

Universitas Negeri Yogyakarta
 untuk Memenuhi Sebagian Persyaratan guna Memperoleh Gelar Sarjana Sains



Disusun Oleh:
 Noviana Ervin Nur Aini
 10305141029

PROGRAM STUDI MATEMATIKA
 JURUSAN PENDIDIKAN MATEMATIKA
 FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
 UNIVERSITAS NEGERI YOGYAKARTA
 2015

BAB I PENDAHULUAN

1.1. Latar Belakang

Geostatistika adalah ilmu yang merupakan gabungan antara geologi, teknik, matematika, dan statistika (Cressie, 1993). Geostatistika merupakan ilmu yang berfokus pada data spasial. Data spasial merupakan data yang menyajikan posisi goegrafis suatu objek, misal data kandungan mineral pada koordinat-koordinat lokasi. Geostatistika dikembangkan untuk melakukan penghitungan cadangan mineral seperti timah (Sn), emas (Au), dan perak (Ag). Teknik analisis geostatistika didasarkan pada variabel random dengan tujuan untuk mengetahui dan mengestimasi besarnya kandungan mineral pada data spasial tersebut. Dengan menggunakan data spasial yang diperoleh akan diestimasi kandungan mineral di lokasi-lokasi yang lain. Besarnya kandungan mineral pada lokasi yang lain disebut variabel teregional (*regionalized variable*). Variabel teregional mempunyai nilai yang bervariasi seiring berubahnya lokasi (Alfiana, 2010).

Metode *kriging* merupakan suatu metode yang dapat digunakan untuk mengestimasi variabel teregional. Istilah *kriging* diambil dari nama seorang ahli, yaitu D. G. Krige, seorang insinyur pertambangan Afrika Selatan. Metode *kriging* kemudian dikembangkan oleh G. Matheron dalam geostatistika. Metode *kriging* tidak hanya digunakan dalam bidang pertambangan, metode ini juga dikembangkan dalam bidang pertanian, kesehatan, dan sebagainya.

Berbagai metode *kriging* dikembangkan untuk menyelesaikan kasus-kasus yang terdapat pada geostatistika. Estimasi variabel teregional dilakukan dengan

METODE ROBUST KRIGING DAN PENERAPANNYA PADA DATA GEOSTATISTIKA

Oleh:
 Noviana Ervin Nur Aini
 NIM. 10305141029

ABSTRAK

Geostatistika adalah ilmu yang merupakan gabungan antara geologi, teknik, matematika, dan statistika. Geostatistika dikembangkan untuk melakukan perhitungan cadangan mineral. Teknik analisis geostatistika didasarkan pada variabel random pada data spasial. Dengan menggunakan data spasial yang diperoleh akan diestimasi kandungan mineral di lokasi-lokasi yang lain. Besarnya mineral pada lokasi yang di lain disebut variabel teregional. Metode *kriging* merupakan suatu metode yang dapat digunakan untuk mengestimasi variabel teregional. Metode *kriging* dibagi menjadi *simple kriging*, *ordinary kriging*, dan *universal kriging*. Metode *robust kriging* merupakan perkembangan dari metode *ordinary kriging* yang mengakomodir adanya pencilan. Pada tulisan ini akan menjelaskan mengenai sifat-sifat estimator dari metode *kriging* dan menerapkan metode *robust kriging* pada data geostatistika beserta langkah-langkah pengestimasian cadangan hasil tambang dan menerapkannya untuk mengestimasi kandungan timah di lokasi penambangan Q.

Data yang digunakan pada metode *robust kriging* merupakan data spasial yang mengandung pencilan. Langkah-langkah dalam mengestimasi dengan menggunakan metode *robust kriging* yaitu 1) analisis deskriptif, 2) analisis stasioneritas, 3) analisis pencilan, 4) analisis semivariogram eksperimental, 5) analisis struktural yaitu membandingkan nilai *Mean Square Error (MSE)* dari nilai semivariogram eksperimental dengan nilai semivariogram teoritis, 6) menghitung estimasi dengan metode *robust kriging*, dan 7) melakukan plot hasil estimasi.

Metode *kriging* menghasilkan estimator yang bersifat Linear, tak bias, dan meminimalkan variansi *error*. Pada penerapan metode *robust kriging*, data spasial yang digunakan sebanyak 530 data sampel timah pada lokasi penambangan Q, yang terdiri dari titik $X(\text{absis})$, $Y(\text{ordinat})$ dan $Z(\text{kandungan timah (Sn)})$. Berdasarkan hasil analisis diperoleh hasil bahwa data spasial tersebut merupakan data yang stasioner dan memiliki pencilan. Untuk melakukan estimasi data spasial maka selanjutnya dilakukan analisis semivariogram eksperimental dengan menggunakan program SAS 9.1 untuk menentukan nilai *siil* dan *range* yang digunakan untuk analisis semivariogram teoritis. Dari hasil semivariogram eksperimental dan teoritis maka selanjutnya adalah analisis struktural yaitu membandingkan nilai *MSE* dari semivariogram eksperimental dan semivariogram teoritis, dan diperoleh model semivariogram teoritis dengan nilai MSE terkecil yang cocok untuk melakukan estimasi dengan metode *kriging* yaitu semivariogram teoritis model eksponensial. Hasil estimasi cadangan timah dengan metode *kriging* diperoleh nilai kandungan timah minimum sebesar 5.079601 g/m³ dan maksimum sebesar 12.37398 g/m³.

Kata kunci: Data Spasial, *Kriging*, Semivariogram
 vii

metode *ordinary kriging* jika data spasial memenuhi asumsi stasioneritas dengan rata-rata (μ) belum diketahui. (Cressie, 1993), (Alfiana,2010), (Fridayani, 2012), (Theodorick, 2013), (Puspita, 2013). Data spasial dikatakan stasioner apabila data spasial tersebut tidak mengandung *trend*. *Trend* dapat terjadi apabila variabel random spasial pada data spasial membentuk lengkungan (Munadi, 2006), (Laksana, 2010).

Pada metode *kriging* dapat dihasilkan nilai estimasi yang kurang tepat jika pada data spasial yang diestimasi terdapat pencilan spasial (*spatial outlier*) (Cressie, 1993), (Fridayani, 2012), (Theodorick, 2013). Pencilan spasial didefinisikan sebagai nilai lokasi yang tidak konsisten atau sangat menyimpang terhadap nilai lokasi yang lainnya. Terjadinya pencilan dapat disebabkan oleh kesalahan pencatatan, kalibrasi alat yang tidak tepat atau kemungkinan lainnya. Untuk mengatasi masalah tersebut maka metode *ordinary kriging* dikembangkan menjadi metode *robust kriging*. Pada metode *robust kriging* variogram eksperimental yang digunakan adalah variogram *robust* karena variogram *robust* dapat mengakomodir adanya pencilan (Cressie, 1993). Pada estimasi data spasial berpencilan metode *robust kriging* menghasilkan *error* yang nilainya lebih kecil dibandingkan dengan metode *ordinary kriging* sehingga hasil yang diperoleh lebih mendekati dengan hasil sebenarnya (Cressie, 1993), (Fridayani, 2012), (Theodorick, 2013).

Penelitian yang pernah dilakukan dengan menggunakan metode *kriging* diantaranya oleh Alfiana (2010) tentang metode *ordinary kriging* pada geostatistika, untuk mengestimasi data kandungan mineral yang tidak

dengan hasil estimasi menggunakan metode *ordinary kriging*. (Theodorick (2013))

menggunakan metode *kriging* untuk memprediksi *peak ground acceleration* berbasis komputer. Dari hasil analisis dengan menggunakan metode *ordinary kriging* diperoleh hasil yang lebih akurat dibandingkan dengan metode *robust kriging* meskipun perbedaannya sangat sedikit.

Terdapat banyak hal estimasi yang menggunakan data spasial. Salah satunya adalah estimasi kandungan mineral yang ada di bumi. Mineral merupakan suatu zat yang terdapat di alam dengan komposisi kimia yang khas. Salah satu jenis mineral yang terkandung didalam bumi adalah timah (Sn). Timah adalah logam berwarna putih keperakan, dengan kekerasan yang rendah. Kegunaan timah banyak sekali terutama untuk bahan baku logam pelapis, solder, cenderamata, dan lain-lain. Potensi timah di Indonesia terdapat di Pulau Bangka, Pulau Belitung, Pulau Singkep, dan Pulau Karimun.

