ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-4-2024
ISPRS TC IV Mid-term Symposium "Spatial Information to Empower the Metaverse", 22–25 October 2024, Fremantle, Perth, Australia

Exploring the Effect of Road Network Structure on Inter-Regional Accessibility in a Diverse Road Network

Renate Thiede[1], Inger Fabris-Rotelli[1], Pravesh Debba[2], Christopher Cleghorn[3]

[1] Department of Statistics, University of Pretoria, Lynnwood Road, Pretoria, South Africa - renate.thiede@up.ac.za, inger.fabris-rotelli@up.ac.za
[2] Council for Scientific and Industrial Research, Meiring Naudé Road, Pretoria, South Africa - pdebba@csir.co.za
[3] School of Computer Science and Applied Mathematics, University of the Witwatersrand, Jan Smuts Ave, Johannesburg, South Africa - chriscleghorn89@gmail.com

Keywords: Road Networks, Accessibility, Homogeneity, Clustering

Abstract

Accessibility analyses quantify the level of access to certain areas or opportunities, such as employment and healthcare facilities. Since public data is often aggregated at the level of regions, such as administrative units, it is useful to quantify accessibility between regions. Many factors influence inter-regional accessibility, most notably the accessibility metric used, and the way in which regions are chosen. This paper investigates the effects of road network structure on accessibility, using a previously developed inter-regional accessibility model that bases its accessibility metric on travel distance via the road network. This paper considers an area within the City of Tshwane municipality in South Africa. We investigate the effects of road structure in two ways. Firstly, regions are chosen based on the road network structure, which is done by extending a previously developed road network clustering algorithm for this novel use. Different spatial scales of regionalisation are considered, and the accessibility between these regions is compared to the accessibility between administrative units within the study area. Secondly, the effect of road network homogeneity on accessibility is investigated, where homogeneity corresponds to a uniform concentration of roads across a region. The results show that although road network homogeneity does not significantly correlate with accessibility, the way in which regions are chosen and their spatial scale has a strong effect on the results of the accessibility model. Our novel method of obtaining regions thus provides fresh insights into road-based accessibility within the City of Tshwane.

1. Introduction

Accessibility analyses provide insight into access to employment, urban green areas, healthcare facilities, and other utilities (Netrdová and Nosek, 2020; Wigley et al., 2020; Quatrini et al., 2019). As such, they are a crucial part of sustainable urban development planning. UN Sustainable Development Goal (SDG) 9 on Industry, Innovation and Infrastructure[1] and SDG 11 on Sustainable Cities and Communities[2] emphasise the need for a better understanding of accessibility to ensure equitable access to opportunities.

Transport infrastructure presents a growing problem in the developing world, particularly in urban areas. The United Nations estimate that 66% of the global population will reside in urban areas by 2050 (Ajami et al., 2019). As of 2018, two thirds of the population in the global south lived in informal settlements (Runsten et al., 2018), which form when a government is unable to provide the housing and infrastructural needs of its population (Kohli et al., 2012). In South Africa, the population living in informal settlements is growing at a faster rate than the population living in formal settlements and towns (Runsten et al., 2018). The creation of informal settlements leads to the formation of informal roads. These roads are created naturally through human and vehicle movement, and are not planned or approved by the government (Thiede et al., 2020). Informal roads in South Africa fill a void created by the legacy of apartheid, during which the transport network was designed to segregate the population rather than integrate it (Giddy, 2019). Figure 1 shows a formal and informal settlement bordering on each other. The informal network exhibits unique spatial characteristics that distinguish it from the formal road network. These include a more irregular pattern, a lack of strictly parallel roads, as well as roads that are shorter and closer together.

Figure 1. An informal settlement (bottom of the image) bordering on a formal settlement (top). The road network is indicated in orange.

Since the characteristics in terms of width, surface type, and network pattern of informal roads differ from formal roads (Nobrega et al., 2006; Thiede et al., 2020), the presence of in-

[1] `https://unstats.un.org/sdgs/metadata/?Text=&Goal=9&Target=9.1`
[2] `https://sdgs.un.org/goals/goal11`

formal roads is likely indicative of inequitable accessibility.

Since municipal community development initiatives typically target regions within the municipality, rather than individual point locations, it is pertinent to study accessibility between regions. Thiede et al. (2023c) developed a model to quantify the relative accessibility between spatial areal units, based on the road network. That study estimated the level of access afforded by an existing road network. The methodology was applied to electoral wards, which are sub-municipal administrative units, in the City of Tshwane municipality, South Africa. However, the boundaries of electoral wards are not designed around the structure of the road network. In reality, people do not consider ward boundaries when travelling. Therefore, the application of (Thiede et al., 2023b), while interesting in its own right, provided a limited understanding of the true state of accessibility in the municipality. Spatial analyses based on regions, or areal units, are typically sensitive to spatial scale and the choice of region boundaries (Viegas et al., 2009), as per the Modifiable Areal Unit Problem (MAUP) (Wong, 2004). This has been shown to affect transport and accessibility studies (Javanmard et al., 2023; Viegas et al., 2009; Clark and Scott, 2014). Alternative regionalisations of the study area, aside from administrative units, should thus be considered for accessibility modelling.

