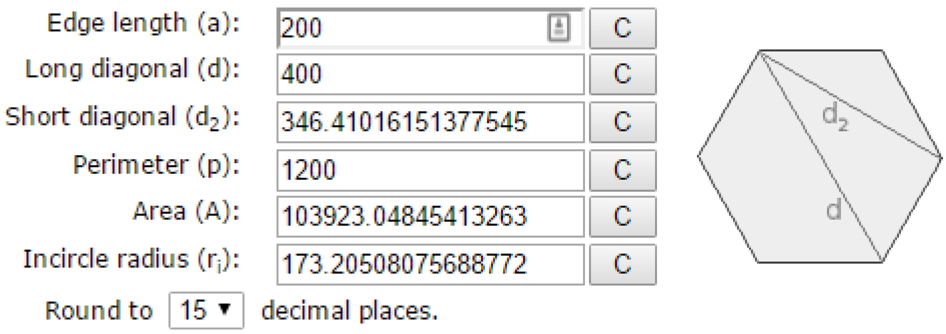
1. Grid

1.1. Hexagon Grid Creation

- Projection Hartbeeshoek Lo29
- A hexagon grid was created using the MMQGIS plugin in QGIS 2.16.2
- This Hexagon Grid covers the shape of Gauteng with a 10km buffer added using the buffer tool in QGIS

1.1.1. 400m Hexagon Dimensions

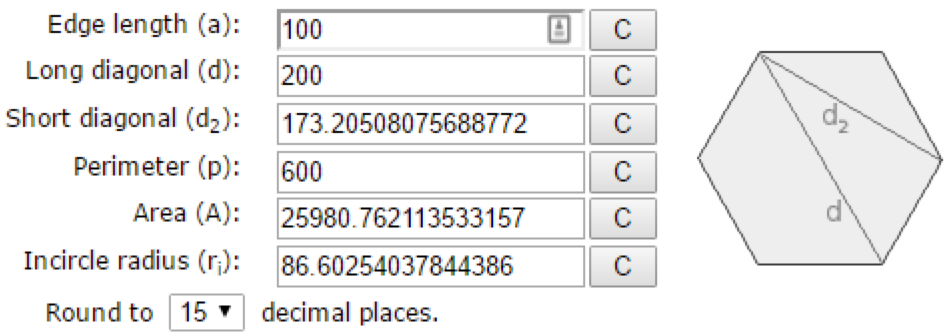
- Y spacing 346.41016151377545m
- X spacing automatic (300.0m)



[1]

1.1.2. 200m Hexagon Dimensions

- Y spacing 173.20508075688772m
- X spacing automatic (150.0m)



1.1.3. Centroids of Hexagons

Centroids of the grid hexagons were calculated in QGIS using the “Polygon Centroids” tool

[1] https://rechneronline.de/pi/hexagon.php

1.1.4. Joburg grid

This grid was created by intersecting the Gauteng grid created above with the outline shape of the CoJ boundary. Any hexagon that intersects the CoJ shape was extracted for the analysis.

2. Network Analysis

- Street data downloaded from Openstreetmap for Lesotho and South Africa on 6 July 2016 from the site: http://download.geofabrik.de/africa/south-africa-and-lesotho.html
- Projection Hartbeeshoek Lo29
- 10km buffer shapefile created for Gauteng Province Boundary
- Roads snipped to Gauteng 10km Buffer shapefile

2.1. Walkability Analysis

- This is used to create a walkability score for each 400m by 400m Hexagon
- For Walkability analysis, certain roads were excluded:
- Using OpenStreetmap data from above, road types deleted:
    - Bridleway
    - Construction
    - Motorway
    - Motorway_access
    - Path
    - Proposed
    - Raceway
    - Service
    - Services
- Sampling Grids Used: See section 1 titled “Grid”
- All Sampling done for Hexagons intersecting with the CoJ border

2.1.1. 1km Service Area: 400m Hexagons, 40m Trim

- This gives an indicator of local walkability
- Facilities for Network Analysis: Centroids of 400m Hexagon Grids
- Locate Facilities: Locate road network within 200m of point. Points more than 200m from a road do not get scored.
- Service area distance: 1000m
- Trim Service areas: 40m (a stretch of road would have a 100m buffer around it)

2.1.2. Method

The square km area of each service area immediately above was calculated. This was divided by 3.14 (the area in km^2 of a 1km diameter circle) to give the walkable area. This can be multiplied by 100 for a percentage or by 10 for a score out of 10.

