Investigating Fine-Scale Residential Segregation

by Means of Local Spatial Statistics

Itzhak Omer and Itzhak Benenson

omery@post.tau.ac.il

Department of Geography and Human Environment, Tel Aviv University

Investigating Fine-Scale Residential Segregation by Means of Local Spatial Statistics

Abstract

The paper presents a GIS-based approach to estimating individual household residential segregation based on three sources of information: a detailed geo-referenced dataset of family characteristics obtained from the 1995 Israeli Census of Population and Housing, subjective data on individuals' estimates of their house's neighborhood, and detailed GIS maps of urban infrastructure. The potential of the proposed approach is illustrated by studying Jewish-Arab residential segregation in the Yaffo area of Tel-Aviv. The combination of detailed objective and subjective geo-referenced data provide the basis for intensive fine-scale urban studies and local planning interventions.

1. Introduction

Geographic Information Systems (GIS) provide powerful and flexible tools for measuring, analysing and displaying urban residential segregation (Wong, 1997a; Wong and Chong, 1998; Lee and Culhane, 1998). The current state of GIS technology is far beyond the demands of the segregation studies, while the data availability is still very limited and, therefore, the majority of these studies continue to utilize population characteristics aggregated over given administrative partitions. Aaggregated data are insufficient for monitoring and predicting urban residential patterns and their relation to the physical environment for several reasons. First, the researches on the 'Modifiable Areal Unit Problem' (MAUP) clearly demonstrate that the conclusions reached on the base of aggregate datasets can change significantly when the same data are considered at different scales (Openshaw, 1984). Second, the relationships between aggregate and individual characteristics remain of uncertain validity – a problem known as the 'ecological fallacy' (Openshaw and Rao, 1995; Wrigley et al., 1996); hence, description of residential patterns is limited to relations between an 'average representative' of the population group and average characteristics of the environment. Third, one cannot directly compare aggregate data to those collected in field surveys by means of personal questionnaires; the latter effectively constrains our ability to investigate the relationship between the individual's perception of the urban environment and the residential patterns found (Cook and Crang, 1995).

The restrictions imposed by aggregate datasets thus have significant conceptual and methodological consequences for the study of urban residential patterns. They are partially responsible for the dichotomy between the humanistic-structuralistic and positivistic streams in human geography and often contribute to their conceptual persuasiveness in specific research contexts (Sayer, 1985). Quantitative positivistic approaches utilize aggregate data extensively, while admitting their inability to estimate variations in individual properties and their relationships to the immediate human and physical environment. Alternatively, idiosyncratic humanistic-structuralistic research concentrates on the unique individual and on the way he/she perceives the social and physical environment (Waldorf, 1993; Schnell, Benjamini, 1999). Research on human geography demand simultaneous use of subjective and objective information; thus detailed population data at the resolution of families and/or individuals are necessary. Aggregative data are consequently of virtually no use for studying humanistic-structuralistic concepts of attachment, sense of place, home area, and so forth.

The dichotomy between aggregate and individual approaches can be also observed in the way existing socio-economic GIS datasets are constructed and maintained. Governments and municipalities provide GIS maps and aggregate databases according to census partitions units. The corresponding individual data are usually also available, but very often unsuitable for geographical analysis. The latter results partially from privacy limitations, but mostly from the long-standing demographic tradition of spatially referencing individual data by the unit of administrative partition only, which makes such data useless for intensive geographical study (Weiner and Harris, 1998; Talen, 1999; Ceccato and Snickars, 2000).

The lack of high-resolution individual data began to be felt in the mid 1990s; in response, the population and housing census in Israel as in several European countries began to accumulate exact geo-referenced data on householders and households (ICBS, 2000). Broadly speaking, detailed GIS census data, by making possible the combination of objective and subjective information, such as residents' socio-economic characteristics on one hand and their judgments and perceptions of the neighborhood on the other, moderate the intensive-extensive dichotomy. The potential inherent in this merging is described by Batty (2000, p. 483) as follows: "Until now, urban geography has been dominated either by aggregate theories of pattern description or by more idiosyncratic case studies of individual behavior in an urban setting. … A new kind of fine scale geography is beginning to emerge from data which are sufficiently intensive to detect detailed patterns and morphologies but also sufficiently extensive to enable these patterns to be generalized to entire metropolitan areas."

This paper is based on detailed geo-referenced household data of the Israeli Census of Population and Housing for 1995 (ICBS, 2000), available for study at the Israeli Central Statistical Bureau (ICBS). The census data are organized within the GIS framework, which contains tables of the characteristics for each person, living in Israel for more than one month before the census and household the person belongs to (see ICBS, 2000 for further details). In GIS frame personal and household records are linked to the polygon representing the house, and, thus, enable the analysis of spatial relationships between individuals and households. The aim of a paper is to further develop the individual-based approach to measuring residential segregation recently suggested by us (Benenson and Omer, 2001), and to relate between high-resolution census data, the built-up environment, and residents' subjective knowledge of the social and physical environment in the vicinity of their homes. The paper begins with a brief presentation of the approach to measuring fine-scale segregation, described in detail in Benenson and Omer (2001). In what follows, we relate objective to subjective estimates of the state of the social and physical environment and determine the "proper" resolution for describing spatial segregation, illustrated by a study of Jewish-Arab residential segregation in the Yaffo area of Tel Aviv.

2. Global and local measurement of residential segregation

2.1 Segregation indices

To describe residential distributions, one must quantitatively characterize the distribution(s) of given population characteristics for one or more groups situated in the urban space. Many *global segregation indices* are used for the description of overall urban residential patterns and for comparing distributions of several groups according to one value. These indices are usually classified according to the dimension of the residential distribution(s) they are meant to disclose, for example, "evenness" of the residential distributions of two population groups or "exposure" of members of one group to members of another (Massey and Denton, 1988). Most popular is the index D of dissimilarity between residential distributions of members of two groups (Duncan and Duncan, 1955), which can be considered as a measure of evenness, while the Lieberson index of exposure (Lieberson, 1981) estimates the probability for a randomly located member of one group to be exposed to a member of another group at the same location. These two indices do not account directly for the neighboring relations between partition units, whereas those estimating “clustering” in fact do so in addition to group fractions and average population characteristics over spatial units (Morrill, 1991). The global measures are bounded to an administrative partitions established long time ago, insensitive to variation within the aggregate unit and ignore spatial relations between individuals located in boundary regions - all that makes their usage problematic for studying residential segregation (Benenson, Omer, 2001). As a by-product, aggregation entails inherent analytical dependency between different global segregation indices (White, 1983; Massey and Denton, 1988; Wong and Chong, 1998).

