UNIVERSITI PUTRA MALAYSIA

PREDICTING PROPERTY RATING VALUES USING GEOGRAPHICALLY WEIGHTED REGRESSION.

SUBASHINI A/P VACLIVELOO

ITMA 2010 1

PREDICTING PROPERTY RATING VALUES USING GEOGRAPHICALLY WEIGHTED REGRESSION.

By

SUBASHINI A/P VACLIVELOO

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Master of Science

March 2010

Specially dedicated

to

Appa, Amma, Ganesh ne, Sunthari ni, Suresh ne, Vicky ni, Papathi ka,

Rajathi ka, Bala ma, Selvan, Mohan, Sharvinny papa, Dinesh,

Thaaranny papa, Jivethan, friends and also deeply devoted to Selva chitapa.

For their love, understanding, endless patience and encouragement

when it was most needed.

"IN FAMILY LIFE, LOVE IS THE OIL THAT EASES FRICTION, THE CEMENT THAT BINDS CLOSER TOGETHER, AND THE MUSIC THAT BRINGS HARMONY."

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Master of Science

PREDICTING PROPERTY RATING VALUES USING GEOGRAPHICALLY WEIGHTED REGRESSION.

By

SUBASHINI A/P VACLIVELOO

March 2010

Chairman: Taher Buyong, PhD

Institute: Institute of Advanced Technology (ITMA)

Currently, the market value for rating valuation applied in Malaysia is the single property valuation technique. This technique is not efficient enough, involving high costs and large labor force because rating involves valuation of large number of properties. Multiple Regression Analysis (MRA) was applied due to these weaknesses. However, the MRA fails to account for the spatial effects (spatial heterogeneity and spatial dependence) inherent in property data. In this study, the Geographically Weighted Regression (GWR) model is introduced as a new method to value rating properties. The GWR model is able to capture spatial heterogeneity by allowing different relationships to occur between variables at different points in space.

This study has two objectives. The first objective is to determine the attributes to be used for MRA and GWR model in this study. Data for this study were collected from two local authorities to represent rent and transaction data-based rating. Data for rent

was obtained from Majlis Perbandaran Kajang (MPKj) and data for transaction was obtained from Majlis Perbandaran Kulai (MPKu). Final attributes for rent–based rating area are land area, main floor area, ancillary floor area, type of ceiling, property position, property type, age of building, distance to centre business of district and neighborhood quality and the attributes for transacted-based rating area are land area, main floor area, additional floor area and floor finishing.

The second objective is to compare the performances of the GWR model with the MRA model in predicting rating values in the study areas. The result of R^2, Adjusted R^2, F-test and standard error of estimates proved that the GWR model provides better fitness compared to the MRA model. Residual analyses also reveal the same conclusion where residual for the GWR model is smaller in absolute values and probability distribution close to normal. The GWR model has also successfully captured spatial heterogeneity in almost all attributes. The prediction assessment of out-sample observations also revealed that the GWR model is able to produce better prediction. The ability of the GWR model to capture spatial effects is the main reason for this model to perform better; the GWR model is able to solve spatial heterogeneity problem explicitly and spatial dependence problem implicitly. Thus, the GWR which has been proven to be able to produce accurate prediction with small number of attributes should be used for rating valuation in Malaysia.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk Ijazah Sarjana Sains

MERAMAL NILAI HARTA TANAH KADARAN MENGUNAKAN "GEOGRAPHICALLY WEIGHTED REGRESSION".

Oleh

SUBASHINI VACLIVELOO

Mac 2010

Pengerusi: Taher Buyong, PhD

Institut: Institut Teknologi Maju (ITMA)

Penilaian kadaran pada masa sekarang menggunakan teknik penilaian tunggal. Teknik ini tidak efisyen, melibatkan kos yang tinggi dan memerlukan tenaga pekerja yang ramai menyebabkan ianya tidak sesuai digunakan untuk penilaian harta tanah yang banyak bagi maksud kadaran. Oleh kerana itu, analisis regrasi berganda (MRA) diaplikasi bagi mengatasi masalah-masalah ini. Walaubagaimanapun, analisis MRA gagal mengambilkira kesan *spatial* (keheterogenan dan kebergantungan) yang sedia wujud dalam data harta tanah. Dalam kajian ini, *Geographically Weighted Regression* (GWR) diperkenalkan sebagai satu kaedah baru untuk digunakan dalam penilaian kadaran. GWR mengambilkira masalah keheterogenan spatial dengan membolehkan pembolehubah-pembolehubah menghasilkan hubungan yang berbeza apabila lokasi berubah.

