Hasanzadeh, Kamyar

IASM : Individualized activity space modeler

Published in:
SoftwareX

DOI:
10.1016/j.softx.2018.04.005

Published: 01/01/2018

Document Version
Publisher's PDF, also known as Version of record

SoftwareX 7 (2018) 138–142

Contents lists available at ScienceDirect

SoftwareX

journal homepage: www.elsevier.com/locate/softx

Original software publication

IASM: Individualized activity space modeler

Kamyar Hasanzadeh

Department of Built environment, Aalto University, PO Box 14100, 00076 Aalto, Finland

ARTICLE INFO

Article history:
Received 10 January 2018
Received in revised form 16 April 2018
Accepted 20 April 2018

Keywords:
Activity space
Neighborhood
GIS
Place exposure

ABSTRACT

Researchers from various disciplines have long been interested in analyzing and describing human mobility patterns. Activity space (AS), defined as an area encapsulating daily human mobility and activities, has been at the center of this interest. However, given the applied nature of research in this field and the complexity that advanced geographical modeling can pose to its users, the proposed models remain simplistic and inaccurate in many cases. Individualized Activity Space Modeler (IASM) is a geographic information system (GIS) toolbox, written in Python programming language using ESRI's Arcpy module, comprising four tools aiming to facilitate the use of advanced activity space models in empirical research. IASM provides individual-based and context-sensitive tools to estimate home range distances, delineate activity spaces, and model place exposures using individualized geographical data. In this paper, we describe the design and functionality of IASM, and provide an example of how it performs on a spatial dataset collected through an online map-based survey.

Code metadata

Current code version	*V 2.0.1*
Permanent link to code/repository used of this code version	https://github.com/ElsevierSoftwareX/SOFTX_2018_35
Legal Code License	*MIT*
Code versioning system used	*none*
Software code languages, tools, and services used	*Python, Arcpy*
Compilation requirements, operating environments & dependencies	*Requires Arcpy module installation*
If available Link to developer documentation/manual	https://github.com/kamyar68/-IASM/blob/master/Instructions%20and%20Tutorial/Instructions.pdf
Support email for questions	kamyar.hasanzadeh@gmail.com

1. Motivation and significance

Researchers coming from different disciplines have long been interested in analyzing and describing human mobility patterns. Understanding human mobility behavior has been especially important in social sciences, as well as in health geography, and environmental health promotion research. Therefore, an ample amount of research has focused on assessing the human exposure to different environmental characteristics to find empirical evidence of their associations with health and wellbeing [1–5].

Although some studies have been able to verify the importance of urban environment on people's lived experiences [6–8], the results from prior research on the environment–health relationship has often revealed marginal influence of environmental factors [9–13]. Researchers increasingly attribute the weakness of the observed associations to the misspecification of the geographical context [4,14–16]. In these studies, it is crucial to have a precise understanding of not only the extents of the geographical context, but also individuals' exposure to the places and their characteristics [17–19].

The growing literature from diverse fields such as public health, health geography, transportation, urban planning, and environmental psychology has brought along different terms for referring to similar concepts related to human activities and mobility in space. These terms include, but are not limited to, activity space [20], local activity space [19], home range [21], territorial range [22], action space [23], home zone [24], and neighborhood [18]. In this companion paper, we use the terms activity space and home range.

Activity space (AS) can be defined as a union of spaces where individuals are in direct contact with as a result of their mobility behavior and daily activities. As it can be implied from this definition, AS is usually deemed as an unrestrained part of the environment, whereas the home range involves a higher focus on

E-mail address: kamyar.hasanzadeh@aalto.fi.

https://doi.org/10.1016/j.softx.2018.04.005

the areas around an individual's place of residence [19]. Therefore, a home range –also known as a local activity space – can be regarded as a subarea of the whole AS which is often restrained by a locality or home range threshold [19]. Accordingly, while both concepts can be referred to using the more inclusive term AS, home range is sometimes used to underline the geographical scale and focus of a model.

