PEMODELAN SEMIPARAMETRIK *GEOGRAPHICAL WEIGHTED LOGISTIC REGRESSION*

SKRIPSI

FITRIATUSAKIAH

H12116015

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

FEBRUARI 2020

PEMODELAN SEMIPARAMETRIK *GEOGRAPHICAL WEIGHTED LOGISTIC REGRESSION*

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin

FITRATUSAKIAH

H12116015

PROGRAM STUDI STATISTIKA DEPARTEMEN STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

MAKASSAR

FEBRUARI 2020

LEMBAR PERNYATAAN KEOTENTIKAN

Saya yang bertanda tangan di bawah ini menyatakan dengan sungguh-sungguh bahwa skripsi yang saya buat dengan judul:

Pemodelan Semiparametrik *Geographical Weighted Logistic Regression*

adalah benar hasil karya saya sendiri, bukan hasil plagiat dan belum pernah dipublikasikan dalam bentuk apapun.

Makassar, 25 Februari 2020

METERAI TEMPEL
82AAHF654826824
6000
ENAM RIBU RUPIAH

FITRIATUSAKIAH

H12116015

PEMODELAN SEMIPARAMETRIK *GEOGRAPHICAL WEIGHTED LOGISTIC REGRESSION*

Disetujui oleh:

Pembimbing Utama	Pembimbing Pertama
Andi Kresna Jaya, S.Si, M.Si	Dr. La Podje Talangko, M.Si
NIP. 19731228 200003 1 001	NIP. 19551219 198701 1 001

Pada Tanggal: 25 Februari 2020

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh :

Nama	:	Fitriatusakiah
NIM	:	H12116015
Program Studi	:	Statistika
Judul Skripsi	:	Pemodelan Semiparametrik *Geographical Weighted Logistic Regression*

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

DEWAN PENGUJI

1. Ketua	: Andi Kresna Jaya, S.Si, M.Si	(..................)
2. Sekretaris	: Dr. La Podje Talangko, M.Si	(..................)
3. Anggota	: Dra. Nasrah Sirajang, M.Si	(..................)
4. Anggota	: Drs. Raupong, M.Si	(..................)

KEMENTERIAN RISET, TEKNOLOGI DAN PENDIDIKAN TINGGI UNIVERSITAS HASANUDDIN

Ditetapkan di : Makassar

: 25 Februari 2020

KATA PENGANTAR

Segala puji bagi Allah *Subhanahu Wa Ta'ala Rabb* semesta Alam, shalawat serta salam semoga senantiasa tercurahkan kepada Nabi yang paling dimuliakan, yakni Rasulullah Muhammad *Shallallahu Alaihi Wasallam* dan kepada para keluarga serta sahabat beliau.

"Maka sesungguhnya sesudah kesulitan ada kemudahan.
Sesungguhnya sesudah kesulitan ada kemudahan"
(QS. Al-Insyirah : 5-6)

Izinkan penulis haturkan terima kasih serta penghargaan setinggi-tingginya untuk orang tua penulis, **Buhaerah** dan **Nurmi** atas doa, kerja keras, nasehat, didikan, motivasi serta cinta dan limpahan kasih sayang yang tiada habisnya dan senantiasa diberikan kepada penulis. Banyak terima kasih juga penulis berikan kepada kakak tersayang **Muh. Rijal** atas doa, nasehat dan motivasi yang senantiasa diberikan kepada penulis.

Penghargaan yang tulus penulis ucapkan kepada:

1. **Rektor Universitas Hasanuddin, Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam, Ketua Departemen Statistika, segenap Dosen Pengajar dan Staf Departemen Statistika** yang telah membekali ilmu kepada penulis dalam berbagai hal selama menjadi mahasiswa di Departemen Statistika.
2. **Bapak Andi Kresna Jaya, S.Si, M.Si**, selaku Pembimbing Utama dan **Bapak Dr. La Podje Talangko, M.Si,** selaku Pembimbing Pertama yang telah meluangkan waktu dan pemikirannya untuk memberikan arahan, dorongan, dan motivasi kepada penulis mulai dari awal hingga selesainya penulisan tugas akhir ini.
3. **Ibu Dra. Nasrah Sirajang, M.Si,** selaku Anggota Tim Penguji dan Penasehat Akademik dan **Bapak Drs. Raupong, M.Si,** selaku Anggota Tim Penguji yang telah memberikan masukan, nasehat, motivasi dan pembelajaran yang bangun dalam penyempurnaan tugas akhir ini serta atas waktu yang di gkan kepada penulis.

Ucapan terimakasih juga penulis sampaikan kepada orang-orang yang telah berperan besar serta istimewa yaitu:

1. **STATISTIKA 2016** Terkhusus kepada **WKND 4EVER: Reski Amalah, Rosdiana, Rusydah Khaerati, Dewi Rahma Ente, Jumrianti, Dewi Santika Upa, Widya Nauli Amalia Puteri, Zhazha Alfkhamulki Ramdhani, Rayhanna Auliya Amin, Ayu Riski Ramadani, Isnawati, Andi Riska Fitriani, Bunga Aprilia, Reski Ulandari, Halniati, Agung Muhammad Takdir, Fajar Affan, Samsul Arifin, Jayzul Usrah, Suritman** dan **Rizki Adiputra**. Terima kasih selalu setia untuk mendengar dan berbagi kisah kehidupan perkuliahan. Terima kasih atas warna warni kehidupan selama ini. terima kasih karena masih selalu ada untuk mendoakan dan berbagi cerita serta saling memberi motivasi untuk mimpi-mimpi kedepannya.
2. Keluarga besar **HIMATIKA FMIPA UNHAS, HIMASTAT FMIPA UNHAS** terkhusus saudara seperjuangan **A16ORITMA,** Teman-Teman **KKN PPM Parepare** yang tidak sempat disebutkan satu persatu, terima kasih untuk cerita suka dan duka perjuangannya.
3. Kakak, adik dan teman-teman Statistika yang tak pernah bosan diganggu dan membantu. Terima kasih atas segala bentuk bantuan yang diberikan kepada penulis.
4. **Kita kita ji ^_^, Moh. Bima Abdillah** dan teman-teman **Pulmo** serta **Apcida** yang selalu memberikan warna warni kehidupan, terimakasih atas segala bantuan, doa, dukungan, hiburan serta kesetiaannya hingga saat ini.

Serta Semua pihak yang telah membantu dalam penyusunan tugas akhir ini. Penulis menyadari bahwa masih banyak kekurangan dalam tugas akhir ini. Oleh karena itu, dengan segala kerendahan hati penulis memohon maaf. Akhir kata semoga tulisan ini memberi manfaat kepada pembaca.

