

**PERBANDINGAN METODE *WEIGHTED Z ALGORITHM*
DAN *ITERATIVE R ALGORITHM* DALAM MENDETEKSI
OUTLIER PADA DATA SPASIAL**

SKRIPSI

Sebagai salah satu syarat untuk memperoleh gelar
Sarjana Sains dalam bidang Statistika

oleh:
ZUMROTUS SA'ADAH
0610950062-95



**PROGRAM STUDI STATISTIKA
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS BRAWIJAYA
MALANG
2013**

LEMBAR PENGESAHAN SKRIPSI

**PERBANDINGAN METODE *WEIGHTED Z ALGORITHM*
DAN *ITERATIVE R ALGORITHM* DALAM MENDETEKSI
OUTLIER PADA DATA SPASIAL**

oleh:

ZUMROTUS SA'ADAH
0610950062-95

Setelah dipertahankan di depan Majelis Penguji
pada tanggal 21 Agustus 2013
dan dinyatakan memenuhi syarat untuk memperoleh gelar
Sarjana Sains dalam bidang Statistika

Pembimbing I

Pembimbing II

Prof. Dr. Ir. Henny Pramodyo, MS.
NIP. 195707051981031009

Prof. Dr. Ir. Loekito Adi S, M.Agr
NIP. 194703271974121001

Mengetahui,
Ketua Jurusan Matematika
Fakultas MIPA Universitas Brawijaya

Dr. Abdul Rouf Alghofari, M.Sc.
NIP. 196709071992031001

LEMBAR PERNYATAAN

Saya yang bertanda tangan di bawah ini :

Nama : ZUMROTUS SA'ADAH
NIM : 0610950062-95
Program Studi : STATISTIKA
Penulis Skripsi Berjudul :

**PERBANDINGAN METODE *WEIGHTED Z ALGORITHM*
DAN *ITERATIVE R ALGORITHM* DALAM MENDETEKSI
OUTLIER PADA DATA SPASIAL**

Dengan ini menyatakan bahwa:

1. Isi dari Skripsi yang saya buat adalah benar-benar karya sendiri dan tidak menjiplak karya orang lain, selain nama-nama yang termaktub di isi dan tertulis di daftar pustaka dalam Skripsi ini.
2. Apabila di kemudian hari ternyata Skripsi yang saya tulis terbukti hasil jiplakan maka saya akan bersedia menanggung segala resiko yang akan saya terima.

Demikian pernyataan ini dibuat dengan segala kesadaran.

Malang, 21 Agustus 2013
Yang menyatakan,

ZUMROTUS SA'ADAH
NIM. 0610950062-95

PERBANDINGAN METODE WEIGHTED Z ALGORITHM DAN ITERATIVE R ALGORITHM DALAM MENDETEKSI OUTLIER PADA DATA SPASIAL

ABSTRAK

Spasial outlier adalah suatu objek yang tidak konsisten dengan tetangga spasialnya sekalipun nilai-nilai nonspasialnya adalah normal untuk sisa dari objek kelas yang sama. Pendeteksian spasial outlier sangat bermanfaat untuk berbagai bidang di sistem informasi geografis (SIG) seperti ekologi, transportasi, kesehatan masyarakat, klimatologi, pelayanan umum dan lain-lain (Lu, dkk., 2003 : 1). Pada penelitian ini akan dibandingkan metode weighted z algorithm dan metode iterative r algorithm untuk mengetahui mana dari kedua metode tersebut yang lebih sensitif dalam mendeteksi spasial outlier dan dapat mengatasi masking effect dan swamping effect. Data yang digunakan pada penelitian ini data sekunder kadar C-organik tanah di Kabupaten Sampang yang diperoleh dari hasil penelitian tanaman tembakau tahun 2008 kerjasama Universitas Brawijaya dengan Balai Penelitian Tanaman Tembakau dan Serat (BALITTAS) yang terdiri atas 56 objek pengamatan. Dari hasil penelitian ini didapatkan bahwa pendeteksian outlier secara tradisional mampu mendeteksi 9 outlier, metode weighted z algorithm mampu mendeteksi 5 spasial outlier lebih sedikit dari yang terdeteksi oleh tradisional outlier. Pada metode iterative r algorithm mampu mendeteksi 10 spasial outlier pada data, lebih banyak dari yang terdeteksi tradisional outlier. Dalam penentuan metode yang lebih baik dan lebih sensitif dapat diketahui dari cara suatu metode dalam mengatasi masking effect dan swamping effect. Dari hasil penelitian didapatkan bahwa pendeteksian outlier secara tradisional dan metode weighted z algorithm tidak cukup mampu dalam mengatasi masking effect sedangkan metode iterative r algorithm mampu mengatasi adanya masking effect dan swamping effect.

Kata kunci : spasial, outlier, bobot, rasio.

COMPARISON METHOD OF WEIGHTED Z ALGORITHM AND ITERATIVE R ALGORITHM IN DETECTING OUTLIER ON SPATIAL DATA

ABSTRACT

Spatial outlier is an object that is inconsistent with its neighborhood though non-spatial values are normal for the other objects in the same class. Detecting spatial outliers is useful in many applications of geographic information systems (GIS) including ecology, transportation, public health, climatology, public safety and location based service (Lu, dkk., 2003 : 1). In this study will be compared with the method of Weighted z Algorithm and Iterative r Algorithm to determine which method is more sensitive and can overcome the masking effect and swamping effect. The data used in this study secondary data C-organic content of soil in Sampang Regency that obtained from the research of tobacco plants in cooperation with University of Brawijaya in 2008 which consist of 56 observations. From the result that traditional outlier detecting 9 outlier, Weighted z Algorithm can detecting 5 spastial outlier less than tradisional outlier detected. Iterative r Algorithm can detected 10 spatial outlier more than tradisional outlier detected. The determination methods better and more sensitive can be known from the method in overcoming masking effect and swamping effect. From the result found that Iterative r Algorithm better able to overcome masking effect and swamping effect than tradisional outlier detecting and Weighted z Algorithm.

Key words : spatial, outlier, weight, ratio.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT yang senantiasa memberikan rahmat, hidayah dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “Perbandingan Metode *Weighted Z Algorithm* dan *Iterative R Algorithm* Dalam Mendeteksi *Outlier* Pada Spasial” sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains dalam bidang Statistika. Pada kesempatan ini penulis mengucapkan terima kasih kepada :

1. Bapak Prof. Dr. Ir. Henny Pramodyo, M. S dan Bapak Prof. Dr. Ir. Loekito Adi S, M. Agr, selaku Dosen Pembimbing atas bimbingan dan masukan yang telah diberikan.
2. Ibu Dr. Rahma Fitriani, M. Sc, selaku Dosen Penguji atas saran dan nasihat yang telah diberikan.
3. Bapak Dr. Abdul Rouf Alghofari, M.Sc. selaku Ketua Jurusan Matematika FMIPA Universitas Brawijaya.
4. Kedua orang tua yang telah memberikan kasih sayang, doa, dan dukungan yang tulus.
5. Kakak-kakak dan adik-adikku tercinta atas dukungan untuk selalu bersemangat dan tidak putus asa dalam menyelesaikan skripsi ini.
6. Teman-temanku Reni, Nining, Dayu, Icha, Weni, Nunun, Sea, Linda, Uyun, Windi, Gilang, seluruh eks 259B dan seluruh Dinoyo 169C atas semangat, dukungan dan kebersamaannya selama ini.
7. Seluruh teman-teman Statistika 2006, 2008 dan 2009 atas semangat, dukungan dan kebersamaannya selama ini.
8. Semua pihak yang tidak dapat penulis sebutkan satu persatu yang telah membantu selama penulisan skripsi.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun demi perbaikan selanjutnya. Semoga skripsi ini dapat bermanfaat bagi pembaca.

Malang, 21 Agustus 2013

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL	x
DAFTAR LAMPIRAN	xi
 BAB I PENDAHULUAN	 1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	2
1.5 Manfaat Penelitian	2
 BAB II TINJAUAN PUSTAKA	 3
2.1 Data Spasial	3
2.1.1 Format Data Spasial	3
2.1.2 Sumber Data Spasial	6
2.2 Sistem Informasi Geografis	7
2.3 <i>Outlier</i>	7
2.4 Spasial <i>Outlier</i>	9
2.5 Metode Pendeteksian Spasial <i>Outlier</i>	10
2.5.1 Formulasi Algoritma	10
2.5.2 k Tetangga Terdekat	11
2.5.3 Ukuran Keruangan (Ukuran Spasial)	11
2.5.4 <i>z Algorithm</i>	13
2.5.5 <i>Weighted z Algorithm</i>	14
2.5.6 <i>Iterative r Algorithm</i>	17
2.6 Pemilihan Metode Terbaik	18
2.7 C-organik	19

BAB III METODE PENELITIAN	22
3.1 Sumber Data	22
3.2 Metode Penelitian	22
3.2.1 Pendeteksian Tradisional <i>Outlier</i>	22
3.2.2 <i>Weighted z Algorithm</i>	23
3.2.3 <i>Iterative r Algorithm</i>	23
3.2.4 Pendeteksian <i>Masking Effect</i> dan <i>Swamping Effect</i>	23
3.2.5 Penentuan Metode Terbaik	24
BAB IV HASIL DAN PEMBAHASAN	26
4.1 Pendeteksian Spasial <i>Outlier</i> Berdasarkan Tradisional <i>Outlier</i>	26
4.2 Pendeteksian Spasial <i>Outlier</i> Berdasarkan <i>Weighted</i> <i>z Algorithm</i>	27
4.3 Pendeteksian Spasial <i>Outlier</i> Berdasarkan <i>Iterative r</i> <i>Algorithm</i>	28
4.4 Perbandingan Pendeteksian Tradisional <i>Outlier</i> , <i>Weighted z Algorithm</i> , dan <i>Iterative r Algorithm</i>	29
4.5 Pendeteksian <i>Masking Effect</i> dan <i>Swamping</i> <i>Effect</i>	30
4.6 Penentuan Metode Terbaik	32
BAB V KESIMPULAN DAN SARAN	34
5.1 Kesimpulan	34
5.2 Saran	34
DAFTAR PUSTAKA	35
LAMPIRAN	39

DAFTAR GAMBAR

Halaman

Gambar 2.1	Contoh data vektor	4
Gambar 2.2	Contoh data raster	5
Gambar 3.1	Diagram alir perbandingan metode Weighted z Algorithm dan Iterative r Algorithm	25



