# A Novel Method to Model Spatial Distribution of Population

**A. B. Jayasinghe, H. W. A. S. Sathsarana, R.M.Y.L. Rathnayake,N. S. Bandara**

*Abstract:This study proposes a simulation approach to model the spatial distribution of population (i.e., population density) in a given regioninto a great deal of detail. The approach is based on two mathematical notions: graph theory and fractal geometry. Accordingly, the approach takesself-similarity of road accessibility into account when simulating the spatial distribution of population. The study has followed three steps; (a) formulatea conceptual framework to model the spatial-distribution of population in a given area (i.e. a city or a region) by employingthe network centrality-based fractal dimension, (b) develop a modeling framework, and (c) calibrate the model utilizing empirical data at for five-selected case study areas: Colombo-Sri Lanka, Hanoi-Vietnam, Kolkata -India, Wuhu-China and Singapore. Network centrality-based fractal dimensions were computed by employing open-data and open-source GIS tools. The study applied threetypes of regression techniques: Robust Regression (RR), Ordinary-Least-Squares Regression (OLSR) and Poisson Regression (PR) for derivingthe most appropriate mathematical model. Then, the study validated the model by testing its prediction accuracy. Results revealed that the developed model to simulate spatial distribution of populationin a given area recorded an accepted level of accuracy ($R^2>0.75$) and predictability (MdAPE< 10%) on a par with the internationalspatial modeling standards. The proposed approach can be adoptedto simulate the spatial-distribution ofpopulation, particularly as a decision-making aid in the domain of urban & regional planning.*

*Keywords:Accessibility,Fractal Geometry,Open Source GIS,Population Density,Spatial Analysis,Spatial Planning,Transport Network,Urban Simulation*

## I. INTRODUCTION

Dueto the rapid growth of population and unprecedented urbanization,it becomes a challenging task to detect the spatial distribution of population at micro-scale spatial units. The limited ability to scale down population data due to its uneven distribution, the long interval betweentwo population census (i.e., typically ten years) as well as the unavailability of accurate, dynamic and cost-effective methods to model the distribution of a given population have become a challenge for decision-makers [1].Conventional population forecasting methods which are based on the birth rate, death rates, and net migration have a limited capability in capturing the influence of socio-economic and physical factors affect on population distribution such as residential attraction, accessibility, carrying capacity, resource availability, economic return and property demand[2].

In such context, this study attempts to introduce a novel approach to model the spatial distribution of population (i.e. population density) based ona theoretical framework that innovatively synergizes two mathematical notions: network centrality and fractal dimensions. The method developed in this study considers the self-similarity of the growth and distribution of the fractals of road segments.

Fractal geometry is a concept evolved through centuries by considering the geometry which includes the properties of space such as the shape, the size and the relative position of figures [3]. Fractal geometry follows the principles of non-Euclidean geometry which, describes non-integer dimensions of fractal properties; and interprets fractals found in nature as scale-invariant, self-similar objects[4]. Sambrook& Voss noted that both city form and the functions are also could be presented as fractals which are irregular, scale-invariant and self-similar[5]. Thus, the city form and functions together could be analyzed based on fractal geometry to explain the relationships between the form and functions. Previous studies on fractal geometry and population have identifiedthat the fractal geometry of road patterns and population distribution has an association which might provide an ability to predict each factorbased on the changes happened to the other[6].

Though the association between accessibility and distribution of population is clear, this fractal association has not been tested yet to find out whether its accuracy isversatile enough for model population density in a given city or a region [9]. This study was intended to develop a spatial simulation approach to model the distribution of population in a given city region, which is applicable for micro-scale spatial units. The study is attempted to develop the approach based on theassociation of spatial distribution of population and spatial variations of accessibility. The research has three (03) main steps: formulate a conceptual framework to model the spatial distribution of population (i.e. population density) in a given region by utilizing the network centrality-based fractal dimensions; develop a modeling framework; and calibrate the model utilizing empirical data in selected case study areas.

**A. B. Jayasinghe** *, Urban Simulation Lab, Department of Town &Country Planning, University of Moratuwa,Mora tuwa,Sri Lanka. Email: amilabj@uom.lk

**H. W. A. S. Sathsarana**, Urban Simulation Lab, Department of Town &Country Planning, University of Moratuwa,Mora tuwa,Sri Lanka.. Email: sahansashi96@gmail.com

**R.M.Y.L. Rathnayake**, Urban Simulation Lab, Department of Town &Country Planning, University of Moratuwa,Mora tuwa,Sri Lanka. Email: yasirulr@gmail.com

**N. S. Bandara**, Osaka City University, Japan. Email: sanj88@gmail.com

## II. METHOD OF STUDY

Temporal data on population and urban settlement for five-selected case studies: Colombo in Sri Lanka, Hanoi in Vietnam, Kolkata in India, Wuhu in China and Singapore were collected from the world pop web portal prepared by University of Southampton. The transport network data were collected from Open-Street-Maps. Network centrality-based fractal dimensions were computed by using open-sourceGeographic Information System (GIS) tools. Trend and pattern analysis tools in GIS were utilized for the relationship identification. Further, the study utilized three types of regression techniques such as Robust Regression (RR), Ordinary-Least-Squares Regression (OLSR) and Poisson Regression (PR) to derive the best mathematical model. Then, the calibrated model was validated by testing its prediction accuracy.