Makassar, 25 Februari 2020

Penulis

PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Hasanuddin, saya yang bertandatangan di bawah ini:

Nama	: Fitriatusakiah
NIM	: H12116015
Program Studi	: Statistika
Departemen	: Statistika
Fakultas	: Matematika dan Ilmu Pengetahuan Alam
Jenis Karya	: Skripsi

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Hasanuddin **Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty- Free Right*)** atas karya ilmiah saya yang berjudul:

"Pemodelan Semiparametrik *Geographical Weighted Logistic Regression*"

Beserta perangkat yang ada (jika diperlukan). Terkait dengan hal di atas, maka pihak universitas berhak menyimpan, mengalih-media/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di Makassar pada tanggal 25 Februari 2020

Yang menyatakan

akiah)

ABSTRAK

Geographical Weighted Logistic Regression (GWLR) adalah salah satu dari model regresi logistik yang memperhatikan lokasi. Model GWLR kemudian dikembangkan menjadi model semiparametrik yang terdiri dari parameter lokal dan global. Penentuan jenis parameter tersebut dalam penelitian ini menggunakan model linier koregionalisasi. Penaksir parameter model yang digunakan adalah metode *Maximum Likelihood Estimation* (MLE) dengan memberikan pembobot yang berbeda pada setiap lokasi pada parameter lokal. Pembobot yang digunakan yaitu *fixed gaussian kernel*. Penelitian ini bertujuan untuk mendapatkan model semiparametrik GWLR dan faktor yang berpengaruh secara signifikan terhadap status kemiskinan kabupaten/kota di Sulawesi Selatan tahun 2017. Berdasarkan hasil analisis, diperoleh bahwa setiap kabupaten/kota memiliki model semiparametrik GWLR yang berbeda-beda. Indikator yang signifikan mempengaruhi nilai status kemiskinan tiap kabupaten/kota di Sulawesi Selatan adalah rumah dengan atap beton, genteng seng dan asbes untuk parameter lokal dan rumah dengan lantai bukan tanah untuk parameter global.

Kata Kunci: Semiparametrik GWLR, MLE, Fungsi Pembobot *Fixed Gaussian Kernel*, Model Linier Koregionalisasi.

ABSTRACT

Geographical Weighted Logistic Regression (GWLR) is one of the logistics regression models that concern the location. The GWLR model was then developed into a semiparametric model consisting of local and global parameters. The determination of these types of parameters in the study used the linear model of coregionalization. The Model parameter estimator used is the Maximum Likelihood Estimation (MLE) method by providing different weight on each location in the local parameters. The weight used is a fixed Gaussian kernel. This research aims to obtain GWLR Semiparametric model and significantly influential factor on the poverty status of the district/city in South Sulawesi in 2017. Based on the results of the analysis, it is obtained that each district/city has a different GWLR semiparametric model. A significant indicator affects the value of poverty status of each district/city in South Sulawesi is a home with concrete roof, zinc tile and asbestos for local parameters and home with non-ground floor for global parameters..

Keywords: GWLR Semiparametric, MLE, Weighting Function of *Fixed Gaussian Kernel, Linear Coregionalization Model*.

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR LAMPIRAN

BAB I
PENDAHULUAN

1.1 Latar Belakang

Salah satu bagian dari analisis data statistika yang digunakan untuk mendapatkan model hubungan antara variabel respon (y) dengan satu atau lebih variabel prediktor (x) adalah analisis regresi (Agresti, 2002). Model dari regresi linier memiliki variabel respon bersifat kontinu dan diasumsikan mengikuti distribusi normal. Akan tetapi, terkadang ditemukan variabel respon yang bersifat diskrit, maka model regresi yang digunakan adalah model regresi logistik.

Model regresi logistik umumnya digunakan ketika unit pengamatan bukan merupakan wilayah atau lokasi, karena model tersebut tidak mempertimbangkan efek lokasi yang selanjutnya disebut efek spasial. Fotheringham & Charlthon dalam Rosa (2015) mengatakan segala sesuatu saling berhubungan satu dengan yang lainnya, tetapi sesuatu yang lebih dekat akan lebih berpengaruh daripada sesuatu yang jauh. Hubungan tersebut dinamakan efek spasial. Oleh karena itu dikembangkan sebuah metode analisis yang dipengaruhi oleh faktor geografis. Faktor geografis diperlukan karena tiap lokasi diduga mempunyai karakteristik yang berbeda, kondisi ini menyebabkan data antar pengamatan sulit untuk memenuhi asumsi regresi sehingga diperlukan analisis yang memperhatikan faktor geografis (Hasriana, 2016).

Model yang dapat mengatasi masalah tersebut adalah model spasial. Salah satu dari model regresi logistik spasial yang memperhatikan lokasi dan diasumsikan bahwa data variabel respon berdistribusi Bernoulli adalah *Geographically Weighted Logistic Regression* (GWLR) (Pradita, 2011). Model GWLR kemudian dikembangkan menjadi model semiparametrik (Fathurahman, 2014). Semiparametrik pada ilmu spasial adalah penggabungan antara regresi spasial dengan regresi biasa (Nakaya dalam Septia Ningsih, 2016). Model semiparametrik GWLR terdiri dari parameter yang dipengaruhi lokasi (*geographically varying* *...t*) dan parameter yang tidak dipengaruhi lokasi (*fixed coefficiet*) (Nakaya, ...enaksir parameter model semiparametrik GWLR diperoleh menggunakan ... *maximum likelihood estimation* dengan memberikan pembobot yang berbeda pada setiap lokasi. Matriks pembobot dapat dibentuk menggunakan suatu fungsi pembobot yang tergantung pada ukuran ketetanggan atau biasa disebut *bandwidth* (Desriwendi, 2015).

Penelitian sebelumnya telah dilakukan oleh Zakiyah Mar’ah (2017) tentang pemodelan regresi terboboti geografis semiparametrik dengan model linier koregionalisasi. Hasil dari penelitian tersebut adalah diperoleh parameter yang dipengaruhi lokasi dan parameter yang tidak dipengaruhi lokasi serta semiparametrik GWR merupakan model terbaik dibandingkan dengan model GWR pada data tingkat kemiskinan di Provinsi Sulawesi Utara. Dewi F.S (2015) telah membahas pemodelan *geographically weighted logistic regression semiparametric* yang menjelaskan bahwa model GWLR *semiparametric* untuk status kesejahteraan kabupaten/kota di Jawa Tengah tahun 2012 memiliki ketepatan klasifikasi yang lebih tinggi dibanding model regresi logistik dan GWR.