ASs have been applied in a diversity of fields to describe the extents of mobility [25,26] or to identify places that people are exposed to [15,27]. However, despite the relatively broad application of ASs, their measurement has not been very advanced either conceptually or technically. That is why most of the efforts to this point are often limited to creating over-simplistic boundaries of these spaces. Buffers of different sizes [13], minimum convex polygons [28], standard deviational ellipses [29], and dynamic home ranges [30], are examples of such so-called *container* approaches [19]. However, some research show that ASs may be more than mere boundaries. Hasanzadeh and his colleagues [19] use a notion of place exposure and argue that not only exposure varies from one person to another in its spatial extents, but also can vary from place to place in its magnitude. The Individualized Residential Exposure Model (IREM) introduced by these researchers is based on the understanding that place exposure is an integral part of ASs and therefore should be part of the modeling.

Although the application of such advanced models have shown to improve the empirical findings, their use has remained limited in research. This can largely be due to the current complications of using geographical information systems (GIS) in this line of research. A geographic information system (GIS) is a framework for gathering, managing, and analyzing spatial data. The readily available GIS tools at this point, do not provide an easy way of implementing many of the discussed concepts related to the modeling of ASs. Moreover, many researchers who are potentially interested in use of such methods do not come from a spatial sciences background and may not have the required expertise to implement such ideas. The software described in this document is motivated by these complications and it aims to provide convenient tools that can support and promote use of more advanced AS models in research.

Individualized Activity Space Modeler (IASM), is a GIS toolbox providing tools for modeling, visualizing, and describing individual ASs. The first tool, Home range distance identifier, uses a mathematical algorithm to estimate an optimal distance for defining home ranges [30]. The second tool, home range modeler, creates an individualized and parametric model of home range boundaries [30]. The third tool, IREModeler, is an implementation of IREM [19]. This tool is an extended version of the home range modeler which not only estimates the extents of AS, but also models the variation of place exposures within the defined boundaries. Finally, the fourth tool provides a new method of modeling AS boundaries incorporating place exposure. The maximum exposure area estimator, identifies and extracts high exposure areas using a minimum exposure value defined by the user.

In this document, the tools' functionalities are demonstrated using data collected from more than 800 participants through an online map-based survey (public participation GIS). However, in order to provide a more focused image of what the toolbox creates, results from an individual randomly selected from the participants is presented in this paper. It is worth noting that the toolbox is also compatible with geographic data collected through other methods such as GPS and mobile phone tracking.

2. Software description

This software is made of four independent, yet closely related and complementing, GIS tools, which carry out specialized tasks for modeling ASs of individuals. The AS models in this toolbox are created using information on participants' home locations, activity points, and optionally travel routes and modes. The tools are developed in Python language using ESRI's Arcpy module and thus, are compatible with ESRI file formats. The modeling tools in this toolbox have the capability of modeling ASs of many individuals in a single run through an implemented iterative process.

2.1. Software architecture

In the following sections, we will provide the overall architecture and implementation details for each of the four tools in this toolbox.

2.1.1. Home range distance identifier

This tool uses distances from activity points to the individual home locations to provide an estimation of the home range distance. This is done through an implementation of Jenk's optimization method, designed to determine the best arrangement of values into different classes [30]. The algorithm seeks to minimize each class's mean deviation from the class average, while maximizing each group's deviation from the means of the other classes [31]. The optimum number of classes in this tool is identified using the Goodness of Variance Fit (GVF), while making sure that the chosen break distance includes at least 80% of activity points.

This approach has been previously implemented in empirical research [30]. More details on the theoretical aspect of the model can be found in the cited paper.

2.1.2. Home range modeler

This tool creates customized convex polygons for each individual using three parameters, D1, D2, and D3. D1 is the circular buffer distance representing the immediate neighborhood of an individual. The default value taken from literature is 500 m [30], however, this can be modified by the user. D2 is the distance used to create circular buffers around visited points. This is meant to represent an estimated area of exposure around each activity point and to take the fuzzy characteristic of ASs into consideration [4]. This value can be determined either by measuring the average size of spatial clusters in the dataset [30], or by calculating the average block sizes in the study area. However, these are some practical suggestions and the user can define this value using other applicable criteria. Finally, D3 is a threshold for defining the home ranges. In other words, if applied, any point beyond this distance will be left out of modeling. D3 can be either systematically defined using the Home range distance identifier tool described above, or by using other criteria defined by the user.

This model has been previously implemented in empirical research [30]. More details on the theoretical aspect of the model can be found in the cited paper.

