Journal of Geographical Systems
https://doi.org/10.1007/s10109-023-00428-7

ORIGINAL ARTICLE

The importance of Arthur Getis to spatial demography

John R. Weeks[1]

Received: 9 June 2023 / Accepted: 17 August 2023

Abstract
My background as a demographer blended very well with Art Getis's knowledge of and research in geography. He taught me what geography is and he taught me how to analyze demographic data from a spatial perspective, especially in the use of the GIS approaches he was instrumental in developing. I am like a lot of people who owe a huge debt to Professor Getis's insight into the spatial aspects of the world around us. In this article I review the way in which he shaped and guided my own research and more generally that of spatial demography, by participating in it even as he was creating important new methods for spatially analyzing data.

Keywords AMOEBA · Getis-Ord statistic · Spatial analysis · Spatial demography

1 Introduction

Demography is an inherently spatial science. The three key demographic phenomena are births, deaths, and migration, the combination of which creates an age-sex structure that helps to define how a society works. Each of these interrelated aspects of demography have spatial (and temporal) components which, when understood, add to our knowledge of how and why change occurs. There are three spatial elements, in particular, that play a role in the different timing and pattern of demographic phenomena (Logan 2016). These are (1) space—demographic changes vary across a region as a function of differences in characteristics such as cultural, economic, and political history, natural environment, and built environment (infrastructure); (2) place ("neighborhood context"—broadly defined–matters when it comes to virtually all aspects of human behavior); and (3) scale (some things are more local in their effects than are others) (for more on this, see Weeks 2004, 2016, 2021).

It is one thing to be aware that spatial characteristics and attributes matter, but it is a very different thing to measure their impact, and the tools to analyze demographic data spatially were just becoming available in the early 1990s, thanks in large part

✉ John R. Weeks
jrweeks@sdsu.edu

[1] Department of Geography, San Diego State University, San Diego, CA 92182, USA

to the huge role that Arthur Getis played in this. Working with him meant that I was catapulted into the early days of what has since become mainstream demographic analysis that incorporates geographic information systems as key research tools. The tools start with georeferenced data, such as x and y coordinates, so that detailed geographic analysis is possible. Then we need statistical methods that can incorporate those geographic locations into the analysis. It is this latter set of tools to which Art Getis made his historic contributions. Those contributions might not have been so historic were it not for the fact that Art was an incredibly well-rounded human geographer—not just a gifted statistician. For evidence of this, you need only read his highly-regarded textbook, *Human Geography: Landscapes of Human Activities* (Getis et al. 1992). All of the social sciences share perspectives about human behavior, but the spatial part of it comes across much stronger and more rigorously when framed by a geographer.

Diving deeply into the georeferenced digital data from censuses and surveys has now become commonplace. Another source of geospatial data that is now regularly incorporated into demographic research is remotely sensed imagery. This allows us to define the physical space in which humans are living and working, in ways that are not possible in ordinary census and survey data gathering (see, for example, Ward et al. 2000). In my own work, my colleagues and I have used such information to characterize geographic areas within cities, and better understand the process of urbanization. Among our important findings were that fertility in many areas within Cairo, Egypt was as high as in rural areas, and these areas of higher than expected levels of reproduction could be identified from their environmental informality, observable from the analysis of remotely sensed imagery (Weeks et al. 2004).

Spatial analysis in the social sciences tests theories that where you are makes a difference in social attitudes and behavior, and that observed differences in the social world are not distributed in a spatially random pattern. The underlying logic is that each random variable (z) is associated with locational attributes (x and y). In spatial data analysis, the researcher uses spatial statistics to glean information from the x and y coordinates, whereas in classical statistical analysis the researcher ignores those coordinates (often not even realizing that they might exist). More to the point, in classical statistical analysis, the locational attributes are considered to be a nuisance, rather than representing useful information–spatial attributes are things to be gotten rid of, or controlled for, whereas in spatial data analysis they become objects of investigation (Anselin and Rey 1991).

There are two key and interconnected aspects of spatial patterns that we must account for: (1) spatial dependence (also known as spatial autocorrelation); and (2) spatial heterogeneity (also known as spatial non-stationarity). Spatial dependence takes us back to Tobler's First Law of Geography (which Getis often said was a key inspiration in his work)– everything is related to everything else, but near things are more related than distant things (Tobler 1970). Proximity is thus a predictor of some aspects of behavior. For example, everywhere we go in the world, fertility is lower among better educated women than among less well-educated women. But in Europe we find that better educated women have fewer children than similarly educated women in sub-Saharan Africa. Where you are matters. This is the essence of spatial dependence. But it is also true that even in the same region, less well-educated women may have fewer children than you might otherwise expect if they live near better educated women, because of the diffusion of attitudes about family size and knowledge of family planning. Indeed, this was an additional finding from our research in Cairo (Weeks et al. 2004). This is an example of spatial heterogeneity.

