

INTERPOLASI ROBUST KRIGING PADA DATA CURAH HUJAN OUTLIER BULAN MARET 2015 DI KABUPATEN MALANG

Suci Astutik¹, Nur Aminah Kusuma Negara²

¹Email: suci_sp@ub.ac.id

Jurusan Matematika FMIPA Universitas Brawijaya

Abstract

Kriging method is one method of data interpolation involving location information (spatial data). The existence of outliers in the spatial data causes the kriging method to be incorrect. To overcome this is required kriging method that is able to handle outliers on spatial data called robust kriging. In general, the weight of the kriging method is determined by the semivariogram that measures spatial correlation. A semivariogram in robust kriging is called a robust semivariogram. There are several robust semivariogram, three of which are robust spherical, exponential, and gaussian semivariogram. The purpose of this research are to detect the location of rainfall precipitation stations including outliers, to determine the most appropriate robust semivariogram among the three semivariograms in robust kriging and to interpolate the rainfall data in March 2015 at Malang Regency. The data used in this research is spatial data of rainfall in March 2015 at 31 rain stations in Malang Regency containing outlier data. Result of research indicate that there are three location of rain station that detected outlier that is Ngantang rain station, Wagir and Bantur (Z value $> 1,96$). The most suitable robust semivariogram is the exponential robust semivariogram seen from the smallest RMSE value compared to the other robust semivariogram of 24.2752. The robust kriging interpolation result with the exponential semivariogram shows that the highest rainfall occurs in Ngantang and Kedungkandang rain stations.

Keywords: *robust semivariogram; rainfall data; kriging robust*

Abstrak

Metode Kriging merupakan salah satu metode interpolasi data yang melibatkan informasi lokasi (data spasial). Adanya pencilan pada data spasial menyebabkan metode kriging menjadi tidak tepat. Untuk mengatasi hal tersebut diperlukan metode kriging yang mampu menangani pencilan pada data spasial yang disebut sebagai robust kriging. Secara umum besarnya bobot pada metode kriging ditentukan oleh semivariogram yang mengukur korelasi spasial. Semivariogram pada pada robust kriging disebut sebagai semivariogram robust. Ada beberapa semivariogram robust, tiga diantaranya adalah semivariogram robust spherical, exponential, dan gaussian Tujuan penelitian ini adalah mendeteksi lokasi stasiun hujan curah hujan yang termasuk outlier, menentukan semivariogram robust yang paling sesuai diantara ketiga semivariogram tersebut pada robust kriging dan menginterpolasi data curah hujan bulan Maret 2015 di Kabupaten Malang. Data yang digunakan dalam penelitian ini adalah data spasial curah hujan bulan Maret 2015 di 31 stasiun hujan di Kabupaten Malang yang memuat data outlier. Hasil penelitian menunjukkan bahwa ada tiga lokasi stasiun hujan yang terdeteksi outlier yaitu stasiun hujan Ngantang, Wagir dan Bantur (Nilai $Z > 1,96$). Semivariogram robust yang paling sesuai adalah semivariogram robust exponential dilihat dari nilai RMSE yang paling kecil dibandingkan semivariogram robust yang lain yaitu sebesar 24,2752.

Hasil interpolasi robust kriging dengan semivariogram exponential menunjukkan bahwa curah hujan tertinggi terjadi di stasiun hujan Ngantang dan Kedungkandang.

Kata Kunci: semivariogram robust; curah hujan; robust kriging

PENDAHULUAN

Stasiun curah hujan dibangun pada beberapa titik pada suatu wilayah untuk mengetahui curah hujan pada daerah tersebut. Adanya keterbatasan baik biaya maupun tempat, maka tidak semua desa atau kecamatan memiliki stasiun curah hujan sendiri sehingga pos hujan dibangun pada daerah yang memiliki potensi dan dapat mewakili daerah disekitarnya. Hal ini menyebabkan curah hujan di daerah sekitar pos-pos hujan tidak bisa diketahui secara pasti. Oleh karena itu untuk menduga curah hujan di daerah sekitar pos-pos hujan yang belum diketahui dibutuhkan suatu metode yang sesuai.

