Areography 2.0. A program to delimit distributional areas of species

A. C. Guasp, M. I. de Torres Curth, D. E. González

Universidad Nacional del Comahue, Unidad Postal UNC, 8400 Bariloche, Argentina

(Received 13 November 1995; accepted 17 November 1996)

Abstract

AREOGRAPHY is a computer program that offers quick and efficient methods to summarize geographical information of spatial point data on a distributional area map. By running this program, users can apply four methods for range compaction: mean propinquity, circular propinquity, grid method and minimum convex polygon. AREOGRAPHY 2.0 can also construct the minimum spanning tree between points. Animal ecologists can also use AREOGRAPHY 2.0 to fit home ranges. Maps can be overlapped and approximate areas can be calculated. Exact localities can be determined on maps by using the coordinates mapping option. Data files can be edited, imported from ASCII format, randomly created, and can include up to one thousand individual data. © 1997 Elsevier Science Ltd. All rights reserved.

Keywords: Area compaction methods; minimum spanning tree; mean propinquity; distributional areas; home range; geographical ranges

Software availability

Program title: AREOGRAPHY 2.0
Developers: A. C. Guasp, M. I. de Torres Curth and D. E. González
Contact address: Universidad Nacional del Comahue, Unidad Postal UNC, 8400 Bariloche, Argentina.
Tel.: +54-944-26368; Fax: +54-944-22111; E-mail: dgonzale@uncomai.edu.ar or guasp@cabctr1.cnea.edu.ar.

Hardware requirements: IBM PC compatible, 640 Kb RAM
Software requirements: Operating System MS-DOS
Program Size: 8000 lines
Cost: $90

1. Introduction

Knowledge about the region a certain species inhabits through time and space is an important tool in population and community studies. It can be used for the analysis of factors that inhibit or facilitate the settlement of a species and for making decisions about its conservation.

The information developed by field ecologists consists of sample localities where a species was or was not found, i.e. presence or absence localities. From this data, distributional boundaries can be delimited. The definition of region and boundaries will basically depend on two aspects: quantity and quality of information. Two very different maps of the distributional area of a single species can be obtained from the same data pool if constructed by two persons using different criteria. Several methods have been proposed in an attempt to unify this kind of work. Some of these methods (maybe the most accurate) have the disadvantage that they are impractical for handling a large amount of information. The objective of computerization is precisely to make abundant information an advantage instead of a difficulty.

With AREOGRAPHY 2.0, distributional areas can be compacted and maps overlapped using four different methods.
2. Area compactation methods

Field data consist of geometrical points distributed on a bi-dimensional region over which a reference system must be located. Points are pairs of x-y coordinates that we call collection localities or simply localities. Starting at these localities, ecologists can obtain maps like planar surfaces that represent the region where the studied species supposedly exists. Obviously the precision of these maps depends on the quality of the information. We will briefly describe the methods but we will not discuss in depth the advantages or disadvantages of their use, that are detailed in other works (Rapoport, 1982; Monjeau et al. in preparation).

AREOGRAPHY applies four methods to field data for distributional area compactation: 'mean propinquity' (Rapoport, 1982), 'circular propinquity' (Monjeau, 1969), 'grid', and 'minimum convex polygon' (Southwood, 1966).

Mean and circular propinquity apply graph theory. Before delimiting the region it is necessary to construct the minimum spanning tree, i.e. a convex graph without cycles. This connects all localities minimizing the inter-localities‘ distance. The object of constructing this graph is to figure out the mean minimum distance value, i.e. the arithmetic mean of lengths of the arcs in the graph. This is used as a radius in drawing a circle centered on each locality. By drawing only the circles, the obtained figure is the distributional area using the circular propinquity method. Mean propinquity takes a step further by smoothing the boundaries of the circular propinquity area. Basically, both methods result in figures with similar shape, but circular propinquity boundaries are whimsical. Natural boundaries are irregular, but this method does not add information either or better the mean propinquity approximation, and resultant maps are visually more complicated. It just shows a regular irregularity.

As the number of data points increase, results will be more precise, with an increased similarity of the figures obtained through both methods. Nevertheless, the circular propinquity approximation of the real area will always result in smaller values than the mean propinquity approximation.

Rapoport (1982) suggests that the field ecologist needs to have a criterion (maybe experimental) to select a radius of the circles. The arithmetic mean of arcs on the graph, or the arithmetic mean plus one, two or three times the standard deviation may be used.