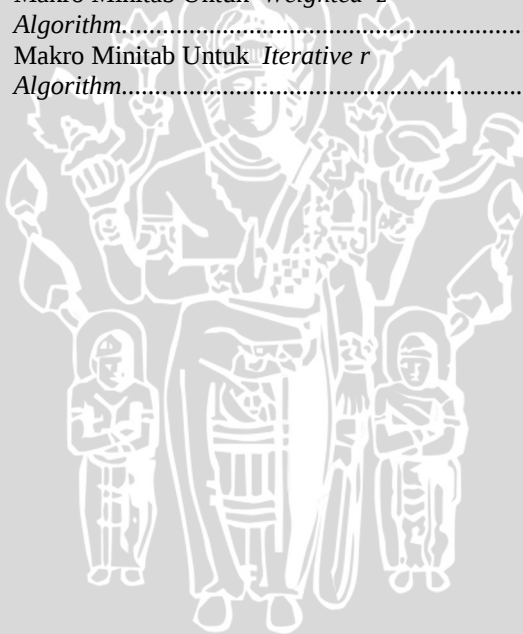
DAFTAR TABEL

		Halaman
Tabel 4.1	Tabel Titik yang Merupakan Spasial Outlier Berdasarkan Pendeteksian Tradisional Outlier.....	26
Tabel 4.2	Tabel Titik yang Merupakan Spasial Outlier Berdasarkan Pendeteksian Weighted z Algorithm.....	27
Tabel 4.3	Tabel Titik yang Merupakan Spasial Outlier Berdasarkan Pendeteksian Iterative r Algorithm.....	28
Tabel 4.4	Perbandingan Pendeteksian Spasial Outlier Berdasarkan Tradisional Outlier, Weighted z Algorithm, dan Iterative r Algorithm.....	29
Tabel 4.5	Pendeteksian Masking Effect dan Swamping Effect Berdasarkan Tradisional outlier	30
Tabel 4.6	Pendeteksian Masking Effect dan Swamping Effect Berdasarkan Metode Weighted z Algorithm	31
Tabel 4.7	Pendeteksian Masking Effect dan Swamping Effect Berdasarkan Metode Iterative r Algorithm.....	31
Tabel 4.8	Kriteria Kebaikan Metode Pendeteksi Spasial Outlier.....	32

DAFTAR LAMPIRAN

Halaman

Lampiran 1	Data Kandungan C-organik Tanah di Kabupaten Sampang.....	39
Lampiran 2	Pendeteksian Spasial <i>Outlier</i> Berdasarkan Tradisional <i>outlier</i>	41
Lampiran 3	Pendeteksian Spasial <i>Outlier</i> Berdasarkan Metode <i>Weighted z Algorithm</i>	45
Lampiran 4	Pendeteksian Spasial <i>Outlier</i> Berdasarkan Metode <i>Iterative r Algorithm</i>	49
Lampiran 5	Peta Kabupaten Sampang Madura.....	53
Lampiran 6	Makro Minitab Untuk <i>Weighted z Algorithm</i>	54
Lampiran 7	Makro Minitab Untuk <i>Iterative r Algorithm</i>	56



BAB I

PENDAHULUAN

1.1. Latar Belakang

Pada suatu data, termasuk data spasial memiliki kondisi yang tidak wajar seperti adanya outlier. Outlier adalah pengamatan yang tidak mengikuti sebagian besar pola dan terletak jauh dari pusat data (Barnett, 1981: 1). Spasial outlier adalah suatu objek yang tidak konsisten dengan tetangga spasialnya sekalipun nilai-nilai nonspasialnya adalah normal untuk sisa dari objek kelas yang sama. Pendeteksian spasial outlier sangat bermanfaat untuk berbagai bidang di sistem informasi geografis (SIG) seperti ekologi, transportasi, kesehatan masyarakat, klimatologi, pelayanan umum dan lain-lain (Lu, dkk., 2003 : 1).

Metode pendeteksian spasial outlier telah banyak dikembangkan oleh ilmuwan. Algoritma pendekatan nilai z (z algorithm) telah dikenalkan oleh Shekhar, dkk (2002). Pada jurnal "*Spatial Weighted Outlier Detection*" (Kou, dkk., 2006), peneliti membandingkan metode z value approach, *weighted z approach* dan *average difference algorithm* untuk mendeteksi spasial outlier dan didapatkan kesimpulan bahwa *weighted z approach* dan *average difference algorithm* lebih baik dalam mendeteksi spasial outlier dibandingkan dengan z value approach. Menurut penelitian yang dilakukan oleh Lu, dkk (2003) "*Algorithms for Spatial Outlier Detection*", didapatkan hasil bahwa metode *iterative z algorithm*, *median algorithm* dan *iterative r algorithm* lebih baik dalam mendeteksi outlier pada data spasial daripada metode z algorithm.

Berdasarkan pada penelitian yang disebutkan di atas diketahui bahwa *weighted z algorithm*, *average difference algorithm*, *median algorithm*, *iterative z algorithm*, dan *iterative r algorithm* merupakan metode yang baik dalam mendeteksi spasial outlier. Dalam penelitian ini, penulis ingin melanjutkan penelitian tersebut yaitu membandingkan metode *weighted z algorithm* dan metode *iterative r algorithm* untuk mengetahui mana dari kedua metode tersebut yang lebih sensitif dalam mendeteksi spasial outlier.

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah yang akan di bahas dalam penelitian ini adalah:

1. Bagaimana menerapkan metode *weighted z algorithm* dan *iterative r algorithm* untuk mendeteksi *outlier* pada data kandungan C-organik tanah di Kabupaten Sampang?
2. Manakah dari kedua metode yang lebih sensitif dan lebih baik dalam mendeteksi *outlier* pada kandungan C-organik tanah di Kabupaten Sampang?

1.3. Batasan Masalah

Data yang digunakan pada penelitian ini data kandungan C-organik tanah di Kabupaten Sampang Madura.

1.4. Tujuan Penelitian

Tujuan dilakukannya penelitian ini adalah:

1. Mempelajari bagaimana penerapan metode *weighted z algorithm* dan *iterative r algorithm* dalam mendeteksi *outlier* pada data kandungan C-organik tanah di Kabupaten Sampang.
2. Membandingkan di antara *weighted z algorithm* dan *iterative r algorithm* untuk mengetahui mana yang lebih sensitif dan lebih baik dalam mendeteksi *outlier* pada data kandungan C-organik tanah di Kabupaten Sampang.

1.5. Manfaat Penelitian

Manfaat penelitian ini adalah didapatkannya informasi mengenai beberapa metode pendeteksian *outlier* yang memiliki ketelitian yang tinggi dalam mendeteksi *outlier* pada data spasial. Bagi masyarakat umum dengan adanya penelitian ini dapat diketahui tempat yang memiliki kandungan C-organik tinggi dan rendah pada tanah di Kabupaten Sampang sehingga dapat diketahui tingkat kesuburan tanah di tempat tersebut.

BAB II

TINJAUAN PUSTAKA

2.1 Data Spasial

Data spasial merupakan sebuah data yang berorientasi geografis dan memiliki sistem koordinat tertentu sebagai dasar referensinya. Data spasial mempunyai dua bagian penting yang membuatnya berbeda dari data lain yaitu :

1. Informasi lokal (spasial), informasi yang berkaitan dengan suatu koordinat baik koordinat geografi (lintang dan bujur) dan koordinat XYZ, termasuk di antaranya informasi datum dan proyeksi.
2. Informasi deskriptif (atribut) atau informasi non spasial yaitu informasi mengenai suatu lokasi yang memiliki beberapa keterangan yang berkaitan dengannya, contohnya : jenis vegetasi, populasi, luasan, kode pos, dan sebagainya (GIS Konsorsium Aceh Nias, 2007 : 2).

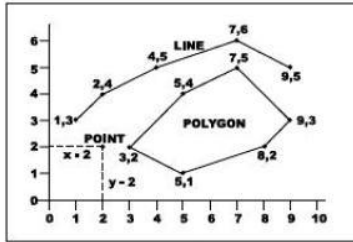
Data spasial kemudian akan diolah dalam sistem informasi geografis (SIG) dan dapat digunakan untuk menjawab beberapa pertanyaan yang berhubungan dengan lokasi, kondisi, pola suatu wilayah.

2.1.1 Format Data Spasial

Secara sederhana format dalam bahasa komputer adalah bentuk dan kode penyimpanan data yang berbeda antara file satu dengan lainnya. Dalam SIG, data spasial dapat direpresentasikan dalam dua format sebagai berikut (Hartoyo, dkk., 2010 : 3-4):

1. Data Vektor

Data vektor merupakan bentuk bumi yang direpresentasikan ke dalam kumpulan garis, area (daerah yang dibatasi oleh garis yang berawal dan berakhir pada titik yang sama), titik dan nodes (merupakan titik perpotongan antara dua buah garis).



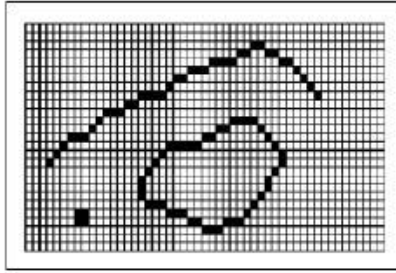
Gambar 2.1. Contoh data vektor

Model data vektor menampilkan, menempatkan, dan menyimpan data spasial dengan menggunakan titik, garis (kurva atau poligon) beserta atributnya. Bentuk dasar representasi data spasial dalam model vektor didefinisikan oleh sistem koordinat kartesian dua dimensi (x , y). Garis atau kurva merupakan sekumpulan titik terurut yang dihubungkan. Sedangkan luasan atau poligon disimpan sebagai kumpulan daftar titik-titik dimana titik awal dan titik akhir poligon memiliki nilai koordinat yang sama.

Keuntungan utama yang didapatkan dari format data vektor adalah ketepatan dalam merepresentasikan fitur titik, batasan dan garis lurus. Hal ini sangat berguna untuk analisis yang membutuhkan ketepatan posisi, misalnya pada basis data batas-batas kadaster. Contoh penggunaan lainnya adalah untuk mendefinisikan hubungan spasial dari beberapa fitur. Kelemahan dari format data vektor yang utama adalah ketidakmampuannya dalam mengakomodasi perubahan gradual.

2. Data Raster

Data raster (disebut juga dengan sel *grid*) adalah data yang dihasilkan dari sistem penginderaan jauh. Pada data raster, objek geografis direpresentasikan sebagai struktur sel *grid* yang disebut dengan *pixel* (*picture element*) (Hartoyo, dkk., 2010 : 2).



Gambar 2.2. Contoh data raster

Pada data raster, resolusi (definisi visual) tergantung pada ukuran *pixel*-nya. Dengan kata lain, resolusi *pixel* menggambarkan ukuran sebenarnya di permukaan bumi yang diwakili oleh setiap *pixel* pada citra. Semakin kecil ukuran permukaan bumi yang direpresentasikan oleh satu sel, semakin tinggi resolusinya. Data raster sangat baik untuk merepresentasikan batas-batas yang berubah secara gradual, seperti jenis tanah, kelembaban tanah, vegetasi, suhu tanah dan sebagainya. Keterbatasan utama dari data raster adalah besarnya ukuran file; semakin tinggi resolusi *grid*-nya semakin besar pula ukuran filenya dan sangat tergantung pada kapasitas perangkat keras yang tersedia (GIS Konsorsium Aceh Nias, 2007 : 3).

Model data raster menampilkan, menempatkan, dan menyimpan data spasial dengan menggunakan struktur matriks atau *pixel-pixel* yang membentuk *grid*. Setiap *pixel* memiliki atribut tersendiri, termasuk koordinatnya yang unik. Akurasi model ini sangat tergantung pada resolusi atau ukuran *pixel*-nya di permukaan bumi. *Entity* spasial raster di dalam *layers* yang secara fungsionalitas direlasikan dengan unsur-unsur petanya. Contoh unsur spasial raster adalah citra satelit (Landsat, Ikonos), citra radar, dan sebagainya.

Masing-masing format data mempunyai kelebihan dan kekurangan. Pemilihan format data yang digunakan sangat tergantung pada tujuan penggunaan, data yang tersedia, volume data yang dihasilkan, ketelitian yang diinginkan, serta kemudahan dalam analisis. Data vektor relatif lebih ekonomis dalam hal ukuran file dan presisi dalam lokasi, tetapi sangat sulit digunakan dalam komputasi matematik. Sedangkan data raster biasanya membutuhkan ruang penyimpanan file yang lebih besar dan presisi lokasinya lebih rendah, tetapi lebih mudah digunakan secara matematis.