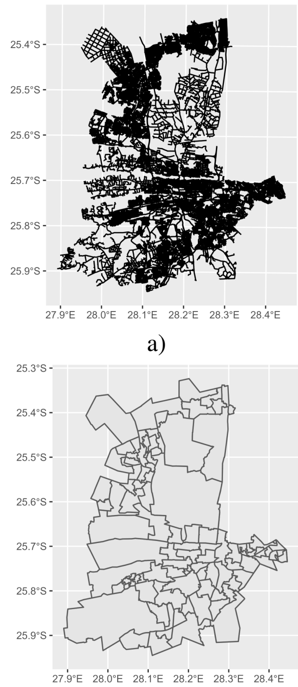
In this paper, we obtain partitions of the municipality based on characteristics of the road network. In particular, this is done using the road network clustering algorithm of (Thiede et al., 2023b). This algorithm divides a spatial area into regions based on the homogeneity of the road network, where a road network is considered homogeneous if its midpoints are distributed with even density across the region under observation (Thiede et al., 2023a). Homogeneity thus encompasses concepts such as uniform road density and length. We propose that obtaining regions based on homogeneity, and then quantifying the accessibility between these regions, will provide new insights into inter-regional accessibility.

Once the study area is subdivided into regions based on the road network, inter-regional accessibility will be estimated using the road network-based accessibility model proposed by (Thiede et al., 2023c). This model makes use of Markov chain theory to reduce the computational cost of traditional network analysis, providing an efficient, computationally simple solution. The model represents the regions as states in a Markov chain, and obtains the probabilities of moving between regions based on the inverse distance via the road network. Relative accessibility from one region to every adjacent region is stored in a Markov chain 1-step transition probability matrix (TPM). In accordance with Markov chain theory, this matrix can be raised to the power n to quantify the accessibility between any two regions in n steps. Letting n tend to infinity results in a matrix with identical rows. This row represents the accessibility of a region regardless of origin, and is called the prominence index (Bavaud, 1998). Finally, accessibility between these regions will be compared to the accessibility between electoral wards, as in (Thiede et al., 2023b).

This paper proceeds as follows. Section 2 describes the data and the nature of the study area. Section 3 outlines the methodology. Section 4 presents and discusses the results, and Section 5 provides the conclusion.

2. Study Area and Data

The study area is the western region of the City of Tshwane municipality in the Gauteng Province, South Africa. The western region consists of mostly urban areas, relevant to this investigation of urban accessibility. The road network within this region is highly diverse, and contains formal and informal urban roads. Informal roads are created by human movement without government permission or planning, and occur mostly in and around informal settlements. The road network data is from Open Street Map. Figure 2(a) shows the road network within the western region of the City of Tshwane municipality. For the sake of comparison to the regions obtained by the homogeneity-based clustering algorithm, we also model accessibility between electoral wards. The region under consideration contains 95 wards, shown in Figure 2(b). The ward boundary data was obtained from the Humanitarian Data Exchange, contributed by the OCHA Regional Office for Southern and Eastern Africa under the Creative Commons Attribution for Intergovernmental Organisations license[3].



a)

b)

Figure 2. The study area, in the western region of the City of Tshwane municipality. a) The road network. b) Electoral ward boundaries within the study area.

An individual road is defined as a stretch of road between either an endpoint and an intersection, or two intersections. Thus, we consider each road segment individually, and represent each road segment by its midpoint. This holds for formal and informal roads.

3. Methodology

This section develops the methodology to identify homogeneous regions and quantify the accessibility between them. Sections 3.1-3.3 provide the required background theory, and Section 3.4 presents the proposed method.

[3] `https://creativecommons.org/licenses/by/3.0/igo/legalcode`

3.1 Road Network Homogeneity

The homogeneity of a road network is tested using the method of (Thiede et al., 2023a), which extended the spatial point pattern homogeneity test of (Kraamwinkel et al., 2018) to line patterns. The road network is represented as a line pattern $L = \{l_1, l_2, ..., l_n\}$, where each line l_i represents a road in the network. Then, the line pattern representation is converted to a point pattern representation $X = \{x_1, x_2, ..., x_n\}$, where x_i corresponds to the midpoint of l_i. Each road is thus represented by its midpoint, and these midpoints constitute a point pattern.

Now, a road network is considered homogeneous if its midpoint pattern X representation is homogeneous. A point pattern is considered homogeneous if it has constant intensity across the spatial domain, where the intensity of a point pattern is the expected number of points per unit area. In other words, a point pattern is homogeneous if points are distributed across an area with more or less equal density throughout. Figure 3 shows an example of homogeneous roads in (a), with their corresponding point pattern in (b), and inhomogeneous roads in (c), with their point pattern representation in (d). With few exceptions, the homogeneous point pattern in (b) exhibits the same level of density throughout, corresponding to roads of generally equal spread and length in (a). The inhomogeneous point pattern in (b), however, exhibits sparser and denser areas, corresponding to roads that are further apart and longer, or closer together and shorter, respectively.