Septia Ningsih (2016) juga telah membahas tentang model *geographically weighted logistic regression semiparametric* dengan pembobot *fixed Gaussian kernel* dan *fixed bisquare kernel.* Model tersebut diterapkan pada data kemiskinan daerah/kabupaten dan kota di provinsi Jawa Barat tahun 2012. Hasil yang diperoleh menjelaskan bahwa pembobot *fixed Gaussian kernel* adalah pembobot terbaik yang dapat diterapkan pada data kemiskinan di Provinsi Jawa Barat.

Namun, tingkat kemiskinan di suatu kabupaten/kota berbeda-beda berdasarkan jumlah penduduk miskin. Provinsi Sulawesi Selatan adalah provinsi yang memiliki tingkat kemiskinan sebesar 9.38% dan memiliki tigkat *Head Count Index* (HCI) sebesar 9.37% pada tahun 2017. Nilai HCI merupakan salah satu pengkategorian tingkat kemiskinan di kabupaten/kota. Suatu kabupaten/kota berstatus miskin jika nilai HCI lebih dari atau sama dengan 9.37% dan berstatus tidak miskin jika nilai HCI kurang dari 9.37%. Pengelompokan status kemiskinan bertujuan untuk mempermudah analisis faktor-faktor yang mempengaruhi kemiskinan secara spasial (Data dan Informasi Kemiskinan, 2018).

Oleh karena itu, dalam penelitian ini akan dibahas mengenai pemodelan iskinan menggunakan semiparametrik *geographical weighted logistic* *n* di Provinsi Sulawesi Selatan tahun 2017 menggunakan model linier alisasi untuk memperoleh parameter yang dipengaruhi lokasi (parameter lokal) dan parameter yang tidak dipengaruhi lokasi (parameter global) dengan pembobot *fixed Gaussian kernel*.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana model semiparametrik GWLR pada data kemiskinan di Provinsi Sulawesi Selatan tahun 2017?
2. Faktor apa saja yang mempengaruhi secara signifikan pada data kemiskinan di Provinsi Sulawesi Selatan tahun 2017?

1.3 Batasan Masalah

Batasan masalah dalam penelitian ini adalah penaksiran parameter menggunakan *Maximum Likelihood Estimation* (MLE) dengan pembobot *fixed Gaussian kernel* pada data kemiskinan di Provinsi Sulawesi Selatan tahun 2017 yang mengandung heterogenitas spasial serta model linier koregionalisasi yang digunakan terbentuk dari model struktur tersarang yang terdiri dari pengaruh *nugget* dan dua model semivariogram dasar.

1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Mendapatkan model semiparametrik GWLR pada data kemiskinan di Provinsi Sulawesi Selatan tahun 2017.
2. Mendapatkan faktor yang berpengaruh secara signifikan pada data kemiskinan di Provinsi Sulawesi Selatan tahun 2017.

1.5 Manfaat Penelitian

Hasil penelitian ini diharapkan dapat menambah wawasan keilmuan mengenai pemodelan data kemiskinan dan faktor yang mempengaruhi kemiskinan bupaten dan kota di Provinsi Sulawesi Selatan dengan menggunakan metrik *geographical weighted logistic regression* serta membantu pemerintah nagmbil arah kebijakan untuk menurunkan tingkat kemiskinan.

BAB II
TINJAUAN PUSTAKA

2.1 Regresi Logistik

Regresi logistik merupakan salah satu metode yang dapat digunakan untuk mencari hubungan variabel respon yang bersifat dikotomus (berskala nominal atau ordinal dengan dua kategori) atau polikotomus (berskala nominal atau ordinal dengan lebih dari dua kategori) dengan satu atau lebih variabel prediktor yang bersifat kontinu atau kategorik (Hosmer & Lemeshow, 2000). Jika hasil pengamatan variabel respon (y) mempunyai 2 kategori yaitu 0 dan 1, maka regresi logistik tersebut menggunakan regresi logistik biner sehingga mengikuti distribusi Bernoulli dengan distribusi peluang (Agresti, 2002):

$$P(Y = y) = \pi^y (1 - \pi)^{1-y}\ , y = 0, 1. \qquad (2.1)$$

Menurut Hosmer & Lemeshow (2000), model probabilitas regresi logistik dengan melibatkan beberapa variabel prediktor (x_i) adalah sebagai berikut:

$$\pi(x_i) \ = \frac{\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik})}, \qquad (2.2)$$

dengan $i = 1,2, \dots, n\ dan\ j = 1,2, \dots, k.$

Fungsi $\pi(x_i)$ merupakan fungsi non linier sehingga perlu dilakukan transformasi logit untuk memperoleh fungsi yang linier agar dapat dilihat hubungan antara variabel respon (y) dengan variabel prediktornya (x_i). Bentuk logit dari $\pi(x_i)$ dinyatakan sebagai $g(x_i)$ yaitu :

$$g(x_i) = ln\left[\frac{\pi(x_i)}{1-\pi(x_i)}\right] = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik}, \qquad (2.3)$$

dengan :

$\pi(x_i)$: peluang untuk variabel prediktor ke-*i*

k : banyaknya variabel prediktor

n : banyaknya pengamatan

β_0 : konstanta (intersep) pada masing-masing lokasi

β_j : parameter model regresi logistik untuk variabel prediktor ke-*j*

ilai pengamatan variabel prediktor ke-*k* pada lokasi pengamatan ke-*i*

enaksir parameter pada regresi logistik dapat dilakukan dengan akan metode *Maximum Likelihood Estimation* (MLE) yaitu memaksimumkan fungsi *likelihood*. Fungsi *likelihood* diperoleh dengan mengalikan fungsi-fungsi kepadatan peluang dari y_i karena setiap pengamatan diasumsikan independen. Sehingga dapat dituliskan dalam bentuk:

$$L(\beta) = \prod_{i=1}^{n} f(y_i), \tag{2.4}$$

dengan $f(y_i) = \pi(x_i)^{y_i}[1 - \pi(x_i)]^{1-y_i}$.

Untuk mendapatkan nilai β yang memaksimumkan nilai fungsi *likelihood*, maka dilakukan transformsi dari fungsi *likelihood*, yang kemudian disebut fungsi log *likelihood*. Sehingga dapat dituliskan dalam bentuk:

$$\ln L(\beta) = \left[\sum_{i=1}^{n} \ln\{\pi(x_i)^{y_i}[1 - \pi(x_i)]^{1-y_i}\}\right]. \tag{2.5}$$

Untuk mendapatkan nilai β yang memaksimumkan nilai fungsi *likelihood* maka nilai parameter β dari turunan pertama fungsi $L(\beta)$ didapatkan melalui suatu prosedur iteratif yang dilakukan dengan metode iterasi Newton Raphson yaitu memaksimumkan fungsi *likelihood* (Agresti, 2002).

