DOMESTIC BUILDING DENSITY ESTIMATES FOR NETWORK PLANNING

J Booysen, M Dekenah, L Sethlogo,
Enerweb - Demand Intelligence Group, Eskom Holdings SOC Ltd.

ABSTRACT

Domestic building density estimates per living standards measure (LSM) in South Africa was needed to produce energy density estimates for network planning by means of the Eskom Geo Load Forecasting system (GLF). Building density estimates for domestic buildings was calculated by linking complex spatial data, using spatial filtering and algorithms to find domestic building density estimates per LSM group in South Africa that is based on reality. Building density estimates can then be combined with domestic load research data (DLR) to calculate domestic energy density. The focus of this paper was on the specific task of calculating domestic building density per LSM in South Africa as no published measures currently exist at a National level.

1. INTRODUCTION

Eskom is a South African electricity public utility. It is the largest producer of electricity in Africa, among the top seven utilities in the world in terms of generation capacity and among the top nine in terms of sales.

The network planner has to anticipate how much power must be delivered as well as where and when it will be needed. A substantial amount of work has gone into the establishment of a load forecasting tool (GLF) that will assist distribution network planning forecast and consolidate load forecasts in a single data repository. This will enable distribution network planning to provide consolidated regional and national views of forecasted future loads.

The established load forecasting tool requires energy density data in conjunction with load profile data that will be used in calculating coincidence factors between loads of different classes and sub-classes over a 24 hour day for different seasons of the year.

The focus of this paper is to create net domestic building density data in order to enable calculation of domestic energy density by combining the building density data with Domestic Load Research (DLR) data.

Building density is a controversial topic and is often misunderstood, perceptions are also linked to personal background, for example: an urban planner from the Indian subcontinent would think that a 100 m^2 erf for low income groups is far too large and will be unaffordable. A planner from East or Southern Africa however will argue that this is far too small, and that it will never be accepted. The response may be “we didn’t fight for independence to reduce our standards.” An Egyptian or a Bolivian planner might think that 100m^2 is acceptable. In Delhi sites and services projects offer erven from 26 m^2 to 90 m^2 in which two story buildings can be constructed, permitting a coverage up to 75% and allowing 2 dwelling units for erven of 48 m^2, 60 m^2 and 90 m^2. In Brasillia the national legislation for low income projects established a minimum erf size of 126 m^2. In Guinea-Bissau erven are conventionally defined as 20m by 25m, covering an area of 500 m^2 and in many African countries erven are larger than 250 m^2.

Domestic or residential building density measures can be expressed as net or gross density. Net residential density refers to the density on a specific site, excluding public roads and public open space, thus only including the area allocated for residential use. Gross residential density refers to the density of a specific site including the land occupied by local facilities such as schools, local shops, open space and roads.

Some misconceptions around building density are that low density creates high quality environments and high densities create low quality environments. The assumption that high density is hazardous to quality of life is misleading as there are other factors that when combined with high density will cause negative impact in the quality of life of a settlement.

It is a misconception that a household type can be created at a certain density. In fact a wide variety of housing types can be provided at most densities. Using data is the only way to know what the general building density trends per income group currently in South Africa is.

A methodology was developed in this paper making use of spatial statistics and spatial data layers to produce net domestic building density recommendations for South Africa based on real data and satellite imagery.

2. DATA USED IN THIS STUDY

GIS (Geospatial Data)

The Eskom SPOT5 Building Count (SBC) product derived from SPOT5 satellite imagery was used to get the location (GPS coordinates) of all domestic dwelling units and all building structures in South Africa. The total buildings available in this dataset were 11,310,724. Classes for the SBC dataset included schools, complex and hostels, mines/quarries, resorts, dense informal and domestic dwellings.

A Spatial LSM dataset (GIS) for all 80,000 enumeration area (EA) polygons with LSM penetration per EA type was purchased for South Africa from AfriScope. The purpose of the LSM is to segment the South African market according to household's living standards using criteria such as access to household assets and type of settlement.

