18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

Modelling urban spatial structure using Geographically Weighted Regression

Noresah M. S. and **Ruslan, R.**

School of Distance Education, Universiti Sains Malaysia
Email: noreshah@usm.my

Abstract: This paper examines urban spatial structure in terms of urban built-up area of the rapidly developing city of Sungai Petani Malaysia. We developed a model of urban built-up area using Geographically Weighted Regression (GWR) to estimate the strength of the relationship between urban built-up area and factors associated with urban change. Ordinary regression models yield only a single estimate of the relationships. In comparison, GWR allows an estimate of the spatial variation of this relationship. In this study twenty explanatory variables describing access and proximity, neighbourhood, zoning and physical factors, were hypothesized to influence the change in the built-up area and analysed using GWR to allow for spatially varying relationships across the study area.

The dataset includes the amount of urban built-up area that has increased from 1992 to 2002 and 20 spatial variables. For the period of 1992-2002, approximately 158 percent of land had been converted to urban use. The spatial variables are generated using geographical information system techniques and calibrated into GWR model. The results of the GWR model are compared to global model. The use of GWR has increased the strength in the relationship especially in terms of the goodness-of-fit statistics (R^2) from 0.29 (OLS global model) to 0.63 (GWR), with individual GWR models ranging from 0.0 to 0.99. Maps of the residuals show that the GWR model fits better in the central region of the study area than the outer region. This is partly due to the better accessibility in the central than the outer region. A Monte Carlo test of the GWR model found that 17 of the 20 explanatory variables displayed significant spatial non-stationarity.

Keywords: *Urban spatial structure, GIS, Geographically Weighted Regression, developing city.*

1. INTRODUCTION

Change in the urban spatial structure of cities and their region is a major theme within urban studies and has long been the subject of academic research (Bertaud, 2004; Ruslan and Noresah, 2004; Soia, 2000; Knox, 1993). An urban spatial structure is a spatial arrangement of a city in which it is a result of the interaction between land markets, topography, infrastructure, taxation, regulations and urban policy over time (Bertaud and Stephen, 2003). Urban spatial structures can be defined using indicators such as average land consumption, spatial distribution of population and daily trips pattern (Bertaud, 2004).

In Malaysia, rapid urbanization for the last two decades due to industrialization, an increase in economic prosperity and urban population has opened opportunities to new developments in urban areas and their periphery. This growth has significantly changed the landscape of many cities. In the context of urbanization, a large amount of agricultural land has been converted to built-up or urban land uses. The dynamics of urban landscape systems of Malaysian cities is driven by complex political, social and economic systems. Development process in Malaysia is driven by 5-year development plans in which at the state level, the structure plan and local plan play roles in every development process. Development in Malaysia can be divided into two stages (Salleh, 2000). Before 1980, rapid urbanization process in the country resulted in urban concentration in order to optimize the agglomeration of economic and transport networks. Major cities such as Kuala Lumpur, Georgetown, Ipoh, and Johor Bahru were primary commodity-based economies. They were located at the confluence of rivers or located at the coastlines or at road or rail intersection points. After 1980, Malaysia experienced a major transformation into an industrial-based economy. Many new urban centres were developed in the urban periphery or non-metropolitan areas. However, little is known about the nature of urban development pattern or the spatial structure of the urban areas in Malaysia.

Modelling an urban development pattern is a prerequisite to understanding the process of urban form and changes. In order to understand the process and change of urban development patterns in Malaysian cities, Geographically Weighted Regression (GWR) modelling approach (Fotheringham et al, 2002) is applied to a set of spatial data of Sungai Petani., Malaysia. The aim is to examine the spatial variation in the relationship between urban built-up area and several determining factors of change and to analyse the implications of the variations on the Malaysian urban planning system.. GWR is based on a geographical weighting function that link spatial data to locations of points under study across the study area. It is assumed that the parameter estimates of GWR model will vary across the study area instead of producing one single estimate as in ordinary regression models. (Fotheringham et al., 2002; Foody, 2003; Platt, 2004; Zhang et al, 2004; Laffan, 2005).