Spatial heterogeneity is really a special case of spatial dependence, in which not only are near things more highly correlated than distant things, but the strength (e.g., strong or weak) and perhaps even the direction of the relationship (e.g., positive or negative) varies from place to place. Spatial dependence does not always include spatial heterogeneity, but spatial heterogeneity always involves spatial dependence. Looking back at the example of fertility by level of education, we can see that knowing a woman's level of education will not let you automatically predict the number of children she has; rather it tells you that a better educated woman likely has fewer children than a less well-educated women in her part of the world. The explanation has to be sought in local cultural norms, which tend to be place-specific, meaning they have a spatial component.

2 The contributions of Arthur Getis to spatial demography

Our ability to draw conclusions about the spatial nature of demographic phenomena rests on the existence of the appropriate statistical algorithms, and this is where Art Getis, typically working in tandem with other colleagues, came to the rescue. Not coincidentally, my research into these relationships came shortly after Art Getis had created, tested, and published the details of the G_i^* statistic (Getis and Ord 1992; Getis 1994, 1995; Ord and Getis 1995). This statistic provides important information about what characteristics under investigation have a spatial component, what those spatial connections are, and how widespread the connections are.

First we test for the presence of spatial dependence in each of the independent/predictor variables by calculating Moran's I, using a spatial weights matrix (such as an inverse of squared distance weights matrix), where distance is typically measured between the centroids of the geographic areas for which we have data (e.g., census tracts or zip codes or neighborhoods measured in a variety of ways (Weeks et al. 2007)). For each spatially dependent independent variable, we then use the Getis-Ord G_i* statistic as a spatial filter to extract the spatially autocorrelated portion of that variable. The difference between the original variable x_i and the filtered variable $x^f{}_i$ is a new variable $x^{sp}{}_i$, that represents the spatial effects embedded in x_i (Getis 1995; Getis and Griffith 2002). These two variables, $x^f{}_i$ and $x^{sp}{}_i$ replace the original variable x_i in the regression equation to produce a spatially filtered regression model in which the contribution of the spatial and filtered (non-spatial) components of each variable can be determined by the beta coefficients in the resulting model. Note that these statistical methods are embedded in commercial software such as ESRI's ArcGIS Pro. Indeed, Lauren Scott, who is responsible for software support, education, documentation, and development of spatial statistics tools in ArcGIS, completed her Ph.D. in Geography in the SDSU/UCSB Joint Doctoral Program under the direction of Art Getis, and has made important contributions to the literature (see, for example, Scott and Janikas 2010).

We can also use the results of the G_i* statistic calculations to map the spatial clustering of variables. Positive values of G_i* that exceed a z-score of 1.96 (the 0.05 level of statistical significance) indicate spatial association of high values, whereas negative values of G_i* that are less than -1.96 indicate spatial association of low values. Identifying areas where values are clustered at high or low levels can then prompt us to investigate what differentiates those places in terms of factors such as the type of economy, prevailing religious orientation, or other sociocultural factors that influence demographic behavior and characteristics.

We also applied these methods to a rural area just to the north of Cairo, confirming the underlying conceptual framework that demographic behavior is a joint function of who people are and where they are (Weeks 2010). We showed that the Getis-Ord G_i* statistic and the Getis spatial filtering methods are very useful geospatial tools for uncovering the spatial patterns of human reproduction in a rural governorate in Egypt that had been assumed by many to be a spatially homogeneous area. We applied the G_i* statistic to dasymetrically mapped data from the 1976, 1986 and 1996 censuses of Egypt to show that there were very distinct spatial patterns in fertility over time in this predominantly rural region of the Nile Delta. The spatial filtering technique allowed us to conclude as well that the spatial component became more important over time as a predictor of fertility levels. Improvements in education represented a key feature of the changing rural social environment driving these spatial changes in fertility, almost certainly facilitated by increases in contraceptive utilization in the region.