2.2. Accessibility Analysis

- No roads excluded
- This is used to assess the amenities and facilities accessible from each 400m by 400m hexagon
- It uses two broad methods:
    - Closest facility (for points like schools etc.)
    - Number of facilities in x distance (for multiple points like GTI data, and for area of land use, like parks)

2.2.1. 2km Service Area: 400m Hexagons, 200m Trim

- This was used to count features and amenities within 2km walking distance of the centre of each hexagon.
- Facilities for Network Analysis: Centroids of 400m Hexagon Grids
- Locate Facilities: Locate road network within 200m of point. Points more than 200m from a road do not get scored.
- Service area distance: 2000m
- Trim Service areas: 200m (a stretch of road would have a 200m buffer around it) This was done so that large commercial, industrial or other facilities may have their centre 200m from a road)
- The square metre area of each service area gives the broad walking area for each hexagon point.

2.2.2. Public Schools Access

- The location of schools was obtained from the Gauteng Department of Education in shapefile format
- These schools were split into high schools and primary schools. High schools included both ‘combined’ and ‘secondary’ schools while primary schools included both ‘combined’ and primary’ schools.
- Access to high schools and primary schools were calculated separately
- Two methods were used to calculate school accessibility

Number of schools in 2km walking distance

Here, the 2000m service areas for each hexagon were used. Using join (between the 2km service area and schools) the number of schools within each 2km walking area was counted. This value was then joined to the original hexagon centroids and then hexagons to give a value for number of schools within 2km.

This was done jointly for all schools (high schools, primary schools, combined schools, and special schools)

Distance to closest school

This was done separately for high schools and primary schools. High schools include secondary and combined schools, primary schools include primary and combined schools.

This was done using OpenStreetMap roads data (mentioned earlier) with no road types excluded. The OD cost matrix was used to calculate the distance, in metres, from each hexagon centroid to the closest school.

2.2.3. Access to other facilities and amenities

To count accessibility to other facilities, the same procedure as for schools above was used. Listed below are the fields calculated, including the data that was used to calculate.

Commercial in 2km

- GeoTeraImage (GTI) 2012 Building Based Land Points used
- Formal Commercial Businesses were isolated from the data, and counted, giving a value of 1 for each point
- The value is a count of formal commercial businesses within 2km walking distance (by road) of the centre of each hexagon.

Industrial in 2km

- GeoTeraImage (GTI) 2012 Building Based Land Points used
- Formal Industrial Businesses were isolated from the data, and counted, giving a value of 1 for each point
- The value is a count of formal industrial businesses within 2km walking distance (by road) of the centre of each hexagon.

Industrial and Commercial in 2km

- This is the sum of commercial in 2km and industrial in 2km

Closest Hospital

- A list of hospitals was obtained from the department of Health, Gauteng
- The same method of measuring distance by road was used as for schools, using an OD Cost Matrix

Clinics

- Geo-referenced clinic data obtained from the Gauteng department of health
- Closest clinic and clinics in 2km were calculated as was done for schools above

Train Stations

- Train stations obtained from openstreetmap, including Gautrain and Prasa Stations. These stations are for all of Gauteng. Data was obtained from https://mapzen.com/data/metro-extracts/
- As stations can often be accessed from two streets, but snap to one street in network analyst, initial results were skewed (Showing access from only one side in Jabulani for example). To overcome this, 150m buffers were drawn around each station using the draw buffer tool in

arcmap. These buffers were intersected with the Gauteng street network creating points around stations that overlap streets. These points were used as facilities in the OD cost matrix.