Untuk mengetahui kadungan mineral yang ada di bumi diperlukan estimasi. Dalam geostatistika khususnya bidang pertambangan, metode yang digunakan untuk mengestimasi kandungan mineral disebut dengan *kriging*. Mengingat pentingnya metode *kriging* untuk mengestimasi kandungan mineral dan kegunaan metode ini untuk menyelesaikan kasus-kasus pada geostatistika, maka dalam tugas akhir ini penulis akan membahas sifat-sifat estimator yang dihasilkan dari

1.5. Manfaat

Manfaat yang diperoleh dari penulisan tugas akhir ini adalah:

1. Penulis dan pembaca dapat mempelajari sifat-sifat estimator yang dihasilkan pada metode *kriging*.
2. Penulis dan pembaca dapat mengetahui penerapan metode *robust kriging* dalam mengestimasi kandungan timah (Sn).

metode *kriging* dan menerapkan metode *robust kriging* untuk mengestimasi kandungan timah (Sn) pada data spasial yang stasioner dan mengandung pencilan di salah satu lokasi penambangan timah.

1.2. Batasan Masalah

Pada tugas akhir ini untuk menghindari masalah yang semakin meluas maka metode *kriging* yang dikembangkan adalah metode *ordinary kriging*. Estimasi dilakukan pada suatu lokasi pertambangan timah dan pengambilan data dilakukan pada tahun 2014. Untuk menjaga kerahasiaan data yang digunakan maka lokasi pertambangan yang digunakan disamarkan menjadi lokasi Q.

1.3. Rumusan Masalah

Berdasarkan latar belakang masalah, maka dapat dirumuskan sebagai berikut?

1. Bagaimana sifat-sifat estimator yang dihasilkan dari metode *kriging*?
2. Bagaimana penerapan metode *robust kriging* dalam mengestimasi kandungan timah (Sn)?

1.4. Tujuan

Berdasarkan rumusan masalah diatas maka tujuan dari penulisan tugas akhir ini adalah:

1. Menjelaskan sifat-sifat estimator yang dihasilkan dari metode *kriging*.
2. Menjelaskan penerapan metode *robust kriging* dalam mengestimasi kandungan timah (Sn).

BAB II LANDASAN TEORI

2.1. Matriks

Matriks merupakan kumpulan angka-angka atau yang disebut elemen yang disusun menurut baris dan kolom sehingga membentuk suatu persegi panjang, dimana panjang suatu matriks ditentukan oleh banyaknya kolom dan lebarnya ditentukan oleh banyaknya baris (Anton, 1995). Suatu matriks dibatasi dengan tanda “[]” atau “()”.

Suatu matriks dinotasikan dengan simbol huruf kapital seperti A, B, atau C dan sebagainya. Suatu matriks yang berukuran $m \times n$ dapat diartikan bahwa matriks tersebut memiliki ukuran panjang atau baris m dan lebar atau kolom n , sehingga dapat ditulis sebagai berikut:

$$A_{m \times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad (2.1.1)$$

atau dapat juga ditulis dengan

$$A = [a_{ij}] \quad i = 1, 2, \dots, m; \quad j = 1, 2, \dots, n \quad (2.1.2)$$

Untuk menyatakan elemen yang terdapat pada baris i dan kolom j yang ada pada suatu matriks A maka digunakan simbol a_{ij} .

Jenis-jenis Matriks:

1. Matriks Kuadrat (Anton, 1995)

$$[a_{n1} \ a_{n2} \ \dots \ a_{nm}]$$

2. Matriks Simetris (Anton, 1995)

Suatu matriks kuadrat $A = [a_{ij}]$; $i, j = 1, 2, \dots, n$ disebut matriks simetris jika elemen dibawah diagonal utama merupakan cermin dari elemen diatas diagonal utama. Matriks simetris jika

$$A^T = A \text{ artinya } a_{ij} = a_{ji} \quad (2.1.4)$$

3. Matriks Identitas (Anton, 1995)

Suatu matriks disebut matriks identitas apabila matriks tersebut merupakan matriks kuadrat dengan

$$A = [a_{ij}] \text{ jika } a_{ij} = 1 \rightarrow i = j; a_{ij} = 0 \rightarrow i \neq j \quad (2.1.5)$$

Matriks identitas dinyatakan dengan I .

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \dots$$

Operasi pada matriks:

1. Perkalian matriks dengan skalar (Anton, 1995)

7

Jika $A = [a_{ij}]$ adalah matriks $m \times n$ maka transpose dari matriks A dinyatakan dalam A^T . A^T adalah matriks $n \times m$ yang diperoleh dari pertukaran antara baris dan kolom pada matriks A yaitu dengan menukar baris pertama pada matriks A dengan kolom pertama pada matriks A , baris kedua pada matriks A dengan kolom kedua pada matriks A , dan seterusnya.

$$A^T = [a_{ji}] \quad (2.1.9)$$

5. Invers matriks (Anton, 1995)

Jika $A = [a_{ij}]$ adalah matriks kuadrat dengan $\det(A) \neq 0$, I suatu matriks identitas, maka A^{-1} merupakan invers dari A dengan $AA^{-1} = A^{-1}A = I$. Sifat dari invers yaitu:

$$a. (A^{-1})^{-1} = A \quad (2.1.10)$$

$$b. (AB)^{-1} = B^{-1}A^{-1} \quad (2.1.11)$$

$$c. (cA)^{-1} = \frac{1}{c}A^{-1}, \text{ untuk suatu skalar } c \in \mathbb{R} \neq 0. \quad (2.1.12)$$

Dalam menyelesaikan kasus pada estimasi data spasial, matriks digunakan untuk menentukan nilai bobot dari masing-masing lokasi tersampel terhadap lokasi yang akan di estimasi.

2.2. Variabel Random

Variabel random Z didefinisikan sebagai suatu fungsi yang memetakan unsur-unsur dalam ruang sampel S dengan titik sampel $s \in S$. Variabel random dinotasikan dengan huruf kapital Z dan huruf kecil z yang menyatakan nilai dari variabel random tersebut.

9

Jika $A = [a_{ij}]$ adalah matriks $m \times n$ dan c adalah suatu skalar, maka hasil kali A dengan c adalah matriks yang diperoleh dengan mengalikan masing-masing entri dari A oleh c .

$$c[a_{ij}] = [ca_{ij}] \quad (2.1.6)$$

2. Penjumlahan matriks (Anton, 1995)

Jika $A = [a_{ij}]$ adalah matriks $m \times n$ dan $B = [b_{ij}]$ adalah matriks $m \times n$ maka penjumlahan dari matriks A dan matriks B yang ditulis dengan $C = [c_{ij}]$.

$$c_{ij} = a_{ij} + b_{ij} \text{ dimana } i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (2.1.7)$$

3. Perkalian matriks dengan matriks (Anton, 1995)

Jika $A = [a_{ij}]$ adalah matriks $m \times p$ dan $B = [b_{jk}]$ adalah matriks $p \times n$ maka hasil kali dari matriks A dan matriks B yang ditulis dengan AB adalah C matriks $m \times n$.