The minimum spanning tree is offered by AREOGRAPHY 2.0 as an option. On the other hand, the grid method is a classical approach for the delimitation of species distributional areas. In the region where localities are defined, the program overlaps a grid, that can be defined by the user, between 2 and 100 divisions per side. Each time a point is inside a cell of the grid, the entire cell is considered occupied by the species, When a point is on a boundary, AREOGRAPHY 2.0 randomly assigns the point to one of the cells sharing the limit. Obviously, the finer the grid, the more similar is the resulting figure to the map with point localities.

The home range method delimits the minimum convex polygon that encloses all localities. This method is very sensitive to sample size and outliers. It has comparatively less precision and cannot deal with fragmented areas.

In practice, animal ecologists assign to each collection locality an influence radius, called activity radius, which depends on the information available on the species. Then the home range is constructed as the covering curve which wraps all the circles. AREOGRAPHY 2.0 offers users the possibility of determining the activity radius. The default radius is zero.

3. How algorithms work

The following paragraphs briefly describe how the main algorithms work in AREOGRAPHY 2.0. They constitute the minimum spanning tree algorithm, a previous step in constructing the circular and the mean propinquity maps, and the home range algorithm.

3.1. Mean propinquity and circular propinquity algorithms

Both methods are based on the construction of the minimum spanning tree. In order to do that, the program uses three auxiliary vectors V, C and G. The vector V, called 'vector of unconnected points', is initially filled with the coordinates of all data points. The vector C, called 'vector of connected points', is initially empty. In each iteration the program uses C to store the points that are connected to the graph G is used to store, in each iteration, the pair of connected points and the length of the connecting arc. Each element in this vector is a term (f,d) where 'f' is defined as the first point in the connection, 'd' is the point which f connects with, and ‘d’ is the Euclidean distance between these two points.

To construct the minimum spanning tree, AREOGRAPHY 2.0 performs the following steps:

Step 1. Define p as the pivot point. Let p = p1, the first element in V. Calculate the Euclidean distance between p and all the other points in V. Select the nearest point p2. Connect p1 and p2. Extract p2 and p1 from V and put them in C. Let p = p2. Store the connected pair of points and the distance between them in G.

Step 2. Calculate the Euclidean distance between the pivot and all the other points in V. Select the nearest point p2. Define d(p,p2) as the distance between p and p2. Calculate the distance between p2 and all points in C. Let p3 be the nearest point to p2 in C. Define
\( d(p_1, p_2) \) as the distance between \( p_1 \) and \( p_2 \). Compute\( d(p_1, p_2) \) to \( d(p_3, p_4) \):

1. If \( d(p_1, p_2) \leq d(p_3, p_4) \), then connect \( p_1 \) with \( p_2 \).
2. If \( d(p_1, p_2) > d(p_3, p_4) \) then connect \( p_1 \) with \( p_4 \).

Extract \( p_1 \) from \( V \) and put it in \( C \). Let \( p = p_3 \). Iterate this procedure until the last element in \( V \) is connected, i.e., all data points are in \( C \). For each iteration, save the connected pair of points and the distance between them.

Step 3. Calculate the mean minimum distance, i.e., the arithmetic mean between all distances in the connection sequence, named \( m \).

Step 4. With \( m \) as the radius, draw a circle centered on each data point. The resulting map is the distributional area compacted with the “circular proximity” method.

If the method previously selected is the mean proximity, the program will follow the next step to construct the mean proximity map.

Step 5. For each term in \( G \), compute the corresponding arc length \( d \) to \( m \). Only when \( d > 2m \), draw two parallel and congruent segments to the one that links the points as seen in Fig. 1.

3.2. Home range algorithm

This algorithm uses the same vector \( V \) of data points. The program will also generate another temporary vector \( G \) with pairs of connected points. Each element in this vector is a tuple \((f, i)\) where \( f \) is defined as the first point in the connection, \( i \) is the point which \( f \) connects with.

Step 1. The points in \( V \) are ordered following an ascendant criterion for the \( x \)-coordinate. If there are points with the same \( x \)-coordinate, the criterion is a descending order for the \( y \)-coordinate. This procedure allows to place the points that will be placed at the left of the screen in the first allocations of \( V \), and at the end of \( V \) those that will be at the extreme right.