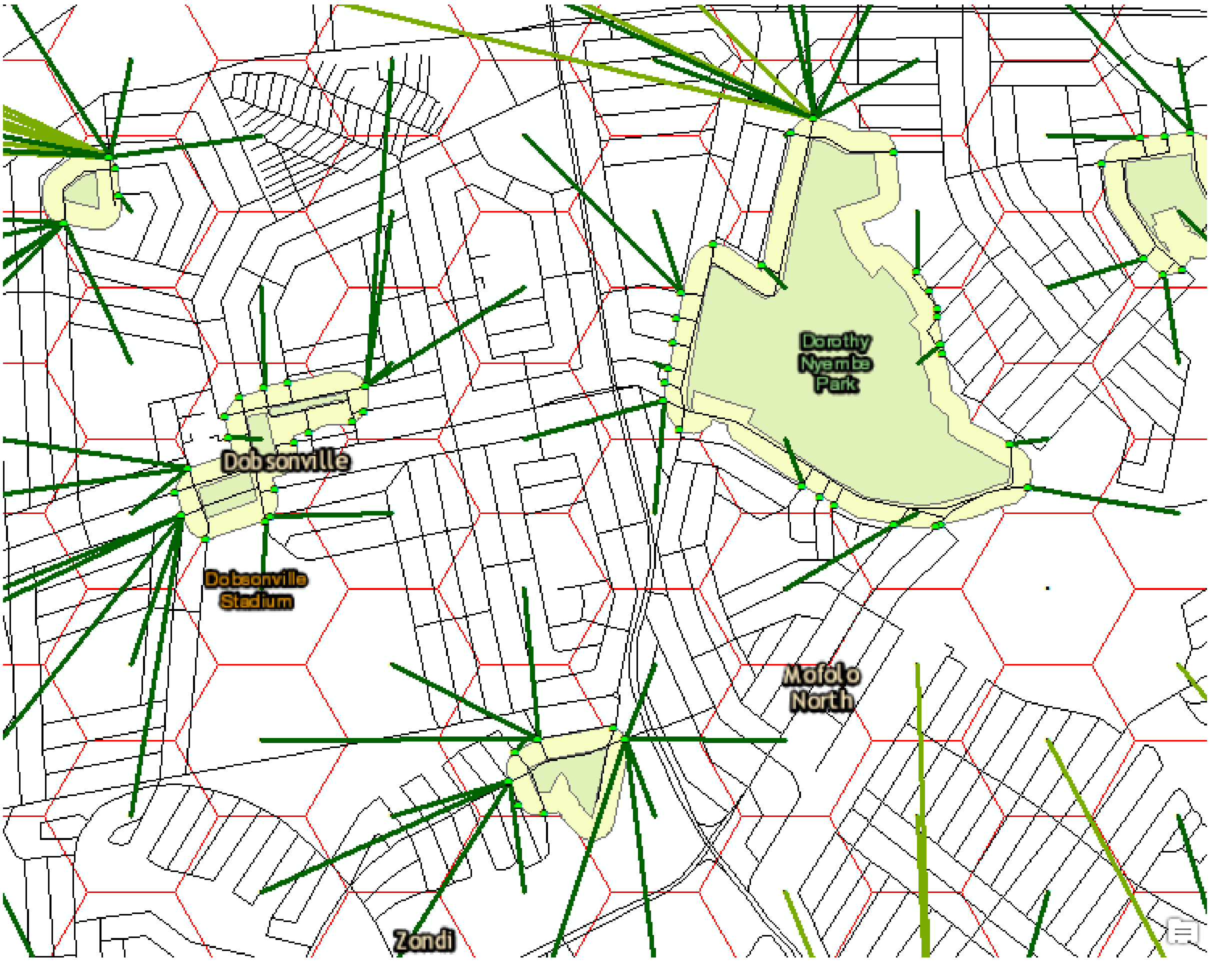
BRT Stations

- Data for BRT Stations was obtained from the Johannesburg transport Department, including Phases 1a and 1b.
- Phase 1b used BRT station data from the Johannesburg SDF 2040 (obtained from the Louis Botha, Empire Perth and Turffontein Strategic Area Frameworks from 2015)
- These points were all located within 200m of streets, and hence no buffers were drawn (as with trains above)
- As with Schools, an OD cost matrix was used to calculate closest BRT station.

Parks

- Data was obtained from Johannesburg City Parks and zoo in May 2017
- The data includes all land owned or managed by city parks and zoo
- Parks within 2km
    - This shows the square m of all parks property within 2km walking distance from the centre of each hexagon cell
    - A field was added to the Parks shapefile dataset, called 100m2 and all values were calculated to 100. The whole parks layer polygon was converted to a raster using “polygon to raster” in Arcmap, and using the 100m2 field (so each cell has a value of 100) and each cell was sized at 10 (giving 10 by 10 cells – 100m2).
    - This raster was converted to a point layer (“raster to point” in ArcMap).
    - The resulting point layer was then spatially joined to the 2km service area shapefile, using sum as the summary statistic. This gives the square metres of parks within each 2km service area, rounded to 100m2.
    - The value for square metres of park was then joined back to each hexagon, giving parks in 2km.
- The following categories were created and the above analysis done for each:
    - All publically accessible open space
        - "OSTYPEF"= 'PARK - FLAGSHIP' OR "OSTYPEF" = 'PARK - DEVELOPED' OR "OSTYPEF" = 'BIRD SANCTUARY' OR "OSTYPEF" = 'BOTANICAL GARDEN' OR "OSTYPEF" = 'FLAGSHIP BIRD SANCTUARY' OR "OSTYPEF" = 'FLAGSHIP NATURE RESERVE' OR "OSTYPEF" = 'NATURE RESERVE' OR "OSTYPEF" = 'ZOO'