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mp} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \dots & b_{pn} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + \dots + a_{1p}b_{p1} & \dots & a_{11}b_{1n} + a_{12}b_{2n} + \dots + a_{1p}b_{pn} \\ a_{21}b_{11} + a_{22}b_{21} + \dots + a_{2p}b_{p1} & \dots & a_{21}b_{1n} + a_{22}b_{2n} + \dots + a_{2p}b_{pn} \\ \vdots & \ddots & \vdots \\ a_{m1}b_{11} + a_{m2}b_{21} + \dots + a_{mp}b_{p1} & \dots & a_{m1}b_{1n} + a_{m2}b_{2n} + \dots + a_{mp}b_{pn} \end{bmatrix}$$

secara matematik dapat ditulis sebagai berikut:

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \dots + a_{ip}b_{pk} \quad (2.1.8)$$

4. Transpose matriks (Anton, 1995)

8

2.2.1. Variabel Random Kontinu (Ross, 2007)

Definisi 2.2.1

Suatu variabel random Z disebut variabel random kontinu jika terdapat fungsi f yang didefinisikan untuk semua bilangan real $z \in (-\infty, \infty)$, memiliki persamaan bahwa, untuk setiap himpunan B dari bilangan real:

$$P\{Z \in B\} = \int_B f(z)dz \quad (2.2.1)$$

Dengan fungsi f merupakan fungsi densitas peluang (fdp) dari variabel random Z . Persamaan (2.2.1) menyatakan bahwa peluang Z pada B dapat diperoleh dengan mengintegrasikan fungsi densitas peluang himpunan B . Dimana Z diasumsikan dengan beberapa nilai, maka f harus memenuhi,

$$P\{Z \in (-\infty, \infty)\} = \int_{-\infty}^{\infty} f(z)dz = 1 \quad (2.2.2)$$

Kemudian untuk $B = [a, b]$, diperoleh

$$P\{a \leq Z \leq b\} = \int_a^b f(z)dz \quad (2.2.3)$$

Jika $a = b$ maka persamaan (2.2.2) diperoleh

$$P\{Z = a\} = \int_a^a f(z)dz = 0 \quad (2.2.4)$$

10

$$P\{Z \leq a_j\} = P\{Z \leq a_j - \int_{-\infty}^a f(z)u(z)dz\} \quad (2.2.5)$$

2.2.2. Ekspektasi (Tirta, 2004)

Definisi 2.2.2

Jika Z adalah sebuah variabel random kontinu dengan fungsi densitas peluang (fdp) di z adalah $f(z)$, dan u adalah fungsi dari Z , maka nilai ekspektasi dari $u(z)$ yang dinotasikan dengan $E[u(Z)]$, didefinisikan sebagai

$$E[u(Z)] = \int_{-\infty}^{\infty} u(z)f(z)dz \quad (2.2.6)$$

Nilai ekspektasi dari $u(z)$ memenuhi sifat-sifat berikut:

a. Jika $u(Z) = m$ dan m adalah konstanta maka $E[Z] = E(m) = m$;

Bukti:

$$\begin{aligned} E[m] &= \int_{-\infty}^{\infty} mf(z)dz \\ &= m \int_{-\infty}^{\infty} f(z)dz \\ &= m \end{aligned} \quad (2.2.7)$$

11

$$\begin{aligned} &= m_1 \int_{-\infty}^{\infty} u_1(Z)f(z)dz \pm m_2 \int_{-\infty}^{\infty} u_2(Z)f(z)dz \\ &= m_1 E(Z_1) \pm m_2 E(Z_2) \end{aligned} \quad (2.2.10)$$

c. $E[u_1(Z)u_2(Z)] = E[u_1(Z)]E[u_2(Z)]$

Bukti:

$$\begin{aligned} E[u_1(Z_1)u_2(Z_2)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u_1(z_1)u_2(z_2)f(z_1, z_2)dz_1 dz_2 \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u_1(z_1)u_2(z_2)f_1(z_1)f_2(z_2)dz_1 dz_2 \\ &= \left[\int_{-\infty}^{\infty} u_1(z_1)f_1(z_1)dz_1 \right] \left[\int_{-\infty}^{\infty} u_2(z_2)f_2(z_2)dz_2 \right] \\ &= E[u_1(Z_1)]E[u_2(Z_2)] \end{aligned} \quad (2.2.11)$$

2.2.3. Variansi

Suatu variabel random diasumsikan terdistribusi normal, sehingga nilai ekspektasi pada variabel random dapat menentukan ukuran penyebaran atau variansi, yang didefinisikan sebagai berikut:

Definisi 2.3.1 (Bain & Engelhart, 1992)

Variansi dari variabel random Z didefinisikan sebagai

$$\text{var}(Z) = E[Z - E(Z)]^2 \quad (2.2.12)$$

Sifat-sifat dari variansi ditunjukkan pada beberapa teorema berikut:

13

b. $E[mu(Z)] = mE[u(Z)]$;

Bukti:

$$\begin{aligned} E[mu(Z)] &= \int_{-\infty}^{\infty} mu(z)f(z)dz \\ &= m \int_{-\infty}^{\infty} u(z)f(z)dz \\ &= mE[u(Z)] \end{aligned} \quad (2.2.8)$$

c. $E\{u_1(Z) \pm u_2(Z)\} = Eu_1(Z) \pm Eu_2(Z)$;

Bukti:

$$\begin{aligned} E\{u_1(Z) \pm u_2(Z)\} &= \int_{-\infty}^{\infty} \{u_1(z) \pm u_2(z)\}f(z)dz \\ &= \int_{-\infty}^{\infty} \{u_1(z)f(z) \pm u_2(z)f(z)\}dz \\ &= \int_{-\infty}^{\infty} \{u_1(z)f(z)dz \pm \int_{-\infty}^{\infty} \{u_2(z)f(z)dz\} \\ &= Eu_1(Z) \pm Eu_2(Z) \end{aligned} \quad (2.2.9)$$

d. $E\{m_1u_1(Z) \pm m_2u_2(Z)\} = m_1E(Z_1) \pm m_2E(Z_2)$

Bukti:

$$\begin{aligned} E\{m_1u_1(Z) \pm m_2u_2(Z)\} &= \int_{-\infty}^{\infty} \{m_1u_1(Z) \pm m_2u_2(Z)\}f(z)dz \\ &= \int_{-\infty}^{\infty} \{m_1u_1(Z)\}f(z)dz \pm \int_{-\infty}^{\infty} \{m_2u_2(Z)\}f(z)dz \end{aligned}$$

12

Teorema 2.3.1 (Bain & Engelhart, 1992)

Jika Z merupakan suatu variabel random, maka:

$$\text{var}(Z) = E(Z^2) - [E(Z)]^2 \quad (2.2.13)$$

Bukti:

$$\begin{aligned} \text{var}(Z) &= \int_{-\infty}^{\infty} [Z - E(Z)]^2 f(z)dz \\ &= \int_{-\infty}^{\infty} [z^2 - 2zE(Z) + (E(Z))^2]f(z)dz \\ &= \int_{-\infty}^{\infty} z^2 f(z)dz - 2E(Z) \int_{-\infty}^{\infty} z f(z)dz + [E(Z)]^2 \int_{-\infty}^{\infty} f(z)dz \\ &= E(Z^2) - 2E(Z)E(Z) + [E(Z)]^2 \\ &= E(Z^2) - [E(Z)]^2 \end{aligned}$$

Teorema 2.3.2 (Bain & Engelhart, 1992)

Jika Z variabel random, a dan b konstanta maka:

$$\text{var}(aZ + b) = a^2 \text{var}(Z) \quad (2.2.14)$$

Bukti:

$$\begin{aligned} \text{var}(aZ + b) &= \int_{-\infty}^{\infty} [az - aE(Z)]^2 f(z)dz + 0 \\ &= \int_{-\infty}^{\infty} [a^2 z^2 - 2a^2 zE(Z) + a^2 [E(Z)]^2]f(z)dz \\ &= \int_{-\infty}^{\infty} a^2 [z^2 - 2zE(Z) + [E(Z)]^2]f(z)dz \end{aligned}$$

14

Kovariansi adalah ukuran korelasi dari dua buah data yang berbeda.

Kovariansi menentukan nilai dari dua buah variabel random dalam bervariasi.

Definisi 2.4.1 (Bain & Engelhardt, 1992)

Nilai kovariansi dari variabel random X dan Y didefinisikan oleh:

$$Cov(X, Y) = E[(X - E(X))(Y - E(Y))] \quad (2.2.15)$$

Kovariansi antara variabel random X dan Y dinotasikan dengan σ_{XY}

Sifat mengenai kovariansi dijelaskan dalam beberapa teorema berikut:

Teorema 2.4.1 (Bain & Engelhart, 1992)

Apabila X dan Y merupakan suatu variabel random, dimana a dan b adalah konstanta, maka:

$$a. \quad Cov(aX, bY) = ab \, Cov(X, Y)$$

Bukti:

$$\begin{aligned} Cov(aX, bY) &= E[a(X - E(X))][b(Y - E(Y))] \\ &= ab \, E[(X - E(X))(Y - E(Y))] \\ &= ab \, Cov(X, Y) \end{aligned} \quad (2.2.16)$$

15

Bukti:

$$\begin{aligned} Cov(X, Y) &= E(XY) - E(X)E(Y) \\ &= E(X)E(Y) - E(X)E(Y) \\ &= 0 \end{aligned} \quad (2.2.20)$$

Teorema 2.4.3 (Bain & Engelhart, 1992)

Jika X dan Y adalah variabel random, maka

$$var(X + Y) = var(X) + var(Y) + 2 \, cov(X, Y) \quad (2.2.21)$$

dan

$$var(X + Y) = var(X) + var(Y) \quad (2.2.22)$$

dengan X dan Y adalah variabel yang independen.