2.1.3. IREModeler

This is an extended version of the Home range modeler described above, which creates a more spatially sensitive model of ASs using a notion of place exposure. In the first step, the spatial delimitation of the AS is defined using an integrated piece of code from Home range modeler tool. In the second step, the place exposures are estimated using home and activity points, frequencies or duration of visit, travel mode, and travel routes. To quantify the level of exposure, weights are assigned to each spatial feature. The weights for point features are calculated in terms of frequency or total time of visit per a defined duration. Accordingly, the highest weight is allocated to home points as the location people typically frequent on a daily basis or spend most of their time in. Depending on the frequency or duration of visit indicated for each visited point, weights are allocated to each of them. A weight for each route is calculated by its estimated frequency or duration of use and the used travel mode. This is operationalized as the geometric mean of destination and origin weights adjusted by the speed of the indicated travel mode.

Following calculation of weights for all geographical features, a sigmoid function is applied to normalize the calculated values. Consequently, all the values are interpolated using inverse distance as the distance decay function to produce a uniform surface. The model outputs are represented as individualized raster files with an estimation of exposure value for each pixel.

This is an implementation of IREM previously introduced in an empirical research [19]. More details on the theoretical aspect of the model can be found in the cited paper.

2.1.4. Maximum exposure area estimator

This tool can work with the IREM raster files created in the tool described above or other similar individualized exposure models. The tool can identify and extract areas of high exposure using a minimum exposure value defined by the user. The extracted areas will be delineated by a polygon for each individual and saved to a feature class i.e., a shapefile as the output. Since the exposure values estimated in IREM tool do no have an upper limit, the minimum exposure value is specified by the user in terms of percentiles. For example a value of 70 set by the user, will extract all the areas which have exposure values of higher than 70% of the maximum exposure limit for a given individual.

2.2. Software requirements

The software is developed using Python language and it employs tools provided within ESRI's Arcpy module. IASM is an ArcGIS toolbox comprising four tools designed to create or facilitate creating individual AS models. The tools can be executed either in Arcmap environment using the defined graphic user interface or independently from a Python IDE, given that ArcGIS is installed and licensed on the workstation. Fig. 1 illustrates an overview of the tool requirements and relationships. More details on the requirements and installation procedure can be found in the software documentation.

2.3. Software functionalities

This software is primarily aimed for place-based research and it can create different models of ASs, namely, home range, IREM, and high exposure area polygon. While trying to keep the software easy to use, IASM is created in a flexible and data-sensitive way allowing for adjustment of parameters for different study environments and contexts. Fig. 2, shows an example of implementation of IASM toolbox for an individual study participant. In this example, the model is created on a dataset collected through a map-based online survey conducted in Helsinki metropolitan area, in Finland. This dataset is collected from more than 800 participants and it includes information on individuals' home location, activity points, as well as frequency of visit and travel mode for each activity points. Travel routes are estimated using network analysis and distances are calculated using GIS measurements. In the following sections, we will present the outputs, and discuss tool-specific functionalities and the potential applications of the toolbox.

Home range distance identifier: This tool provides an estimation of home range distance based on dataset characteristics. Running this tool on our dataset from Helsinki area revealed 4km as an optimum home range distance.

This tool provides a data and context sensitive way of defining neighborhoods while avoiding making purely arbitrary decisions, which are common in research. The distance estimated in this tool can be applied in any research aiming to delineate and scrutinize the living environments of individuals.

Home range modeler: This tool provides delineations of individual home ranges using information on living location and activity places. By specifying a home range distance as D3, this model will only map the home range of individuals, whereas without a threshold distance, the whole AS will be mapped. In the example data from Helsinki area, a threshold of 4km – as identified by the home range distance identifier – was applied to the model. As shown in Fig. 2(B), the tool created individual home range polygons for all study participants. The polygons can considerably vary in shape and size depending on individual mobility patterns.

This tool can be used in a wide range of fields where an individual-specific geographical container is needed for analyzing environmental characteristics and investigating person–environment relationships.

IREModeler: This tool provides a more refined image of ASs through a place-based estimation of exposures. Fig. 2(C) illustrates IREM for the same individual from Helsinki Metropolitan area. This model is created based on the home range boundary and distance obtained from the two previous tools.

This tool can be applied in different fields such as environmental sciences, urban studies, and health studies, to evaluate the level of exposure to different environmental factors within individual ASs. Further, the sum of exposure values can be used as an indicative of intensity of an individual AS in general.

