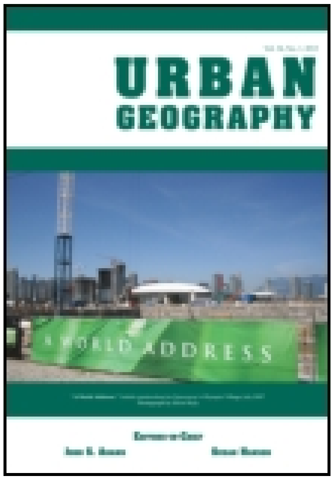

Urban Geography
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rurb20

Spatial Measures of Segregation and Gis[1]

David W. S. Wong [a]
[a] George Mason University
Published online: 16 May 2013.

RESEARCH NOTE

SPATIAL MEASURES OF SEGREGATION AND GIS[1]

David W. S. Wong[2]
Department of Geography
George Mason University
Fairfax, Virginia 22030

Abstract: Traditional segregation measures have limitations in discerning different patterns of population distributions. Spatial measures of segregation have been introduced, but have not been widely adopted partly because of the difficulties in using them. A recent effort is to implement several spatial segregation measures as additional GIS tools in a popular desktop GIS package so that researchers and practitioners not savvy in GIS can use these tools to compute spatial segregation indices. This paper provides a concise review of these measures and elaborates the new tools developed. [Key words: spatial segregation, dissimilarity index, spatial information, ArcView GIS.]

Most studies of segregation use primarily the dissimilarity index *D* (Duncan and Duncan, 1955). Even though the effectiveness of the *D* index in measuring segregation has been questioned, its usefulness and advantages were later reaffirmed from statistical and sociological perspectives (Massey and Denton, 1988). However, from the spatial perspective, the *D* index has serious limitations. In the past decade, a series of geographical studies has highlighted the inadequacies of the dissimilarity index. Subsequently, a set of spatial segregation measures based upon the two-group *D* index has been proposed. The multigroup version of the *D* index was also modified to become a spatial measure of segregation for multigroup comparisons. In addition, a spatial correlation index was introduced to indicate the level of spatial segregation among population groups. All these measures are regarded as spatial because either their formulations utilize geographical information explicitly, or the spatial interaction among population groups across areal unit boundaries is incorporated in modeling the level of segregation.

However, almost a decade after some of the spatial measures were first introduced, they have not been used extensively. Besides the legacy of the traditional measures such as *D*, part of the reason that spatial measures were not widely adopted is because these measures are difficult to implement and use. Most popular traditional segregation measures such as the *D* index and several others (Kaplan and Holloway, 1998) can easily be

[1] This project is partially supported by the National Institute of Health/Child Health and Human Development (NICHD) under National Institute of Health (NIH) grant number 1 R03 HD38292-01.
[2] Correspondence concerning this article should be addressed to David W. S. Wong, Department of Geography, George Mason University, Fairfax, VA 22030; telephone: 703-993-1212; fax: 703-993-1216; e-mail: dwong2@gmu.edu

Urban Geography, 2002, **23**, 1, pp. 85–92.

computed using generic software packages such as spreadsheet or database programs. These computing tools are more than adequate to perform the calculation of traditional measures even if the study region covers an extensive area and involves a relatively large database. But to compute even the simplest spatial index, a certain type of geographic information is required. If the study area involves only a few areal units, the geographic information may be derived manually. But for any realistic study using a data set of reasonable size, various types of geographic information have to be derived in an automatic fashion and Geographic Information Systems (GIS) technology becomes indispensable. Unfortunately, the development of GIS packages has not advanced to the stage that users can use standard built-in GIS functions to compute spatial segregation indices. Therefore, these spatial segregation measures are not widely used among researchers and practitioners even if GIS are available to them.

