

**PERBANDINGAN INTERPOLASI SPASIAL DENGAN
METODE *ORDINARY* DAN *ROBUST KRIGING* PADA DATA
SPASIAL BERPENCILAN
(Studi Kasus: Curah Hujan di Kabupaten Karangasem)**

**NI MADE SUMA FRIDAYANI¹, I PUTU EKA NILA KENCANA²,
KOMANG GDE SUKARSA³**

^{1,2,3}Jurusan Matematika, Fakultas MIPA, Universitas Udayana
e-mail: ¹ma_laft_pink@yahoo.com, ²i.putu.enk@gmail.com,
³sukarsakomang@yahoo.com

Abstract

Kriging as optimal spatial interpolation can produce less precise predictive value if there are outliers among the data. Outliers defined as extreme observation value of the other observation values that may be caused by faulty record keeping, improper calibration equipment or other possibilities. Development of Ordinary Kriging method is Robust Kriging which transforms weight of classic variogram thus become variogram that robust to outlier. The spatial data that used in this research is the spatial data that contains outliers and meet the assumptions of Ordinary Kriging. The analysis showed that the estimation value of Ordinary Kriging and Robust Kriging method is not much different in terms of Mean Absolute Deviation values which generated by both methods. An increase value of Mean Absolute Deviation on Robust Kriging estimation does not indicate that the Ordinary Kriging method is more precise than Robust Kriging method in the rainfall estimates of Amlapura control point remind that Robust Kriging does not eliminate the data of observation such as the Ordinary Kriging method. In general, Ordinary Kriging and Robust Kriging method can estimate the rainfall value of Amlapura control point quite well although it is not able to cover the changes in rainfall value that occurs due to the behavior geographic data.

Keywords: *Spatial interpolation, Kriging, Ordinary Kriging, Robust Kriging, outliers*

1. Pendahuluan

Hujan merupakan unsur iklim yang dapat diukur dengan suatu ukuran yaitu curah hujan. Curah hujan adalah ketinggian air hujan yang terkumpul pada tempat yang datar, tidak menguap, tidak meresap, dan tidak mengalir. Unsur hujan satu milimeter artinya dalam luasan satu meter persegi pada tempat yang datar tertampung air hujan setinggi satu milimeter atau tertampung air hujan sebanyak satu liter [1]. Pemerintah telah membangun pos-pos hujan untuk mengukur curah hujan di titik atau lokasi tertentu di

¹ Mahasiswa Jurusan Matematika FMIPA Universitas Udayana

^{2,3} Staf Pengajar Jurusan Matematika FMIPA Universitas Udayana

berbagai daerah di Indonesia yang dianggap memiliki potensi dan dapat mewakili daerah sekitarnya. Namun curah hujan di daerah sekitar pos-pos hujan tidak diketahui secara pasti. Diperlukan suatu metode interpolasi untuk memprediksi curah hujan di daerah yang tidak memiliki pos hujan.

Kriging adalah salah satu metode interpolasi spasial yang memanfaatkan nilai spasial pada lokasi tersampel untuk memprediksi nilai pada lokasi lain yang belum dan/atau tidak tersampel. *Kriging* dibawah asumsi kestasioneran sehingga jika asumsi kestasioneran tersebut dilanggar maka *Kriging* menghasilkan nilai prediksi yang kurang presisi. Selain itu, nilai prediksi kurang presisi juga dapat dihasilkan jika di antara data yang ada terdapat pencilan (*outlier*). *Kriging* yang umum digunakan diantaranya adalah *Ordinary Kriging* yang tidak mengakomodir adanya pencilan. Lebih lanjut *Ordinary Kriging* dikembangkan menjadi *Robust Kriging* yang mentransformasi bobot variogram sehingga menjadi variogram yang *robust* terhadap pencilan (Darmanto dan Soepraptini, [2]).