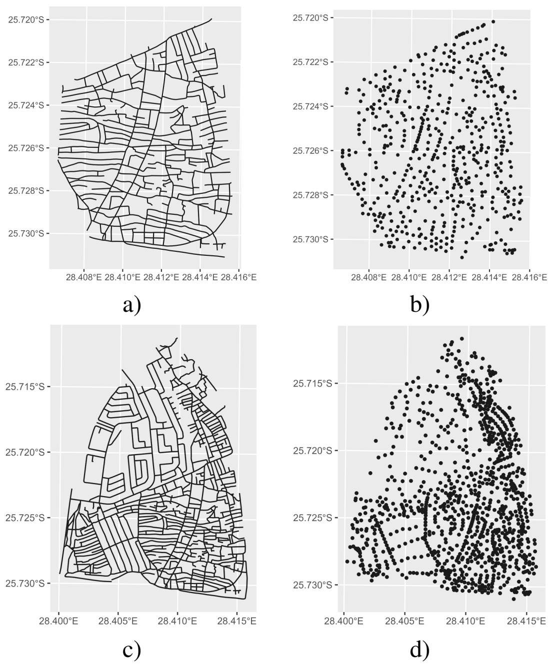


Figure 3. Examples of homogeneous and inhomogeneous road networks with their point patterns. a) A homogeneous road network, with its point pattern in b). c) An inhomogeneous road network, with its point pattern in d).

The hypothesis of homogeneity is tested using the approach of (Kraamwinkel et al., 2018), which modifies the Pearson χ^2 goodness-of-fit test (Potthoff and Whittinghill, 1966) to account for spatial dependence. The spatial area is divided into a grid of $m \times m$ quadrats, where the quadrats act as the categories of the χ^2 test, and the number of points in each quadrat is equivalent to the number of observations per category. In order to overcome the spatial dependence inherent to spatial point patterns, a sample of 50% of the quadrats is used to calculate the test statistic. Figure 4 illustrates how an area containing a point pattern can be divided into quadrats.

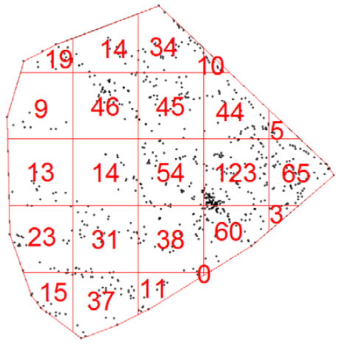


Figure 4. A point pattern domain divided into a grid of 5×5 quadrats. The numbers in each quadrat shows the number of observed points in that quadrat. The figure was first presented in (Kraamwinkel et al., 2018).

The result of a homogeneity test is given as a p-value. Here, a point pattern is considered homogeneous if its p-value is greater than 0.05, and inhomogeneous if its p-value is less than 0.05.

3.2 Homogeneity-Based Clustering

The unsupervised clustering method of (Thiede et al., 2023b) clusters roads according to their midpoint-based homogeneity. The method requires an initial network distance-based clustering of the roads, provided by a method such as network k-means, and proceeds to merge adjacent homogeneous clusters based on their homogeneity.

3.2.1 Initial Clustering (Thiede et al., 2023b) used network k-means to determine the initial clustering. The number of clusters was determined as $k = \frac{\text{number of points per quadrat}}{(m \times m) \times 5}$. Here, $m \times m$ are the dimensions of the grid used for the homogeneity test (Section 3.1), and 5 is the minimum number of expected points per quadrat for the homogeneity test to be valid, under the assumptions of the χ^2 test.

For the dataset herein, however, network k-means was impractical. Firstly, there were over 80 000 points in the dataset, and therefore calculating the network distance between these points and k was incredibly computational for any reasonable value of k. Secondly, network k-means clustering on this dataset, calculating k as recommended in (Thiede et al., 2023b), resulted in a large number of uninformative clusters.

We therefore developed a novel method of initial clustering, called isochrone clustering, which incorporates the network distance and remains computationally simple. First, the points are clustered using Euclidean k-means clustering, where k is determined as described above. The centroids of these clusters are then calculated, and isochrones are determined around the centroids, to replace the initial Euclidean k-means clusters. The isochrones are calculated such that they create a partition of the road network domain. The midpoints falling within an isochrone are assigned to the same cluster.

3.2.2 Cluster Merging Once the initial clustering is obtained, the clusters are merged iteratively based on their homogeneity. This method is developed and discussed in full in

(Thiede et al., 2023b). For the sake of completeness, we provide a summary of the steps below.

