

PEMODELAN MENGGUNAKAN METODE SPASIAL DURBIN MODEL UNTUK DATA ANGKA PUTUS SEKOLAH USIA PENDIDIKAN DASAR

Luh Putu Safitri Pratiwi¹, Shofwan Hanief², I Ketut Putu Suniantara³

Program Studi Sistem Informasi, STMIK STIKOM Bali^{1,2,3}

e-mail: safitri.pratiwi@yahoo.com

Abstrak

Masalah anak yang putus sekolah perlu mendapatkan perhatian karena salah satu indikator yang berguna untuk mengukur kemajuan sumber daya manusia pada bidang pendidikan. Untuk menekan laju pertumbuhan jumlah anak putus sekolah tersebut dapat dilakukan dengan cara mengetahui faktor-faktor yang berpengaruh terhadap jumlah anak putus sekolah dan berpotensi dalam meningkatkan laju pertumbuhan anak yang putus sekolah. Pemodelan yang menggunakan pengaruh daerah (area) disebut pemodelan spasial. Ciri dari pemodelan spasial adalah adanya matriks pembobot yang merupakan penanda adanya hubungan antara suatu wilayah dengan wilayah lain. Salah satu model spasial yaitu Spasial Durbin Model (SDM). Penelitian ini bertujuan untuk mengetahui jumlah anak putus sekolah di wilayah Bali dengan menggunakan metode SDM dan ingin mengetahui faktor-faktor yang mempengaruhi anak putus sekolah di wilayah Bali. Model yang didapat ialah pemodelan SDM menghasilkan nilai AICc yang lebih kecil dibandingkan pemodelan dengan OLS. Tidak adanya lag variabel independen yang signifikan menyebabkan hasil estimasi parameter menggunakan SDM menjadi tidak signifikan akan tetapi pada identifikasi nilai Moran's I mengidentifikasi adanya dependensi spasial pada variabel independen yang artinya ada kemiripan sifat untuk lokasi yang saling berdekatan.

Kata kunci: Regresi, Spasial, OLS, SDM

I. PENDAHULUAN

Sumber daya manusia bermutu yang merupakan produk pendidikan adalah kunci keberhasilan pembangunan suatu negara. Undang-Undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional, menjamin hak atas "pendidikan dasar" bagi warga negara Indonesia yang berusia 7-15 tahun. Salah satu upaya untuk meningkatkan taraf pendidikan penduduk Indonesia dengan menyelesaikan Program Wajib Belajar Pendidikan Dasar 9 Tahun. Berlawanan dengan harapan yang ada, kenyataannya angka putus sekolah di jenjang pendidikan dasar hingga saat ini masih tinggi. Putus sekolah merupakan salah satu indikator yang berguna untuk mengukur kemajuan sumber daya manusia pada bidang pendidikan pada tiap wilayah.

Pemodelan yang menggunakan pengaruh daerah (area) disebut pemodelan spasial. Ciri dari

pemodelan spasial adalah adanya matriks pembobot yang merupakan penanda adanya hubungan antara suatu wilayah dengan wilayah lain. Metode spasial merupakan metode untuk mendapatkan informasi pengamatan yang dipengaruhi efek ruang atau lokasi. Pengaruh efek ruang tersebut disajikan dalam bentuk koordinat lokasi (longitude, latitude) atau pembobotan. [1] menyatakan bahwa pemodelan spasial dilakukan dengan proses autoregressive, yaitu ditunjukkan dengan hubungan ketergantungan sekumpulan pengamatan atau lokasi.

Menurut [2] salah satu model spasial autoregressive adalah model spasial Mixed Regressive-Autoregressive memiliki bentuk persamaan seperti Spatial Autoregressive Model (SAR), dengan pengaruh spasial lag hanya pada variabel dependen. Hubungan spasial setiap

pengamatan tersebut dinyatakan dalam matriks pembobot (W). Spatial Durbin Model (SDM) merupakan salah satu jenis model tersebut, dikembangkan karena dalam beberapa kasus hubungan dependensi dalam spasial tidak hanya terjadi pada variabel dependen, tetapi juga pada variabel independen sehingga ditambahkan spasial lag WX pada model.

Kasus putus sekolah sesuai digunakan dalam metode spasial dikarenakan untuk melihat jumlah anak putus sekolah dalam suatu wilayah dipengaruhi oleh daerah lainnya yang berdekatan. Tujuan dari penelitian ini adalah untuk mendeskripsikan jumlah anak putus sekolah di Bali serta faktor – faktor yang diduga mempengaruhinya. serta memodelkan jumlah anak putus sekolah di wilayah dengan menggunakan metode SDM.

II. TINJAUAN PUSTAKA

A. Penelitian yang Terkait

Beberapa penelitian tentang regresi spasial yaitu:

- [3] dalam penelitiannya yang berjudul “Pemodelan Angka Putus Sekolah Usia Wajib Belajar Menggunakan Metode Regresi Spasial di Jawa Timur” menerapkan aspek spasial untuk melihat apakah keragaman karakteristik antar kabupaten/kota di Jawa Timur menentukan kualitas pendidikan pada daerah tersebut, selain itu Fitriani juga ingin mengetahui faktor-faktor yang berpengaruh secara nyata terhadap angka putus sekolah di Jawa Timur. Hasil akhir dari penelitian ini menunjukkan bahwa APS (Angka Putus Sekolah) tingkat SD tidak terdapat dependensi spasial. Hal ini menandakan bahwa pada tingkat SD di setiap kabupaten/kota di Jawa Timur sudah berjalan mandiri dan tidak memiliki ketergantungan dengan daerah yang lainnya. Faktor-faktor yang diperoleh, yaitu rasio guru dan murid, PDRB per kapita, dan rasio penduduk tamatan maksimal SD.
- [4] dalam penelitiannya yang berjudul “Pemodelan Jumlah Anak Putus Sekolah Di Provinsi Bali Dengan Pendekatan “*Semi-Parametric Geographically Weighted Poisson Regression*” mengatakan bahwa faktor-faktor yang mempengaruhi jumlah putus sekolah pada tiap wilayah berbeda-beda tergantung pada karakteristik dari

masing-masing daerah tersebut, sehingga diperlukan suatu analisis statistika yang memperhitungkan faktor geografis pada tiap pengamatannya. Berdasarkan hasil analisis dan pembahasan yang dilakukan, Astari menyimpulkan bahwa faktor spasial/lokasi tidak berpengaruh terhadap data jumlah anak putus sekolah usia pendidikan dasar di Provinsi Bali tahun 2010.