There are different methods in calculating fractal geometry. Box counting method is the most famous one which includes two sub-methods [8]. The second one is the Mass-radius method. Box counting method provides the fractal geometry of a city as a whole whereas the mass-radius method measures the fractal geometry of city as zones of different sizes. Therefore, the mass-radius method provides a better spatial understanding on the changes to the fractal geometry made by any physical development or other physical changes occurred in the city as a whole as well as within its subsystems. The mass-radius method has the capability to explain the spatial dynamics of a city [9]. By considering these factors, this study is opted for mass radius method in measuring fractal geometry of the roads and population distribution.

The mass-radius method calculates the fractal dimension based on the dependency of black and white pixels on a given circle-shaped plain at different sizes of radius.The fractal dimension measures as the ratio of the change of the different sizes of radius. It assumes that all fractal geographical entities demonstrate adimension relationship, i.e., a self-similar pattern with each other (Eq. 1).

$$L^{1/1} \alpha\ A^{1/2}\ \alpha\ V^{1/3} (1)$$

Wherein,'L' is the length of a geographical entity, 'A' is the area, 'V' is the volume [4]. The study utilized the closeness centrality of road segments to represent L value which represents the notion of accessibility. Closeness centrality measures how close a given location to all other locations along the shortest path in a given network[10] (Eq. 2).

$$CC_i = \frac{(N-1)}{\sum_{j \in N, j \neq i} d_{ij}} (2)$$

Wherein, $'CC_i'$ is the closeness centrality of the road segment [or links] '$i$'; 'N', is the total number of road-segments [or links] in a given road network and $'d_{ij}'$, is the distance [i.e. travel cost] between road segment [or links] '$i$' and '$j$' along the shortest distance path.

The study utilized the angular distance [11] to compute distance between two locations.'Spatial-Design-Network-Analysis' (sDNA) tool was utilizedto compute the CC values of road segments. Fig. 1 illustrates the key steps followed in the study to compute the closeness centrality values of road segments (refer Fig. 1).

The study drew buffers from the geographic center to the peripheral edge of the city at 500m intervals (refer Fig. 2). Then, the study computed the fractal dimension of road segments by utilizing Eq. 3.
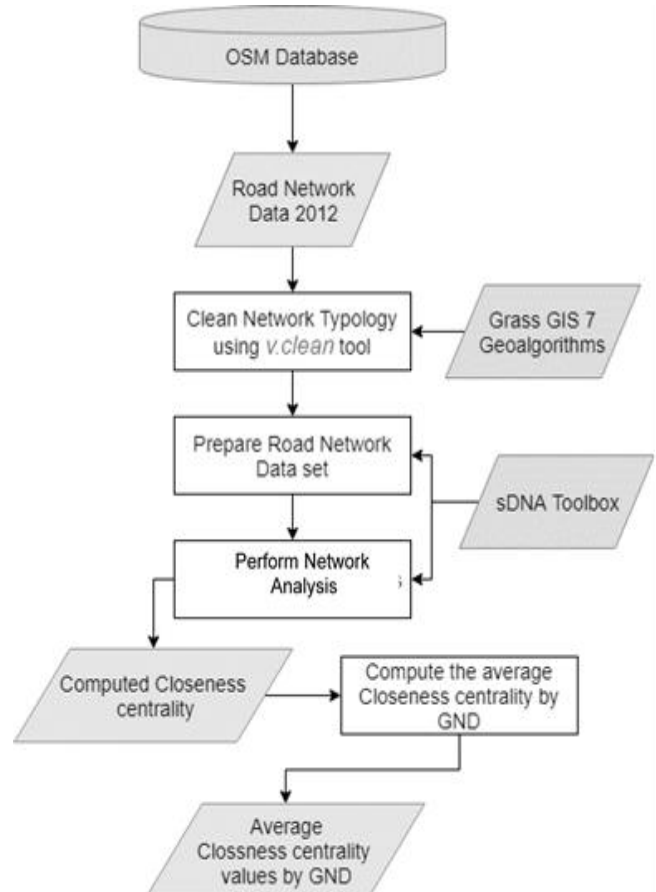


**Fig. 1: Key stepsfollow in the study to compute closeness centrality values of road segments**

$$D(Li) = \frac{Log\left[\frac{L(R_i)}{L(R_{i-1})}\right]}{Log\left(\frac{R_i}{R_{i-1}}\right)} (3)$$

Wherein, 'D(L$i$)' is the fractal dimension of the road segmentswithin the buffer at the $i^{th}$radius, 'R$i$' is the radius,'L (R$i$ )' is the average closeness centrality at the $i^{th}$radius.