Maximum exposure area estimator: This tool provides a delineation of the high exposure areas within an individual AS to enable identification and scrutiny of the potentially most significant areas of the AS. The user has the possibility to define 'high' exposure by providing a percentile value. In the example data from Helsinki area, as illustrated in Fig. 2(D), we extracted areas of IREM that had an exposure value of more than 50% of the maximum range.

This tool can have a wide range of applications in different fields of research. One of the main features of this model is that it is an exposure-based model without having the complexity of working with raster files. As illustrated in Fig. 2(D), the tool allows researchers to identify the most intense areas of an individual AS and extract them as polygons. This facilitates a more focused approach to the environmental studies. Furthermore, this model can be used to study the complete geographical extents of an AS without facing some of the limitations of container approaches. Container approaches uninformed of actual individual place exposures, are prone to overestimation of exposure areas or arbitrary choice of thresholds, both of which may affect the findings.

3. Impact

In many disciplines such as environmental health, urban planning, transportation engineering, and social sciences, researchers need a working definition of geographical context in order to evaluate the human–environment relationships. Depending on the research and the level of complexity required, the geographical model can be as simple as a polygon or of more complexity such as a spatially sensitive exposure model. However, creating complex models of ASs has proven to be far from easy due to a lack of specialized spatial tools and expertise. As a result of these challenges, researchers have mostly relied on models which are often overly simplistic and do not adequately account for the individual differences in how people experience and use their environment.

The software introduced in this document is motivated by these challenges and it aims to provide a feasible way of creating complex AS models. IASM includes four GIS tools to create several types of AS models suitable for different research purposes. The home range modeler creates a functional and individual specific model

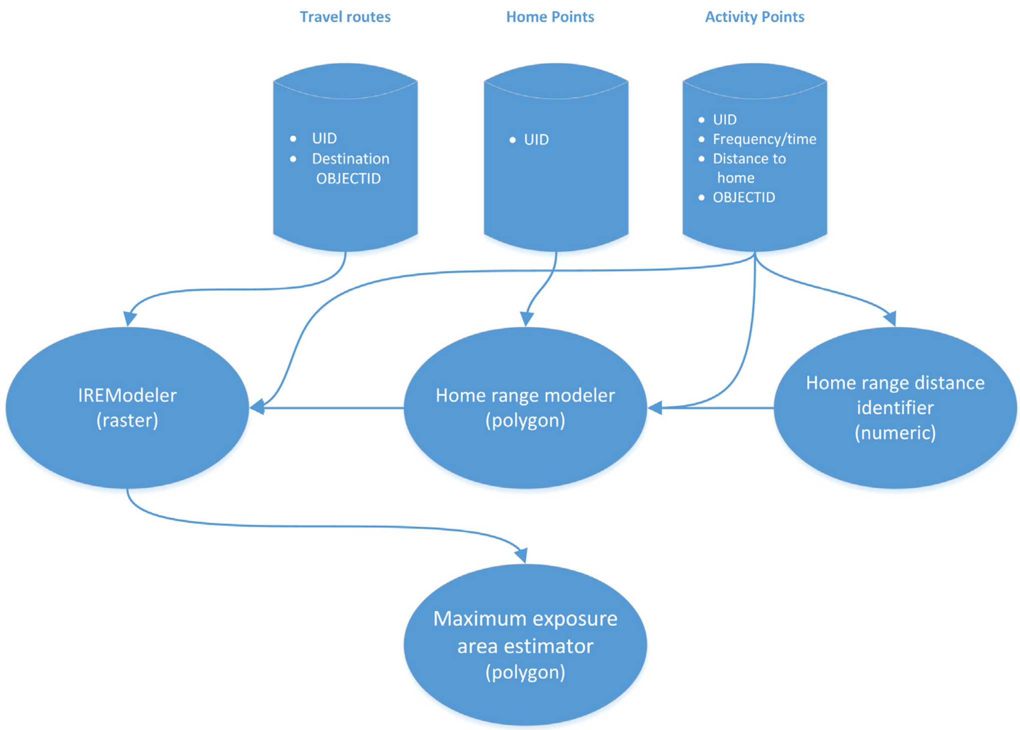

Fig. 1. An overview of the toolbox requirements, relationships, and outputs. Output formats are presented in parenthesis. UID: unique user identifier.

