

***NONPARAMETRIC GEOGRAPHICALLY WEIGHTED
REGRESSION DENGAN PENDEKATAN SPLINE TRUNCATED
PADA DATA PRODUKTIVITAS TANAMAN PADI***

Nur Fitri¹, Sifriyani^{1*}, Desi Yuniarti¹

¹Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas
Mulawarman, Indonesia

Corresponding author: sifriyani@fmipa.unmul.ac.id

Abstrak. Penelitian ini menggunakan metode *Nonparametric Geographically Weighted Regression* dengan pendekatan *Spline Truncated* (NGWR-TS). Metode tersebut merupakan pengembangan dari regresi nonparametrik untuk data spasial dengan estimator parameter bersifat lokal untuk setiap lokasi pengamatan. Spline truncated merupakan model polinomial tersegmen yang memberikan sifat fleksibilitas terhadap karakteristik lokal dari data. Pendekatan spline truncated digunakan untuk menyelesaikan permasalahan analisis spasial yang bentuk kurva regresinya tidak diketahui. Data yang digunakan adalah produktivitas tanaman padi di 56 Kabupaten/Kota di Kalimantan pada Tahun 2017. Fungsi pembobot yang digunakan adalah Kernel Gaussian dan Kernel Bisquare. Dalam penentuan besarnya pembobot, digunakan metode *Generalised Cross Validation* (GCV). Hasil dari penelitian ini adalah model *Nonparametric Geographically Weighted Regression* dengan pendekatan spline truncated yang terbaik dengan nilai orde $m = 1$ dan titik knot $h = 1$ serta memiliki nilai R-Squared sebesar 90,015. Hal ini mengindikasikan bahwa model NGWR-TS mampu menjelaskan pengaruh variabel prediktor terhadap variabel respon produktivitas tanaman padi sebesar 90,015%. Variabel prediktor yang berpengaruh untuk setiap lokasi berbeda-beda, sehingga dapat dikelompokkan menjadi 9 kelompok berdasarkan variabel prediktor yang berpengaruh.

Kata Kunci: *generalised cross validation, nonparametric geographically weighted regression, pembobot kernel gaussian, pembobot kernel bisquare, produktivitas tanaman padi, spline truncated, titik knot.*

1 PENDAHULUAN

Produktivitas tanaman padi di Kalimantan, rata-rata mengalami penurunan mulai dari tahun 2014 hingga tahun 2016, namun pada tahun 2017 jumlah produktivitas padi mengalami kenaikan. Badan Pusat Statistik mencatat bahwa terdapat tiga Provinsi di Kalimantan yang mengalami penurunan terus menerus mulai dari tahun 2014 hingga tahun 2016, yaitu pada Provinsi Kalimantan Barat mengalami penurunan sebesar 2,86 Ku/Ha dari 30 Ku/Ha pada tahun 2014 menjadi 27,49 Ku/Ha pada tahun 2016, kemudian bertambah menjadi 27,54 Ku/Ha pada tahun 2017, Provinsi Kalimantan Timur mengalami penurunan 4,55 Ku/Ha dari 42,55 Ku/Ha pada tahun 2014 menjadi 38 Ku/Ha pada tahun 2016, kemudian pada tahun 2017 bertambah menjadi 42,39 Ku/Ha, dan Provinsi Kalimantan Utara mengalami penurunan 9,27 Ku/Ha dari 36,05 Ku/Ha pada tahun 2014 menjadi 26,75 Ku/Ha pada tahun 2016, kemudian pada tahun 2017 bertambah menjadi 31,62 Ku/Ha. Jumlah produktivitas tanaman padi di Provinsi Kalimantan Tengah mengalami kenaikan 0,5 Ku/Ha dari 34,57 Ku/Ha pada tahun 2014 menjadi 35,07 Ku/Ha pada tahun 2015 kemudian turun menjadi 29,01 Ku/Ha pada tahun 2016, setelah itu pada tahun 2017 bertambah menjadi 31,51 Ku/Ha. Jumlah produktivitas tanaman padi di Provinsi Kalimantan Selatan mengalami penurunan sebesar 0,18 Ku/Ha dari 42,05 Ku/Ha pada tahun 2014 menjadi 41,87 Ku/Ha, kemudian mengalami peningkatan terus menerus dari tahun 2016 hingga tahun 2017 menjadi 43,02 Ku/Ha [1][2][3][4][5][6][7][8][9][10].