2.2 Regresi Spasial

Regresi spasial adalah analisis yang mengevaluasi hubungan antara satu variabel dengan beberapa variabel lain dengan memberikan efek spasial pada beberapa lokasi yang menjadi pusat pengamatan (Fatati I.F, 2017). Regresi spasial merupakan hasil pengembangan dari metode regresi klasik. Pengembangan itu berdasarkan adanya pengaruh tempat atau spasial pada data yang dianalisis. Data spasial adalah suatu data yang mengacu pada posisi, objek dan hubungan diantaranya dalam ruang bumi (Salmawaty dkk, 2016). Data spasial merupakan sebuah data yang berorientasi geografis, memiliki sistem koordinat (lintang dan bujur) tertentu sebagai dasar referensinya.

2.3 Geographically Weighted Logistic Regression

Geographically Weighted Logistic Regression (GWLR) merupakan salah ode nonparametrik pada regresi yang dapat mempertimbangkan faktor ehingga akan dihasilkan nilai parameter bagi masing-masing titik atau mana data tersebut diamati. GWLR digunakan untuk memprediksi atau menduga model dari kumpulan data yang memiliki variabel respon biner melalui model logistik. Model umum GWLR adalah sebagai berikut:

$$\pi(x_i) = \frac{\exp\left(\beta_0(u_i, v_i) + \sum_{j=1}^{k} \beta_j(u_i, v_i) x_{ij}\right)}{1 + \exp\left(\beta_0(u_i, v_i) + \sum_{j=1}^{k} \beta_j(u_i, v_i) x_{ij}\right)}, \tag{2.6}$$

dengan $i = 1,2, \dots, n$; $j = 1,2, \dots, k$.

Bentuk logit model GWLR untuk $i = 1,2, \dots, n$ yang dinyatakan dengan $g(x_i)$ adalah sebagai berikut :

$$g(x_i) = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)x_{i1} + \beta_2(u_i, v_i)x_{i2} + \cdots + \beta_k(u_i, v_i)x_{ik}, \tag{2.7}$$

dengan:

k : banyaknya variabel prediktor

$\beta_0(u_i, v_i)$: konstanta (intersep) pada masing-masing lokasi

$\beta_j(u_i, v_i)$: parameter model variabel prediktor ke-*j* pada masing-masing lokasi

(u_i, v_i) : koordinat *latitutde* dan *longitude* dari titik ke-*i* pada suatu lokasi geografis.

x_{ij} : nilai pengamatan variabel prediktor ke-*j* pada lokasi pengamatan ke- i.

Parameter model GWLR diestimasi menggunakan metode MLE dengan menambahkan bobot yang berbeda untuk setiap lokasi pengamatan dalam persamaan log *likelihood*. Model GWLR menghasilkan bentuk penaksir parameter model yang bersifat lokal untuk setiap lokasi.

2.4 Semiparametrik *Geographical Weighted Logistic Regression*

Semiparametrik GWLR merupakan perluasan dari metode GWLR (Fathurahman, 2014). Nakaya *et al.* (2005) menyatakan semiparametrik GWLR terdiri dari parameter yang dipengaruhi lokasi (*geographically varying coefficient*) dan parameter yang tidak dipengaruhi lokasi (*fixed coefficient*). Parameter yang tidak dipengaruhi lokasi bernilai konstan untuk seluruh lokasi pengamatan. Pada model semiparametrik GWLR, variabel respon (y) diprediksi berdasarkan variabel prediktor (x) yang masing-masing koefisien regresinya $\beta_j(u_i, v_i)$ bergantung pada ... ta tersebut diamati dan koefisien regresi α_m tidak bergantung pada lokasi ... . Model semiparametrik GWLR dapat dituliskan seperti persamaan (2.8).

$$\pi(x_i) = \frac{\exp\left(\beta_0(u_i, v_i) + \sum_{j=1}^{k^*} \beta_j(u_i, v_i) x_{ij} + \sum_{m=k^*+1}^{k} \alpha_m x_{im}\right)}{1 + \exp\left(\beta_0(u_i, v_i) + \sum_{j=1}^{k^*} \beta_j(u_i, v_i) x_{ij} + \sum_{m=k^*+1}^{k} \alpha_m x_{im}\right)}, \quad (2.8)$$

dengan $i = 1,2, \dots, n;\ j = 1,2, \dots, k$.

Bentuk logit untuk model semiparametrik GWLR adalah :

$$g(x_i) = \beta_0(u_i, v_i) + \sum_{j=1}^{k^*} \beta_j(u_i, v_i) x_{ij} + \sum_{m=k^*+1}^{k} \alpha_m x_{im}, \quad (2.9)$$

dengan :

k	: banyaknya variabel prediktor
k^*	: banyaknya variabel prediktor pada model yang bersifat lokal
$\beta_j(u_i, v_i)$	: parameter model yang bersifat lokal ke-*j* untuk setiap lokasi
$\beta_0(u_i, v_i)$	: konstanta (intersep) pada masing-masing lokasi
α_m	: parameter model yang bersifat global ke-*m* untuk setiap lokasi
x_{im}	: nilai pengamatan variabel prediktor yang bersifat global ke-*m* pada lokasi ke-*i*
x_{ij}	: nilai pengamatan variabel prediktor yang bersifat lokal ke-*j* pada lokasi ke-*i*

Model semiparametrik GWLR menghasilkan bentuk penaksir parameter model yang bersifat lokal untuk setiap lokasi dan penaksir parameter yang bersifat global atau tidak dipengaruhi lokasi. Prosedur penaksiran parameter pada model semiparametrik GWLR menggabungkan mekanisme parametrik untuk menaksir parameter global atau parameter yang bernilai konstan dan non parametrik untuk menaksir parameter lokal atau parameter yang dipengaruhi lokasi (Nakaya *et al*. 2005). Pada model semiparametrik GWLR, metode penaksir parameter yang digunakan adalah metode MLE. Fungsi *likelihood* yang terbentuk yaitu sebagai berikut:

$$L(\beta(u_i, v_i), \alpha_m) = \left\{ exp \sum_{i=1}^{n} y_i \left(\beta_0(u_i, v_i) + \sum_{j=1}^{k^*} \beta_j(u_i, v_i) x_{ij} + \sum_{m=k^*+1}^{k} \alpha_m x_{im} \right) \right\}$$

$$\left\{ \prod_{i=1}^{n} \left(1 + \exp\left(\beta_0(u_i, v_i) + \sum_{j=1}^{k^*} \beta_j(u_i, v_i) x_{ij} + \sum_{m=k^*+1}^{k} \alpha_m x_{im} \right) \right)^{-1} \right\}. \quad (2.10)$$

