

**MODEL GABUNGAN REGRESI BERPEMBERAT SECARA
GEOGRAFI : BERPENDEKATAN TEGUH**

NOR HASLIZA BT ASNGARI

**SARJANA SAINS
UNIVERSITI MALAYSIA TERENGGANU
2012**

**MODEL GABUNGAN REGRESI BERPEMBERAT SECARA GEOGRAFI: BERPENDEKATAN
TEGUH**

NOR HASLIZA BT. ASNGARI

**Tesis Ini Dikemukakan Sebagai Memenuhi Syarat Untuk Memperolehi Ijazah Sarjana
Sains di Fakulti Sains dan Teknologi
Universiti Malaysia Terengganu**

Januari 2012

Abstrak tesis yang dikemukakan kepada Senat Universiti Malaysia Terengganu sebagai memenuhi keperluan untuk ijazah Master Sains.

**MODEL GABUNGAN REGRESI BERPEMBERAT SECARA GEOGRAFI:
BERPENDEKATAN TEGUH**

NOR HASLIZA BT. ASNGARI

Januari 2012

Pengerusi : Muhamad Safiih bin Lola, Ph.D.

Ahli : Profesor Madya Mohd Lazim bin Abdullah, Ph.D.

Fakulti : Sains dan Teknologi

Kaedah Regresi Berpemberat Secara Geografi (RBSG) adalah kaedah penganggaran data berpengaruh secara geografi yang amat meluas digunakan terutamanya untuk menilai hubungan di antara pembolehubah yang berpengaruh secara geografi. Walau bagaimanapun, penganggaran ini tidak mengambil kira tentang kewujudan situasi data bercampur iaitu sebahagian pembolehubah dipengaruhi secara geografi dan selebihnya dipengaruhi secara global. Oleh sebab itu, satu kaedah penganggaran berdasarkan Gabungan Regresi Berpemberat Secara Geografi (GRBSG) telah diperkenalkan. Amnya, model GRBSG adalah gabungan di antara model regresi biasa dengan kaedah penganggaran kuasadua terkecil (KDT) dan model RBSG. Masalah utama penggunaan kaedah regresi adalah kewujudan data terpencil dalam data yang dicerap. Dengan itu, kaedah penganggaran KDT telah diperbaiki berdasarkan kaedah teguh. Dalam kajian ini, kaedah penganggar- M yang tidak mudah dipengaruhi oleh kehadiran data terpencil dikacukkan ke dalam model GRBSG iaitu terhadap penganggaran KDT. Penggabungan ini menghasilkan satu model teguh yang dikenali sebagai Gabungan Regresi Teguh Berpemberat Secara Geografi (GRTBSG). Untuk melihat keberkesanan penganggaran

model yang dicadangkan, satu kajian simulasi Monte Carlo telah dijalankan. Seterusnya, model yang dicadangkan akan dibandingkan terhadap model sedia ada berdasarkan kepada dua keadaan; pertama terhadap data bersih dan kedua terhadap data tercemar masing-masing dengan pencemaran 10%, 20% dan 50%. Selain itu, dua kelompok saiz data dipertimbangkan iaitu kelompok data bersaiz kecil, ($n < 30$) dan kelompok data bersaiz besar, ($n \geq 30$). Bagi data bersaiz kecil, tiga saiz dipertimbangkan iaitu $n = 9, n = 16$ dan $n = 25$. Manakala untuk data bersaiz besar pula, $n = 36, n = 64$ dan $n = 169$. Pemilihan saiz sampel yang berbeza adalah untuk melihat keupayaan menganggar GRTBSG terhadap saiz sampel. Hasil kajian simulasi menunjukkan bahawa kaedah GRTBSG yang dicadangkan adalah lebih konsisten dan efisien berbanding model GRBSG yang sedia ada. Ini dapat dibuktikan apabila nilai n meningkat (kes sampel bersaiz kecil dan besar), didapati bahawa min bagi nilai suaian yang diperolehi semakin menghampiri nilai sebenar. Manakala, nilai Ralat Piawai dan Punca Min Kuasadua Ralat masing-masing menjadi semakin kecil selari dengan peningkatan nilai n (saiz sampel).

Abstract of thesis presented to the Senate of University Malaysia Terengganu in fulfillment of the requirement for the degree of Master of Science.

MIXED GEOGRAPHICALLY WEIGHTED REGRESSION MODEL: ROBUST APPROACH

NOR HASLIZA BT. ASNGARI

January 2012

Chairperson : Muhamad Safiih bin Lola, Ph.D.

Member : Associate Professor Mohd Lazim bin Abdullah, Ph.D.

Faculty : Science and Technology

Geographically Weighted Regression (GWR) is widely used estimation method for spatial data and generally used to assess the relationship between variables that influence geographically. However, this estimation does not take into account the existence of mixed data situation, where some variables are locally affect and the rest is a globally effect. Therefore, an estimation method based on Mixed Geographically Weighted Regression was introduced. In general, MGWR model is a combination between the basic regression model with ordinary least square (OLS) estimation method and MGWR model. The main problem that often encountered in the application of the regression is the presence of outliers. With the OLS estimation method has been improved based on robust method. In this study, *M*-estimator method which not easily influenced by the presence of outliers was hybrid into the MGWR model, which is in OLS estimation part. This consolidation has resulted in a robust model known as the Robust Mixed Geographically Weighted Regression (RMGWR). To examine the effectiveness of the proposed estimation model, a Monte Carlo simulation study was carried out. Next, the proposed model compared against the existing models based on

two conditions: first of clean data and the second one is a contaminated data with 10%, 20% and 50% contaminations respectively. In addition, two groups of sample size are considered, small-sized data groups, ($n < 30$) and large-sized data groups, ($n \geq 30$). For small-sized data, three different size are considered which is $n = 9$, $n = 16$ and $n = 25$. While for large data, the $n = 36$, $n = 64$ and $n = 169$ are considered. The variation of sample size selection is to look at the estimation ability of RMGWR against sample size. From simulation study findings, it shows that the proposed method, RMGWR are consistent as well efficient than the existing model, MGWR. This can be proved when n is increased (the case of small and large sample size), it was found that the obtained estimation values are approximating the true. Where, the value of SE and RMSE becomes smaller as n increased, respectively.