Jumlah produktivitas tanaman padi di Kalimantan mengalami ketidakstabilan setiap tahunnya, sehingga diperlukan suatu solusi untuk mengatasi masalah tersebut. Salah satu upaya evaluasi yang dapat dilakukan adalah dengan mencari faktor-faktor apa saja yang mempengaruhi jumlah produktivitas tanaman padi.

Penelitian ini menggunakan jumlah produktivitas tanaman padi sebagai variabel respon dan beberapa variabel prediktor yang diduga mempunyai pengaruh terhadap jumlah produktivitas tanaman padi. Hubungan yang terbentuk antara jumlah produktivitas tanaman padi dengan masing-masing variabel prediktor diduga tidak mengikuti pola tertentu dan terdapat pola yang berubah-ubah. Perubahan ini dikarenakan setiap wilayah memiliki karakteristik yang berbeda-beda. Oleh karena itu, regresi *Nonparametric Geographically Weighted Regression* dengan pendekatan spline *truncated* (NGWR-TS) diduga dapat digunakan dalam studi kasus ini.

Berdasarkan latar belakang yang telah dipaparkan, mendorong peneliti untuk meneliti tentang *Nonparametric Geographically Weighted Regression* dengan pendekatan spline *truncated*, yang diaplikasikan pada data produktivitas padi di Kalimantan.

2 SUMBER DATA DAN METODOLOGI

2.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Badan Pusat Statistik [1][2][3][4][5][6][7][8][9][10]. Data penelitian ini terdiri dari produktivitas tanaman padi di 56 Kabupaten/Kota Tahun 2017 di

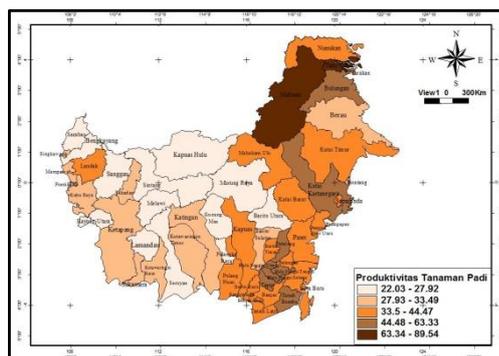
Kalimantan sebagai variabel respon dan jumlah produksi padi (x_1), Laju Pertumbuhan Ekonomi (LPE) (x_2), Produk Domestik Regiona Bruto (PDRB) (x_3), luas lahan sawah (x_4), jumlah tenaga kerja di bidang pertanian (x_5) serta luas panen lahan sawah (x_6) sebagai variabel prediktor.

2.2 Metode Penelitian

Metode penelitian yang digunakan adalah *Nonparametric Geographically Weighted Regression* dengan pendekatan *spline truncated* (NGWR-TS). Langkah analisis yang dilakukan terdapat empat tahap. Tahap-tahap yang dilakukan adalah estimasi model [11], pengujian hipotesis kesesuaian model [12], pengujian hipotesis simultan [13] dan pengujian signifikansi parameter secara parsial [14].

2.3 Eksplorasi Data

Rata-rata produktivitas tanaman padi Tahun 2017 di Kalimantan adalah sebesar 37,19 Ku/Ha, dimana produktivitas tanaman padi tertinggi berada di Kabupaten Malinau Provinsi Kalimantan Utara sebesar 89,54 Ku/Ha dan produktivitas tanaman padi terendah berada di Kota Palangka Raya Provinsi Kalimantan Tengah. Penyebaran dan pemetaan produktivitas tanaman padi di Kalimantan ditunjukkan pada Gambar 1.