For AfriScope to produce a spatial LSM, Information on the LSMs was obtained from nationally representative household surveys. AfricaScope has over many years used LSM data from "Caxton's ROOTS" survey, the "Human Sciences Research Council's (HSRC) South African Social Attitude Survey (SASAS)" and "FinMark Trust's FinScope" survey.

To map the LSMs down to a local level, small area estimation techniques were used. Afriscope used general regression neural networks (GRNN) to impute the LSMs from areas where the survey was conducted to all other areas. Demographic estimates at a local level were used to borrow strength in the imputation of the LSMs. The advantage of using artificial neural networks is that in creates an intricate link between the demographic and LSM variables to create strong models for the imputing of the data at a local level.

The 80,000 Enumeration areas are classified with LSM percentage breakdowns and classified of being of type: farms, urban settlements, industrial areas, vacant land, hostels, institutions, recreational, small holdings, informal settlements and tribal settlements.

An Open Street map extract with public open space and other polygons for South Africa was obtained and transformed into PostgresSQL (GIS) database and then loaded into an Oracle Spatial database for use in this study.

3. DATA FILTERING

Using a spatial LSM layer containing 80000 enumeration areas (EA) in South Africa and covering the whole of South Africa, areas were filtered based on EA type (See all possible types in data section). EA types chosen for the study were urban settlements, small holdings, informal settlements and tribal settlements adding up to a total of 66165 enumeration areas after filtering.

On the Eskom SPOT5 building count (SBC) building GIS layer, only buildings of class dwelling (residential) were included. (See data section for all possible classes).

Each enumeration area was inspected and the dominant LSM identified as well as the percentage of homogeneity inside each EA. For example in an area that was LSM6 dominant and 70% homogenous, meant that the majority, 70% of households belonged to LSM6 in that EA. In an ideal world only the areas per LSM that were 100% homogenous to a LSM group would be relevant, but because in reality this is never the case the percentage homogeneity to the LSM was inspected and the appropriate cut-off percentage decided based on the amount of areas available per LSM, this was done only after the density estimates were determined for all residential enumeration areas in South Africa.

Open Street Map's free tagging system allows the map to contain unlimited data about its elements, like roads and buildings. The community agrees on certain key and value combinations for tags that are informal standards.

Even though these polygons relies on the community to update it and are not 100% complete, the accuracy was found to be of good quality, and it was further supplemented with a spatial algorithm described later in this report to exclude public open space in areas where public open spaces where not flagged in open street maps.

The following public open spaces where excluded in computing of domestic building density making use of open street maps polygons: aero ways, natural boundaries including lakes and water bodies, open grasslands, amenities like toilets, telephones, banks, pharmacies and schools, all Land use identified as other than residential, leisure areas and buildings identified as other than residential, house, terrace or sheds.

4. SPATIAL ALGORITMN

Below is the zoomed GIS view of an example enumeration area (EA) boundary (the thick black border line). The black crosses indicate building positions by means of the spot building count layer (SBC), the red areas within the EA boundary below represent "parks" from the open street maps layer and were excluded in the area building density calculation per LSM.

On observing the image above, the public open space park area indicated by the green arrow is not highlighted in red, this is an example of an area that has not been marked in open street maps but can clearly be seen to be a park of some sort. This area is excluded by means of a spatial buffering algorithm described next.

Even though building density varies widely between areas, it is rather consistent within most enumeration areas. This translates to buildings being evenly spread out between each other within an EA (small area).

For each enumeration area, the nearest neighbour distance between each building in the area was computed as a distribution of nearest neighbour distances. Buildings that had a nearest neighbour distance that was more than 2 standard deviations were considered outliers and excluded in the building density estimation. Buildings that were less than 2 meter apart were also filtered out to clean the data further.