2.1.2 Sumber Data Spasial

Data spasial merupakan salah satu syarat sistem informasi geografis , yang dapat diperoleh dari beberapa sumber antara lain (Hartoyo, dkk., 2010 : 4):

1. Peta analog
Peta analog yaitu peta dalam bentuk cetak, contohnya antara lain peta topografi, peta tanah dan sebagainya. Pada umumnya peta analog dibuat dengan menggunakan tehnik kartografi, di mana kemungkinan besar memiliki referensi spasial seperti koordinat, skala, arah mata angin dan sebagainya. Dalam tahapan SIG sebagai keperluan sumber data, peta analog dikonversi menjadi peta digital dengan cara merubah format raster menjadi format vektor melalui proses digitasi sehingga dapat menunjukkan koordinat sebenarnya di permukaan bumi.
2. Data sistem penginderaan jauh
Data penginderaan jauh merupakan sumber data yang terpenting bagi SIG karena ketersediaannya secara berkala dan mencakup area tertentu. Adapun contoh dari data sistem penginderaan jauh antara lain citra satelit, foto udara dan sebagainya. Dengan adanya bermacam-macam satelit di ruang angkasa dengan spesifikasinya masing-masing, kita bisa memperoleh berbagai jenis citra satelit untuk beragam tujuan pemakaian. Data sistem penginderaan jauh biasanya direpresentasikan dalam format raster.
3. Data hasil pengukuran lapangan
Data pengukuran lapangan yang dihasilkan berdasarkan tehnik perhitungan tersendiri, pada umumnya data ini merupakan sumber data atribut contohnya: batas administrasi, batas kepemilikan lahan, batas persil, batas hak pengusahaan hutan dan lain-lain.
4. Data GPS (*Global Positioning System*)
Teknologi GPS memberikan terobosan penting dalam menyediakan data bagi SIG. Keakuratan pengukuran GPS semakin tinggi dengan berkembangnya teknologi. Data ini biasanya direpresentasikan dalam format vektor.

2.2 Sistem Informasi Geografis

Menurut Hartoyo (2010 : 1), sistem informasi geografis adalah sistem berbasis komputer digunakan untuk memasukkan, menyimpan, memperbaiki, memperbaharui, mengelola, memanipulasi, dan menganalisis sistem informasi geografis. Sistem informasi geografis memerlukan data masukan agar dapat berfungsi dan memberikan informasi lain sebagai hasil analisisnya. Data masukan tersebut dapat diperoleh dari tiga sumber, yaitu:

1. Data Lapangan
Data ini diperoleh langsung dari pengukuran lapangan secara langsung, seperti pH tanah, salinitas air, curah hujan suatu wilayah, dan sebagainya.
2. Data Peta
Informasi yang telah disimpan pada peta dikonversikan ke dalam bentuk digital. Misalnya peta geologi, peta tanah dan sebagainya. Apabila data sudah terekam dalam bentuk peta, tidak lagi diperlukan data lapangan, kecuali untuk pengecekan kebenarannya.
3. Data Citra Penginderaan Jauh
Citra penginderaan jauh yang berupa foto udara dapat diinterpretasikan terlebih dahulu sebelum dikonversi ke dalam bentuk digital. Sedangkan citra yang diperoleh dari satelit yang sudah dalam bentuk dapat langsung digunakan setelah diadakan koreksi seperlunya.

Sistem informasi geografis memiliki kemampuan untuk menghubungkan berbagai data pada suatu titik tertentu di bumi, kemudian menggabungkannya, menganalisa, dan akhirnya memetakan hasilnya. Dalam sistem informasi geografis, diperlukan kesatuan sistem yang terdiri dari perangkat keras, perangkat lunak, data geografis yang benar dan sumber daya manusia untuk memformulasikan suatu masalah geografis agar dapat dilakukan analisis lebih lanjut (GIS Konsorsium Aceh Nias, 2007 : 2).

2.3 Outlier

Sebuah *outlier* didefinisikan sebagai sebuah titik data pada suatu database di mana sangat berbeda dibandingkan dengan titik data pada database secara umum dengan suatu ukuran tertentu. Titik ini seringkali mempunyai informasi yang sangat berguna yang

didefinisikan data pada kelakuan sistem yang abnormal (Handriyani, dkk., 2009 : 1).

Berikut ini adalah beberapa definisi *outlier* (Soemartini, 2007 : 6):

1. *Outlier* adalah suatu data yang menyimpang dari sekumpulan data yang lain.
2. *Outlier* adalah pengamatan yang tidak mengikuti sebagian besar pola dan terletak jauh dari pusat data.
3. *Outlier* adalah pengamatan yang jauh dari pusat data yang mungkin berpengaruh besar terhadap koefisien regresi.

Jika terdapat masalah yang berkaitan dengan *outlier*, maka diperlukan alat diagnosis yang dapat mengidentifikasi masalah *outlier*.

Berdasarkan perbandingan dengan objek lainnya, *outlier* dibagi menjadi dua yaitu *global outlier* dan *local outlier*. *Global outlier* adalah suatu objek pada data yang berbeda nyata dengan objek lain secara keseluruhan, sedangkan *local outlier* adalah suatu objek yang berbeda nyata dengan objek lain pada suatu kelompok walaupun objek tersebut sebenarnya tidak berbeda nyata terhadap data secara keseluruhan (Shekhar, dkk., 2003 : 1-2).

Menurut Shekhar, dkk. (2002), *outlier* pada data spasial dapat dikelompokkan menjadi tiga kelas yaitu:

- *Set-based outlier* adalah objek data yang memiliki atribut non-spasial berbeda nyata dengan objek lain tanpa memperhatikan hubungan spasial.
- *Multi-dimensional space-based outlier* adalah *spatial outliers* yang objek spasial yang memiliki nilai atribut non-spasial berbeda nyata dengan nilai atribut objek lain yang masih satu dalam lingkungan spasial, dimana lingkungan spasial didefinisikan berdasarkan pada jarak Euclidean.
- *Graph-based outlier*, memiliki pengertian yang sama dengan *Multi-dimensional space-based outliers*, namun pada *graph-based outliers*, lingkungan spasial didasarkan pada *graph connectivity*.

Pada penelitian ini, *spasial outlier* yang dimaksud adalah *multi-dimensional space-based outliers*.

2.4 Spasial *Outlier*

Spasial *outlier* adalah suatu objek yang tidak konsisten dengan tetangga spasialnya sekalipun nilai-nilai non-spasialnya normal untuk sisa dari objek kelas yang sama. Sehubungan dengan hal ini, metode pendeteksian *outlier* non-spasial (tradisional *outlier*) tidak mendeteksi dengan akurat tanpa mempertimbangkan penempatan lokasi keruangannya (lokasi spasialnya) (Han, dkk., 2001 : 2-3).

Menurut Haining (1990 : 13) spasial *outlier* digambarkan sebagai suatu pengamatan yang tidak biasa berkaitan dengan nilai-nilai tetangganya. Sementara menurut Barnett dan Lewis (1994) spasial *outlier* ditunjuk sebagai nilai-nilai *pixel* yang berbeda nyata dengan tetangganya meskipun itu bukanlah suatu kasus didalam keseluruhan skala perumpamaan. Secara keseluruhan, Shekhar, dkk (2003 : 4) memulai memperkenalkan suatu definisi umum mengenai spasial *outlier*, yaitu suatu titik spasial tertentu dimana nilai-nilai atribut non-spasialnya berbeda nyata dari titik-titik lain yang masih menjadi tetangganya.

Menurut Kou, dkk (2006 : 613), terdapat beberapa hal yang membedakan spasial *outlier* dengan non-spasial *outlier* (tradisional *outlier*) yaitu:

- a. Tradisional *outlier* difokuskan pada perbandingan secara global pada semua objek yang ada, sedangkan spasial *outlier* lebih difokuskan pada perbandingan secara lokal diantara lingkungan spasialnya (*spatial neighborhood*).
- b. Pendeteksian tradisional *outlier* berkaitan dengan nilai, sifat, dan kategori, sedangkan pada spasial *outlier* dibutuhkan proses yang lebih rumit karena data spasial dapat berupa titik, garis, poligon ataupun objek tiga dimensi.
- c. Pada pendeteksian spasial *outlier*, hubungan secara spasial harus diperhitungkan karena dalam salah satu aturan dalam bidang geografi dikatakan bahwa “segala sesuatu berkaitan dengan sesuatu yang lain, dua benda yang berdekatan akan memiliki hubungan yang lebih kuat daripada dua benda yang saling berjauhan”, sedangkan pada pendeteksian tradisional *outlier* tidak ada hubungan yang perlu dipertimbangkan.

Namun, apabila data spasial hanya terdiri atas satu cluster (*neighborhood*) maka pendeteksian spasial *outlier* memiliki prinsip sama dengan tradisional *outlier*.

2.5 Metode Pendeteksian Spasial Outlier

2.5.1 Formulasi Algoritma

Formulasi yang digunakan dalam pendeteksian spasial *outlier* dinyatakan sebagai berikut (Lu, dkk., 2003 : 2):

1. Dinyatakan set data spasial, $X = (x_1, x_2, \dots, x_n)$ dalam ruang dengan dimensi $p \geq 1$.
2. Suatu fungsi atribut f digambarkan sebagai suatu pemetaan dari X untuk R (set dari angka riil). Fungsi atribut $f(x_i)$ menyatakan nilai atribut dari titik spasial x_i .
3. Untuk titik yang ditentukan x_i , $NN_k(x_i)$ menandakan k tetangga terdekat x_i , dimana $k = k(x_i)$ tergantung pada nilai x_i , untuk $i = 1, 2, \dots, n$.
4. Fungsi tetangga g digambarkan sebagai suatu pemetaan dari X untuk R , dimana untuk masing-masing x_i , $g(x_i)$ mengembalikan suatu ringkasan statistik nilai atribut dari semua titik spasial di dalam $NN_k(x_i)$. Sebagai contoh, $g(x_i)$ berupa rata-rata nilai atribut k tetangga terdekat x_i . Untuk mendeteksi spasial *outlier*, kita bandingkan nilai atribut dari tiap titik x_i dengan nilai atribut tetangganya $NN_k(x_i)$.
5. Perbandingan dilaksanakan melalui suatu fungsi pembanding h , yang merupakan suatu fungsi dari f dan g . Ada banyak pilihan untuk format h . Sebagai contoh, h dapat berupa selisih, $f - g$, atau rasio, $\frac{f}{g}$.
6. Pandang $y_i = h(x_i)$ untuk $i = 1, 2, \dots, n$. y_i merupakan standardisasi dari fungsi $h(x_i)$ yang diberi tanda mutlak.
7. Titik x_i adalah spasial *outlier* jika y_i adalah suatu nilai yang ekstrim pada set $(y_1, y_2, \dots, y_n)$.

Dapat digarisbawahi definisi tersebut tergantung pada pemilihan fungsi k , g dan h yang digunakan (Lu, dkk., 2003 : 614).

2.5.2 k Tetangga Terdekat

Metode *k-nearest neighbor* atau yang sering disebut sebagai *k* tetangga terdekat (*k* – *NN* atau *KNN*) adalah sebuah pendekatan pengklasifikasian data dengan prinsip kerja yang mudah dan efisien. Metode tersebut mengasumsikan data yang saling berdekatan sebagai tetangga terdekat dan dimasukkan dalam kelas yang sama. Tetangga terdekat tersebut ditentukan berdasarkan perhitungan vektor dan jarak. Jarak yang sering digunakan untuk menghitung *k* tetangga terdekat adalah jarak Euclid (Jiang Sheng, 2002 : 1).