Kajian ini mengandungi dua objektif. Objektif pertama adalah untuk menentukan ciri-ciri harta tanah untuk digunakan bagi Model MRA dan Model GWR dikawasan kajian. Data untuk kajian ini diperolehi dari dua kerajaan tempatan untuk mewakili data sewa dan data transaksi. Data sewa diperolehi dari Majlis Perbandaran Kajang (MPKj) manakala data transaksi diperolehi dari Majlis Perbandaran Kulai (MPKu). Ciri-ciri harta tanah yang dikenalpasti bagi penilaian berasaskan sewa adalah luas tanah, luas bangunan utama, luas bangunan sokongan, jenis siling, kedudukan harta tanah, jenis harta tanah, umur bangunan, jarak dari bandar terdekat dan kualiti persekitaran (*neighborhood quality*) dan ciri-ciri bagi penilaian berasakan transaksi adalah luas tanah, luas bangunan utama, tambahan luas bangunan dan jenis lantai.

Objektif kedua adalah untuk membandingkan prestasi Model GWR dengan Model MRA dalam meramal nilai kadaran di kawasan kajian. Hasil R^2, *adjusted* R^2, ujian F dan *standard error of estimate* membuktikan bahawa model GWR lebih baik jika dibandingkan dengan model MRA. Kesimpulan yang sama digambarkan oleh analisis *residual* dimana nilai mutlak *residual* menjadi kecil dan plot taburan kebarangkalian menghampiri normal. Model GWR juga telah berjaya mengambilkira sifat heterogenan pada hampir kesemua ciri-ciri harta tanah. Hasil taksiran *out-sample* juga menyatakan GWR mampu menghasilkan taksiran yang lebih baik. Keupayaan GWR untuk mengambilkira kesan *spatial* merupakan faktor utama bagi model ini berupaya menghasilkan pretasi yang lebih baik; model GWR mampu mengatasi masalah heterogenan secara nyata dan masalah pergantungan secara tersirat. Oleh itu, model GWR yang telah dibuktikan bahawa berupaya menghasilkan tafsiran yang lebih tepat dengan hanya mengambilkira beberapa ciri-ciri harta tanah seharusnya digunakan dalam penilaian kadaran di Malaysia.

ACKNOWLEDGEMENTS

I would like to thank to God to permit HIS blessings on me and to give me strength to finish my master studies successfully at UPM.

I wish to express my profound gratitude and thanks to my supervisor Dr. Taher Buyong for his unlimited guidance, supervision and encouragement throughout the duration of the study. His precious advice has changed my perceptions in studies and also of the outside world. Many things I have learnt and only God can pay him for it. Gratitude is also extended to co-supervisor, Assoc. Prof. Dr. Ahmad Rodzi Mahmud for sharing their knowledge and guidance.

Sincere appreciation and gratitude is expressed to Mr. Mohd. Mahfuzan (Aput) and Mr. Andrew (Majlis Perbandaran Kajang), Mr. Hermi, Mrs. Roslina, Ms. Normalina and Mrs. Norita from administration for their excellent co-operation, support and assistance. Thanks to all my friends for their support and encouragement especially to Ms.Haemamalar, Ms.Renuga, Ms.Kavin, Ms. Risyalini, Mr. Suparamaniam, Mrs.Rohana, Mr.Logeswaran, Mr.Rohsidul, Mr.Mubbarak, Ms.Bawani, Mr.Sohel, Mr.Fahkri, Mr.Firdaus and K2 juniors. It is an honor to have met all of you and hope our lovely friendship will never end.

Finally, I am forever in dept to my parents, my entire family and my beloved Mohan for their prayers, moral support, understanding, endless encouragement, patience and love in molding me into who I am today. Thank you all and May God bless all of you with good health and great life.

I certify that an Examination Committee has met on 25th March 2010 to conduct the final examination of **Subashini a/p Vacliveloo** on her **Master of Science** thesis "**Predicting property values in rating using Geographically Weighted Regression**" in accordance with Universiti Pertanian Malaysia (Higher Degree) Act 1980 and Universiti Pertanian Malaysia (Higher Degree) Regulations 1981. The Committee recommends that the student be awarded the Master of Science.