Bukti:

$$\begin{aligned} var(X + Y) &= E[(X + Y) - (E(X) + E(Y))]^2 \\ &= E[(X - E(X)) + (Y - E(Y))]^2 \\ &= E[(X - E(X))^2 + 2E(X - E(X))(Y - E(Y)) + (Y - E(Y))^2] \\ &= var(X) + var(Y) + 2 \, cov(X, Y) \end{aligned}$$

Untuk variabel random X dan Y yang independen maka $cov(X, Y) = 0$ sehingga,

$$var(X + Y) = var(X) + var(Y)$$

2.3. Data Spasial

Data spasial adalah data yang disajikan dalam posisi geografis dari suatu obyek, berkaitan dengan lokasi dalam ruang bumi. Data spasial merupakan data dependen terhadap lokasi, karena berasal dari lokasi spasial yang berbeda yang mengindikasikan adanya keterkaitan antara pengukuran dengan lokasi (Cressie,

17

$$b. \quad Cov(X + a, Y + b) = Cov(X, Y)$$

Bukti:

$$\begin{aligned} Cov(X + a, Y + b) &= E[(X + a - E(X + a))][(Y + b - E(Y + b))] \\ &= E[(X + a - a - E(X))][(Y + b - b - E(Y))] \\ &= E[(X - E(X))(Y - E(Y))] \\ &= Cov(X, Y) \end{aligned} \quad (2.2.17)$$

$$c. \quad cov(X, aX + b) = a \, Var(X)$$

Bukti:

$$\begin{aligned} Cov(X, aX + b) &= E[(X - E(X))[(aX + b) - E(aX + b)]] \\ &= E[(X - E(X))(aX + b - aE(X) - b)] \\ &= aE(X - E(X))(X - E(X)) \\ &= a \, Var(X) \end{aligned} \quad (2.2.18)$$

Teorema 2.4.2 (Bain & Engelhart, 1992)

Apabila X dan Y merupakan suatu variabel random, kemudian

$$Cov(X, Y) = E(XY) - E(X)E(Y) \quad (2.2.19)$$

Bukti:

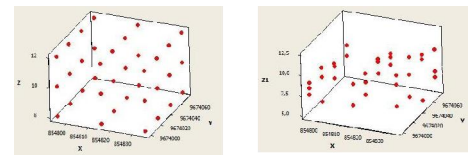
$$\begin{aligned} Cov(X, Y) &= E[(X - E(X))(Y - E(Y))] \\ &= E(XY - XE(Y) - E(X)Y + E(X)E(Y)) \\ &= E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

Kemudian jika X dan Y yang independen maka $Cov(X, Y) = 0$

16

1993), (Umbara, 2007). Data spasial merupakan data yang memuat informasi lokasi, dengan $z(s_i)$ adalah nilai observasi pada lokasi i atau koordinat yang dinyatakan dengan vektor (s_i) .

Data spasial dapat memiliki lokasi spasial beraturan (*regular*) atau tak beraturan (*irregular*). Data spasial disebut memiliki lokasi spasial beraturan (*regular*) apabila pada lokasi yang berdekatan memiliki posisi yang beraturan dengan jarak sama besar, sedangkan lokasi spasial tak beraturan (*irregular*) apabila pada lokasi yang berdekatan memiliki posisi yang tidak beraturan dengan jarak yang berbeda (Anantia, 2010). Gambar 2.1 merupakan salah satu contoh data spasial pada ruang dimensi 3



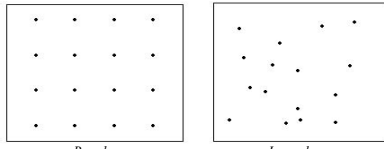
Regular
Irregular
Gambar 2.1 Contoh plot data spasial pada ruang dimensi 3

Nilai acak di suatu lokasi s dinotasikan dengan $Z(s)$ dimana $Z(s)$ merupakan variabel random spasial dan $s \in \mathbb{R}^d$ merupakan lokasi data dalam dimensi- d ruang euclidean dan $Z(s)$ merupakan variabel random pada lokasi $s \in D$, karena selain dapat dinyatakan dalam koordinat satu dimensi, lokasi juga dapat dinyatakan dalam koordinat dua atau tiga dimensi. Lambang D merupakan simbol

18

1. Data Geostatistika (Geostatistical Data)

Pada awal tahun 1980 geostatistika dikenal sebagai disiplin dari geologi, teknik, matematika dan statistika (Cressie, 1993). Kelebihan dari geostatistika adalah geostatistika mampu memodelkan kecenderungan spasial maupun korelasi spasial dengan lebih baik. Data geostatistika dari setiap sampel titik didefinisikan oleh lokasi dan nilai pengukuran objek yang diamati. Data geostatistika mengarah pada data sampel yang berupa titik, baik beraturan (*regular*) maupun tak beraturan (*irregular*) dari suatu distribusi spasial kontinu (Cressie, 1993). Prinsip dasar geostatistika adalah bahwa area yang saling berdekatan cenderung memiliki nilai bobot yang tidak jauh berbeda dibandingkan dengan area yang tidak berdekatan (Alfiana, 2010). Data geostatistik dapat diilustrasikan seperti gambar 2.2 dibawah ini



Gambar 2.2 Contoh plot Data Geostatistika *regular* dan *irregular*.

2. Data Area (Lattice Data)

19

penentuan posisi pohon-pohon dengan ukuran tertentu. Apakah pohon-pohon tersebut membentuk pola keteracakan spasial, *clustering*, atau keteraturan. Analisis data pola titik dilakukan karena untuk mengetahui apakah daerah titik yang menjadi objek penelitian membentuk daerah beraturan, sehingga dapat diketahui apakah terjadi ketergantungan antar titik atau tidak.

Pada estimasi data spasial, teknik analisis data geostatistika bertujuan untuk mengetahui dan mengestimasi nilai variabel teregional pada lokasi s . Nilai dari suatu variabel yang diamati dapat dinyatakan sebagai variabel random spasial $Z(s)$ dengan s adalah vektor lokasi $D \in \mathbb{R}^d$. Metode *kriging* merupakan metode untuk mengestimasi nilai dari variabel teregional $\hat{Z}(s)$ pada suatu lokasi s berdasar variabel random spasial $Z(s)$. Variabel random spasial $Z(s)$ pada data geostatistika merupakan suatu variabel random Z di lokasi s .

Pada data spasial variabel random X didefinisikan sebagai variabel random spasial $Z(s)$ di lokasi s dan variabel random Y didefinisikan sebagai variabel random spasial $Z(s+h)$ di lokasi $s+h$.

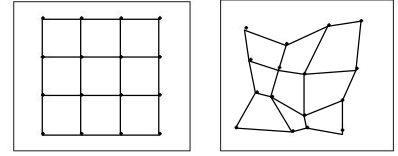
Pada analisis data geostatistika variansi digunakan untuk menentukan korelasi antara variabel random spasial $Z(s)$ dan $Z(s+h)$. Nilai variansi dari variabel random spasial $Z(s)$ dan $Z(s+h)$ pada lokasi s dan $s+h$ dapat ditentukan sebagai berikut:

$$\begin{aligned} \text{var}(z(s) - z(s+h)) &= \text{var}(z(s)) + \text{var}(z(s+h)) - 2\text{cov}(z(s), z(s+h)) \\ &= E \left[(Z(s) - Z(s+h)) - (E(Z(s)) - E(Z(s+h))) \right]^2 \end{aligned}$$

21

Data area merupakan konsep dari garis yang teratur dalam ruang \mathbb{R}^d , terkait dengan titik terdekat, titik terdekat kedua, dan seterusnya yang disebut sebagai data area beraturan dan data tak beraturan yang didukung oleh informasi lingkungan dan batas-batas tertentu. Data area berhubungan dengan daerah spasial karena merupakan kumpulan data diskrit yang merupakan hasil pengukuran pada wilayah tertentu (Cressie, 1993). Data untuk setiap area didefinisikan oleh lokasi dan nilai pengukurannya.