Pada penelitian ini, penulis tertarik untuk meneliti bagaimana pendugaan nilai curah hujan suatu titik pada data spasial yang mengandung pencilan menggunakan metode *Ordinary Kriging* dan *Robust Kriging*. Pengujian pencilan pada penelitian ini dilakukan dengan *spatial Z-test* yang mempertimbangkan jarak antar titik amatan. Titik yang diduga adalah pos hujan Amlapura yang terletak di Kabupaten Karangasem. Data set yang digunakan adalah data set yang mengandung pencilan yang diduga dengan metode *Robust Kriging*. Kemudian metode *Ordinary Kriging* digunakan untuk menduga data set yang sama dengan terlebih dahulu menghilangkan data pencilan. Dari pendugaan dengan kedua metode tersebut diperoleh suatu nilai galat (*error*) yang kemudian dibandingkan sehingga diketahui metode yang lebih presisi.

Variogram dan Semivariogram

Variogram adalah suatu fungsi acak intrinsik yang menunjukkan seberapa besar perubahan perbedaan sebanding dengan peningkatan jarak. Apabila suatu variabel pengukuran $z(s_1)$ dan $z(s_2)$ semakin jauh maka nilai $z(s_1) - z(s_2)$ akan semakin besar. Tujuan utama dari variogram adalah untuk memahami perbedaan kuadrat antar setiap titik:

$$(z(s_1) - z(s_2))^2$$

Jika varians dari perbedaan $[z(s_1) - z(s_2)]$ bergantung hanya pada perbedaan $(s_1 - s_2)$ maka:

$$var[z(s_1) - z(s_2)] = 2\gamma(s_1 - s_2) \quad (1)$$

Fungsi 2γ disebut variogram sedangkan fungsi γ disebut semivariogram.

Semivariogram Eksperimental Klasik

Semivariogram eksperimental adalah semovariogram yang diperoleh dan dihitung dari sampel. Semivariogram eksperimental ini sering juga disebut dengan semivariogram sampel (Oliver and Webster, [3]). Semivariogram ini dapat dihitung dengan menggunakan persamaan :

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

- Dengan s_i : lokasi titik sampel
 $z(s_i)$: nilai pengamatan pada lokasi
 h : jarak antara dua titik sampel
 $(s_i, s_i + h)$: pasangan titik sampel yang berjarak h
 $N(h)$: banyaknya pasangan berbeda yang memiliki jarak h

Spasial Outlier

Spasial Outlier (pencilan spasial) didefinisikan sebagai nilai lokasi observasi yang tidak konsisten atau sangat menyimpang (ekstrem) terhadap nilai lokasi observasi yang lainnya. Munculnya pencilan dapat disebabkan oleh mekanisme pengambilan nilai observasi yang berbeda dengan yang lainnya (Shekhar, S., dkk. [4]). Salah satu metode yang digunakan untuk mendeteksi adanya pencilan adalah dengan *spatial statistics Z test*[5]. *Spatial statistics Z test*, didefinisikan sebagai:

$$Z_{s(x)} = \left| \frac{s(x) - \mu_s}{\sigma_s} \right| > \theta \quad (3)$$

dengan

μ_s : nilai mean dari $s(x)$

σ_s : standar deviasi dari $s(x)$

θ : nilai Z tabel untuk tingkat signifikansi tertentu

$s(x)$ didefinisikan sebagai selisih antara nilai amatan dari lokasi x dengan rata-rata nilai amatan lokasi yang dekat dengan x :

$$s(x) = f(x) - E_{y \in N(x)} f(y) \quad (4)$$

Jika $Z_{s(x)} > \theta$, maka x dideteksi sebagai pencilan (*outlier*). Untuk tingkat signifikansi 5%, nilai $\theta \approx 2$.

Robust Kriging

Kriging sebagai interpolasi spasial yang optimum dapat menghasilkan nilai prediksi kurang presisi jika terdapat pencilan pada data. Menghapus pencilan ketika mengestimasi variogram mungkin masuk akal, tetapi ketika memprediksi suatu titik pengamatan, diperlukan sebuah cara alternatif untuk menangani pencilan. *Robust Kriging* merupakan pengembangan *Ordinary Kriging* yang mentransformasi bobot variogram menjadi variogram yang *robust* terhadap pencilan (Cressie, [6]).