The purpose of this note is to report a recent effort supported by the National Institute of Child Health and Human Development (NICHD) of the National Institute of Health (NIH) (Grant Number 1 R03 HD38292-01) to implement a set of spatial segregation measures within ArcView, one of the most popular desktop GIS packages. These spatial measures were implemented as additional spatial analytical functions in ArcView, and these functions will be available to GIS users and anyone who is interested to apply these spatial measures to study segregation. The next section provides a condensed overview of the set of spatial segregation measures. Interested readers can refer to existing literature for more detailed discussions. Spatial measures for both the traditional two-group and multi-group settings will be included. Then this note provides an overview of the new spatial tools developed for calculating spatial segregation measures.

SPATIAL SEGREGATION MEASURES

Two-Group Measures

The dissimilarity index *D* advocated by Duncan and Duncan (1955) has dominated the population and urban literature in segregation studies for several decades. This index is easy to compute and its intuitive interpretations are favored by many sociologists and population researchers. Among several desirable properties, *D* ranges from zero to one, with zero indicating no segregation and one, indicating maximum segregation. The index has received strong support from recent findings that it is effective in capturing the evenness dimension, which is the most important dimension in measuring segregation (Massey and Denton, 1988).

It is true that *D* is effective to capture the evenness of population, but only to the extent that the spatial arrangement of population among areal units is not considered. As long as each areal unit within the study area is dominated by one group or the other exclusively, the *D* index will return a “1,” indicating a perfectly segregated situation. The dissimilarity index cannot distinguish the situation in which a region has a checkerboard pattern, with adjacent units occupied by two groups, from the situation in which the two groups form two subregions and each of which is dominated by one group exclusively. In both situations, *D* will return a “1,” indicating perfectly segregated patterns. This limitation of *D* and other traditional measures of segregation is attributable to the fact that interaction

TABLE 1.—SPATIAL SEGREGATION INDICES AND THEIR SPECIAL FEATURES

	Indices	Special features
Two-group measures	*D*	Aspatial
	D(adj)	Compares racial proportions of neighboring units
	D(w)	D(adj) with a boundary length element
	D(s)	D(w) compactness measures (P/A)
Multigroup measures	*D(m)*	Aspatial
	SD(m)	Uses composite population counts
	S	Uses ellipses to describe population groups and overlay ellipses

between population groups across unit boundaries is not considered as a variable to lower the level of segregation (Wong, 1993).

To overcome this limitation of traditional measures, a series of spatial measures for two-group analyses has been introduced (Table 1). Most of these spatial segregation measures were based upon Newby's argument that segregation involves spatial separation of population groups (Newby, 1982), and spatial separation reduces interaction among groups over space. In other words, spatial interaction between population groups can diminish segregation level, and therefore spatial measures should include not just the evenness of population distribution but also the magnitude of potential interaction between population groups across enumeration unit boundaries. Morrill (1991) introduced the first index of this type, *D(adj)*, to reflect the level of spatial segregation based upon the original dissimilarity index less the amount of potential interaction across areal unit boundaries. The amount of potential interaction between any pair of neighboring units (i and j in Fig. 1) is reflected by the difference in the racial mixes of neighboring units.

Subsequently, Wong (1993) modified the *D(adj)* index in several directions. He argued that the intensity of interaction across a boundary is not a simple function of adjacency, but likely the length of the shared boundary (d_{ij} in Fig. 1). The *D(adj)* index is slightly rewritten to incorporate a boundary-length component such that the potential interaction between neighboring units is scaled by the portion of the boundary length shared by the neighbors and the length of the entire boundary. Furthermore, Wong (1993) argued that the intensity of interaction between areal units is also a function of the size and shape, or the compactness of the two adjacent areal units. To incorporate the geometric characteristics of areal units into the modeling exercise, a compactness measure based upon the perimeter-area ratio (P_i/A_i in Fig. 1) was incorporated in the proposed *D(s)* measure. The compactness values of the neighboring units are averaged and then scaled by the maximum compactness value found in the entire study region. These two-group indices and their corresponding special features are summarized in Table 1.