Persamaan fungsi log *likelihood* adalah sebagai berikut:

$$\ln L\,(\beta(u_i,v_i),\alpha_m) = \sum_{i=1}^{n} y_i\left(\beta_0(u_i,v_i) + \sum_{j=1}^{k^*}\beta_j(u_i,v_i)x_{ij} + \sum_{m=k^*+1}^{k}\alpha_m x_{im}\right)$$

$$-\sum_{i=1}^{n} ln\left\{1+\exp\left(\beta_0(u_i,v_i)+\sum_{j=1}^{k^*}\beta_j(u_i,v_i)x_{ik}+\sum_{m=k^*+1}^{k}\alpha_m x_{im}\right)\right\}. \quad (2.11)$$

Faktor letak geografis merupakan faktor pembobot pada model semiparametrik GWLR. Pembobot dimasukkan pada persamaan fungsi log *likelihood* yaitu:

$$\ln L^*\,(\beta(u_i,v_i),\alpha_m) = \sum_{i=1}^{n} w_{ij}(u_i,v_i)y_i\left(\beta_0(u_i,v_i) + \sum_{j=1}^{k^*}\beta_j(u_i,v_i)x_{ij} + \sum_{m=k^*+1}^{k}\alpha_m x_{im}\right)$$

$$-\sum_{i=1}^{n} w_{ij}(u_i,v_i) ln\left\{1+\exp\left(\beta_0(u_i,v_i)+\sum_{j=1}^{k^*}\beta_j(u_i,v_i)x_{ij}+\sum_{m=k^*+1}^{k}\alpha_m x_{im}\right)\right\} \quad (2.12)$$

Faktor letak geografis merupakan faktor pembobot pada model semiparametrik GWLR. Nilai $\ln L^*\,(\beta(u_i,v_i),\alpha_m)$ kemudian diturunkan terhadap $\beta(u_i,v_i)$ dan α_m. Hasil yang diperoleh disamakan dengan nol agar diperoleh nilai β dan α_m yang dapat memaksimumkan $L(\beta(u_i,v_i),\alpha_m)$. Selanjutnya karena hasil turunan pertama fungsi log *likelihood* terhadap masing-masing parameter tidak dapat diselesaikan secara analitik, maka digunakan metode iterasi Newton Raphson sebagai berikut:

$$\left(\boldsymbol{\beta}^{(t+1)}(u_i,v_i),\boldsymbol{\alpha}^{(t+1)}\right) = \boldsymbol{\beta}^{(t)}(u_i,v_i),\boldsymbol{\alpha}^{(t)}$$
$$-(\boldsymbol{H}^{(t)^{-1}}(\boldsymbol{\beta}^{(t)}(u_i,v_i),\boldsymbol{\alpha}^{(t)}))\boldsymbol{g}^{(t)}(\boldsymbol{\beta}^{(t)}(u_i,v_i),\boldsymbol{\alpha}^{(t)})), \quad (2.13)$$

dengan :

$$\boldsymbol{g}^{(t)}(\boldsymbol{\beta}^{(t)}(u_i,v_i),\boldsymbol{\alpha}^{(t)}) = \begin{bmatrix} \dfrac{\partial \ln L^*(\beta(u_i,v_i),\alpha)}{\partial \beta_0(u_i,v_i)} \\ \dfrac{\partial \ln L^*(\beta(u_i,v_i)\alpha)}{\partial \beta_1(u_i,v_i)} \\ \vdots \\ \dfrac{\partial \ln L^*(\beta(u_i,v_i)\alpha)}{\partial \beta_{k^*}(u_i,v_i)} \\ \dfrac{\partial \ln L^*(\beta(u_i,v_i)\alpha)}{\partial \alpha_{k^*+1}} \\ \vdots \\ \dfrac{\partial \ln L^*(\beta(u_i,v_i)\alpha)}{\partial \alpha_k} \end{bmatrix},$$

$$\boldsymbol{H}^{(t)}(\boldsymbol{\beta}^{(t)}(u_i,v_i),\boldsymbol{\alpha}^{(t)}) = \begin{bmatrix} \dfrac{\partial^2 \ln L^*(\beta(u_i,v_i),\alpha)}{\partial \boldsymbol{\beta}^{\boldsymbol{T}}(u_i,v_i)\partial \boldsymbol{\beta}(u_i,v_i)} & \dfrac{\partial^2 \ln L^*(\beta(u_i,v_i),\alpha)}{\partial \boldsymbol{\beta}^{\boldsymbol{T}}(u_i,v_i)\partial \boldsymbol{\alpha}} \\ \dfrac{\partial^2 \ln L^*(\beta(u_i,v_i),\alpha)}{\partial \boldsymbol{\beta}^{\boldsymbol{T}}(u_i,v_i)\partial \boldsymbol{\alpha}} & \dfrac{\partial^2 \ln L^*(\beta(u_i,v_i),\alpha)}{\partial \boldsymbol{\alpha}^{\boldsymbol{T}}\partial \boldsymbol{\alpha}} \end{bmatrix},$$

$(\boldsymbol{\beta}^T(u_i,v_i)) = [\beta_0(u_i,v_i) \quad \beta_1(u_i,v_i) \quad \cdots \quad \beta_{k^*}(u_i,v_i)]$ dan $\boldsymbol{\alpha}^T = [\alpha_{k^*+1} \quad \cdots \quad \alpha_k]$.

Iterasi berhenti pada saat konvergen, yaitu pada saat $\left\|\boldsymbol{\beta}^{(t+1)}(u_i,v_i),\boldsymbol{\alpha}^{(t+1)}\right\| = \left\|\boldsymbol{\beta}^{(t)}(u_i,v_i),\boldsymbol{\alpha}^{(t)}\right\|$.

Pengujian parameter model semiparametrik GWLR dilakukan dengan menguji parameter secara parsial. Pengujian ini dilakukan untuk mengetahui parameter variabel prediktor yang bersifat global dan lokal yang berpengaruh secara signifikan terhadap variabel respon.

a. Uji parsial model lokal

Hipotesis:

$H_0 : \beta_j(u_i,v_i) = 0$ (parameter variabel yang bersifat lokal tidak signifikan)

$H_1 : \beta_j(u_i,v_i) \neq 0$ (parameter variabel yang bersifat lokal signifikan)

Statistik uji:

$$Z_{hit} = \frac{\hat{\beta}_j(u_i,v_i)}{se(\hat{\beta}_j(u_i,v_i))} \tag{2.14}$$

dengan:

$$se\left(\hat{\beta}_j(u_i,v_i)\right) = \sqrt{var(\hat{\beta}_j(u_i,v_i))}$$

Daerah penolakan: Tolak H0 jika $|Z_{hit}| > Z_{\alpha/2}$

b. Uji parsial model global

Hipotesis:

$H_0 : \alpha_m = 0$ (parameter variabel yang bersifat global tidak signifikan)

$H_1 : \alpha_m \neq 0$ (parameter variabel yang bersifat global signifikan)

Statistik uji:

$$Z_{hit} = \frac{\hat{\alpha}_m}{se(\hat{\alpha}_m)} \tag{2.15}$$

n :

$) = \sqrt{var(\hat{\alpha}_m)}$

n penolakan: Tolak H0 jika $|Z_{hit}| > Z_{\alpha/2}$.