Untuk menduga curah hujan di sekitar titik-titik yang diketahui diperlukan metode interpolasi data spasial. Menurut Anderson (2001) interpolasi spasial adalah suatu metode atau fungsi matematika yang menduga nilai atribut pada lokasi-lokasi yang datanya tidak diketahui. Interpolasi spasial mengasumsikan bahwa atribut data bersifat kontinu di dalam ruang (*space*) dan atribut ini saling berhubungan (*dependence*) secara spasial. Kedua asumsi tersebut yang mendasari bahwa pendugaan atribut pada suatu lokasi dapat dilakukan berdasarkan data dari lokasi-lokasi disekitarnya dan nilai pada titik-titik yang berdekatan akan lebih mirip dari pada titik-titik yang berjauhan (Prasasti, Wijayanto dan Christanto, 2005). Ada beberapa metode interpolasi spasial yang dikenal salah satunya yaitu metode *kriging*. Metode ini ditemukan oleh D.L. Krige untuk memperkirakan nilai dari bahan tambang. Kemudian metode ini dikembangkan oleh G.Matheron dalam bidang geostatistika. Menurut Oliver dan Webster (2015) *kriging* adalah teknik geostatistik yang bertujuan untuk menduga nilai pada lokasi yang tidak terukur dari nilai-nilai lokasi contoh yang terukur. *Kriging* bekerja berdasarkan pembobot yang tergantung pada *variogram* dan susunan dari lokasi contoh di lingkungan sekitar target. *Kriging* bersifat *best linear unbiased predictor*, *best* dalam meminimumkan ragam galat prediksi.

Pada beberapa penelitian dalam dunia geostatistika, para ahli telah banyak membuktikan bahwa metode *kriging* menghasilkan estimasi yang lebih baik dibandingkan dengan metode estimasi yang lainnya. Hal ini disebabkan karena dalam prosedur analisisnya, metode *kriging* bertujuan meminimalkan variansi dari galat. Terdapat berbagai macam model *kriging* diantaranya *ordinary kriging* dimana nilai tengah contoh belum diketahui. Metode *ordinary kriging* menghendaki adanya kestasioneran dalam rata-rata (μ) dan ragam (σ^2), sehingga jika kestasioneran tersebut dilanggar maka *ordinary Kriging* menghasilkan nilai prediksi yang kurang tepat. Seperti halnya pada semua metode analisis statistika, *ordinary kriging* juga akan menghasilkan nilai prediksi yang kurang tepat jika terdapat pencilan (*outlier*) pada data yang dianalisis. Menurut Lu, Chen dan Kou (2003) *Spatial outlier* didefinisikan sebagai objek spasial di mana nilai atribut non spasialnya berbeda secara signifikan dengan nilai-nilai dalam lingkungan tetangganya. Penyebab adanya *spatial outlier* antara lain adanya kesalahan alat pencatat, kesalahan dalam pencatatan atau memang benar suatu data yang teramati.

Untuk mengatasi adanya pencilan, digunakan metode *robust kriging*, metode yang mengakomodir adanya *spatial outlier*. Pada *robust kriging*, *semivariogram* eksperimental yang digunakan adalah *semivariogram robust*. Menurut Cressie (1993)

robust kriging lebih baik digunakan apabila data mengandung *spatial outlier*. Berdasarkan penelitian yang dilakukan oleh Darmanto dkk (2009) tentang *Robust Kriging* untuk Interpolasi Spasial pada Data Spasial Berpencilan (*Outlier*) telah membuktikan bahwa model *Robust Kriging* jauh lebih presisi dibandingkan dengan *ordinary kriging* dalam mengestimasi titik-titik spasial yang mengandung pencilan. Namun pada penelitian tersebut terdapat keterbatasan, yaitu belum bisa menginterpolasi titik yang belum tersampel. Sehingga peneliti ingin melakukan interpolasi *robust kriging* pada data spasial curah hujan bulanan Kabupaten Malang tahun 2015 yang mengandung *spatial outlier*.

METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah data sekunder tentang data Curah Hujan Bulanan (mm) di Kabupaten Malang pada Bulan Maret tahun 2015 yang terdiri dari 31 stasiun hujan. Beberapa tahapan yang dilakukan dalam penelitian ini adalah: (1) Identifikasi asumsi stasioneritas data spasial, (2) Pendeteksian *spatial outlier*, (3) Penghitungan Semivariogram eksperimental, (4) Fitting semivariogram teoritis dengan kriteria RMSE, (5) estimasi parameter Robust Kriging, (6) Pemetaan interpolasi *robust kriging*. Algoritma pendeteksian *spatial outlier* dengan metode *Spatial Z test* adalah sebagai berikut (Shekhar dkk, 2003):