1. Let $C = \{c_1, c_2, ..., c_k\}$ be the initial clustering.
2. Obtain the set of convex hulls $H = \{h_1, h_2, ..., h_k\}$ of the clusters, such that h_i is the convex hull of cluster c_i.
3. Calculate $P = \{p_1, p_2, ..., p_k\}$, the set of homogeneity p-values. The p-value p_i is calculated for each c_i using h_i as the cluster domain, $i = 1, 2, ..., k$.
4. Obtain $E = \{(c_i, c_j)\}$, the set of all pairs of adjacent homogeneous clusters. Two clusters c_i and c_j are considered adjacent if their convex hulls overlap, i.e. if $h_i \cap h_j \neq \emptyset$, and c_i is homogeneous if $p_i > \alpha$, where α is the significance level.
5. Calculate the homogeneity p-value, p_{ij}, for each pair of adjacent homogeneous clusters, (c_i, c_j) using the convex hull around both clusters, h_{ij}, as the cluster domain.
6. Merge the pair of adjacent homogeneous clusters (c_i, c_j) with the highest associated p-value p_{ij}, given $p_{ij} > \alpha$.
7. Repeat steps 3-6 until no more adjacent clusters can be merged.

3.3 The Markov Chain-Based Geographical Accessibility Model

A summary of the steps to construct the Markov chain-based accessibility model is presented below, referring the reader to (Thiede et al., 2023c) for further details. The input to the accessibility model consists of a road network represented as a spatial linear network L, and some partition of the road network's domain, say $V = \{V_1, V_2, ..., V_n\}$. Accessibility is then quantified between the V_i's.

1. For $i = 1, 2, ..., n$, intersect each V_i with the road network L to obtain l_i, the road network within V_i.
2. Obtain representative nodes within each road sub-network l_i via Louvain clustering (Blondel et al., 2008), as explained in (Thiede et al., 2023c). These representative nodes are called the Louvain nodes.
3. Calculate the average inverse distance between all pairs of Louvain nodes within each area V_i, based on the road network l_i.
4. For each pair of adjacent areas V_i and V_j, calculate the average inverse distance between each Louvain node in V_i and each Louvain node in V_j, based on the combined road network $l_{ij} = l_i \cup l_j$. Two areas V_i and V_j are considered adjacent if they share a boundary.
5. Create a matrix containing all the average inverse distances between adjacent V_is.
6. Row-standardise the inverse distance matrix to create the 1-step transition probability matrix (TPM).

Raising the 1-step TPM to the power $n > 0$ gives the n-step TPM, which quantifies the relative accessibility between areas V_i and V_j in n steps. Letting n tend to infinity results in the prominence index, which quantifies the accessibility of the V_i's for an infinite journey, regardless of the V_i in which the journey originated.

3.4 Proposed Approach

The previous subsections provided the necessary theory to explain each component of the proposed approach to quantify accessibility between homogeneous regions. Figure 5 now outlines this proposed approach. The method takes a road network as its input, and outputs a 1-step TPM between the most homogeneous subdivision of the road network's domain, based on the initial clustering.

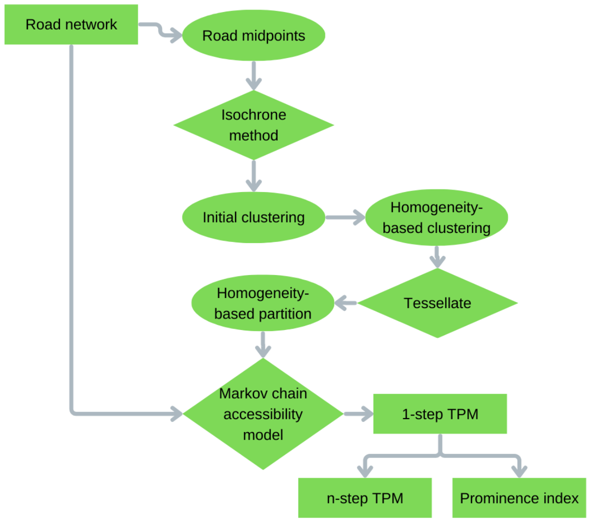


Figure 5. Flowchart illustrating the process for obtaining the 1-step TPM, n-step TPM and prominence index.

The method begins by obtaining the midpoints of the road network. These midpoints are then clustered using the isochrone method. Once the initial clustering is obtained, it is used as the basis for the homogeneity-based clustering method outlined in Section 3.2. This clustering method results in clusters that may overlap, or have gaps between them; the convex hulls of the final clusters thus do not create a strict partition of the road network's domain. In order to create a partition, overlaps between the convex hulls are removed, and the convex hulls are extended until there are no gaps between their boundaries.

Finally, the homogeneity-based partition and the road network are fed into the Markov chain accessibility model, resulting in the 1-step TPM between the homogeneity-based partitioned areas. This also allows for the calculation of the n-step TPM and the prominence index.

4. Results and Discussion

The proposed approach is now applied to the road network in the western City of Tshwane. This section presents and discusses the results, and discusses limitations and future work.