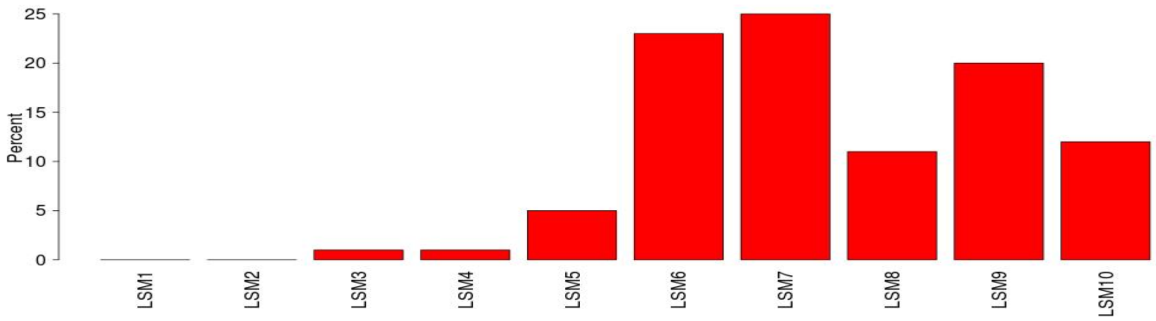
A spatial buffering algorithm using the mean nearest neighbour distance plus 1.96 standard deviations (95th percentile) risk as the buffer width were chosen. Spatial buffering creates circle buffers with a buffer width radius around each building merging the circle polygons where they touch or overlap to form the area used in determining building density.

Below is a view of the spatial buffering algorithm within the example EA in action. The red area is the area included in estimating the building density. The buffering uses the 95th percentile nearest neighbour distance as buffer width per EA. The open street maps layers are then excluded over and above the buffering algorithm for computing Net residential building density per area.

You can clearly see that the spatial algorithm automatically excluded the area that was not found in the open street map system, increasing the accuracy of the household building density estimates.

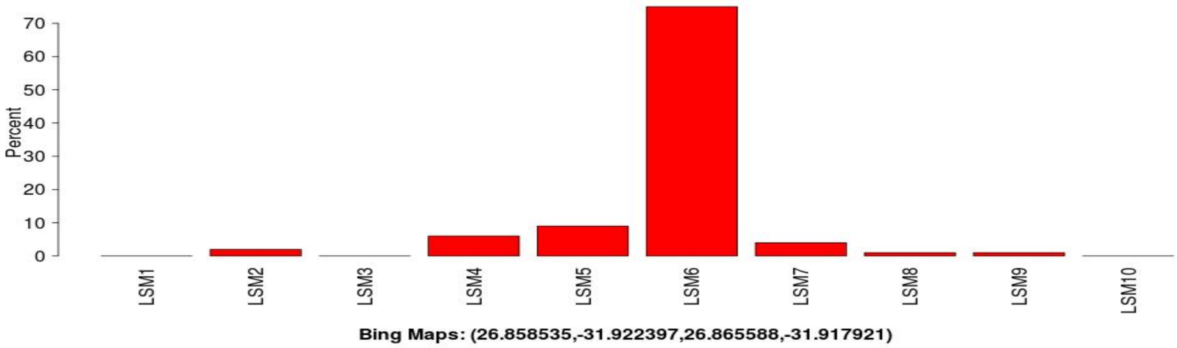
This spatial algorithm was applied to all 66165 residential enumerations areas in South Africa and the domestic (residential) building density computed as buildings per hectare making use of the buffered area size and the amount of buildings inside the areas.