1. Untuk masing-masing titik spasial s_i , hitung k tetangga terdekat $NN_k(s_i)$, fungsi rata-rata nilai atribut tetangga dengan persamaan (1) :

$$g(s_i) = \frac{1}{k} \sum_{s_j \in NN_k(s_i)} Z(s_j) \quad (1)$$

Fungsi pembanding memiliki persamaan :

$$h_i = h(s_i) = Z(s_i) - g(s_i) \quad (2)$$

2. μ_h dan σ_h berturut-turut merupakan rata-rata dan simpangan baku dari h_1, h_2, \dots, h_n , yang kemudian digunakan untuk menghitung statistik uji z di mana :

$$z_{h(s_i)} = \left| \frac{h(s_i) - \mu_h}{\sigma_h} \right| > \theta \quad (3)$$

Jika nilai $z_{h(s_i)} > \theta$, maka dideteksi sebagai pencilan (*outlier*), untuk tingkat signifikansi 5%, nilai θ mendekati 2.

Langkah-langkah pemodelan *robust kriging* adalah sebagai berikut (Paul dan Delfiner, 1999) :

1. menghitung *semivariogram robust*, dan mencocokkan model yang sesuai berdasarkan nilai RMSE yang terkecil.
2. menggunakan *semivariogram robust* untuk menghitung bobot *kriging* dan melakukan estimasi setiap lokasi.

$$\hat{Z}(s_i) = \sum_{\substack{j=1 \\ i \neq j}}^n \lambda_j Z(s_j) \quad (4)$$

dengan ragam dari *kriging* adalah σ_K^2 , dimana λ_i adalah bobot *kriging*.

3. Sebuah penduga *robust* dari $Z(s_i)$ memperhatikan *weighted median* pada persamaan (5) adalah :

$$\tilde{Z}(s_i) = \text{weighted median}(\{Z(s_j), i \neq j\}; \{\lambda_{ij}, i \neq j\}) \quad (5)$$

4. mengubah titik lokasi $Z(s_i)$ yang diidentifikasi sebagai *spatial outlier* dengan *winsorized version* :

$$Z^{(e)}(s_i) = \begin{cases} \tilde{Z}(s_i) + c\sigma_K & \text{jika } Z(s_i) - \tilde{Z}(s_i) > c\sigma_K \\ Z(s_i) & \text{jika } |Z(s_i) - \tilde{Z}(s_i)| \leq c\sigma_K \\ \tilde{Z}(s_i) - c\sigma_K & \text{jika } Z(s_i) - \tilde{Z}(s_i) < -c\sigma_K \end{cases} \quad (6)$$

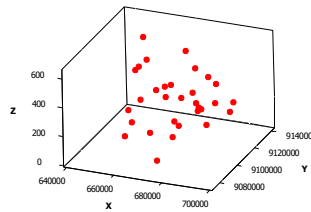
Dengan nilai $1,5 \leq c \leq 2,5$ akan mengontrol nilai σ_K tertinggi. Nilai c yang terpilih akan mengubah *outlier* yang tidak diinginkan tanpa mengubah rata-rata global lebih dari $\pm 5\%$. $Z^{(e)}(s_i)$ disebut *robust edited values*.

5. Kemudian melakukan interpolasi *ordinary kriging*, dengan menggunakan *semivariogram robust* pada lokasi s_j dengan menggunakan data yang telah diubah :

$$\hat{Z}^*(s_j) = \sum_{\substack{i=1 \\ i \neq j}}^n \lambda_{ij} Z^{(e)}(s_i) \quad (7)$$

HASIL DAN PEMBAHASAN

Hasil identifikasi stasioneritas dapat dilihat melalui plot antara koordinat lokasi (X,Y) dengan data curah hujan (Z) sebagaimana disajikan di Gambar 1. Berdasarkan Gambar 1 terlihat bahwa tidak terdapat trend antara lokasi dan data curah hujan bulan Maret tahun 2015 sehingga data memenuhi asumsi stasioneritas pada metode kriging.