Gambar 1: Pemetaan Produktivitas Padi Tahun 2017 di Kalimantan

Berdasarkan Gambar 1 produktivitas padi terbesar berada di sekitar 63,34 Ku/Ha sampai dengan 89,54 Ku/Ha, dimana Kabupaten/Kota yang termasuk dalam kelompok ini adalah Kabupaten Malinau dan Kabupaten Tana Tidung.

3 HASIL DAN PEMBAHASAN

a. Heterogenitas Spasial dan Pembobot Geografis

Heterogenitas spasial disebabkan karena adanya perbedaan karakteristik antar titik lokasi pengamatan, sehingga dibutuhkan pembobot spasial. Pembobot spasial yang terbaik dibutuhkan nilai *bandwidth* yang optimum. Pada penelitian ini digunakan *Generalised Cross Validation* (GCV) untuk menentukan *bandwidth* yang optimum. Berikut hasil pengujian heterogenitas spasial dan pemilihan *bandwidth* terbaik.

Tabel 1: *Breusch-Pagan Test* dan *Bandwidth Optimum*

Uji	<i>Breusch-Pagan</i>	db	<i>P-Value</i>	Keputusan
Heterogenitas Spasial	141,871	12	0,000	Tolak H_0
	Fungsi Pembobot	Nilai <i>Bandwidth</i>	<i>Generalised Cross Validation (GCV)</i>	
Pemilihan Pembobot	Kernel <i>Gaussian</i>	1,482	162,699	
	Kernel <i>Bisquare</i>	6,144	199,496	

Berdasarkan Tabel 1 dapat dilihat bahwa nilai *Breusch-Pagan* (BP)=141,871 > $\chi^2_{(0,05;12)} = 21,026$ maka dapat disimpulkan terdapat heterogenitas spasial, dengan nilai *bandwidth* optimum 1,482 menggunakan fungsi pembobot Kernel *Gaussian*. Hal ini berdasarkan nilai GCV terkecil.

b. Uji Hipotesis Kesesuaian Model

Uji hipotesis untuk kesesuaian model antara NGWR-TS dengan regresi nonparametrik spline *truncated* adalah [12]

$$H_0 : \beta_{pk}(u_i, v_i) = \beta_{pk} \text{ dan } \delta_{p,m+h}(u_i, v_i) = \delta_{p,m+h}, p = 1, 2, \dots, 6; k = 1; h = 1; i = 1, 2, \dots, 56$$

$$H_1 : \text{Minimal ada satu } \beta_{pk}(u_i, v_i) \neq \beta_{pk} \text{ atau } \delta_{p,m+h}(u_i, v_i) \neq \delta_{p,m+h},$$

Statistik *test* untuk menguji hipotesis H_0 lawan H_1 di atas adalah sebagai berikut

$$V = \frac{\mathbf{y}^T \tilde{\mathbf{S}} \mathbf{y}}{\frac{n - lm - 1}{\mathbf{y}^T D(u_i, v_i) \mathbf{y}} \text{tr}((\mathbf{I} - \xi)^T (\mathbf{I} - \xi))} \quad (1)$$

Hasil perhitungan didapatkan nilai $F_{hitung}(V) = 6,986 > F_{(0,05;34;29)} = 1,832$, maka dapat disimpulkan bahwa terdapat perbedaan yang signifikan antara model NGWR-TS dengan regresi nonparametrik spline *truncated*.

c. Metode *Nonparametric Geographically Weighted Regression* dengan Pendekatan *Spline Truncated*

Langkah selanjutnya adalah pemilihan titik knot yang optimum dengan menggunakan metode GCV, dimana semakin kecil nilai GCV maka semakin

optimum titik knot yang dipilih. Nilai GCV, *R-Squared*, AIC dan RMSE untuk model NGWR-TS ditunjukkan pada Tabel 2.