Berbeda dengan semivariogram klasik, untuk mengakomodir adanya *outlier*, variogram untuk robust kriging dirumuskan sebagai

$$2\bar{\gamma}(h) \equiv \frac{\left\{ \frac{1}{N(h)} \sum_{n=1}^{N(h)} |z(s_i) - z(s_j)|^2 \right\}^4}{(0,457 + \frac{0,494}{N(h)})} \quad (5)$$

Dengan h : jarak antara dua titik sampel

$z(s_i) - z(s_j)$: selisih nilai pengamatan pada pasangan titik sampel yang berjarak h

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

2. Metode Penelitian

Data yang digunakan adalah data set curah hujan bulanan di Kabupaten Karangasem Tahun 2009 yang mengandung pencilan dan memenuhi asumsi *stationer intrinsik*. Tahapan penelitian diawali dengan memeriksa adanya pencilan pada seluruh data set tahun 2009 menggunakan statistik spasial *Z-test*. Kemudian dilanjutkan dengan memeriksa kestasioneran untuk data set yang mengandung pencilan dengan terlebih dahulu menghilangkan data pencilannya. Pada data yang memenuhi asumsi *stationer intrinsik* kemudian dilakukan Analisis dengan metode *Ordinary Kriging* yang tidak mengikutsertakan data pencilan kemudian metode *Robust Kriging* dilakukan dengan mengikutsertakan data pencilan.

Analisis pada metode *Ordinary Kriging* diawali dengan Membuat plot semivariogram eksperimental klasik untuk memperoleh model variogram yang sesuai dengan keadaan data menggunakan *software Matlab 7.0 package Easy Krig 3.0* dilanjutkan dengan melakukan uji validasi silang untuk menentukan model semivariogram terbaik. Dari model semivariogram terbaik yang diperoleh kemudian dihitung nilai bobot masing-masing titik amatan terhadap titik kontrol dilanjutkan dengan menghitung nilai dugaan pada titik control menggunakan bobot masing-masing titik amatan yang diperoleh. Sedangkan pada metode *Robust Kriging* plot semivariogram terboboti (*robust semivariogram*) diperoleh menggunakan fungsi “*variog()*”, uji validasi silang dilakukan menggunakan fungsi “*xvalid()*”, nilai bobot masing-masing titik amatan dan nilai dugaan pada titik kontrol diperoleh menggunakan fungsi “*krweights()*” dan fungsi “*krige.conv()*” yang terdapat pada *software R 2.14.0package geoR 1.7-2*. Langkah-langkah tersebut dilakukan pada semua data set sampel. Setelah diperoleh nilai dugaan titik kontrol pada semua data set sampel dilanjutkan dengan menghitung nilai *Mean Absolute Deviation (MAD)* dari metode *Ordinary Kriging* dan *Robust Kriging* yang kemudian dibandingkan sehingga diketahui metode yang lebih presisif.

3. Hasil dan Pembahasan

Pengujian Data Pencilan

Untuk melakukan pengujian pencilan terlebih dahulu ditentukan pos-pos hujan yang saling bertetangga. Pada penelitian ini pos-pos hujan diasumsikan bertetangga apabila jarak antar pos hujan kurang dari atau sama dengan $|\mu - \sigma|$. Dari jarak masing-masing pos hujan diperoleh rata-rata (μ) sebesar 16,62 dan simpangan baku (σ) sebesar 7,52. Pos hujan dikatakan bertetangga apabila jarak antar pos hujan tersebut kurang dari atau sama dengan nilai $|\mu - \sigma| = 16,62 - 7,52 = 9,10$. Nilai Z untuk data yang merupakan pencilan pada masing-masing data set dirangkum dalam tabel 1.