    - Developed and Flagship parks
    - Undeveloped parks
    - Nature reserves, bird sanctuaries, zoo and botanical gardens
    - Developed and Flagship parks, Nature reserves, bird sanctuaries, zoo and botanical gardens
    - Koppies, Ridges
- Closest Park
    - For this, the following category of parks were used:
        - Developed and Flagship parks, Nature reserves, bird sanctuaries, zoo and botanical gardens

- Undeveloped parks
  - A 50m buffer was then drawn around each of these shapefile categories, and the 50m buffers were intersected with the road network to give points where parks have close access to roads.
  - Using these points an OD Cost Matrix was used as with schools described earlier in the document. Example shown below.



CoJ Capital Investment in 2km walking distance

- For this analysis, shapefiles were obtained from the Johannesburg Strategic Infrastructure Platform (JSIP) final results for the MTREF capital budget 2017/18. The data is in point form, so must be aggregated to show impact over areas. As such, 2km service areas (described earlier in the document) were used to analyse them.
- A new field was added to the data called 3_yr_bud (2 year budget). This was calculated as a sum of the 2017/18, 2018/19 and 2019/20 allocated budget, to give a single figure for all three years. Budget items with no specific spatial location are located well outside of the city (deliberately), so would not be included in the analysis (as they are located further than 2km outside of the city, to avoid them having a specific spatial location).
- No budget points were excluded; all capital budget is included in the analysis.
- Only items included in the 2017/18 budget book were included

- This 3 year budget data was then used to calculate allocated budget spend over the next three years, within 2km walk of each hexagon centre.