In the next step, the study computed the fractal dimension of population utilizingEq. 4. Wherein, 'D(P$i$)' is the fractal dimension of population density withthe $i^{th}$radius buffer, 'R$i$' is the radius and 'P (R$i$ )' is average population density with in the $i^{th}$ radius buffer.
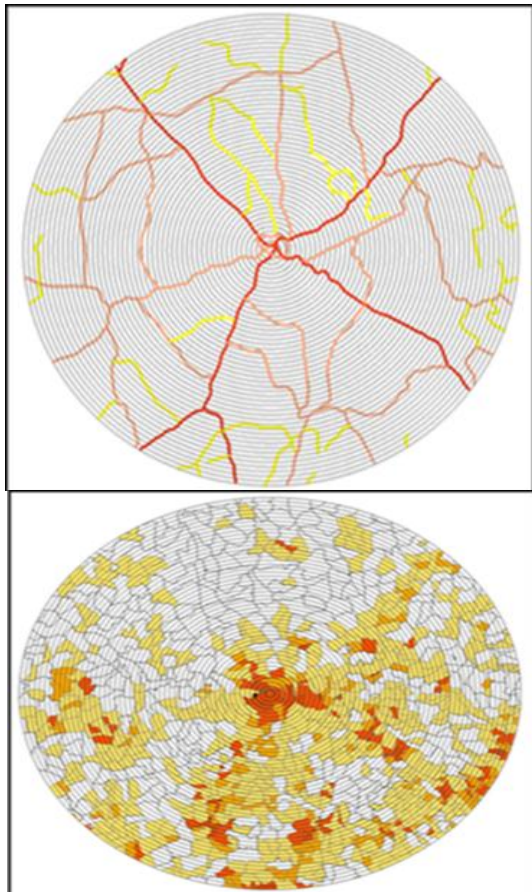
$$D(P_i) = 2 + \frac{Log\ P(R_i)}{Log\ (R_i)} (4)$$

35

**Fig. 2: 500 buffers on the road (top) and 500m buffers on the population density (bottom)**



**Fig. 3: Distribution of fractal dimension of roads closeness centrality D(L$i$) (Top) and distribution of fractal dimension of population density distribution D (P$i$) (Bottom).**

## III.  ANALYSIS AND RESULTS

Fig. 3 depicts the distribution of the fractal dimension of the road segments values i.e.,D(L$i$) and the distribution of fractal dimension of population density valuesi.e.,D (P$i$).The pattern of two graphs visualy depicts the self-similarity between the fractal dimension of the road segmentsD(L$i$) (Top) and the fractal dimension of population density D (P$i$).This supports the relationship prposed inEq. 1.

The study analyzed theabove-mentiond patternutilizing visual analysis and spatial statistical analysis methods availble in the refered open-source GIS envrionment (Refer Fig. 4, 6, 8, 10 and 12). Further, the study analyzes the distribution of fractal dimension of population density D (P$i$) values and fractal dimension of the road segments D(L$i$) values of each cityas depicted inline graphs(Refer Fig. 5, 7, 9, and 11).D (P$i$) values are depicted in blue color and D(L$i$)values are depicted in orange color for the purpose of visual comparison.

Fig. 5, 7, 9, and 11depicts a recognizable pattern of self-similarity between the distribution of the fractal dimension of the road segments and the fractal dimension of correspondent population density values.
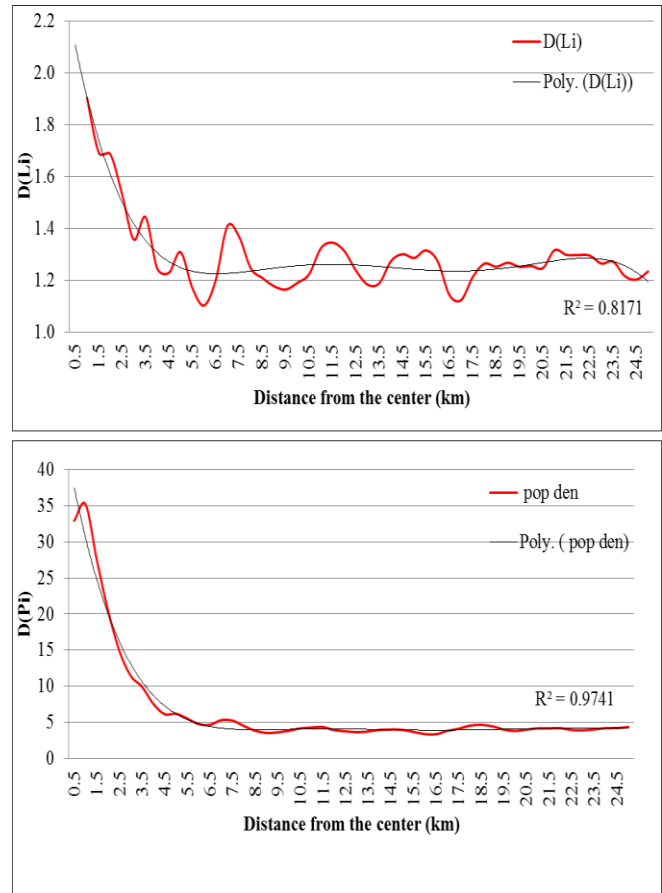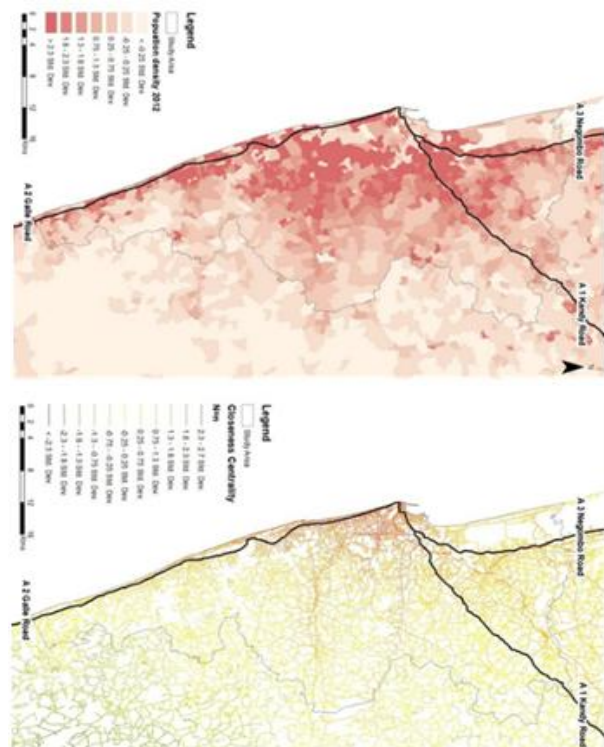