B. Analisis Regresi

Analisis Regresi merupakan metode yang digunakan untuk mengetahui pola hubungan antara variabel respon dan variabel bebas. Model persamaan regresi sebagai berikut.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon \quad (1)$$

dengan :

a) Pengujian Serentak

Hipotesis yang digunakan untuk pengujian serentak adalah sebagai berikut.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$$

H_1 : minimal ada satu $\beta_k \neq 0$; $k=1,2,\dots,p$
statistik uji untuk menghitung F_{hit} dengan rumus (2) dibawah ini:

$$F_{hit} = \frac{MSR}{MSE} \quad (2)$$

b) Pengujian Parsial

Pengujian parameter regresi secara parsial dilakukan untuk mengetahui parameter yang signifikan memengaruhi variabel dependen. Hipotesis untuk pengujianya sebagai berikut.

$$H_0: \beta_k = 0$$

$$H_1: \beta_k \neq 0; k=1,2,\dots,p$$

statistik uji sesuai rumus (3) di bawah ini:

$$t_{hit} = \frac{\hat{\beta}_k}{S(\hat{\beta}_k)} \quad (3)$$

C. Regresi Spasial

Hukum pertama tentang geografi menyatakan bahwa segala sesuatu saling berhubungan satu dengan yang lainnya, tetapi sesuatu yang dekat mempunyai pengaruh lebih besar daripada sesuatu yang jauh [2]. Hukum tersebut merupakan dasar pengkajian permasalahan berdasarkan efek lokasi atau metode spasial. Metode spasial merupakan metode untuk mendapatkan informasi pengamatan yang dipengaruhi efek ruang atau lokasi. Pengaruh efek ruang tersebut disajikan dalam bentuk koordinat lokasi (*longitude, latitude*) atau pembobotan.

Pemodelan spasial mempunyai dua tipe data yaitu dapat dibedakan menjadi pemodelan dengan pendekatan titik dan area. Pengujian efek spasial dilakukan dengan uji heterogenitas dan dependensi spasial. Penyelesaian jika ada efek heterogenitas adalah dengan menggunakan pendekatan titik. Jenis pendekatan titik antara lain *Geographically Weighted Regression* (GWR), *Geographically Weighted Poisson Regression* (GWPR), *Space-Time Autoregressive* (STAR), dan *Generalized Space-Time Autoregressive* (GSTAR). Penyelesaian jika ada efek dependensi spasial adalah dengan menggunakan pendekatan area. Pendekatan area antara lain *Spatial Autoregressive Model* (SAR), *Spatial Error Model* (SERM), *Spatial Durbin Model* (SDM), *Conditional Autoregressive Model* (CAR), dan *Spatial Autoregressive Moving Average* (SARMA). SDM merupakan kasus khusus dari model SAR dengan menambahkan pengaruh lag pada variabel independen, sehingga pembobotan dilakukan pada variabel independen maupun dependen.

D. Model Regresi Spasial

Menurut [2] bahwa model regresi yang melibatkan pengaruh spasial disebut model regresi spasial. Salah satu pengaruh spasial yaitu autokorelasi spasial, adanya unsur autokorelasi spasial mengakibatkan terbentuknya parameter spasial autoregresif dan *moving average*, sehingga terbentuk proses spasial sebagai berikut:

$$\begin{aligned} y &= \rho W_1 y + X\beta + u \\ u_t &= \lambda W_2 u_{t-1} + \varepsilon \\ \varepsilon &\sim N(0, \sigma^2 I) \end{aligned}$$

sehingga model umum yang terbentuk adalah:

$$y = \rho W_1 y + X\beta + \lambda W_2 u + \varepsilon$$

dengan y merupakan vektor variabel dependen yang berukuran $n \times 1$, ρ adalah parameter koefisien spasial lag variabel dependen, λ adalah parameter koefisien spasial lag *error*, u merupakan vektor *error* pada persamaan (2.6) yang berukuran $n \times 1$, dan ε merupakan vektor *error* pada persamaan (2.7) berukuran $n \times 1$, yang berdistribusi normal dengan mean nol dan varians $\sigma^2 I$. Sementara itu, W_1 dan W_2 merupakan matriks pembobot dengan ukuran $n \times n$ dimana $W_1 = W_2 = W$. Vektor parameter koefisien regresi dilambangkan dengan β yang berukuran $(p+1) \times 1$. Matriks X merupakan matriks variabel independen yang berukuran $n \times (p+1)$.

merupakan matriks identitas yang berukuran $n \times n$, n banyak amatan atau lokasi ($i = 1, 2, 3, \dots, n$), dan p banyak variabel independen.

E. Spatial Durbin Model (SDM)

Model spasial *Mixed Regressive-Autoregressive* memiliki bentuk persamaan seperti *Spatial Autoregressive model* (SAR) dengan hanya ada pengaruh spasial lag pada variabel dependen. Namun *Spatial Durbin Model* (SDM) merupakan kasus khusus dari model SAR dengan menambahkan pengaruh lag pada variabel independen. Oleh karena itu perlu ditambahkan spasial lag pada model sehingga pembobotan dilakukan pada variabel independen maupun dependen. Bentuk model SDM adalah sebagai berikut [2]:

$$\begin{aligned} y_i &= \rho \sum_{j=1}^n w_{ij} y_j + \beta_0 + (\beta_{11} x_{1i} + \beta_{12} x_{2i} + \dots + \beta_{1k} x_{ki} + \dots + \beta_{1l} x_{li}) + \\ &\quad \left(\beta_{21} \sum_{j=1}^n w_{ij} x_{1j} + \beta_{22} \sum_{j=1}^n w_{ij} x_{2j} + \dots + \beta_{2k} \sum_{j=1}^n w_{ij} x_{kj} + \dots + \beta_{2l} \sum_{j=1}^n w_{ij} x_{lj} \right) + \varepsilon_i \quad (4) \\ y_i &= \rho \sum_{j=1}^n w_{ij} y_j + \beta_0 + \sum_{k=1}^l \beta_{1k} x_{ki} + \sum_{k=1}^l \beta_{2k} \sum_{j=1}^n w_{ij} x_{kj} + \varepsilon_i \end{aligned}$$

dengan k menyatakan banyaknya variabel dan l menyatakan amatan ke- i .