In most cases, researchers are still forced to use global indices due to access to aggregated datasets only. Nonetheless, the alternative *local* approach, which describes residential distributions by characterizing their properties at each location, has been developing during the last two decades. Until recently, applications of the local approach were quite scarce, but the situation is changing rapidly following the growing availability of new GIS datasets. According to the local approach, properties of a residential distribution are represented by means of *local indices of spatial association*, which could be considered as *local segregation indices*. These indices are based on the comparison of the characteristic of a given spatially located object and its neighbors (Anselin, 1995). In our case, the census GIS makes possible to compare the characteristics of the householders populating the given house to the characteristics of the householders populating the houses within the house’s neighborhood. Local indices are calculated for each distinct location, that is, for each house in our case; the size of the neighborhood serves as the parameter in each formula. One goal of this paper is to establish the size of the neighborhood that best fits those indices for the study of segregation phenomena.

Local indices can be differentiated into three standard groups according to their statistical properties – those based on local average(s), variance(s) or spatial (auto)correlation of the investigated characteristics (Benenson and Omer, 2001). The results obtained by different local indices of the same group are necessarily close, eliminating the need to use more than one index from each group. In this paper we use the Getis’ local index G* (Getis and Ord, 1992, 1996) based on local average, and Geary index K (Anselin, 1995; Getis and Ord, 1996) based on local variance. Formally, we consider 2D distribution of points with predefined neighborhood relationships and suppose that each point has some characteristic f; running ahead, the points represent centers of houses and f represents fraction of the ethnic group in a house.

The local Getis index G^*_i, calculated at location i regarding the neighbors within i's neighborhood U(i), is based on a weighted average of characteristic f over locations j within U(i), normalized for the average and variance of f over the entire area studied:

$$G^*_i = \Sigma_{j\in U(i)}(w_{ij}f_j - <f>)/(W_i\ s) \qquad (1),$$

where f_i denotes the value of characteristic f at i, U(i) denotes the neighborhood of i, <f> is an average, and s^2 is the variance of f over the entire area . The set of non-negative weights $\mathbf{w_{ij}}$ (where $w_{\mathbf{ii}} \equiv 0$ for each i) defines the a priori "influence" of neighboring locations on location i, where $W_i = \Sigma_{j\in U(i)}w_{ij}$.

It follows from (1) that high positive values of G^*_i at location i is obtained when the weighted mean of f over the neighborhood U(i) is essentially higher than the global mean; low negative value is obtained when the local weighted mean falls far below the global mean. Consequently, either high positive or low negative G^*_i values over continuous domains identify relatively homogeneous areas, where the characteristic f is close to either global maximum or minimum over most of the locations.

Geary index K_i estimates the local variance of characteristic f:

$$K_i = \Sigma_{j\in U(i)}w_{ij}|f_j - f_i|/(W_i s) \qquad (2)$$

where all the variables are as in (1). K_i is always positive, and its value represents the heterogeneity of the neighborhood U(i) of i.

Local Moran index I_i estimates local autocorrelation of characteristic f as a product between the value of f_i – <f> at location i and the weighted average of f_j – <f> over the neighborhood U(i):

$$I_i = (f_i - <f>)\Sigma_{j\in U(i)}w_{ij}(f_j - <f>)/(W_i s^2) \qquad (3)$$

where all the variables are as in (1). We do not use it in this paper because in our case it adds little to the results obtained with G^*_i and K_i, and present it here for completeness.

It is worth noting that each global index can be localized either by decomposition into local components or by natural reformulation of its analytical expression (Anselin, 1995; Benenson and Omer, 2001). One can explain the inherent relationship between different global indices on this basis, that is the latter are related if they use the same local components, whether as averages, variances or autocorrelations.

2.2 The scale of residential segregation

A latent but very important dimension of segregation indices is the partition of the urban space that determines the scale of the data presentation. A partition is necessary both to estimate the population fractions or characteristics of the population groups in each unit as well as to define neighborhood relations, necessary for local segregation indices (1) – (3). As each partition represents the same data, the results obtained on the basis of partitions at different scales are interdependent (Green and Flowerdew, 1996; Wong, 1997b).

We have already mentioned that until now, administrative partitions are almost exclusively used for the representation of population data. The units of these partitions – boroughs, electoral districts, statistical areas, and urban regions – are usually organized in a "container" hierarchy (Coffey, 1981; Ahl and Allen, 1996). This rigid and presumed hierarchy is insufficient to understand the individual's situation. The focus on the individual, proposed by the humanistic approach (Norberg-Schulz, 1971; Relph, 1976), argues that units of geographic space are determined by the personal experience. It rests on the claim that humans inherently need spatial as well as social segmentation (Tuan, 1977; Johnston, 1991). A hierarchy of spatial units exists, but it is defined by the individual's location and environment. For example, the house serves as the basic level, then the "home area" (Lee, 1968), the neighborhood, and so forth.

To formalize the humanistic perspective, we construct the hierarchy of places for *each* location. To be specific, for individual A, located at house H, we determine H as the elementary spatial unit of the lowest level L_0. The unit $U_1(H)$ at the level L_1 is defined as an individual's house H together with the set of closest neighboring houses. The rule for determining the houses that are the neighbors of the given house depends on the researcher's objectives; we consider one such rule below. The units of the upper levels are defined recursively; the unit $U_2(H)$ consists of $U_1(H)$ and the set of houses closest to one of the houses belonging to $U_1(H)$ and so on. The units $U_2(H)$, $U_3(H)$, $U_4(H)$ … of levels L_2, L_3, L_4 … are, then, the neighborhoods of H of increasing size, whose boundaries extend until the largest one, the city as a whole, is reached. For each individual A, therefore, we consider the series of embedded spatial units H, $U_1(H)$, $U_2(H)$, $U_3(H)$, … as representing different levels of a nested hierarchy with A's house H as the center.

The last step, necessary for the establishing individual-based hierarchy of urban space is the definition of the closest neighbors of house H. In this paper, we define the spatial proximity of houses based on coverage of Voronoi polygons (Benenson, Omer, Hetna, , forthcoming; Halls et al., 2001), constructed around their centroids. Two houses are considered as adjacent if their Voronoi polygons have a common edge, the criterion that reflects house visibility. To represent human perception of the neighboring house we account for too more conditions. Namely, we define the minimal (first-order) neighborhood of house H as a set of houses $U_1(H)$, where each $G \in U_1(H)$ satisfies three conditions (Fig. 1): G is adjacent to H; the distance between centroids of G and H is below a given threshold value (we set it equal to 100 m for the case of Yaffo); G and H are on the same side of a street if the street has two or more traffic lanes. In this way, we construct the units $U_1(H)$ of the level L_1.