Members of the Examination Committee were as follows:

Prof. Madya Dr. Ir. Nor Mariah Adam Ph.D, P. Eng
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

Prof. Madya Dr. Abdul Rashid Mohamed Shariff
Spatial and Numerical Lab (SNML)
Universiti Putra Malaysia (UPM)
(Internal Examiner I)

Dr. Zaiton Ali
Faculty of Economics and Management
Universiti Putra Malaysia (UPM)
(Internal Examiner II)

Prof. Ir. Dr. Miswan @ Abdul Hakim Mohammed
Faculty of Engineering and Science Geoinformation
Universiti Teknologi Malaysia (UTM)
(External Examiner)

PROF. DR. BUJANG KIM HUAT
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia
Date:

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Master of Science. The members of the Supervisory Committee were as follows:

Taher Buyong, PhD
Research Fellow
Institute of Advanced Technology (ITMA)
Universiti Putra Malaysia
(Chairman)

Ahmad Rodzi Mahmud, PhD
Associate Professor
Institute of Advanced Technology (ITMA)
Universiti Putra Malaysia
(Member)

HASANAH MOHD. GHAZALI, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 15th July 2010

DECLARATION

I declare that the thesis is my original work expect for quotations and citations which have been duly acknowledged. I also declare that it has not been previously, and is not concurrently, submitted for any other degree at Universiti Putra Malaysia or at any other institution.

SUBASHINI A/P VACLIVELOO

Date: 7th May 2010

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADD_F	Additional Floor Area
AFA	Ancillary Floor Area
AGE	Property Age
AIC	Akaike Information Criterion
ANN	Artificial Neural Network
ANOVA	Analysis of Variances
B_ROOM	Bathroom
BBB	Bandar Baru Bangi
BBM	Bandar Bukit Mahkota
CBD	Central Business District
CEIL	Ceiling
COD	Coefficient of Dispersion
COMC	Commercial Centre
CSM	Comparable Sales Method
CV	Cross-Validation
FENCE	Fence
FLOOR	Floor Finishing
GIS	Geographic Information System
GWR	Geographically Weighted Regression
IAAO	International Association of Assessing Officers
JPPH	Valuation and Property Services Department
KL	Kuala Lumpur

LA	Land Area
MAPE	Mean Absolute Percent Error
MFA	Main Floor Area
MPKj	Majlis Perbandaran Kajang
MPKu	Majlis Perbandaran Kulai
MRA	Multiple Regression Analysis
MWK	Moving Window Kriging
MWR	Moving Window Regression
NEIG_Q	Neighborhood Quality
nnb	nearest neighbor
OLS	Odinary Least Square
PM	Particulate Matter
POSI	Property Position
PRD	Price Related Differential
PUTRAJY	Putrajaya
RMSE	Root Mean Squared Error
ROOF	Roof
SHM	Spatial Hedonic Model
SPSS	Statistical Package for Social Science
sqm	Square meter
Unstd. Coe.	Unstandardized coefficients
VIF	Variance Inflation Factor

CHAPTER 1

PREAMBLE

1.1 Introduction

Generally, property taxes are one of the main sources of revenue for the government. In Malaysia, property taxes have a very significant contribution towards local authorities revenue, as they contribute more than 50% of the total revenue for local authorities (Hizam et al, 1999). The income from taxes is use to manage and develop the local authority territory. This includes aspects of physical development, environment, health, social services and economy. The activities comprise the planning of new growth areas, managing of the developed areas in terms of their compliance to established laws and regulations as well as ensuring of a clean, safe and orderly environment for the citizens.

1.2 Problem Statement

Currently, the valuation for rating purposes, applied the single property appraisal or “fee” appraisal technique to determine the market value of a property. There are three methods normally use to derive the market value for rating purposes, which are comparison method, cost method and investment method. However, the comparison method is the most common method among them (Arifian & Hasmah, 2001).

Through the comparison method, the market value for a property is accessed by comparing the property attributes with other property attributes surrounding it. Adjustments are made accordingly to the dissimilarity between properties attributes. Normally, adjustments are executed for the types of property, locations, transaction dates, tenures and physical factors. As such, the major problem in comparison method is the rationality in determining the adjustment amount among the differences that occur in properties attributes.

For the time being, the amount of adjustments in the comparison method depends on the property valuer's (valuation officer) assumptions. These assumptions are based on their opinions, knowledge and experiences. Therefore, this causes inaccurate and inefficient estimations due to various opinions between valuers regarding a property, which could lead to different estimations on the same property. As a result, the accuracy of the market value (fair market value) determined by using this method, is unreliable.

The valuation using the single valuation technique was found not suitable for rating purposes because the mass valuation rating requires hundreds of thousands of properties to be valued. Mass valuation processes also necessitate high costs and large labors force (Mustafa, 2004). As a result, most of the local authorities have failed to revaluate their properties for every five years as required by the law and the current property taxes do not point out the current tax rates (Marbeck, 1986).