2. THE STUDY AREA – SUNGAI PETANI

Sungai Petani is the capital of Kuala Muda district and is located in the northwestern region of Peninsular Malaysia, in the state of Kedah. Refer to Figure 1. The district extends from 5° 54′ to 6° 06′ N and 100° 48′ and 100° 57’ E. Sungai Petani is located about 35 km to the north of the metropolitan city of Georgetown, Penang. The district covers an area of 925 km^2. Sungai Petani was a small town with a population of 116,977 in 1991. However the town has gradually developed over the years especially after the opening of the North-South Expressway in 1991. This has increased access to other areas particularly to Penang State in the south. It is one of the examples of new urban growth centres in a non-metropolitan area. With the rapid growth of urban extent and urban population, Sungai Petani has been categorized as a Semi-Regional Centre with a population of 174,609 in 2000, estimated to increase to 230,675 by 2010 (Government of Malaysia, 1996). Sungai Petani is a major administrative and commercial centre for Kuala Muda and the surrounding areas. The town is located at the intersection of two federal roads and a federal railway which provides greater accessibility to the study area. The town can also be access via the North South Expressway through three toll intersections

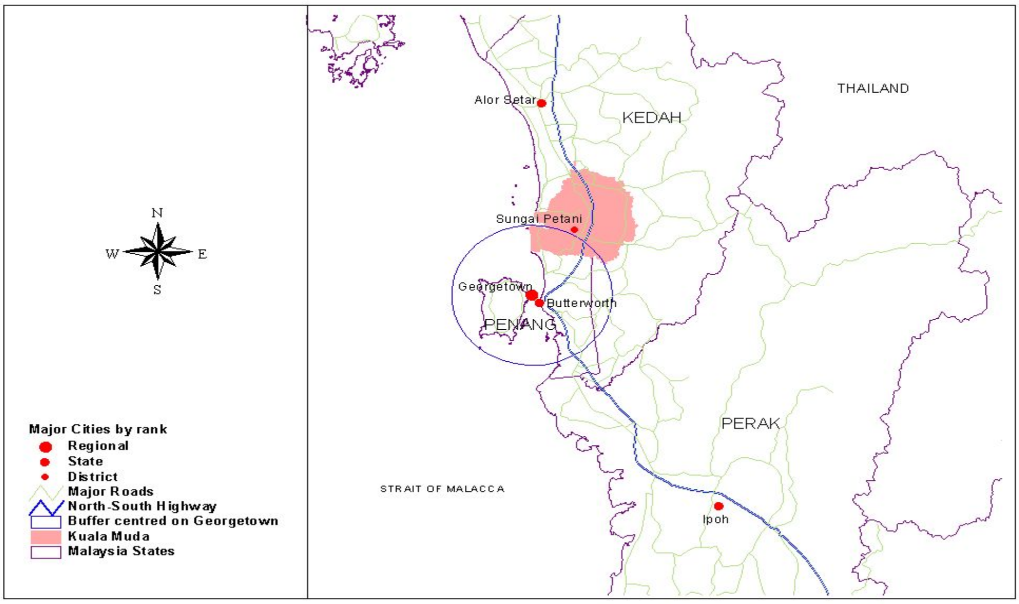

Figure 1: Location of the study area, Sungai Petani, Malaysia.

3. METHODS

3.1. Variables and spatial analysis

The dependent variable used in the analyses is the increase in the built-up area over the ten year period from 1992 to 2002. The amount of change in the built-up area is calculated in hectares by subtracting built-up area of 1992 from that in 2002. The built-up area of Sungai Petani has increased from 3490 hectares in 1992 to 8990 hectares in 2002, an increase of 150 percent over the period. The spatial structure of built-up area of Sungai Petani is shown in Figure 2, from which it is evident that urban development mainly occurs along major thoroughfares.