Pada umumnya data area digunakan untuk studi *epidemiologi*, untuk mengetahui pertumbuhan suatu penyakit pada suatu wilayah yang terbagi menjadi area-area tertentu. Data area dapat diilustrasikan seperti gambar 2.3 dibawah ini



Gambar 2.3 Contoh plot data area *regular* dan *irregular*.

3. Pola titik (point pattern)

Pola titik (*point pattern*) merupakan pola yang muncul dari variabel yang dianalisis pada daerah tersampel (Cressie, 1993). Sampel yang digunakan adalah sampel yang tidak beraturan atau sampel yang memiliki jarak berbeda. Daerah tersebut diperoleh dari data koordinat kartesius (x,y) dari titik yang diamati. Data pola titik spasial dapat diperoleh dari informasi apakah pola yang diperoleh menggambarkan keteracakan spasial, *clustering*, atau keteraturan. Misalnya

20

$$\begin{aligned} &= E \left((Z(s) - Z(s+h))^2 \right) + \left(E(Z(s) - Z(s+h)) \right)^2 \\ &\quad - 2E(Z(s) - Z(s+h))E(Z(s)) - E(Z(s+h)) \\ &= E \left((Z(s) - Z(s+h))^2 \right) + \left(E(Z(s) - Z(s+h)) \right)^2 \\ &\quad - 2 \left(E(Z(s) - Z(s+h)) \right)^2 \\ &= E(Z(s) - Z(s+h))^2 - \left(E(Z(s) - Z(s+h)) \right)^2 \quad (2.3.1) \end{aligned}$$

Nilai kovariansi dari variabel spasial $Z(s)$ dan $Z(s+h)$ di lokasi s dan $s+h$ yang bergantung pada jarak h dapat didefinisikan sebagai:

$$\text{Cov}(Z(s), Z(s+h)) = E \left[(Z(s) - E(Z(s))) (Z(s+h) - E(Z(s+h))) \right] \quad (2.3.2)$$

2.4. Stasioneritas pada Data Spasial

Dalam geostatistika analisis stasioneritas dilakukan untuk menentukan analisis data spasial lebih lanjut. Suatu data spasial dikatakan memiliki sifat stasioner apabila data spasial tersebut tidak mengandung *trend*. Pada geostatistika terdapat 3 macam stasioneritas, yaitu (Cressie, 1993):

1. Stasioner Kuat (*Stricly Stationarity*) (Cressie, 1993)

Suatu variabel random spasial $Z(s)$ dikatakan stasioner kuat apabila untuk sebarang koordinat titik s_1, s_2, \dots, s_n dan untuk sebarang vektor h yang berdimensi sama dengan s_i , maka berlaku

$$F(z(s_1), z(s_2), \dots, z(s_n)) = F(z(s_1+h), z(s_2+h), \dots, z(s_n+h)) \quad (2.4.1)$$

22

Variabel random $Z(s)$ dikatakan stasioneritas orde kedua Jika:

- a. Rata-rata untuk semua lokasi, didefinisikan dengan

$$E[Z(s)] = m(s) = m \text{ untuk semua } s \in \mathbb{D} \quad (2.4.2)$$

berarti untuk semua lokasi s memiliki nilai rata-rata yang sama, sehingga mengakibatkan $E[Z(s)] = E[Z(s+h)]$.

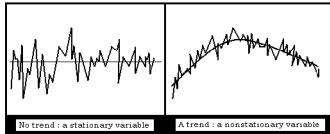
- b. Nilai kovariansi $Cov(Z(s), Z(s+h))$ antara lokasi s dan $s+h$ ada, dan hanya tergantung pada jarak h , didefinisikan sebagai berikut:

$$\begin{aligned} Cov(Z(s), Z(s+h)) &= E[Z(s) - m](Z(s+h) - m) \text{ untuk semua } h \\ &= E[Z(s)Z(s+h)] - m^2 \end{aligned} \quad (2.4.3)$$

Kovariansi untuk dua data yang berjarak 0, atau $h = 0$ didefinisikan sebagai berikut:

$$\begin{aligned} Cov(Z(s), Z(s)) &= E[(Z(s) - m)(Z(s) - m)] \\ &= E[Z(s)^2 - 2mZ(s) + m^2] \\ &= E[Z(s)^2 - 2E(Z(s))(Z(s)) + (E(Z(s)))^2] \\ &= E[Z(s)^2 - 2[E(Z(s))]^2 + (E(Z(s)))^2] \end{aligned}$$

23



Gambar 2.4 Contoh plot stasioneritas data spasial

Dari gambar 2.5 dapat dijelaskan bahwa grafik sebelah kiri merupakan grafik yang menyatakan bahwa data spasial merupakan data stasioner, sedangkan grafik sebelah kanan menyatakan bahwa data spasial merupakan data nonstasioner karena variabel spasial pada data spasial tersebut membentuk lengkungan.

2.5. Variogram dan Semivariogram

Pada data spasial dua buah nilai spasial yang letaknya berdekatan memiliki kecenderungan lebih besar untuk bernilai sama dibandingkan dengan dua buah nilai spasial yang letaknya berjauhan, oleh karena itu untuk melakukan estimasi pada data spasial diperlukan perangkat statistik berupa variogram. Variogram adalah karakteristik dari korelasi spasial, yaitu korelasi antara dua buah data spasial tersebut menjadi kurang berkorelasi ataupun tidak berkorelasi yang disebabkan bertambahnya jarak dari data yang diambil (Cressie, 1993).

Variogram didefinisikan sebagai berikut (Cressie, 1993):

$$\begin{aligned} 2\gamma(h) &= var(Z(s) - Z(s+h)) \\ 2\gamma(h) &= E[Z(s) - Z(s+h)]^2 - [E[Z(s) - Z(s+h)]]^2 \\ 2\gamma(h) &= E[Z(s) - Z(s+h)]^2 \end{aligned} \quad (2.5.1)$$

25

$$\begin{aligned} &= E[Z(s)^2 - (E(Z(s)))^2] \\ &= Var(Z(s)) \\ &= \sigma^2 \end{aligned} \quad (2.4.4)$$

sehingga nilai kovariansi dua data yang berjarak 0 sama dengan variansi dari populasi.

3. Stasioneritas intrinsik (Cressie, 1993)

Definisi 2.4.1

Variabel spasial $Z(s)$ pada lokasi $s \in \mathbb{D}$ dikatakan stasioner intrinsik apabila untuk sebarang jarak h mempunyai nilai ekspektasi antara lokasi s dan $s+h$ mendekati nol

$$E[Z(s) - Z(s+h)] = 0 \quad (2.4.5)$$

Berdasarkan definisi 2.4.1 dan persamaan 2.3.1 dapat diperoleh:

$$\begin{aligned} var[Z(s) - Z(s+h)] &= E\{[Z(s) - Z(s+h)]^2\} - \{E[Z(s) - Z(s+h)]\}^2 \\ var[Z(s) - Z(s+h)] &= E[Z(s) - Z(s+h)]^2 \end{aligned} \quad (2.4.6)$$

Pada analisis geostatistika, suatu data spasial dikatakan stasioner apabila pada data spasial tersebut tidak mengandung *trend*. Sedangkan suatu data spasial dikatakan nonstasioner apabila data spasial tersebut mengandung suatu *trend* yaitu dimana variabel pada data spasial tersebut membentuk suatu lengkungan (Suprajitno, 2006). Kestasioneran data spasial dapat dilihat dalam gambar 2.4 (Suprajitno, 2006):

24

Pada estimasi data spasial, suatu perangkat yang digunakan untuk menggambarkan, memodelkan, dan menghitung korelasi spasial antara variabel random $Z(s)$ dan $Z(s+h)$ disebut semivariogram (Cressie, 1993). Besarnya nilai semivariogram merupakan setengah dari nilai variogram (Cressie, 1993), (Umbara, 2007) dan (Ghanim, 2013). Semivariogram didefinisikan sebagai berikut (Cressie, 1993):

$$\gamma(h) = \frac{1}{2} E[Z(s) - Z(s+h)]^2 \quad (2.5.2)$$

Variogram terdiri atas:

1. *Sill*

Sill adalah saat dimana nilai *semivariogram* cenderung mencapai nilai yang stabil(Cressie, 1993). Nilai *sill* sama dengan nilai variansi dari data spasial.

2. *Range*

Range merupakan jarak pada saat semivariogram mencapai nilai *sill* (Cressie, 1993).