2.5 Analisis Variogram

Variogram merupakan statistik deskriptif yang dapat memperlihatkan secara grafik kontinuitas spasial, karena jika ada dua buah nilai spasial yang letaknya berdekatan, maka akan relatif bernilai sama dibandingkan dengan dua buah nilai spasial yang letaknya berjauhan (Cressie, 1993). Variogram dapat menggambarkan hubungan spasial antar variabel dan dinyatakan sebagai $2\gamma(h)$. Variogram dirumuskan sebagai berikut :

$$2\gamma(h) = E[(x_i - x_{i^*})^2], \tag{2.16}$$

dengan $\gamma(h)$ disebut sebagai semivariogram. Penduga semivariogram dirumuskan sebagai berikut (Cressie, 1993):

$$\hat{\gamma}_x(h) = \frac{1}{2N(h)} \sum_{(i,i^*)=1}^{N(h)} (x_i - x_{i^*})^2, \tag{2.17}$$

dan semivariogram silang yang dapat menggambarkan hubungan spasial setiap pasangan variabel didefinisikan sebagai berikut:

$$\hat{\gamma}_{xy}(h) = \frac{1}{2N(h)} \sum_{(i,i^*)=1}^{N(h)} (x_i - x_{i^*})(y_i - y_{i^*}), \quad i \neq i^* \tag{2.18}$$

dengan :

i, i^* : $1, 2, \ldots, n$

n : banyaknya pengamatan

$N(h)$: banyaknya pasangan data yang memiliki jarak h

h : jarak antara dua titik sampel

x_i : pengamatan ke-i pada variabel x

y_i : pengamatan ke-i pada variabel y

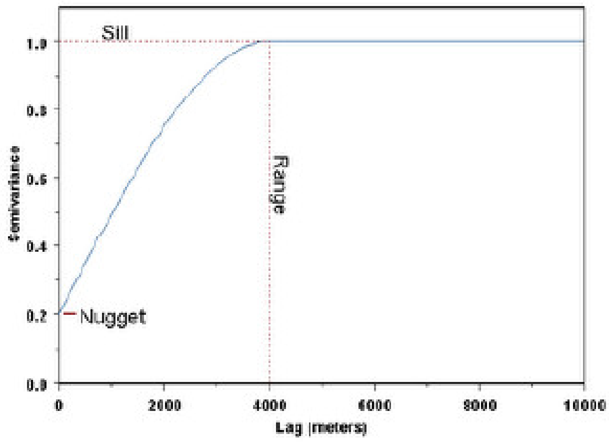
x_i , x_{i^*} : pasangan titik sampel yang berjarak h

Untuk mencari nilai semivariogram, banyak pasangan data akan di bagi menjadi beberapa kelas (k). Setelah nilai semivariogram diperoleh, dilanjutkan dengan menghitung nilai parameter yang akan digunakan pada perhitungan model semivariogram. Beberapa parameter yang digunakan untuk mencari nilai dalam model semivariogram yaitu:

1. *Nugget effect* merupakan pendekatan nilai semivariogram pada jarak sekitar nol. Pengaruh *nugget* menyatakan keragaman yang dikarenakan satu atau beberapa , seperti sifat keragaman spasial variabel, faktor eksternal yang tidak , kesalahan pengukuran dan ketiadaan informasi karena jarak yang sangat Selain itu, ukuran pengaruh non-*nugget* (skala spasial kecil dan skala

spasial besar) menyatakan persentase pengaruh spasial yang dapat dimodelkan (Ribeiro *et al*. 2016).

2. *Sill* (c) adalah nilai struktur korelasi spasial tertinggi saat dimana nilai semivariogram pada jarak tertentu cenderung mencapai nilai yang stabil (Issaks & Srivastava, 1989).
3. *Range* (a) merupakan jarak pada saat semivariogram mencapai nilai *sill* atau jarak pada saat semivariogram membentuk sebuah dataran tinggi atau sebuah masa stabil (Issaks & Srivastava, 1989). Dalam grafik semivariogram, *range* yaitu jarak pada sumbu horizontal mulai dari titik nol sampai titik proyeksi perubahan variogram dari miring ke mendatar.



Gambar 2. 1 Semivarigram

Semivariogram dihitung setelah nilai *nugget*, *sill* dan *range* diperoleh. Nilai yang diperoleh dari model semivariogram akan digunakan untuk membandingkan nilai MSE. Dari perbandingan tersebut akan dipilih model yang memiliki nilai MSE paling kecil, yang nantinya model tersebut akan digunakan untuk melakukan pendugaan data spasial.(Rozalia, 2015). Terdapat beberapa jenis model semivariogram dasar yang sering digunakan, yaitu (Cressie, 1993) :

1. Model *Nugget*

   Semivariogram untuk model *nugget* dirumuskan sebagai berikut :

$$g(h) = \begin{cases} 0, & untuk\ h = 0, \\ 1, & untuk\ h > 0. \end{cases} \quad (2.19)$$

l *Spherical*

variogram untuk model *spherical* dirumuskan sebagai berikut :

$$g(h) = \begin{cases} \left[\left(\frac{3h}{2a}\right) - 0{,}5\left(\frac{h}{a}\right)^3\right], & untuk\ 0 \le h < a, \\ 1, & untuk\ h \ge a. \end{cases} \tag{2.20}$$

