Using spatial indices to measure dynamic racial residential segregation in Gauteng province (South Africa)

Samy Katumba [a], Serena Coetzee [a], Alfred Stein [b,c] and Inger Fabris-Rotelli [c]

[a]Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, South Africa; [b]Faculty of Geo-information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands; [c]Department of Statistics, University of Pretoria, Pretoria, South Africa

ABSTRACT

Apartheid laws resulted in racial residential segregation that became entrenched into the urban morphology of South Africa. When apar-theid ended in the 1990's, the new South African democratic government was resolved to bring about social and spatial justice, address inequalities and promote social cohesion. To determine progress towards racial residential integration, aspatial indices of segregation are widely employed despite their shortcomings and limi-tations. This study employs two spatial indices of segregation: the spatial information theory index and the spatial exposure/isolation index in order to measure and quantify the levels of racial residential segregation that individuals living in Gauteng province (South Africa) would experience on average within radii of 500m, 4km and 8km of their respective residential locations. The analysis is based on the 1996, 2001 and 2011 Census data. The results indicate that the levels of racial residential segregation have steadily declined between 1996 and 2011. The levels of exposure of Whites to Black Africans continue to increase while the levels of exposure of Black Africans to Whites have remained unchanged over the same period. These observations are true for the different geographic scales (i.e. within radii of 500m, 4km and 8km) of analysis considered in this study.

KEYWORDS

Segregation; spatial segregation index; spatial information theory index; spatial exposure/isolation index; segregation profile curve

1. Introduction

The introduction of apartheid laws such as the Group Areas Act in the early 1950s resulted in racial residential segregation that indelibly marked the history of South Africa. Such laws, which were mainly based on race and led by the state, brought about the isolation of other population groups from urban areas reserved for the White population group (Lemanski et al., 2008; Wellings & Black, 1986). The Population Registration Act No 30 (1950) stipulated that all South Africans be racially classified into one of three categories: White, Black or Coloured. At that stage, anyone who was not White or Black, including Asians and Indians, was classified as Coloured. However, Indians were treated

CONTACT Samy Katumba samy.katumba@up.ac.za Department of Geography, Geoinformatics and MeteorologyUniversity of Pretoria

as a separate population group. For example, through the Extension of University Education Act No 45 (1959), separate tertiary institutions for Blacks, Indians, Coloureds and Whites were established, and through the Group Areas Act No 41 (1950) semi-urban townships were set up for Black, Indian and Coloured population groups, respectively (SAHO, 2020).

In the early 1990s, apartheid was abolished and the new democratic South African government was set to bring about social and spatial justice, address inequalities and promote social cohesion. This also meant doing away with the laws that restricted where one could live by race. Ironically, in order to measure progress with redress, today, South Africans still have to state their race on many forms, including the Census questionnaire where respondents have to describe themselves in terms of one of five population groups: Black African, Coloured, Indian or Asian, White, and Other.

Although post-apartheid policies pursue developmental reforms that favour socio-economic and spatial integration, their implementation have been inhibited by existing socio-economic circumstances such as social attitudes, poverty, inequality and unemployment (Hamann & Horn, 2015). Harrison and Todes (2015) noted the lack of deliberate attempts of current urban policies to address racial segregation proactively. For example, around three million housing units have been built in South Africa since 1994 under the housing programme of the Reconstruction and Development Programme (RDP) policy framework meant for reintegrating South African cities (Harrison & Todes, 2015). Most of these housing units have been built on cheaper land located at the fringes of economic urban centres (Todes et al., 2018). To some extent, this perpetuates the legacy of apartheid in that Black Africans, who are mostly the beneficiaries of the RDP housing programme, are still living far away from economic urban centres (Lemanski et al., 2008).

The Gauteng Spatial Development Framework (GSDF) of 2011 published by the Gauteng Provincial Government (Gauteng Provincial Government (GPG), 2011, p. 10) states that:

> The municipalities in the province all have plans, strategies and frameworks in place that seek to ensure integration to redress the apartheid fragmented socio-spatial structure and space economy.

Harrison and Todes (2015) note that several changes to the urban form in Johannesburg have been due to housing markets enabled by the relaxation of planning regulation. Places where such housing developments are taking place include middle- to high-income areas and inner-city spaces inhabited by diverse population groups (Todes et al., 2018). As a consequence, private residential complexes also known as 'gated communities' are being established, and the process of gentrification is also taking its course. Although these kinds of residential developments may lead to desegregation and be beneficial to those individuals occupying such spaces, their compounding effects result in a new form of residential segregation based on social class rather than race (Lemanski et al., 2008). Furthermore, some areas in the province have also experienced resegregation due to the so-called 'White flight' phenomenon whereby certain neighbourhoods in the province have gone from being predominantly White to be becoming predominantly Black African (Hamann & Horn, 2015; Horn & Ngcobo, 2003).

It is against the backdrop of these observed nuances that this study aims to provide a quantitative analysis which is essential in support of current and future qualitative studies for deepening the understanding of racial residential segregation, and ultimately the formulation of suitable urban policies in the context of Gauteng (South Africa). Furthermore, it is worthwhile to measure the levels of racial residential segregation on a regular basis in order to evaluate progress with respect to the South African ideals of the post-apartheid era.

Studies that have measured the levels of racial residential segregation in the context of South Africa widely used aspatial indices of segregation despite their limitations (Parry & Van Eeden, 2015). Aspatial indices of segregation assume that the social environment is uniform across a predefined spatial unit of aggregation (e.g., ward, mainplace, subplace, etc.), and ignore the proximity of neighbourhoods (i.e. spatial units of aggregation) across space (Massey & Denton, 1988; Yao et al., 2019). In other words, by not considering the spatial proximity among spatial units of data aggregation across a study area, aspatial indices of segregation ignore the chequerboard problem. Furthermore, aspatial indices of segregation (and even some of the spatial ones) are affected by the Modifiable Areal Unit Problem (MAUP), which refers to the sensitivity of such indices to the size and shape of the spatial units of analysis that might be arbitrarily chosen or might not accurately reflect the actual racial composition of the local neighbourhoods (Openshaw, 1984; Wong, 2004). The MAUP occurs when the results of the analysis depend on the level or scale of the spatial unit of aggregation of interest (i.e., the scale effect of the MAUP), and/or when the results of the analysis depend on the way in which the borders between the spatial units of aggregation have been drawn (i.e., the zoning effect of the MAUP) (Hennerdal & Nielsen, 2017; Weir-Smith, 2016).

The chequerboard problem and the MAUP both introduce the possibility of obtaining inaccurate measures of racial residential segregation and of being unable to compare the results at different scales of analysis (Reardon & O'Sullivan, 2004). To address these shortcomings, spatial indices of segregation have been proposed in the literature (Barros & Feitosa, 2018; Monkkonen & Zhang, 2014; O'Sullivan & Wong, 2007; Oka & Wong, 2015; White, 1983).

In South Africa, Parry and Van Eeden (2015) acknowledged the importance of employing spatial indices of segregation that minimize the MAUP and the chequerboard problem even though they did not explicitly use one due to the computational complexity of implementing such indices. Horn (2005) employed the spatial dissimilarity index for multiple ethnic groups (*SD*), the adjusted dissimilarity index (D(adj)) and the adjusted dissimilarity index which incorporates the length of shared boundaries between adjacent areal units (D(w)), in order to measure the levels of racial residential segregation in Pretoria (Gauteng, South Africa). Wong (2004) had noted that these spatial indices were affected by the MAUP.

Given the fact that indices of segregation are useful for understanding the problems of segregation and formulating public policies (Feitosa et al., 2007), it is therefore critical to ascertain the accuracy and reliability of these indices that quantify urban residential segregation. The purpose of this research is to measure and quantify the levels of racial residential segregation across space and time in Gauteng (South Africa) by employing two spatial indices of segregation, namely the spatial information theory index ($\tilde{H}$) and the spatial exposure/isolation index ($\tilde{P}^*$). The two indices are employed in order to explore the spatial evenness/clustering and the spatial exposure/isolation dimensions of residential segregation in Gauteng. According to Reardon and O'Sullivan (2004, p. 122), $\tilde{H}$ and $\tilde{P}^*$ 'are the most conceptually and mathematically satisfactory of the proposed spatial indices of segregation'. In theory, the two spatial indices minimize the MAUP and the chequerboard problem. Furthermore, the results obtained from the two indices can be compared at different scales of analysis as defined by the user/analyst irrespective of the underlying geographic unit of aggregation (e.g., small area, subplace, mainplace, ward, etc.) employed.

The remainder of this paper is structured as follows: section 2 provides a literature review on the measurement and quantification of racial residential segregation. A description of the study area, data and methodology employed in this paper is provided in section 3. The results are described and discussed in section 4. Section 5 concludes this paper by also providing some recommendations and future plans.

2. Measuring racial residential segregation

2.1 Spatial and aspatial indices of segregation

Massey and Denton (1988, p. 282) defined residential segregation as 'the degree to which two or more groups live separately from one another in different parts of the urban environment'. They referred to residential segregation as a multidimensional phenomenon that can be conceptualized into five different dimensions: evenness, exposure, concentration, centralization and clustering. Each of these dimensions implicitly represents a spatial variation of how segregation manifests itself. Massey and Denton (1988) surveyed 20 indices of segregation and grouped them into five dimensions. Their aim was to shed light on fierce debates scholars had about the merit of existing segregation measures at the time. To put an end to these debates, Massey and Denton (1988) recommended that a multidimensional socio-economic phenomenon such as residential segregation be measured by more than one single index depending on the dimensions of segregation being investigated. However, limitations of existing traditional indices of segregation continued to inspire scholars interested in the study of segregation to develop new indices. Traditional indices of segregation are referred to as 'aspatial' indices of segregation because they do not satisfactorily take into consideration the spatial arrangements of residential locations (Morrill, 1991; Reardon & O'Sullivan, 2004; Wong, 2002).

Notable limitations that have plagued aspatial indices of segregation as reported in the literature include the chequerboard problem and the modifiable areal unit problem (MAUP). Indices of segregation that ignore the chequerboard problem fail to consider the spatial proximity of neighbourhoods in their formulation (Reardon & O'Sullivan, 2004). In other words, an index of segregation is affected by the chequerboard problem when the levels of segregation it measures remain unchanged irrespective of changes in the spatial arrangement of the different population groups across the study region. To illustrate the chequerboard problem, Figure 1 provides two hypothetical study regions showing the distribution patterns of two population groups (black and white each occupying a grid cell exclusively). When used, an aspatial measure of segregation such as the dissimilarity index D would compute the same value for the two study regions even though they do not appear to display the same levels of segregation overall.

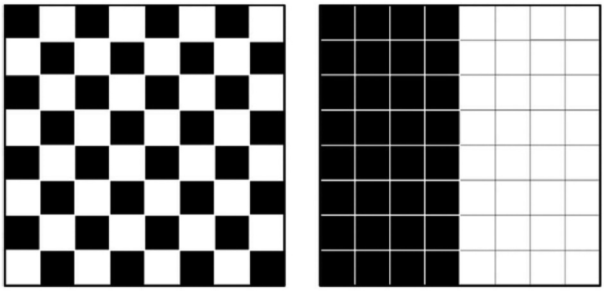

Figure 1. Illustration of the chequerboard problem.

Furthermore, ignoring the MAUP may result in spatial and aspatial segregation indices that are difficult to interpret and compare (Reardon et al., 2008). For a measure of segregation to be free from the MAUP, it needs to be computed from data collected at the exact locations of the individuals and those (individuals) located in their proximities (Reardon & O'Sullivan, 2004). Appendix A (section 6) provides a list of selected aspatial indices of segregation and their corresponding spatial versions commonly used in the literature in the context of South Africa.