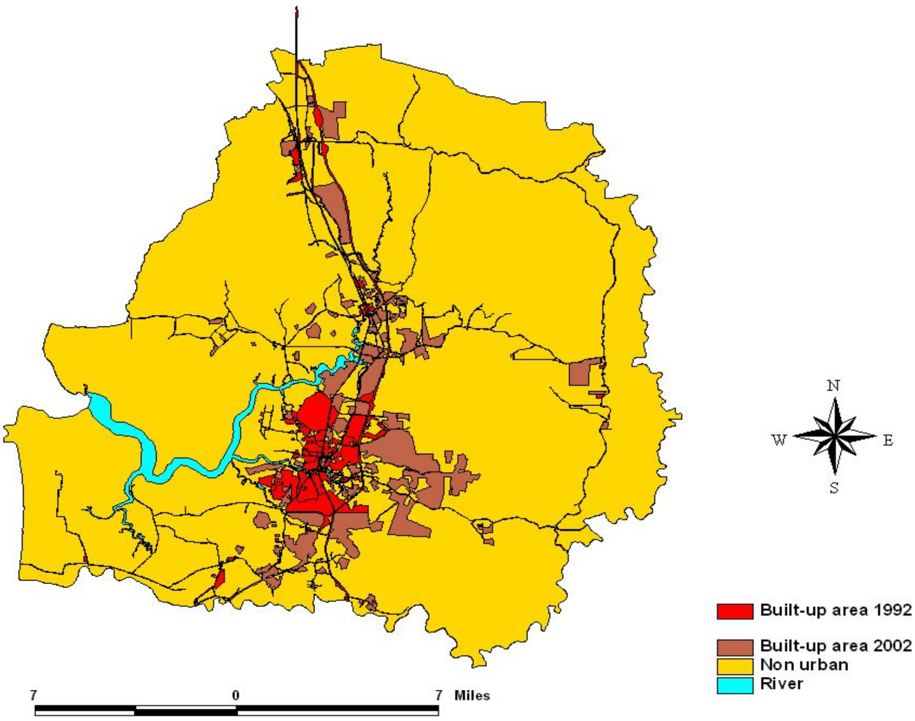

Figure 2: Urban Spatial Structure of Sungai Petani, 1992 and 2002.

The independent variables are divided into four groups: location/proximity, neighbourhood, zoning and physical factors. Table 2 shows variables used in this study and their descriptions. A set of 3702 points on a square lattice with a spacing of 500m is used as the basis for spatial and statistical analysis.

GIS spatial analysis is carried out for each variable and stored as attribute for each point and summarised in the data set's attribute table. The relationship between urban built-up areas and location and proximity variables are calculated using road network travel distance, an indicator of accessibility from the study area to other areas via tolled interchanges, and also using Euclidean distance. This analysis provides the quantity of built-up area at some distance from the nearest interchanges. The spatial neighbouring effect was measured by means of a neighbourhood urban index, which is the average percent of urbanized land around the point of observation. Zoning variables are analysed using zoning plan of Sungai Petani. This plan is intersected with points of observation and amount of land zoned for urban uses is recorded. Land available for development is derived from agricultural land and vacant land, while slope is calculated in degrees.

Table 2 Variables used in the analysis

Variables		Descriptions
Dependent	CHG_URB	Amount of urban built-up area in 2002 (ha)
Independent:		
Access and proximity variables	APV_INTC APV_CISP APV_EMP APV_RGC APV_RSTN APV_IND APV_PORT APV_APORT APV_URB APV_SPT APV_FEDRD1 APV_FEDRD67	Road network distance to nearest interchanges Road network distance to centres in the study Road network distance to nearest employment centres Road network distance to nearest regional centres Road network distance to nearest train station Road network distance to nearest industrial areas Road network distance to nearest sea port Road network distance to nearest airport Proximity (Euclidean distance) to nearest urban built-up areas Proximity (Euclidean distance) to town of Sungai Petani Proximity (Euclidean distance) to nearest federal route 1 Proximity (Euclidean distance) to nearest federal route 67
Neighbourhood variables	NV_RES NV_COM NV_IND	Proportion of residential land neighbouring cells Proportion of residential land neighbouring cells Proportion of residential land neighbouring cells
Zoning variables	ZV_RES ZV_COM ZV_IND	Cells fall on land zoned for residential use Cells fall on land zoned for commercial use Cells fall on land zoned for industrial use
Physical variables	PV_SLOPE PV_AVLAND	Slope steepness (degree) Amount of land available for new development (ha)