The wider neighborhoods – units of the higher levels L_k, (k = 2, 3, …) - are defined recursively. Namely, neighborhood $U_k(H)$ of the order k is defined as $U_{k-1}(H)$ plus houses next to houses of $U_{k-1}(H)$ (Fig. 1):

$$U_k(H) = \{F \mid F \in U_1(G) \text{ AND } G \in U_{k-1}(H)\} \qquad (4).$$

The above definitions and consequent estimates demand intensive use of GIS algorithms. We use the MapInfo™ 5.0 GIS, Vertical Mapper™ 1.5 and MapBasic™ application working within a MapInfo environment for constructing Voronoi coverage, determining polygon adjacency, and constructing the set of first- and higher-order neighbors for each house. After the neighborhoods of all orders are defined, each local segregation index (1) – (3) can be calculated at each scale k. For example, the Getis index G^*_i generates the series of $G^*_{i,k}$, calculated regarding neighbors within neighborhoods of order k:

$$G^*_{i,k} = \Sigma_{j\in Uk(i)}(w_{ij}f_j - \langle f \rangle)/(W_i s) \qquad (5).$$

Below we set all the weights w_{ij}, $i \neq j$, equal to unit (to recall, $w_{ii} \equiv 0$). In this case the formulae for the series of Getis and Geary indices are as follows:

$$G^*_{i,k} = \Sigma_{j\in Uk(i)}(f_j - \langle f \rangle)/(s(N_k - 1)) \qquad (6),$$
$$K_{i,k} = \Sigma_{j\in Uk(i)}|f_j - f_i|/(s(N_k - 1)) \qquad (7),$$

where N_k is the number of houses within $U_k(i)$. As noted above, we do not discuss in this paper the problem of statistical inference and do not account for the significance of $G^*_{i,k}$ and $K_{i,k}$.

The hierarchy above helps to avoid uncontrollable variation of global indices of segregation calculated at different levels of aggregation (Jones and McEvoy, 1978; Morrill, 1991), and, thus, the modifiable areal unit problem (Stell and Holt, 1996; Carrington and Troske, 1997; Wu and Sui, 2001). The estimates of segregation over neighborhoods of increasing size straightforwardly shed light on changes in an individual's segregation state with respect to different environments perceived simultaneously. In this paper, we concentrate on local indices as reflecting the relationships between householders and their neighbors and do not consider relation between the values of indices calculated at different hierarchical levels and indices' statistical properties, both still lack investigation.

As stated, we define neighborhoods on the base of adjacency between Voronoi polygons, which reflects visibility, whereas the traditional definition is based on distances between houses. A GIS environment allows relating these definitions to one another. The calculations, based on coverage of Voronoi polygons of Yaffo's houses (Fig. 1), provide the following estimates: mean area of the Voronoi polygon of a house (L_0 – unit) equals 1,500 sq. m.; mean area of neighborhoods U_1 – U_4, that is, units at levels L_1 – L_4 equals 6,900, 21,000, 69,000 and 110,000 sq. m., respectively. In terms of proximity, the houses of the minimal neighborhood are located at a distance of an order of $\sqrt{6900} \sim 80 - 90$ m from the vicinity of a central house, the houses of the $U_2(H)$ at a distance of $\sqrt{21000} \sim 130 - 150$ m and so on. Below, we compare the size of formally defined neighborhoods $U_i(H)$ with the size of the neighborhoods subjectively perceived by Yaffo inhabitants. Given these results, we then selected the levels of the individual-based hierarchy that best fit the descriptions of individual residential segregation.

We now illustrate the above local approach by applying it to the estimation of ethnic segregation in Yaffo.

3. Arab-Jewish residential distribution in Yaffo

3.1 Background and data

Yaffo, a region in Tel-Aviv, is spread over an area of approximately 7 km^2; it has 39,000 residents. In 1995, the Jewish majority comprised about 70% and the Arab minority the other 30% of the population. Until the 1948 war, Yaffo was an Arab city of 70,000. After the War, only 3,000 remained, almost all of whom where concentrated within the small Adjami neighborhood. Jewish immigrants came to populate the other parts of Yaffo (Portugali, 1991; Omer, 1996). During the period 1948-1995, the Arab population of Yaffo continuously grew and spread throughout the region, whereas the Jewish majority gradually dwindled (Omer, 1996).

The main source of data for the study is the Israeli Census of Population and Housing for 1995 (ICBS, 2000). Unlike the previous census, the 1995 data are geo-referenced at the level of householders. That is, one field of the database record for each householder contains a unique identifier of the polygon representing the building in which he/she lives. The ethnic residential distribution we investigate is based solely on the householder’s “religion”. For the purposes of the current study, we combined the categories “Moslem”, “Druze” and “Christian” Arabs into one category, “Arab”. The census GIS also contains layers of streets, statistical areas, housing land use and open spaces (ICBS, 2000), which make it possible to apply the above definition of the neighborhood. In addition, we integrated subjective data regarding the perceived environment of 36 individuals, obtained via personal questionnaires in Yaffo in 1999 into the GIS.

3.2 The scale of perceived neighborhoods in Yaffo

The proposed algorithm for constructing increasing embedded units $U_i(H)$ is intentionally very detailed and, consequently, the number of levels of the hierarchy of spatial units is very high. Nothing but intuitive arguments can be used to select the levels for specific study or judge the importance of the obtained estimates of homogeneity/heterogeneity at the levels of neighborhoods of the first, second or higher orders. The only way to substantiate such judgments is to identify salient scales based on a study of individual perceptions of the urban environment.

As mentioned, the blending of qualitative and quantitative information within GIS applications is still being developed (Weiner and Harris, 1998). Few studies are available that concentrate on residents' judgments of residential areas and their perceptions of the borders of those neighborhoods (Talen, 1999; Ceccato and Snickars, 2000). The goal here is to decide what scales are relevant for measuring residential segregation and which are important from the individual's viewpoint.

To obtain information on perceived neighborhoods, one should ask residents to indicate their boundaries; several methods have been used for that purpose (Lee, 1973; Pacione, 1983). We asked the respondents to outline the neighborhood on a 1:4000 GIS-produced map, using their house as the center. Thirty-six randomly selected inhabitants (28 Jewish, 8 Arab), all having both Arab and Jewish neighbors in their immediate environment, were interviewed. They were asked (in Hebrew or Arabic) to outline two kinds of areas. First, following Pacioni (1983), they were asked, about the sense of "being at home (area)": "Please outline the area around your house where you feel yourself at home (the area which you feel most attached to)". Second, they were asked about their residential choice area: "Please outline the area around your house, which you had considered when you looked for an apartment." The answers that are the polygons of home and residential choice areas, sketched by the respondent on the map, were digitized and superimposed on a digital map of Yaffo. The maps of individuals' sketches and the distributions of home and residential areas size are presented in Fig. 2.

As we noted above, we compare the size of the units of the levels of neighborhood hierarchy built on the base of Voronoi coverage to the size of individual's estimates of home and residential areas. The mean size of the home area according to the individual estimates equals 10,500 sq. m, while the mean size of the residential choice area equals 71,000 sq. m. We did not find any significant differences between the size of the home or of residential area when controlled for ethnic affiliation, sex, or age of the respondent, maybe because of high variance of both estimate and low number of respondents.