As originally proposed by Duncan and Duncan (1955), the dissimilarity index *D* measures segregation only between two population groups within an areal subunit of a study region of interest. Given that in most realistic settings more than two population groups are involved, Sakoda (1981) proposed the generalized index of dissimilarity *D*(m) that consider all the population groups within a study region. Although these two indices (i.e. *D* and *D*(m)) are the most mentioned aspatial indices of segregation in the literature, they are not immune to the MAUP and the chequerboard problem (Wong, 1998). As a way of addressing these shortcomings, Morrill (1991) proposed the adjusted dissimi-larity index *D* (adj), Wong (1993) proposed the boundary-adjusted dissimilarity index *D*(w) and the *D* (s) index, and Wong (1998) proposed the multigroup spatial index of segregation *SD*(m). These are all spatial versions of the traditional aspatial indices of segregation *D* and *D*(m). The *D* (adj) incorporates a spatial term that defines the spatial connectivity between any two adjacent areal subunits within the study region. The *D*(w) index includes in its formulation a weight term which computes a ratio between the length of the common boundary between two adjacent areal subunits and the total length of all the common boundaries between adjacent areal subunits in the entire study region. Besides the weight term as defined in the case of the boundary-adjusted dissimi-larity index, the *D*(s) also considers the perimeter and area value of each of the areal subunits in the study region. Finally, the multigroup spatial index of segregation incorporates in its formulation, the composite population count of a population group within a given areal subunit and the population counts of the same group in all surrounding areal subunits. However, when comparing the results of aspatial and spatial indices of segregation at different scales, Wong (2004) noted that all these segregation indices (e.g., *D, D*(m), *D*(adj), *D*(w), *D*(s) and *SD* (m)) were affected by the scale effect of the MAUP.

Other efforts towards addressing the MAUP and the chequerboard problem are also discussed. For example, Reardon and O'Sullivan (2004) proposed a generalized approach for constructing spatial segregation indices based on spatial proximity that takes a functional form of how segregation is experienced by individuals both locally and globally. Theoretically, this would require ‘point-to-point proximity functions’ that define the local environments or scales of analysis instead of relying on edge contiguity matrices that some spatial segregation indices use, especially with data that have been aggregated by polygon areas or subareas (Reardon & O’Sullivan, 2004). Similar approaches can be found in the work of Wong (2005) who further proposed a generalized spatial segregation measure which incorporates spatial adjacency as suggested by Morrill (1991). By adopting a flexible definition of neighbourhood and a multiscale approach, Wong (2005) addressed the concentration and clustering dimensions of segregation by employing the concept of composite population counts to capture the interaction of different population groups within the neighbourhoods of a given area or subarea. However, the generalized spatial segregation index as proposed by Wong (2005) was only applied to two groups (Black and White populations) therein.

When studying residential patterns of foreign population in Southern European’s cities, Benassi et al. (2020), employed two aspatial segregation indices: the Theil’s information theory index H, the dissimilarity index for multiple groups D (m) and a spatial index of segregation: the 'local J evenness index' which is a modified version of H for modelling spatial segregation at the local level. Although Benassi et al. (2020) also employed a local version of H, given that the three indices (i.e. H, D(m) and J *are* computed based on data aggregated per grid cells, they are not immune to the chequerboard problem and the MAUP, which negatively affect the results of aspatial indices of segregation (Chodrow, 2017; Reardon et al., 2008; Wong & Chong, 1998).

The two spatial indices of segregation employed in this study as proposed by Reardon and O’Sullivan (2004), are the spatial information theory Index ($\tilde{H}$) and the spatial exposure/isolation ($\tilde{P}^*$).These two indices ($\tilde{H}$ and $\tilde{P}^*$) are spatial versions of the Theil information theory index (H) and the exposure/isolation aspatial indices (xP^*_y/xP^*_x) respectively. The two spatial indices ($\tilde{H}$ and $\tilde{P}^*$) are discussed in detail in the methodology section of this paper (section 3). The reasons for employing the two indices have also been provided.

2.2 Selected studies on measuring racial residential segregation in South Africa

2.2.1 Use of aspatial indices of segregation

Christopher (2001) used the 1996 Census data collected at the small area and enumeration (tract) area levels, applied the dissimilarity index as formulated by Duncan and Duncan (1955) to study racial segregation in post-apartheid South Africa and observed an overall decline in urban racial segregation. Considering the results in detail, Christopher (2001) noted that Whites remained more segregated compared to the other population groups made of Black Africans, Coloured and Indians or Asians. Although the levels of integration of Black Africans had improved, they had limited options in terms of residential mobility due to poverty. Indians or Asians and Coloureds had experienced the highest levels of integration as they had begun to return to areas from which they were removed by force under apartheid laws.

Christopher (2005) used the dissimilarity index (D) despite its limitations to measure the levels of segregation in all the provinces of South Africa between 1996 and 2001 using the 1996 and 2001 Census population data. Christopher (2005) noted an overall slow pace of residential integration compared to the post-apartheid period between 1991 and 1996. However, high levels of integration were observed within areas which were predominantly occupied by the White population group.

Hamann and Horn (2015) provided insights into the understanding of racial residential segregation (or lack thereof) in Tshwane (Pretoria) by applying aspatial indices of segregation, namely, the dissimilarity index (*D*), the dissimilarity index for multiple groups (*D* (m)), interaction/exposure index $\left(xP_y^*\right)$ and isolation index $\left(xP_x^*\right)$, on the 1991 and 2011 population Census data aggregated at the subplace level. They argued that Tshwane still displayed high levels of residential segregation in 2011. Such a conclusion could have been influenced by the fact that their analysis was only performed at the subplace level, a smaller areal unit than the mainplace or municipality, which would have presented lower levels of racial residential segregation. Furthermore, Hamann and Horn (2015) noted low levels of interaction amongst the four population groups (i.e. Black Africans, Whites, Coloureds and Indians or Asians). Black Africans were the most isolated population group based on the 2011 Census, followed by Whites. Indians or Asians and Coloureds were the least segregated from the other population groups based on the 2011 Census. The levels of residential interaction between Whites and Black Africans were the lowest compared to other population groups. Although it might be tempting to liken the results obtained from Tshwane to other municipalities in the province (i.e. Gauteng), Hamann and Horn (2015) hinted that studies of racial residential segregation need to be context-specific. This is because they found very few similarities when comparing the results obtained in Tshwane with other selected municipalities in South Africa.

2.2.2 Use of spatial indices of segregation

Horn (2005) employed spatial indices of segregation suggested by Wong and Chong (1998) and Wong (2003) when analysing the 1991 and 1996 population Census data at the subplace level. Horn (2005) explored the *D* (adj), *D* (w) and *SD* (m) and recommended the use of *SD* (m) as the 'standard global index' for measuring racial segregation in 'late-modern' South African cities. Although Horn's (2005) study is among the first to have explicitly employed spatial indices in the measurement of racial segregation in the context of South Africa, it should be noted that Wong (2004) had already mentioned that these spatial indices (i.e. *D* (adj), *D* (w), *D* (s) and *SD* (m)) failed to address the scale effect of the MAUP.

More recently, Parry and Van Eeden (2015) acknowledged the MAUP when analysing changes in racial diversity and racial segregation between 1991 and 2011 in the cities of Cape Town and Johannesburg at different geographic scales. After employing Theil's entropy index of segregation on Census data collected in 1991, 1996, 2001 and 2011, Parry and Van Eeden (2015) observed an overall steady decline in the levels of segregation in both cities. Parry and Van Eeden (2015) found that Cape Town was more racially segregated at a lower geographic scale while Johannesburg was more segregated at a higher geographic scale. They divided the study areas (i.e. Cape Town and Johannesburg) into square grids of different cell sizes to compute the segregation profile curves at different geographic scales of data aggregation. Nevertheless, Parry and Van Eeden (2015) hinted that future studies on racial residential segregation should employ the complete method proposed by Reardon et al. (2008) which had been originally proposed by Reardon and O'Sullivan (2004).

3. Data and methodology

3.1 Study area

The study area considered in this paper is Gauteng, the smallest province in terms of land area in South Africa. With an estimated population of 13.3 million, Gauteng is the most densely populated province in South Africa (StatsSA Community Survey, 2016). It is the most urbanized province and has three metropolitan municipalities: City of Johannesburg (Johannesburg), City of Tshwane (Tshwane) and City of Ekurhuleni (Ekurhuleni). Additionally, there are six local municipalities in Gauteng. The map in Figure 2 illustrates the geographic location of Gauteng province and its municipalities in South Africa at the time of the 2011 Census (i.e. three metropolitan municipalities and seven local municipalities). The province generates more than a third (33.8%) of the country's Gross Domestic Product (StatsSA, 2019), making Gauteng an economic hub of Southern Africa.

Despite Gauteng being an economically vibrant province, many residents do not benefit from the economic and social opportunities in the province. Such inequalities were already evident well before the introduction of apartheid laws that officially came to promote racial residential segregation and the economic exclusion of mostly the Black African section of the population (Parnell & Pirie, 1991).

3.2 Data description

To understand the spatial patterns of racial residential segregation, Statistics South Africa (StatsSA) Census datasets related to racial population groups were analysed. The differ-

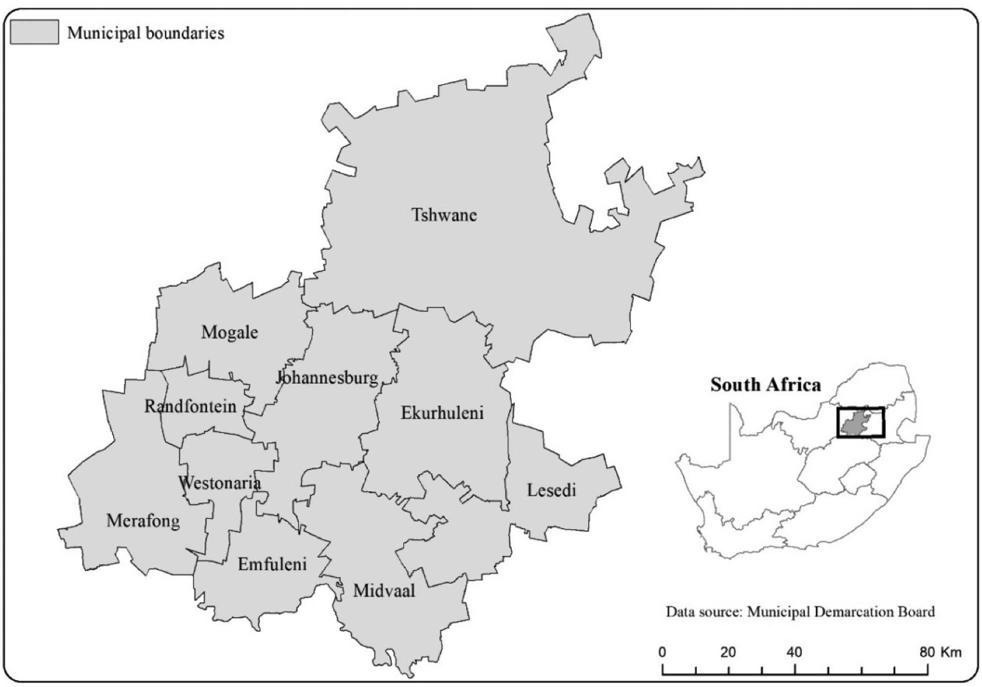

Figure 2. Geographic location of gauteng province (as per StatsSA Census 2011).

ent levels of data aggregation considered in the study are as follows:

- The 2011 Census data aggregated at small area, subplace and ward levels.
- The 1996, 2001 and 2011 Census data aggregated at the ward level, respectively.

The small area is the lowest spatial unit of aggregation at which StatsSA makes Census data available to the public (http://www.statssa.gov.za/?page_id=4086). A typical small area unit in Gauteng occupies a land area of less than 578 km^2. Several small areas are combined to form a subplace. Subplaces provide recognizable names of neighbourhoods in Gauteng. Given that the ward boundaries were harmonized, it is possible to compare the results of spatial segregation indices at the ward level across three iterations of Census datasets (1996, 2001, and 2011). Four population groups, namely the Black African (BA), White (W), Coloured (C) and Indian or Asian (IA) were considered in the analysis, following the classification of population groups in Censuses conducted by StatsSA since 1996 (Christopher, 2005). The 2011 population Census also adopted the same classification. These population groups were mainly considered for the purpose of comparing this study's results with the ones of previous studies which have mostly analysed the four population groups (i.e. BA, W, IA and C).