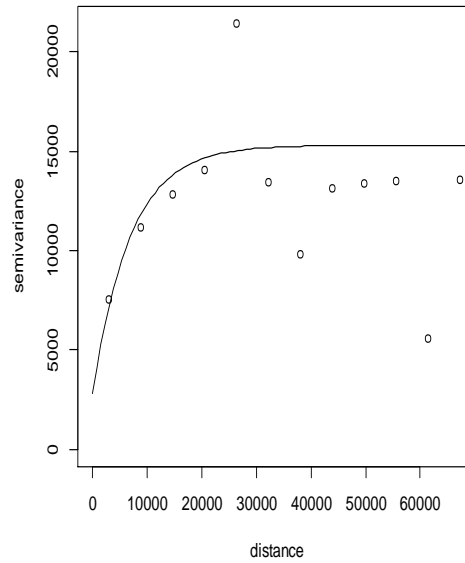
Gambar 1. Plot antara koordinat lokasi stasiun hujan dan data curah hujan bulan Maret 2015 di Kabupaten Malang

Pendeteksian spasial *outlier* dengan metode *Spatial Z test* disajikan pada Tabel 1. Berdasarkan Tabel 1 terlihat bahwa terdapat 3 dari 31 stasiun hujan di Kabupaten Malang yang terdeteksi *outlier* (Nilai $Z > Z^{0,025} = 1,96$). Oleh karena itu diperlukan metode *robust kriging* untuk menangani *outlier* dalam menginterpolasi data curah hujan.

Tabel 1. Stasiun Hujan yang Terdeteksi Outlier dengan Metode *Spatial Z Test*

Stasiun Hujan	Longitude (Bujur)	Latitude (Lintang)	Data	Nilai Z
Ngantang	651354,47	9131269,13	558	2,0073
Wagir	675503,09	9113778,39	631	2,6738
Bantur	674355,86	9081607,90	59	2,4221

Selanjutnya dalam pembentukan model kriging *robust*, diperlukan *semivariogram robust* yang mengukur korelasi antar lokasi pada data curah hujan yang memuat *outlier*. *Semivariogram robust* eksperimental data curah hujan bulan Maret 2015 di Kabupaten Malang sebagaimana disajikan di Gambar 2.



Gambar 2. Semivariogram Robust Data Curah Hujan Bulan Maret 2015 di Kabupaten Malang

Selanjutnya semivariogram robust eksperimental di Gambar 2 di *fitting* dengan semivariogram robust teori. Pada penelitian ini digunakan tiga semivariogram robust teori yaitu semivariogram robust *spherical*, *exponential* dan *Gaussian* dengan hasil sebagai berikut:

a. Model *spherical*

$$\gamma(\mathbf{h}) = \begin{cases} 6445,023 + 9030,89 \left[\frac{3\|\mathbf{h}\|}{2(26662,882)} - \frac{1}{2} \left(\frac{\|\mathbf{h}\|}{26662,882} \right)^2 \right], & 0 < \|\mathbf{h}\| < 26662,882 \\ 9030,89 & \text{otherwise} \end{cases}$$

b. Model *exponential*

$$\gamma(\mathbf{h}) = 2805,388 + 12487,310 \left(1 - \exp \left(-3\|\mathbf{h}\|/6890,007 \right) \right)$$

c. Model *gaussian*

$$\gamma(\mathbf{h}) = 7268,292 + 8025,475 \left[1 - \exp \left(-3 \frac{\|\mathbf{h}\|^2}{11567,611^2} \right) \right]$$

Untuk mengetahui semivariogram robust teori yang paling sesuai dengan semivariogram eksperimental, digunakan kriteria RMSE sebagaimana disajikan di Tabel 2.

Tabel 2. Nilai RMSE dari Tiga Semivariogram

Bulan	Model	RMSE
Maret	<i>Spherical</i>	56,4297
	<i>Exponential</i>	24,2752
	<i>Gaussian</i>	66,0070

Berdasarkan Tabel 2, menunjukkan bahwa nilai RMSE Semivariogram robust exponential paling kecil dibandingkan dengan dua semivariogram robust yang lain. Hal ini berarti semivariogram exponential yang paling sesuai dengan semivariogram eksperimental data curah hujan bulan Maret 2018. Selanjutnya semivariogram exponential yang telah diperoleh, digunakan untuk memperoleh pembobot robust kriging bagi tiga stasiun hujan yang terdeteksi outlier sebagaimana disajikan pada Tabel 3.

