figures, three factors were displayed: (1) locations of occurrence, (2) abundance (proportional to symbol diameter), and (3) season (winter, spring, summer and fall). No attempt was made to present seasonal variation in occurrence and abundance in a single figure; instead, four separate figures were generated, one for each season. This type of figure is representative of many geographic visualizations that appear in the marine literature (e.g., Certain et al. 2007, Fig. 3a; Serra-Sogas et al. 2008, Fig.4).

Referring to the criteria of Williamson and McGuinness (1990), two respondents reported that it was difficult to distinguish relative differences in abundance because the "maps were too small" and points overlapped too much. Overlap seemed most confusing for Fig. 1a, where many individual occurrences resulted in a larger number of uniform symbols. Two respondents expressed a difficulty with distinguishing these smaller points from islands. This seemed to be less of a problem for the Thick-billed

Murre visualization (Fig.1b), as one respondent described the proportional symbols to be “quite effective for showing hotspots”. In this case, winter, summer and fall occurrences were uncommon and sharply contrasted with the heavier spring usage. A number of respondents correctly described the seasonal differences in occurrence but were also able to pinpoint such geographic locations as the Halifax Harbour and the Cabot Strait, neither of which were labeled or indicated in the original visualizations. This suggests that prior geographic familiarity may contribute to map interpretability. Some respondents also reported discomfort with the absence of information about survey effort, particularly in the case of Fig. 1b where absences in winter, summer and fall may have been partially attributable to uneven sampling coverage.

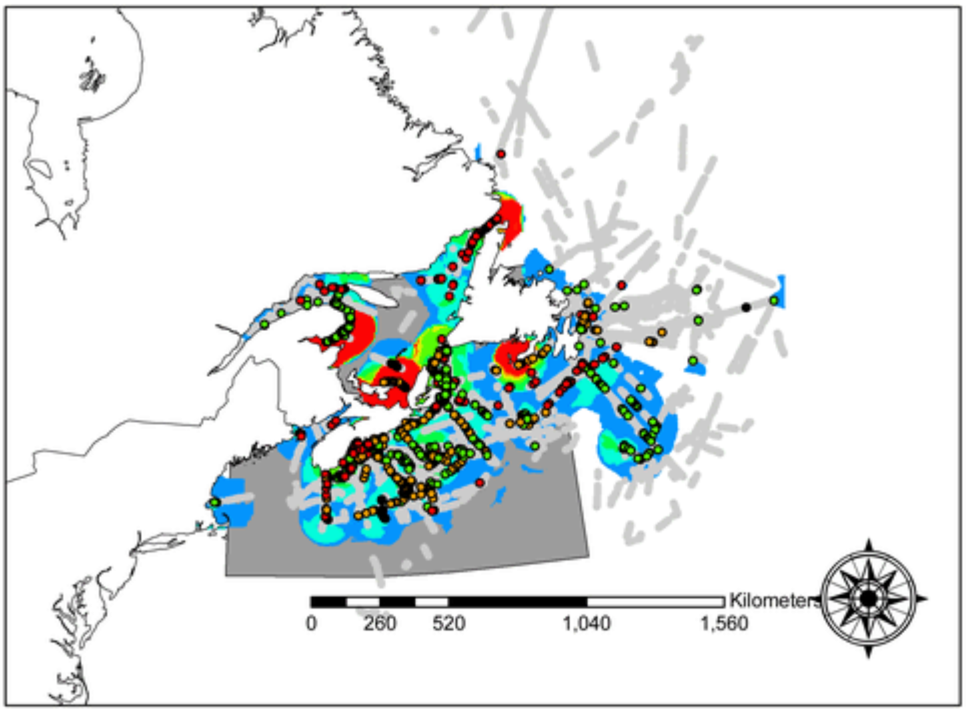


Fig. 2a.