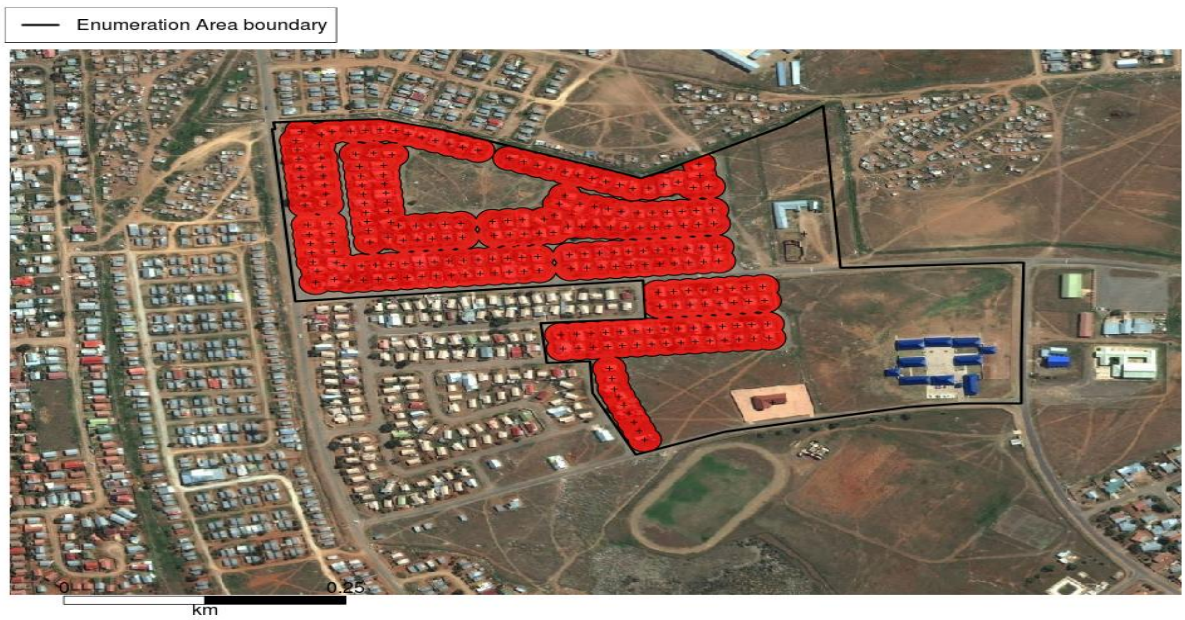
An enumeration area with id='17102473' in the Village 2 North suburb opposite Khayelitsha No 3 Secondary School, Cape Town, has a net building density of **45.1 buildings per Ha** and is 25% dominant to LSM7. The total enumeration area is 5.9 Ha, and the new filtered and buffered area with open space excluded is 2.7 Ha having 124 buildings.



Bing Maps: (18.673071,-34.034706,18.677466,-34.032209)

An enumeration area with id='22100223' in Queenstown across Aloe Park Pre-Primary School has a net building density of **21.6 buildings per Ha** and is 75% dominant to LSM6. The total enumeration area is 18.1 Ha, and the new filtered and buffered area with open space excluded is 10.4 Ha having 226 buildings.





5. RESULTS

By applying the spatial filtering and algorithm as described in the previous section, the net building density was calculated for all 64 506 residential enumeration areas (EA) in South Africa.

All areas that are homogenous to a specific LSM is inspected at different homogenous percentages and the amount of EA areas taken into consideration, the highest level of homogeneousness was chosen where the amount of enumeration areas in the sample was bigger than 100. Recommended density estimates in buildings per Ha that is based on reality are highlighted.

Dominant LSM	mean buildings/HA	95th percentile	max	min	EA areas(n)	Percentage Homogenous
1	1.27	5.61	5.79	0.31	6	10
1	1.7	7.04	5.79	0.31	4	20
1	5.79	5.79	5.79	5.79	1	30

For areas that are dominant to LSM1, there were 6. Being less than 100, there are no reasonable recommendations available for LSM1.