3. *Nugget Effect*

Nugget Effect merupakan kediskontinuan pada pusat semivariogram terhadap garis vertical yang melompat dari nilai 0 pada pusat ke nilai semivariogram pada pemisahan jarak terkecil (Cressie, 1993).

Gambar 2.6 berikut merupakan ilustrasi dari semivariogram

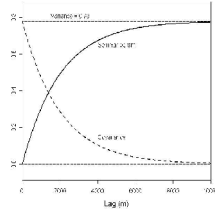
26

Dari gambar diatas dapat dijelaskan bahwa kenaikan nilai $\gamma(h)$ akan berlangsung selama masih terdapat pengaruh nilai h antar variabel random spasial. Daerah ini dinamakan daerah pengaruh suatu variabel random spasial, sampai akhirnya konstan di suatu nilai C (*sill*) yang merupakan variansi dari variabel random spasial. Daerah pengaruh suatu variabel random spasial mempunyai jarak (α) yang disebut *range*, di luar jarak ini maka rata-rata variansi nilai $Z(s)$ dan $Z(s+h)$ tidak lagi tergantung dengan jarak, dengan kata lain $Z(s)$ dan $Z(s+h)$ tidak berkorelasi satu dengan yang lainnya.

Untuk suatu variabel spasial $Z(s)$ yang memenuhi asumsi stasioner orde dua, terdapat hubungan antara semivariogram dengan fungsi kovariansi sebagai berikut:

$$\begin{aligned}
 2\gamma(h) &= \text{var}[Z(s) - Z(s+h)] \\
 &= E[Z(s) - Z(s+h)]^2 \\
 &= E[\{Z(s)\}^2 - 2Z(s+h)Z(s) + \{Z(s+h)\}^2] \\
 &= E[Z(s)]^2 - 2E[Z(s+h)Z(s)] + E[Z(s+h)]^2 \\
 &= \text{var}[Z(s)] + [E[Z(s)]]^2 + 2E[Z(s+h)Z(s)] + \text{var}[Z(s+h)] + \\
 &\quad [E[Z(s+h)]]^2
 \end{aligned}$$

27



Gambar 2.5 Contoh plot hubungan Semivariogram dan Kovariansi

2.5.1. Variogram dan Semivariogram Eksperimental

Variogram eksperimental merupakan variogram yang diperoleh dari hasil pengukuran data spasial yang ada di lapangan. Variogram eksperimental dibuat berdasarkan nilai korelasi spasial antara dua buah variabel yang dipisahkan oleh suatu jarak tertentu sebesar h . Variogram eksperimental dirumuskan sebagai berikut (Cressie, 1993):

$$2\gamma(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} [Z(s_i) - Z(s_i+h)]^2 \quad (2.5.5)$$

Dan semivariogram eksperimental dirumuskan sebagai berikut (Cressie 1993):

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(s_i) - Z(s_i+h)]^2 \quad (2.5.6)$$

Bukti:

$$2\gamma(h) = E[Z(s) - Z(s+h)]^2$$

$$\begin{aligned}
 &= 2 \text{var}[Z(s)] - 2E[\{Z(s) - E[Z(s)]\} \{Z(s+h) - E[Z(s+h)]\}] \\
 &= 2 \text{var}[Z(s)] - 2[E[Z(s)Z(s+h)] - m^2] \\
 &= 2\sigma^2 - 2\text{Cov}(Z(s), Z(s+h)) \\
 \gamma(h) &= \sigma^2 - \text{Cov}(Z(s), Z(s+h)) \quad (2.5.3)
 \end{aligned}$$

Karena $\sigma^2 = \text{Cov}(Z(s), Z(s))$ sehingga hubungan antara kovariansi dengan semivariogram dapat dituliskan sebagai berikut:

$$\begin{aligned}
 \text{Var}[Z(s)] &= \text{Cov}(Z(s), Z(s)) = \sigma^2 \\
 \gamma(h) &= \text{Cov}(Z(s), Z(s)) - \text{Cov}(Z(s), Z(s+h)) \quad (2.5.4)
 \end{aligned}$$

Bukti

$$\begin{aligned}
 2\gamma(h) &= E[Z(s) - Z(s+h)]^2 \\
 &= E[\{Z(s)\}^2 - 2Z(s+h)Z(s) + \{Z(s+h)\}^2] \\
 &= E[Z(s)]^2 - 2E[Z(s+h)Z(s)] + E[Z(s+h)]^2 \\
 &= \sigma^2 + \sigma^2 - 2\text{Cov}(Z(s), Z(s+h)) \\
 &= 2\sigma^2 - 2\text{Cov}(Z(s), Z(s+h)) \\
 &= 2\text{Cov}(Z(s), Z(s)) - 2\text{Cov}(Z(s), Z(s+h)) \\
 \gamma(h) &= \text{Cov}(Z(s), Z(s)) - \text{Cov}(Z(s), Z(s+h))
 \end{aligned}$$

Dari persamaan diatas maka grafik hubungan antara semivariogram dan kovariansi saling bertolak belakang. Pada saat semivariogram bergerak dari nilai rendah ke nilai tinggi, maka nilai kovariansi bergerak sebaliknya, seperti yang diilustrasikan pada gambar 2.5 berikut:

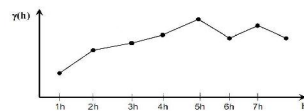
28

Misalkan $V = [Z(s) - Z(s+h)]^2$ dan diketahui bahwa nilai ekspektasi untuk semua lokasi adalah sama, sehingga menurut persamaan (2.4.2) $E(V) = m$ sehingga diperoleh

$$\begin{aligned}
 2\gamma(h) &= E[V] \\
 &= \frac{1}{N(h)} \sum_{i=1}^{N(h)} V \\
 &= \frac{1}{N(h)} \sum_{i=1}^{N(h)} [Z(s_i) - Z(s_i+h)]^2 \\
 \gamma(h) &= \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(s_i) - Z(s_i+h)]^2
 \end{aligned}$$

dengan,
 $2\gamma(h)$: nilai variogram dengan jarak h ,
 $\gamma(h)$: nilai semivariogram dengan jarak h ,
 $Z(s_i)$: nilai pengamatan di titik s_i ,
 $Z(s_i+h)$: nilai pengamatan di titik s_i+h ,
 $N(h)$: banyaknya pasangan titik yang mempunyai jarak h .

Gambar 2.7 berikut merupakan ilustrasi plot dari perhitungan semivariogram eksperimental pada suatu koordinat kartesian antar jarak antar pasangan data spasial (h) dan semivariogram $\gamma(h)$



Gambar 2.7 Contoh plot semivariogram eksperimental

Dari gambar diatas dapat dijelaskan bahwa pada arah tertentu terdapat n buah data dengan jarak tertentu sebesar h, dimana terdapat (n-1) pasangan data untuk

29

30

variogram *robust* merupakan variogram yang mengakomodir adanya pencilan. Variogram *robust* didefinisikan sebagai berikut (Cressie,1993):

$$2\gamma(h) = \frac{\left\{ \frac{1}{[N(h)]} \sum_{i=1}^{N(h)} [Z(s_i) - Z(s_i + h)]^2 \right\}^{\frac{1}{2}}}{\left(0.457 + \frac{0.494}{[N(h)]} \right)} \quad (2.5.7)$$

salah satu parameter untuk mengestimasi variogram *robust* untuk data yang berdistribusi normal adalah median, sehingga estimator variogram *robust* dapat didefinisikan sebagai berikut, (Cressie, 1993).

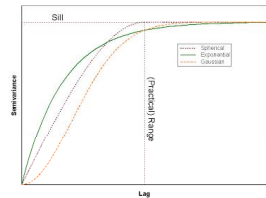
$$2\gamma(h) = \frac{\left[\text{median} \left\{ |Z(s_i) - Z(s_i + h)|^{\frac{1}{2}}; (s_i, s_i + h) \in N(h) \right\} \right]^{\frac{1}{2}}}{B(h)} \quad (2.5.8)$$

Dengan $B(h)$ merupakan koreksi bias dan nilai $B(h) \approx 0.457$.