Manual valuations are prone to mistakes which often occur in calculation parts, especially when it involves complicated calculations. The valuation for rating involves bundles of files as it involves a large number of properties. Therefore, the valuer sometimes fails to notice certain data which gives a significant impact to the property value. Additionally, updating the existing information is very complicated because of the difficulty faced in dealing with many bundles of files and papers.

Due to the weakness of this current manual valuation process, many researchers' have attempted to develop and use mass approach models since early 1920's for mass valuation. Most of the researchers adopted the Multiple Regression Analysis (MRA) as a mass approach. In Malaysia, the mass valuation model began to use as the only viable alternative in 1980's (Azahari, 1992). Mustafa (2004) has supported the MRA approach as a suitable method for mass appraisal. This approach is more systematic and more objective compared to the single valuation technique. The MRA approach also consumes less time and a computerized valuation was obtained using this model.

Unfortunately, the basic concept of the MRA (global model), produces "average" or "global" parameters estimates, which are applied equally over the whole region (Fotheringham & Charlton, 1998). This means only one regression is applied for the whole area and the whole area was treated equally same (Fotheringham et al., 2002). As a result, all local particularities are overlooked and the MRA method has not been able to take into account of spatial effects inherent in property data. According to Anselin (1988), spatial effects are further divided into two types: spatial dependence and spatial heterogeneity. Spatial dependence and spatial heterogeneity are the main reasons why spatial stationarity occurs. Spatial stationarity is a circumstance in which a simple global regression model cannot effectively clarify the relationship between various sets of variables over a geographical region (Fotheringham et al., 2002).

Spatial dependence or spatial autocorrelation is a circumstance where observations at one location depend on other observations at different locations (LeSage, 1999). In other word, observations in dataset lack of the independencies among it which tends to show similar patterns within certain geographical area (Anselin, 1988). Spatial dependence problem violates the basic assumptions of MRA models where all errors should be randomly distributed and not related to each other. But, spatially errors were related to each other and show the dependences among it.

Meanwhile, spatial heterogeneity refers to variation in relationship over space (LeSage, 1999). Different points in space hold different relationships. Anselin (1988) has pointed out spatial heterogeneity as a situation where functional form and parameters vary with location and are not homogeneous throughout the data set. This obviously shows that the MRA models have failed to meet other assumptions of the MRA which require a homogeneous area.

In addition, parameters being estimated are implicit to be immobile over space (Fotheringham & Charlton, 1998). However, in reality, especially in social science, the measurement of relationship depends on its location and mobile.

Alternatively, some researchers had given their opinion to model spatial effects within a global framework to avoid spatial non-stationarity. However, Brunsdon et al. (1999) and Fotheringham et al. (2000) have proven that if spatially varying relationships were modeled within a global framework such as the standard regression, then the error terms in global regression model would exhibit spatial autocorrelation.

Recently, several advanced mass regression methods are being studied. This study was based on the Geographically Weighted Regression (GWR) method because this method can capture spatial variation by calibrating a multiple regression model that allows different relationships between variables to exist at different points in space. The GWR model is also able to produce a local statistic that measures the relationship of a location within the study region. As the location changes, the statistic can take on different values (Fotheringham et al., 2002). The study issue now is, how far the GWR models can accurately predict property value for mass valuation compared to the MRA models.

1.3 Aim

The aim of this study is to determine the suitability of the GWR method to use in mass appraisal for rating in Malaysia. The suitability of this study is clarified in terms of how accurate this model can predict property price and how practical and easy it is to apply this model in Malaysia.

1.4 Objectives

The objectives of this study are as follows:-

i. To determine the attributes to be used for MRA and GWR model in this study.

ii. To compare the performances of the GWR model with the MRA model in predicting rating values in the study areas.

1.5 Scope and Limitation of the Study

This study was conducted within the Majlis Perbandaran Kajang (MPKj) and Majlis Perbandaran Kulai (MPKu) areas. These two areas are chosen to represent the basis of assessment rates. There are two basis of assessment which are annual value and improved value. So, MPKj area represents the annual values and MPKu area represents the improved value.

1.6 Significant of the Study

Through this study, a new method (the GWR method) was introduced to predict property value, particularly in mass appraisal. This method has the ability to estimate property values accurately compared to the current method use by the local authorities. As a result, fair and equitable taxes are able to determine and this study will be beneficial to the local authorities, property valuers, researchers, and tax payers.