3.3 Spatial information theory index ($\widetilde{H}$), Spatial exposure/isolation index ($\widetilde{P}^*$) indices and segregation profile curves

The two measures of segregation employed in this study: the spatial information theory index ($\tilde{H}$) and spatial exposure/isolation index ($\tilde{P}^*$) were computed by following the methodology proposed in Reardon and O'Sullivan (2004). The two indices of segregation are operationalized in R statistical software under the 'seg' package (Hong et al., 2014). While $\tilde{P}^*$ measures the extent of the encounter between members of two different population groups within their local spatial environments, $\tilde{H}$ quantifies how diverse individual local environments are, on average, with respect to the total population in the study region of interest (Lee et al., 2008; Reardon & O'Sullivan, 2004). The racial composition of local environments is computed at different geographic scales with radii between 500 m and 8 km. Figure 3 provides an illustration of how a local environment is determined. As can be seen in Figure 3, the centre (or centroid to be precise) of a given subunit of aggregation (as in this case the mainplace) is determined. It is assumed that each individual in the subunit lives at its centroid, and the centroid constitutes the central point of the local environment of that particular individual. The racial composition of the local environment is calculated by considering the distances that separate its centroid with all the other surrounding centroids which fall within a circle of a given radius (i.e. the radius of the circle defines the scale or size of the local environment). Local environments are determined as circles which are referred to as different geographic scales of neighbourhoods (or different scales of analysis) considered in this study. Lee et al. (2008) refer to the racial composition of the local environment as 'the proximity-weighted racial average composition of each individual's local environment'. For the case of the spatial information theory index ($\widetilde{H}$), once the racial compositions of individuals belonging to the different considered population groups within their respective local environments have

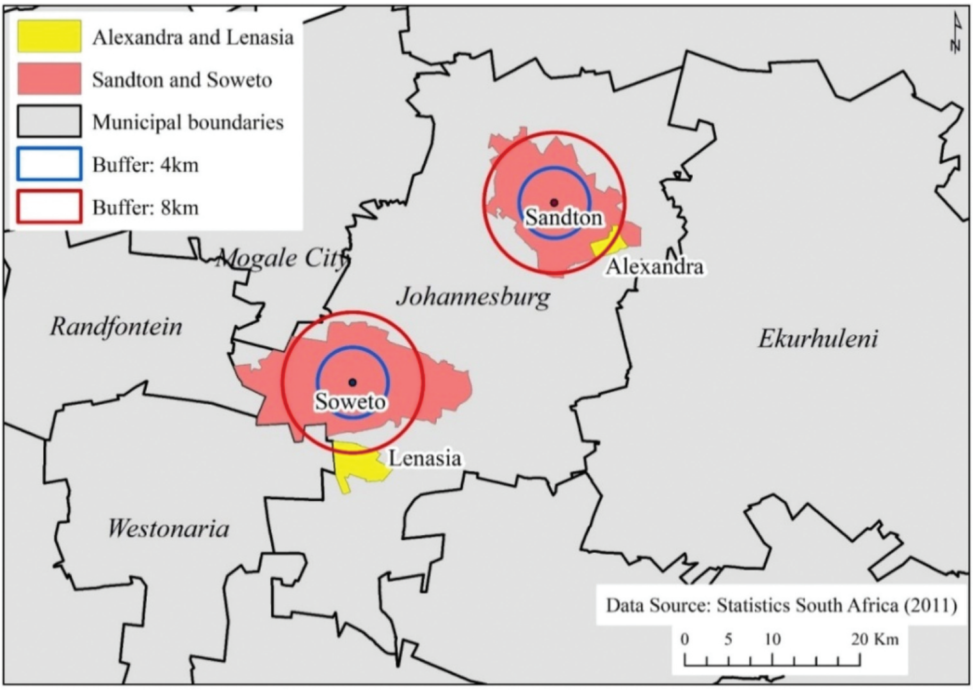

Figure 3. Scale of local environments (Illustration).

been determined, then the racial entropy of each local environment in the study area can be calculated. $\tilde{H}$ is computed by taking into account the average of the entropy values computed based on local environments of the same size (or scale) across the study region. The spatial exposure/isolation index ($\tilde{P}^*$) is computed by summing the product of the proportions of any two population groups as calculated within local environments of the same size. This allows $\tilde{H}$ and $\tilde{P}^*$ to be computed at different geographic scales of analysis.

Given that the population might not be evenly distributed within certain areal subunits (e.g., small area, subplace or ward), the idea of fixing the entire population of an areal subunit to its geometric centroid might have its own limitations. This is a concern inherent in data aggregated per a given spatial unit as it is the case with Census datasets used in this study. However, the spatial proximity among the different areal subunits (e.g., small area, subplace and ward) have been considered in the computation of the two indices ($\tilde{H}$ and $\tilde{P}^*$) through the use of a distance decay function. That is, the levels of interaction and clustering among population groups within their local environments have been considered in the computation of the racial composition in those local environments. During the process of computing the racial composition of local environments, population groups that are far apart would have less weight based on the distance separating them (i.e. between centroids) within their respective local environments compared to population groups that are closer to each other.

The mathematical formulations of $\tilde{H}$ and $\tilde{P}^*$ have been provided by the equations 1 to 4 and 5 to 6, respectively, as follows:

$$\tilde{H} = 1 - \frac{1}{TE} \int_{p \in R} T_p E_p dp \tag{1}$$

$$E_p = -\sum_{m=1}^{M} (\tilde{\pi}_{pm}) log_M (\tilde{\pi}_{pm}) \tag{2}$$

$$E = -\sum_{m=1}^{M} (\pi_m) log_M (\pi_m) \tag{3}$$

$$\tilde{\pi}_{pm} = \frac{\int_{q \in R} T_{qm} W(d) dq}{\int_{q \in R} T_q W(d) dq} \tag{4}$$

Where:

- p: the pth centroid within study area R,
- q: the qth centroid within study area R,
- M: the number of population groups in study area R,
- m: a given population group (e.g., Black African),
- T_p: population density at centroid p,
- T_q: population density at centroid q,
- T_{pm}: population density of population group m at centroid p,
- T_{qm}: population density of population group m at centroid q,
- $\tilde{T}_{pm}$: population density of population group m in the local environment of centroid p,
- $\tilde{T}_{qm}$: population density of population group m in the local environment of centroid q,
- T: total population within study area R,
- π_{pm} : proportion of population group m at centroid p,
- $\tilde{\pi}_{pm}$: proportion of population group m in the local environment of centroid p,
- π_m : proportion of population group m of the total population in the study area R,
- E_p: entropy (diversity) of the local environment of centroid p,
- E: overall entropy of the study area R,
- $W(d) = 1/d^{\alpha}$: power function (with α: the rate of decay set to 1 and d is the Euclidean distance between two centroids). Maximum values of d are set to determine the different scales of neighbours (i.e. local environments) at which segregation is measured.

Total racial residential integration corresponds to $\tilde{H} = 0$ and complete racial segregation to $\tilde{H} = 1$.

The spatial exposure/isolation segregation index ($\tilde{P}^*$) which measures the extent of exposure amongst members of the same population group (i.e. spatial isolation: $m\tilde{P}^*{}_m$), and the extent of exposure amongst members of two different population groups (i.e. spatial exposure: $m\tilde{P}^*{}_n$) is given by:

$$mP_n^* = \int_{q \in R} \frac{T_{qm}}{T_m} \tilde{\pi}_{qn} dq \quad (5)$$

$$mP_m^* = \int_{q \in R} \frac{T_{qm}}{T_m} \tilde{\pi}_{qm} dq \quad (6)$$

Where:

- $m\tilde{P}^*{}_n$: the spatial exposure of population group m to the proportion of group n as the average proportion of group n in the local environment of each member of group m,
- $m\tilde{P}^*{}_m$: the spatial isolation of a given population group m, specifying the exposure of population group m to itself.
- $\tilde{\pi}_{qn}$: proportion of population group n in the local environment of centroid q,
- $\tilde{\pi}_{qm}$: proportion of population group m in the local environment of centroid q,
- T_m: Total population of population group m in the study area R.

The mathematical integral sign is employed in order to conform to the original formulation of the two indices ($\tilde{H}$ and $\tilde{P}^*$) in Reardon and O'Sullivan (2004) where population density surfaces are computed across the study region, allowing for the population density of each population group to be estimated at any point on the surface (e.g., in O'Sullivan & Wong, 2007). One of the approaches for computing a population density surface is to attribute the population density of a spatial unit (e.g., Census small area unit, Census subplace, Census ward or Census tract, etc.) to any point within that particular spatial unit (Reardon & O'Sullivan, 2004).

Given that the analysis is based on an irregular polygon dataset, the racial composition of local spatial environments was calculated using the 'localenv()' function which was passed to the 'spatseg()' function as recommended by Hong et al. (2014) when using the R package 'seg'. By default, the population was assigned to the centroid of each small area unit, subplace or ward in the study area. The centroid represents the entire small area unit, subplace or ward.

The 'seg' R statistical software package was used because it caters for the computation of indices of spatial segregation at different scales of analysis, while also considering the proportion of the racial composition of the local neighbourhoods (or environments) as proposed by Reardon and O'Sullivan (2004). That is, instead of being constrained to polygon boundaries (e.g., grids), the approach adopted in this paper minimizes the MAUP and the chequerboard problem in its analytical process. This is because the computation of the indices of segregation is based on a method of aggregating population within concentric circles as users' defined scales of analysis. This provides a realistic approximation of the actual racial composition of local spatial environments. The local spatial environment defines the extent of the area within which an individual would experience segregation.

To obtain the segregation profile curve which is a plot of values of $\tilde{H}$ against corresponding radii (i.e. scales of analysis/geographic scales of neighbourhood) as defined by Reardon and O'Sullivan (2004), the minimum radius was set to 71 m. This is the smallest distance between two centroids in the study area when using the subplace dataset. The two subplaces with centroids separated by a distance of 71 m are Alexandra Ext 37 and Alexandra Ext 29 located in Alexandra mainplace in Johannesburg (see Figure 4). Given that it was computationally demanding to calculate distances between centroids of small area polygons in a matrix, a distance of 71 m was used as the minimum starting radius of the local environments.

As illustrated by the segregation profile curve in Figure 5, the value of the spatial information theory index ($\tilde{H}$) decreases as the radius increases from 71 m to 38 km. $\tilde{H}$ reaches its maximum (0.61) when r = 71 m is used to delimit the area of the local environment of each of the centroids that represent small areas. The value of $\tilde{H}$ decreases to 0.57 when the subplace dataset is used. At r = 38 km, $\tilde{H}$ approaches zero (see Figure 5). This means that there is a total racial integration at r = 38 km.

3.4 Macro/Micro segregation ratios

Racial residential segregation that occurs at small geographic scales is referred to as micro-segregation and is often modelled by a steep segregation profile curve (Reardon &

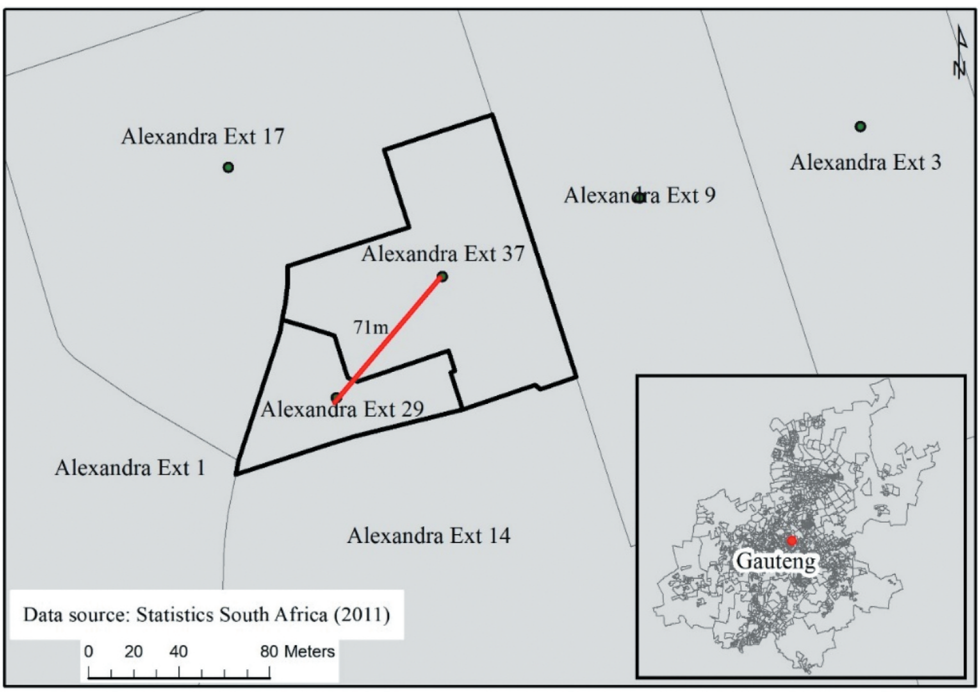

Figure 4. Shortest distance between two centroids (of subplaces) in the study area.