These subjective estimates of neighborhood size allow us to determine the most important levels of the individual hierarchy. According to the estimates obtained of the mean area of units at levels L_1 - L_4, subjective estimates of the size of the home areas correspond to the size of units at levels L_1- L_2, while subjective estimates of the size of residential choice areas correspond to the size of units at levels L_3 - L_4 (Fig. 2). In this paper we apply Getis and Geary local indices for levels 1, 2 and 4, where level 4 is used in additional analyses of the influence of the built and social environment on the ethnic identity of residents in a house. It is worth noting, that the size of neighborhoods at levels L_1 - L_4 is much below the average size of a statistical area, the latter equals 260,000 sq m.

3.3 Fine-scale residential segregation in Yaffo and the segregation state of the individual

We are now prepared to estimate the segregation state of Yaffo inhabitants, as based on distributions of indices $G^*_{i,k}$ and $K_{i,k}$ constructed for houses versus neighborhoods of the relevant levels L_1, L_2 and L_4. Before doing so, we review, in brief, the importance of individual-based measures by comparing the residential patterns discerned according to the aggregate and the individual-based approach, using level L_4 for the latter (Fig. 3). To make the maps directly comparable, we use in Fig. 3 the fraction of a population group over the statistical area and neighborhoods $U_4(i)$. According to (5), Getis local index $G^*_{i,4}$ is simply a normalized average fraction of population group over the $U_4(i)$. Comparing the map of fractions of the Arab population in $U_4(i)$ to a map of these fractions by statistical area, one can identify an essential bias in ethnic representation in the latter. The statistical area outlined by a dashed red line in Fig. 3 is an extreme example: Here population of two spatially separated groups of houses, each populated almost exclusively by Arab or Jewish residents, appear as if living in an ethnically mixed statistical area.

To return to the local indices (Fig. 4), recall that values of $G^*_{i,k}$ close to the maximum or the minimum for all k distinguish domains of homogeneous populations, whether of Arabs or Jews (Fig. 4a). The rest of the Yaffo area, containing about 20% of the populated houses, where the values of $G^*_{i,k}$ are far from the extremes for all evaluated k, cannot be uniquely classified because moderate values of $G^*_{i,k}$ are obtained over these areas. Moderate value of $G^*_{i,k}$ is characteristic for neighborhoods heterogeneous in two senses. First, the fractions of the Arab population in all the houses of the neighborhood are similar and close to the global fraction of Arab residents in Yaffo; or second, the houses themselves essentially differ in the fraction of

Arab residents. We can better understand the segregation situation of the individuals inhabiting these areas by analyzing the local value of Geary index $K_{i,k}$ (Fig. 4b). According to (2), $K_{i,k}$ is low when the fraction of Arab population is the same in all the houses within $U_k(i)$. Thus, in addition to homogeneous Arab and Jewish areas, disclosed by extreme values of $G^*_{i,k}$, low values of $K_{i,k}$ mark the locations where the ethnic content of the individual houses within the neighborhood is similar. The latter is not typical of Yaffo's mixed areas where, according to the Fig. 4b, the values of $K_{i,k}$ over the areas of intermediate $G^*_{i,k}$ are high at almost all locations and for almost all orders 1 – 4. That means that the fraction of Arab residents in the houses within Yaffo's mixed area varies to a high degree within each neighborhood.

Accepting the finding that neighborhoods of order 4 and below are most important from the individual's point of view, we can describe the segregation situations of a hypothetical member A of, say, an Arab ethnic group located in house i in terms of the values of $G^*_{i,k}$ and $K_{i,k}$. The simplest situation is that of individual A residing in house i, which is located within a domain having the highest or lowest values of $G^*_{i,k}$ and close to zero $K_{i,k}$ for all k. Such an individual resides within a homogeneous environment both within the house, home area, and the residential choice area. High values of $G^*_{i,k}$ indicate that A is not exposed to the Jewish population until visiting locations beyond the area of residential choice. When A resides in a house located within an area of close to minimal $G^*_{i,k}$ (and, again, close to zero $K_{i,k}$) for all k, the situation is reversed, that is, A is found within a homogeneous Jewish environment within the house, home area, and the residential choice area. If A resides in house i, which is located within a domain having intermediate values of $G^*_{i,k}$ and far from zero $K_{i,k}$, then A is exposed to an ethnically mixed population within the home and residential choice areas and encounters houses with varying population compositions, from purely Arab to purely Jewish. Empirically, there are no situations of intermediate $G^*_{i,k}$ and close to zero $K_{i,k}$ in Yaffo.

3.4 Factors shaping Arab-Jewish distribution in Yaffo

The knowledge regarding local properties of the residential distribution makes possible to estimate the influence of the physical and human environment on residential segregation in Yaffo. We consider architectural style of the buildings as the factor of the physical environment and characterize the state of human environment by means of local segregation indices $G^*_{i,4}$ and $K_{i,4}$. Based on Yaffo GIS database, we estimate the correlations between the fraction of Arabs residents in a house on the one hand and the style of the building and the state of the neighborhood on the other.

The structure of Yaffo residential distribution, as revealed by $G^*_{i,4}$ and $K_{i,4}$, is already discussed above. Regarding houses' style, Omer (1996) has demonstrated that during 1961 – 1995, the Arab ethnic minority had diffused from the Adjami core outward (Fig. 5), and this process can be considered in relation to the architectural style of Yaffo's houses. Other studies likewise confirm the relation between residential segregation and house architectural type in mixed cities in Israel (Ben Artzi, 1980; Portugali, 1991) and abroad (Boal, 1982; Woods, 1980). In Yaffo, about 90% of the buildings can be characterized as either "Oriental", marked in the GIS database by 1, or "Block", marked by 0. The remaining 10% approach one of these two styles and are marked within the GIS database either by 0.9 or by 0.1. As one can see (Fig. 5), the style of the buildings in Adjami and in its outer ring is mostly Oriental, while Blocks are located on the Yaffo's periphery.

For Yaffo as a whole, houses' style and neighborhood ethnic structure strongly correlate. In 1995, 82% of the Arab population resided in Oriental houses and most of their neighbors were Arabs, while 72% of the Jewish population resided in Blocks and most of their neighbors were Jews (compare Fig. 4 and 5). Yet, this relation might be the consequence of fact that Arab families always stayed in Adjami Oriental houses and Jewish families occupied Blocks after the war of 1948 and did not migrate later. The influence of these factors should be tested then in the areas of mixed population only, where a process of residential redistribution really occurs. We select the mixed areas by means of local Geary index K, and consider houses with $K_{i,4} > 0.1$ as marking these areas. To quantify factors, which potentially shape residential distribution within this area, we test the correlation between the fraction of Arab residents in a house and two factors - house architectural style and the value of Getis index $G^*_{i,4}$ - for the same houses. The first equals 0.24 and the second 0.40 (for both $p < 0.001$), while the correlation between two independent variables - house style and Getis index $G^*_{i,4}$ is close to zero - 0.06 ($p \sim 0.04$). We can state, then, that both factors and almost independently influence the fraction of Arab residents in a house within an area of mixed population. Representing the same results based on R^2, we have $R^2 = 0.203$ for the regression of the fraction of Arab residents in a house on both variables, the latter being very close to sum of two values, $R^2 = 0.157$, obtained for regression on $G^*_{i,4}$ only, and $R^2 = 0.057$, obtained for regression on style of a house.