4.1 Initial Clustering

For the initial clustering, the choice of k was determined as specified in Section 3.2.1, namely $k = \frac{\text{number of points per quadrat}}{(m \times m) \times 5}$. The number of points per quadrat is influenced by the choice of m, which is the number of quadrats. The values of m considered, along with the resulting value of k, before and after data cleaning, is given in Table 1. Data cleaning was required since most of the initial clusters contained too few points for the homogeneity test, and others overlapped to such degree that they were combined.

m	k (initial)	k (after cleaning)
4	1130	384
5	720	279
6	502	202
7	370	159

Table 1. Number of initial clusters k for the dataset as determined by the choice of m.

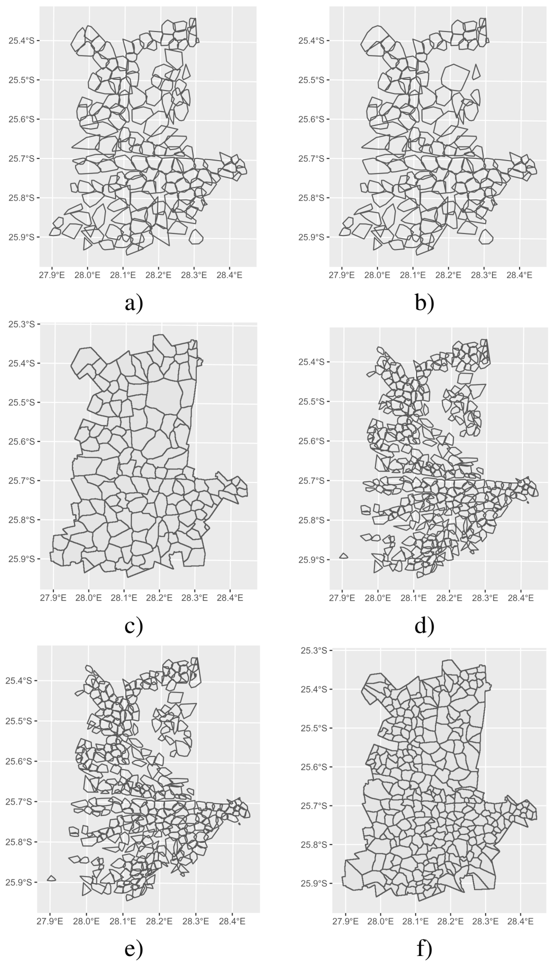


Figure 6. Results of the clustering process for $k = 159$ ((a)-(c)) and $k = 384$ ((d)-(f)). a) Original clusters based on isochrone method for $k = 159$. b) Clusters after homogeneity merging for $k = 159$. c) Results of tessellation for $k = 159$. d) Original clusters based on isochrone method for $k = 384$. e) Clusters after homogeneity merging for $k = 384$. f) Results of tessellation for $k = 384$.

4.2 Analysis

In the interest of brevity, we analysed only the finest and coarsest clusterings, namely $k = 159$ and $k = 384$. Presenting the results for all clusterings would not necessarily add value to the discussion. Figure 6 shows the clustering steps for $k = 159$ and $k = 384$. Figure 6(a) and (d), respectively, show the convex hulls of the initial clusters, obtained via the isochrone method. Figure 6(b) and (e) respectively show the convex hulls of the clusters after homogeneity-based merging. For $k = 159$, the number of clusters reduced by 5%, and for $k = 384$, by 4%.

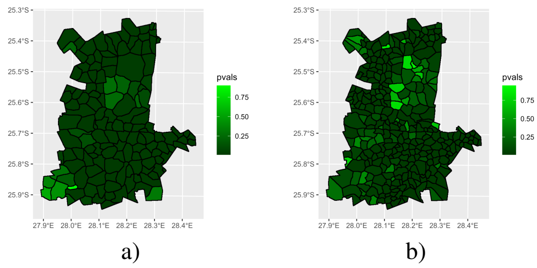


Figure 7. Results of homogeneity clustering. a) The p-values associated with the clusters for $k = 159$. A lighter colour corresponds to a more homogeneous region. b) The p-values associated with the clusters for $k = 384$.

Figure 7(a) and (b) show the p-values of the regions, for $k = 159$ and $k = 384$ respectively. Recall that a p-value greater than 0.05 indicates homogeneity, while a p-value less than 0.05 indicates inhomogeneity.

In both cases, there is a concentration of homogeneous wards in the northern central area, in the southeast, and in the northeast. The areas of homogeneity for $k = 159$ generally agree with the areas of homogeneity for $k = 384$. However, for $k = 159$, the homogeneous regions generally have lower p-values than for $k = 384$. This stands to reason, as larger areas may be expected to contain more roads and hence more diverse road networks, leading to a lower p-value. The finer partition of $k = 384$ allows for smaller areas, each containing fewer roads and thus having a higher chance to contain less diverse, more homogeneous road networks.