3. Model *Eksponensial*

   Semivariogram untuk model *eksponensial* dirumuskan sebagai berikut:

$$g(h) = \left[1 - \exp\left(-\frac{3h}{a}\right)\right]. \tag{2.21}$$

4. Model *Gaussian*

   Semivariogram untuk model *Gaussian* dirumuskan sebagai berikut:

$$g(h) = \left[1 - \exp\frac{-3h^2}{a^2}\right], \tag{2.22}$$

2.6 Model Linier Koregionalisasi

Model Linier Koregionalisasi (MLK) merupakan alat yang berguna untuk menggambarkan hubungan spasial antar variabel (Goulard, 1992). Isaaks & Srivastava (1989) menyatakan MLK terdiri dari semivariogram dan semivariogram silang dari dua atau lebih variabel. Setiap variabel dicirikan dengan semivariogramnya dan setiap pasangan variabel dicirikan oleh semivariogram silangnya. MLK terbentuk dari model struktur tersarang dari kombinasi linier model semivariogram dasar sebagai berikut:

$$\gamma_{ij}(h) = \sum_{s=0}^{m} c_{ij,s}\, g_s(h_s), \tag{2.23}$$

dengan $i = 1,2,\dots,p, j = 1,2,\dots,p, p$ merupakan banyaknya variabel, s merupakan struktur model dengan $s = 0,1,2,\dots,m$ dan m merupakan jumlah struktur tersarang model yang digunakan. MLK yang digunakan terbentuk dari model struktur tersarang yang terdiri dari pengaruh *nugget* dan dua model semivariogram dasar sebagai berikut (Zakiyah Mar'ah, 2017):

$$\gamma_{ij}(h) = c_{ij,0}\, g_0(h_0) + c_{ij,1}\, g_1(h_1) + c_{ij,2}\, g_2(h_2), \tag{2.24}$$

dengan :

$c_{ij,0}$: koefisien semivariogram (*sill*) untuk pengaruh *nugget*

$c_{ij,1}$: koefisien semivariogram untuk model semivariogram dasar pertama

koefisien semivariogram untuk model semivariogram dasar kedua

pengaruh *nugget*

model semivariogram dasar pertama

model semivariogram dasar kedua

(h_0) : jarak untuk model semivariogram pengaruh *nugget*

(h_1) : jarak untuk model semivariogram dasar pertama

(h_2) : jarak untuk model semivariogram dasar kedua

Setiap model semivariogram dan semivariogram silang dibangun menggunakan model semivariogram dasar yang sama. Wackernagel (2003) menyatakan bahwa beberapa matriks MLK menggambarkan struktur korelasi pada jarak spasial yang berbeda. Bentuk matriks pada persamaan (2.22) sebagai berikut:

$$\boldsymbol{\Gamma}(h) = \begin{bmatrix} \gamma_{11}(h) & \cdots & \gamma_{1p}(h) \\ \vdots & \ddots & \vdots \\ \gamma_{np}(h) & \cdots & \gamma_{pp}(h) \end{bmatrix} = \sum_{s=0}^{m} \boldsymbol{c}_s g_s(h_s), \qquad (2.25)$$

dengan :

$\boldsymbol{c}_s$ **:** matriks persegi koefisien semivariogram (matriks koregionalisasi),

$$\boldsymbol{c}_s = \begin{bmatrix} c_{11,s} & \cdots & c_{1p,s} \\ \vdots & \ddots & \vdots \\ c_{p1,s} & \cdots & c_{pp,s} \end{bmatrix}$$

$\boldsymbol{\Gamma}(h)$: matriks semivariogram pada jarak *h*, diagonalnya merupakan semivariogram dan *off*-diagonal merupakan semivariogram silang

g_s : dugaan fungsi semivariogram

Matriks koregionalisasi $\boldsymbol{c}_s$ adalah matriks korelasi yang dapat dipandang sebagai partisi matriks ragam-peragam yang menggambarkan struktur korelasi pada jarak spasial yang berbeda. Proporsi keragaman spasial tiap variabel disetiap jarak spasial yang berbeda dihitung berdasarkan persentase keragaman spasial tiap variabel disetiap jarak spasial dibagi dengan jumlah keragaman spasial tiap variabel, sedangkan pengaruh spasial masing-masing variabel diperoleh dari proporsi pengaruh non *nugget* (skala spasial kecil dan skala spasial besar) (Pebesma, 2004).

2.7 Pengujian Multikolinieritas

Pada analisis regresi logistik tidak diperkenankan terjadi kasus multikolinieritas. Uji multikolinieritas yaitu adanya hubungan linier antara variabel prediktor dalam model regresi. Metode untuk menguji adanya multikolinieritas hat pada *tolerance value* atau V*ariance Inflation Factor* (VIF). Nilai VIF eroleh dengan rumus berikut :

$$VIF = \frac{1}{tolerance} = \left(1 - R_j{}^2\right)^{-1}, \qquad (2.26)$$

dengan ${R_j}^2$ adalah koefisien detrminasi antara x dengan variabel prediktor lain. Rumus koefisien determinasi adalah:

$$R^2 = \frac{\left[\sum(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})\right]^2}{\sum(y_i - \bar{y})^2 \left(\hat{y}_i - \bar{\hat{y}}\right)^2}. \tag{2.27}$$

Berdasarkan persamaan (2.27), maka rumus untuk koefisien determinasi antara X_j dengan variabel prediktor lain adalah (Hasriana, 2016).:

$$R_j{}^2 = \frac{\left[\sum(X_j - \bar{X}_j)(X_{j'} - \bar{X}_{j'})\right]^2}{\sum(X_j - \bar{X}_j)^2 \left(X_{j'} - \bar{X}_{j'}\right)^2}, \tag{2.28}$$

dengan :

X_j : variable prediktor ke-j; $j = 1,2,\ldots,k, j \neq j'$.

Batas *tolerance* adalah 0,10 atau nilai VIF adalah 10. Jika nilai VIF < 10 dan nilai *tolerance* > 0,10, maka dapat diartikan tidak terdapat multikolinieritas pada penelitian tersebut (Santoso, 2012).

2.8 Pengujian Heterogenitas Spasial

Heterogenitas spasial terjadi akibat adanya perbedaan antara satu wilayah dengan wilayah lainnya. Penyimpanagan heterogenitas biasa disebut dengan homoskedastisistas. Homoskedastisitas ditandai dengan variasi residu yang sama untuk semua pengamatan. Adanya penyimpangan asumsi ini (heteroskedastisitas) dalam regresi dapat diketahui dengan menggunakan beberapa cara, salah satunya adalah uji Breusch-Pagan (Anselin, 1998).

Hipotesis:

H_0 : ${\sigma_1}^2 = {\sigma_2}^2 = \cdots = {\sigma_n}^2 = \sigma^2$(terjadi homoskedastisitas)

H_1: Minimal ada satu ${\sigma_1}^2 \neq \sigma^2$ (terjadi heteroskedastisitas)

Statistik uji:

$$BP = \left(\frac{1}{2}\right)\boldsymbol{f}^T\boldsymbol{Z}(\boldsymbol{Z}^T\boldsymbol{Z})^{-1}\boldsymbol{Z}^T\boldsymbol{f} \sim \chi^2_{(k-1)}, \tag{2.29}$$

dengan elemen vector f adalah $f_i = \left(\frac{e_i^2}{\sigma^2}\right) - 1$ dimana $e_i = y_i - \bar{y}$adalah *residual* ngamatan ke-i, σ^2 adalah ragam *residual* dan Z merupakan matriks n $(n \times (k+1)$ berisi vektor yang sudah di normal standarkan untuk tiap tan.