We can thus arrive at the preliminary conclusion that the fraction of Arab or Jewish residents in a house is positively related to both local ethnic environment and architectural style of the house. Hence, the residential diffusion of Arabs in Yaffo is directed towards neighborhoods dominated by Arab ethnicity and houses of Oriental style. This conclusion is arrived indirectly; it is not based on direct evidence because the residential distribution observed in 1995 is the result of dynamics, which initiated in 1948. The static data for 1995 do not allow exploration of the question whether the two factors considered, directly determine Yaffo's residential distribution. To test this suggestion, we recently developed agent-based model of Yaffo's residential dynamics, where the above factors determine the utility of a residence for migrating householder agents (Benenson, Omer, Hetna, forthcoming). Model householders, belonging to either the Arab or Jewish residential group, tend to settle in a house of the style they prefer and within a neighborhood populated by householders of similar ethnic identity. An agent-based model clearly demonstrates that weak but continuing effects of a neighborhood's ethnic structure and a house's architectural style are sufficient to enable for very good simulation of Yaffo's spatial residential dynamics in the 40-year period 1955-1995.

Conclusions

The paper proposes a GIS-based framework for the analysis of relations between social and physical properties of an urban system at the level of individuals. The analysis is based on three sources of the information: detailed GIS of census data, detailed GIS maps and databases of urban infrastructure and subjective spatial data collected with individual questionnaires. The framework combines objective and subjective data and, in addition to new tools for analyzing residential distribution, lessens the gap between the positivistic-extensive and the intensive-humanistic-structuralistic approaches.

The correspondence between the size of the perceived neighborhoods and the units of the individual-based hierarchy makes it possible to measure segregation at the resolution, determined by individuals themselves and recognize factors that influence residential behavior of the householders. By that, the structuration process that creates and preserves individuals' places (Gregory, 1989), can be investigated. The analysis of all kinds of information regarding Arab-Jewish residential distribution illustrates that the segregation process occurs at a fine scale of buildings and neighborhoods and depends on a specific constellation of human and physical characteristics. The using of local spatial statistics is essential for identifying and evaluating factors of Yaffo's residential dynamics. The immediate application of this and similar studies will eventually strengthen the social constituent of urban planning and management.

References

Ahl, V., Allen, T. F. H. (1996) *Hierarchy theory: a vision, vocabulary, and epistemology.* New York: Columbia University Press.

Anselin, L. (1995) Local Indicators of Spatial Association – LISA. *Geographical Analysis,* 27(2):93–115.

Batty, M. (2000) The new urban geography of the third dimension. *Environment and Planning B,* **27:**483-484.

Ben Artzi, Y. (1980) The creation of the residential pattern and residential mobility characteristics of the Haifa's Arab (in Hebrew). *Jewish-Arab Center, Haifa,* (Haifa University).

Benenson, I., Omer, I. (2001) Individual-Based Approach to Measuring Spatial Segregation: Formal Representation and the Case Study. In Schnell, I. (ed.) *Patterns of Segregation and Desegregation*: Avebury.

Benenson I., Omer I., Hetna E. (forthcoming) Entity-Based Modeling of Urban Residential Dynamics – The Case of Yaffo, Tel-Aviv, *Environment and Planning B.*

Boal, F. W. (1982) Segregation and Mixing: Space and Residence in Belfast. In Boal, F. W., Douglas, J. N. H. (eds.) *Integration and Division.* London: Academic Press, pp. 249-280.

Carrington, W. J., Troske, K. R. (1997) On measuring segregation in samples with small units. *Journal of Business & Economic Statistics,* 15(4):402-409.

Ceccato, V. A., Snickars, F. (2000) Adapting GIS technology to the needs of local planning. *Environment and Planning B,* 27:923-937

Coffey, W. J. (1981*) Geography: Towards a General Spatial Systems Approach.* London: Methuen.

Cook, I., Crang, M. (1995) Doing Ethnograpies, *CATMOG,* 58.

Duncan, O. Duncan, B. (1955) A Methodological Analysis of Segregation Indices, *American Sociological Review*, 20: 210 -217. Getis, A., Ord, J. K. (1992) The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis,* 24:189-206.

Getis, A., Ord, J. K. (1996) Local Spatial Statistics: an Overview. In Longley P., Batty M. (eds.) *Spatial Analysis: Modeling in a GIS Environment*. Cambridge (UK): Geoinformation International, pp. 269-285.

Green, M., Flowerdew, R. (1996) New evidence on the modifiable areal unit problem. In Longley, P., Batty, M. (eds) *Spatial Analysis: Modeling in a GIS Environment*. Cambridge (UK): Geoinformation International, pp. 41–54.

Gregory, D. (1989) Presences and Absences: Time-Space Relations and Structuration Theory. In Held, D., Thompson, J. (eds.) *Social Theory of Modern Societies: Anthony Giddens and his Critics.* Cambridge: Cambridge University Press, pp. 183-214.

Halls P J, Bulling M, White P C L., Garland L, Harris S. (2001) Dirichlet neighbors: revisiting Dirichlet tessellation for neighborhood analysis. *Computers, Environment and Urban Systems* 25:105-117

ICBS (Israeli Central Bureau of Statistics) (2000) Socio-Economic Characteristics of Population and Households in Localities and Statistical Areas. *Pub. No 8 in the 1995 Census of Population and Housing series*. Jerusalem: State of Israel, Central Bureau of Statistics Publications.

Johnston, R. J. (1991) People and places in the behavioral environment. In Boal, F. W., Livingstone, D. N. (eds.) *The Behavioral Environment.* London: Routledge.

Jones, T. P, McEvoy, D. (1978) Race and Space in Cloud-Cuckoo Lane. *Area,* 10:162-166.

Lee, C. M., Culhane, D. P. (1998) A perimeter-based clustering index for measuring spatial segregation: a cognitive GIS approach. *Environment and Planning B,* 25(3): 327-343.

Lee, T. R. (1973) Psychology and Living Space. In Downs R., Stea, D. (eds.) *The Image of the Environment.* Chicago: Aldine Publishing.

Lee, T. R. (1968) Urban Neighborhood as a Social-Spatial Scheme. *Human Relations,* 21:241-268.

Lee, T. R. (1978) Race, Space and Scale. *Area*, 10:365-367.

Lieberson, S. (1981) An Asymmetrical Approach to Segregation. In Peach C., Robinson, V. and Smith, S. (eds.) *Ethnic Segregation in Cities.* Croom Helm: London.