O'Sullivan, 2004). Micro-segregation may be caused by factors related to the spatial location of public amenities such as schools, clinics and shops (Parry & Van Eeden, 2015; Reardon et al., 2008). In contrast, macro-segregation is characterized by a separation of population groups at large geographic scales. Macro segregation is often modelled by a flat segregation profile curve (Reardon et al., 2008). As Parry and Van Eeden (2015) suggest, the factors that may cause macro-segregation relate to larger structural processes such as labour, demographic characteristics or historical settlement patterns.

In order to gain more insight into the geographic scales of racial residential segregation (whether large or small), macro/micro segregation ratios at well-defined geographic scales were employed. Macro/micro-segregation ratios provide insight into the possible causes or consequences of segregation (Reardon et al., 2008). A macro/micro segregation ratio close to zero implies that the factors that influence micro-segregation are only partially due to the patterns of racial residential segregation observed at a macro-scale level (Parry & Van Eeden, 2015; Reardon et al., 2008). A macro/micro segregation ratio close to one suggests that a large extent of the observed micro-segregation is due to macro-segregation patterns. In other words, a similar level of segregation is observed whether within small or large distances. Local environments with radii equal to 500 m, 4 km and 8 km were chosen to compute the macro/micro segregation ratios. The interval between 500 m and 4 km approximates the distance within which an individual would experience or interact with his or her neighbourhood (Reardon et al., 2008). However, since the racial composition of most Gauteng's neighbourhoods is homogenous, the radius was increased to 8 km in order to capture the levels of racial diversity amongst contiguous neighbourhoods. At shorter radii with less variability in the racial composition of small areas, high levels of segregation can be expected. As the radius of local environments increases beyond 4 km, the levels of racial residential segregation start to decrease.

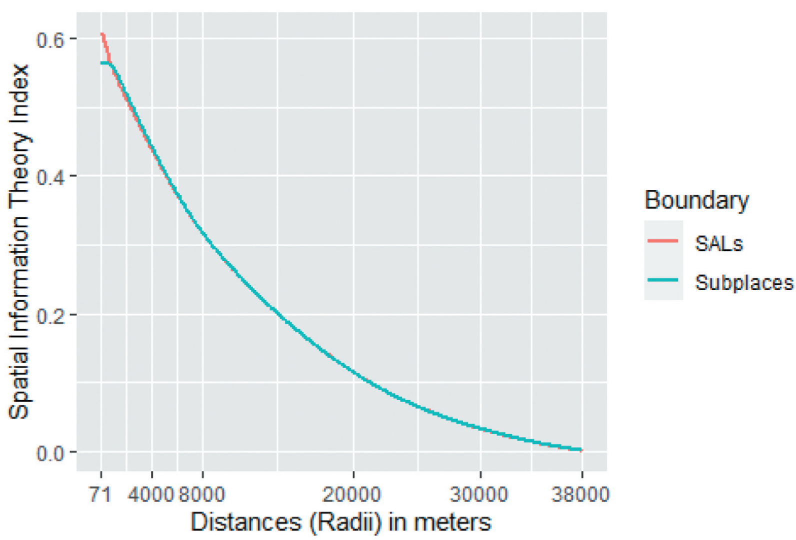

Figure 5. Segregation profile curves – Small area layer/unit (SAL) and Subplace based on 2011 Census data.

To provide an illustration, Figure 3 (mentioned in section 3.3) shows the neighbourhoods of Soweto and Sandton in the metropolitan municipality of Johannesburg. Based on StatsSA Census 2011, aggregated at the mainplace level, Soweto is predominantly occupied by the Black African population group. A typical neighbourhood of 4 km radius within Soweto would result in a homogenous racial composition when measuring the levels of segregation. A neighbourhood of 8 km radius would then include Lenasia, a predominantly Indian or Asian neighbourhood located to the south of Soweto. This also applies to Sandton, predominantly occupied by the White population group, and Alexandra, predominantly occupied by the Black African population group (see Figure 3).

3.5 Dot density maps

Dot density maps are used to depict the spatial distribution of the four population groups considered in this study. They are specifically employed to provide a visual (non-quantitative) means for validating the quantitative results obtained from the computation of the $\tilde{H}$ and $\tilde{P}^*$ segregation indices, the macro/micro segregation ratios and the segregation profile curves. The availability of StatsSA Census data at the ward level collected in 1996, 2001 and 2011, allowed for the temporal comparison of the distribution of the four population groups (i.e. BA, W, IA, and C) across space using dot density maps. The dot density maps were obtained by using the 'dot density symbology' available in ArcMap GIS software (Esri, 2008). For each polygon (i.e. ward) in the study area, quantitative values of fields representing the four population groups (i.e. BA, W, IA, and C) are represented as dots of equal size. For example, if there are 8 people of group BA in a given ward, the dot density symbology can be set up to represent 8 dots (i.e. 1 dot per person).

4. Results and discussion

Besides reporting the results of the two spatial indices ($\tilde{H}$ and $\tilde{P}^*$), the segregation profile curves and the macro/micro segregation ratios, by basing the analyses on the small area, subplace and ward (as spatial units of analysis), section 4.1 also illustrates how the MAUP was overcome for local environments of radii equal or greater than 4 km. The 2011 StatsSA Census data was used to performed the analysis reported in section 4.1.

Section 4.2 reports on the results of the two spatial indices of segregation (i.e., $\tilde{H}$ and $\tilde{P}^*$), the macro/micro segregation ratios and the segregation profile curves for each of the municipalities in Gauteng in order to understand some of the racial composition dynamics in Gauteng's municipalities. The 2011 Census data aggregated at the small area level was used for the analysis in section 4.2.

Finally, the analysis reported in section 4.3 was specifically based on the ward (as the spatial unit of analysis) in order to understand how the levels of segregation have changed (or not) across time in Gauteng province. Longitudinal StatsSA Census data (collected in 1996, 2001 and 2011) was only harmonized at the ward level. The dot density maps are included as a visual (non-quantitative) means for validating the results of the segregation profile curves and the two spatial indices of segregation ($\tilde{H}$ and $\tilde{P}^*$).

4.1 Racial residential segregation when using data aggregated per small area, subplace and ward levels

Even though the original datasets have been aggregated per small area, subplace and ward, the analysis has been performed based on defined scales (i.e. 500 m, 4 km and 8 km) as per the method described in the data and methodology section. For example, as displayed on Table 1, with local environments of radius 500 m, $\tilde{H}$ has a value of 0.58 when the population group dataset aggregated by small area has been used. This suggests that in Gauteng, an individual would experience a level of segregation equal to 0.58 on average within a radius of 500 m of his/her local environment. Table 1 also provides the macro/micro segregation ratios when the small area, subplace and ward datasets are used to compute values for the $\tilde{H}$ index. The ratios of 75%, 79%, and 89% when the small area, subplace and ward datasets have been employed, respectively, quantify the extent to which micro-segregation, that is, segregation observed among the '500 m-radius' local environments, in Gauteng, can be attributed to the racial residential patterns observed at a macro-scale level (i.e. within and beyond 4 km). Although these three macro/micro segregation ratios decrease when computed between 500 m and 8 km, they remain nevertheless relatively above 50%. In other words, these observed ratios (i.e. 75%, 79%, and 89%) suggest that variations in the racial composition of Gauteng's residential spaces occur over larger distances.

Another observation is, the levels of segregation within 8 km radius when computed based on the three datasets, (i.e. small area, subplace and ward), are approximately the same (i.e. 0.32, 0.32 and 0.30 respectively) (refer to Table 1). On a methodological note, this observation is a demonstration of how the MAUP has been overcome with the $\tilde{H}$ index as computed in this study, when the radius of local environments is greater than 4 km. That is, at a scale of local environments beyond 4 km radius, the level at which the population group dataset has been aggregated (whether small area, subplace or ward), does not affect the values of $\tilde{H}$.

Even though the computation of the levels of segregation at different scales of analysis combined with the inspection of macro/micro segregation ratios improve the measurement of segregation at different geographic scales, the availability of datasets at the microdata level or aggregated per a relatively small spatial unit, is of great value. This is important in order to understand some of the dynamics in the racial composition of neighbourhoods at a micro-scale level that might be overlooked when population group data are aggregated per a large spatial unit such as the ward, municipality or some large subplaces.

The spatial exposure/isolation segregation index results can further supplement and/or confirm the results already obtained at different scales of local environments with $\tilde{H}$ (see Table 2). Table 2 presents the results of the $\tilde{P}^*$ index. The figures highlighted in bold

Table 1. Macro/Micro segregation ratios at different geographic scales (2011).

Geographic scale	$\tilde{H}$ (500 m)	$\tilde{H}$ (4 km)	$\tilde{H}$ (8 km)	Macro/Micro ratio: $\tilde{H}$ (4 km)/$\tilde{H}$ (500 m)	Macro/Micro ratio: $\tilde{H}$ (8 km)/$\tilde{H}$ (500 m)
Small area	0.58	0.44	0.32	0.75 (75%)	0.54 (54%)
Subplaces	0.56	0.45	0.32	0.79 (79%)	0.57 (57%)
Wards	0.46	0.41	0.30	0.89 (89%)	0.65 (65%)

represent the degree of isolation of a given population group. The higher the magnitude of the figure displayed in bold, the higher the degree of isolation of that particular population group. The remaining figures (not in bold) represent the degree of exposure of one population group to the other. The higher the magnitude of the figure (not in bold) the higher the extent of exposure across space. Table 2 should be read horizontally. For example, the levels of exposure of Black Africans to Black Africans (i.e. isolation) within a radius of 500 m is 0.92 while their levels of exposure to Whites is 0.05. Within a radius of 8 km, the levels of exposure of Whites to Black Africans is 0.49 while their levels of exposure to themselves (i.e. isolation) is 0.43.

With high values of the spatial isolation segregation index and low values of spatial exposure segregation index, in 2011, the Black African population group was the most isolated and least exposed to other population groups in Gauteng province. Besides considering the spatial arrangement of residential households by race in Gauteng, this observation can partly be explained by the fact that the Black Africans constitute the majority of the population in all the municipalities in Gauteng. The Indian/Asian and Coloured population groups are more exposed to the White population group as compared to the Black African population. In other words, there are more Whites in the local environments of Coloureds and Indians/Asians than there are in the local environments of Black Africans. Conversely, the White population group is more exposed to the Black African population group compared to the Indian/Asian and Coloured population groups. In general, spatial isolation amongst the population groups is higher within areas (local environments) of 500 m radius as compared to areas within 4 km or 8 km radii. As the distance across neighbourhoods increases, there is more exposure amongst the different population groups. This latter observation corroborates some of the results already obtained from employing the spatial information theory index ($\tilde{H}$) which suggest that there is less variations in the racial composition of Gauteng's residential spaces at shorter distances and high levels of racial diversity as the sizes/scales of residential neighbourhoods become large.

Table 2. Spatial exposure/isolation segregation index ($\tilde{P}^*$) based on small area dataset at three different geographic scales (2011).

	Population groups			
500 m	Black African	Coloured	Indian/Asian	White
Black African	**0.92**	0.02	0.01	0.05
Coloured	0.35	**0.46**	0.04	0.14
Indian/Asian	0.32	0.05	**0.39**	0.24
White	0.27	0.03	0.04	**0.65**
4 km	Black African	Coloured	Indian/Asian	White
Black African	**0.89**	0.02	0.02	0.07
Coloured	0.57	**0.23**	0.05	0.16
Indian/Asian	0.49	0.06	**0.21**	0.24
White	0.37	0.04	0.04	**0.55**
8 km	Black African	Coloured	Indian/Asian	White
Black African	**0.87**	0.03	0.02	0.08
Coloured	0.67	**0.12**	0.05	0.16
Indian/Asian	0.62	0.06	**0.11**	0.22
White	0.49	0.04	0.04	**0.43**

4.2 Evaluating the levels of segregation in Gauteng's municipalities

Interesting dynamics in the patterns of racial composition in Gauteng's municipalities are illustrated in Table 3. The macro/micro segregation ratios for each municipality are provided. All the municipalities have a macro/micro segregation ratio of least 60% between 500 m and 4 km radii except for Randfontein with 58%. These ratios drop considerably when computed between 500 m and 8 km radii. In both cases, metropolitan municipalities (i.e. Johannesburg, Tshwane and Ekurhuleni) have higher macro/micro segregation ratios (above 70% between 500 m and 4 km and above 50% between 500 m and 8 km) compared to the remaining local municipalities except for Emfuleni, which has a 77% macro/micro segregation ratio between 500 m and 4 km and 55% ratio between 500 m and 8 km). Such observations come as no surprise partly because the computation of $\tilde{H}$ is weighted by population density. Furthermore, the four municipalities (i.e. Johannesburg, Tshwane, Ekurhuleni and Emfuleni) are the most densely populated municipalities in Gauteng compared to the rest.