Bentuk model SDM (4) dapat dinyatakan dalam matriks dengan vektor parameter koefisien spasial lag variabel independen dinyatakan dalam β_2 .

$$y = \rho W_1 y + \beta_0 + X\beta_1 + WX\beta_2 + \varepsilon$$

atau

$$y = \rho W_1 y + Z\beta + \varepsilon$$

dengan, $Z = [1 \quad X \quad WX]$ $\beta = [\beta_0 \quad \beta_1 \quad \beta_2]^T$

F. Penduga Parameter Spatial Durbin Model (SDM)

Penduga parameter model SDM menggunakan *Maximum Likelihood Estimation* (MLE). Fungsi *likelihood* terbentuk melalui *error* (ε) sehingga menjadi persamaan (5).

$$y = \rho W_1 y + Z\beta + \varepsilon \quad (5)$$

$$\varepsilon = y - \rho W_1 y - Z\beta$$

$$\varepsilon = (I - \rho W_1)y - Z\beta$$

$$L(\sigma^2; \varepsilon) = \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} \exp \left(-\frac{1}{2\sigma^2} (\varepsilon^T \varepsilon) \right)$$

$$L(\rho, \beta, \sigma^2 | y) =$$

$$\left(\frac{1}{2\pi\sigma^2} \right)^{n/2} (J) \exp \left(-\frac{1}{2\sigma^2} (\varepsilon^T \varepsilon) \right) \quad (7)$$

$$J = \left| \frac{\partial \varepsilon}{\partial y} \right| = |I - \rho W_1|.$$

Fungsi *Jacobian* merupakan turunan terhadap y dari persamaan (5). Substitusikan persamaan (5) yang merupakan fungsi *Jacobian* ke persamaan (7), sehingga menghasilkan persamaan (8).

$$L(\rho, \beta, \sigma^2|y) = \left(\frac{1}{2\pi\sigma^2}\right)^{n/2} |I - \rho W_1| \exp\left(-\frac{1}{2\sigma^2} ((I - \rho W_1)y - Z\beta)^T ((I - \rho W_1)y - Z\beta)\right). \quad (8)$$

G. Pengujian Hipotesis Signifikansi Penduga Parameter

Pengujian terhadap parameter model dilakukan untuk mengetahui peranan variabel bebas dalam model. Uji Wald digunakan untuk menguji parameter β_j . Rumus untuk uji Wald berdasarkan hipotesis (Anselin,1998):

$$H_0: \hat{\beta}_j = 0$$

$$H_1: \hat{\beta}_j \neq 0$$

statistik uji yang digunakan adalah pada persamaan (2.29).

$$Wald = \left[\frac{\hat{\beta}_j^2}{var(\hat{\beta}_j)} \right]$$

dengan $\hat{\beta}_j$ menyatakan penduga parameter ke- j , dan $var(\hat{\beta}_j)$ menyatakan varian parameter ke- j . Kriteria pengambilan keputusan adalah mengikuti sebaran $\chi^2_{\alpha,1}$. H_0 ditolak jika nilai $Wald > \chi^2_{\alpha,1}$.

H. Uji Efek Spasial

[2] membedakan efek spasial menjadi dua bagian yaitu dependensi spasial dan heterogenitas spasial. Dependensi spasial ditunjukkan dengan kemiripan sifat untuk lokasi yang saling berdekatan, sedangkan heterogenitas spasial ditunjukkan oleh perbedaan sifat antara satu lokasi dengan lokasi lainnya.

1) Uji Dependensi Spasial

[2] menyatakan bahwa uji untuk mengetahui dependensi spasial di dalam *error* suatu model dengan menggunakan statistik Moran's I.

Hipotesis yang digunakan adalah:

$$H_0 : I_M = 0 \text{ (tidak ada dependensi antarlokasi)}$$

$$H_1 : I_M \neq 0 \text{ (ada dependensi antarlokasi)}$$

statistik uji [5] disajikan pada persamaan yaitu:

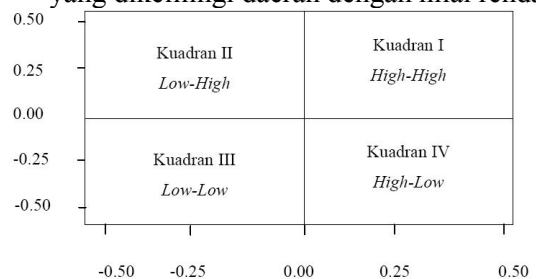
$$Z_{hitung} = \frac{I_M - I_{M0}}{\sqrt{var(I_M)}}$$

dengan x_i merupakan data ke- i ($i = 1, \dots, n$), x_j merupakan data ke- j ($j = 1, \dots, n$), \bar{X} merupakan rata-rata data, $var(I_M)$ merupakan varians Moran's I, $E(I_M)$ merupakan *expected value* Moran's I.

Jika $|z_{hit}| > z_{\alpha/2}$ maka keputusan diambil adalah menolak H_0 . Nilai dari indeks I_M adalah antara -1 dan 1, jika $I_M > I_{M0}$ maka data memiliki autokorelasi positif, jika $I_M < I_{M0}$ maka data memiliki autokorelasi negatif. Pola penyebaran antarlokasi dapat disajikan Moran's *Scatterplot* Gambar 1.

[6] menyebutkan bahwa Moran's *Scatterplot* dibagi atas empat kuadran yang cocok untuk empat pola kumpulan spasial lokal setiap daerah yang bertetangga. Berikut adalah penjelasan dari masing-masing kuadran [5].