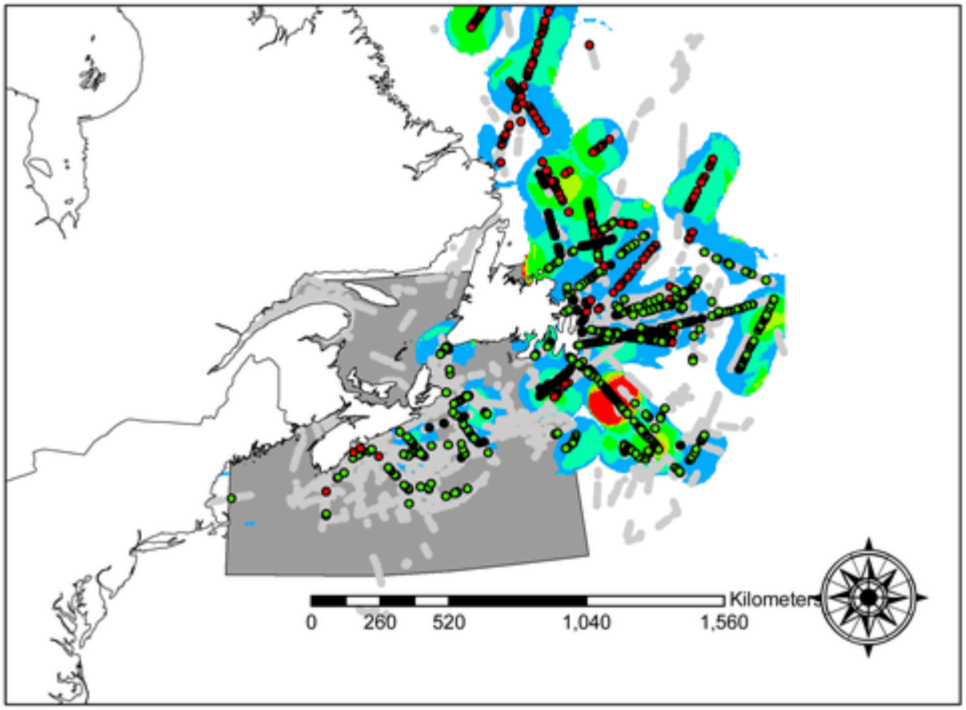


Fig. 2b.

Receptivity to 2D, Interpolated Surface Maps

Figures 2a,b presented two-dimensional interpolated surface maps with all four seasons shown simultaneously. In these figures, four factors were displayed: (1) locations of occurrence, (2) abundance (proportional to the colour gradient of the interpolation surface), (3) season (winter, spring, summer and fall, indicated by solid symbol colour), and (4) survey tracks (light gray symbols). These figures addressed a deficiency expressed for Figs. 1a,b, i.e., the provisioning of survey tracks, but they also used two different colour gradients and one single panel.

Spatial trends were described as "unclear", with Figs.2a,b showing too many colours and too much information. As with Figs.1a,b, three respondents reported that there was too much overlap in the occurrence points, further obscuring spatial trends. One respondent reported a preference for separate seasonal visualizations, while another responded favourably to the use of colour-shaded interpolation. Respondents were comfortable describing the reliance of the Northern Gannet on the shallower water of the Gulf of St. Lawrence, Scotian Shelf and the Newfoundland Grand Banks during most seasons, and correctly described the predominantly Newfoundland offshore distribution of the Thick-billed Murre. Despite expressing reservations about the complexity of the figures, respondents correctly interpreted the seasonal differences in distribution.

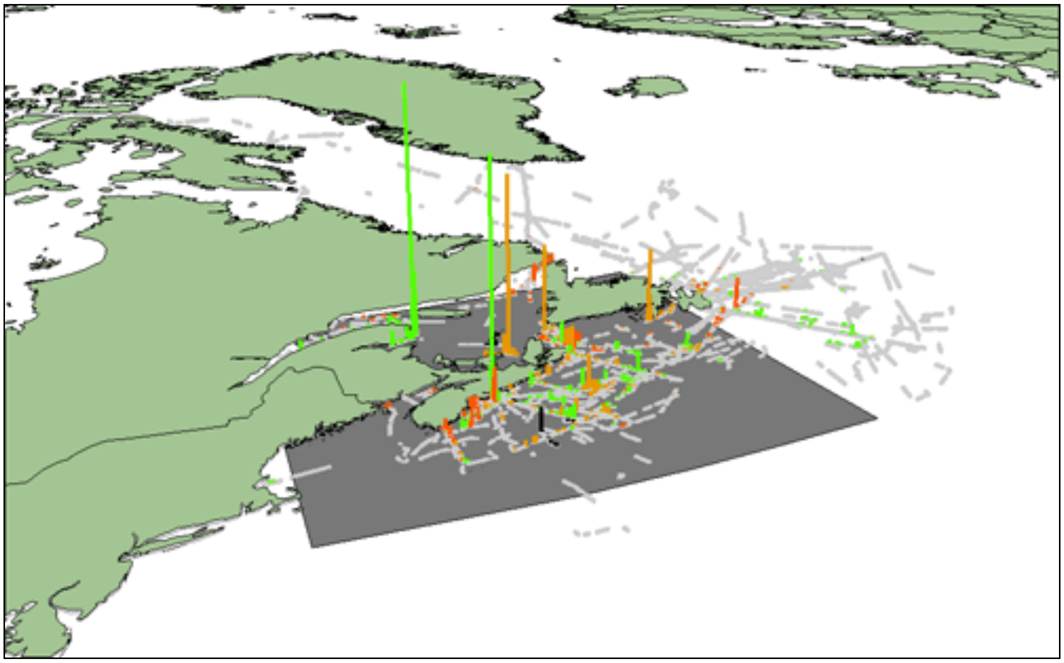


Fig. 3a.