Due to the fact that they occupy large surface areas, respectively, the four municipalities (i.e. Johannesburg, Tshwane, Ekurhuleni and Emfuleni) also display higher levels of segregation compared to the other six remaining municipalities even at a large radius of 8 km. For example, although Randfontein has similar levels of segregation as Johannesburg within a radius of 500 m (i.e. $\tilde{H}$ = 0.55), between 500 m and 4 km radii, the macro/micro segregation ratio of Johannesburg (0.72) is higher than that of Randfontein (0.58). The observed macro/micro ratio between 500 m and 8 km is even more reduced for Randfontein (0.19) than it is for Johannesburg (0.51). A macro/micro segregation ratio of 0.72 for Johannesburg means that 72% of segregation within 500 m radius in Johannesburg is due to the variation in the racial residential segregation observed at 4 km-radius or more. This explanation illustrates how the results in Table 3 should be interpreted. In general, except for Emfuleni, a large extent of segregation observed within small distances (i.e. 500 m radius) in metropolitan municipalities in Gauteng, can be explained by the factors that influence the patterns of racial composition at large distances (i.e. 4 km and beyond). This is a confirmation of the results already highlighted using the macro/micro segregation ratio, that is, the patterns of racial residential segregation observed in metropolitan municipalities are influenced by variations in the racial composition of population groups observed at macro-scale levels. Similar results were obtained by Parry and Van Eeden (2015) who observed that

Table 3. Spatial information theory indices ($\tilde{H}$) and macro/micro ratios of segregation of municipalities in Gauteng.

Geographic scale	$\tilde{H}$ (500 m)	$\tilde{H}$ (4 km)	$\tilde{H}$ (8 km)	Macro/Micro ratio: $\tilde{H}$ (4 km)/ $\tilde{H}$ (500 m)	Macro/Micro ratio: $\tilde{H}$ (8 km)/ $\tilde{H}$ (500 m)
Randfontein	0.55	0.32	0.11	0.58	0.19
Lesedi	0.56	0.34	0.20	0.60	0.35
Mogale	0.63	0.38	0.22	0.60	0.35
Midvaal	0.41	0.25	0.14	0.61	0.34
Merafong	0.50	0.34	0.18	0.68	0.35
Westonaria	0.41	0.29	0.18	0.69	0.45
Johannesburg	0.55	0.40	0.28	0.72	0.51
Ekurhuleni	0.61	0.44	0.31	0.72	0.51
Emfuleni	0.63	0.48	0.34	0.77	0.55
Tshwane	0.60	0.50	0.39	0.83	0.65

Johannesburg displayed patterns of macro-segregation based on the 2011 Census data. One obvious cause of these observed patterns is the legacy of historical settlement patterns imposed by the laws of apartheid pre-1994. It is worth investigating in future studies other structural processes that relate to labour and other socio-economic factors that can cause macro-segregation as observed in the metropolitan municipalities of Gauteng. Compared to metropolitan municipalities, local municipalities are relatively less influenced by the patterns of segregation observed at macro-scale levels.

The spatial segregation profile curves in Figure 6 illustrate the levels of segregation at a range of distances *r* between 500 m and 8 km increasing at 20 m radius increment for each of the municipalities. Although, the spatial segregation curves suggest a decline in the levels of segregation as *r* increases, in general terms, there are some subtle micro-scale variation in the levels of segregation that might not be easily explained by the segregation profile curves or the macro/micro segregation ratios.

Table 4 presents the results of the spatial exposure/isolation segregation index for neighbourhoods within 500 m and 4 km radii to explain some of the subtle variations that might be occurring at smaller geographic scales. As illustrated in Table 4, the levels of isolation of the Black African population group have remained relatively high (above 0.70) for neighbourhoods whether within 500 m or 4 km radii. The levels of Black Africans' exposure to Whites has also remained relatively unchanged in all the municipalities. However, the levels of exposure of the White population to Black Africans has remarkably increased in neighbourhoods within 4 km radius in all municipalities in Gauteng. These results imply that at smaller geographic scales, there are more Black Africans in the local environments of Whites than there are Whites in the local environments of Black Africans. Even though this observation might not be a new insight while referring to the literature on racial segregation in Gauteng, it, however, illustrates how these observed patterns of segregation and racial integration can be quantified and explained objectively based on a spatial index of segregation.

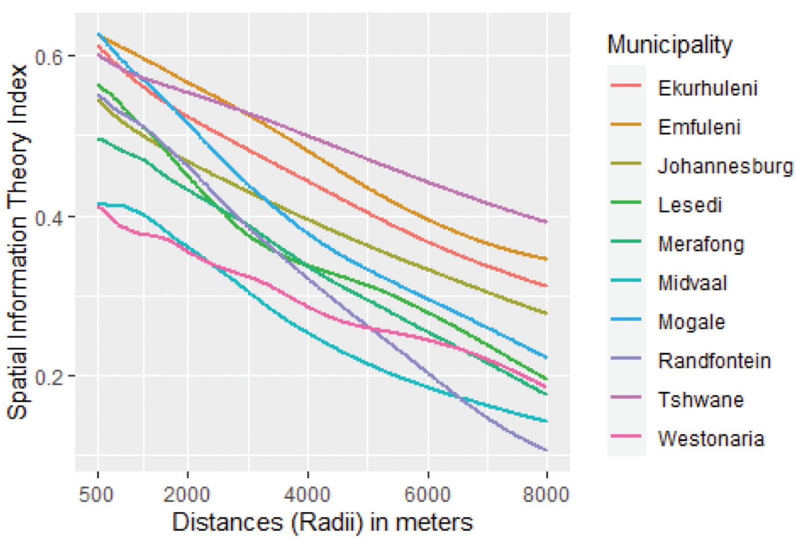

Figure 6. Segregation profile curve between 500 m and 8 km radius for individual municipalities in Gauteng based on StatsSA 2011 Census.

Table 4. Spatial exposure/isolation segregation index per municipality at 500 m radius based on small area dataset.

	500 m					4 km			
		Ekurhuleni					**Ekurhuleni**		
	Black African	Coloured	Indian/Asian	White		Black African	Coloured	Indian/Asian	White
Black African	**0.93**	0.01	0.01	0.05	Black African	0.90	0.02	0.01	0.06
Coloured	0.37	**0.44**	0.04	0.15	Coloured	0.64	0.15	0.03	0.18
Indian/Asian	0.32	0.05	**0.28**	0.34	Indian/Asian	0.51	0.04	0.11	0.34
White	0.23	0.03	0.05	**0.69**	White	0.36	0.03	0.05	0.56
		Emfuleni					**Emfuleni**		
	Black African	Coloured	Indian/Asian	White		Black African	Coloured	Indian/Asian	White
Black African	**0.95**	0.01	0.00	0.04	Black African	0.93	0.01	0.00	0.06
Coloured	0.59	**0.23**	0.02	0.16	Coloured	0.59	0.13	0.13	0.15
Indian/Asian	0.28	0.02	**0.54**	0.15	Indian/Asian	0.39	0.16	0.30	0.14
White	0.30	0.02	0.01	**0.67**	White	0.43	0.02	0.01	0.54
		Johannesburg					**Johannesburg**		
	Black African	Coloured	Indian/Asian	White		Black African	Coloured	Indian/Asian	White
Black African	**0.91**	0.02	0.02	0.05	Black African	**0.88**	0.04	0.03	0.06
Coloured	0.33	**0.50**	0.05	0.11	Coloured	0.55	**0.27**	0.06	0.13
Indian/Asian	0.33	0.06	**0.41**	0.21	Indian/Asian	0.48	0.07	**0.25**	0.20
White	0.31	0.05	0.08	**0.55**	White	0.41	0.06	0.08	**0.45**
		Lesedi					**Lesedi**		
	Black African	Coloured	Indian/Asian	White		Black African	Coloured	Indian/Asian	White
Black African	**0.91**	0.01	0.01	0.07	Black African	**0.88**	0.01	0.01	0.10
Coloured	0.63	**0.03**	0.03	0.30	Coloured	0.65	**0.02**	0.03	0.31
Indian/Asian	0.32	0.03	**0.32**	0.33	Indian/Asian	0.53	0.02	**0.05**	0.41
White	0.28	0.02	0.02	**0.68**	White	0.46	0.02	0.03	**0.49**
		Merafong					**Merafong**		
	Black African	Coloured	Indian/Asian	White		Black African	Coloured	Indian/Asian	White
Black African	**0.93**	0.01	0.00	0.06	Black African	**0.91**	0.01	0.00	0.08
Coloured	0.71	**0.16**	0.00	0.13	Coloured	0.65	**0.07**	0.00	0.28
Indian/Asian	0.66	0.01	**0.02**	0.31	Indian/Asian	0.71	0.01	**0.01**	0.27
White	0.44	0.01	0.01	**0.54**	White	0.58	0.02	0.01	**0.39**
		Midvaal					**Midvaal**		
	Black African	Coloured	Indian/Asian	White		Black African	Coloured	Indian/Asian	White
Black African	**0.79**	0.02	0.01	0.19	Black African	**0.74**	0.02	0.01	0.24
Coloured	0.54	**0.05**	0.01	0.39	Coloured	0.56	**0.04**	0.01	0.39
Indian/Asian	0.47	0.03	**0.05**	0.45	Indian/Asian	0.53	0.03	**0.03**	0.42
White	0.29	0.02	0.01	**0.68**	White	0.38	0.02	0.01	**0.59**

(*Continued*)

Table 4. (Continued).

	500 m					4 km			
		Mogale					**Mogale**		
	Black African	Coloured	Indian/Asian	White		Black African	Coloured	Indian/Asian	White
Black African	**0.92**	0.01	0.01	0.07	Black African	**0.86**	0.01	0.02	0.11
Coloured	0.58	**0.02**	0.03	0.36	Coloured	0.62	**0.01**	0.02	0.35
Indian/Asian	0.20	0.01	**0.70**	0.08	Indian/Asian	0.80	0.01	**0.11**	0.08
White	0.25	0.01	0.01	**0.73**	White	0.42	0.01	0.01	**0.56**
		Randfontein					**Randfontein**		
	Black African	Coloured	Indian/Asian	White		Black African	Coloured	Indian/Asian	White
Black African	**0.86**	0.05	0.00	0.09	Black African	**0.80**	0.08	0.00	0.12
Coloured	0.32	**0.62**	0.01	0.06	Coloured	0.64	**0.27**	0.00	0.09
Indian/Asian	0.49	0.12	**0.02**	0.37	Indian/Asian	0.53	0.10	**0.01**	0.37
White	0.32	0.03	0.01	**0.65**	White	0.41	0.05	0.01	**0.53**
		Tshwane					**Tshwane**		
	Black African	Coloured	Indian/Asian	White		Black African	Coloured	Indian/Asian	White
Black African	**0.92**	0.01	0.01	0.07	Black African	**0.90**	0.01	0.01	0.08
Coloured	0.37	**0.39**	0.03	0.21	Coloured	0.52	**0.20**	0.02	0.25
Indian/Asian	0.32	0.03	**0.38**	0.27	Indian/Asian	0.48	0.02	**0.22**	0.28
White	0.25	0.02	0.02	**0.71**	White	0.31	0.03	0.03	**0.63**
		Westonaria					**Westonaria**		
	Black African	Coloured	Indian/Asian	White		Black African	Coloured	Indian/Asian	White
Black African	**0.95**	0.01	0.00	0.05	Black African	**0.94**	0.01	0.00	0.05
Coloured	0.84	**0.02**	0.01	0.13	Coloured	0.87	**0.01**	0.00	0.11
Indian/Asian	0.80	0.01	**0.02**	0.17	Indian/Asian	0.85	0.01	**0.01**	0.13
White	0.59	0.01	0.01	**0.39**	White	0.72	0.01	0.01	**0.27**

4.3 Racial residential segregation at ward level in 1996, 2001 and 2011

To compare the levels of racial residential segregation across time, the population group dataset aggregated by ward for the years 1996, 2001 and 2011 was analysed. By examining the segregation profile curves drawn based on the results of the spatial information theory index ($\tilde{H}$) for the three periods (i.e. 1996, 2001 and 2011), it is evident that the levels of segregation have steadily declined between 1996 and 2011.