Tabel 1 Titik Amatan yang Merupakan Pencilan pada Masing-masing Data Set

Data Set	Titik Amatan	Nilai Z	Data Set	Titik Amatan	Nilai Z
Januari	Besakih	2,7142	September	Ulakan	-2,0707
Februari	Besakih	2,3031		Besakih	2,5549
Maret	Besakih	2,7059	Oktober	Singarata	2,2468
April	Besakih	2,4849		Besakih	2,0185
Mei	Besakih	2,3805	November	Besakih	3,1359
Juni	Besakih	2,4288	Desember	Besakih	2,4308
Agustus	Ulakan	-2,3958			

Tabel 1 menunjukkan tidak terdapatnya pencilan pada data set Juli kemungkinan dikarenakan pada bulan Juli merupakan bulan kering dimana nilai curah hujan relatif kecil sehingga perbedaan nilai curah hujan antar daerah tidak terlalu signifikan. Nilai curah hujan pos hujan Besakih merupakan pencilan hampir pada semua data set. Ditinjau dari jarak antar titik amatan hal ini kemungkinan disebabkan oleh letak pos hujan Besakih yang jauh dari pos hujan lainnya mengingat pengujian pencilan dengan statistik uji *Z-test* mempertimbangkan jarak antar titik amatan.

Pengujian Asumsi Stationer Intrinsik

Pengujian asumsi kestasioneran dilakukan menggunakan *scatterplot* data curah hujan yang terstandarisasi seluruh titik amatan selain titik amatan yang merupakan pencilan pada masing-masing data set. Asumsi tersebut terpenuhi apabila data menyebar secara acak dan tidak membentuk pola tertentu. Untuk lebih menguatkan analisa, dilakukan pengujian kenormalan data terhadap masing-masing data set. Hasil uji kenormalan menggunakan uji *Anderson Darling* dengan $\alpha = 0,05$ menunjukkan data set Agustus dan Oktober tidak menyebar normal sehingga dapat dikatakan bahwa keacakan tidak signifikan. Oleh karena itu asumsi stasioneritas tidak terpenuhi pada data set tersebut sehingga pada analisis selanjutnya data set yang digunakan adalah data set Januari, Februari, Maret, April, Mei, Juni, September, November, dan Desember.

Metode Ordinary Kriging

Pada pendugaan nilai curah hujan dengan metode ini, data yang merupakan pencilan pada masing-masing data set tidak diikutsertakan dalam analisis. Analisis dengan metode *Ordinary Kriging* diawali dengan uji validitas model semivariogram klasik. Model semivariogram yang akan diujikan adalah model *Spherical*, *Eksponential*, *Gaussian* dan *Linear*. Statistik uji yang digunakan adalah statistik uji Q_1 . Model terbaik dipilih berdasarkan nilai $|Q_1|$ terkecil. Nilai $|Q_1|$ dari masing-masing model semivariogram untuk masing-masing data set dirangkum pada tabel 2.

Tabel 2. Nilai $|Q_1|$ Masing-Masing Model Semivariogram Klasik untuk Masing-masing Data Set

Model Semi-variogram	Data Set								
	Januari	Februari	Maret	April	Mei	Juni	September	November	Desember
<i>Spherical</i>	-0,5780	0,1160	-0,5623	-0,3087	-0,1908	0,1047	-0,2739	-0,5421	0,2489
<i>Exponential</i>	-0,5631	0,0977	-0,6968	-0,2210	-0,0826	-0,0543	-0,1754	-0,5413	-0,0244
<i>Gaussian</i>	-0,5976	0,0540	-1,5288	-0,1391	-0,1126	0,1253	-0,0118	-0,1926	-0,0441
<i>Linear</i>	-0,6460	0,0946	-0,8768	-0,1537	0,0721	-0,1092	-0,2333	-0,6585	-0,1353

Berdasarkan tabel 2 dapat dilihat bahwa semivariogram klasik model *spherical* digunakan pada data set Maret. Model *exponential* digunakan pada data set Januari, Juni, dan Desember. Model *gaussian* digunakan pada dataset Februari, April, September, dan November sedangkan Model *linear* digunakan pada data set Mei. Dari fungsi semivariogram terbaik diperoleh bobot dan nilai curah hujan dugaan titik kontrol Amlapura untuk masing-masing data set.