2.5.3. Semivariogram Teoritis

Untuk analisis lebih lanjut dalam estimasi data geostatistika diperlukan analisis struktural, yaitu membandingkan nilai *MSE* antara semivariogram eksperimental dengan semivariogram teoritis, dan dari perbandingan tersebut dipilih model semivariogram teoritis dengan nilai *MSE* paling kecil untuk menentukan bobot yang digunakan untuk melakukan estimasi data spasial. Berikut ini adalah beberapa model semivariogram teoritis yang digunakan sebagai

31



Gambar 2.8 Contoh plot semivariogram teoritis

2.6. Kriging

Kriging adalah suatu metode untuk mengestimasi variabel tereregional dengan pendekatan bahwa data yang dianalisis merupakan variabel random, dari variabel random tersebut maka akan membentuk suatu fungsi random menggunakan model struktural variogram (Alfiana, 2010). Istilah *kriging* diambil dari nama seorang ahli, yaitu D. G. Kridge, seorang insinyur pertambangan Afrika Selatan. Metode *Kriging* dikembangkan dalam bidang geostatistika oleh G. Matheron pada tahun 1960-an. Metode *kriging* kemudian dikembangkan untuk mengestimasi data geologi yang terdistribusi secara spasial (Martin, 2010). *Kriging* merupakan metode yang secara umum digunakan untuk menganalisis data geostatistik, yaitu metode yang secara umum digunakan untuk menganalisis data geostatistik, yaitu untuk mengestimasi kandungan mineral berdasarkan data sampel. Data sampel biasanya diambil dari lokasi-lokasi atau titik-titik yang tidak beraturan. Metode ini digunakan untuk mengestimasi besarnya nilai variabel tereregional \hat{Z} pada titik tidak tersampel berdasarkan informasi titik tersampel Z yang berada di sekitarnya dengan mempertimbangkan korelasi spasial yang ada dalam data tersebut. $Z(s_i)$

33

pembanding dari semivariogram eksperimental yang telah dihitung (Cressie, 1993):

1. Model Spherical

Semivariogram untuk Model *Spherical* dirumuskan sebagai::

$$\gamma(h) = \begin{cases} C \left[\left(\frac{3h}{2a} \right) - \left(\frac{h}{2a} \right)^3 \right] & \text{untuk } h \leq a \\ C & \text{untuk } h > a \end{cases} \quad (2.5.9)$$

dengan,
 h = jarak lokasi antar sampel,
 C = *sill*, yaitu nilai semivariogram untuk jarak pada saat besarnya konstan,
 a = *range*, yaitu jarak pada saat nilai semivariogram mencapai *sill*.

2. Model Eksponensial

Semivariogram model eksponensial dirumuskan sebagai berikut:

$$\gamma(h) = 1 - \exp\left(-\frac{3h}{a}\right) \quad (2.5.10)$$

3. Model Gaussian

Semivariogram model gaussian merupakan bentuk kuadrat dari eksponensial, dan dirumuskan sebagai berikut:

$$\gamma(h) = \left[1 - \exp\left(-\frac{3h}{a}\right)^2 \right] \quad (2.5.11)$$

Gambar 2.8 berikut merupakan ilustrasi dari ketiga model semivariogram teoritis:

32

didefinisikan sebagai variabel random pada titik $s_i, i = 1, 2, \dots, n$. Estimator *kriging* $\hat{Z}(s)$ dari $Z(s)$ dengan bobot λ_i adalah sebagai berikut (Bohling 2005):

$$\hat{Z}(s) - m(s) = \sum_{i=1}^n \lambda_i [Z(s_i) - m(s_i)] \quad (2.6.1)$$

dengan,
 s : lokasi untuk estimasi,
 s_i : salah satu lokasi data yang berdekatan,
 $m(s)$: nilai eksptasi dari $Z(s)$,
 $m(s_i)$: nilai eksptasi dari $Z(s_i)$,
 λ_i : Pembobot yang menentukan ukuran jarak antar titik,
 n : banyaknya data sampel yang digunakan untuk estimasi.

Apabila pada setiap lokasi $Z(s)$ terdapat estimator *error* $\tilde{e}(s)$, maka perbedaan antara nilai estimasi $\hat{Z}(s)$ dengan nilai $Z(s)$ didefinisikan sebagai berikut (Alfiana, 2010):

$$\tilde{e}(s) = \hat{Z}(s) - Z(s) \quad (2.6.2)$$

Metode Kriging terbagi menjadi tiga jenis kriging pokok yaitu *simple kriging*, *ordinary kriging* dan *universal kriging*.

1. Simple kriging

Simple kriging merupakan metode *kriging* dengan asumsi bahwa rata-rata dari populasi telah diketahui dan bernilai konstan. Pengembangan dari metode *simple kriging* adalah metode *sequential kriging* dimana pada metode ini data spasial yang akan di estimasi dipartisi menjadi beberapa bagian.

2. Ordinary kriging

Ordinary kriging merupakan metode kriging dengan asumsi bahwa rata-rata dari populasi tidak diketahui, dan pada data spasial tersebut tidak mengandung *trend*. Apabila pada data spasial yang akan di estimasi terdapat pencilan, maka

34

data yang tidak stasioner atau mengandung *trend*.

2.6.1. Ordinary Kriging

Ordinary kriging adalah salah satu metode geostatistika yang sederhana. Pada metode ini diasumsikan bahwa rata-rata (*mean*) tidak diketahui dan bernilai konstan. Pada metode *ordinary kriging* $m(s)$ diasumsikan sebagai rata-rata dari $Z(s)$, yaitu $m(s) = E(Z(s))$, dimana $E(Z(s)) = \mu$.

Ordinary kriging merupakan metode yang berhubungan dengan estimasi spasial dengan dua asumsi (Journel dan Huijberg, 1978), (Cressie, 1993), dan (Ghanim, 2010):

Asumsi Model

$$Z(s) = \mu + e(s); \quad se D, \mu \in \mathbb{R} \text{ dan } \mu \text{ tak diketahui} \quad (2.6.3)$$

Asumsi Estimasi

$$\hat{Z}(s) = \sum_{i=1}^n \lambda_i Z(s_i) \quad \text{dengan} \quad \sum_{i=1}^n \lambda_i = 1 \quad (2.6.4)$$

dengan,

$\hat{Z}(s)$: Nilai prediksi pada variabel s ,
 $e(s)$: nilai *error* pada $Z(s)$,
 λ_i : Pembobot yang menentukan ukuran jarak antar titik,
 i : 1, 2, ..., n, dimana n banyaknya data sampel yang digunakan untuk estimasi,
 $Z(s_i)$: Nilai *Actual* pada variabel s pada data ke- i .

Model yang mendasari *robust kriging* adalah (Cressie, 1993) :

Asumsi Model

$$Z(s) = \mu + e(s) + w(s); \quad se D, \mu \in \mathbb{R} \text{ dan } \mu \text{ tak diketahui} \quad (2.6.5)$$

Asumsi Estimasi

$$\hat{Z}(s) = \sum_{i=1}^n \lambda_i Z(s_i) w(Z(s_i)) \quad (2.6.7)$$

Dengan,

$\hat{Z}(s)$: Nilai prediksi pada variabel s ,
 $w(Z(s_i))$: transformasi dari bobot semivariogram yang berfungsi mengurangi nilai ekstrim,
 λ_i : Pembobot yang menentukan ukuran jarak antar titik,
 i : 1, 2, ..., n, dimana n banyaknya data sampel yang digunakan untuk estimasi,
 $Z(s_i)$: Nilai *Actual* pada variabel s pada data ke- i .

$w(\cdot)$ merupakan fungsi terboboti yang diasumsikan stasioner. Fungsi $w(\cdot)$ adalah untuk mengatasi adanya pencilan yang terdapat pada data spasial, sehingga asumsi estimasi pada persamaan (3.2.2) dapat dituliskan sebagai berikut (Cressie, 1993):

$$\hat{Z}(s) = \sum_{i=1}^n \lambda_i Z^{(o)}(s_i) \quad \text{dengan} \quad \sum_{i=1}^n \lambda_i = 1 \quad (2.6.8)$$

Dengan $Z^{(o)}(s_i)$ merupakan transformasi dari $Z(s_i)$ untuk nilai $Z(s_i)$ yang terpencil.

Nilai estimasi variabel terregional $\hat{Z}(s)$ pada titik yang tidak tersampel ditentukan berdasarkan nilai $Z(s_i)$ yang berada di sekitarnya dengan mempertimbangkan korelasi spasial pada data tersebut. Korelasi spasial pada data spasial ditentukan oleh nilai semivariogram teoritis. Semivariogram teoritis yang

Ketepatan prediksi *kriging* bergantung terhadap model semivariogram yang digunakan untuk menentukan bobot *kriging*.