A closer look at the segregation profile curves between 500 m and 4 km radii highlights some interesting dynamics in the observed patterns of segregation with the ward dataset (see Figure 7). Between 500 m and 1 km, the profile curves have a flat shape suggesting the presence of macro-segregation at play (refer Figure 7). This means that the levels of segregation experienced by an individual within a radius of 500 m would also be the same as within a radius of 1 km. Although these results corroborate some of the findings already highlighted with the small area and subplace datasets, they may hide some subtle variations at micro-scale levels that can be uncovered with the availability of fine-grained demographic data at a microdata level. However, decreasing levels of segregation can be observed at scales beyond 2 km radius where the slope of the segregation profile curve as illustrated in Figure 7 starts to decrease.

Figures 8, 9, 10 and 11 illustrated by dot density maps based on the Census datasets collected in 1996 and 2011 aggregated per ward, provide visual illustrations which corroborate these observed dynamic racial residential patterns. Figures 10 and 11 are provided in Appendix B (see section 7) to visualize changes in the racial composition of Johannesburg from 1996 to 2011. The four maps (i.e. Figures 8, 9, 10 and 11) provide a compelling depiction that summarizes the discussions on the spatial segregation profile curves and the macro/micro segregation ratios. The 'BA' population group is depicted in black colour, the 'W' in red, the 'IA' in blue and the 'C' in green. Compared to 1996 (Figure 8), in 2011

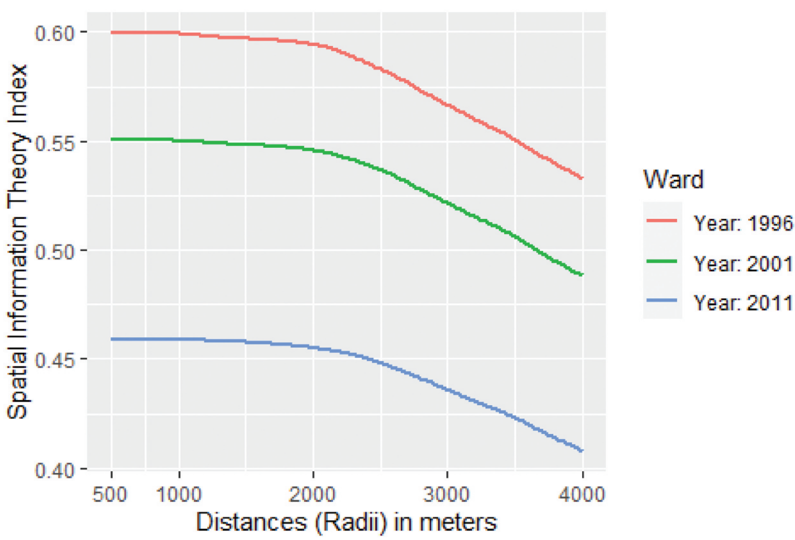

Figure 7. Segregation profile curves between 500 m and 4 km radius based on Census 1996, 2001 and 2011 dataset aggregated per ward for Gauteng.

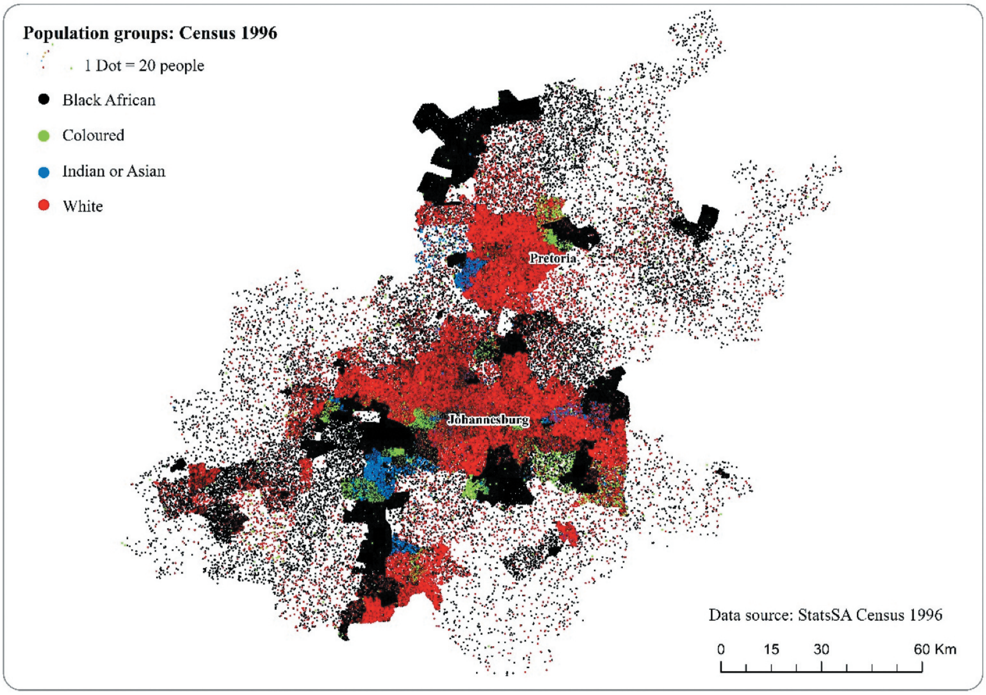

Figure 8. Dot density map of population (Census 1996).

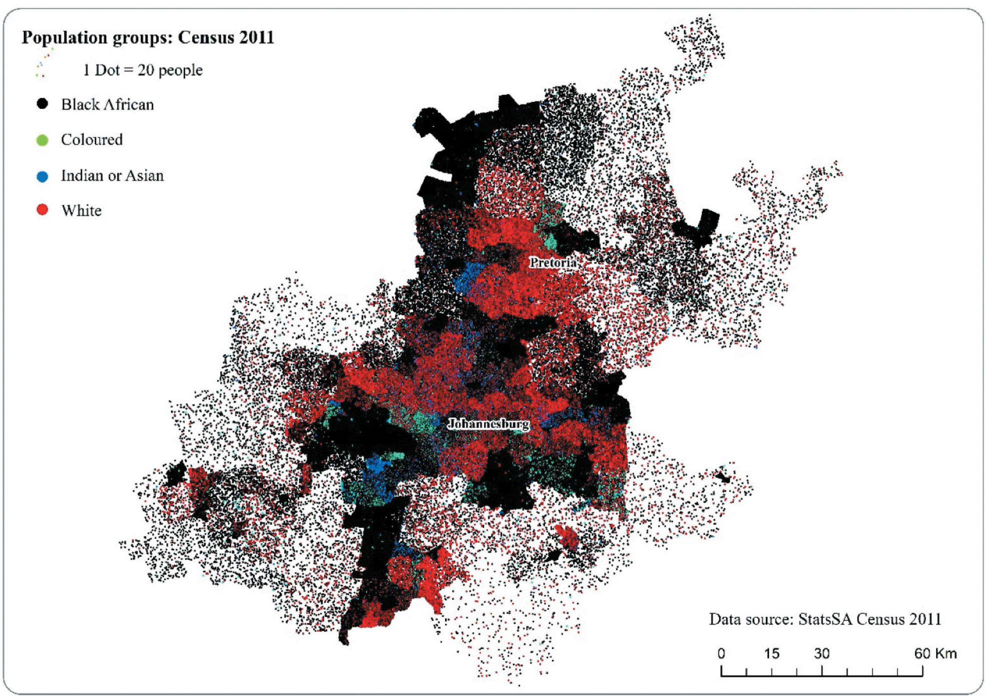

Figure 9. Dot density map of population (Census 2011).

Table 5. Spatial exposure/isolation segregation index per municipality at 500 m and 4 km radii based on ward datasets: 1996, 2001 & 2011.

500 m					4 km				
Year: 1996	Black African	Coloured	Indian/ Asian	White	Year: 1996	Black African	Coloured	Indian/ Asian	White
Black African	**0.89**	0.02	0.01	0.08	Black African	**0.89**	0.02	0.01	0.08
Coloured	0.33	**0.46**	0.04	0.17	Coloured	0.46	**0.32**	0.04	0.18
Indian/ Asian	0.35	0.07	**0.39**	0.19	Indian/ Asian	0.44	0.07	**0.31**	0.19
White	0.26	0.03	0.02	**0.70**	White	0.29	0.03	0.02	**0.66**
500 m					**4 km**				
Year: 2001	Black African	Coloured	Indian/ Asian	White	Year: 2001	Black African	Coloured	Indian/ Asian	White
Black African	**0.89**	0.02	0.01	0.08	Black African	**0.88**	0.02	0.01	0.08
Coloured	0.40	**0.40**	0.03	0.16	Coloured	0.51	**0.28**	0.04	0.17
Indian/ Asian	0.41	0.05	**0.31**	0.22	Indian/ Asian	0.49	0.06	**0.24**	0.21
White	0.30	0.03	0.03	**0.64**	White	0.33	0.04	0.03	**0.61**
500 m					**4 km**				
Year: 2011	Black African	Coloured	Indian/ Asian	White	Year: 2011	Black African	Coloured	Indian/ Asian	White
Black African	**0.88**	0.02	0.02	0.08	Black African	**0.88**	0.02	0.02	0.08
Coloured	0.48	**0.31**	0.04	0.16	Coloured	0.57	**0.22**	0.04	0.17
Indian/ Asian	0.49	0.05	**0.22**	0.24	Indian/ Asian	0.54	0.05	**0.17**	0.23
White	0.38	0.04	0.05	**0.54**	White	0.41	0.04	0.04	**0.51**

(Figure 9), there is an observed increased amount of a mixture of colours that can be seen particularly in the central parts of Tshwane, in subplaces such as Akasia and Pretoria mainplace, and in the southern parts of Tshwane, in mainplaces such as Centurion. The northern parts of Johannesburg have also experienced a steady increase in the levels of racial residential integration in mainplaces such as Midrand, Randburg and Sandton (see Figures 10 and 11). The same patterns can also be observed in the central and northern parts of Ekurhuleni in mainplaces such as Kempton Park, Benoni and Germiston, suggesting some form of racial integration among the four population groups taking place in those identified areas. This evidently explains the decline in the levels of racial residential segregation in Gauteng between 1996 and 2011 as also illustrated by the segregation profile curves in Figure 7. Distinct large patches of homogenous colours representing each of the different population groups as can also be seen in the maps (Figures 8, 9, 10 and 11), could also be explained by relatively high macro/micro segregation ratios. Broadly speaking, the province of Gauteng is characterized by racial residential segregation occurring at a macro-scale level.

A further investigation of the spatial exposure/isolation segregation index results across space (i.e. for scales of analysis within 500 m and 4 km) and time (i.e. in 1996, 2001 and 2011) sheds some light in terms of understanding the patterns of racial integration that are taking place in certain areas of Gauteng province. By examining Table 5, it can readily be observed that the levels of isolation for the 'BA' population group, has remained the same (0.89) across space (i.e. whether for neighbourhoods within 500 m or 4 km) and across time (i.e. whether in 1996, 2001 or 2011).

Differences are very marginal. While the levels of 'BA' exposure to 'W' has remained the same across space (for scale of analysis within 500 m and 4 km) and time (i.e. in 1996, 2001 and 2011), the levels of exposure of 'W' to 'BA' has steadily increased across both space and time. This suggests that the observed patterns of racial integration are taking place in spaces that were previously reserved for the 'W' population group. This observation is confirmed by the maps illustrated in Figures 8, 9, 10, 11 and 12 showing patterns of racial integration happening in areas (e.g., central and southern parts of Tshwane and also in the northern parts of Johannesburg and central and northern parts of Ekurhuleni) that were previously predominantly occupied by the 'W' population group. Christopher (2005) also observed similar patterns of racial integration in the former 'White' areas when measuring the levels of racial segregation in the provinces of South Africa between 1996 and 2001. It is important to investigate the upcoming Census 2021 data to establish whether this trend will be sustained.

5. Conclusion

This study quantified racial residential segregation based on four racial population groups in Gauteng (South Africa) between 1996 and 2011. The spatial information theory segregation index $\tilde{H}$ was used to explore the evenness and clustering dimensions of segregation. The spatial segregation profiles and macro/micro segregation ratios based on $\tilde{H}$ provided a better understanding of whether the observed racial residential segregation types were caused by racial residential patterns observed at a macro or micro-scale level. The exposure and isolation dimensions of racial residential segregation were explored using the spatial exposure/isolation segregation index $\tilde{P}^*$. The results of both indices ($\tilde{P}^*$ and $\tilde{H}$) were confirmed by population density maps illustrating the distribution patterns of the four population groups in the study area. Although the datasets employed in this study have been aggregated at the small area, subplace and ward levels, the methodology used in the computation of the two indices of segregation made it possible to determine on average the levels of segregation an individual in Gauteng would experience within radii of 500 m, 4 km and 8 km of his/her residential location. Overall, the levels of racial residential segregation steadily declined between 1996 and 2011. While the levels of exposure of Whites to Black Africans continue to increase, the levels of exposure of Black Africans to Whites remained unchanged.