**Fig. 4: The fractal dimension of the distributionof population density D (P$i$) (top) and the fractal dimension of the closeness centrality values of roadsegments D(L$i$) (bottom) in Colombo-Sri Lanka case study area**
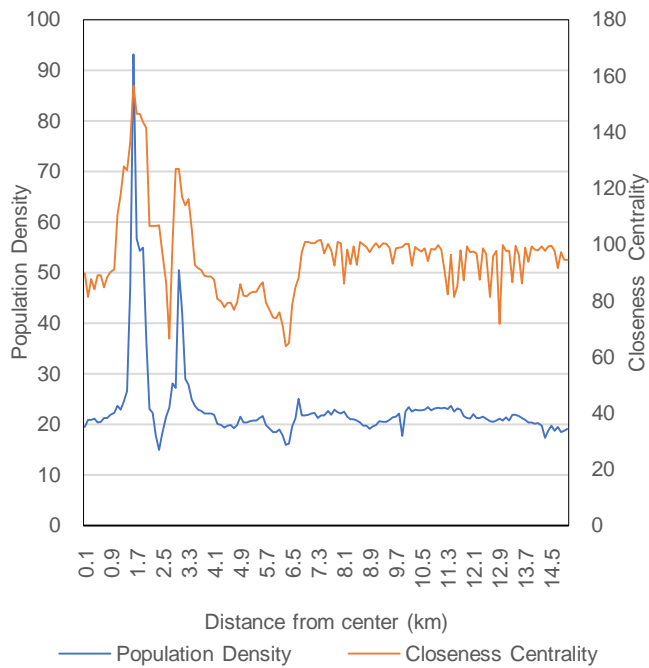
**Fig. 5: The pattern of the fractal dimension of population density and the pattern of the fractal dimension of the closeness centrality values of roads in Hanoi-Vietnam case study area**
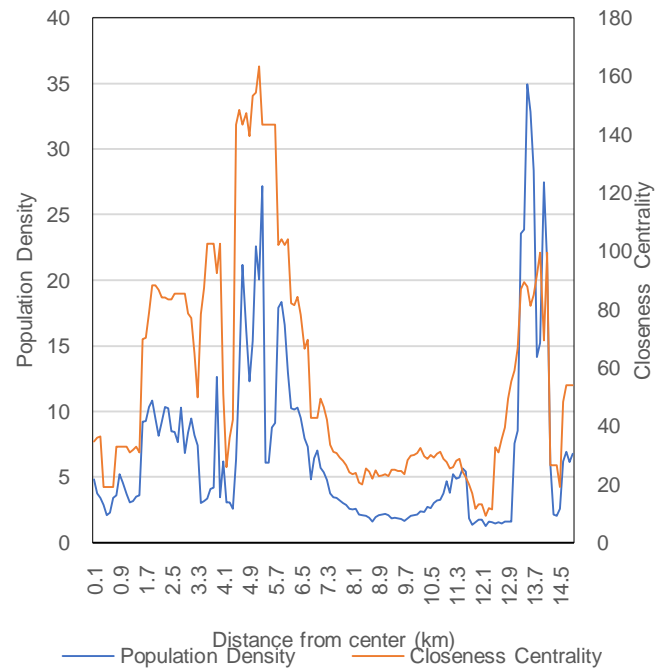