Private Schools

- Here the same data as was used for public schools was used, but with only 'independent' schools included. This includes both subsidised and non-subsidised independent schools.
- Closest private primary and high school and private schools in 2km walking distance were calculated as was done for public schools above

Private Hospitals and Clinics

- Data for this layer could not be obtained from a government source, so it was obtained from the Hospital Association of South Africa website[2]. From the website, the following service providers' hospitals were mapped, selecting only hospitals within 100m of Marshalltown, Johannesburg:
    - Life Healthcare
    - Mediclinic
    - National Hospital Network
    - Netcare
    - Mooimed Clinic
    - Joint Medical Holdings
    - Akeso Clinics
- Other hospital groups added include:
    - Lennmed Health[3]
    - Clinix[4]
- All addresses and names of hospitals above were inserted into a table and geocoded using the Google Maps API through Google sheets, using the plugin "Geocode by Awesome Table"[5]
- Closest private hospitals and private hospitals in 2km walking distance were calculated as was done for public schools above.

Taxi Routes and Ranks

- Data for taxi ranks and for taxi routes was sourced from the Johannesburg Transport Department, with the data marked 2013
- For taxi routes, the lines of the routes were converted to points using the Arcmap plugun Create Points From Lines [6]. This was done with points every 10m.
- These points were used to calculate the closest taxi route to each hexagon center in the same way that closest public school was calculated above.

---

[2] https://www.hasa.co.za/
[3] http://www.lenmedhealth.com/en/
[4] https://www.clinix.co.za/
[5]https://chrome.google.com/webstore/detail/geocode-by-awesome-table/cnhboknahecjdnlkjnlodacdjelippfg?utm_source=permalink
[6] https://www.arcgis.com/home/item.html?id=a2a41c8345e24ab6a9dd2ae215710b39

- For closest taxi rank, only 'formal taxi ranks' was included. Using these points, an OD Cost Matrix was used as with schools, described earlier in the document, to define the closest police station to each hexagon centre

Police Stations

- Data for police station locations for the whole of South Africa was downloaded from the SAPS web site at: https://www.saps.gov.za/services/boundary.php
- Using these points, an OD Cost Matrix was used as with schools, described earlier in the document, to define the closest police station to each hexagon centre

Tertiary Education

- All data for tertiary education was obtained from the Department of Higher Education and Training website.
- Three categories of tertiary education institution were obtained and mapped:
  - Public Universities: The list of public universities in Gauteng was obtained from the website http://www.dhet.gov.za. Addresses and names of each campus were obtained from the individual websites of each institution.
  - Registered Institutions: This list is from the document "List of Accredited Higher Education Institutions" dated 6 July 2017 from the website http://www.dhet.gov.za. Registered institutions include all Gauteng based institutions in the list "registered Institutions"
  - Provisionally Registered Institutions: This list is from the document "List of Accredited Higher Education Institutions" dated 6 July 2017 from the website http://www.dhet.gov.za. Registered institutions include all Gauteng based institutions in the list "Provisionally Registered Institutions"
- All addresses and names of institutions above were inserted into a table and geocoded using the Google Maps API through Google sheets, using the plugin "Geocode by Awesome Table"[7]
- Distances to all three categories were separately calculated. Using these points, an OD Cost Matrix was used as with schools, described earlier in the document, to define the closest tertiary education institution in each category to each hexagon centre.

2.3. Population density

For Population, the Sub Place Layer was used from the Census 2011.

2.3.1. Population her 400m Hexagon Cell

- Density was calculated for each SP polygon by a) calculating the area in km2 of each then b) dividing population by area. This gave a value in people per square kilometre.
- The entire shapefile was converted to a raster of 5m by 5m cells, using the population density variable.

[7] https://chrome.google.com/webstore/detail/geocode-by-awesome-table/cnhboknahecjdnlkjnlodacdjelippfg?utm_source=permalink

- Density in each cell was calculated using the "zonal statistics as table" tool in arcmap. Calculating the mean density for each cell.
- Calculated values were joined to the 400m hexagon shapefile, and the population of each cell was calculated by multiplying the density of each cell (in people per square km) by the area of each (in square km).