Massey, D., Denton, N. (1988) The Dimensions of Residential Segregation. *Social Forces,* 76:281-315.

Morrill, R. L. (1991) On the Measure of geographic Segregation. *Geography Research Forum*, 11:25-36.

Norberg-Schulz, C. (1971) *Social Order and the Theory of Class.* New York: Harcourt.

Omer, I. (1996) *Ethnic Residential Segregation as a Structuration Process,* Unpublished Ph.D Thesis, Tel-Aviv University, Tel-Aviv.

Openshaw, S., Rao, L. (1995) Algorithms for re-engineering 1991 census geography. *Environment and Planning A*, 27:425-446.

Pacione, M. (1983) The temporal stability of perceived neighborhood areas in Glasgow. *The Professional Geographer*, 35(1): 66-72.

Portugali, J. (1991) An Arab Segregated Neighborhood in Tel -Aviv: The Case of Adjami. *Geography Research Forum,* 11:37-50.

Relph, E. (1976) *Place and Placelessness.* London: Pion.

Sayer, A. (1985) The Difference that Space Makes. In Gregory D., Urry J. (eds.) *Social Relations and Spatial Structure.* London: Macmillan, pp. 45-66.

Schnell, I., Benjamini, Y. (1999) Socio-Spatial Lifestyles and Segregation, www.cybergeo.presse.fr, April, 1999.

Steel, D. G., Holt, D. (1996) Rules for random aggregation. *Environment and Planning A,* 28(6):957-978.

Talen, E. (1999) Constructing neighborhoods from the bottom up: the case of residents-generted GIS. *Environment and Planning B,* 26:533-554.

Tuan, Y. F. (1977) *Space and Place: The Perspective of Experience*. Minneapolis: University of Minnesota Press.

Waldorf, B. S. (1993) Segregation in Urban Space: A New Measurement Approach. *Urban Studies*, 30(7):1151-1164.

White, M.J. (1983) The measurement of spatial segregation. *American Journal of Sociology,* 88:1008-1018.

Weiner, D., Harris, T. (1998) Community-integrated GIS for land reform in Mpumalanga province, South Africa, *www.ncgia.ucsb.edu/varenius/ppgis/papers/weiner.html.*

Wong, D. (1997a) Enhancing segregation studies using GIS. *Computers, Environment and Urban Systems,* 20(2): 99-109.

Wong, D. (1997b) Spatial Dependency of Segregation Indices. *Canadian Geographer,* 41(2):128-136.

Wong, D., Chong, W. (1998) Using spatial segregation measures in GIS and statistical modeling packages. *Urban Geography*, 19(5):477-485.

Woods, R. I. (1980) Spatiotemporal Models of Ethnic Segregation and their Implications for Housing Policy. *Environment and Planning A*, 13:1415-1433.

Wrigley, N., Holt, T., Steel, D., and Tranmer, M. (1996) Analyzing, modeling, and resolving the ecological fallacy. In Longley P., Batty M. (eds) *Spatial Analysis: Modeling in a GIS Environment*. Cambridge (UK): Geoinformation International, pp. 25-40.

Wu, X. B., Sui, D. Z. (2001) An initial exploration of a lacunarity-based segregation measure. *Environment and Planning B*, 28:433-446.

Figure captions:

Figure 1: Definition of neighborhoods according to adjacency of Voronoi polygons. House H, marked in black, denotes the vicinity of the neighborhood; neighborhood $U_1(H)$ consists of H and houses marked by dark gray; neighborhood $U_2(H)$ consists of H and houses marked by dark gray and houses marked by light gray.

Figure 2: Neighborhood boundaries marked by the respondents and distributions of the perceived neighborhoods by size.

a. Home area
b. Residential choice area

Figure 3: Map of fraction of Arab population found in neighborhoods of level 4 superimposed on a map of the fraction of Arab population by statistical areas.

Figure 4: Local segregation indices at different levels of individual hierarchy in Yaffo

a. Getis index G*
b. Geary index K

Figure 5: Distribution of houses in Yaffo by architectural style.

Figure 1

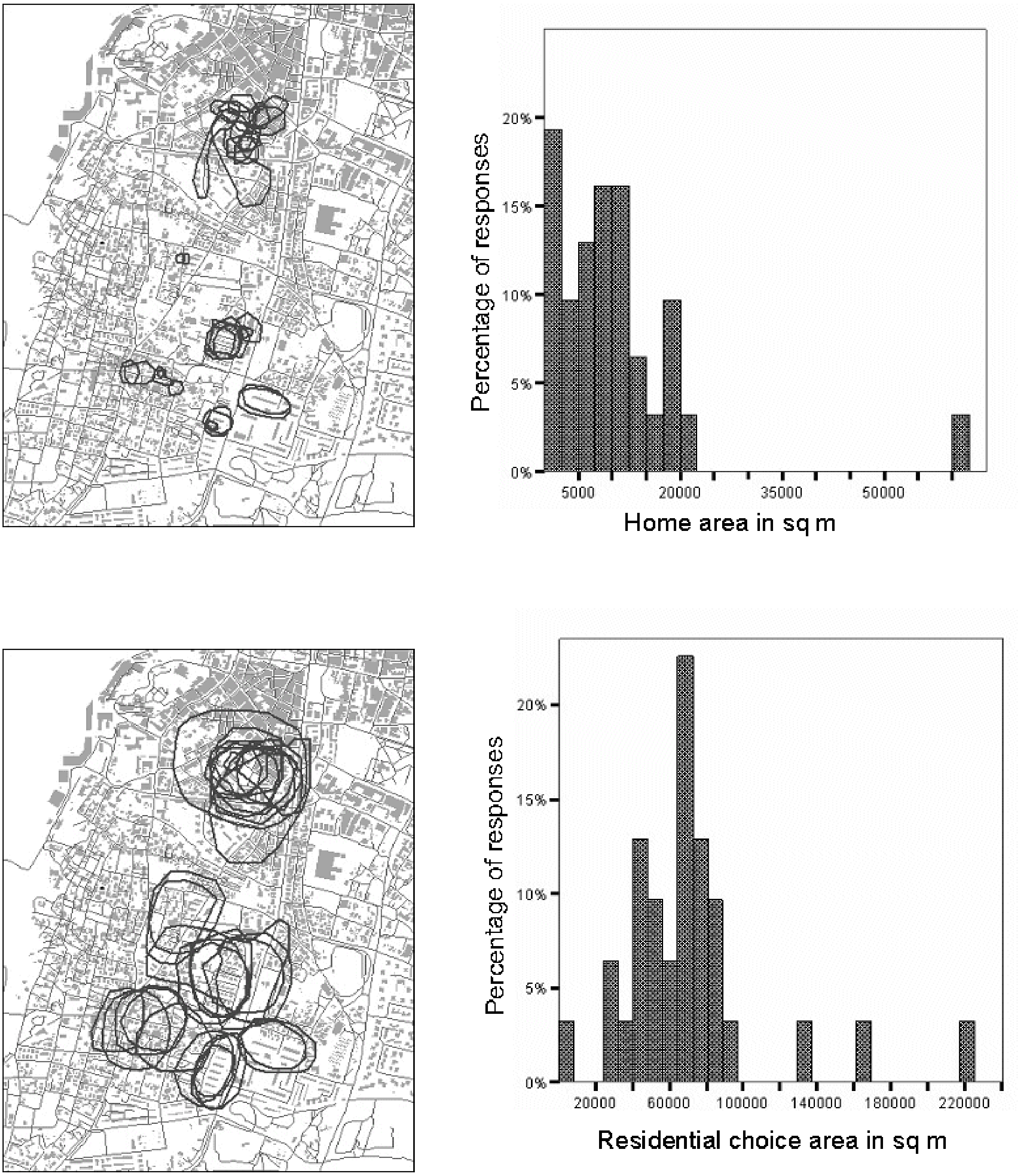
Percentage of responses
20%
15%
10%
5%
0%
5000
20000
35000
50000
Home area in sq m
Percentage of responses
20%
15%
10%
5%
0%
20000
60000
100000
140000
180000
220000
Residential choice area in sq m