Menurut Shekhar, dkk (2003 : 155), metode *Connectivity Clustered Acces Method* (CCAM) disarankan sebagai alternatif metode clustering yang digunakan untuk menentukan *k* tetangga terdekat. Pada metode CCAM menggunakan *edge* yaitu garis koneksi dari titik-titik spasial sebagai ukuran kedekatan dalam menentukan cluster.

Edge dapat dihitung dengan menggunakan jarak Euclid. Proses algoritma CCAM diawali dengan menganggap seluruh titik pada data set sebagai satu cluster. Kemudian, titik dengan konektivitas terdekat membentuk cluster tersendiri. *Edge* tersebut merupakan ukuran spasial (Lin dan Ye, 2008 : 6-8).

2.5.3 Ukuran Keruangan (Ukuran Spasial).

Keterikatan atau hubungan spasial adalah salah satu sifat dalam lingkungan geografi, karakter dari lokasi yang saling berdekatan akan menjadi berkorelasi. Terdapat beberapa alasan terkait dengan adanya suatu hubungan spasial. Kemungkinan pertama, karena terdapat hubungan korelasi spasial sederhana; apapun yang menyebabkannya, objek pada suatu lokasi memiliki kemiripan dengan objek lain yang berdekatan. Kemungkinan kedua adalah hubungan kausalitas spasial, objek pada suatu lokasi tertentu akan secara langsung mempengaruhi objek lain di sekitarnya. Kemungkinan ketiga adalah adanya interaksi spasial, mobilitas masyarakat, barang dan informasi menciptakan hubungan yang nyata antara beberapa lokasi (Upton dan Fingleton, 1985: 214).

Ukuran kedekatan didapatkan dari perhitungan hubungan spasial antara titik yang berbeda. Setiap objek spasial memiliki posisi dalam sistem koordinat tertentu dan masing-masing memiliki

keterkaitan atau hubungan spasial. Hubungan spasial tersebut dapat diukur dengan tipologi, jarak, dan hubungan langsung (Schwering dan Raubal, 2005 : 2-3). Jarak yang dapat menggambarkan kedekatan tersebut adalah jarak *Euclidean* dan *Manhattan*.

Jarak *Euclidean* adalah jarak lurus antara dua titik dengan menggunakan fungsi *Pythagoras*. Jika titik $P=(x_1, x_2)$ pada suatu permukaan, jarak *Euclidean*, $d=(O, P)$ dari titik P ke titik pusat $O=(0,0)$ berdasarkan pada teorema *Pythagoras* adalah:

$$d(O, P) = \sqrt{x_1^2 + x_2^2} \quad (2.1)$$

Secara umum, jika titik P memiliki koordinat sehingga $P=(x_1, x_2, \dots, x_p)$, maka jarak *Euclidean* dari titik P ke titik pusat $O=(0, 0, \dots, 0)$ adalah:

$$d(O, P) = \sqrt{x_1^2 + x_2^2 + \dots + x_p^2} \quad (2.2)$$

Sedangkan jarak *Euclidean* antara dua titik P dan Q dengan koordinat $P=(x_1, x_2, \dots, x_p)$ dan $Q=(y_1, y_2, \dots, y_p)$ untuk n -dimensi adalah:

$$d(P, Q) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_p - y_p)^2} \quad (2.3)$$

(Johnson dan Winchern, 2002 : 30).

Jarak *Euclidean* antara dua titik P dan Q dengan koordinat $P=(x_1, x_2)$ dan $Q=(y_1, y_2)$ untuk dua dimensi adalah :

$$d(P, Q) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \quad (2.4)$$

(Wikipedia, 2013).

Jarak *Manhattan* adalah jarak yang dibatasi untuk berjalan sejajar dengan sumbu (Upton dan Fingleton, 1985 : 214). Pada jarak *Manhattan* jarak antara dua objek adalah jumlah dari harga mutlak selisih koordinat kedua objek. Jarak *Manhattan* antara dua objek (P dan Q) pada dimensi- p adalah sebagai berikut (Krause dalam Wikipedia, 1987):

$$d(P, Q) = \sum_{i=1}^p |x_i - y_i| \quad (2.5)$$

di mana :

$d(P, Q)$: jarak *Manhattan* antara objek P dengan objek Q

x_i : objek P pada dimensi- p di mana $x_i = (x_1, x_2, \dots, x_p)$

y_i : objek Q pada dimensi- p di mana $y_i = (y_1, y_2, \dots, y_p)$

Sehingga pada ruang dua dimensi, jarak *Manhattan* antara objek p dengan objek q adalah (Krause dalam Wikipedia, 1987):

$$d(p, q) = |p_1 - q_1| + |p_2 - q_2| \quad (2.6)$$

Dibandingkan dengan jarak *Euclidean*, jarak *Manhattan* lebih baik digunakan di daerah yang memiliki banyak batasan yang tidak dapat dilanggar seperti di perkotaan yang banyak dibatasi oleh gedung. Oleh karena itulah, jarak *Manhattan* dapat disebut sebagai jarak *City Blok* (Upton dan Fingleton, 1985 : 214).

Dalam penelitian ini, digunakan jarak *Euclidean* untuk menggambarkan kedekatan antara objek spasial. Untuk menentukan lingkungan spasial dari suatu objek spasial perlu diketahui k tetangga terdekat dan jarak *Euclidean* mampu mewakili jarak terpendek antara dua objek spasial.

2.5.4 *z Algorithm*

Spasial *outlier* mengindikasikan bahwa dapat atau tidaknya suatu titik dikatakan sebagai spasial *outlier* tergantung pada selisih antara nilai atribut suatu titik tersebut dengan rata-rata nilai atribut dari seluruh titik yang merupakan tetangganya (Shekhar, dkk., 2003). Algoritma yang diusulkan Shekhar, dkk (2003) menggunakan atribut non-spasial tunggal yang membandingkan selisih antara tetangga-tetangga suatu titik dan mengidentifikasi spasial *outlier* melalui perhitungan yang efisien dengan cara menghitung fungsi *algebraic aggregate* umum. Selain itu, Shekhar, dkk (2003) juga mempertimbangkan struktur grafik data spasial dan memanfaatkan suatu metode grafis untuk mendeteksi spasial *outlier*. Secara rinci, metode ini membandingkan selisih antara nilai atribut dari suatu titik data dan rata-rata atribut tetangganya, dan memeriksa keseluruhan data ke dalam distribusi normal. Jika nilai pengujian pada titik-titik tersebut lebih besar dari suatu taraf kepercayaan yang ditetapkan maka titik tersebut dinyatakan sebagai spasial *outlier*.

Algoritma pendeteksian spasial *outlier* dengan metode pendeteksian nilai z dinyatakan sebagai berikut (Shekhar, dkk., 2003 : 141):

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

$$g(x_i) = \frac{1}{k} \sum_{x \in NN_k(x_i)} f(x) \quad (2.7)$$

Dan fungsi pembandingan dengan persamaan :

$$h_i = h(x_i) = f(x_i) - g(x_i) \quad (2.8)$$

2. μ_h dan σ_h secara berturut-turut merupakan rata-rata dan standar deviasi dari fungsi pembandingan $h(x_i)$, yang kemudian digunakan untuk menghitung statistic uji z di mana:

$$Z_{h(x_i)} = \frac{h(x_i) - \mu_h}{\sigma_h} \quad (2.9)$$

3. Objek spasial yang dinyatakan sebagai spasial *outlier* adalah objek yang memiliki statistik uji $Z_{h(x_i)} > \theta$ di mana $\theta = 2$ untuk selang kepercayaan 95%.

2.5.5 *Weighted z Algorithm*

Weighted z algorithm merupakan pengembangan dari *z algorithm*. Perbedaan *weighted z algorithm* dengan *z algorithm* terletak pada fungsi tetangga (*neighborhood function*) g . Jika pada *z algorithm*, fungsi tetangga (*neighborhood function*) yang digunakan merupakan rata-rata nilai atribut non-spasial dari masing-masing lingkungan spasial, sedangkan pada *weighted z algorithm* menggunakan *neighborhood function* berupa *weighted mean* (rata-rata terboboti) dari masing-masing lingkungan spasial (Kou, dkk., 2006 : 614).

Weighted mean (rata-rata terboboti) adalah salah satu ukuran pemusatan. Rata-rata terboboti (M_w) dari n objek pengamatan ($x_1, x_2, \dots, x_n$) adalah (Glodberg dan Kercheval, 2002 : 3):

$$M_w(x_1, x_2, \dots, x_n) = \frac{\sum_{j=1}^n w_j x_j}{\sum_{j=1}^n w_j} \quad (2.10)$$

atau

$$M_w(x_1, x_2, \dots, x_n) = \frac{w_1 x_1 + w_2 x_2 + \dots + w_n x_n}{w_1 + w_2 + \dots + w_n} \quad (2.11)$$

Di mana w_i adalah *weight* (bobot) dari objek x_i . Semakin besar bobot w_i , maka semakin besar pula kontribusi objek x_i terhadap rata-rata terboboti. Pada umumnya, bobot w_i menggambarkan besarnya kepentingan masing-masing objek x_i .

Rata-rata terboboti digunakan pada saat ada beberapa objek x_i yang memberikan kontribusi berbeda dengan objek lain terhadap rata-rata. Rata-rata terboboti sama dengan rata-rata aritmatik pada saat masing-masing objek x_i memberikan kontribusi yang sama pada rata-rata, dengan kata lain jika semua bobot w_i memiliki nilai yang sama, maka rata-rata terboboti sama dengan rata-rata aritmatik. Rata-rata terboboti pada umumnya digunakan pada saat akan mengkombinasikan beberapa nilai rata-rata dari suatu populasi dengan banyak sampel yang berbeda (Hardy, dkk., 1988 : 13-14).

Menurut Kou, dkk (2006 : 614), objek spasial pada data spasial bukan merupakan suatu objek yang terisolasi dari lingkungan sekitarnya. Masing-masing objek spasial memiliki sifat saling mempengaruhi objek spasial lain yang masih berada dalam satu lingkungan spasial yang sama. Semakin dekat objek spasial dengan objek spasial yang lain, maka semakin besar pengaruh objek spasial tersebut bagi objek spasial yang lain. Berdasarkan sifat tersebut, maka dibuat suatu algoritma pendeteksi spasial *outlier*. Bobot ditentukan berdasarkan hubungan spasial yaitu jarak *Euclidean* antara dua objek spasial.