1. Kuadran I (terletak di kanan atas) disebut *High-High* (HH) menunjukkan daerah yang memiliki pengamatan tinggi yang dikelilingi oleh daerah yang juga memiliki pengamatan tinggi untuk variabel yang dianalisis.
2. Kuadran II (terletak di kiri atas) disebut *Low-High* (LH) menunjukkan daerah dengan nilai rendah tapi dikelilingi daerah dengan nilai tinggi.
3. Kuadran III (terletak di kiri bawah) disebut *Low-Low* (LL) menunjukkan daerah dengan nilai pengamatan rendah dan dikelilingi oleh daerah yang juga mempunyai nilai pengamatan rendah.
4. Kuadran IV disebut *High-Low* (HL) menunjukkan daerah dengan nilai tinggi yang dikelilingi daerah dengan nilai rendah.



Gambar 1 Moran's *Scatterplot*

2) Uji Heterogenitas Spasial

Uji heterogenitas spasial digunakan untuk menunjukkan adanya keragaman antarlokasi. Heterogenitas data secara spasial dapat diuji dengan menggunakan *Breusch-Pagan Test* [2] yang hipotesisnya sebagai berikut:

$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$ (homoskedastisitas)

H_1 : minimal ada satu $\sigma_i^2 \neq \sigma^2$ (heterokedastisitas)

nilai *Breusch-Pagan Test* (BP Test) adalah:

$$BP = (1/2)f^T Z(Z^T Z)^{-1} Z^T f \sim \chi_k^2 \quad (2.31)$$

dengan elemen vektor f adalah:

$$f_i = \left(\frac{e_i^2}{\sigma^2} - 1 \right)$$

dengan e_i merupakan *least squares residual* untuk observasi ke- i , Z merupakan matriks berukuran $n \times (k + 1)$ yang sudah dinormalstandarkan (z) untuk setiap observasi.

Tolak H_0 jika $BP > \chi_k^2$.

III. METODOLOGI PENELITIAN

A. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Dinas Kesehatan Provinsi Bali tahun 2016 dan data dari Survei Sosial Ekonomi Nasional (SUSENAS) Provinsi Bali tahun 2016.

B. Variabel Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder. Variabel yang digunakan dalam penelitian ini terdiri dari 1 variabel respon dan 4 variabel prediktor. Serta menggunakan variabel lintang selatan (u_i) dan letak bujur timur (v_i) tiap kecamatan di Provinsi Bali yang digunakan untuk penentuan pembobot pada model GWR. Variabel respon pada penelitian ini adalah jumlah anak putus sekolah usia pendidikan dasar pada tiap kecamatan di Provinsi Bali

Sedangkan variabel bebas yang digunakan terdiri dari rasio siswa terhadap sekolah ($X1$), rasio siswa terhadap guru ($X2$), jumlah kepala keluarga dengan pendidikan terakhir ayah SD atau SMP ($X3$), angka buta huruf ($X4$).

C. Alur Analisis

Adapun tahapan analisis data yang akan dilakukan dalam penelitian ini adalah sebagai berikut:

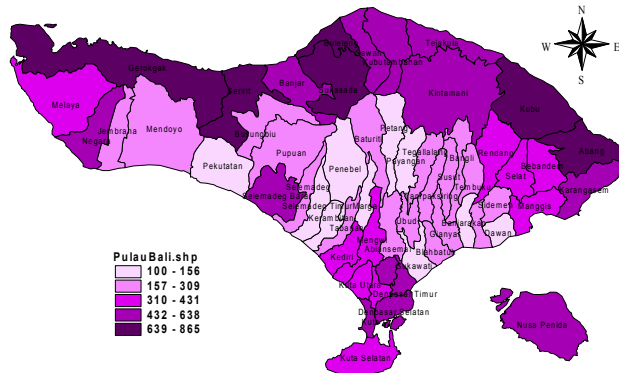
1. Melakukan deskripsi sebagai gambaran awal untuk mendeskriptifkan data jumlah anak putus sekolah
2. Pemodelan variabel persentase rumah tangga miskin di wilayah Kabupaten Gianyar dan faktor-faktor memengaruhinya. Langkah-langkahnya sebagai berikut.
 - a. Mengidentifikasi pola hubungan dengan menggunakan scatterplot dan analisis korelasi.

asi.

- b. Uji dependensi spasial dengan menggunakan statistik uji Moran's I pada setiap variabel kemudian membentuk Moran's scatterplot untuk mengetahui penyebaran antarlokasi.
3. Melakukan pemodelan Ordinary Least Square (OLS), terdiri dari:
 - a. Melakukan penduga parameter pada variabel independen terhadap variabel dependen serta melakukan uji hipotesis signifikansi parameter secara serentak dan parsial
 - b. Pemeriksaan asumsi residual, untuk mengetahui asumsi residual memenuhi identik, independen, dan berdistribusi normal serta memeriksa multikolinieritas antara variabel-variabelprediktor dengan menggunakan nilai VIF
 4. Memeriksa aspek spasial pada data yang digunakan
 5. Pemodelan Regresi Spasial jumlah anak putus sekolah digunakan untuk mengidentifikasi dependensi spasial yang lebih spesifik
 6. Melakukan pemodelan SDM, terdiri dari:
 - a. Melakukan penduga parameter untuk melihat dependensi lag pada variabel dependen maupun variabel independen.
 - b. Melakukan uji hipotesis signifikansi parameter
 - c. Memilih model terbaik
 7. Melakukan Pemodelan *Geographically Weighted Rgression* dengan runtutan sebagai berikut;
 - a. Mengestimasi nilai *bandwith* kernel dengan *cross validation* (CV) Menghitung nilai pembobot kernel W untuk masing-masing lokasi fungsi jarak *Gaussian* (*Gaussian Distance Function*).
 - b. Mengestimasi nilai koefisien regresi untuk masing-masing lokasi berdasarkan nilai *bandwith* dan pembobot kernel yang diperoleh pada langkah sebelumnya.
 - c. Membandingkan hasil analisis antara regresi klasik (OLS), regresi spasial dan GWR
 8. Mendapatkan model regresi terbaik untuk pemodelan jumlah anak putus sekolah dengan kriteria AIC.