As we were completing our analysis of the fertility transition in Egypt, one of our collaborators, Allan Hill from Harvard University, had begun a UN-funded household survey of women's health in Accra, Ghana. That led to discussion about how we could apply the kind of spatial analysis we had just completed in Cairo to what was happening in Accra. In 2004, he and I and Arthur Getis and Douglas Stow received a three-year grant from the Ethel Kennedy Shriver National Institute of Child Health and Human Development to study "Intraurban Health Assessed by Remote Sensing and GIS." That was followed by a four-year grant, also from NICHD, titled "Health, Poverty, and Place: Modeling Inequalities in Accra Using RS and GIS," for which Art was once again a co-investigator.

Technically, Arthur Getis retired from teaching in 2004, but that actually opened up some of his time to collaborate even more intensively on our joint research. Indeed, he traveled to Accra to get a first-hand look at our research there. Our work focused especially on issues of urban health and fertility, and Art's scholarly background was very strong in those areas. So, again it was not just his spatial statistical methods that counted, it was also his ability to make good sense of our findings, teaching us how to do that in the process.

A good example of that was our research on the relationship between malaria prevalence and urban agriculture in Accra– "Distance Threshold for the Effect of Urban Agriculture on Elevated Self-reported Malaria Prevalence in Accra, Ghana" (Stoler et al. 2009). Irrigated urban agriculture (UA) has helped alleviate poverty and increase food security in rapidly urbanizing sub-Saharan Africa but our research found that it may inadvertently support malaria vectors. We applied the Getis-Ord G_i* statistic to study spatial clusters of self-reported malaria rates aggregated at the enumeration area (EA) scale independently of a presumed source. EAs were represented geographically by their centroids. The analysis suggested that being within a 1 km distance of urban agriculture, controlling for household characteristics, was associated with an elevated prevalence of self-reported malaria. The 1 km distance is important because it is known that this is the approximate flight range for female anopheline mosquitoes seeking a blood meal. Although household locations were approximated by the centroid of its encompassing EA, the spatial decay of self-report rates beyond 1 km was strong, and the city-wide pattern of elevated self-reports within that range indicated an important health disparity.

It was also during the time we were conducting our research in Accra that Art and his doctoral student, Jared Aldstadt, developed the AMOEBA method of spatial agglomeration (A Multidirectional Optimal Ecotope-Based Algorithm) (Aldstadt and Getis 2006). It is based on the concept of spatial autocorrelation (Fischer and Getis 2010). "In brief, this algorithm starts with an initial area to which neighboring areas are iteratively attached until the addition of any neighboring area fails to increase the magnitude of the local G_i^* of Getis and Ord (1992) and Ord and Getis (1995). The resulting region is considered an ecotope. This procedure is executed for all areas, and final ecotopes are defined after resolving overlaps and asserting nonrandomness" (Duque et al. 2011:356). Thus, it builds on the idea of spatial clustering by creating "neighborhoods" based on statistical similarity of contiguous regions.

We first used AMOEBA to test the hypothesis that fertility levels in Accra, Ghana, are shaped and influenced by the neighborhood contexts in which women live, even when controlling for the individual characteristics of women. Our initial geographical unit of analysis was what we called vernacular neighborhoods, referring to neighborhood boundaries that are broadly recognized and agreed to by residents of a given city—in this case Accra, Ghana—even if they may have no premeditated and formal definition. These are the place names, for example, that would be provided to a taxi driver, especially since there is no comprehensive street address system in Accra (Weeks et al. 2010). These boundaries are not unlike those generated on the basis of local knowledge without access to census data and are similar to what one would find in printed tourist maps of Accra. The original boundaries of the neighborhoods were, I should note, drawn by Ghana Statistical Service (GSS), but in our research fieldwork we validated and reconciled differing neighborhood boundaries by visiting each neighborhood and talking to people about the name that they and their neighbors use. The result of this effort was a modification of the original GSS vernacular neighborhoods to reflect the perceptions of residents of the local boundaries. We called these the field modified vernacular (FMV) neighborhoods. Most of the difference between the original and FMV neighborhood definitions is that the latter provide a more nuanced and finer gradation. The one constraint on boundaries was that they could not divide the census-based enumeration areas (EAs) which are similar to a U.S. census tract, and for which data are summarized by GSS, and which then form the statistical data for each vernacular neighborhood.