Figure 2

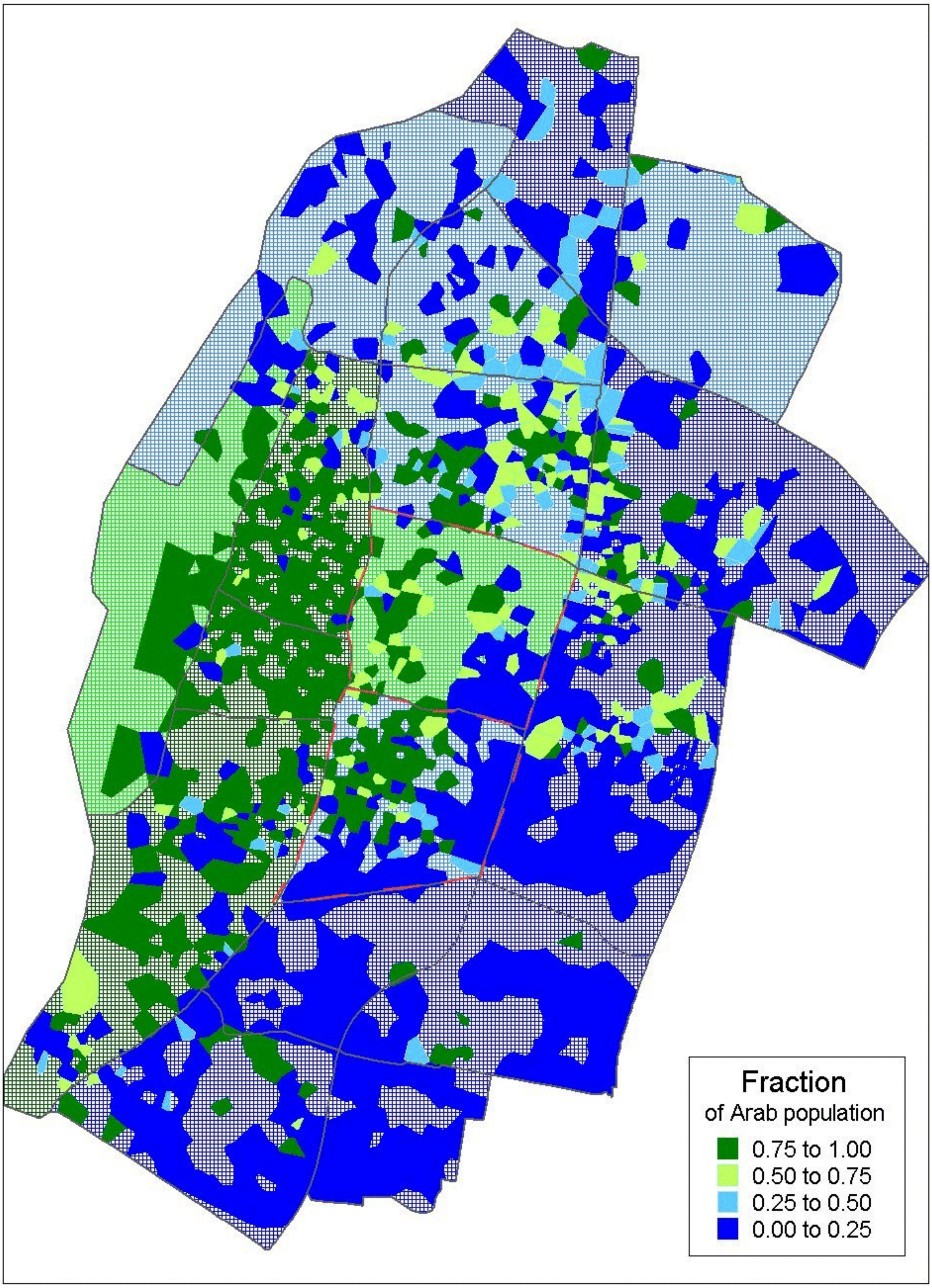
Fraction
of Arab population
0.75 to 1.00
0.50 to 0.75
0.25 to 0.50
0.00 to 0.25