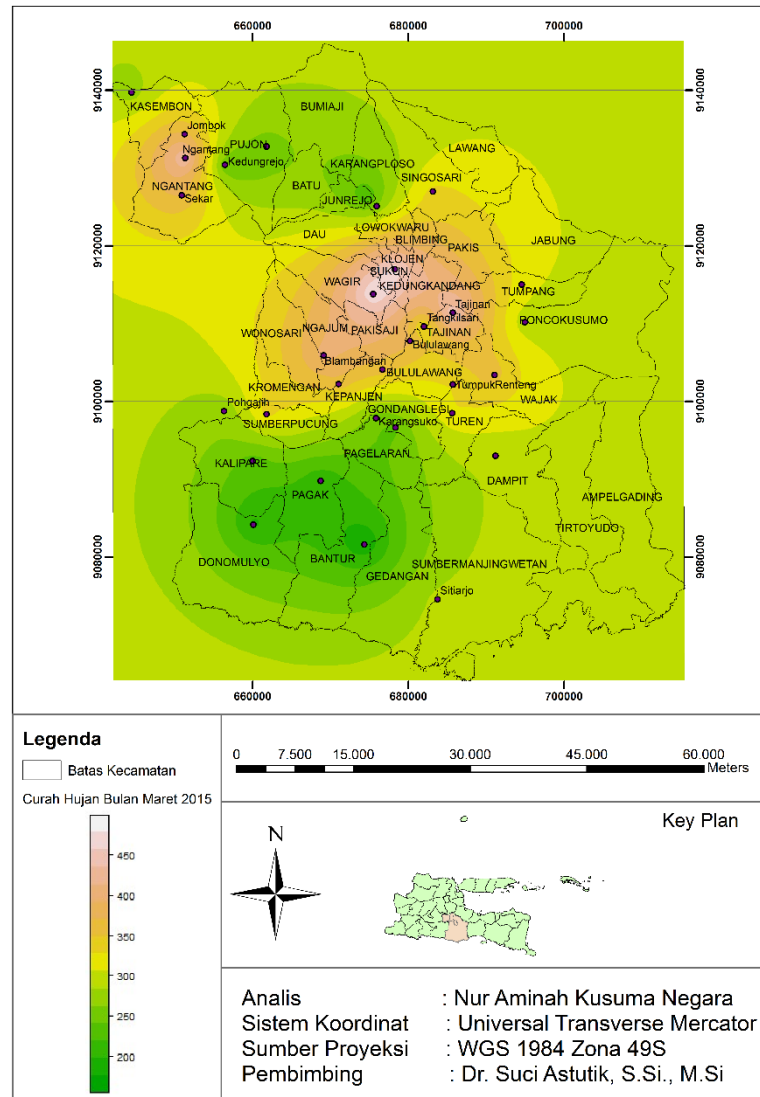
Tabel 3. Bobot Robust Kriging dari Tiga Stasiun yang Outlier

Lokasi 3: stasiun Ngantang					
No.	Lokasi	Bobot <i>kriging</i>	No.	Lokasi	Bobot <i>kriging</i>
1	Kasembon	0,0456	17	Karangsuko	0,0021
2	Jombok	0,3961	18	Blambangan	0,0017
4	Sekar	0,2523	19	Bantur	0,0067
5	Kedungrejo	0,1822	20	Turen	0,0027
6	Pujon	0,0180	21	Wajak	0,0038
7	Sukun	0,0033	22	Dampit	0,0071
8	Wagir	0,0034	23	Sitiarjo	0,0094
9	Tumpang	0,0061	24	TumpukRenteng	0,0017
10	Poncokusumo	0,0049	25	Kepanjen	0,0023
11	Singosari	0,0071	26	Sumberpucung	0,0035
12	Karangploso	0,0057	27	Pohgajih	0,0067
13	Bululawang	0,0014	28	Donomulyo	0,0073
14	Tajinan	0,0026	29	Kalipare	0,0040
15	Tangkilsari	0,0014	30	Pagak	0,0041
16	Gondanglegi	0,0028	31	Ngajum	0,0042

Lokasi 8: stasiun Wagir					
No.	Lokasi	Bobot <i>kriging</i>	No.	Lokasi	Bobot <i>kriging</i>
1	Kasembon	0,0139	17	Karangsuko	-0,0009
2	Jombok	0,0057	18	Blambangan	0,0627
3	Ngantang	0,0041	19	Bantur	0,0105
4	Sekar	0,0138	20	Turen	0,0015
5	Kedungrejo	0,0102	21	Wajak	0,0025
6	Pujon	0,0166	22	Dampit	0,0111
7	Sukun	0,3772	23	Sitiarjo	0,0150
9	Tumpang	0,0090	24	TumpukRenteng	-0,0056
10	Poncokusumo	0,0057	25	Kepanjen	0,0196
11	Singosari	0,0149	26	Sumberpucung	0,0061
12	Karangploso	0,0539	27	Pohgajih	0,0127
13	Bululawang	0,1036	28	Donomulyo	0,0116
14	Tajinan	0,0196	29	Kalipare	0,0054
15	Tangkilsari	0,0862	30	Pagak	0,0049
16	Gondanglegi	0,0006	31	Ngajum	0,1078