Penentuan bobot melalui hubungan spasial antara objek x_i dan objek tetangganya x_j . Nilai bobot untuk x_j terhadap x_i sebesar 0

sampai 1 dan jumlah dari bobot semua objek tetangga x_i adalah 1. Jika x_j adalah tetangga ke r dari x_i , maka bobot dari x_j adalah (Kou, dkk, 2006):

$$w_j = \sum_{p=1}^q \alpha_p \cdot \frac{S_{pr}}{\sum_{l=1}^k S_{pl}} \quad (2.12)$$

Di mana :

- w_j : bobot dari objek x_j terhadap objek x_i
- q : banyaknya sifat spasial yang menentukan bobot.
- S_{pr} : *particular spatial property* atau atribut spasial (S_p) untuk objek tetangga ke r dari x_i , dalm hal ini S_{pr} merupakan *invers* jarak antara x_i dengan x_r .
- S_{pl} : *particular spatial property* atau atribut spasial (S_p) untuk objek tetangga ke l dari x_i , dalm hal ini S_{pl} merupakan *invers* jarak antara x_i dengan x_l .
- α_p : faktor yang menyatakan tingkat kepentingan sifat spasial

$$S_p \text{ dan } \sum_{p=1}^q \alpha_p = 1.$$

Rata-rata terboboti (M_w) didapatkan dengan mengalikan bobot dengan nilai atribut non-spasial y_j untuk masing-masing *neighbor* x_j (Kou, dkk, 2006).

$$M_w(x_i) = \sum_{j=1}^k f(x_j) * w_j \quad (2.13)$$

Di mana :

- $M_w(x_i)$: rata-rata terboboti untuk objek spasial x_i
- k : banyak tetangga terdekat dengan objek x_i
- $f(x_j)$: fungsi atribut tetangga, x_j
- w_j : bobot dari objek tetangga (x_j) terhadap objek x_i

Berikut ini algoritma pendeteksian spasial *outlier* dengan *weighted z algorithm* (Kou, dkk., 2006 : 614-615):

1. Untuk masing-masing objek, menentukan k tetangga terdekat menggunakan jarak *Euclidean* antara dua objek berdasarkan persamaan 2.3.
2. Untuk masing-masing objek x_i , menghitung *neighborhood function* $g(x_i)$ yang merupakan rata-rata terboboti (M_w) dari atribut non-spasial untuk semua objek spasial x_i berdasarkan persamaan 2.13.
3. Menghitung fungsi pembandingan yaitu selisih antara nilai atribut non-spasial x_i dengan *neighborhood function* $g(x_i)$ atau dilambangkan dengan $h(x_i)$.

$$h(x_i) = f(x_i) - g(x_i) \quad (2.14)$$

4. Menghitung $h(x_i)$ yang telah distandardisasi berdasarkan rata-rata dan standar deviasi dari $h(x_i)$ sesuai dengan persamaan 2.9 yang kemudian akan dijadikan faktor untuk menentukan spasial *outlier*.
5. Mengidentifikasi objek dengan nilai $Z(x_i) > \theta$ sebagai spasial *outlier*, di mana $\theta = 2$ untuk selang kepercayaan 95% .

2.5.6 Iterative r Algorithm

Seperti pada algoritma pendeteksian spasial *outlier* yang lain, pendeteksian spasial *outlier* tergantung pada *neighborhood function* $g(x_i)$ dan fungsi pembandingan $h(x_i)$. Pemilihan dari $g(x_i)$ dan $h(x_i)$ menentukan kinerja setiap algoritma. Pada *iterative r algorithm*, *neighborhood function* $g(x_i)$ menduga titik spasial x_i yang diambil menjadi nilai rata-rata atribut dari semua k tetangga terdekat ($NN_k(x_i)$). Fungsi pembandingan $h(x_i)$ diambil dari rasio antara $f(x_i)$ dan $g(x_i)$.

Algoritma pendeteksian spasial *outlier* dengan metode *iterative r algorithm* adalah sebagai berikut (Lu, dkk., 2003 : 2):

1. Untuk masing-masing titik x_i , hitung k tetangga terdekat ($NN_k(x_i)$) dengan cara menghitung jarak *Euclidean* antara dua objek.
2. Menghitung *neighborhood function* $g(x_i)$ dengan persamaan

$$g(x_i) = \frac{1}{k} \sum_{x \in NN_k(x_i)} f(x) \quad (2.15)$$

dan fungsi pembanding dengan persamaan

$$h_i = h(x_i) = \frac{f(x_i)}{g(x_i)} \quad (2.16)$$

3. Pandang h_q atau h_q^{-1} yang menandakan nilai maksimum dari $h_1, h_2, \dots, h_n, h_1^{-1}, h_2^{-1}, \dots, h_n^{-1}$. Untuk tresshold yang diberikan θ , jika h_q atau $h_q^{-1} \geq \theta$, maka x_q merupakan spasial *outlier*.
4. Ganti $f(x_q)$ dengan $g(x_q)$. Untuk setiap titik spasial x_i dengan $NN_k(x_i)$ yang beranggotakan x_q , hitung kembali $g(x_i)$ dan $h(x_i)$.
5. Ulangi langkah 2 dan 3 sampai kondisi threshold tidak dijumpai atau banyaknya spasial *outlier* sama dengan m.

Menurut Lu, dkk (2003 : 2), pemilihan paling sederhana untuk θ pada *iterative r algorithm* adalah $\theta = 1$. Nilai θ dapat bernilai lebih dari 1 tergantung pada konteks yang berbeda.

2.6 Pemilihan Metode Terbaik

Dalam penelitian ini, pemilihan metode terbaik didasarkan pada dua kriteria. Kriteria pertama, kesamaan spasial *outlier* yang dideteksi kedua metode dengan tradisional *outlier* yang telah ditentukan dan kriteria yang kedua adalah kemampuan kedua metode dalam mengatasi adanya *masking effect* dan *swamping effect*.

Masking effect adalah suatu gejala tertutupnya suatu *outlier* (*outlier* kedua) oleh *outlier* yang lain (*outlier* pertama) sehingga *outlier* kedua tidak terdeteksi sebagai *outlier*. *Outlier* kedua baru

dapat terdeteksi sebagai *outlier* apabila *outlier* pertama dihapus. *Masking effect* terjadi saat suatu *outlier* mampu mencodongkan rata-rata dan ragam data ke arah *outlier*, sehingga jarak antara rata-rata dengan *outlier* menjadi kecil (Gal-Irad, 2005 : 7-8).

Swamping effect adalah suatu gejala di mana suatu *outlier* (*outlier* pertama) mampu membawa atau menenggelamkan objek yang lain sehingga menjadi *outlier* (*outlier* kedua). Sehingga dapat dikatakan bahwa apabila *outlier* pertama tidak ada maka tidak akan muncul *outlier* kedua. *Swamping effect* terjadi saat *outlier* mampu mencondongkan rata-rata dan ragam data menjauh dari objek non-*outlier*, sehingga jarak antara objek non-*outlier* dengan rata-rata menjadi besar (Davies dan Gather, 1993 : 784).

Pada kriteria yang pertama, suatu metode dikatakan lebih baik dari metode yang lainnya bila *top outlier* yang terdeteksi sama dengan tradisional *outlier* yang telah ditentukan pada data spasial, dari jumlah maupun lokasi titiknya. Pada kriteria kedua, suatu metode dikatakan lebih baik dari metode yang lainnya apabila metode tersebut mampu mengatasi *masking effect* yang berarti bisa menghindari amatan *outlier* dideteksi sebagai amatan normal dan mampu mengatasi *swamping effect* yang berarti bisa menghindari amatan normal terdeteksi sebagai amatan *outlier*.

2.7 C-organik.

Tanah di Madura umumnya terbentuk dari bahan induk batu kapur di bawah pengaruh iklim (curah hujan) yang tegas antara bulan basah dan kering. Sehingga tanah yang terbentuk berbeda karakteristiknya dengan tanah yang ada daerah yang tidak mengenal bulan kering. Tanah-tanah di Madura mempunyai reaksi tanah netral hingga alkalis, kandungan bahan organik dan nitrogen total rendah, P total sedang hingga tinggi dan basa kalsium tinggi. Tanah-tanah semacam ini mempunyai kadar liat yang cukup tinggi, dan dengan pengelolaan yang baik dapat diarahkan untuk mempunyai produktivitas tanah yang tinggi. Kandungan bahan organik dan nitrogen sangat rendah, sehingga dalam pengelolaan tanah kedua hal tersebut menjadi perhatian utama.

Kesuburan tanah adalah kemampuan tanah untuk menyediakan hara, air dan oksigen dalam keadaan yang seimbang bagi tanaman. Kemampuan ini dipengaruhi oleh sifat fisika, kimia dan biologi tanah. Dari sudut kimia, kesuburan tanah diartikan

kemampuan tanah untuk menyediakan hara yang cukup bagi tanaman. Penilaian status kesuburan tanah biasanya didasarkan kandungan Nitrogen, Fosfor dan Kalium, karena nutrisi makro ini dibutuhkan dalam jumlah banyak. Selanjutnya ketersediaan suatu unsur hara dipengaruhi oleh faktor tanah seperti tekstur, kapasitas tukar kation, kandungan bahan organik dan pH tanah (Supriyadi, 2007 : 124-125).

Bahan organik adalah bagian dari tanah yang merupakan suatu sistem kompleks dan dinamis, yang bersumber dari sisa tanaman dan atau binatang yang terdapat di dalam tanah yang terus menerus mengalami perubahan bentuk karena dipengaruhi oleh faktor biologi, fisika dan kimia. Bahan organik tanah adalah semua jenis senyawa organik yang terdapat di dalam tanah, termasuk serasah, fraksi bahan organik ringan, biomassa mikroorganisme, bahan terlarut di dalam air, dan bahan organik yang stabil atau humus. Bahan organik merupakan presentase kesuburan dalam tanah yang terdiri dari berbagai ikatan karbon (Suryani, 2007 : 18-19).

Bahan organik memiliki peranan penting dalam menentukan kemampuan tanah untuk mendukung tanaman sehingga jika kadar bahan organik tanah menurun kemampuan tanah dalam mendukung produktivitas tanaman juga menurun. Selain itu, bahan organik juga mempunyai peranan penting dalam mempengaruhi tingkat kesuburan tanah. Makin tinggi kandungan bahan organik maka relatif tinggi pula tingkat kesuburan tanahnya. Bahan organik dapat mempengaruhi tingkat ketersediaan unsur hara bagi tanaman, misalnya nitrogen, fosfor dan sulfur. Disamping itu bahan organik juga berperan dalam meningkatkan KTK (kapasitas tukar kation), memperbaiki struktur tanah dan meningkatkan ketersediaan air. Presentase C-organik mencirikan kandungan bahan organik, sedangkan C/N menunjukkan tingkat dekomposisi bahan organik (Saidi, 2004 : 136).

Bahan organik tanah berpengaruh terhadap sifat-sifat kimia, fisik maupun biologi tanah. Fungsi bahan organik di dalam tanah antara lain :

1. Berpengaruh langsung maupun tidak langsung terhadap ketersediaan hara.
2. Membentuk agregat tanah yang lebih baik dan memantapkan agregat yang telah terbentuk sehingga aerasi, permeabilitas dan infiltrasi menjadi lebih baik.

3. Meningkatkan retensi air yang dibutuhkan bagi pertumbuhan tanaman.
4. Meningkatkan retensi unsur hara melalui peningkatan muatan di dalam tanah.
5. Meningkatkan suhu tanah.
6. Menyediakan energi bagi organisme tanah.
7. Meningkatkan organisme saprofit dan menekan organisme parasit bagi tanah.
8. Memperbaiki karakteristik lingkungan.
9. Produktivitas tanah tinggi secara berkelanjutan.
10. Dapat mengurangi keracunan unsur hara mikro.
11. Memperbaiki tata kehidupan jasad tanah, khususnya bakteri sehingga seluruh proses mikrobiologis dalam tanah berjalan sempurna.