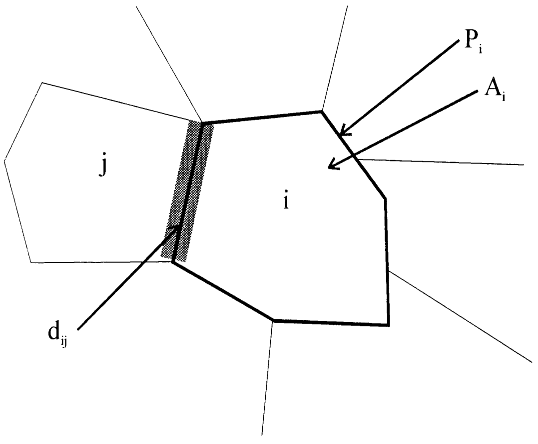

Fig. 1. Geographical and geometric elements used in various spatial segregation indices.

Multigroup Measures

All measures described above are for comparing two population groups. But in today's societies, multiethnicity is the norm rather than the exception. In measuring segregation, the two-group measures can still be used in a multiethnic setting by comparing all possible groups in a pair-wise manner (Morrill, 1995), but the comparison cannot be conducted for multiple groups simultaneously. The limitation of the *D* index to the two-group situation is removed by the introduction of a multigroup segregation measure. Based upon the concept of the dissimilarity index *D*, Morgan (1975) and later Sakoda (1981) introduced *D(m)*, a multigroup version of *D*. Detailed interpretations of the index can be found in Morgan (1975), and Wong (1998) further elaborated certain aspects of the index.

However, this *D(m)* multigroup measure of segregation shares the same limitation with the dissimilarity index *D*. That is, when populations in different areal units are rearranged spatially, the overall level of segregation will not change as indicated by *D(m)*. To overcome the limitation that spatial interaction between population groups across unit boundaries is not accounted for, Wong (1998) proposed to use the composite population counts instead of the original population counts to compute the multigroup index. The composite population count of a given areal unit includes all people in the neighborhood. This formulation implicitly allows people in the neighborhood to interact and mingle together as if in the same areal units. Using the composite population counts, the spatial version of the multigroup index, *SD(m)*, is derived, and it shares the same mathematical properties with *D(m)*.

Different from all the measures mentioned so far that rely on the concept of dissimilarity among population groups, another spatial segregation measure for multigroup situations was introduced based upon the concept of spatial congruence. If different groups have similar distribution patterns, they are not likely separated in a geographical sense. To evaluate the spatial correlation among multiple population groups, Wong (1999) suggested using the centrographic measure of a standard deviational ellipse to capture the overall spatial distribution of each population group. After multiple ellipses are derived for different groups, they are then compared and combined to derive an index of spatial segregation *S* based upon the ratio of the intersection and union of all ellipses. Special characteristics of these multigroup measures are also summarized in Table 1.

All spatial measures discussed above incorporate spatial information implicitly or explicitly in their formulations. Because GIS are efficient to store, retrieve and analyze spatial data from which values of spatial segregation indices are derived, therefore, it is logical to use GIS to compute spatial segregation measures.

GIS TOOLS FOR SPATIAL SEGREGATION MEASURES

It has been demonstrated that GIS are potentially powerful in segregation studies (Wong, 1996). But in the context of measuring spatial segregation, using GIS is necessary. Different types of spatial information pertinent to the computation of spatial segregation measures are captured by spatial data. GIS have to be used to extract information from the data. Specifically, the computation of the *D(adj)* index suggested by Morrill (1991) requires the neighborhood or adjacency information for each enumeration unit. To compute the *D(w)* index, the length of the shared boundary (d_{ij} in Fig. 1) for each pair of enumeration units and the perimeter (P_i) of each unit are required in addition to the adjacency information. For the *D(s)* index, the area of the enumeration unit (A_i) is also needed in addition to the perimeter in order to compute the compactness measure. To compute the multigroup *SD(m)* index, neighborhood information is used to derive the composite population counts on which the calculation of the *SD(m)* is based. The basic process to derive the ellipse-based *S* index is to fit a standard deviational ellipse (Ebdon, 1988) for each population group. The spatial information required to compute the ellipse involves the centroid locations of polygons assuming that the data are in the vector format and census areal units are represented by polygons. All these types of spatial information can be extracted from the spatial data through GIS.