Metode Robust Kriging

Berdasarkan hasil dari validasi silang dengan metode *Robust Kriging*, model semivariogram terbaik untuk masing-masing data set dipilih berdasarkan nilai rataan residual standar terkecil. Nilai rataan residual standar dari masing-masing model semivariogram untuk masing-masing data set dirangkum pada tabel 3.

Tabel 3 Nilai Rataan Residual Standar Masing-masing Model *Robust* Semivariogram untuk Masing-masing Data Set

Model Semivariogram	Data Set								
	Januari	Februari	Maret	April	Mei	Juni	September	November	Desember
<i>Spherical</i>	-0,5780	0,1160	-0,5623	0,0111	-0,0102	0,0099	-0,0078	0,0215	-0,0145
<i>Exponential</i>	$0,5 \times 10^{-4}$	-0,0185	0,0003	-0,0186	-0,0091	$2,1 \times 10^{-17}$	-0,0077	0,0144	$-8,5 \times 10^{-18}$
<i>Gaussian</i>	0,0006	-0,0145	-0,0015	-0,0280	-0,0123	$2,1 \times 10^{-17}$	-0,0098	-0,0082	$-7,7 \times 10^{-9}$

Berdasarkan tabel 3 dapat dilihat bahwa *robust* semivariogram model *spherical* digunakan pada data set April. Model *exponential* digunakan pada data set Januari, Maret, Mei, Juni, September dan, Desember, sedangkan model *gaussian* digunakan pada dataset Februari, Juni dan November. Sementara model *linear* tidak valid untuk semua data set. Dari model semivariogram terbaik diperoleh bobot dan nilai curah hujan dugaan titik kontrol Amlapura untuk masing-masing data set.

Perbandingan Metode *Ordinary Kriging* dan *Robust Kriging*

Tabel 4. Perbandingan Metode *Robust Kriging* dan *Ordinary Kriging*

Data Set	Nilai Amatan	Metode <i>Ordinary Kriging</i>			Metode <i>Robust Kriging</i>		
		Nilai Dugaan	Galat	Galat Mutlak	Nilai Dugaan	Galat	Galat Mutlak
Januari	391	401,6580	-10,6580	10,6580	391,6260	-0,6260	0,6260
Februari	542	420,5193	121,4807	121,4807	370,5237	171,4763	171,4763
Maret	63	80,0427	-17,0427	17,0427	80,4117	-17,4117	17,4117
April	111	175,9016	-64,9016	64,9016	161,5594	-50,5594	50,5594
Mei	161	102,3733	58,6267	58,6267	106,4269	54,5731	54,5731
Juni	14	17,4859	-3,4859	3,4859	13,6923	0,3077	0,3077
September	58	86,3704	-28,3704	28,3704	73,4677	-15,4677	15,4677
November	15	-13,7290	28,7290	28,7290	42,0010	-27,0010	27,0010
Desember	90	247,8648	-157,8648	157,8648	259,2154	-169,2150	169,2150

Dari tabel 4 terlihat bahwa terdapat perbedaan yang cukup besar antara nilai amatan dengan nilai dugaan pada beberapa data set seperti data set Februari, April, Mei, dan Desember. Perbedaan yang tidak begitu besar terjadi pada data set Maret, September, dan November. Sedangkan perbedaan yang sangat kecil terjadi pada data set Januari dan Juni. Nilai *MAD* yang dihasilkan dari metode *Ordinary Kriging* adalah sebesar **54,5733** sedangkan nilai *MAD* dari metode *Robust Kriging* adalah sebesar **56,2931**.