2.6.2 Metode Robust Kriging

Kriging salah satu metode estimasi pada data spasial. Pada estimasi data spasial, variabel terregional dapat menghasilkan nilai estimasi yang kurang tepat jika pada data spasial yang diestimasi tersebut terdapat pencilan (*outlier*). Pencilan didefinisikan sebagai nilai yang ekstrim dari nilai pengamatan yang lainnya yang dapat disebabkan oleh kesalahan pencatatan, kalibrasi alat yang tidak tepat, atau kemungkinan yang lainnya.

Untuk mengestimasi data spasial yang mengandung pencilan maka digunakan metode *robust kriging* yang merupakan pengembangan dari metode *ordinary kriging*. Pada analisisnya, metode *robust kriging* mempertimbangkan adanya pencilan (*outlier*). Dalam estimasi data spasial dengan menggunakan metode *robust kriging* maka data spasial yang digunakan merupakan data spasial yang berpencilan. Dari data spasial berpencilan yang diperoleh dari lapangan dapat analisis semivariogram eksperimental. Pada analisis semivariogram eksperimental diperoleh nilai *siil* dan *range* yang selanjutnya digunakan untuk melakukan analisis semivariogram teoritis. Dari hasil analisis semivariogram eksperimen dan semivariogram teoritis maka selanjutnya adalah melakukan analisis struktural, yaitu membandingkan nilai semivariogram eksperimental dengan nilai semivariogram teoritis, kemudian dipilih model semivariogram teoritis dengan nilai MSE terkecil untuk menentukan bobot *kriging* yang selanjutnya digunakan untuk melakukan estimasi dengan metode *robust kriging*.

digunakan adalah semivariogram teoritis dengan nilai *MSE* terkecil pada analisis struktural. Analisis struktural dilakukan dengan membandingkan nilai *MSE* antara semivariogram teoritis dengan semivariogram eksperimental. Semivariogram eksperimental diperoleh dari analisis data spasial yang ada di lapangan.

Pada proses estimasi data spasial, semivariogram teoritis hasil analisis struktural digunakan untuk menentukan bobot dari masing-masing lokasi tersampel terhadap lokasi yang akan di estimasi. Besarnya nilai bobot pada masing-masing lokasi tersampel dapat ditentukan menggunakan operasi matriks sebagai berikut:

$$\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_i \\ \vdots \\ \lambda_n \\ p \end{bmatrix} = \begin{bmatrix} CZ_{11} & \cdots & CZ_{1n} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ CZ_{n1} & \cdots & CZ_{nn} & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} CZ_{10} \\ \vdots \\ CZ_{n0} \\ 1 \end{bmatrix}$$

Dengan $\sum_{i=1}^n \lambda_i = 1$.

Pada metode *robust kriging* untuk mengatasi adanya data yang mengandung pencilan maka dalam proses estimasi diperlukan transformasi dari bobot semivariogram $w(Z(s_i))$ yang berfungsi untuk mengurangi nilai ekstrim, sehingga dapat diperoleh hasil estimasi yang memiliki variansi *error* yang kecil dengan metode *robust kriging*.

2.7. Pencilan Spasial (Spatial Outlier)

Pencilan spasial (*Spatial Outlier*) didefinisikan sebagai nilai lokasi yang tidak konsisten atau sangat menyimpang terhadap nilai lokasi yang lainnya (Fridayani, 2012). Pencilan spasial dapat terjadi karena nilai observasi yang berbeda dengan yang lainnya yang disebabkan oleh kesalahan pada saat pengukuran atau



PDF

Complete

Your complimentary
use period has ended.
Thank you for using
PDF Complete.

Click Here to upgrade to
Unlimited Pages and Expanded Features

μ : nilai rata-rata dari $s(x)$,
 σ_s : standar deviasi dari $s(x)$,
 θ : nilai Z tabel untuk signifikansi tertentu.

$s(x)$ didefinisikan sebagai selisih antara nilai pengamatan dari lokasi x dengan rata-rata nilai pengamatan yang dekat dengan x , yaitu (Fridayani, 2012). Jika $Z_{s(x)} > \theta$, maka x dideteksi sebagai pencilan, untuk tingkat signifikansi 5% nilai $\theta \approx 3$ (Widhiarso, 2001).

2.8. Metode Lagrange

Metode *Lagrange* digunakan untuk menyelesaikan kasus optimasi (penentuan nilai ekstrim) dengan batasan-batasan(*Constraints*) tertentu (Alfiana, 2010). Jika akan ditentukan suatu nilai ekstrim dari fungsi $f(x,y)$ dengan batasan (*Constrain*) tertentu maka harus dipenuhi $g(x,y) = 0$ sehingga diperoleh fungsi *Lagrange* sebagai berikut:

$$F(x,y,\lambda) = f(x,y) + \lambda(g(x,y) - c) \tag{2.8 .1}$$

Dimana λ merupakan suatu pengali *Lagrange* dengan syarat

$$\frac{\partial F}{\partial x} = 0, \quad \frac{\partial F}{\partial y} = 0, \quad \frac{\partial F}{\partial \lambda} = 0$$

Pada analisis *kriging* metode lagrange digunakan untuk meminimalkan variansi *error*.

DAFTAR PUSTAKA

Alfiana, Anantia N. (2010). *Metode Ordinary Kriging pada Geostatistika*. (Skripsi sarjana pada FMIPA Universitas Negeri Yogyakarta). Yogyakarta: Tidak diterbitkan.

Anton, Howard. (1995). *Aljabar Linear Elementer (edisi kelima)*. (Terjemah oleh Pantur Silaban & I Nyoman Susila). Jakarta: Erlangga.

Bain & Engelhart. (1992). *Introduction to Probability and Mathematical Statistics 2nd Edition*. California: Duxbury Press.

Bohling, G. (2005). *Kriging*. [Online]. Tersedia: <http://people.ku.edu/~gbohling> diakses tanggal 24 Februari 2014.

Cressie, Noel A. (1993). *Statistics for Spatial Data*. New York: John Wiley & Sons, Inc.

Fridayani, Ni Made Suma, Kencana, Putu Eka Nila & Sukarsa, Komang Gede. (2012). Perbandingan Interpolasi Spasial Dengan Metode Ordinary dan Robust Kriging pada Data Spasial Berpencilan (Studi Kasus: Curah Hujan di Kabupaten Karangasem). *E-Jurnal Matematika*. 1(1): 68-74.

Ghanim Mhmood Dhaher, Muhammad Hisyam Lee. (2013). Robust Estimation for Two different Sets of Spatial Data with Application. *Australian Journal of Basic and Applied Sciences*, 7(10):562-569.

Journal, A.G., and C. Huijbregts. 1978. Mining Geostatistics, Academic Press, 600 pp.

Laksana, Endra L. (2010). *Analisis Data Geostatistik dengan Universal Kriging*. (Skripsi sarjana pada FMIPA Universitas Negeri Yogyakarta). Yogyakarta: Tidak diterbitkan.

Lasmiasih, Katarina. (2013). *Karakteristik Penduga Variogram untuk Data Nonstationer*. (Skripsi sarjana pada FMIPA Universitas Lampung)

Martin, Jay. D. (2010). Robust Kriging Models. 51st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference-
 18th.

Puspita, Wira dkk. (2012). *Analisis Data Geostatistik Menggunakan Metode Ordinary Kriging*. (Skripsi sarjana pada FMIPA Universitas Pendidikan Indonesia). Bandung: Tidak diterbitkan.

Suprajitno Munadi. (2005). *Pengantar geostatistik*. Jakarta: Universitas Indonesia.

Theodorick, dkk. (2013). *Metode Kriging untuk Prediksi Peak Ground Acceleration Berbasis Komputer*. (Thesis). Jakarta: Tidak diterbitkan.

Tirta, I Made. (2004). *Pengantar Statistika Matematika*. FMIPA Universitas Jember.

Umbara, Rian Febrian. (2007). *Model Semivariogram Kopula dan Reduksi Komputasi pada Algoritma Sequential Kriging*. (Thesis pada Institut Teknologi Bandung). Bandung: Tidak diterbitkan.