Dominant LSM	mean buildings/HA	95th percentile	max	min	EA areas(n)	Percentage Homogenous
2	2.25	6.24	18.06	0.26	566	10
2	2.22	6.01	18.06	0.26	562	20
2	2.21	5.34	17.51	0.39	383	30
2	1.39	2.86	2.8	0.47	11	40

Dominant LSM	mean buildings/HA	95th percentile	max	min	EA areas(n)	Percentage Homogenous
3	3.36	17.58	106.75	0.01	8525	10
3	3.28	17.01	106.75	0.01	8494	20
3	3.17	16.57	106.75	0.01	8189	30
3	3.68	18.71	91.7	0.01	3456	40
3	10.53	37.24	83.38	0.01	423	50
3	10.15	31.75	43.89	0.43	62	60
3	16.12	39.93	43.89	2.3	32	70
3	14.49	35.61	42.94	2.4	25	80

Dominant LSM	mean buildings/HA	95th percentile	max	min	EA areas(n)	Percentage Homogenous
4	4.64	18.37	173.34	0.16	5390	10
4	4.67	18.45	173.34	0.16	5323	20
4	4.96	19.34	173.34	0.16	4504	30
4	5.22	15.81	79.07	0.2	2142	40
4	6.8	23.05	39.82	0.25	302	50
4	13.7	30.32	39.82	0.62	42	60

Dominant LSM	mean buildings/HA	95th percentile	max	min	EA areas(n)	Percentage Homogenous
5	7.96	29.4	185.17	0.23	11869	10
5	7.98	29.43	185.17	0.23	11834	20
5	8	29.35	185.17	0.23	11256	30
5	9.86	34.55	168.68	0.24	5377	40
5	9.47	32.84	133.85	0.29	2402	50
5	8.59	22.38	50.41	0.43	637	60

5	9.16	29.94	25.18	1.12	5	70

Dominant LSM	mean buildings/HA	95th percentile	max	min	EA areas(n)	Percentage Homogenous
6	19.67	51.22	235.21	0.2	22047	10
6	19.68	51.18	235.21	0.2	21962	20
6	19.9	51.72	235.21	0.2	17949	30
6	19.91	46.25	235.21	0.2	10136	40
6	20.91	51.04	235.21	0.31	3582	50
6	21.04	57.75	235.21	0.34	1044	60
6	22.01	68.97	179.14	0.61	187	70
6	19.67	68.81	157.38	0.77	72	80

Dominant LSM	mean buildings/HA	95th percentile	max	min	EA areas(n)	Percentage Homogenous
7	15.82	43.79	144.71	0.08	3037	10
7	15.85	43.85	144.71	0.08	3026	20
7	17.51	46.25	144.71	0.17	1884	30
7	17.6	46.33	144.71	0.17	941	40
7	19.12	48.1	54.92	0.71	56	50

Dominant LSM	mean buildings/HA	95th percentile	max	min	EA areas(n)	Percentage Homogenous
8	12.09	31.23	60.98	0.3	633	10
8	12.04	31.15	60.98	0.3	631	20
8	12.19	30.99	60.98	0.3	435	30
8	13.88	34.54	60.98	0.3	129	40
8	22.01	22.01	22.01	22.01	1	50

Dominant LSM	mean buildings/HA	95th percentile	max	min	EA areas(n)	Percentage Homogenous
9	13.1	40.03	115.06	0.24	2683	10
9	13.19	40.14	115.06	0.26	2664	20
9	10.27	32.49	115.06	0.26	1518	30
9	9.62	29.62	78.16	0.3	830	40
9	8.48	26.77	75.62	0.3	262	50
9	10.7	31.39	46.44	0.69	43	60
9	9.72	21.18	13.85	5.58	2	70

Dominant LSM	mean buildings/HA	95th percentile	max	min	EA areas(n)	Percentage Homogenous
10	8.4	25.09	98.36	0.16	6732	10
10	8.39	25.05	98.36	0.16	6728	20
10	8.17	24.45	81.49	0.16	6229	30
10	8.28	24.51	81.49	0.16	4746	40
10	8.92	27.12	81.49	0.34	2158	50
10	8.81	26.99	59.67	0.34	893	60
10	9.38	30.41	57.96	0.34	328	70
10	11.1	36.67	57.96	0.34	88	80
10	15.6	39.41	51.31	1.33	28	90
10	21.09	52.75	51.31	6.44	10	100

6. CONCLUSION

Net residential building estimates are now available, based on satellite imagery and various surveys and estimation techniques to give a peak into reality on the subject of domestic building density per LSM in South Africa.