Metropolitan municipalities and the local municipality of Emfuleni display high levels of racial residential segregation, mainly caused by macro-scale factors. The remaining local municipalities have lower levels of segregation as compared to the metropolitan municipalities. Segregation profile curves and macro/micro segregation ratios provided means for determining the levels of segregation at well-defined scales of analysis. The results of the spatial exposure/isolation segregation index across time and geographic scales combined with dot-density maps helped with the identification of areas where racial residential integration occurred. Hence, the two measures of spatial segregation enabled the identification of patterns of racial residential segregation in Gauteng province (South Africa) across space and time in a way that has not been explored in previous studies.

Similar to other studies discussed in this paper, the results of this study have demonstrated that racial residential segregation manifests itself over large scales of neighbourhoods in Gauteng province. Besides some of the housing policies (e.g., housing programme of the RDP) that have permitted housing provisions for the poor (who are mostly Black Africans) to take place in isolated areas far from the centres of economic activities in the province, this observed pattern of macro-segregation can also be attributed to the legacy of segregated settlements established during the apartheid era. Nevertheless, some areas, particularly those that were reserved for the White population group during apartheid, have observed a steady increase in the levels of racial residential integration with residents (mostly middle to high-income earners) of all population groups. These observations have prompted authors such as Hamann and Horn (2015) and Parry and Van Eeden (2015) to suggest policies that would redress socio-economic imbalances in order to address racial residential segregation more effectively. For example, careful replications of successful implementations of existing inclusionary/mixed-income housing policies hold the potential of achieving meaningful and sustainable residential integration. However, there is a need to monitor the implementation and outcomes of such policies on a continuous basis. This study puts at the disposal of policy-makers or any interested party, a quantitative tool for monitoring and evaluating policies that are meant for bringing about racial residential integration in neighbourhoods where it is needed across the province of Gauteng. This is because the application of spatial indices of segregation as employed in this study enables the measurement, quantification and comparison of the levels of racial residential segregation at varying defined scales of analysis across the province of Gauteng. This is a novel methodology for measuring the levels of segregation in the context of South Africa. The results of this study can serve as baseline for comparison with future analyses which will be based on up-to-date Census data in order to evaluate the effectiveness of policies meant for addressing racial residential segregation. Furthermore, the quantitative analysis performed in this paper provides some background against which future qualitative studies on racial residential segregation can be contextualized.

This study analysed residential segregation, i.e. based on where someone lives. To better understand the level of exposure of different population groups to each other, one could analyse segregation based on the locations that people visit daily or weekly, e.g., discernible from mobility data from cell phone service providers. Further work will investigate the causes and factors that influence the patterns of residential segregation observed in Gauteng and in other provinces of South Africa.

Acknowledgments

Samy Katumba would like to thank Dr Richard Ballard for the comments he provided in the early versions of the manuscript. The authors would also like to thank the anonymous reviewers for their insightful comments and suggestions.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Samy Katumba http://orcid.org/0000-0002-6530-7690
Serena Coetzee http://orcid.org/0000-0001-8683-8047
Alfred Stein http://orcid.org/0000-0002-9456-1233
Inger Fabris-Rotelli http://orcid.org/0000-0002-2192-4873

References

Barros, J., & Feitosa, F. F. (2018). Uneven geographies: Exploring the sensitivity of spatial indices of residential segregation. *Environment and Planning B: Urban Analytics and City Science*, *45*(6), 1073–1089. https://doi.org/10.1177/2399808318760572

Benassi, F., Iglesias-Pascual, R., & Salvati, L. (2020). Residential segregation & social diversification: Exploring spatial settlement patterns of foreign population in Southern European cities. *Habitat International*, 101(2020), 102200. https://doi.org/10.1016/j.habitatint.2020.102200

Chodrow, P. S. (2017). Structure and information in spatial segregation. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, *114*(44), 11591-11596. https://doi.org/10.1073/pnas.1708201114

Christopher, A. J. (2001). Urban segregation in post-apartheid South Africa. *Urban Studies*, *38*(3), 449–466. https://doi.org/10.1080/00420980120080031

Christopher, A. J. (2005). The slow pace of desegregation in South African cities, 1996-2001. *Urban Studies*, *42*(12), 2305–2320. https://doi.org/10.1080/00420980500332122

Duncan, O. T., & Duncan, B. (1955). A methodological analysis of segregation indices. *American Sociological Review*, *20*(2), 210–217. https://doi.org/10.2307/2088328

Esri (2008). *Dot density mapping with arcmap: Part 1*, https://www.esri.com/arcgis-blog/products/product/mapping/dot-density-mapping-with-arcmap-part-1/,Accessed,11/02/2019

Feitosa, F. F., Camara, G., Monteiro, A. M. V., Koschitzki, T., & Silva, M. P. S. (2007). Global & local spatial indices of urban segregation. *International Journal of Geographical Information Science*, *21*(3), 299–323. https://doi.org/10.1080/13658810600911903

Gauteng Provincial Government (GPG) (2011). *Gauteng Spatial Development Framework (GSDF)*, February.

Hamann, C., & Horn, A. C. (2015). Continuity or discontinuity? Evaluating the changing socio-spatial structure of the city of Tshwane, South Africa. *Urban Forum*, *26*(1), 39–57. https://doi.org/10.1007/s12132-014-9231-7

Harrison, P., & Todes, A. (2015). Spatial transformation in a "loosening state": South Africa in a comparative perspective. *Geoforum*, *61*(2015), 148–162. https://doi.org/10.1016/j.geoforum.2015.03.003

Hennerdal, P., & Nielsen, M. M. (2017). A multiscalar approach for identifying clusters and segregation patterns that avoids the modifiable areal unit problem. *Annals of the American Association of Geographers*, *107*(3), 555–574. https://doi.org/10.1080/24694452.2016.1261685

Hong, S. Y., O'Sullivan, D., & Sadahiro, Y. (2014). Implementing spatial segregation measures in R. *PLOS ONE*, *9*(11), e113767. https://doi.org/10.1371/journal.pone.0113767

Horn, A. (2005). Measuring multi-ethnic spatial segregation in South African cities. *South African Geographical Journal*, *87*(1), 58–72. https://doi.org/10.1080/03736245.2005.9713827

Horn, A. C., & Ngcobo, J. R. B. (2003). The suburban challenge: (de)segregation, opportunity and community in akasia, city of Tshwane. *Urban Forum*, *14*(4), 320–346. https://doi.org/10.1007/s12132-003-0017-6

Lee, B. A., Matthews, S. A., Matthews, S. A., Reardon, S. F., Farrell, C. R., & O'Sullivan, D. (2008). Beyond the census tract: Patterns and determinants of racial segregation at multiples geographic scales. *American Sociological Review*, *73*(5), 766–791. https://doi.org/10.1177/000312240807300504

Lemanski, C., Landman, K., & Durington, M. (2008). Divergent and similar experiences of 'gating' in South Africa: Johannesburg, durban and cape town. *Urban Forum*, *19*(2), 133–158. https://doi.org/10.1007/s12132-008-9030-0

Massey, D. S., & Denton, N. A. (1988). The dimensions of residential segregation. *Social Forces*, *67* (2), 281–315. https://doi.org/10.2307/2579183

Monkkonen, P., & Zhang, X. (2014). Innovative measurement of spatial segregation: Comparative evidence from Hong Kong and San Francisco. *Regional Science and Urban Economics*, *47*(2014), 99–111. https://doi.org/10.1016/j.regsciurbeco.2013.09.016

Morrill, R. L. (1991). On the measure of geographical segregation. *Geography Research Forum*, *11* (1991), 25–36. https://grf.bgu.ac.il/index.php/GRF/article/view/91

O'Sullivan, D., & Wong, D. (2007). A Surface based approach to measuring spatial segregation. *Geographical Analysis*, *39*(2), 147–168. https://doi.org/10.1111/j.1538-4632.2007.00699.x

Oka, M., & Wong, D. W. S. (2015). Spatializing segregation measures: An approach to better depict social relationships. *Cityscape: A Journal of Policy Development and Research*, *17*(1), 97–113. https://www.huduser.gov/portal/periodicals/cityscpe/vol17num1/ch8.pdf

Openshaw, S. (1984). *The modifiable areal unit problem: Concepts and techniques in modern geography*. Geoabstracts.

Parnell, S., & Pirie, G. (1991). Johannesburg. In *Anthony Lemon (eds), Homes apart: South Africa's segregated cities* (pp. 129–145). Bloomington: Indiana University Press.

Parry, K., & Van Eeden, A. (2015). Measuring racial residential segregation at different geographic scales in cape town and johannesburg. *South African Geographical Journal*, *97*(1), 31–49. https://doi.org/10.1080/03736245.2014.924868

Reardon, S. F., Matthews, S. A., O'Sullivan, D., Lee, B. A., Firebaugh, G., Farrell, C. R., & Bischoff, K. (2008). The geographic scale of metropolitan racial segregation. *Demography*, *45* (3), 489–514. https://doi.org/10.1353/dem.0.0019

Reardon, S. F., & O'Sullivan, D. (2004). Measures of Spatial Segregation. *Sociological Methodology*, *34*(1), 121–162. https://doi.org/10.1111/j.0081-1750.2004.00150.x

Sakoda, J. M. (1981). A Generalized index of dissimilarity. *Demography*, *18*(2), 245–250. https://doi.org/10.2307/2061096

South African History Online (SAHO) (2020). *Apartheid legislation 1850s-1970s* https://www.sahistory.org.za/ [Accessed 09/ December/2020]

StatsSA. (2011). *Census 2011 questionnaire A*. http://www.statssa.gov.za/?page_id=3852 [Accessed 09/ December/2020]

StatsSA. 2019. *Four facts about our provincial economies*, http://www.statssa.gov.za/?p=12056, [Access 02/December/2019]

StatsSA Community Survey. (2016). *Statistical release P0301: Community survey*, http://cs2016.statssa.gov.za/wp-content/uploads/2016/07/NT-30-06-2016-RELEASE-for-CS-2016-_Statistical-releas_1-July-2016.pdf, [Accessed 02/December/2019]

Todes, A., Weakly, D., & Harrison, P. (2018). Densifying johannesburg: Context, policy and diversity. *Journal of Housing and the Built Environment*, *33*(2), 281–299. https://doi.org/10.1007/s10901-017-9561-6

Weir-Smith, G. (2016). Changing boundaries: Overcoming modifiable areal unit problems related to unemployment data in South Africa. *South African Journal of Science*, *112*(3/4), 1–8. https://doi.org/10.17159/sajs.2016/20150115

Wellings, P., & Black, A. (1986). Industrial decentralization under apartheid: The relocation of industry to the South African periphery. *World Development*, *14*(1), 1–38. https://doi.org/10.1016/0305-750X(86)90094-X

White, M. J. (1983). The measurement of spatial segregation. *The American Journal of Sociology*, *88* (5), 1008–1018. https://doi.org/10.1086/227768

Wong, D. W. S. (1993). Spatial indices of segregation. *Urban Studies*, *30*(3), 559–572. https://doi.org/10.1080/00420989320080551

Wong, D. W. S. (1998). Measuring multiethnic spatial segregation. *Urban Geography*, *19*(1), 77–78. https://doi.org/10.2747/0272-3638.19.1.77

Wong, D. W. S. (2002). Spatial measures of segregation and GIS, urban geography. *23*(1), 85–92. https://doi.org/10.2747/0272-3638.23.1.85
Wong, D. W. S. (2003). Implementing spatial segregation measures in GIS. *Computers, Environmental and Urban Systems*, *27*(1), 53–70. https://doi.org/10.1016/S0198-9715(01)00018-7
Wong, D. W. S. (2004). Comparing traditional and spatial segregation measures: A spatial scale perspective. *Urban Geography*, *25*(1), 66–82. https://doi.org/10.2747/0272-3638.25.1.66
Wong, D. W. S. (2005). Formulating a general spatial segregation measure. *The Professional Geographer*, *57*(2), 285–294. https://doi.org/10.1111/j.0033-0124.2005.00478.x
Wong, D. W. S., & Chong, W. K. (1998). Using spatial segregation measures in GIS and statistical modeling packages. *Urban Geography*, *19*(5), 477–485. https://doi.org/10.2747/0272-3638.19.5.477
Yao, J., Wong, D. W. S., Bailey, N., & Minton, J. (2019). Spatial segregation measures: A methodological review. *Journal of Economic and Social Geography*, *110*(3), 235–250. https://doi.org/10.1111/tesg.12305