We wanted to compare the analysis of the vernacular neighborhoods with that of what Getis and Aldstadt labeled "organic neighborhoods," which are contiguous agglomerations of census-based enumeration areas that are similar to one another with respect to contextual characteristics, using the AMOEBA algorithm to create these neighborhoods, as noted above. In its original form AMOEBA was designed to identify "hot spots" in regions in which there might be found statistically significant spatial clusters of a variable of interest for which a simple distance or contiguity spatial weights matrix did not adequately describe the pattern of clustering. For this work on identifying Accra neighborhoods, the algorithm was expanded to exhaustively classify all subareas (EAs) into clusters regardless of statistical significance. In this way, areas of homogeneity of a variable can be delimited across the entire city without the restriction that the areas must be hot or cold spots. Briefly, for a variable, the technique requires that each EA be evaluated for the strength of its association with contiguous EAs. The association is measured using any one of the local spatial autocorrelation statistics such as local Moran's I, local Geary's C, or the Getis-Ord G_i* statistic (which was employed in this analysis). The EAs are then ordered from highest to lowest association with their neighbors. The highest contiguous association is selected as the seed to begin a process in which through a sequential operation the contiguous neighbors of the highest EA are included in a cluster if those contiguous neighbors raise the level of association by their inclusion in a cluster. The sequential operation continues by selecting the contiguous neighbors of the previously selected contiguous neighbors that increase the association of the EAs already selected. When the level of association is reduced by the addition of a contiguous neighbor, the process comes to an end and the boundary of the group of associated neighbors is identified.

This first region of homogeneity is ineligible for the selection of the next possible high association between an EA and its contiguous neighbors, and so on. In this way the algorithm continues to find associated neighbors until all EAs are included within clusters. There is no restriction on the shape or size of the delimited neighborhoods. The distinguishing feature of the AMOEBA approach is its flexibility in identifying spatial association of nearby units regardless of the configuration of those units.

Our results confirmed the validity of the organic neighborhoods in the sense that we found very similar results for the two ways of defining neighborhoods. Since researchers may not have the same kind of government census information that allowed us to us to create our vernacular neighborhoods in Accra, this is very reassuring.

The purpose of AMOEBA is to identify spatial clustering of variables, but Art Getis realized that the resulting organic neighborhoods may not all have the same level of consistency with respect to the characteristics under investigation. This can produce a situation of heteroscedasticity, and he and Cliff Ord devised another technique to measure this phenomenon. LOSH stands for local spatial heteroscedasticity, and is measured by the H statistic, as defined by Ord and Getis (2012), and designed to better understand the relationship between spatially overlapping variables. The statistic allows us to focus on the nature of the pattern of possibly spatially related variables. Values of the H statistic, in conjunction with the G_i* statistic, reveal the pattern of homogeneity or heterogeneity. Using our data from Accra, Art employed LOSH to identify neighborhoods that are transitional, in the sense that they are undergoing more change than other parts of the city (Getis 2015).

During the time that Art was working with us on our data for Egypt and Ghana, the rise of spatial analytic techniques was dramatic, including the introduction of geographically weighted regression (GWR) (Brunsdon et al. 1996, Fotheringham et al. 1998, Fotheringham et al. 2002). Although Art was not directly involved in the creation of GWR, it built on the very same concepts that he had been working on and he encouraged us to employ it in our research, which we did. For example, we used it to improve our understanding of the spatial variability of fertility within Accra (Weeks et al. 2013). Using GWR, among several other spatial analysis tools, we concluded that Accra is a spatially complex city, and the patterns of reproduction reflect that complexity in ways that are not typical of western cities. Increasing levels of education have almost certainly contributed to the decline in fertility over time, but employment insecurity and housing insecurity may have contributed to a delay in marriage among women, and to a desire within marriage to postpone children while waiting for economic circumstances to improve. At the same time, these factors (which are admittedly difficult to measure) vary in their importance from neighborhood to neighborhood, perhaps being influenced by a variety of cultural factors including religion, ethnicity, and region of the country from which residents have been drawn.

3 Conclusion

My research over the past 30 years would not have happened had it not been for Art Getis, because the ability to undertake spatial demography was dependent upon his contributions to the field. He was my collaborator and my teacher, helping me and countless other researchers transition from being "spatially aware" to being "spatially analytical." Throughout our time together, Art and I shared a large number of graduate students (many of whom are co-authors on research publications discussed in this paper), and I guarantee you that each one of those students has a better life today because of Art's influence. Art was a brilliant man, and he wanted very much to share his insights with others, and I and my own students quickly became his students. Overall, Art's passion for geography, and for life, combined with his wonderful sense of humor and overall generosity, are remembered fondly and gratefully. Spatial demography is certainly a better science because of his contributions.