In both cases, more homogeneous areas coincide with less populated, more rural regions. These tend to have sparse, spread out roads. Furthermore, the most homogeneous areas in both cases are some of the smallest. This may in part be due to actual homogeneity of the road networks in these areas, but could also be caused by the low number of roads and hence points within these small regions. The fewer points there are in a region, the smaller the number of possible quadrats. The fewer quadrats there are, the less the influence of spatial dependence is accounted for, thereby potentially inflating the p-value (Kraamwinkel et al., 2018). Although we found no significant correlation between the number of points and homogeneity, the number of quadrats should still be taken into account when interpreting homogeneity p-values.

Figure 8(a) and (b) show the prominence index values of the regions for $k = 159$ and $k = 384$ respectively. Recall that a higher prominence index indicates that a region is more accessible, regardless of the origin of the journey. Here, there is very little agreement between the cases where $k = 159$ and $k = 384$, unlike for the homogeneity. For $k = 159$, the most

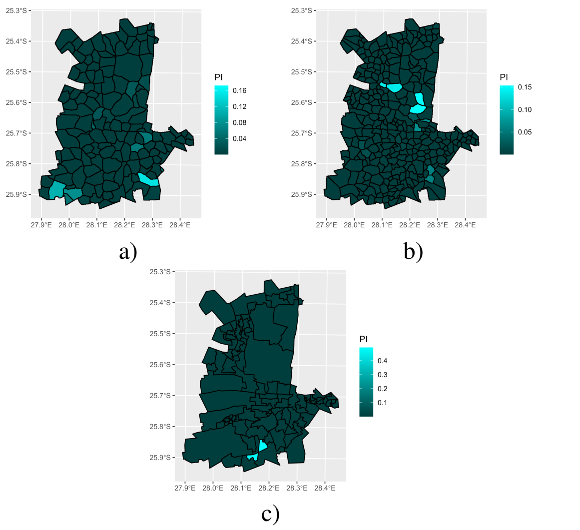


Figure 8. Results of accessibility modelling. a) The prominence index for $k = 159$. A lighter colour corresponds to a more prominent region. b) The prominence index for $k = 384$. c) The prominence index calculated on the wards.

accessible regions are one in the southwestern area, and two in the southeastern area. The southwestern region corresponds to an area called Elarduspark, a growing residential and commercial centre. Its high accessibility is thus expected. The two southeastern regions contain mainly motorways or primary roads, and few other roads. Thus, they are well-connected, and contain few nodes.

For $k = 384$, there are two sets of adjacent regions in the upper middle of the study area with a very high prominence index. The more western regions coincide with the Tshwane University of Technology, which should indeed be easily accessible to students and staff. As with the accessible southeastern regions for $k = 159$, all of these regions contain mainly motorways or primary roads, and few other roads. It is of note that the two adjacent regions in the southwest, with a prominence index of approximately 0.08, intersect the area covered by the Elarduspark region in (a). This confirms the prominence of Elarduspark in terms of accessibility.

The scale of the prominence index is similar for both regionalisations, ranging from just over 0 to just over 0.16 for $k = 159$, and 0.15 for $k = 384$.

Figure 8(c) shows the prominence index calculated on the administrative units (electoral wards). This differs from the prominence index calculated on the homogeneity-based regionalisations. Firstly, the scale of the prominence index is greater, ranging from just over 0 to nearly 0.5. Here, all the wards have a prominence index well below 0.2, except two wards in the south, with a prominence index of nearly 0.5. These wards contain part of Centurion, a residential and economic hub within the area, well-connected to the rest of the study area via residential and major roads.

We found no significant linear correlation between the prominence index and homogeneity for either $k = 159$ or $k = 384$. Choosing a homogeneous region as the origin of a journey also did not result in faster convergence to the prominence index. Figure 9 shows the top 10 most accessible regions for the case where $k = 159$, for $n = 5, 2000, 3000$. In Figure 9(a)-(c), the results are given with respect to a homogeneous region, outlined in pink. Figure 9(d)-(f) shows the results with respect to an inhomogeneous region, outlined in pink. This shows that the rate and pattern of convergence is not influenced by the homogeneity of the origin region.