2.3.2. Population within 2km walking distance

- Instead of calculating the people within a 400m hexagon, the number of people who can walk to the centre of that hexagon in 2km was calculated here.
- Data for population per hexagon from 2.3.1 above was joined to the point layer of centroids of each 400m hexagon.
- This layer was spatially joined to the 2km service area layer using the sum function to add up the population points that fall within each 2km service area. The summed population within 2km walking distance was then joined to the hexagon layer to give the cell a value of population in 2km.
- A different method was used to 2.3.1 (zonal statistics as table) as that method does not work for overlapping counting-polygons.

3. Index Calculations

3.1. Normalising values

Calculations were conducted on all fields to convert them to a scale from 0 to 1, with 1 being considered the best and 0 the worst.

As in some fields a high value indicates good access (such as schools in 2km) and in others a low value indicates good access (such as closest school in m) two methods were used to normalise the values.

Note: In all cases, items that score 0 are either not located on the road network in the Network analysis above, or do not have any features (clinics etc.) within the 2km service area. For this reason, in all cases, 0 is the worst score, indicating either no access to roads, or no features within 2km.

3.1.1. Normalising values where high numbers equate to good access

Here, in each field, the highest value (most schools, clinics, BRT etc. accessible in 2km) was ascertained. This was done by creating a table in excel, and sorting from highest to lowest. Each value was divided by the highest value, to give a score from 0 to 1 with 1 representing the cell with the best access to whichever facility is being measured. These scores are only relative to other scores within the field, and not across fields.

3.1.2. Normalising values where low numbers equate to good access

Here, scores had to be altered, so that the smallest values give the highest score.

Initially, the following method was used, however this was then replaced.

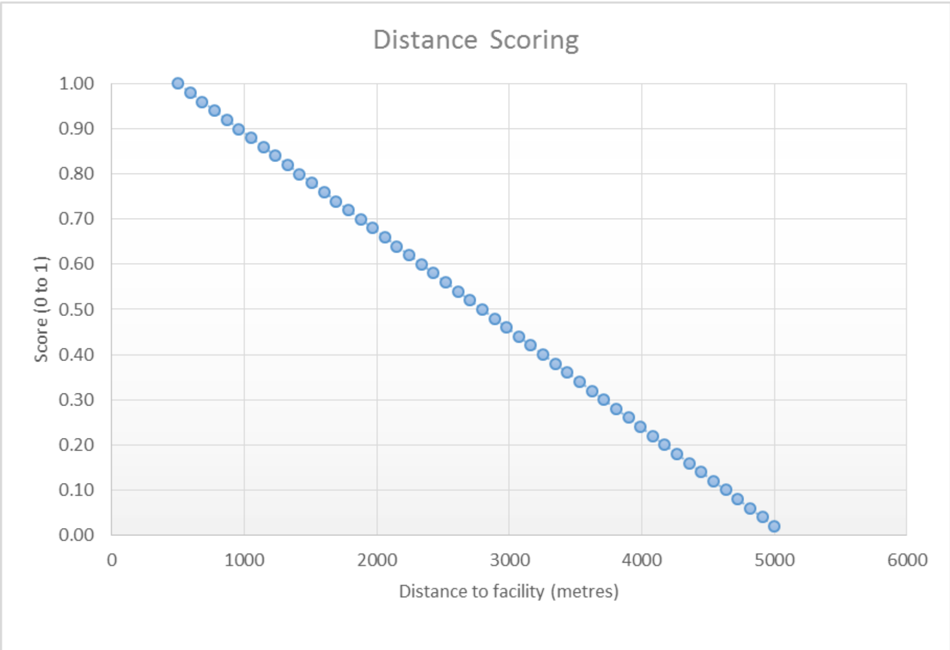
Annexure 2: Urban Potential Modelling Method

Method 1 (not used)

a) First, is the same as the section 3.1.1 above. Values were sorted from highest to lowest (to obtain the highest value) and all values were divided by the highest value. This gives a scale of 1 to 0, but with anything above 0 being the best access, and 1 being the worst.
b) Second, the value immediately above was subtracted from 1. This means that a cell scoring 0.1 in a) would give a final score of 0.9.