Pemberian bahan organik ke dalam tanah memberikan dampak yang baik terhadap tanah, tempat tumbuh tanaman. Tanaman akan memberikan respon yang positif apabila tempat tanaman tersebut tumbuh memberikan kondisi yang baik bagi pertumbuhan dan perkembangannya. Bahan organik yang ditambahkan ke dalam tanah menyediakan zat pengatur tumbuh tanaman yang memberikan keuntungan bagi pertumbuhan tanaman seperti vitamin, asam amino, auksin dan giberelin yang terbentuk melalui dekomposisi bahan organik. Bahan organik yang ditambahkan ke dalam tanah mengandung karbon yang tinggi. Pengaturan jumlah karbon di dalam tanah meningkatkan produktivitas tanaman dan kelanjutan umur tanaman karena dapat meningkatkan kesuburan tanah dan penggunaan hara secara efisien (Suryani, 2007 : 19-22).

BAB III

METODE PENELITIAN

3.1. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder kadar C-organik tanah di Kabupaten Sampang yang diperoleh dari hasil penelitian tanaman tembakau tahun 2008 kerjasama Universitas Brawijaya dengan Balai Tanaman Tembakau dan Serat (BALITTAS). Data tersebut merupakan kondisi kandungan organik pada titik-titik wilayah tertentu di wilayah Kabupaten Sampang. Pada lokasi di Kabupaten Sampang merupakan lokasi pada daerah dataran rendah karena daerah Sampang merupakan daerah di sekitar perairan (pantai) dengan tinggi permukaan lokasi yang rendah. Kadar C-organik adalah penyusun dari bahan organik tanah yang didefinisikan sebagai ukuran kesuburan tanah pada suatu lokasi (Setyani, 2011). Prosentase kandungan C-organik di Kabupaten Sampang terletak antara 0.38–1.13 (Supriyadi, 2007 : 128). Data selengkapnya dapat dilihat pada Lampiran 1.

3.2. Metode Penelitian

Tahapan-tahapan analisis pada penelitian ini adalah sebagai berikut:

1. Menentukan *tradisional outlier* pada data kadar C-organik untuk dijadikan sebagai pembanding terhadap spasial *outlier* yang akan ditemukan dengan menggunakan metode *weighted z algorithm* dan *iterative r algorithm*.
2. Menentukan spasial *outlier* berdasarkan metode *weighted z algorithm* dan *iterative r algorithm*. Penerapan metode tersebut menggunakan *software* Minitab 15, dengan cara memanggil macro minitab untuk masing-masing metode.
3. Memilih metode terbaik berdasarkan tingkat kesensitifan metode dalam mendeteksi spasial *outlier* dan kemampuan kedua metode dalam mengatasi adanya *masking effect* dan *swamping effect*.

3.2.1 Pendeteksian Tradisional Outlier

Tahapan-tahapan dalam mendeteksi *outlier* secara tradisional adalah sebagai berikut mencari *outlier* pada data di mana dalam hal

ini pengamatan dikatakan *outlier* apabila nilai prosentase C-organik kurang dari 0.38 atau lebih dari 1.13.

3.2.2 *Weighted z Algorithm*

Berikut ini adalah tahapan mendeteksi spasial *outlier* berdasarkan metode *weighted z algorithm*:

1. Dihitung bobot (*weight*) untuk masing-masing objek spasial x_i berdasarkan persamaan 2.12.
2. Dihitung *neighborhood function* $g(x_i)$ yaitu rata-rata terboboti untuk masing-masing objek spasial x_i berdasarkan persamaan 2.13.
3. Dihitung fungsi pembanding $h(x_i)$ berdasarkan persamaan 2.14.
4. Dihitung statistik uji $Z(x_i)$ yang didapatkan hasil pembakuan fungsi pembanding $h(x_i)$ sesuai dengan persamaan 2.9.
5. Objek spasial yang diidentifikasi sebagai spasial *outlier* adalah objek spasial yang memiliki nilai atribut $Z(x_i) > 2$.

3.2.3 *Iterative r Algorithm*

Berikut ini adalah tahapan mendeteksi spasial *outlier* berdasarkan metode *iterative r algorithm*:

1. Dihitung fungsi tetangga $g(x_i)$ berdasarkan persamaan 2.15.
2. Dihitung fungsi pembanding $h(x_i)$ berdasarkan persamaan 2.16.
3. Ditentukan h_q atau h_q^{-1} dan menghitung kembali fungsi tetangga $g(x_i)$ dan fungsi pembanding $h(x_i)$.
4. Objek spasial yang diidentifikasi sebagai spasial *outlier* adalah objek yang memiliki nilai h_q atau $h_q^{-1} \geq \theta$. Nilai θ yang digunakan adalah 1.

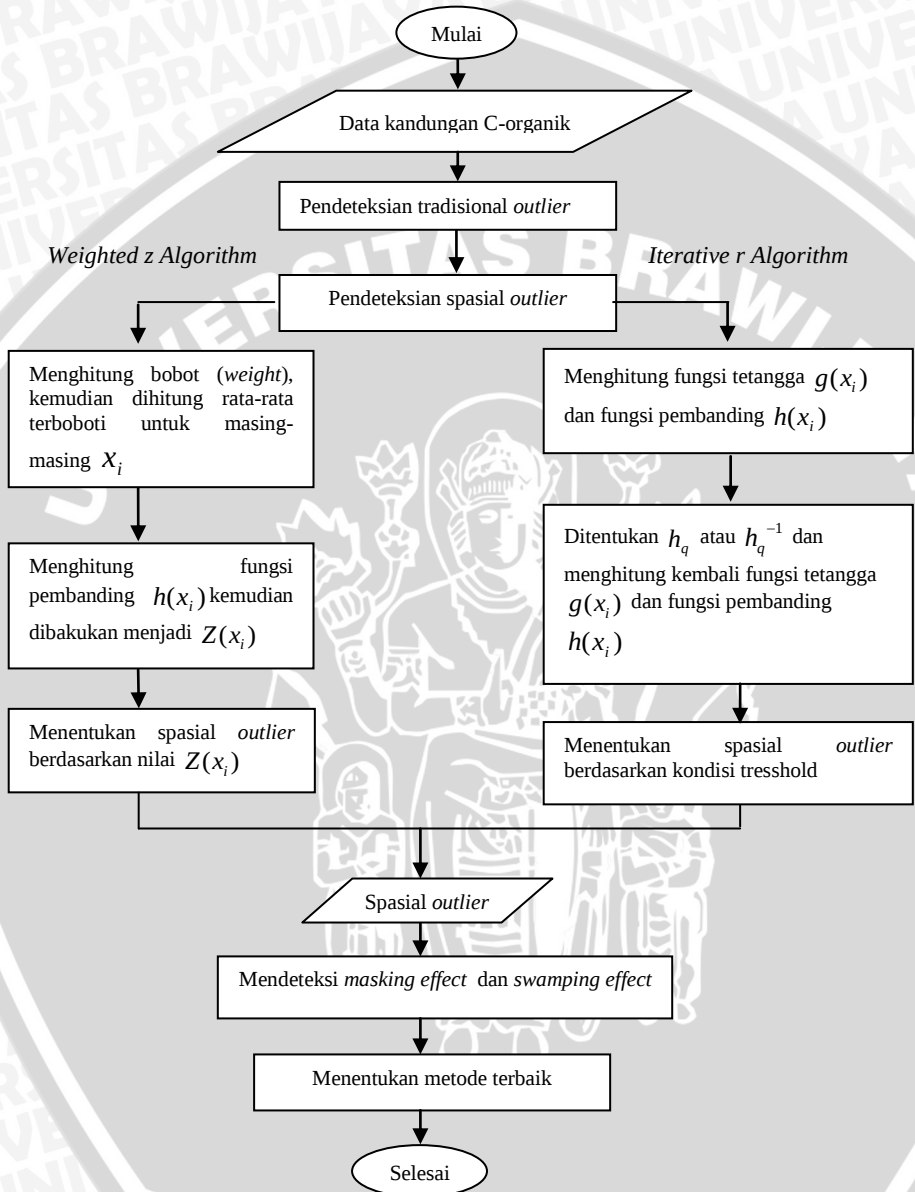
3.2.4 *Pendeteksian Masking Effect dan Swamping Effect*

Tahapan-tahapan pendeteksian gejala masking effect dan swamping effect adalah sebagai berikut:

1. Dipilih *outlier* yang menyimpang paling jauh dari *neighborhood function* (fungsi tetangga) atau memiliki nilai atribut Z terbesar dari kedua metode pendeteksian spasial *outlier*, kemudian *outlier* tersebut dihapus.
2. Dilakukan pendeteksian spasial *outlier* ulang. Jika pendeteksian spasial *outlier* yang kedua (setelah *outlier* terjauh dihapus), ada objek non-*outlier* (berdasarkan pendeteksian spasial *outlier* yang pertama) yang dideteksi sebagai *outlier*, maka dapat dikatakan terdapat *masking effect* pada data tersebut. Jika pada pendeteksian *outlier* yang kedua terdapat *outlier* (berdasarkan pada pendeteksian spasial *outlier* yang pertama yaitu sebelum *outlier* terjauh dihapus) yang tidak dideteksi sebagai *outlier*, maka dapat dikatakan bahwa pada data tersebut terdapat *swamping effect*. Jika hasil pendeteksian pada pendeteksian spasial *outlier* kedua tidak berbeda dengan pendeteksian spasial *outlier* pertama, maka pada data tersebut tidak terdapat *masking effect* dan *swamping effect*.

3.2.5 Penentuan Metode Terbaik

Pada pendeteksian spasial *outlier*, metode dikatakan metode yang terbaik apabila metode tersebut lebih sensitif dalam mendeteksi spasial *outlier*, dalam hal ini metode mampu menghasilkan pendugaan spasial *outlier* sama dengan hasil pendugaan tradisional *outlier* dan metode yang mampu mengatasi *masking effect* dan *swamping effect* sehingga dapat mendeteksi spasial *outlier* dengan benar.



Gambar 3.1. Diagram alir perbandingan metode *weighted z algorithm* dan *iterative r algorithm*.

BAB IV HASIL DAN PEMBAHASAN

4.1 Pendeteksian Spasial *Outlier* Berdasarkan Tradisional *Outlier*

Metode pendeteksian tradisional *outlier* menyatakan suatu titik sebagai spasial *outlier* apabila objek spasial yang diidentifikasi sebagai spasial *outlier* adalah objek spasial yang memiliki prosentase kurang dari 0.38 atau lebih besar dari 1.13 (Supriyadi, 2007). Hasil pendeteksian spasial *outlier* dengan metode tradisional *outlier* dapat dilihat pada Tabel 4.1.

Tabel 4.1. Titik yang merupakan spasial *outlier* berdasarkan pendeteksian *outlier* secara tradisional.

No	Kode	Kecamatan	Tradisional <i>Outlier</i>	
			Titik	C-organik (%)
1	SPG 007	Sampang	4	1.46
2	SPG 105	Ketapang	40	1.41
3	SPG 030	Robatal	9	1.38
4	SPG 082	Torjun	25	1.34
5	SPG 123	Sampang	48	0.3
6	SPG 100	Kedundung	36	0.18
7	SPG 103	Sokobanah	38	0.18
8	SPG 068	Ketapang	20	0.13
9	SPG 114	Banyu Ates	43	0.06

Tabel 4.1 menunjukkan hasil pendeteksian *outlier* secara tradisional didapatkan 9 *outlier*. Titik yang dikatakan sebagai spasial *outlier* apabila memiliki prosentase C-organik kurang dari 0.38 atau lebih besar 1.13. Seperti pada kecamatan Sampang, prosentase C-organik sebesar $1.46 > 1.13$, sehingga titik 4 dideteksi sebagai spasial *outlier*. Begitu pula pada kecamatan Banyu Ates prosentase C-organik sebesar $0.06 < 0.38$, sehingga titik 43 dideteksi sebagai spasial *outlier*, dan seterusnya. Selain itu, dari hasil pendeteksian *outlier* secara tradisional dapat diketahui bahwa Kecamatan

Sampang Desa Banyu Anyar mempunyai prosentase kandungan C-organik paling tinggi yaitu sebesar 1.46 dan Kecamatan Banyu Ates Desa Trapang mempunyai prosentase kandungan C-organik paling rendah yaitu sebesar 0.06.