3.2. Geographically Weighted Regression Analysis

Geographically Weighted Regression (Fotheringham et al., 2002) is used to incorporate data in each point of observation into a regression model using a series of distance related weights. The relationship between urban built-up area and location characteristics for a particular point for example, is given a higher weight than for points far from that point. The GWR regression model for urban change (CHG_URB) in Sungai Petani is summarized in the following equation

$$CHG\_URB(ui, vi) = \beta 0(ui, vi) + \beta L(ui, vi)\, Li + \beta N(ui, vi)\, Ni + \beta Z(ui, vi)\, Zi + \beta P(ui, vi)\, Pi + ei$$

where β0 is the intercept term, βL, βN, βZ, βP are spatially varying coefficients of location (L), neighbourhood (N), zoning (Z) and physical (P) attributes respectively, and ei is an error term at point i, (ui, vi) represents the coordinates of the i'th point in space.

The GWR model is fitted to the data using the GWR 3.0 package which allows the use of a variety of calibration techniques to specify regression weights and to optimise bandwidth parameters. In this study, a fixed defined kernel with a bi-square function in which the bandwidth was determined by minimization of the Akaike Information Criterion (Fotheringham et al., 2002) is used. The reason is that the points of analysis used are in regular and equal sizes. Monte Carlo tests (Fotheringham et. al., 2002) were also done to determine the significance of the spatial variability in the local parameter estimates.

4. RESULTS AND DISCUSSION

4.1. Sungai Petani GWR Urban Built-up Model

Table 3 summarises the results of the global and GWR analysis of urban built-up areas of Sungai Petani. The Monte Carlo test calibrated for the GWR model found that 17 of the 20 explanatory variables displayed significant spatial non-stationarity. Furthermore, in the GWR model, the explanatory variables explain 63 percent of the variance, thus the GWR model provides better explanatory ability than the global model (adjusted R^2 = 0.29). Figures 3a to 3d show the spatial variations of the factors that influence change in urban built-up areas in Sungai Petani. The local R^2 of each individual GWR model (Fig 3a) ranges from 0.0 to 0.9, with a mean of 0.4. Only about 17.5% of the local R^2 values are lower than the global value, and approximately 27% are higher than 0.5. It can be inferred that the relationship between the selected factors and urban built-up is better captured by the GWR model in those regions. The growth of urban use in region with a low R^2 may be affected more by other factors not considered in this study, and possibly also edge effects which were not considered in this study.

Table 3: Summary Results of the Sungai Petani Global and GWR Urban built-up Models.

		Urban built-up Model Coefficients		
		GLOBAL		GWR
Variables		β	t	p-value [a]
	Intercept	41.57	4.26***	0.00***
Access and Proximity Variables (APV)				
	APV_RGC	-27.02	-1.71***	0.10***
	APV_PORT	22.00	3.97***	0.00***
	APV_APORT	4.69	0.31 n/s	0.21 n/s
	APV_IND	0.91	3.53***	0.00***
	APV_FDR67	-0.44	-1.91***	0.00***
	APV_INTC	-0.47	-2.15***	0.00***
	APV_SPT	-1.31	-5.25***	0.00***
	APV_EMP	-0.18	-0.86n/s	0.00***
	APV_FRD1	0.58	3.20***	0.00***
	APV_URB	-0.45	-4.51***	0.00***
	APV_RAIL	-0.02	-0.26 n/s	0.94 n/s
	APV_DCTR	-0.08	-0.67 n/s	0.00***
Neighbourhood Variables (NV)				
	NV_COM	-0.15	-2.24***	0.00***
	NV_RES	0.15	1.05 n/s	0.00***
	NV_IND	-0.16	-2.03***	0.00***
Zoning Variables (ZV)				
	ZV_COM	1.58	3.91***	0.87 n/s
	ZV_IND	2.66	7.10***	0.31 n/s
	ZV_RES	0.52	1.85***	0.00***
Physical Variables (PV)				
	PV_AVLAND	0.66	5.19***	0.00***
	PV_SLOPE	0.02	0.28 n/s	0.00***
N=		3702		
Adjusted R^2		0.29		0.63