IV. HASIL DAN PEMBAHASAN

A. Statistika Deskriptif



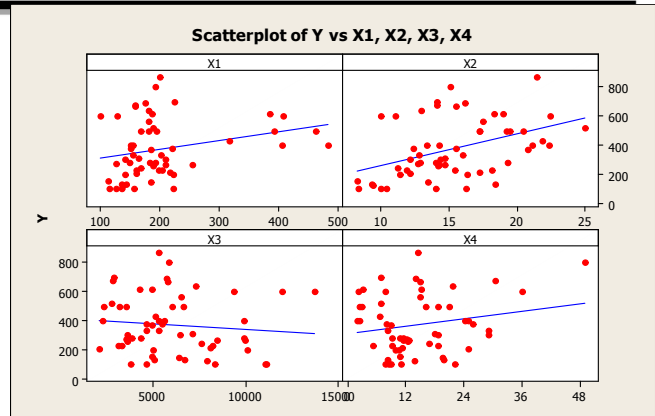
Gambar 2 Persebaran Jumlah Anak Putus Sekolah (Y) per Kecamatan di Provinsi Bali

Seperti tampak pada Gambar 2, dapat dilihat pola penyebaran jumlah anak putus sekolah di Provinsi Bali. Berdasarkan Gambar 2 diketahui warna lokasi yang semakin gelap, mengidentifikasi semakin tinggi jumlah anak putus sekolah di lokasi tersebut. Terlihat bahwa kecamatan dengan kategori jumlah anak putus sekolah sangat tinggi (639—865) terdapat di Kecamatan Gerogak, Seririt, Buleleng, Sukasada (Kabupaten Buleleng), Kecamatan Kubu dan Abang (Kabupaten Karangasem).

Kategori jumlah angka putus sekolah sangat rendah (100—156) terdapat di Kecamatan Pekutatan (Kabupaten Jembrana), Selamadeg Timur, Kerambitan, Penebel, Marga (Kabupaten Tabanan), Kecamatan Petang (Kabupaten Badung), Payangan, Sukawati (Kabupaten Gianyar), Kecamatan Banjarangkan, Dawan (Kabupaten Bangli).

B. Identifikasi Pola Hubungan Antar Variabel

Sebelum melakukan pemodelan *Ordinary Least Square* (OLS) dan pemodelan *Spatial Durbin*, maka dilakukan identifikasi pola hubungan antara variabel kemiskinan dan faktor-faktor yang memengaruhidengan menggunakan *Scatterplot*, dapat dilihat pada Gambar 3.



Gambar 3 Pola Hubungan antara Jumlah Anak Putus Sekolah dan Faktor yang Memengaruhinya

Seperti tampak pada Gambar 4.2 serta hasil pengujian korelasi dengan taraf signifikansi 5% yang dapat dilihat pada Lampiran 1, menunjukkan bahwa variabel independen yang berkorelasi positif terhadap variabel dependen yaitu X_1 , X_2 , dan X_4 , ini berarti semakin tinggi nilai variabel independen yang berpengaruh maka semakin tinggi angka putus sekolah dan begitu pula sebaliknya.

C. Nilai Moran's I

Uji dependensi spasial dilakukan untuk mengidentifikasi apakah ada hubungan antarlokasi terhadap masing-masing variabel dengan menggunakan Moran's I. Pengambilan keputusan dilakukan jika $|Z_{hitung}| > Z_{\alpha/2}$ maka tolak H_0 atau terdapat dependensi antarlokasi

Tabel 1 Pengujian Moran's I

Kode	Variabel	Moran's I	$ Z_{hitung} $	p-value
Y	Jumlah anak putus sekolah usia pendidikan dasar pada tiap kecamatan di Provinsi Bali	0,352	1,302	0,193
X_1	Rasio siswa terhadap sekolah pada tiap kecamatan di Provinsi Bali	0,747	1,166	0,243
X_2	Rasio siswa terhadap guru pada tiap kecamatan di Provinsi Bali	0,642	2,711	0,007*

X_3	Jumlah kepala keluarga dengan pendidikan terakhir ayah SD atau SMP pada tiap kecamatan di Provinsi Bali	0,173	-0,245	0,806
X_4	Angka Buta Huruf usia 15 tahun ke atas (ABH 15+) pada tiap kecamatan di Provinsi Bali	0,621	3,457	0,0005*

Total 221959
0 56

2) Uji Parsial

Untuk mengetahui parameter mana yang berpengaruh signifikan secara global dilakukan pengujian signifikansi parameter secara parsial dengan hipotesis sebagai berikut.

$$H_0 : \beta_k = 0$$

$$H_1 : \beta_k \neq 0; k = 1, 2, \dots, p$$

Hasil uji parsial dapat dilihat pada Tabel 4.4, dengan menggunakan metode OLS pengujian p-value dengan tingkat signifikansi 5% diperoleh variabel yang signifikan yaitu Rasio siswa terhadap guru pada tiap kecamatan di Provinsi Bali (X_2) dan Angka Buta Huruf usia 15 tahun ke atas (ABH 15+) (X_4)

Sumber: data diolah 2013, Ket: *) signifikan pada $\alpha = 5\%$; $Z_{0,025} = 1,96$

Berdasarkan hasil pengujian $|Z_{hitung}| > Z_{\alpha/2}$ serta pengujian p-value dengan tingkat signifikansi 5% seperti pada Tabel 4.1, maka pengambilan keputusan adalah tolak H_0 atau terdapat dependensi setiap desa pada dua variabel, yaitu Rasio siswa terhadap guru pada tiap kecamatan di Provinsi Bali (X_2) dan Angka Buta Huruf usia 15 tahun ke atas (ABH 15+) pada tiap kecamatan di Provinsi Bali (X_4).

Pada Tabel semua variabel memiliki nilai Moran's I yang lebih besar dari $I_{MO} = -\frac{1}{n-1} = -0,019$ menunjukkan pola data mengelompok sehingga setiap desa memiliki karakteristik yang sama.

D. Pemodelan Ordinary Least Square (OLS)

1) Pengujian Serentak

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{minimal ada satu } \beta_k \neq 0; k = 1, 2, \dots, p$$

Berdasarkan Tabel 4.2 menghasilkan nilai $F_{hitung} = 6,91 > F_{0,05;4;52} = 2,55$ sehingga keputusan diambil adalah tolak H_0 . Hal ini berarti bahwa variabel independen secara serentak berpengaruh terhadap variabel dependen.