Daerah penolakan:

Tolak H$_0$ jika $BP > \chi^2_{(k-1)(\alpha)}$ atau nilai p-value < α, dengan $\chi^2_{(k-1)(\alpha)}$ merupakan titik kritis uji χ^2 dengan taraf nyata α dan derajat bebas $(k-1)$, dengan menolak H$_0$ dapat dikatakan bahwa terdapat heterogenitas spasial (Breuesch & Pagan, 1979 dalam Zakiyah Mar'ah, 2017).

2.9 Pemilihan Pembobot

Pembobot digunakan untuk memberikan nilai yang berbeda disetiap lokasi karena akan berpengaruh pada parameter regresinya. Proses pembobotan dalam penaksiran parameter mengikuti *Tobler's First Law of Geography*, yaitu data yang lebih dekat dengan lokasi *i* akan mempunyai pengaruh yang lebih kuat dalam memprediksi parameter di lokasi ke-*i* dibandingkan data yang lebih jauh. Sebelum pembobot ditentukan, harus dihitung dahulu d_{ij} yang merupakan jarak lokasi (u_i, v_i) menggunakan jarak *euclidian* yaitu (Chasco dkk.2007):

$$d_{ij} = \sqrt{\left(u_i - u_j\right)^2 + \left(v_i - v_j\right)^2}\ . \tag{2.30}$$

Pada penelitian ini, fungsi pembobot spasial yang digunakan adalah *fixed Gaussian kernel. Fixed Gaussian kernel* digunakan karena titik-titik pada data tersebar secara beraturan di wilayah penelitian (Hasriana, 2016). Fungsi *fixed Gaussian kernel* :

$$w_{ij}(u_i, v_i) = exp\left[-\frac{1}{2}\left(\frac{d_{ij}}{h}\right)^2\right], \tag{2.31}$$

dengan :

h : *bandwidth*

d_{ij} : jarak *euclidean* antara lokasi pengamatan ke-*i* dengan lokasi ke-*j*

w_{ij} : nilai bobot dari pengamatan pada lokasi ke-*i* dengan lokasi ke-*j*

u_i : *latitude* pada lokasi ke-*i*

u_j : *latitude* pada lokasi ke-.*j*

v_i : *longitude* pada lokasi ke-*i*

v_j : *longitude* pada lokasi ke-*j*

2.10 Penentuan *Bandwidth*

Fungsi dari *bandwidth* adalah untuk menentukan bobot dari suatu lokasi terhadap lokasi lain yang digunakan sebagai pusat. Semakin dekat wilayah dengan daerah pusat, akan semakin besar pula pengaruh yang diberikan. Ada beberapa metode yang digunakan untuk memilih *bandwidth* optimum, salah satunya adalah validasi silang atau c*ross validation* (CV) dengan rumus:

$$CV(h) = \sum_{i=1}^{n} \left(y_i - \hat{y}_{\neq i}(h)\right)^2, \tag{2.32}$$

dengan:

$\hat{y}_{\neq i}(h)$: nilai penaksir y_i (*fitting value*) dengan pengamatan di lokasi (u_i, v_i) dihilangkan dalam proses penaksiran.

y_i : pengamatan ke-*i*

n : jumlah sampel (Rosa, 2015)

Untuk mendapatkan nilai h yang optimal, maka diperoleh dari h yang menghasilkan nilai CV yang minimum (Fortheringham *et al*. 2002).

2.11 Data Kemiskinan

Kemiskinan merupakan masalah multidimensi, yang bukan hanya mencakup kondisi ekonomi tetapi juga sosial, budaya, dan pollitik. Kemiskinan menjadi masalah utama yang terjadi di setiap provinsi, termasuk Sulawesi Selatan. Sejak 1984 sampai sekarang, pendekatan pengeluaran juga dilakukan Badan Pusat Statistik (BPS) untuk mengukur kemiskinan Indonesia. Dengan pendekatan ini, kemiskinan Indonesia dipandang sebagai ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran, yang kemudian batasan dari sisi pengeluaran inilah disebut sebagai garis kemiskinan (Badan Pusat Statistik, 2018).

Penduduk miskin adalah penduduk yang memiliki rata-rata pengeluaran per kapita per bulan dibawah garis kemiskinan. Garis kemiskinan terdiri dari Garis Kemiskinan Makanan (GKM) dan Garis Kemiskinan Bukan Makanan (GKBM). ...erupakan nilai pengeluaran kebutuhan minimum makanan yang ...an dengan 2100 kalori per kapita perhari. GKBM adalah kebutuhan minimum untuk perumahan, sandang, pendidikan, dan kesehatan (Badan Pusat Statistik, 2018).

Untuk mengukur kemiskinan, BPS menggunakan konsep kemampuan memenuhi kebutuhan dasar (*basic needs approach*). Dengan pendekatan ini, dapat dihitung *Head Count Index* (HCI), yaitu persentase penduduk miskin terhadap total penduduk. HCI secara sederhana mengukur proporsi yang dikategorikan miskin menggunakan rumus:

$$P_0 = \frac{1}{n}\sum_{i=1}^{q}\left[\frac{z - y_i}{z}\right]^0$$

dengan:

$i = 1,2, \dots, q; y_i < z$

P_0 : persentase penduduk miskin (*head count index)*

z : garis kemiskinan

y_i : rata-rata pengeluaran perkapita sebulan penduduk yang berada dibawah garis kemiskinan

q : banyaknya penduduk yang berada di bawah garis kemiskinan

n : jumlah penduduk

Garis Kemiskinan digunakan sebagai batas untuk mengelompokkan penduduk miskin dan tidak miskin. Penduduk miskin adalah penduduk yang memiliki rata-rata pengeluaran perkapita per bulan di bawah garis kemiskinan. Garis kemiskinan Sulawesi Selatan sendiri mengalami peningkatan setiap tahun selama periode 2016-2018. Tahun 2016 sebesar Rp 270.601 dan di tahun 2018 sebesar Rp 306.545. Penduduk miskin di Sulawesi Selatan Maret 2017 berjumlah 813,07 ribu, meningkat sebesar 16,26 ribu jiwa jika dibandingkan kondisi September 2016 yang berjumlah 796,81 ribu jiwa.

Karakteristik kemiskinan Sulawesi Selatan dapat dilihat dari kondisi demografi, pendidikan dan ketenagakerjaan dari kepala rumah tangga; kondisi perumahan; dan persebarannya menurut pulau. Pemahaman mengenai karakteristik kemiskinan penting sebagai dasar dalam penyusunan kebijakan dan program [illegible]san kemiskinan agar tepat sasaran (Badan Pusat Statistik, 2018).