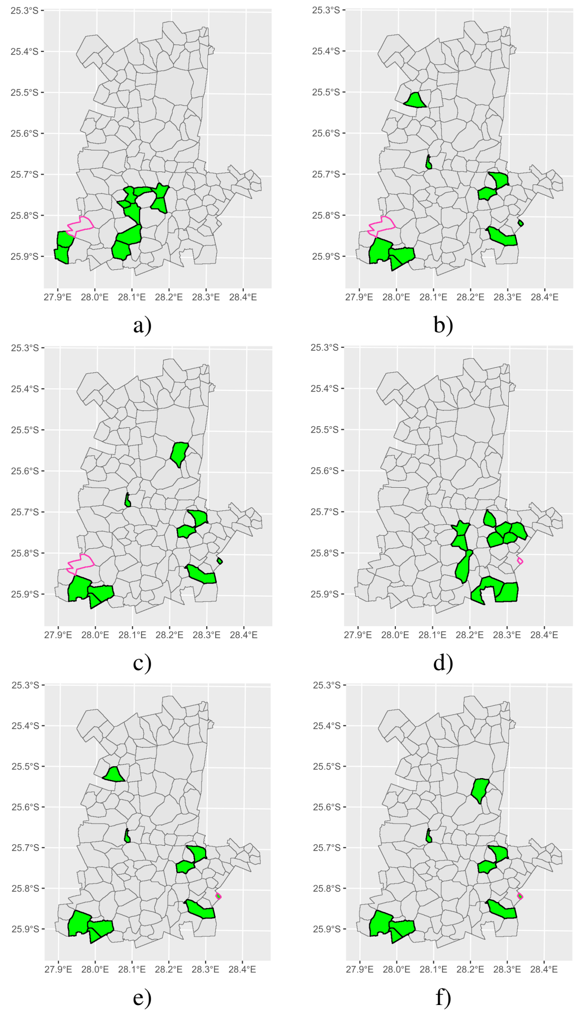


Figure 9. Top 10 most accessible regions from homogeneous and inhomogeneous regions. a) Top 10 most accessible regions starting from a homogeneous region, for $n = 5$. The origin region is outlined in pink. b) $n = 2000$. c) $n = 3000$. d) Top 10 most accessible regions starting from an inhomogeneous region, for $n = 5$. The origin region is outlined in pink. e) $n = 2000$. f) $n = 3000$.

For the case where $k = 384$, convergence to the prominence index was also not influenced by whether or not the origin region was homogeneous. However, convergence to the prominence index is much slower for this regionalisation. Figure 10 shows the convergence to the prominence index for $k = 384$. At $n = 2000$ (Figure 10(a)), the most accessible regions are still far from the prominence index.

These results thus demonstrate that the homogeneity clustering method is able to obtain meaningful regionalisations of the road network, and that these regionalisations, in conjunction with the accessibility model, can provide new insights into the state of accessibility via the road network.

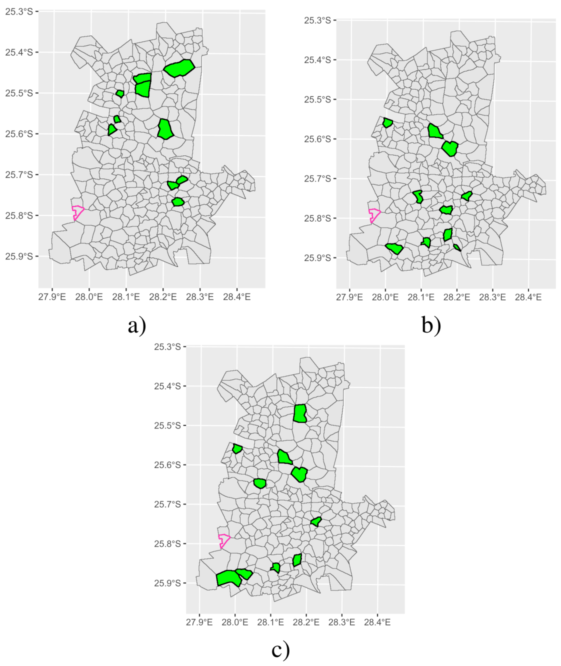


Figure 10. Top 10 most accessible regions for $k = 384$. The origin ward is outlined in pink. a) $n = 2000$. b) $n = 200000$. c) $n = 500000$. Even at $n = 500000$, the results have not yet converged to the prominence index (Figure 8(b)).

4.3 Limitations and Future Work

The method has some limitations. Firstly, the accessibility model assumes that the Markov property holds; this is a limitation acknowledged in (Thiede et al., 2023c). Secondly, the homogeneity clustering method is very sensitive to the initial clustering. Lastly, neither the clustering method nor the accessibility model considered edge effects. In reality, road networks do not terminate abruptly at the edge of any given area, nor is travel confined by the borders of municipal areas (except where these align with national borders).

Future work could address some of these concerns. Alternative methods of initial clustering could be developed that explicitly use homogeneity criteria to cluster. Edge effects should be explored and mitigated as far as possible.

The method could also be expanded in future. Homogeneity is only one aspect of road structure; other characteristics such as density, gridlikeness, connectivity and more could be considered during clustering and accessibility modelling. Work is already in progress to incorporate traffic data into the accessibility model in order to base accessibility on travel time instead of distance. Traffic data could also be incorporated into the clustering steps. Finally, it should be acknowledged that the results of the clustering and accessibility model are heavily dependent on the geographical scale and level of aggregation, represented here by the size of the initial clusters. This is a clear example of the modifiable areal unit problem. Herein, we compared two spatial scales: $k = 159$, which resulted in regions that were comparable to the electoral wards within the City of Tshwane, and $k = 384$, the finest subdivision that still allowed reasonably sound results for the homogeneity test. Future work could further explore the effects of the spatial scale on homogeneity and accessibility.