Figure 3

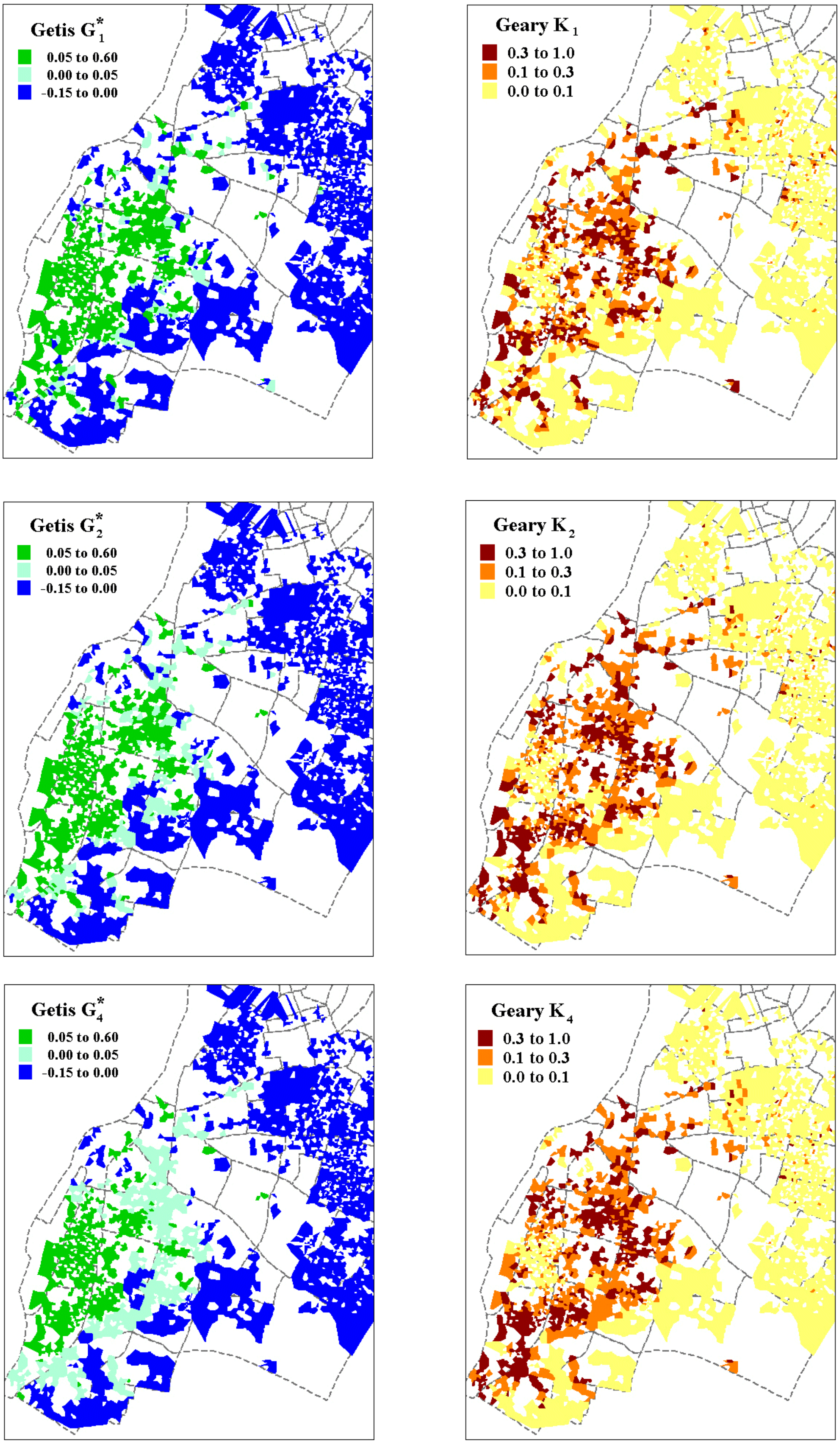
Getis G_1^*
0.05 to 0.60
0.00 to 0.05
-0.15 to 0.00
Geary K_1
0.3 to 1.0
0.1 to 0.3
0.0 to 0.1
Getis G_2^*
0.05 to 0.60
0.00 to 0.05
-0.15 to 0.00
Geary K_2
0.3 to 1.0
0.1 to 0.3
0.0 to 0.1
Getis G_4^*
0.05 to 0.60
0.00 to 0.05
-0.15 to 0.00
Geary K_4
0.3 to 1.0
0.1 to 0.3
0.0 to 0.1

Figure 4

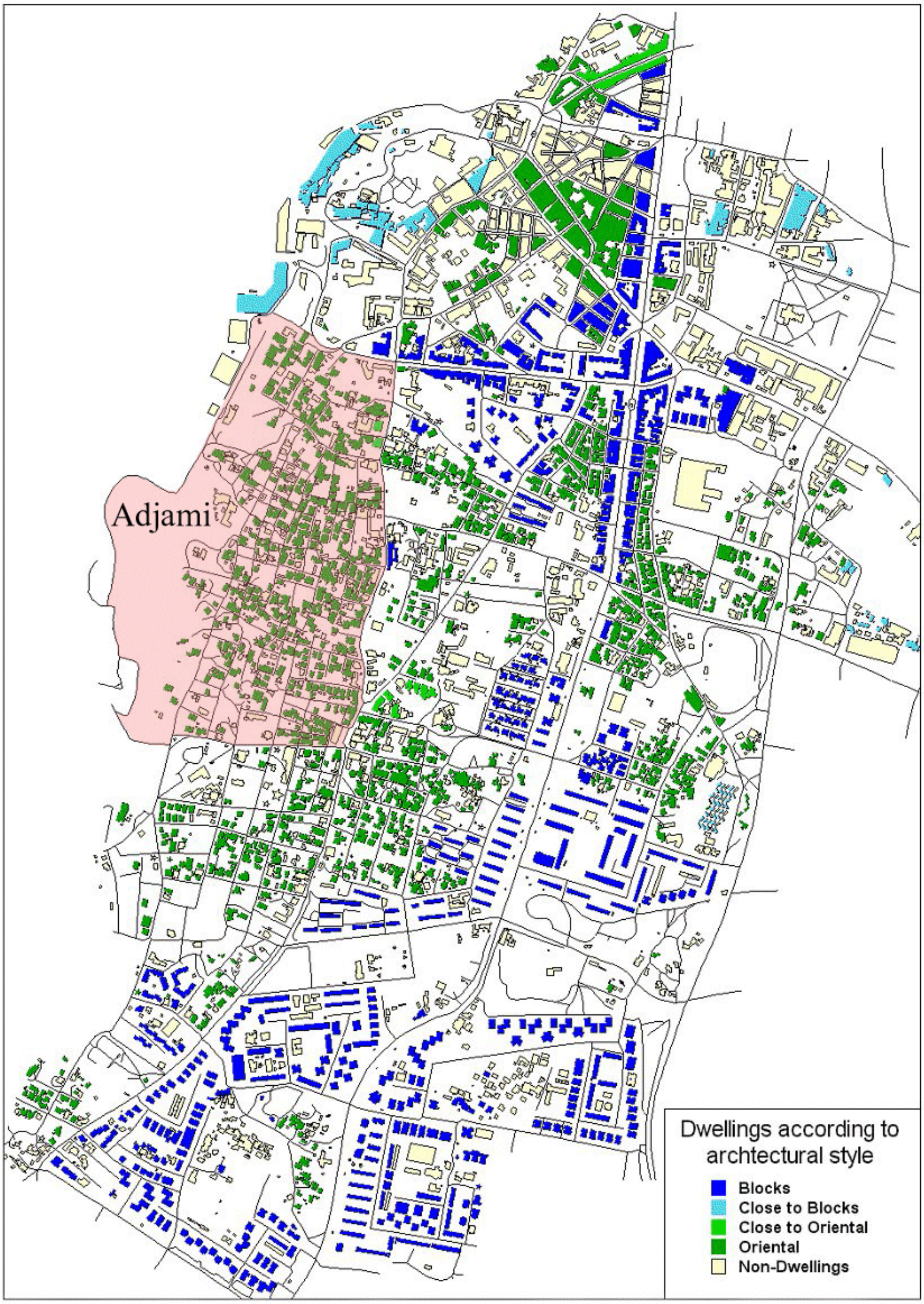
Adjami
Dwellings according to archtectural style
Blocks
Close to Blocks
Close to Oriental
Oriental
Non-Dwellings

Figure 5