Tabel 2 Analisis Varian Model Regresi Linier Berganda

Sumber Variansi	Jumlah Kuadrat	Derajat Bebas	Kuadrat Tengah	F	P-value
Regresi	770272	4	192568	6,91	0,000
Error	1449318	52	27872		

Tabel 3 Penduga Parameter dengan Metode OLS

Parameter	Penduga/Estimasi	$ t_{hitung} $	p-value	VIF
β_0	-194.8	-1.48	0,145	
β_1	0.3715	1.03	0,309	1,9
β_2	25.058	3.32	0,002	1,8
β_3	-0.00492	-0.58	0,562	1,0
β_4	9.830	3.52	0.001	1.3
R-square				34.7%
F_{hitung}				6.91
AICc				749.94
				1

Sumber: data diolah 2013, Ket: *) signifikan pada $\alpha = 5\%$; $t_{0,025;59} = 2,001$; $F_{0,05;10;59} = 2,00$

Oleh karena ada dua variabel yang signifikan pada uji parsial dilakukan penduga parameter ulang pada kedua variabel yang signifikan tersebut untuk dibandingkan dengan metode SDM, dapat dilihat pada Tabel 6

Tabel 6 Penduga Parameter dengan Metode OLS

Parameter	Penduga/Estimasi	$ t_{hitung} $	p-value	VIF
β_0	-209.6	-1,83	0,073	
β_2	29.763	4.87	0,000	1,2
β_4	9.005	3.45	0,001	1,2
R-square				33.1%
F_{hitung}				13.35
				747.3
				42

AICc

maka model terbaik dengan OLS adalah:

$$\hat{y} = -209,6 + 29,763X_2 - 9,005X_4$$

Intepretasi dari model tersebut yaitu apabila rasio siswa terhadap guru berkurang sebanyak satu satuan maka jumlah anak putus sekolah akan cenderung bertambah sebanyak 29,763 kali dengan asumsi variabel yang lain konstan, apabila Angka Buta Huruf usia 15 tahun ke atas berkurang sebanyak satu satuan maka akan cenderung mengurangi jumlah anak putus sekolah sebanyak 9,005 kali dengan asumsi variabel yang lain konstan.

Setelah mendapatkan model OLS selanjutnya dilakukan pengujian asumsi *residual*.

D. Perbandingan Titik Knot Optimal dengan metode GCV dan CV

1) Pengujian Asumsi Residual

Homogenitas

Uji homogenitas dilakukan untuk mengetahui apakah *residual* mempunyai nilai yang sama (homokesdasitas). Pengujian asumsi residual identik pada model regresi dilakukan dengan uji Glejser dengan hipotesis sebagai berikut:

$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2$ (kesamaan varians residual/ homoskedastisitas)

$H_1 : \sigma_i^2 \neq \sigma_j^2$ (heterokedastisitas)

Tabel 5 Uji Glejser

Predictor	Coef	T	Sig
Constant	-4,949	-0,061	
X ₁	-0,285	-1,276	0,208
X ₂	6,238	1,339	0,186
X ₃	0,013	2,567	0,013*
X ₄	0,416	0,241	0,810

Uji Glejser ini diperoleh dengan meregresikan nilai absolut residual dari model awal dengan semua variabel prediktor yang digunakan. Hasil pengujian pada Tabel 4.5 menunjukkan bahwa terdapat satu variabel (X₃) yang nyata berpengaruh pada taraf signifikansi $\alpha=0,05$ sehingga dapat disimpulkan bahwa asumsi varians residual homogen (identik) tidak terpenuhi.

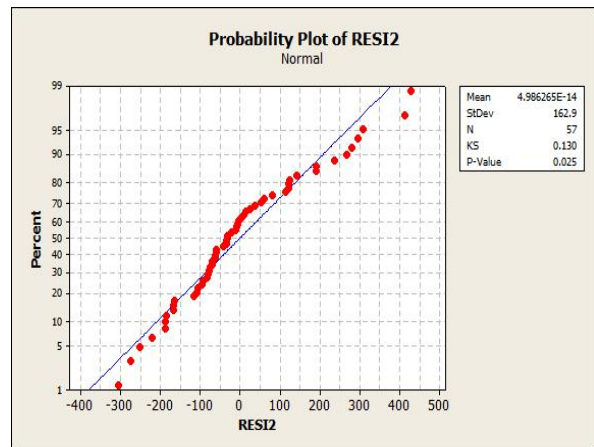
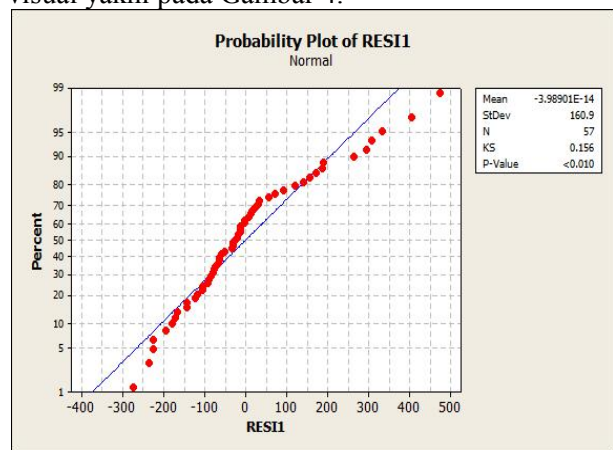
Normalitas

Pengujian asumsi residual berdistribusi normal model regresi dilakukan dengan uji Kolmogorov-Smirnov dengan hipotesis sebagai berikut:

H_0 : residual berdistribusi normal

H_1 : residual tidak berdistribusi normal

Berdasarkan hasil pengujian dengan *software* Minitab diperoleh nilai statistik uji Kolmogorov-Smirnov untuk semua variabel sebesar 0.156 dengan *p-value* sebesar <0,010. Sehingga, H_0 ditolak, dengan kata lain residual tidak berdistribusi normal. Sedangkan untuk dua variable yang signifikan nilai statistik uji Kolmogorov-Smirnov sebesar 0.130 dengan *p-value* sebesar 0.025 Sehingga, H_0 ditolak, dengan kata lain residual tidak berdistribusi normal. Pengujian asumsi residual berdistribusi normal dengan uji Kolmogorov-Smirnov menggunakan *software* Minitab dapat dilihat tampilan secara visual yakni pada Gambar 4.