A previous effort in implementing spatial segregation measures has been successful (Wong and Chong, 1998), but it involved two computing environments: a GIS package (ARC/INFO) loosely coupling with a statistical package (S-plus). ARC/INFO was used to extract spatial information from the spatial database and S-plus was used subsequently to handle the computation. The major drawbacks of this approach are that users have to use two different packages and information has to be passed between two computing environments. With the recent advances in GIS, it is preferable to implement spatial segregation measures completely within a GIS environment and take advantage of the powerful programming tools such that researchers (geographers, sociologists, and demographers) and practitioners (planners, policy analysts, and local activists) alike can calculate spatial segregation measures without knowing the ins-and-outs of GIS as long as they have access to a minimal GIS setup.

A recent development effort funded by the National Institute of Health (NIH) in the U.S. has implemented all spatial segregation measures discussed above within ArcView. ArcView is chosen because it is one of the most popular GIS packages. It also includes a powerful programming language, Avenue scripts, that works within the ArcView environment. The author had designed several algorithms to implement various spatial segregation measures. These algorithms were translated into Avenue scripts that can be executed

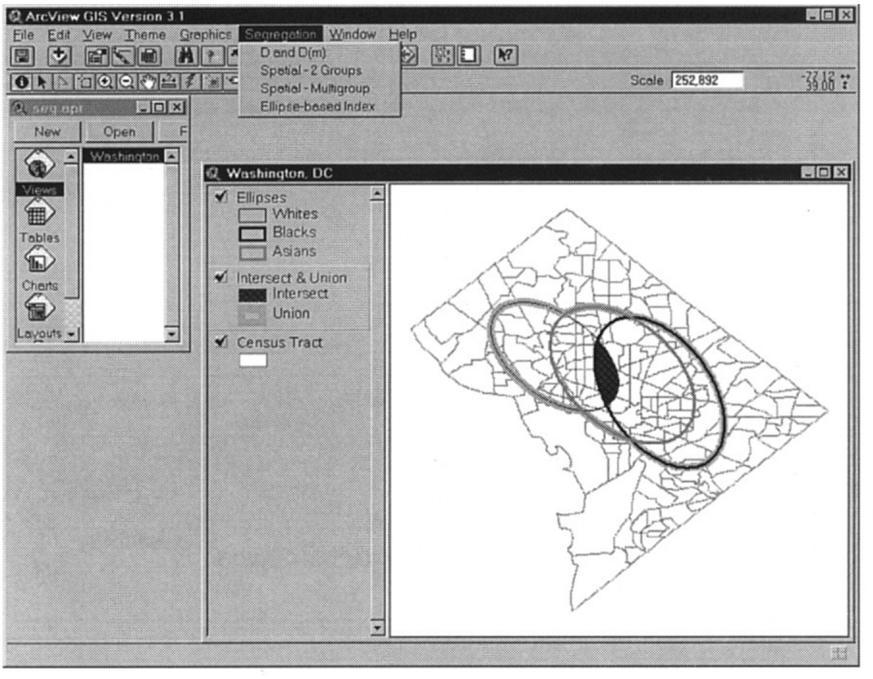

Fig 2. ArcView GUI with the new segregation tools and results fromWashington, DC.

within ArcView. Instead of having the users handle individual programs or scripts, these scripts are linked to the ArcView graphic user interface (GUI) as additional menu items. Users can select these items to execute the calculations of indices. Fig. 2 shows an ArcView user interface with the new menu items added.

In the modified user interface, a new menu "Segregation" was added. Under that menu, four items were included. Each menu item is linked to a specific Avenue script to compute a specific index or indices. For the first menu item, the traditional *D* index for two groups or *D(m)* for multiple groups will be computed. When this menu item is selected, users will be asked to select population groups from a drop-down window. If only two population groups are selected, the *D* index will be computed. If more than two groups are selected, the program will automatically compute the multigroup *D(m)*. The second menu item (Spatial-2 groups) activates the program to calculate the set of spatial segregation indices for the two-group situations. Users will be asked to choose two population groups and ArcView will compute and report the *D(w)*, *D(adj)*, and *D(s)* indices.