Tabel 2: Pemilihan Titik Knot

No.	x_1	x_2	x_3	x_4	x_5	x_6	GCV
1	4.026,27	0,7531	1.189.139	1.255,24	2.749	1.047,923	102,372*
2	11.818,81	0,8993	2.562.353	3.725,72	5.109	3.071,769	117,288
3	15.715,08	0,9724	4.748.960	4.960,96	6.289	4.083,692	137,3403

Metode	Titik Knot	<i>R-Squared</i>	AIC	RMSE
NGWR -TS Orde $m = 1$	1	90,015*	7,971*	2,738*
	2	69,486	4,518	6,492
	3	71,817	4,723	6,167

Berdasarkan Tabel 2 dapat diketahui bahwa titik-titik knot yang optimum memiliki nilai GCV 102,372, dimana nilai GCV yang didapatkan merupakan yang terkecil. Model dengan satu titik knot memiliki nilai *R-Squared* terbesar, sehingga model NGWR-TS dengan satu titik knot lebih tepat digunakan, dimana nilai-nilai titik knot optimumnya adalah sebagai berikut:

$$K_{11} = 4.026, 27; K_{21} = 0, 7531; K_{31} = 1.189.139; K_{41} = 1.255, 24; K_{51} = 2.749; K_{61} = 1.047, 923$$

d. Uji Signifikansi Parameter Simultan

Hipotesis untuk pengujian parameter model NGWR-TS secara simultan adalah sebagai berikut [13]

$$H_0 : \beta_{pk}(u_i, v_i) = 0 \text{ dan } \delta_{p,m+h}(u_i, v_i) = 0$$

$$H_1 : \text{Minimal ada satu } \beta_{pk}(u_i, v_i) \neq 0 \text{ atau } \delta_{p,m+h}(u_i, v_i) \neq 0$$

Statistik *test* untuk menguji hipotesis H_0 lawan H_1 di atas adalah sebagai berikut

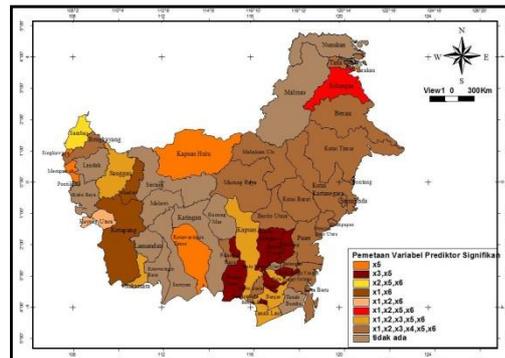
$$V^* = \frac{\left(\frac{\mathbf{y}^T M(u_i, v_i) \mathbf{y}}{\text{tr}((\mathbf{I} - \mathbf{B}_a)^T (\mathbf{I} - \mathbf{B}_a))} \right)}{\left(\frac{\mathbf{y}^T D(u_i, v_i) \mathbf{y}}{\text{tr}((\mathbf{I} - \xi)^T (\mathbf{I} - \xi))} \right)} \quad (2)$$

Hasil statistik hitung didapatkan nilai $F_{hitung}(V^*) = 5,028 > F_{(0,05;52;29)} = 1,77$, maka diputuskan tolak H_0 dan dapat disimpulkan bahwa terdapat parameter yang

berpengaruh signifikan terhadap produktivitas padi dalam model *Nonparametric Geographically Weighted Regression* dengan pendekatan *spline truncated*.

e. Uji Signifikansi Parameter Parsial

Hasil perhitungan dari pengujian signifikansi parameter secara parsial menunjukkan bahwa variabel prediktor yang berpengaruh berbeda-beda untuk setiap area. Hal ini mengakibatkan terdapat 9 kelompok yang terbentuk. Pemetaan Kabupaten/Kota berdasarkan variabel prediktor yang berpengaruh ditunjukkan pada Gambar 2.