**Fig. 6: The pattern of the fractal dimension of population density and the pattern of the fractal dimension of the closeness centrality values of roads in Wuhu-China case study area**
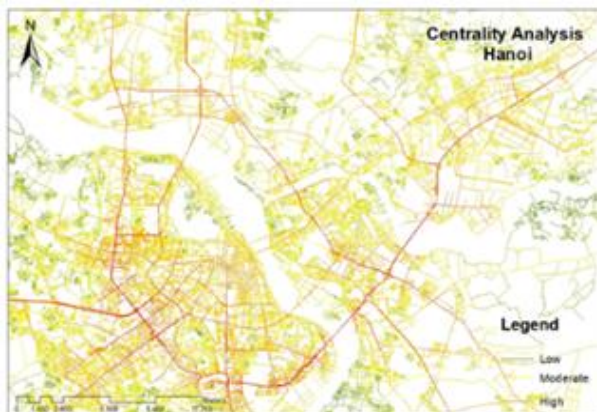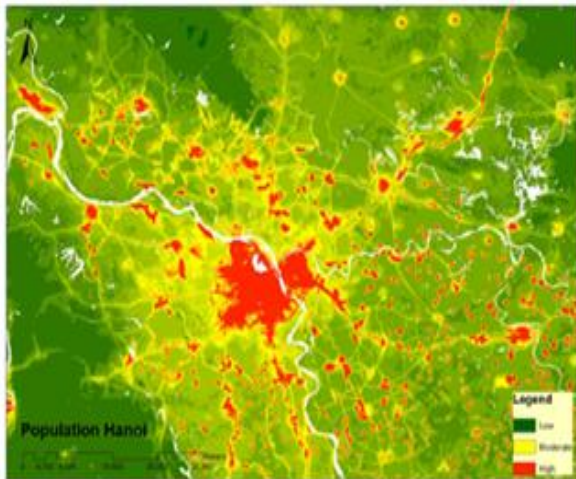


**Fig. 6: The fractal dimension of the distribution of population density D (Pi) (top) and the fractal dimension of the closeness centrality values of road segments D (Li) (bottom) in Hanoi-Vietnam case study area**
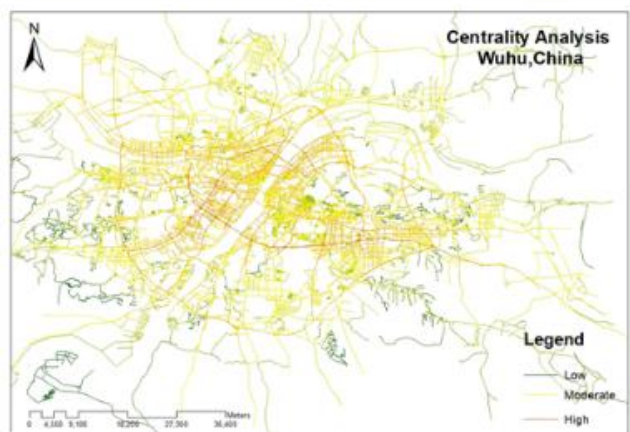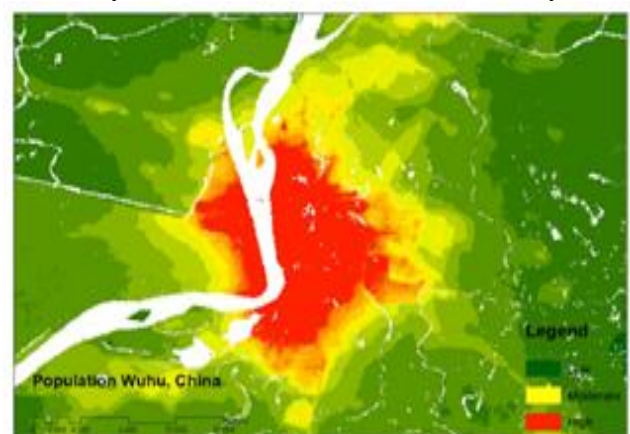


**Fig. 7: The fractal dimension of the distribution of population density D (Pi) (top) and the fractal dimension of the closeness centrality values of road segments D (Li) (bottom) in Wuhu-China case study area**