References

Aldstadt J, Getis A (2006) Using amoeba to create a spatial weights matrix and identify spatial clusters. Geogr Anal 38:327–343

Anselin L, Rey S (1991) Properties of tests for spatial dependence in linear regression models. Geogr Anal 23(2):112–132

Brunsdon C, Stewart Fotheringham A, Charlton ME (1996) Geographically weighted regression: a method for exploring spatial nonstationarity. Geogr Anal 28(4):281–298

Duque JC, Aldstadt J, Velasquez E, Franco JL, Betancourt A (2011) A computationally efficient method for delineating irregularly shaped spatial clusters. J Geogr Syst 13(4):355–372

Fischer MM, Getis A (eds) (2010) Handbook of applied spatial analysis. Springer, Berlin
Fotheringham AS, Charlton ME, Brunsdon C (1998) Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. Environ Plan A 30:1905–1927
Fotheringham AS, Charlton M, Brundson C (2002) Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. John Wiley & Sons Ltd, Chichester, UK
Getis Arthur (1995) Spatial filtering in a regression framework: examples using data on urban crime, regional inequality, and government expenditures. In: Anselin L, Florax RJGM (eds) New directions in spatial econometrics. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 172–185. https://doi.org/10.1007/978-3-642-79877-1_8
Getis A (2015) Analytically derived Neighborhoods in a rapidly growing west African city: the case of Accra, Ghana. Habitat Int 45:126–134
Getis A, Griffith DA (2002) Comparative spatial filtering in regression analysis. Geogr Anal 34(2):130–140
Getis A, Ord JK (1992) The analysis of spatial association by use of distance statistics. Geogr Anal 24(3):189–206
Getis A, Getis J, Fellman JD (1992) Human geography: landscapes of human activities, 3rd edn. William C. Brown, Dubuque
Getis, Arthur. 1994. "Spatial Dependence and Heterogeneity and Proximal Databases." in *Spatial Analysis and Gis*, edited by A. S. Fotheringham and P. A. Rogerson. London: Taylor & Francis.
Logan JR (2016) Challenges of spatial thinking. In: Howell F, Porter J, Matthews SA (eds) Recapturing space: new middle-range theory in spatial demography. Springer, Dordrecht
Ord JK, Getis A (1995) Local spatial autocorrelation statistics: distributional issues and an application. Geogr Anal 27(4):286–306
Ord JK, Getis A (2012) Local spatial heteroscedasticity (Losh). Ann Reg Sci 48:529–539
Scott L, Janikas M (2010) Spatial statistics in arcgis. In: Fischer CS, Getis A (eds) Handbook of applied spatial analysis: software tools, methods and applications. Springer-Verlag, Berlin
Stoler J, Weeks JR, Getis A, Hill AG (2009) Distance threshold for the effect of urban agriculture on elevated self-reported malaria prevalence in accra, ghana. Am J Trop Med Hyg 80(4):547–54
Tobler WR (1970) A computer movie simulating urban growth in the Detroit region. Econ Geogr 46:234–240
Ward D, Phinn SR, Murray AT (2000) Monitoring growth in rapidly urbanizing areas using remotely sensed data. Prof Geogr 52(3):371–385
Weeks John R (2004) The role of spatial analysis in demographic research. In: Goodchild MF, Janelle DG (eds) Spatially integrated social science: examples in best practice. Oxford University Press, New York
Weeks JR (2010) Spatial patterns of fertility change in rural Egypt. In: Anselin L, Rey SJ (eds) Perspectives on spatial data analysis. Springer Publishing Co, New York
Weeks John R (2016) Demography is an inherently spatial science. In: Howell F, Porter J, Matthews SA (eds) Recapturing space: new middle-range theory in spatial demography. Springer, Dordrecht, The Netherlands
Weeks JR (2021) Population: an introduction to concepts and issues, 13th edn. Cengage, Boston
Weeks JR, Getis A, Hill AG, Saad Gadalla M, Rashed T (2004) The fertility transition in Egypt: intra-urban patterns in Cairo. Ann Assoc Am Geogr 94(1):74–93
Weeks JR, Hill AG, Stow D, Getis A, Fugate D (2007) Can you spot a neighborhood from the air? Defining neighborhood structure in accra, ghana. Geo J 69:9–22
Weeks JR, Getis A, Hill AG, Agyei-Mensah S, Rain D (2010) Neighborhoods and fertility in accra, ghana: an amoeba-based approach. Ann Associat Am Geogr 100(3):558–78
Weeks JR, Hill AG, Stoler J (eds) (2013) Spatial inequalities: health, poverty, and place in Accra, Ghana. Springer Netherlands, Dordrecht