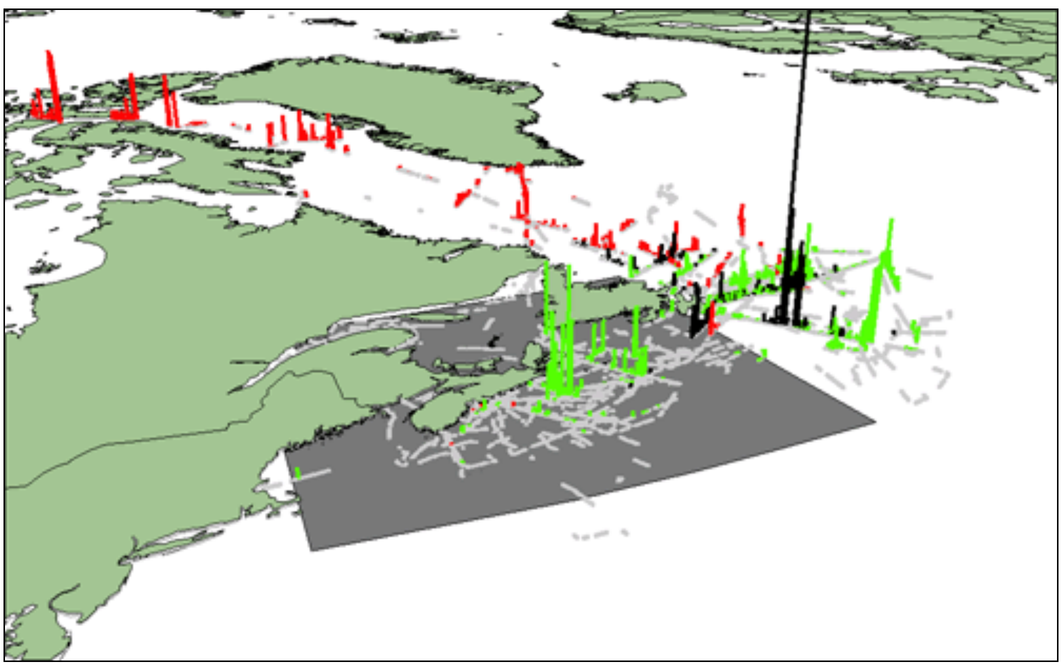


Fig. 3b.

Receptivity to Static 3D Maps

Figures 3a,b displayed: (1) locations of occurrence, (2) season (indicated by a solid symbol colour), and (3) survey tracks (light grey symbols). Relative abundance was represented by the degree of vertical exaggeration rather than by a 2D interpolation surface as with figs.2a,b. In this way a third dimension was added to this series of visualizations.

Several respondents described these visualizations as "useful", and declared that trends more easily discernable due to a lack in overlap in the points and the fact that relative differences were easy to see. One respondent reported the scale was too large to make out particular densities. One respondent noted the tendency for the Northern Gannet to occur in the shallower waters of the Gulf of St. Lawrence and in the vicinity of Halifax Harbour. Differences in spring and fall occurrence were noted, as were the tendency for wintering occurrences to be south of the Nova Scotian shelf break. One respondent declared that their interpretations would have been assisted by visualization of the survey tracks by season rather than pooled across the entire period.

Receptivity to Dynamic 3D Maps

Dynamic visualizations were constructed using the same 3D maps presented in figs.3a,b. While users could not directly control the scale or the vantage point for each, the visualizations took the viewer through an automated pan of the entire study area which allowed for multiple perspectives. As with the static 3D maps, locations of occurrence, season (indicated by solid symbol colour), and survey tracks (light grey symbols) were all plotted simultaneously, and vertical exaggeration was used to communicate differences in relative abundance.

The large size of the animation files meant that fewer participants were able to download and view the dynamic 3D maps, but the two who did were both declared the distribution of abundances to be more "precisely" shown and "significantly clearer" than the static 3D displays. These visualizations proved sufficiently stimulating that one participant drew attention to new knowledge not reported elsewhere: for example, the role of coastal PEI as an important area of the Gulf of St. Lawrence for fall Northern Gannet. Confidence in the interpretation of seasonal patterns was enhanced, and was reflected in the volume and quality of the inferences submitted to the website. Attention was drawn to summer occurrences of the Northern Gannet, something not remarked on following the viewing of previous visualizations, in the vicinity of coastal areas of the Gulf of St. Lawrence and Nova Scotia. One respondent attributed the lower summer abundances of the Northern Gannet to be due to birds being tied to their breeding colonies, an explanation/implication not raised in a previous visualization. Some deficiencies were identified: difficulty distinguishing orange and red bars, survey tracks that were not seasonally specific, and the absence of interpolation surfaces like those of figs.2a,b (implying this respondent found them useful).