**Fig. 8: The pattern of the fractal dimension of population density and the pattern of the fractal dimension of the closeness centrality values of roads in Singapore case study area**
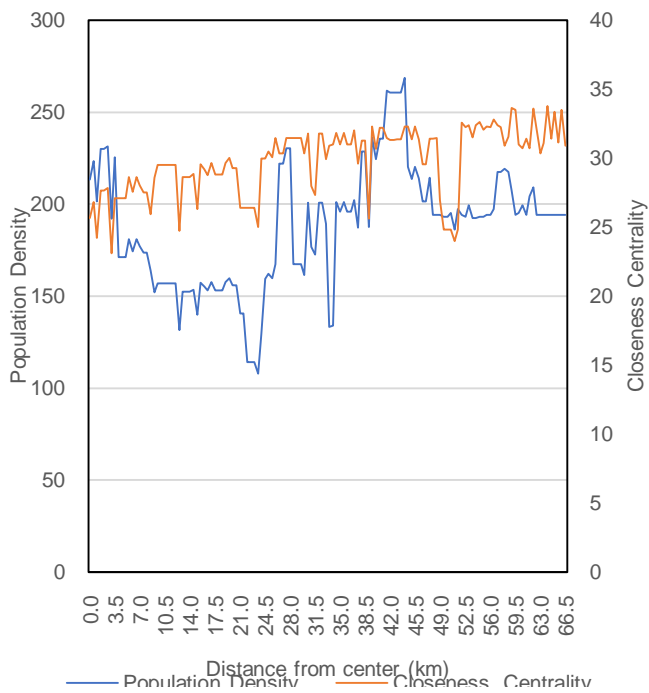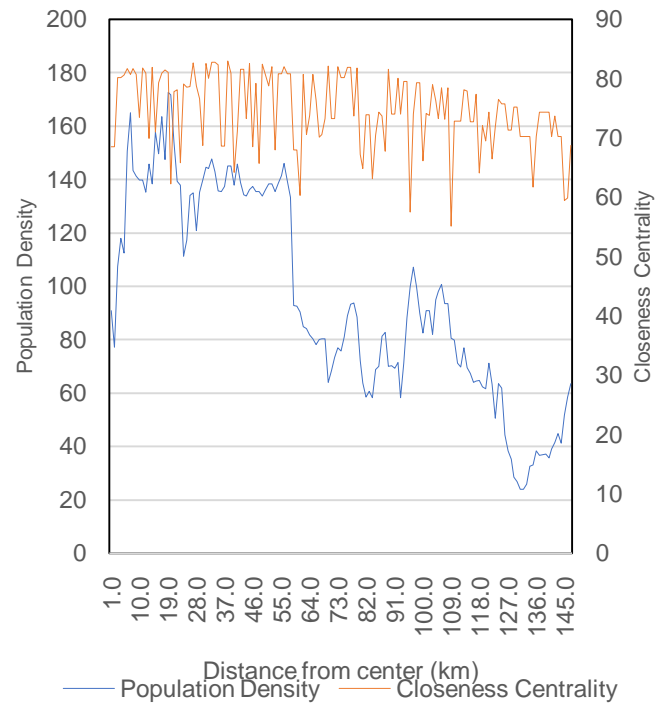


**Fig. 10: The pattern of the fractal dimension of population density and the pattern of the fractal dimension of the closeness centrality values of roads in Kolkata case study area**



**Fig. 9: The fractal dimension of the distribution of population density D (Pi) (top) and the fractal dimension of the closeness centrality values of road segments D (Li) (bottom) in Singapore case study area**
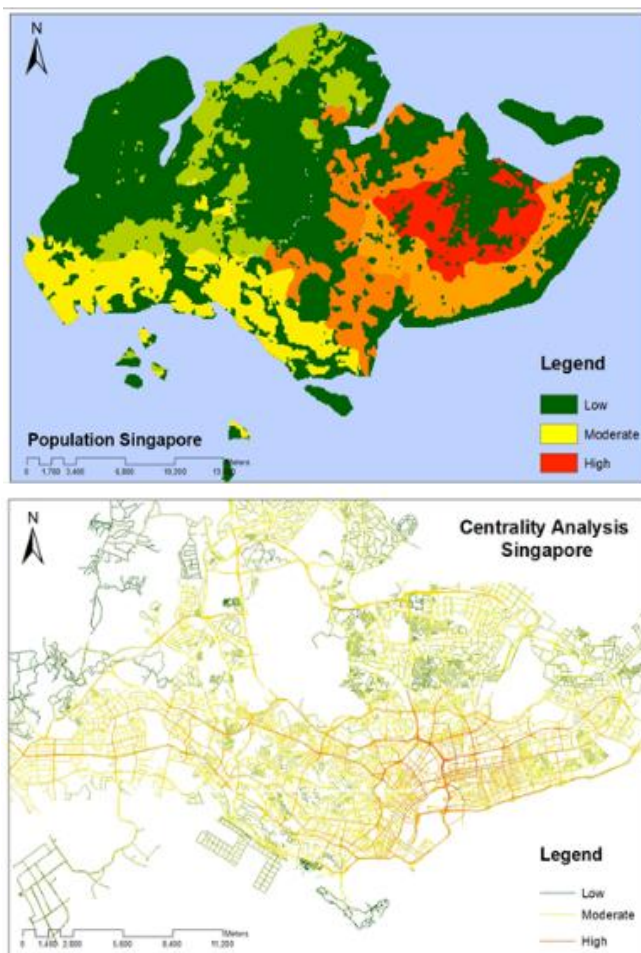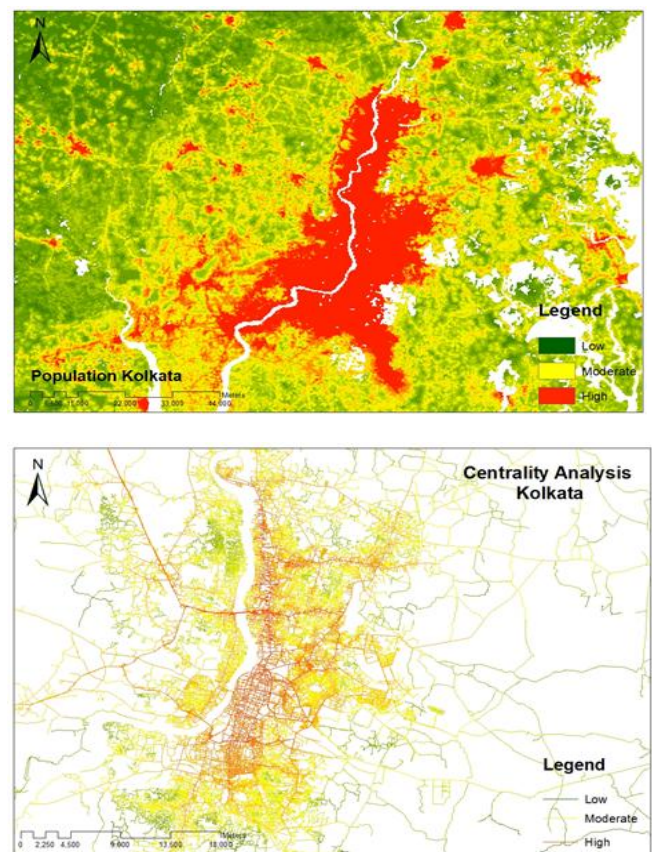