*** = significant at 1% level
n/s = not significant
[a] Results of Monte Carlo test for spatial non-stationarity .(Fotheringham, 2002)

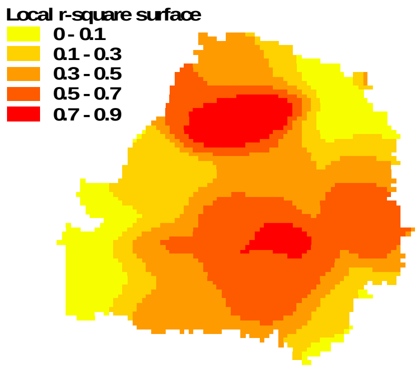

Figure 3a: spatial variation of local r-square

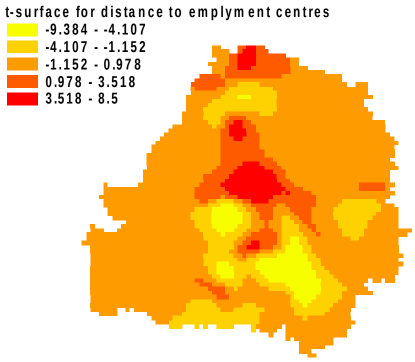

Figure 3b: t-surface for distance to employment centres parameter

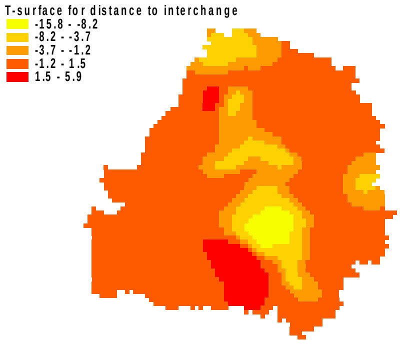

Figure 3c: t-surface for distance to Interchanges parameter

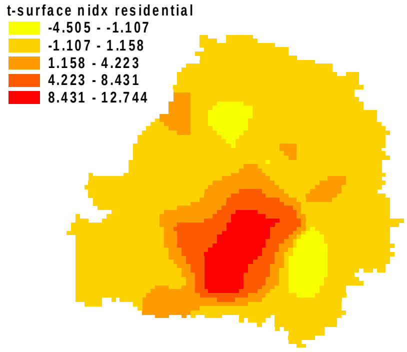

Figure 3d: t-surface for residential neighbourhood index

In GWR, F-test is also used to test whether spatial variation exists in the relationship under study (Brunsdon et al., 1996), specifically testing whether the GWR model offers an improvement over, and describes the relationship significantly better than, the ordinary global model. This concern is addressed via an ANOVA test implemented. For Sungai Petani GWR urban built-up model, the F-value is 8.2. High F-value suggesting that the GWR model has a significant improvement over the global model in determining the relationship between urban growth and the various determinant factors. In addition, the Akaike Information Criterion (AIC) of the GWR model (18747.4) is far less than the one of the global model (20571.1) indicating that GWR model performs better than the OLS model (refer Table 4).