Gambar 2: Pemetaan Kabupaten/Kota berdasarkan Variabel Prediktor yang Signifikan

Pembagian 9 kelompok yang terbentuk berdasarkan variabel prediktor yang signifikan dapat dijelaskan sebagai berikut:

1. Produktivitas tanaman padi di Kabupaten Paser, Kabupaten Kutai Barat, Kabupaten Kutai Kartanegara, Kabupaten Kutai Timur, Kabupaten Berau, Kabupaten Penajam Paser Utara, Kabupaten Mahakam Ulu, Kota Balikpapan, Kabupaten Tana Tidung, Kabupaten Kota Baru, Kabupaten Barito Kuala, Kabupaten Balangan, Kabupaten Barito Utara, Kabupaten Murung Raya dan Kota Palangka Raya dipengaruhi oleh besar produksi padi, LPE, PDRB, luas lahan sawah, jumlah tenaga kerja di bidang pertanian dan luas panen padi.
2. Produktivitas tanaman padi di Kota Samarinda, Kota Bontang, Kabupaten Bengkayang, Kabupaten Sanggau, Kota Pontianak, Kota Tarakan, Kabupaten Tanah Laut, Kabupaten Banjar, Kabupaten Hulu Sungai Selatan, Kota Banjarmasin, Kabupaten Kapuas dan Kabupaten Sukamara dipengaruhi oleh besar produksi padi, LPE, PDRB, jumlah tenaga kerja di bidang pertanian dan luas panen padi.
3. Produktivitas tanaman padi di Kabupaten Bulungan dipengaruhi oleh besar produksi padi, LPE, jumlah tenaga kerja di bidang pertanian dan luas panen padi.
4. Produktivitas tanaman padi di Kabupaten Kayong Utara dipengaruhi oleh besar produksi padi, LPE dan luas panen padi.
5. Produktivitas tanaman padi di Kabupaten Sambas dipengaruhi oleh LPE, jumlah tenaga kerja di bidang pertanian dan luas panen padi.
6. Produktivitas tanaman padi di Kabupaten Ketapang dan Kabupaten Sekadau dipengaruhi oleh besar produksi padi dan luas panen padi.
7. Produktivitas tanaman padi di Kabupaten Tapin, Kabupaten Hulu Sungai Tengah, Kabupaten Tabalong, Kota Banjarbaru, Kabupaten Barito Selatan,

Kabupaten Pulang Pisau dan Kabupaten Barito Timur dipengaruhi oleh PDRB dan jumlah tenaga kerja di bidang pertanian.

8. Produktivitas tanaman padi di Kabupaten Mempawah, Kabupaten Kapuas Hulu, Kota Singkawang dan Kabupaten Kotawaringin Timur dipengaruhi oleh jumlah tenaga kerja di bidang pertanian.
9. Produktivitas tanaman padi di Kabupaten Landak, Kabupaten Sintang, Kabupaten Melawi, Kabupaten Kubu Raya, Kabupaten Malinau, Kabupaten Nunukan, Kabupaten Hulu Sungai Utara, Kabupaten Tanah Bumbu, Kabupaten Kotawaringin Barat, Kabupaten Lamandau, Kabupaten Seruyan, Kabupaten Katingan dan Kabupaten Gunung Mas tidak dipengaruhi oleh besar produksi padi, LPE, PDRB, luas lahan sawah, jumlah tenaga kerja di bidang pertanian maupun luas panen padi.