4.2 Pendeteksian Spasial *Outlier* Berdasarkan *Weighted Z Algorithm*

Metode *Weighted z algorithm* menyatakan suatu titik spasial sebagai spasial *outlier* jika titik spasial (nilai atribut) tersebut memiliki nilai mutlak *outlierness factor* lebih besar dari 2 ($|y(x_q)| > 2$), dimana *outlierness factor* didapatkan dari persamaan 2.9 (Kou, dkk., 2006 : 614). Hasil pendeteksian spasial *outlier* dengan metode *Weighted z algorithm* yang secara ringkas ditunjukkan pada Tabel 4.2.

Tabel 4.2. Titik yang merupakan spasial *outlier* berdasarkan pendeteksian *Weighted z algorithm*.

No	Kode	Kecamatan	<i>Weighted z Algorithm</i>		
			Titik	C-organik (%)	$ y(x_q) $
1	SPG 105	Ketapang	40	1.41	4.181067
2	SPG 009b	Robatal	6	0.61	4.31847
3	SPG 007	Sampang	4	1.46	3.241645
4	SPG 063	Robatal	18	0.94	2.88377
5	SPG 045	Pangarengan	12	1.14	2.749567

Tabel 4.2, menunjukkan hasil pendeteksian dengan metode *Weighted z algorithm* didapatkan 5 *outlier* yaitu titik 4, 6, 12, 18 dan 40. Pada metode *Weighted z algorithm* titik yang dikatakan sebagai spasial *outlier* apabila $|y(x_q)| > 2$. Seperti pada Kecamatan Robatal didapatkan $|y(x_q)| = 2.88377 > 2$, maka titik 18 dideteksi sebagai spasial *outlier*. Pada Kecamatan Sampang didapatkan $|y(x_q)| = 3.241645 > 2$, maka titik 4 dideteksi sebagai spasial *outlier*. Nilai $|y(x_q)|$ yang paling tinggi terletak pada Kecamatan Ketapang sebesar 4.181067 sedangkan nilai $|y(x_q)|$ yang paling rendah terletak pada Kecamatan Pangarengan sebesar 2.749567. Dari Tabel 4.2 dapat

disimpulkan bahwa metode *Weighted z algorithm* mendeteksi spasial *outlier* kurang dari banyaknya *outlier* yang terdeteksi oleh tradisional *outlier*.

4.3 Pendeteksian Spasial Outlier Berdasarkan *Iterative R Algorithm*

Metode *Iterative R Algorithm* menyatakan suatu titik spasial sebagai spasial *outlier* jika titik spasial (nilai atribut) tersebut memiliki h_q atau $h_q^{-1} \geq \theta$, dimana $\theta = 1$ (Lu, dkk., 2003 : 2). Hasil pendeteksian *spatial outlier* dengan metode *Iterative r Algorithm* yang secara ringkas ditunjukkan pada Tabel 4.3.

Tabel 4.3. Titik yang merupakan spasial *outlier* berdasarkan pendeteksian *Iterative r algorithm*.

No	Kode	Kecamatan	<i>Iterative r Algorithm</i>		
			Titik	C-organik (%)	$ y(x_q) $
1	SPG 007	Sampang	4	1,46	1,459999432
2	SPG 105	Ketapang	40	1,41	1,409999431
3	SPG 030	Robatal	9	1,38	1,379999431
4	SPG 082	Torjun	25	1,34	1,33999943
5	SPG 123	Sampang	48	1,3	1,29999943
6	SPG 045	Pangarengan	12	1,14	1,139999427
7	SPG 096	Robatal	34	1,12	1,119999427
8	SPG 068	Ketapang	20	0,13	1,129999411
9	SPG 019	Kedundung	8	1,08	1,079999426
10	SPG 114	Banyu Ates	43	0,06	1,05999941

Pada Tabel 4.3 menunjukkan hasil pendeteksian dengan metode *Iterative r algorithm* didapatkan 10 *outlier* yang terdeteksi yaitu titik 4, 40, 9, 25, 48, 12, 34, 20, 8, dan 43. Pada pendeteksian dengan metode *Iterative r algorithm* titik yang dikatakan sebagai spasial *outlier* apabila $|y(x_q)| > 1$. Seperti pada Kecamatan Torjun didapatkan $|y(x_q)| = 1,33999943 > 1$, maka titik 25 dideteksi sebagai spasial *outlier*. Pada Kecamatan Sampang didapatkan $|$

$y(x_q) = 1.29999943 > 1$, maka titik 48 dideteksi sebagai spasial *outlier*. Nilai $|y(x_q)|$ yang paling tinggi terletak pada Kecamatan Sampang sebesar 1.459999432 sedangkan nilai $|y(x_q)|$ yang paling rendah terletak pada Kecamatan Banyu Ates sebesar 1.05999941. Dari Tabel 4.3 dapat disimpulkan bahwa metode *Iterative r algorithm* mendeteksi spasial *outlier* lebih banyak dari banyaknya *outlier* yang terdeteksi oleh tradisional *outlier* dan juga metode *Weighted z algorithm*.

4.4 Perbandingan Pendeteksian Tradisional *Outlier*, *Weighted z Algorithm*, dan *Iterative r Algorithm*.

Perbandingan hasil pendeteksian ketiga metode perlu dilakukan untuk mengetahui sejauh mana sensitifitas metode *Weighted z algorithm* dan *Iterative R Algorithm* dalam mendeteksi spasial *outlier*. Perbandingan ini dilakukan dengan cara melihat kesamaan banyaknya spasial *outlier* maupun lokasi titiknya dan kesamaan urutan top *outlier* (urutan derajat *outlier*).

Tabel 4.4. Perbandingan Pendeteksian Spasial *Outlier* Berdasarkan Tradisional *Outlier*, *Weighted z algorithm*, dan *Iterative R Algorithm*.

Tradisional <i>Outlier</i>		<i>Weighted z Algorithm</i>		<i>Iterative r Algorithm</i>	
Jumlah <i>outlier</i>	Titik	Jumlah <i>outlier</i>	Titik	Jumlah <i>outlier</i>	Titik
9	4; 40; 9; 25; 48; 36; 38; 20; 43	5	40; 6; 4; 18; 12	10	4; 40; 9; 25; 48; 12; 34; 20; 8; 43

Dari Tabel 4.4, dapat diketahui bahwa untuk banyak dan titik lokasi spasial *outlier* yang terdeteksi, metode *Weighted z Algorithm* kurang mampu mendeteksi *outlier* sama banyak dengan tradisional *outlier*. Hal ini dapat dilihat dari banyaknya *outlier* yang terdeteksi oleh metode *Weighted z Algorithm* sebanyak 5 sedangkan yang terdeteksi oleh tradisional *outlier* sebanyak 9 *outlier*, berbeda dengan metode *Iterative r Algorithm* yang mendeteksi *outlier* lebih banyak dengan tradisional *outlier* pada 10 titik yaitu 4, 9, 12, 20, 25, 34, 40, 43, 8 dan 48 . Dalam hal ini metode *Iterative r Algorithm* dapat dikatakan lebih sensitif daripada metode *Weighted z Algorithm*,

karena *Iterative r Algorithm* mendeteksi *outlier* lebih banyak dari tradisional *outlier* sementara *Weighted z Algorithm* lebih sedikit.

4.5 Pendeteksian *Masking Effect* dan *Swamping Effect*

Masking effect dan *swamping effect* dapat dideteksi dengan cara menghapus titik *outlier* yang memiliki derajat *outlier* paling tinggi (urutan paling tinggi). Setelah dilakukan penghapusan *outlier* yang memiliki derajat *outlier* paling tinggi, dilakukan pendeteksian *outlier* ulang atau dilakukan pendeteksian spasial *outlier* kedua. Ada tidaknya *masking effect* dan *swamping effect* dapat dilihat dari perbandingan hasil antara pendeteksi spasial *outlier* pertama (sebelum *outlier* terjauh dihapus) dengan pendeteksian spasial *outlier* kedua (setelah *outlier* terjauh dihapus). Apabila pada pendeteksian kedua muncul titik baru yang terdeteksi sebagai *outlier*, maka pada proses pendeteksian tersebut terjadi *masking effect*. Apabila pada pendeteksian kedua titik yang semula terdeteksi sebagai *outlier* tidak terdeteksi lagi sebagai *outlier* maka pada proses pendeteksian tersebut terjadi *swamping effect*. Untuk menunjukkan pendeteksian *masking effect* dan *swamping effect* berdasarkan metode pendeteksian tradisional *outlier*, *Weighted z Algorithm* dan *Iterative r Algorithm* secara berturut-turut dapat dilihat pada Tabel 4.5, Tabel 4.6 dan Tabel 4.7.

Tabel 4.5. Pendeteksian *Masking Effect* dan *Swamping Effect* Berdasarkan Tradisional *outlier*.

Pendeteksian Pertama		Titik yang dihapus	Pendeteksian Kedua		Ket*
Jumlah	Titik		Jumlah	Titik	
9	4; 40; 9; 25; 48; 36; 38; 20; 43	4	9	40; 9; 25; 48; 2; 36; 38; 20; 43	M

*Keterangan : M = terdapat *masking effect*
 S = terdapat *swamping effect*
 MS = terdapat *masking* dan *swamping effect*
 A = tidak terdapat *masking* dan *swamping effect*

Dari Tabel 4.5, dapat diketahui bahwa tidak terdapat *swamping effect* pada seluruh data. Namun, ketika titik 4 dihapus muncul *outlier* baru pada titik 2 sehingga terjadi *masking effect*. Hal

tersebut mengindikasikan bahwa pendeteksian tradisional *outlier* tidak cukup mampu dalam mengatasi adanya *masking effect*, meskipun mampu dalam mengatasi *swamping effect*.

Tabel 4.6. Pendeteksian *Masking Effect* dan *Swamping Effect* Berdasarkan Metode *Weighted z Algorithm*.

Pendeteksian Pertama		Titik yang dihapus	Pendeteksian Kedua		Ket*
Jumlah	Titik		Jumlah	Titik	
5	40; 6; 4; 18; 12	40	5	6; 4; 18; 12; 43	M

*Keterangan : M = terdapat *masking effect*
 S = terdapat *swamping effect*
 MS = terdapat *masking* dan *swamping effect*
 A = tidak terdapat *masking* dan *swamping effect*

Dari Tabel 4.6, dapat diketahui bahwa tidak terdapat *swamping effect* pada seluruh data. Namun, ketika titik 40 dihapus muncul *outlier* baru pada titik 43 sehingga terjadi *masking effect*. Hal tersebut mengindikasikan bahwa metode *Weighted z Algorithm* tidak cukup mampu dalam mengatasi adanya *masking effect*, meskipun mampu dalam mengatasi *swamping effect*.

Tabel 4.7. Pendeteksian *Masking Effect* dan *Swamping Effect* Berdasarkan Metode *Iterative r Algorithm*.