Table 4: ANOVA test of the SP GWR over the OLS regression model

Source	SS	DF	MS	F
OLS Residuals	55316.1	26.0		
GWR Improvement	30064.9	464.46	4.73	
GWR Residuals	25251.2	3211.54	7.86	8.232

GWR Akaike Information Criterion: 18747.4 (OLS: 20571.1)

5. CONCLUSIONS

The above results indicate the spatial variations in the local parameter estimates of the 17 variables in influencing the urban change in the study area. Due to these variations, the relationships between the explanatory variables and the urban built-up area are better modelled in some parts of the study area than some other areas. The use of GWR approaches has increased the strength in the relationship especially in terms of the goodness-of-fit statistics (R^2) from 0.29 (OLS) to 0.63 (GWR). While the adjusted R^2 value of the global OLS is 0.2, the value for local GWR models ranges from 0.1 to 0.99 across the study area. The results are mapped in the GIS environment in order to visualize the spatial variation of the factors that significantly influenced urban development. They allow local land use decision makers to see clearly the results of past land use choices and to make appropriate decisions for future planning. GWR can be used as a tool to help determine problem areas so that proper planning can be focused on.

ACKNOWLEDGMENTS

The authors would like to thank the Universiti Sains Malaysia for funding this research. The authors would like to thank the anonymous referees for their comments and observations on draft of this paper.

REFERENCES

Bertaud, A. and Stephen M. (2003), The Spatial Distribution of Population in 48 World Cities: Implications for Economies in Transition. Online. The Centre for Urban Land economic research. University of Wisconsin. http://alain-bertaud.com/AB_Files/Spatia_Distribution_of_Pop_50_Cities.pdf

Bertaud, A. (2004), The spatial organization of cities: Deliberate outcome or unforeseen consequence? Online. The Centre for Urban Land economic research. University of Wisconsin. http://alain-bertaud.com/AB_Files/

Brunsdon, C et. al. (1996), Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis*, 28(4),281-298.

Chapin, F. S. and Kaiser, E. J. (1979), Urban Land Use Planning. Urbana, University of Illinois Press.

Foody, G.M. (2003), Geographical weighting as a further refinement to regression modelling: An example focused on the NDVI-rainfall relationship. *Remote Sensing of Environment*, 88 (3), 283-293.

Fotheringham, A.S. et. al. (2002), Geographically Weighted Regression: the Analysis of Spatially Varying Relationships. West Sussex, John Wiley and Sons Ltd., England.

Government of Malaysia (1992), *1991 Population and Housing Census of Malaysia.* Kuala Lumpur: Department of Statistics.

Government of Malaysia (1996), Sungai Petani Local Plan, Sungai Petani, Municipal Council of Sungai Petani

Government of Malaysia (2001), *2000 Population and Housing Census of Malaysia.* Kuala Lumpur: Department of Statistics.

Knox, P. (1993) The restless urban landscape, Englewood Cliffs, NJ: Prentice Hall.

Laffan, S.W. and Bickford, S. A. (2005) Using spatial randomisations to improve the utility of Geographically Weighted Regression model results. In Zerger, A. and Argent, R.M. (eds) MODSIM 2005 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December 2005, pp. 170-176. ISBN: 0-9758400-2-9.

Nakaya, T. (2001) Local spatial interaction modelling based on the Geographically Weighted Regression approach. GeoJournal, 347-358.

Osborne , P. E and Suarez-Seoane, S. (2002), Should data be partitioned spatially before building large-scale distribution models? *Ecological Modelling,* 157: 249–259.

Platt, R. V. (2004), Global and local analysis of fragmentation in a mountain region of Colorado. *Agriculture, Ecosystem and Environment*, 101, 207-218.

Ruslan, R**.** and Noresah, M. S. (2004) Modelling urban land use dynamic using GIS and Markov Chain Model. *Geoinformation Science Journal* 4 (1): pp. 72-81. ISSN 1551-9491.

Salleh, G. (2000), *Urbanization and Regional Development in Malaysia*, Utusan Publications & Distributors Sdn. Bhd., Kuala Lumpur.

Soia, E. (2000), Postmetropolis: Critical Studies of Cities and Regions. Oxford: Blackwell Publishers.

Zhang, L., Bi HuiQuan, Cheng, P. F., and Davis, C. J. (2004) Modeling spatial variation in tree diameter-height relationships. *Forest Ecology and Management*, 189, 317-329.