Method 2 (preferred method, and used in models)

a) Method 1 resulted in skewed results, as all values were divided by the smallest, which was a single digit distances in metres. This did give a relative range, however places that were for example 80m away from a school, rather than 4m, scored significantly lower, even though 80m is a very good walking distance, and should score very high. For this reason, the following distance chart was used to define distance from facilities, on a scale of 0 to 1 (with 0 being bad, and 1 being good). Facilities more than 5000m away were considered not walkable, and scored 0.



Note, as mentioned earlier, 0m in the closest categories (closest school, clinic, hospital etc.) does not mean 0m but rather a null value, as the centroid of the hexagon cell in question is not located on the road network (more than 200m from a road segment, as discussed in section 2.2.1)

3.2. Calculating Indexes

Using the Normalised Index scores from above, final indexes were created.

First, individual indexes were created for different categories, indicated and explained under headings below (Education, Health, Jobs, Parks, Public Transport, and Population). Then joint indexes were created for different categories (Residential Densification, Economic Growth, Industrial growth).

3.2.1. Individual Indexes

Walkability

Here, the 1km walking distance score was used. The 2km walking distance score is built into most of the other indexes (in that it was used to calculate them) so there was no need to include it again here. The 1km walkability score is included to highlight the inherent potential of good, walkable street networks in the city, and cells that correlate to them. It shows the potential of existing street networks, irrespective of whether they are well serviced with various amenities.

Education (schools)

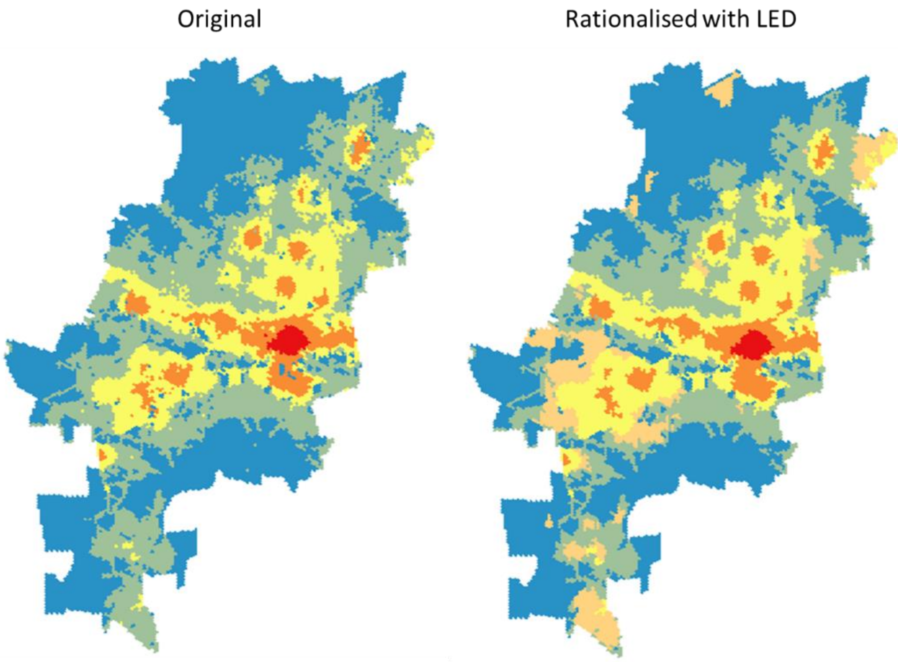
3.2.2. Economic Nodes Index

The economic nodes index was calculated, using normalised values for the: public transit index, commercial buildings in 2km index and commercial buildings in hexagon indexes. A weighted average was used weighting the three at 15%, 70% and 15% respectively.