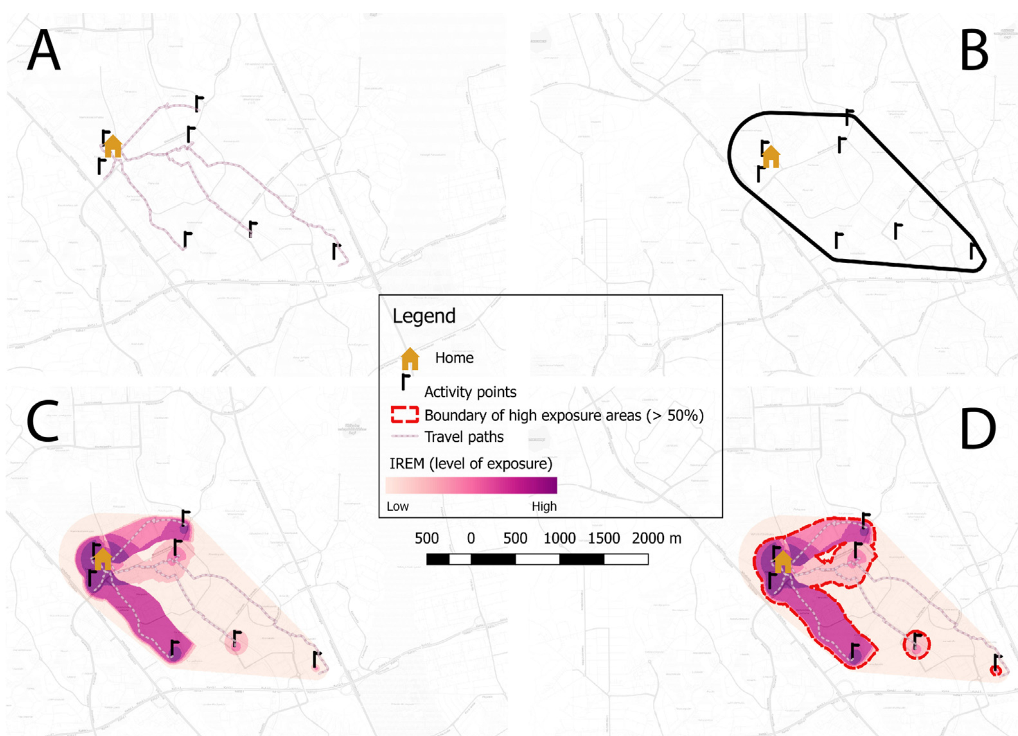

Fig. 2. Exemplary illustrations of model outputs for an individual from Helsinki area. A: Input data: home, activity points, and travel routes, B: Home range model, C: IREM, D: High exposure areas (>50%) extracted from IREM.

of ASs, which can serve as an improved geographical container for environmental studies. IREModeler provides users with a more realistic and complex model of ASs based on a place-based assessment of environmental exposures using a multilevel approach. Although the model is more complex and the output may be more difficult to work with, it can provide researchers with a more refined picture of ASs that can enable a more spatially sensitive understanding, and assessment of contextual effects. Further, the maximum exposure area estimator, lowers the complexity of IREM by extracting an easy to use polygon of areas of highest exposure. And complimentary to all these tools, the first tool, home range distance identifier, provides a systematic way of defining a locality threshold to make it possible to focus on the residential neighborhoods.

Despite the complexity of the models and operations, this software provides the users with convenient, iterative, and individual-based spatial modeling tools through a simplified interface. Furthermore, this versatile software provides users with the possibility to adjust model parameters to meet with different project-specific needs and preferences. The use of these tools can facilitate place and individual-based environmental research. This can help researchers find answers for a wide range of research questions related to use of urban environments as well as the contextual effects on active living, travel behavior, and public health.

4. Conclusions

Activity space (AS) is a concept used to describe spatial behavior of individuals and it is widely used in several fields of research. The software introduced in this companion document, Individualized Activity Space Modeler (IASM), is a GIS toolbox written in Python programming language and is designed to work with ESRI's ArcGIS software. IASM consists of four practical tools for modeling ASs of individuals. The first tool, Home range distance identifier, provides a systematic way of estimating an optimal distance for defining home ranges. The second tool, home range modeler, optionally uses the above value to create an individualized and parametric model of AS boundaries. The third tool, IREModeler, is an extended version of the home range modeler which creates a description of each individual AS using a raster file modeling the variation of place exposures within the space. Finally, the fourth tool, maximum exposure area estimator, extracts the areas of high exposure for each individual to provide a simplified exposure-based container.