Step 2. Connect the points vertically aligned (if there is any) to the left of the screen. Place the pivot \( p_0 \) at the point with the lowest \( x \)-coordinate (and the lowest \( y \)-coordinate if there is more than one).

Step 3. Calculate the angle \( \alpha \) \((-\pi/2 < \alpha < \pi/2\)) between the horizontal axis and the \( p_0, p \) segment for each \( p \) in \( V \) to the right of \( p_0 \). Select \( p \), the point that has the lowest angle. Connect \( p_0, p \). Let \( p_0 = p \). Repeat the procedure until the last point in \( V \) is reached. Put the connection \((p_0, p)\) to \( G \).

Step 4. Connect all points vertically aligned (if there is any) to the right of the screen. Place pivot at the point with the highest \( x \)-coordinate (and with the highest \( y \)-coordinate, if there is more than one), named \( p_{1r} \).

Step 5. Calculate the angle \( \alpha' \) \((\pi/2 < \alpha' < 3\pi/2)\) between the horizontal axis and the \( p_{1r}, p \) segment for each \( p \) in \( V \) at the left of \( p_{1r} \). Select \( p \), the point that has the lowest angle. Connect \( p_{1r}, p \). Let \( p_{1r} = p \). Repeat the procedure until the first point in \( V \) is reached. Put the connection \((p_{1r}, p)\) to \( G \).

Let \( r \) be the user-defined radius. If \( r \) is zero (default radius), trace a line joining each pair in \( G \). This connecting line is the minimum convex polygon. If \( r \) is higher than zero, then draw a circle with the radius \( r \) centered on each point reached by the connecting line.

Step 6. Draw the enveloping curve following the connection order.

4. Aereography’s numerical and graphic output

Once a data file is processed, AEROGRAPHY 2.0 constructs a graphic output, i.e., a map that represents the distributional area delimited with the selected method. The map is presented on a square graphic window that ranges between the minimum and maximum values of the \( x \) and \( y \) coordinates. When the method allows it, AEROGRAPHY 2.0 gives the mean minimum distance between points.

Once the map is ready, the possibilities are: (a) to map coordinates of any point on the map (AEROGRAPHY 2.0 provides a cursor that can travel over the whole map, showing while it moves the coordinates of its position, referred to the reference system); (b) to calculate distances between any two points on the map; and (c) to change the map scale. Generally, a change on the map scale, represents a loss of resolution, but allows maps obtained in the same region to be compared.

The area of the obtained figure can be calculated.
This process is approximate, because AREOGRAPHY 2.0 overlaps a very fine grid on the map and calculates the area. The area is presented as a percentage in reference to the whole graphic screen.

One objective of the work with distributional areas is to estimate both size and shape of the region of influence of a species. Sometimes it is necessary to compare distributional areas. Maps created from two different data files can be overlapped with AREOGRAPHY 2.0, provided they were compacted by the same method. Users obtain three maps, one for each data pool, on which it is possible to work, and a third map for the overlapping area. In the latter case, the program yields four area values: (a) the overlapping area, i.e. the coexistence region, (b) the proper area of each individual species and (c) the area that is not occupied. All values are expressed as a percentage.

5. Data manipulated by Areography

AREOGRAPHY 2.0 works with data files organized in x-y coordinates form. It offers a friendly file manager that allows users to create, modify, merge, and delete data files. It is also possible to import ASCII files. Creating data files allows to input field data or create randomly distributed points by selecting x-y coordinates. A maximum of 1000 data can be treated with AREOGRAPHY 2.0.

6. Two examples

Two methods from those discussed above are illustrated with real data. For ‘home range’, we use capture and recapture data for Abrothrix santhorhinus, a murid rodent from northern Patagonia. Data provided by Guthmann (unpublished data), were obtained using Sherman live traps on a 10 × 10 grid covering 1 ha during 10 consecutive days in March 1991. The capture coordinates data of the two pre-reproductive females (female no. 18 and female no. 24), were input into AREOGRAPHY 2.0 in two separated files. The object of the work was to decide if this species has a territorial behavior during different reproductive stages.