5. Conclusion

This paper considered the effects of the homogeneity of road networks on inter-regional accessibility. This was done by extending a homogeneity-based road network clustering algorithm to subdivide a road network into regions. The homogeneity and accessibility of the regions was calculated. The method was applied to the road network within the City of Tshwane municipality, South Africa. The accessibility based on our method of creating regions, at two spatial scales, was compared to accessibility based on electoral wards. The results showed that, although accessibility is not directly correlated with homogeneity, it is highly dependent on the way in which regions are constructed. Our novel construction of regions provided information on accessibility within the municipality that complements the use of administrative units, and offers a broad range of future research opportunities into the effects of road structure on accessibility.

References

Ajami, A., Kuffer, M., Persello, C., Pfeffer, K., 2019. Identifying a slums’ degree of deprivation from VHR images using convolutional neural networks. *Remote Sensing*, 11(11), 1282.

Bavaud, F., 1998. Models for spatial weights: A systematic look. *Geographical Analysis*, 30(2), 153–171.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), P10008.

Clark, A., Scott, D., 2014. Understanding the impact of the modifiable areal unit problem on the relationship between active travel and the built environment. *Urban Studies*, 51(2), 284–299.

Giddy, J. K., 2019. The influence of e-hailing apps on urban mobilities in South Africa. *African Geographical Review*, 38(3), 227–239.

Javanmard, R., Lee, J., Kim, J., Liu, L., Diab, E., 2023. The impacts of the modifiable areal unit problem (MAUP) on social equity analysis of public transit reliability. *Journal of Transport Geography*, 106, 103500.

Kohli, D., Sliuzas, R., Kerle, N., Stein, A., 2012. An ontology of slums for image-based classification. *Computers, Environment and Urban Systems*, 36(2), 154–163.

Kraamwinkel, C., Fabris-Rotelli, I., Stein, A., 2018. Bootstrap testing for first-order stationarity on irregular windows in spatial point patterns. *Spatial Statistics*, 28, 194–215.

Netrdová, P., Nosek, V., 2020. Spatial dimension of unemployment: Space-time analysis using real-time accessibility in Czechia. *ISPRS International Journal of Geo-Information*, 9(6), 401.

Nobrega, R., O’hara, C., Quintanilha, J., 2006. Detecting roads in informal settlements surrounding Sao Paulo city by using object-based classification. *Proceedings of the 1st International Conference on Object-based Image Analysis (OBIA 2006), Salzburg, Austria*, 4–5.

Potthoff, R. F., Whittinghill, M., 1966. Testing for homogeneity: II. The Poisson distribution. *Biometrika*, 53(1/2), 183–190.

Quatrini, V., Tomao, A., Corona, P., Ferrari, B., Masini, E., Agrimi, M., 2019. Is new always better than old? Accessibility and usability of the urban green areas of the municipality of Rome. *Urban Forestry & Urban Greening*, 37, 126–134.

Runsten, S., Nerini, F. F., Tait, L., 2018. Energy provision in South African informal urban Settlements-A multi-criteria sustainability analysis. *Energy Strategy Reviews*, 19, 76–84.

Thiede, R., Fabris-Rotelli, I., Debba, P., Cleghorn, C. W., 2023a. Measuring homogeneity of spatial line patterns. *Under review*.

Thiede, R., Fabris-Rotelli, I., Debba, P., Cleghorn, C. W., 2023b. Unsupervised clustering of road networks based on spatial homogeneity. *To be submitted*.

Thiede, R., Fabris-Rotelli, I., Debba, P., Li, M., Stein, A., 2020. Uncertainty quantification for the extraction of informal roads from remote sensing images of South Africa. *South African Geographical Journal*, 102(2), 249–272.

Thiede, R. N., Fabris-Rotelli, I. N., Debba, P., Cleghorn, C. W., 2023c. A Markov chain model for geographical accessibility. *Spatial Statistics*, 55, 100748.

Viegas, J. M., Martinez, L. M., Silva, E. A., 2009. Effects of the modifiable areal unit problem on the delineation of traffic analysis zones. *Environment and Planning B: Planning and Design*, 36(4), 625–643.

Wigley, A., Tejedor-Garavito, N., Alegana, V., Carioli, A., Ruktanonchai, C. W., Pezzulo, C., Matthews, Z., Tatem, A., Nilsen, K., 2020. Measuring the availability and geographical accessibility of maternal health services across Sub-Saharan Africa. *BMC Medicine*, 18(1), 1–10.

Wong, D. W., 2004. The modifiable areal unit problem (MAUP). *WorldMinds: Geographical Perspectives on 100 Problems: Commemorating the 100th Anniversary of the Association of American Geographers 1904–2004*, Springer, 571–575.

https://doi.org/10.5194/isprs-annals-X-4-2024-357-2024