When the third menu item is chosen, the system will compute *SD(m)*, the spatial index for multiple groups after asking for population group inputs. The final menu item on the new ArcView user interface is for computing the ellipse-based *S* index. After selecting population groups for the analysis, the program will compute a standard deviational ellipse for each population group. Then, it will ask if the user would like to save all ellipses as geographical objects or features that can be displayed by ArcView subsequently. After this step, the program will overlay the ellipses as polygon features to compute the interaction and union of the ellipses. Then the program will ask if the user would like to save the intersection and union of all ellipses as geographical features for further analysis. If the user provides affirmative answers to these two questions, polygons representing the ellipses, and the intersection and union of all ellipses can be added to the map display as additional layers or themes. Figure 2 uses Washington, DC as an example. After the *S* index was calculated for Whites, Blacks, and Asians, three ellipses representing the three groups were stored and added to the View window. The polygons that show the intersection and union of the ellipses were also added into the View together with the census tracts.

These new functions in ArcView are incorporated into an ArcView project file named SEG.APR. Interested readers can download this project file from the website maintained by the author (http://geog.gmu.edu/seg). After the project file is opened by ArcView, users need to add data (as themes) of their study area into the View window and make the theme active before using the additional spatial segregation tools for the analysis. If multiple themes or layers are in the View window, the first theme will be used for the calculation in the default setting.

SUMMARY

Traditional measures of segregation are useful to compare the characteristics of population groups within the same enumeration units. But when the characteristics of population across enumeration boundaries are the concern, spatial measures of segregation can offer valuable insights. It has been unfortunate that spatial segregation measures have not been used more, in part because of the difficulties in using them. It is hoped that by implementing them in one of the most popular GIS packages, researchers, academics, and practitioners will have free access to the necessary tools to compute these spatial indices and more studies in the future will incorporate these spatial measures. The author recognizes the new development in ArcView and an effort is underway to convert the existing algorithms in Avenue into Visual Basic for Applications (VBA) for the newer version of ArcView and ArcInfo8.

REFERENCES

Duncan, O. D. and Duncan, B., 1955, A methodological analysis of segregation indexes. *American Sociological Review,* Vol. 20, 210–217.

Ebdon, D., 1988, *Statistics in Geography.* New York, NY: Blackwell.

Kaplan, D. H. and Holloway, S. R., 1998, *Segregation in Cities*. Washington, DC: Association of American Geographers.

Massey, D. S. and Denton, N. A., 1988, The dimensions of residential segregation. *Social Forces,* Vol. 67, 281–315.

Morgan, B. S., 1975, The segregation of socioeconomic groups in urban areas. *Urban Studies,* Vol. 12, 47–60.

Morrill, R. L., 1991, On the measure of geographical segregation. *Geography Research Forum,* Vol. 11, 25–36.

Morrill, R. L., 1995, Racial segregation and class in a liberal metropolis. *Geographical Analysis,* Vol. 27. 22–41.

Newby, R. G., 1982, Segregation, desegregation, and racial balance: Status implications of these concepts. *The Urban Review*, Vol. 14, 17–24.

Sakoda, J. N., 1981, A generalized index of dissimilarity. *Demography,* Vol. 18, 245–250.

Wong. D. W. S., 1993, Spatial indices of segregation. *Urban Studies*, Vol. 30, 559–572.

Wong, D. W. S., 1996, Enhancing segregation studies using GIS. *Computers, Environment and Urban Systems,* Vol. 20, No. 2, 99–109.

Wong, D. W. S., 1998, Measuring multiethnic spatial segregation. *Urban Geography,* Vol. 19, 77–87.

Wong, D. W. S., 1999, Geostatistics as measures of spatial segregation. *Urban Geography,* Vol. 20, No. 7, 635–647.

Wong, D. W. S. and Chong, W. K., 1998, Using spatial segregation measures in GIS and statistical modeling packages. *Urban Geography,* Vol. 19, No. 5, 477–485.