To construct the graphs with the minimum convex polygon method, the activity radius was fixed at 5 m, half the distance between traps. Figure 2 represents the graphic output of AREOGRAPHY 2.0 for the overlapped maps obtained from each data pool. The area values are shown on the left. AREOGRAPHY 2.0 displays four area values, labelled as (a) Map I Area, (b) Map II Area, (c) Overlapping Area and (d) Absence Area. The first two represent the percentage of the whole hectare that is occupied exclusively by one female. The third value, 3.91%, is the percentage of the whole map that is occupied simultaneously by both females. When expressed in square meters, it means that 391 m² are of common use by the captured females. Finally, the last area value is the fraction of the hectare that is not used by those females.

By analyzing these results, it can be concluded that pre-reproductive females (in this example) overlap their home ranges.

A second example where one method was applied is shown in Fig. 3. Data consist of the location of 62 redwood seedlings in a 23-m-square area (Dieggle, 1983). Point coordinates were introduced in the program in a reference system, ranging from 0 to 100. Here the obtained map shows the distributional boundaries as delimited with the ‘mean propinquity method’. The minimum spanning tree map is shown overlapping the first one. Mean minimum distance is 6.164 and the standard deviation 7.946, both referred to the map dimensions.

Once the map is constructed, the distributional area can be calculated. In this example, the estimate area is 33.18% as referred to the whole map, i.e. 175.5 m² approximately.

Fig. 2. AREOGRAPHY’s graphic output for the home range of Abrothrix santhorhinus. On the left proper and common areas for both maps are displayed.

Fig. 3. AREOGRAPHY’s graphic output for the distributional area of redwood seedlings using the mean propinquity method. Mean minimum distance, standard deviation and percent area are displayed on the left.
7. Discussion

While biogeography is interested in the delimitation of faunistic or floristic sets and in the origin of their different elements, areography focuses its attention on the form and size of the geographical ranges of species and other taxa. It represents the study of the spatial distribution of taxa, but at a geographical level (Rapport, 1987).

The delimitation of distributional areas of species, i.e. setting the geographical ranges where the species inhabits, offers some difficulties however. Two kinds of errors may occur place in this work. On one hand, the criteria can vary among the scientists. Very different results can be obtained from the same data pool. Then the approximations to reality can be very different. On the other hand, once the method is selected and the geographical range is delimited (using a more or less sophisticated method), the difficulty in getting a measure of the obtained surface subsists. As the precision of the method increases, the manual work becomes more difficult. Additionally, with more elaborated methods of area computation, the probability of making errors in the approximation of area values is presumably increased. Border irregularity is a reason for this fact. For example, area calculation by using planimeters gives serious errors if the user is not very precise, and besides it can be biased.

AREOGRAPHY 2.0 was developed in order to offer scientists working on geographical aspects of species distributions a reliable tool to delimit and calculate distributional areas in a quick, efficient, and unbiased manner.

Borders and shapes are two of the most important descriptors in the delimitation of distributional areas. However, sometimes it is not enough to describe the real use of the region by the species. To be sure that a given area is really the distributional area of a species, it is important to have a real knowledge of the ecology of this species. Therefore, the construction of the area with a collection of presence and absence data is just an approximation that will improve with a better sample design. A possibility is to construct a map with the presence data and superpose it on a second map obtained with absence data. Four regions will be determined: (a) A region with no information about the species, (b) two regions determined only by presence or absence information, and (c) a region where presence and absence maps are superimposed.

There are several possibilities to solve this conflict. A solution proposed in the bibliography is to split the conflict region using a certain criterion, for example Voronoi's tessellates (Ripley, 1981). This method proposes to divide the figure of the overlapping area in two by means of segments joining the points where the perimeter crosses. Each of the resultant figures is assigned to presence or absence maps by proximity. Another possibility is to reinforce the sampling in this area, and to reconstruct the map until the overlapped area can be neglected.

Another interesting issue in order to determine the activity area of animal species, is to overlap the obtained map with a fixed one so as to explain the absence data. It will allow the user to determine the sites where sampling should be reinforced, for example, when working at microscale, maps with vegetation patterns can be used, at medium or macroscale, topographic or climatic patterns can be employed. The possibility to input fixed maps with topographical or vegetational characters will soon be incorporated in an advanced version of AREOGRAPHY.

Acknowledgements

We would like to thank Dr E. H. Rapport for his continuous support, to Lic. N. Guthmann for rodent data, to Ing. H. Fontanini for his assistance in developing AREOGRAPHY's installation and the protection system of the software and to C. Ezcurra, J. A. Smith-Fluëk, and B. Draus for suggestions about the English manuscript.

References