**Fig. 11: The fractal dimension of the distribution of population density D (Pi) (top) and the fractal dimension of the closeness centrality values of road segments D (Li) (bottom) in Kolkata case study area**

In the next step, the studydeveloped a model to simulatethe fractal dimension of population densityD (Pi) based on the fractal dimension of the road segments D(Li). For this purpose, study utilized three regression methods:Robust-Regression (RR), Ordinary-Least-Squares Regression (OSR)and Poisson-Regression (PR). This study tests the goodness of fit andthe predictability error with reference to $R^2$ (refer Eq. 4) andMedian-Absolute-Percent-Error (MdAPE) (refer Eq. 5) respectively.

$$R^2 = 1 - \frac{\Sigma(y_n - \hat{y}_n)^2}{\Sigma(y_n - \bar{y})^2} (5)$$

Wherein,'$R^2$' is the coefficient of determination, '$y_n$' is the actual population density, '$\hat{y}_n$' is the estimated population density and '$\bar{y}$' is the mean of the actual population density.

$$MdAPE = Median\left(\frac{|y_n - \hat{y}_n|}{y_n}\right) * 100 \qquad (6)$$

Wherein,MdAPE is the Median Absolute Percent Error,$y_n$is the actual population density, $\hat{y}_n$ is theestimated population density and N is thetotal number of data points. Summary results of thegoodness of fit of each cased study area is depicted in table 1. Regression analysis betweenthefractal dimension of population density distribution and the fractal dimension of road segments provide a convincingresult. $R^2$value of all five case study areas were recordedover 0.75(i.e. 75%) accuracy. MdAPE value of all five case study areas is less than 10%. It indicates that there is a strong associationbetween two variables and a strong possibility of predicting one variable utilizingthe other. Accordingly, the results proved the study hypothesis that the fractal dimension of the distribution of population density D(P$i$) can be modeled by the fractal dimension of the distribution of the road segments D(L$i$). Accordingly, it can be concludedasthe distribution of population density can be modeledbyEq. 7.

**Table-I: The accuracy of thedeveloped model by case study area**

| Case Study Areas | R square | Adjusted R Square | MdAPE |
|---|---|---|---|
| Colombo-Sri Lanka | 0.892 | 0.860 | 7.48% |
| Hanoi-Vietnam | 0.834 | 0.811 | 8.01% |
| Kolkata -India | 0.780 | 0.758 | 8.85% |
| Wuhu-China | 0.843 | 0.820 | 6.78% |
| Singapore | 0.778 | 0.753 | 9.48% |

$$D(P_i) = a \, D(Li)^b \qquad (7)$$

Wherein, 'D(Pi)' is the fractal dimension of the distribution of population density within the$i^{th}$ radius, 'D(Li)' is the fractal dimension of the distribution of road segments at i$^{th}$ radius, which is measured by closeness values. 'a' and 'b' are constant values that need to be calibrated for each case study area.

## IV. CONCLUSION

The study developeda simulation approach to model the distribution of population (i.e. population density) in a given region. The model is developed based on two mathematical notions i.e., graph theory and fractal geometry. The study tested the developed modeling framework in five-selected case study areas: Colombo-Sri Lanka, Hanoi-Vietnam, Kolkata -India, Wuhu-China and Singapore. Results indicated that the proposed model for simulating the distribution of population density has obtained a level of predictability and accuracy ($R^2$> 0.75 and MdAPE< 10%)up to an internationally accepted standard in the domain of spatial modelling. Given the successful results, the study recommends that the proposed simulation approach can be employed as a population simulation tool which would certainly benefit spatial planning decision-makers in the domains of urban and regional planning.