Lokasi 19: stasiun Bantur					
No.	Lokasi	Bobot <i>kriging</i>	No.	Lokasi	Bobot <i>kriging</i>
1	Kasembon	0,0359	16	Gondanglegi	0,0533
2	Jombok	0,0161	17	Karangsuko	0,0267
3	Ngantang	0,0104	18	Blambangan	0,0037
4	Sekar	0,0246	20	Turen	0,0228
5	Kedungrejo	0,0142	21	Wajak	0,0145
6	Pujon	0,0267	22	Dampit	0,0520
7	Sukun	0,0136	23	Sitiarjo	0,1803
8	Wagir	0,0135	24	TumpukRenteng	0,0054
9	Tumpang	0,0248	25	Kepanjen	0,0093
10	Poncokusumo	0,0198	26	Sumberpucung	0,0127
11	Singosari	0,0296	27	Pohgajih	0,0244
12	Karangploso	0,0256	28	Donomulyo	0,0954
13	Bululawang	0,0046	29	Kalipare	0,0208
14	Tajinan	0,0103	30	Pagak	0,1883
15	Tangkilsari	0,0053	31	Ngajum	0,0155

Kemudian bobot kriging yang diperoleh untuk menginterpolasi data curah hujan di lokasi lain sebagaimana disajikan pada Gambar 3. Berdasarkan Gambar 3 terlihat bahwa interpolasi data curah hujan tertinggi ada di stasiun hujan Ngantang dan Kedungkandang



Gambar 3. Peta Hasil Interpolasi Data Curah Hujan bulan Maret 2015

SIMPULAN DAN SARAN

Berdasarkan hasil dan pembahasan dapat diperoleh beberapa simpulan yaitu (1) terdapat tiga stasiun hujan yang terdeteksi sebagai outlier yaitu stasiun hujan Ngantang, Wagir dan Bantur, (2) Semivariogram robust exponential merupakan semivariogram yang paling sesuai dibandingkan dengan semivariogram spherical dan Gaussian dilihat dari nilai RMSE, dan (3) data curah hujan tertinggi dari hasil interpolasi robust kriging bulan Maret 2015 di Kabupaten Malang adalah stasiun hujan Ngantang dan Kedungkandang.

Pada penelitian selanjutnya perlu dikembangkan metode interpolasi curah hujan di lokasi lain pada bulan yang lain. Disamping itu juga perlu dikembangkan metode estimasi parameter robust kriging yang lain seperti Bayesian.

DAFTAR PUSTAKA

- Anderson, S. 2001. *An evaluation of Spatial interpolation methods on air temperature in Phoenix, AZ*. Department of Geography, Arizona State University Tempe, AZ 85287-0104
- Cressie, N. A. C. 1993. *Statistic for Spatial Data*. John Wiley and Sons, Inc. New York
- Lu, C.T., D. Chen, dan Y. Kou. 2003. Algorithms for spatial outlier detection. In *Proceedings of the 3rd IEEE International Conference on Data Mining (ICDM 2003)*. 19-22 December 2003. Melbourne, Florida, USA, IEEE Computer Society
- Oliver, M.A., R. Webster. 2015. *Basic Steps in Geostatistics: The Variogram and Kriging*. Springer. London
- Paul, J.C., dan P. Delfiner. 1999. *Geostatistics Modelling Spatial Uncertainty*. John Wiley & Sons, Inc. Newyork
- Prasasti, I. H. Wijayanto, dan M. Christanto. 2005. Analisis Penerapan Metode Kriging Dan Invers Distance Pada Interpolasi Data Dugaan Suhu, Air Mampu Curah (Amc) Dan Indeks Stabilitas Atmosfer (Isa) Dari Data Noaa-Tovs. *Pertemuan Ilmiah Tahunan MAPIN XIV*; Surabaya 14-15 September 2005. *Institut Teknologi Sepuluh November*
- Shekar, S., C.T Lu, dan P. Zhang . 2003. Unified Approach to Detecting *SpatialOutlier*. *Geoinformatics* 7(2): 141