Given the importance of AS concept to researchers from various disciplines and the existing spatial modeling limitations, this tool can help researchers understand individual ASs better and facilitate the assessment of contextual effects. This toolbox aims to provide a feasible way of applying more advanced AS models in empirical research. IASM can be widely used for spatial data collected through different methods such as public participation GIS, GPS, and mobile tracking.

Acknowledgments

This research is primarily funded by Finnish academy as part of PLANhealth project (13297753). We would also like to thank Finnish ministry of education and culture for partially funding this research.

References

[1] van Kamp I, van Loon J, Droomers M, de Hollander A. Rev Environ Health 2004;19:381–401.
[2] Mitchell R, Popham F. Lancet 2008;372:1655–60.
[3] Zenk SN, Schulz AJ, Matthews SA, Odoms-Young A, Wilbur J, Wegrzyn L, Gibbs K, Braunschweig C, Stokes C. Health Place 2011;17:1150–61.
[4] Perchoux C, Chaix B, Cummins S, Kestens Y. Heal Place 2013;21:86–93.
[5] Kyttä M, Broberg A. Wellbeing. Chichester, UK: John Wiley & Sons, Ltd.; 2014. p. 1–54.
[6] Dalgard OS, Tambs K. Br J Psychiatry 1997;171:530–6.
[7] Frank LD, Engelke PO. J Plan Lit 2001;16:202–18.
[8] Dye C. Science 2008;319:766–9.
[9] Pickett KE, Pearl M. J. Epidemiol. Commun. Heal. Heal. 2001;55:111–22.
[10] Diez Roux AV. Soc Sci Med 2004;58:1953–60.
[11] Subramanian S. Soc Sci Med 2004;58:1961–7.
[12] Adams RJ, Howard N, Tucker G, Appleton S, Taylor AW, Chittleborough C, Gill T, Ruffin RE, Wilson DH. Int J Publ Health 2009;54:183–92.
[13] Kytta M, Broberg A, Haybatollahi M, Schmidt-Thome K. Environ Plan B Plan Des 2015;43:34–57.
[14] Spielman S, Yoo E. Soc Sci Med 2009;68:1098–105.
[15] Kwan M-P. Ann GIS 2012;18:245–55.
[16] Vallée J, Le Roux G, Chaix B, Kestens Y, Chauvin P. Urban Stud 2014;1–20.
[17] Matthews SA. Communities, neighborhoods, heal. expand boundaries place. 2011. p. 35–55.
[18] Vallée J, Cadot E, Roustit C, Parizot I, Chauvin P. Soc Sci Med 2011;73:1133–44.
[19] Hasanzadeh K, Laatikainen T, Kyttä M. J Geogr Syst 2018.
[20] Schönfelder S, Axhausen K. Arbeitsberichte Verkehrs-Und Raumplan 2004; 258:1–40.
[21] Hasanzadeh K, Broberg A, Kyttä M. Appl Geogr 2017;84.
[22] Kyttä AM, Broberg AK, Kahila MH. Am J Health Promot 2012;26:e137–48.
[23] Dijst M. Netherlands J Hous Built Environ 1999;14:163–82.
[24] Biddulph M. J Urban Des 2003;8:217–41.
[25] Sherman JE, Spencer J, Preisser JS, Gesler WM, Arcury TA. Int J Health Geogr 2005;4:24.
[26] Perchoux C, Kestens Y, Thomas F, Van Hulst A, Thierry B, Chaix B. Soc Sci Med 2014;119:64–73.
[27] F.E.H., Reynolds DR. Econ Geogr 1971;47:36–48.
[28] Buliung RN, Kanaroglou PS. Growth Change 2006;37:172–99.
[29] Arcury TA, Gesler WM, Preisser JS, Sherman J, Spencer J, Perin J. Health Serv Res 2005;40:135–55.
[30] Hasanzadeh K, Broberg A, Kyttä M. Appl Geogr 2017;84:1–10.
[31] Jenks GF. Int Yearb Cartogr 1967;7:186–90.