The proposed method can be considered as an accurate, technically simpler, less data consuming, and financially affordable approach to model thedistribution of population density in a given region. This is especially beneficial for situations with limited data and resource availability as typically in developing countries. This method can be realistically employeddecision makingfor spatial planning purposes such as modelling the spatial distribution of population density, identification of the impact of changes in transportation network on the distribution of population density, and identification of the required accessibility improvements based on an estimatedpopulation in a given region.

Though this study has been completed by successfully achieving the desired objectives, it paves a path for further studies on the population density distribution particularly in spatial planning and transport planning applications. A few of the future research areas may include, but not limited to the followings; modellingtemporal variationsuch as daytime population and nighttime populationin thegiven region andin modellingthe regional migration.

### REFERENCES

1. A. Jayasinghe, K. Sano, C. C. Abenayake and P. K. S. & Mahanama, "A novel approach to model traffic on road segments of large-scale urban road networks," *MethodsX*, 2019, vol. 6, pp. 1147-1163.
2. A. Jayasinghe, N. T. Tenzin, C. Abenayake and P. Mahanama, "Simulation of Population Growth and Distribution based on Centrality of Road Network," *International Journal of Innovative Technology and Exploring Engineering*, 2019, vol. 8, no. 10S, pp. 389-393.
3. S. H. Lv and C. Cao, "The Evolution of Mathematics Curriculum and Teaching Materials in Secondary Schools in the Twenty-First Century," The 21st Century Mathematics Education in China, Berlin: Springer, 2018, pp. 147-169.
4. M. Batty, "Cities and complexity- understanding cities with cellular automata agent-based models and fractals," Cambridge: MIT press, 2017.

5. R. C. Sambrook and R. F. Voss, "Fractal analysis of US settlement patterns," *Fractals*, 2001, vol. 9, no. 3, pp. 241-250.
6. Y. Lu and J. Tang, "Fractal dimension of a transportation-network and its' relationship with urban growth: a study of the Dallas Fort Worth area," *Environment and Planning B: Planning and Design*, 2004, vol. 31, no. 6, pp. 895-911.
7. X. Shi, J. Pan, Y. Jin, Z. Wang, Q. Niu and M. Li, "Micrometer scale fractures in coal related to coal rank based on micro CT scanning and fractal theory," *Fuel*, 2018, vol. 212, pp. 162-172.
8. Y. Chen and J. Wang, "Multifractal characterization of urban form and growth: the case of Beijing, China," Environment and Planning B: Planning and Design, vol. 40, no. 5, pp. 884-904, 2013.
9. A. Albert, J. Kaur, E. Strano and M. Gonzalez, "Spatial sensitivity analysis for urban land-use prediction with physics constrained conditional generative adversarial networks," *ArXiv*, 2019, vol. 1907.09543.
10. S. Porta, V. Latora, S. Rueda, E. Strano, S. Scellato, A. Cardillo, E. Belli, F. Cardenas, B. Cormenzana and L. Latora, "Street centrality and the location of economic activities in Barcelona," *Urban Studies*, 2012, vol. 49, no. 7, pp. 1471-1488.
11. A. Jayasinghe, K. Sano, R. Kasemsri and H. Nishiuchi, "Travelers' route choice: comparing relative importance of metric, topological and geometric distance," *Procedia Engineering*, 2016, vol. 142, pp. 18-25.
12. T. Elmqvist, M. Fragkias, B. Güneralp, P. Marcotullio, R. McDonald, S. Parnell, M. Schewenius, M. Sendstad, K. Seto and C. Wilkinson, A global outlook on urbanization. In Urbanization, biodiversity and ecosystem services: Challenges and opportunities, Dordrecht: Springer, 2013.

## AUTHORS PROFILE

**A. B. Jayasinghe**
B.Sc.[Hons] (Town & Country Planning), SL;
M.Plan.(Infrastructure Planning) India;
D.Eng.(NUT, Japan);
MITP(Sri Lanka) Chartered Town Planner.
Dr. Jayasinghe is working as a senior lecturer at the Department of Town & Country Planning, University of Moratuwa, Sri Lanka. His research works are focused on infrastructure planning, spatial simulation and open-source GIS applications in spatial planning.

**H. W. A. S. Sathsarana**
Undergraduate - B.Sc.[Hons] (Town & Country Planning), SL.
Mr. Sathsarana is currently working at the Urban Simulation Lab, Department of Town & Country Planning, University of Moratuwa as a research assistant. His main research interests are spatial modeling and open-source GIS applications.

**R.M.Y.L. Rathnayake**
B.Sc.[Hons] (Town & Country Planning, SL.
Mr. Rathnayake is a Graduate studentcurrently working at the Urban Simulation Lab, Department of Town & Country Planning, University of Moratuwa, Sri Lanka as a research scholar. His main research interests are spatial planning and design.

**N. S. Bandara**
B.Sc.[Hons] (Town & Country Planning), SL;
M.Urban Informatics (OCU, Japan).
Mr. Bandarais a doctoral student at Osaka City University, Japan. His research works are focused on developing generic software platformsand frugal engineering approaches for spatial datacollection and manipulation with open-sourcesynergy. He is a chartered member of OpenSource Geospatial Foundation (OSGeo).