7. FUTURE WORK

Building density was investigated on a national level in South Africa, but there may be regional density distributions that are distinct and significant to a specific region. This can be tested and more optimally identified with a Bayesian Information Criteria trade-off technique that weights regional classifications systems and the overall model complexity (amount of regional estimates) with the increase in accuracy provided for regional building density estimates.

8. ACKNOWLEDGMENTS

This work was funded by Eskom, and directed by the DSM, Energy Efficiency & Customer Research Steering Committee under Amal Khatri, (Activity Leader: Lloyd Setlhogo) undertaken through the Eskom GLF working group headed by Simphiwe Hashe.

9. REFERENCES

[1] Open Street map Project. http://www.openstreetmap.org/
[2] Adri de la Rey. Enabling decision making with the SPOT5 building count – Adri de la Rey, Eskom ESI-GIS
[3] Guideline document for higher density residential development – housing department Ekurhuleni metropolitan municipality – July 2005
[4] Craig Schwabe . Living Standards Measures continue to change in South Africa –Afriscope
[5] R Project. http://www.r-project.org/
[6] Oracle Spatial Developer's Guide, 11g Release 1
[7] Landcom - Residential Density Guide http://www.landcom.com.au/news/publications-and-programs/residential-density-guide.aspx
[8] Guidelines for human settlement planning and design Volume 1 – CSIR under the patronage of the department of housing.
[9] SHEN Wei - Building boundary extraction based on lidar point clouds data. Ocean College of Shanghai Fisheries University, Shanghai 200090, China - shen_wei@sina.com

10. AUTHOR(S)

Principal Author:

Jacques Booysen holds a B.Sc (Hons) Computer Science and B.Sc Applied Mathematics 1998-2001 from the University of Johannesburg. Jacques has been involved in the research and technical development of various systems and research projects for Eskom since 2002, systems varying from operational SCADA networks and automated load shedding over thin networks, business intelligence warehouses, remote protocol telecommunication gateways, statistical modelling and load research. Currently he does load research for Eskom as a data scientist working at Enerweb.

Co-author:

Marcus Dekenah holds a NHD ElecEng HC (Cape Technikon, 1983), M. Dip. Tech. Eng. (Cape Technikon, 1987), a BSc Eng (1st Cl Hons, UCT, 1991) and MBL II (Unisa, 1996).

Marcus Started his career at the Electricity department at Cape Town City Council involved in a wide range of distribution operations (1980-1988).

From 1991- 1995 Marcus was executing contract research projects in HV testing, Energy management and Load research at Electric Power programme, Division of Energy Technology, CSIR.

In 1996 Marcus formed his own company and has since conducted contract load research (domestic, and non-domestic) both inside and outside South Africa with multi-disciplinary teams of specialists from industry, private, and the academic sector.

Marcus is currently assisting the Demand Intelligence Group of Enerweb with the execution of load research. Enerweb is a Southern African company focused on delivering solutions and services that enable utilities to manage demand flexibility.

Co-author:

Lloyd Setlhogo holds a ND ElecEng HC (Pretoria Technikon, 1997), B. Tech Power Eng (Unisa, Final Year).

Lloyd started his career at Technology Services International (TSI) a Division of Eskom Enterprises in 2000 as a Senior Technician. In 2005 Lloyd became the Load Research Project Leader. Lloyd was then promoted to Senior Advisor Engineering at Eskom Research, Testing and Development (2012).

Lloyd is currently managing the Load Research Project in Eskom Research (Demand Management Section) and has a wealth of experience in Eskom of 14 years.

Presenter: The paper is presented by Jacques Booysen.