Gambar 4 Uji Normalitas Residual

Uji Autokorelasi

Pengujian independensi residual dapat dilakukan dengan menggunakan plot ACF (*Autocorrelation Function*) residual dan menggunakan uji Durbin Watson

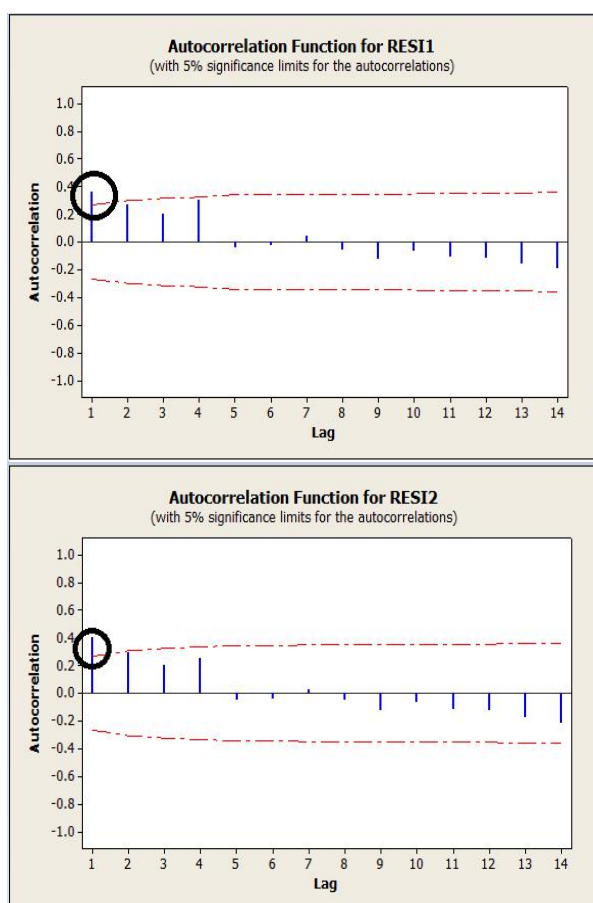
Pengujian asumsi residual independen model regresi dilakukan dengan uji Durbin-Watson dengan hipotesis sebagai berikut:

H_0 : $\rho = 0$ (tidak terdapat kasus autokorelasi)

H_1 : $\rho \neq 0$ (terdapat kasus autokorelasi)

Nilai DW 1.24783, nilai ini akan kita bandingkan dengan nilai tabel signifikansi 5%, jumlah sampel 57 (n) dan jumlah variabel independen 4 (K=4), maka diperoleh nilai du 1.7253. Nilai DW 1.24783 lebih kecil dari batas atas (du) yakni 1.7253 dan kurang dari (4-du) $4 - 1.7253 = 2,27$ dapat disimpulkan bahwa terdapat autokorelasi (tidak independen).

Uji independensi juga bisa menggunakan plot ACF (*Autocorrelation Function*) residual seperti tampak pada gambar 5



Gambar 5 Uji Autokorelasi Residual

Hasil dari uji asumsi *residual* untuk semua variabel dan dua variabel yang signifikan untuk mengetahui ada tidaknya autokorelasi pada *residual*, dapat dilihat melalui plot *Autocorrelation Function* (ACF) pada kedua *residual*. Plot tersebut menunjukkan bahwa *residual* tidak independen karena terdapat nilai lag *residual* yang keluar dari garis pada plot *Autocorrelation Function* (ACF).

Non-Multikolinieritas

Pengujian asumsi multikolinieritas bertujuan untuk mengetahui apakah terdapat korelasi antar variabel prediktor. Uji asumsi multikolinieritas dapat dilihat pada Tabel 4, untuk semua variabel dan untuk dua variabel yang signifikan menunjukkan hasil dari uji tersebut telah terpenuhi, ditunjukkan oleh nilai *Variance Inflation Factors* (VIF). Batasan VIF = 4, menandai adanya kemungkinan permasalahan multikolinieritas. Pada VIF = 10 atau lebih, multikolinieritas dinyatakan sangat parah dan membahayakan (*harmful*).

2) Memeriksa efek spasial pada data

[2] membedakan efek spasial menjadi dua bagian yaitu dependensi spasial dan heterogenitas spasial. Dependensi spasial ditunjukkan dengan kemiripan sifat untuk lokasi yang saling berdekatan, sedangkan heterogenitas spasial ditunjukkan oleh perbedaan sifat antara satu lokasi dengan lokasi lainnya.

Uji Dependensi Spasial

[2] menyatakan bahwa uji untuk mengetahui dependensi spasial di dalam *error* suatu model dengan menggunakan statistik Moran's I.

Hipotesis yang digunakan adalah:

H_0 : $I_M = 0$ (tidak ada dependensi antarlokasi)

H_1 : $I_M \neq 0$ (ada dependensi antarlokasi)

Nilai dari Moran's I yaitu: 2.6174080

Nilai P-value untuk Moran's I, yaitu : 0.0088601

Dengan tingkat kesalahan α sebesar 5% maka keputusan dari pengujian Moran's I ini adalah tolak H_0 , maka dapat disimpulkan bahwa terdapat dependensi spasial antar lokasi untuk jumlah anak putus sekolah di tiap Kecamatan di Provinsi Bali

Uji Heterogenitas Spasial

Uji heterogenitas Spasial menggunakan metode uji Breusch-Pagan [2]. Hipotesisnya adalah:

H_0 : $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$ (kesamaan varian atau dikenal *homoskedastisitas*)

H_1 : minimal ada satu $\sigma_i^2 \neq \sigma^2$ (*heterokedastisitas*)

Nilai dari BP-test yaitu: 12.52307

Nilai P-value untuk BP-test, yaitu : 0.0138573

Dengan tingkat kesalahan α sebesar 5% maka keputusan dari pengujian Breusch Pagan ini adalah tolak H_0 , maka dapat disimpulkan bahwa terdapat heterogenitas spasial untuk jumlah anak putus sekolah di tiap Kecamatan di Provinsi Bali.