Metode *Ordinary Kriging* dan *Robust Kriging* memberikan nilai pendugaan yang tidak jauh berbeda ditinjau dari nilai *MAD* yang dihasilkan oleh kedua metode. Dilihat dari nilai mutlak galat yang dihasilkan oleh kedua metode, metode *Robust Kriging* lebih unggul dari *Ordinary Kriging* pada hampir semua data set kecuali data set Februari, Maret, dan Desember. Pada data set Februari, metode *Ordinary Kriging* memiliki nilai galat mutlak yang jauh lebih kecil dari nilai galat mutlak pada metode *Robust Kriging* menunjukkan bahwa metode *Ordinary Kriging* dapat menduga nilai curah hujan jauh lebih baik untuk data set Februari. Hal ini yang menyebabkan nilai *MAD* *Robust Kriging* menjadi lebih besar dari nilai *MAD* *Ordinary Kriging*. Kemudian lebih lanjut pada data

set Februari dengan metode *Robust Kriging*, validasi silang menghasilkan rata-rata residual standar yang tidak jauh berbeda antara ketiga model *robust* semivariogram yang valid. Sementara jika dibandingkan dengan rata-rata residual standar pada data set lain, rata-rata residual standar pada data set Februari bernilai cukup besar, hal ini mengindikasikan bahwa ketiga model *robust* semivariogram tidak dapat mewakili keadaan spasial data dengan cukup baik untuk data set Februari. Secara umum, metode *Ordinary Kriging* maupun *Robust Kriging* dapat melakukan pendugaan curah hujan titik kontrol Amlapura dengan cukup baik walaupun belum dapat meng-cover perubahan-perubahan nilai curah hujan yang terjadi akibat perilaku geografis data.

4. Kesimpulan

Pendugaan curah hujan terhadap titik kontrol Amlapura dengan metode *Robust Kriging* menghasilkan nilai MAD yang sedikit lebih besar dari nilai MAD yang dihasilkan pada pendugaan dengan metode *Ordinary Kriging*. Menurut penulis, terjadinya peningkatan nilai MAD pada pendugaan dengan metode *Robust Kriging* tidak mengindikasikan bahwa metode *Ordinary Kriging* lebih akurat daripada metode *Robust Kriging* dalam pendugaan curah hujan terhadap titik kontrol Amlapura mengingat metode *Robust Kriging* tidak meniadakan data amatan seperti pada metode *Ordinary Kriging*.

Metode *Kriging* yang digunakan dalam penelitian ini adalah metode *Kriging* dengan semivariogram isotropik sehingga untuk penelitian lebih lanjut dapat dikembangkan dengan menggunakan semivariogram anisotropik yang memperhatikan arah dan sudut. Selain itu dapat digunakan jenis *Kriging* lain seperti *Co-Kriging* yang merupakan suatu teknik khusus dalam interpolasi dengan memakai dua variabel yang berbeda akan tetapi secara spasial saling berhubungan maupun *Universal Kriging* yang merupakan *Kriging* untuk data yang mempunyai kecenderungan *trend* tertentu.

Daftar Pustaka

- [1] Balai Besar Meteorologi Klimatologi dan Geofisika Wilayah III Denpasar. 2010. "Informasi Cuaca, Iklim, dan Gempabumi Provinsi Bali". *Buletin Meteorologi Klimatologi dan Geofisika Tahun IV No. 04*.
- [2] Darmanto dan Soepraptini. 2009. "*Robust Kriging* untuk Interpolasi Spasial pada Data Spasial Berpencilan (*Outlier*)". *Laporan Penelitian*. Departemen Matematika FMIPA Universitas Brawijaya. Tidak diterbitkan.
- [3] Oliver, M. A and R. Webster 2007. *Geostatistic for Environmental Statistic*, Second Edition. United Kingdom: Willey.
- [4] Shekhar, S., dkk. 2007. "Unified Approach to Detecting Spatial Outlier". *Research Report*. Pekka Maksimainen, University of Helsinki. Tidak diterbitkan.
- [5] Oh, Jisu and Shan Huang. 2005. "Spatial Outlier Detection and Implementation in WEKA". *Research Report*. University of Minnesota. USA. Tidak diterbitkan.
- [6] Cressie, N. 1993. *Statistic for Spatial Data*, Revised Edition. New York: Willey.