4 KESIMPULAN

Model NGWR-TS dengan pendekatan spline *truncated* yang terbaik pada data produktivitas tanaman padi di Kalimantan Tahun 2017 adalah orde $m = 1$ dan titik knot $h = 1$. Nilai $R^2 = 90,015$, hal ini mengindikasikan bahwa model mampu menjelaskan pengaruh variabel prediktor terhadap variabel respon produktivitas tanaman padi sebesar 90,015%. Hasil pengujian parsial dari estimasi parameter mengelompokkan Kabupaten/Kota menjadi 9 grup berdasarkan variabel prediktor yang berpengaruh terhadap produktivitas tanaman padi di Kalimantan Tahun 2017.

DAFTAR PUSTAKA

- [1] Badan Pusat Statistik Provinsi Kalimantan Barat. *Hasil Survei Pertanian antar Sensus SUTAS 2018 Provinsi Kalimantan Barat*. Badan Pusat Statistik: Kalimantan Barat. (2018)
- [2] Badan Pusat Statistik Provinsi Kalimantan Selatan. *Hasil Survei Pertanian antar Sensus SUTAS 2018 Provinsi Kalimantan Selatan*. Badan Pusat Statistik: Kalimantan Selatan. (2018)
- [3] Badan Pusat Statistik Provinsi Kalimantan Utara. *Hasil Survei Pertanian antar Sensus SUTAS 2018 Provinsi Kalimantan Utara*. Badan Pusat Statistik: Kalimantan Utara. (2018)
- [4] Badan Pusat Statistik Provinsi Kalimantan Tengah. *Hasil Survei Pertanian antar Sensus SUTAS 2018 Provinsi Kalimantan Tengah*. Badan Pusat Statistik: Kalimantan Tengah. (2018)
- [5] Badan Pusat Statistik Provinsi Kalimantan Timur. *Hasil Survei Pertanian antar Sensus SUTAS 2018 Provinsi Kalimantan Timur*. Badan Pusat Statistik: Kalimantan Timur. (2018)
- [6] Badan Pusat Statistik Provinsi Kalimantan Barat. *Provinsi Kalimantan Barat Dalam Angka 2018*. Badan Pusat Statistik: Kalimantan Barat. (2018)
- [7] Badan Pusat Statistik Provinsi Kalimantan Selatan. *Provinsi Kalimantan Selatan Dalam Angka 2018*. Badan Pusat Statistik: Kalimantan Selatan. (2018)

- [8] Badan Pusat Statistik Provinsi Kalimantan Utara. *Provinsi Kalimantan Utara Dalam Angka 2018*. Badan Pusat Statistik: Kalimantan Utara. (2018)
- [9] Badan Pusat Statistik Provinsi Kalimantan Tengah. *Provinsi Kalimantan Tengah Dalam Angka 2018*. Badan Pusat Statistik: Kalimantan Tengah. (2018)
- [10] Badan Pusat Statistik Provinsi Kalimantan Timur. *Provinsi Kalimantan Timur Dalam Angka 2018*. Badan Pusat Statistik: Kalimantan Timur. (2018)
- [11] Sifriyani, Haryatami, I.N. Budiantara, dan Gunardi. “Geographically Weighted Regression with Spline Approach”. *Far East Journal of Mathematical Sciences (FJMS)*. Vol. 101 No. 6, pp. 1183-1196. Doi:10.17654/MS101061183. (2017)
- [12] Sifriyani, I. N. Budiantara, S.H. Kartiko, dan Gunardi. “A New Method of Hypotesis Test for Truncated Spline Nonparametric Regression Influenced by Spatial Heterogeneity and Application”. *Hindawi Abstract and Apllied Analysis*. Vol. 2018. Doi: 10.1155/2018/9769150. (2018b)
- [13] Sifriyani. “Simultaneous Hypothesis Testing of Multivariable Nonparametric Spline Regression in the GWR Model”. *International Journal of Satistics and Probability*. 8(4). Doi: 10.5539/ijsp.v8n4p32. (2019)
- [14] Fotheringham, A.S., C. Brudson, dan M. Charlton. *Geographically Wiegthed Regression: The Analysis of Spatial Varying Relationships*. John Wiley & Sons: England. (2002)