N_Public Transport Index	15%
SC_Comc_2km	70%
SC_Com_HX	15%
	100%

Once the index was completed, it was joined back to the arcmap hexagon model. In arcmap, the data was split into 5 categories using "Jenks natural breaks" and this created the 5 levels for the nodal review.

Once the index was finalised, the data was "rationalised" manually. Here, whenever there was a single cell of one node level, surrounded by cells of another level, it was changed to reflect the surrounding cells. LED nodes were also added to the final model, where they scored as LED and scored as level 4 or lower. This means that level three remained level 3, and were not replaced by LED. The calculation for LED is shown under the following heading. The example and outcome are shown below:



3.2.3. LED Node Index

LED nodes are those nodes that have a high population density and low access to jobs and economic they are areas like Diepsloot that need economic development, but that do not score high on the economic nodes index because of a lack of current economic activity. Hexagons scored as LED if they scored in the 80th percentile or higher of population in 2km, AND were below the 80th percentile in terms of industrial and commercial buildings in 2km.

3.2.4. Residential Densification Index

Normalised individual index values were used. A weighted index was created, with the following weightings to give the residential densification (normalised) index, from 0 to 1.

N_Walkability	12.0%
N_Parks Index	12.0%
N_Education index	12.0%
N_Health Index	12.0%
N_Public Transport Index	25.0%
N_Jobs Index	25.0%
SC_Cpx_2k	2.0%
	100.0%

4. Redefining Nodes / Creating final layer

4.1. Mapping Existing Nodes

An in depth exercise was undertaken to map all existing nodes in the City based on the SDF, RSDF's, Precinct Plans, Urban Design Frameworks and Strategic Area Frameworks. Neighbourhood nodes were not included, die to the inconsistency as to how they are recorded. Rather, a condition for any existing neighbourhood node will be added to the final nodal review document.

4.2. Redefining nodes

Once the economic nodes index was finalised and rationalised (as above) the existing nodes in the City (based on Previous RSDF's and SDF's) was analysed to re- classify and change their shape if needed.

a) Inner City/ Metropolitan Core: As the index indicated a central area of the CBD as node level 1, and the existing CBD shape fully covered the areas under "node level 1" of the index, the existing CBD shape was retained as the only part of the city that classifies in the highest node level, "Primary Metropolitan Centre"
b) Level 2, Principle Metropolitan Sub-Centres: Here Level 2 of the economic nodes index was compared to existing nodes. Those existing nodes whose centres fall within level 2 economic nodes as per the index, or that connect two areas defined as level 2 (Jabulani was the only Example), were classified to level 2, Principle Metropolitan sub-centres. The rest of the area from the economic nodes index for level 2, were used either to expand nodes, or define new ones.
c) For Level three nodes, general urban zone, the entire area that the model defined was used as level 3, general urban, as well as existing nodes (that didn't qualify above as level 1 or 2) that either intersect with level 3 or level 4 as per the index. This means that the only existing nodes that wouldn't qualify as level 3, are those that are totally located in level 5 within the commercial node index.
d) LED Nodes: All hexagons that qualified as LED, and that were ranked as below 3 (so 4 or 5) in the commercial index, were registered as LED areas. Any existing node that overlaps this LED area, was registered as level 3.
e) Existing nodes that do not overlap with any hexagons from the model that are level 1 to 4 (i.e. the nodes are completely ranked 5, and are surrounded only by level 5) were ranked level 4.

4.3. Joining property layer to nodal definitions

Once the final layers were created, containing existing nodes and their redefinitions (based on the analysis) and the rest of the final categorisation using the hexagon model, individual properties in the city were joined, so as to give each property an individual node level/zone.

4.3.1. Property Layer

- Property layer obtained from LIS (land information system) on 11 December 2017
- Layer was cleaned as follows:
    - ONLY SG Approved and Registered properties were included
    - All properties were then joined before doing any further cleaning
- Property was joined to the final nodal review layer based on the development zone in which it predominantly falls.