Conclusions

Based on this small sample of respondents, multiple figures were not considered a drawback, and some users expressed a preference for multiple panels if it enlarged the maps or simplified their ability to discern main trends. The use of two different sets of symbols -- each with their own color classes – combined with a color-graded interpolation surface seemed to overload Figures 2a and 2b. This suggested that there was an upper limit to the complexity that could be effectively communicated using a single image. While the impact of 3D visualizations on interpretation time was not assessed, study respondents appeared to be more confident in their ability to make inferences. Spatial patterns were described for 3D visualizations that were not noted for 2D ones. In assessing the role of visualization in fostering collaboration it can be concluded that they readily provided "talking points" for discussion. Unusual occurrences were readily noted, and the implications of distributions centered, for example, on shelf breaks were highlighted for further inquiry. As part of a wider analysis of seabird distribution it is felt that these visualizations offered a critical, "non-parametric" summary of trends at different spatial and temporal scales.

References

Andrienko, G., N. Andrienko, Fischer, R., Mues, V., and Schuck, A. 2006. Reactions to geovisualization: an experience from a European project. International Journal of Geographical Information Science 20: 1149-1171.

Brunsdon, C., J. Corcoran, and G. Higgs. 2007. Visualising space and time in crime patterns: a comparison of methods. Computers, Environment and Urban Systems 31: 52-75.

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

DiBiase, D., C. Reeves, A.M. MacEachren, M. von Wyss, J.B. Krygier, J.L. Sloan, and M.C. Detweiler. 1994. Multivariate display of geographic data: applications in earth system science. Pages 287-312 *in* Visualization in modern cartography, A.M. MacEachren and D.R. Fraser Taylor (eds). Pergamon.

Durazo, R., Harrison, N.M., Hill, A.E. 1998. Seabird observations at a tidal mixing front in the Irish Sea. Estuarine, Coastal and Shelf Science 47: 153-164.

Environmental Systems Research Institute. 2006a. ArcMap version 9.2. Environmental Systems Research Institute, Redlands.

Environmental Systems Research Institute. 2006b. ArcScene version 9.2. Environmental Systems Research Institute, Redlands.

Gahegan, M. 2008. Multivariate geovisualization. Pages 292-316 *In* Handbook of Geographic Information Science, J.P. Wilson and A.S. Fotheringham (eds.). Blackwell Publishing, Ltd.

Ihaka, R. and R. Gentleman. 1996. R: A language for data analysis and graphics. Journal of Computational and Graphical Statistics 5: 299-314.

MacEachren, A.M. and Kraak, M.J., 2001, Research challenges in geovisualization. Cartography and Geographic Information Science 28: 3-12.

PHP Group, 2008. Pre-hypertext programming language. http://www.php.net.

Serra-Sogas, N., P.D. O'Hara, R. Canessa, P. Keller, and R. Pelot. 2008. Visualization of spatial patterns and temporal trends in aerial surveillance of illegal oil discharges in western Canadian marine waters. Marine Pollution Bulletin 56: 825-833.

Skov, H., and Durinck, J., 1998, Constancy of frontal aggregations of seabirds at the shelf break in the Skagerrak. Journal of Sea Research 39: 305-311.

Sun Microsystems, 2008. MySQL version 5. http://www.mysql.org.

Weimerskirch, H., 2007, Are seabirds foraging for unpredictable resources? Deep-Sea Research II 54: 211-223.

Wilhelm, S., C. Gjerdrum, and D.A. Fifield. 2008. Standardized protocol for pelagic seabird surveys conducted in eastern Canada from moving and stationary platforms. Draft report of the Canadian Wildlife Service.

Williamson, J., and C. McGuinness. 1990. The role of schemata in the comprehension of maps. Pages 29-40 *in* Lines of thinking: reflections on the psychology of thought. K.J. Gilhooly, M.T.G. Keane, R.H. Logie, and G. Erdos (eds). John Wiley & Sons, Toronto.

Yen, P.P.W., Huettmann, F. and Cooke, F., 2004a, A large-scale model for the at-sea distribution and abundance of Marbled Murrelets (*Brachyramphus marmoratus*) during the breeding season in coastal British Columbia, Canada. Ecological Modelling 171: 395-413.

Yen, P.P.W., Sydeman, W.J., and Hyrenbach, K.D., 2004b, Marine bird and cetacean associations with bathymetric habitats and shallow-water topographies: implications for trophic transfer and conservation. Journal of Marine Systems 50: 79-99.