Appendix A

Table A1. Spatial and Aspatial indices of segregation

Dimension	Aspatial Indices	(Related) Spatial Indices
Evenness/ Clustering	Dissimilarity Index (D) $$D = \frac{1}{2}\sum_i \left\lvert \frac{a_i}{A} - \frac{b_i}{B} \right\rvert$$ a_i: population of group a in areal unit i. b_i: rest of the other population groups in areal unit i. A: total population of group a in the entire study area. B: total population of other groups in the entire study area.	Adjusted Dissimilarity Index (D(adj)) $$D(adj) = D - \frac{\sum_i \sum_j \lvert c_{ij}(z_i - z_j) \rvert}{\sum_i \sum_j c_{ij}}$$ z_i: proportion of population group z in areal unit i which is adjacent to areal unit j z_j z_j: *proportion of population group z in areal unit j*which is adjacent to areal unit *i* c_{ij}: is a spatial term in the connectivity matrix. c_{ij} has a value 1 when *i* and *j* are adjacent and 0 otherwise. Boundary-Adjusted Segregation Index D(w) $$D(w) = D - \frac{\sum_i \sum_j w_{ij} \lvert z_i - z_j \rvert}{\sum_i \sum_j w_{ij}}$$ $$w_{ij} = \frac{d_{ij}}{\sum_j d_{ij}}$$ d_{ij}: is the length of the boundary separating areal unit i and j Spatial segregation Index D(s) $$D(s) = D - \sum_i \sum_j \left\{ \frac{w_{ij} \lvert z_i - z_j \rvert}{\sum_i \sum_j w_{ij}} x \frac{\frac{1}{2}\left(\frac{P_i}{A_i} + \frac{P_j}{A_j}\right)}{MAX\left(\frac{P}{A}\right)} \right\}$$ A_i: areaofarealunit *i* P_i: perimeter of areal unit *i* A_j: area of areal unit *j* P_j: perimeter of areal unit *j* $MAX\left(\frac{P}{A}\right)$: *maximum perimeter − area ratio in the study region*

(*Continued*)

Table A1. (Continued).

Dimension	Aspatial Indices	(Related) Spatial Indices
	Generalised Index of Dissimilarity *D*(m) for multiple groups. $$D(m) = \frac{1}{2}\frac{\sum_i \sum_j \lvert N_{ij} - E_{ij} \rvert}{\sum_j NP_{.j}(1 - P_{.J})}$$ $$E_{ij} = \frac{N_{i.}N_{.J}}{N}$$ N_{ij}: population count of population group j in areal unit i. $N_{i.}$: total population count in areal unit i $N_{.J}$: total population count of group *j* *N*: total population count in the entire study area $P_{.J}$: proportion of population in group *j* E_{ij}: expected population size of group j in areal unit *i* assuming a uniform population distribution Theil Information Theory Index (*H*) $$H = \sum_i \frac{T_i(E_s - E_i)}{T_s E_s}$$ E_s: entropy score for the entire study area E_i: entropy score for areal unit *i* T_i: total population in areal unit *i* T_s: total population of the entire study area $$Entropy = \sum_k P_k \ln\left(\frac{1}{P_k}\right)$$ P_k: proportion of population of group *k* (whether with respect to the total population for a given areal unit or with respect to the entire study area).	Multigroup spatial index of segregation, *SD* (m) $$SD(m) = \frac{1}{2}\frac{\sum_i \sum_j \lvert CN_{ij} - CE_{ij} \rvert}{\sum_j CNxCP_{.j}(1 - CNP_{.j})}$$ $$CE_{IJ} = \frac{CN_{i.}CN_{.j}}{CN}$$ CN_{ij}: Composite population count of group *j* in areal unit *i*. It includes population counts of the same group in all neighbouring units. $CN_{i.}$: total composite population count in areal unit *i* $CN_{.j}$: total composite count of group *j* *CN*: total composite population in the entire study area $CP_{.j}$: proportion of composite population in group *j*
Exposure/ Isolation	Interaction Index(xP_y^*) $$xP_y^* = \sum_i \left[\left(\frac{x_i}{X}\right)\left(\frac{y_i}{t_i}\right)\right]$$ Isolation Index(xP_x^*) $$xP_x^* = \sum_i \left[\left(\frac{x_i}{X}\right)\left(\frac{x_i}{t_i}\right)\right]$$ x_i: total minority population in areal unit *i* *X*: total minority population in areal unit *i* y_i: total majority population in areal unit *i* t_i: total population in areal unit *i*	

Appendix B

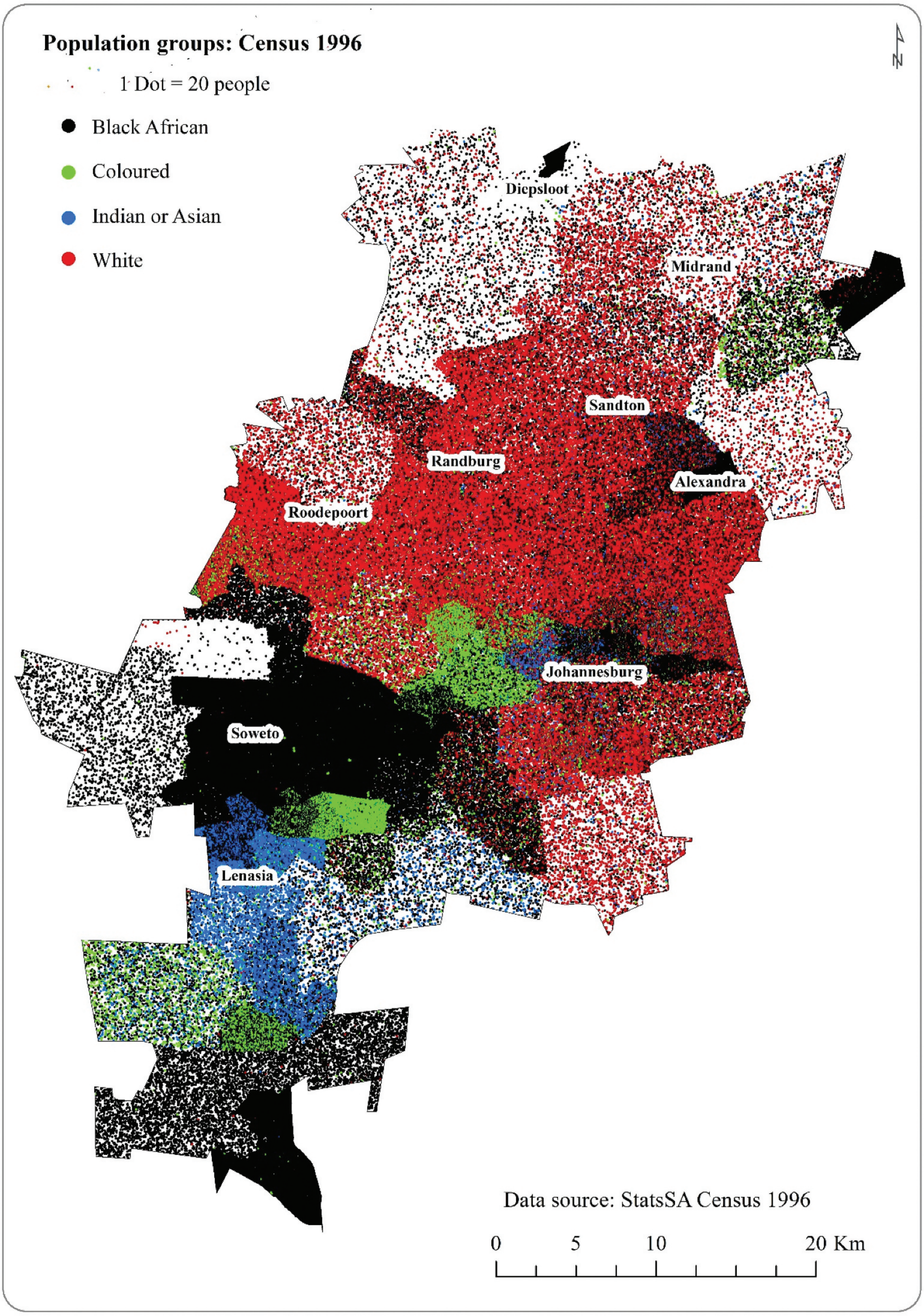

Figure 10. Dot density map of population-Johannesburg (Census 1996).

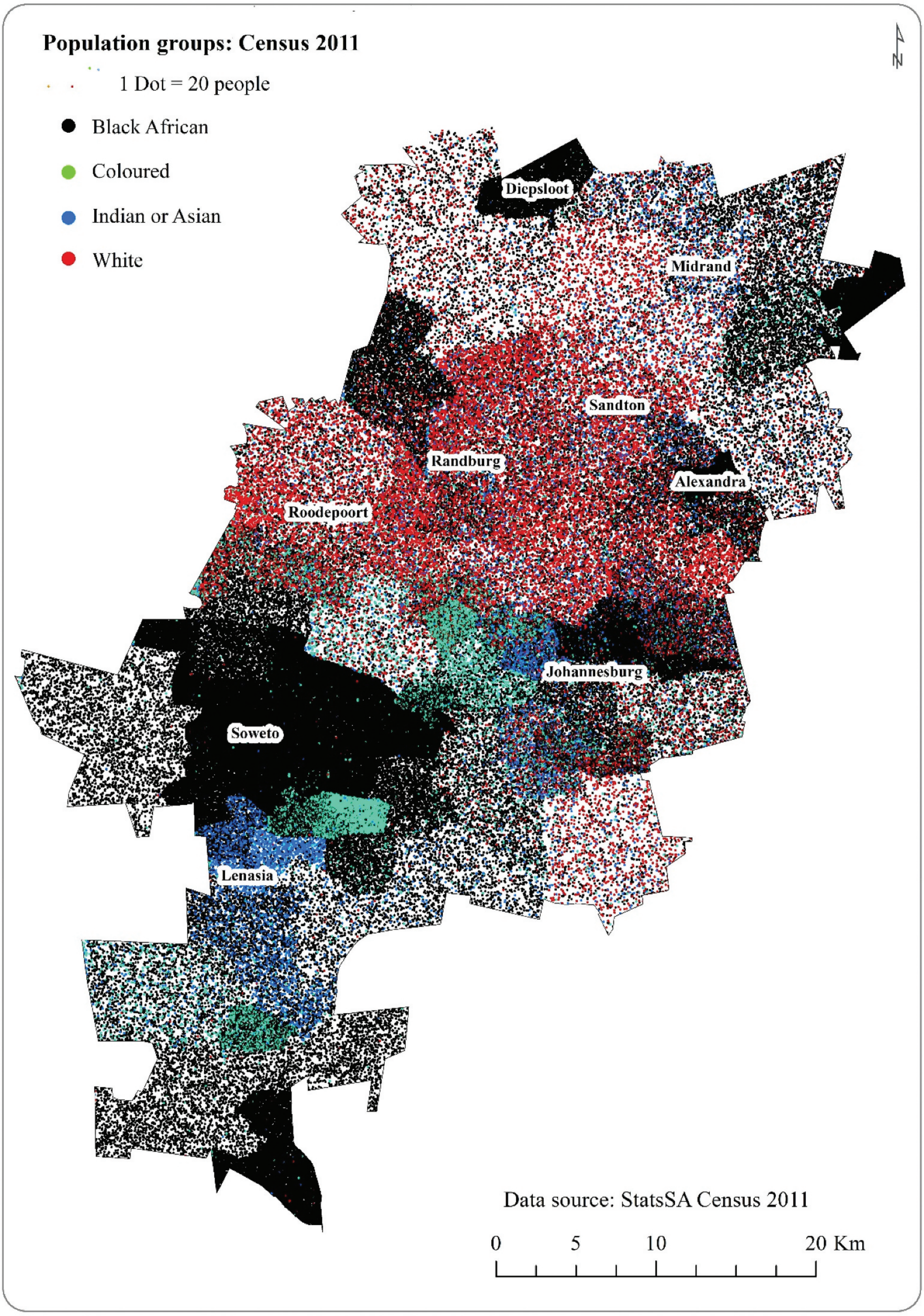

Figure 11. Dot density map of population-Johannesburg (Census 2011).