Uji asumsi

E. Spasial Durbin Model (SDM)

Tabel 7 Penduga Parameter dengan Metode SDM

Parameter	Variabel	Penduga/Estimasi	Wald
β_0	intersep	-0.0027	0.0007
β_{12}	Rasio siswa terhadap guru	0.6282	31.929*
β_{14}	Angka Buta Huruf usia 15 tahun ke atas (ABH 15+)	0.4051	13.613*
β_{22}	Rasio siswa terhadap guru	-0.1131	3.598
β_{24}	Angka Buta Huruf usia 15 tahun ke atas (ABH 15+)	0.0390	0.435
ρ	Jumlah anak putus sekolah usia pendidikan dasar	0.0066	0.0245
R-square AICc			42,15 275,4899

Sumber: data diolah 2013

(Ket: *) $\chi^2_{0,05;1} = 3,841$

Model yang terbentuk dari SDM adalah:

$$\hat{y}_i = 0,0066 \sum_{j=1}^n W_{ij}y_j - 0,0027 + 0,6282X_{2i} + 0,4051X_{4i} - 0,1131 \sum_{j=1}^n W_{ij}y_{2j} + 0,039 \sum_{j=1}^n W_{ij}y_{4j}$$

Model SDM tersebut dapat diinterpretasikan:

- 1) 0,0066 menjelaskan jumlah anak putus sekolah akan mengalami kenaikan sebesar 0,0066 apabila pada kecamatan tetangga naik sebesar satu satuan, apabila faktor lain dianggap konstan
- 2) Intersep 0,0027 artinya jumlah anak putus sekolah sebesar 0,0027, apabila tidak terdapat faktor-faktor yang memengaruhi jumlah anak putus sekolah seperti Rasio siswa terhadap guru dan Angka Buta Huruf usia 15 tahun ke atas (ABH 15+) (rasio siswa terhadap guru = angka buta huruf usia 15 tahun ke atas (ABH 15+) = 0).

- 3) 0,6282 menjelaskan jika rasio siswa terhadap guru di suatu kecamatan naik satu satuan, maka menaikkan jumlah anak putus sekolah di kecamatan tersebut sebesar 0,6282 apabila faktor lain dianggap konstan
- 4) 0,4051 menjelaskan jika angka buta huruf usia 15 tahun ke atas (abh 15+) di suatu kecamatan naik satu satuan, maka menaikkan jumlah anak putus sekolah di kecamatan tersebut sebesar 0,4051 apabila faktor lain dianggap konstan
- 5) 0,1131 menjelaskan jumlah anak putus sekolah di suatu desa akan mengalami penurunan sebesar 0,1131 jika rasio siswa terhadap guru dari kecamatan tetangga naik satu satuan apabila faktor lain dianggap konstan
- 6) 0,039 menjelaskan jumlah anak putus sekolah di suatu kecamatan akan mengalami kenaikan sebesar 0,039 jika rasio siswa terhadap guru dari kecamatan tetangga naik satu satuan apabila faktor lain dianggap konstan,

Nilai ρ yang tidak signifikan yaitu sebesar 0,0066 menunjukkan tidak adanya dependensi spasial lag pada variable jumlah anak putus sekolah, dan $\sum_{j=1}^n W_{ij}y_j$ menunjukkan adanya pengaruh letak kecamatan yang berdekatan (i) dengan kecamatan yang diamati (j) terhadap jumlah anak putus sekolah

Berdasarkan analisis pada kedua model yaitu model OLS dan Model SDM dapat disimpulkan bahwa pemodelan SDM menghasilkan nilai R^2 yang lebih besar dan nilai AICc yang lebih kecil dibandingkan pemodelan dengan OLS. Tidak adanya lag variabel independen yang signifikan menyebabkan hasil estimasi parameter menggunakan SDM menjadi tidak signifikan akan tetapi pada identifikasi nilai Moran's I mengidentifikasi adanya dependensi spasial pada variabel independen yang artinya ada kemiripan sifat untuk lokasi yang saling berdekatan.

V. KESIMPULAN DAN SARAN

Berdasarkan analisis pada kedua model yaitu model OLS dan SDM, model SDM dapat disimpulkan bahwa pemodelan SDM menghasilkan nilai AICc yang lebih kecil dibandingkan pemodelan dengan OLS. Tidak adanya lag variabel independen yang signifikan menyebabkan hasil estimasi

parameter menggunakan SDM menjadi tidak signifikan akan tetapi pada identifikasi nilai Moran's I mengidentifikasi adanya dependensi spasial pada variabel independen yang artinya ada kemiripan sifat untuk lokasi yang saling berdekatan.

Tabel 8 Perbandingan Metode OLS dan Metode SDM

Model	AICc	Rsq
OLS	747,342	33,1%
SDM	275,4899	42,15%

Saran yang dapat diberikan untuk penelitian selanjutnya adalah dengan menggunakan jenis pembobot selain jenis *row standardization*, kode biner, dan *Contiguity* untuk mendapatkan pemodelan yang lebih baik. Salah satu jenis pembobot tersebut adalah dengan menggunakan matriks pembobot berdasarkan *Distance Band*.

VI. DAFTAR PUSTAKA

- [1] Lesage dan Pace, R. K. 2009. *Introduction to Spatial Econometrics*. CRC Press. Boca Ration
- [2] Anselin, L. 1988. *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers. Netherlands.
- [3] Fitroni, B. N. (2013). *Pemodelan Angka Putus Sekolah Usia Wajib Belajar Menggunakan Metode Regresi Spasial di Jawa Timur, Skripsi*, Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember, Surabaya.
- [4] Astari, G. R., & Srinadi, G. M. (2013). Pemodelan Jumlah Anak Putus Sekolah di Provinsi Bali dengan Pendekatan Semi-Parametric Geographically Weighted Poisson Regression. *E-Jurnal Matematika Universitas Udayana Bali*, Vol.2 No.3, 29-34.
- [5] Lee J. and Wong S.W.D. 2001. *Statistical Analysis with Arcview GIS*, John Willey & Sons, Inc., United Stated of America
- [6] Perobelli, F. S. dan Haddad, E. 2003. *An Exploratory Spatial Data Analysis of Brazilian Interregional Trade (1985-1996)*. <http://www.uiuc.edu/unit/real.pdf>. Diakses pada tanggal 27 Februari 2013.