Pendeteksian Pertama		Titik yang dihapus	Pendeteksian Kedua		Ket*
Jumlah	Titik		Jumlah	Titik	
10	4; 40; 9; 25; 48; 12; 34; 20; 43; 8	4	9	40; 9; 25; 48; 12; 34; 20; 43; 8	A

*Keterangan : M = terdapat *masking effect*
 S = terdapat *swamping effect*
 MS = terdapat *masking* dan *swamping effect*
 A = tidak terdapat *masking* dan *swamping effect*

Dari Tabel 4.7, dapat diketahui bahwa tidak terdapat *masking effect* dan *swamping effect* pada seluruh data. Hal tersebut

mengindikasikan bahwa metode *Iterative r Algorithm* mampu dalam mengatasi adanya *masking effect* dan *swamping effect*.

4.6 Penentuan Metode Terbaik

Suatu metode dikatakan lebih baik dari metode lainnya bila sebanyak m outlier dengan *outlierness factor* tertinggi (m top outlier) yang terdeteksi sama dengan tradisional outlier, dari jumlah maupun lokasi dan urutannya. Pada kriteria kedua, suatu metode dikatakan lebih baik dari metode lainnya bila metode tersebut mampu mengatasi *masking effect* yang berarti bisa menghindari amatan outlier dideteksi sebagai amatan normal dan mampu mengatasi *swamping effect* yang berarti bisa menghindari amatan normal terdeteksi sebagai amatan outlier. Karakteristik kebaikan kedua metode dapat dilihat pada Tabel 4.8.

Tabel 4.8. Kriteria Kebaikan Metode Pendeteksi Spasial Outlier.

Kesamaan Dengan Tradisional Outlier						Kemampuan Mengatasi <i>Masking effect</i> dan <i>Swamping Effect</i>					
Jumlah outlier		Titik lokasi		Urutan top outlier		<i>Masking effect</i>			<i>Swamping Effect</i>		
I	W	I	W	I	W	T	I	W	T	I	W
TS	TS	TS	TS	TS	TS	TM	M	TM	M	M	M

Keterangan : I = *Iterative r Algorithm*
W = *Weighted z algorithm*
T = Tradisional outlier
S = sama dengan tradisional outlier
TS = tidak sama dengan tradisional outlier
M = mampu mengatasi
TM = tidak mampu mengatasi

Untuk kriteria kebaikan metode yang pertama apabila salah satu kriteria kesamaan dengan tradisional outlier (jumlah outlier, titik lokasi yaitu lokasi titik yang dideteksi sebagai outlier, dan urutan top outlier) pada data tertentu menyatakan tidak sama dengan metode yang diteliti, maka metode tersebut dinyatakan berbeda dengan tradisional outlier. Berdasarkan pada Tabel 4.8 di atas, pada metode *Weighted z algorithm* dan *Iterative r Algorithm* dari jumlah outlier, titik lokasi dan urutan top outlier tidak sama dengan tradisional outlier, dimana outlier yang dideteksi oleh metode

Weighted z algorithm lebih sedikit daripada tradisional *outlier* sedangkan pada metode *Iterative r Algorithm outlier* yang terdeteksi lebih banyak daripada tradisional *outlier*.

Pada kriteria kebaikan metode yang kedua dijelaskan bahwa suatu metode dikatakan lebih baik dan lebih sensitif apabila memiliki kemampuan dalam mengatasi adanya *masking effect* dan *swamping effect*. Berdasarkan hasil yang ditunjukkan pada Tabel 4.8, didapatkan bahwa pendeteksian tradisional *outlier* dan metode *Weighted z algorithm* tidak cukup mampu mengatasi *masking effect* sedangkan metode *Iterative r Algorithm* memiliki kemampuan yang lebih baik dalam mengatasi *masking effect* dan *swamping effect*. Dapat disimpulkan bahwa metode *Iterative r Algorithm* lebih sensitif dan lebih baik daripada pendeteksian tradisional *outlier* dan *Weighted z algorithm* dalam mendeteksi *outlier* pada data spasial.



BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil dan pembahasan dapat ditarik beberapa kesimpulan yaitu :

1. Spasial *outlier* yang dideteksi *Weighted z Algorithm* lebih sedikit daripada *Iterative r Algorithm*.
2. Pada Kabupaten Sampang prosentase kandungan C-organik yang paling tinggi sebesar 1.46 terletak di Kecamatan Sampang Desa Banyu Anyar sedangkan prosentase kandungan C-organik paling rendah sebesar 0.06 terletak di Kecamatan Banyu Ates Desa Trapang.
3. Pendeteksian tradisional *outlier* mendeteksi sebanyak 9 *outlier* dan tidak memiliki kemampuan dalam mengatasi adanya *masking effect*. Metode *weighted z algorithm* lebih sedikit dibandingkan dengan tradisional *outlier* dalam mendeteksi spasial *outlier* dan tidak cukup memiliki kemampuan dalam mengatasi *masking effect*. *Iterative r Algorithm* lebih banyak dibandingkan dengan tradisional *outlier* dalam mendeteksi spasial *outlier* dan memiliki kemampuan yang lebih baik dalam mengatasi *masking effect*.
4. Metode *Iterative r Algorithm* lebih sensitif dan lebih baik daripada *weighted z algorithm* dalam mendeteksi spasial *outlier* pada data spasial.

5.2 Saran

Dari penelitian yang telah dilakukan, terdapat saran yang dapat digunakan untuk penyempurnaan hasil penelitian ini, yaitu : untuk penelitian selanjutnya selain metode *Iterative r Algorithm* dan *Weighted z Algorithm* masih ada metode lain untuk mendeteksi spasial *outlier* seperti *average difference algorithm*, *Scatterplot* dan *Moran Scatterplot* untuk mengetahui metode pendeteksian spasial *outlier* yang lebih baik.

DAFTAR PUSTAKA

- Bailey, T. C. dan Gatrell, A. C. 1995. *Interactive Spatial Data Analysis*. Prentice Hall.
- Barnett, V. dan Lewis, T. 1994. *Outliers in Statistical Data*, 3rd edition. John Wiley. New York
- Davies L. dan Gather U. 1993. *The Identification of Multiple Outliers*. Journal of The American Statistical Association. 88(423), 782-792. Di akses pada tanggal 23 Januari 2013.
- Gal-Ben, I. 2005. *Outlier Detection*. Tel- Aviv University. Israel. Hal 1-16
- GIS Konsorsium Aceh Nias. 2007. *Modul Pelatihan ArcGIS Tingkat Dasar*. GTZ-SLGSR. Banda Aceh.
- Goldberg, L. R dan Kercheval, A. N. 2002. *T-Statistics for Weighted Means With Application to Risk Factor Models*. Hal 1-3.
- Haining, R.. 1990. *Spatial Data Analysis in The Social and Environmental Sciences*. Cambridge University Press. Cambridge.
- Handriyani, D., Bijaksana, M. A dan Setiawan, E. B. 2009. *Analisis Perbandingan Clustering-Based dan Density-Based dalam Mendeteksi Outlier*. IT Telkom. Bandung.

- Han, J., Kamber, M dan Tung. A. 2001. *Spatial Clustering Methods In Data Mining : A Survey*. In H. Miller and J. Han, editors Geographic Data Mining and Knowledge Discovery. Taylor and Francis.
- Hartoyo, E, Nugroho, Y, Bhirowo, A, dan Khalil, B. 2010. *Modul Pelatihan Sistem Informasi Geografis (GIS) Tingkat Dasar*. Tropenbos International Indonesia Programme (TBI Indonesia. Balikpapan
- Jiangsheng, Y. 2002. *Method of k-nearest Neighbors*. Institute of Computational Linguistics. China.
- Johnson, R. A. dan Winchern, D. W. 2002. *Applied Multivariate Statistical: Analysis, Fifth Edition*. Prentice Hall. New Jersey.
- Kou, Y., Lu, C. T., dan Chen, D. 2006. *Spatial Weighted Outlier Detection*. IEEE Computer Society.
- Lestari, T. 2010. *Metode Iterative z Algorithm dan Weighted z Algorithm dalam Mendeteksi Outlier Pada Data Spasial*. Skripsi Program Studi Statistika Jurusan Matematika Fakultas MIPA Universitas Brawijaya Malang (Tidak dipublikasikan).
- Lin, J dan Ye. D. Y. 2008. *Minimum Spanning Tree Based Spatial Outlier Mining and Its Applications*. Fuzhou University. China.
- Lu, C. T, Chen, D dan Kou, Y. 2003. *Algorithms for Spatial Outlier Detection*. In Proc. Of The 3rd IEEE International Conference on Data Mining. Melbourne.
- Prahasta, E. . 2001. *Konsep-Konsep Dasar Sistem Informasi Geografis*. Press Inc. Bandung

- Ru, Li X., Ming, Li X., Ling, Li H., dan Yong Cao Mao. 2009. *Rejecting Outlier Based On Correspondence Manifold*. Volume 35. Hal 17.
- Saidi, A. 2004. *Kajian Potensi Kesuburan Tanah Pada Lahan Sentra Pertanian Hortikultura Di Sumatera Barat*. Fakultas Pertanian Universitas Andalas Padang. Stigma Vol. XXI No. 2. ISSN 0853-3776. Hal 134-139.
- Setyani, D. 2011. *Penggunaan Metode Iterpolasi Spasial Inverse Distance Weighted (IDW) dan Thin-Plate Spline Pada Data Spasial*. Skripsi Jurusan Statistika Fakultas MIPA Universitas Brawijaya. Universitas Brawijaya. Malang
- Soemartini. 2007. *Pencilan (Outlier)*. Skripsi Jurusan Statistika Fakultas MIPA Universitas Padjadjaran. Universitas Padjadjaran. Jatinangor. Diakses pada tanggal 23 Januari 2013.
- Shekhar, S., Lu, C. T, dan Zhang, P. 2002. *Detecting Graph-Based Spatial Outliers*. University of Minnesota. Minneapolis.
- Shekhar, S., Lu, C. T., dan Zhang, P. 2003. *A Unified Approach to Detecting Spatial Outliers*. University of Minnesota. Minneapolis. Hal 139-140.
- Schwering, A dan Raubal, M. 2005. *Spatial Relations for Semantic Similarity Measurement*. Ordnance Survey of Great Britain. United Kingdom.
- Supriyadi, S. 2007. *Kesuburan Tanah Di Lahan Kering Madura*. Jurusan Budidaya Pertanian Fakultas Pertanian Universitas Trunojoyo. Embryo Vol.4 No. 2. ISSN 0216-0188. Hal 124-131.

Suryani, A. 2007. *Perbaikan Tanah Media Tanaman Jeruk Dengan Berbagai Bahan Organik Dalam Bentuk Kompos*. Skripsi Sekolah Pasca Sarjana. Institut Pertanian Bogor. Bogor. Diakses pada tanggal 19 Agustus 2013.

Upton, G. J. dan Fingleton, B. 1985. *Spatial Data Analysis by Example Volume 1: Point Pattern and Quantitative Data*. John Wiley & Sons. New York.

Wikipedia, Inc. 2013. Taxicab Geometry. http://en.wikipedia.org/wiki/Taxicab_geometry. Di akses pada tanggal 10 Februari 2013

Wikipedia, Inc. 2013. Distance Transform. http://en.wikipedia.org/wiki/Distance_transform. Di akses pada tanggal 10 Februari 2013.

Wikipedia, Inc. 2013. Euclidean Distance. http://en.wikipedia.org/wiki/Euclidean_distance. Di akses pada tanggal 10 Februari 2013.

Lampiran 1. Data kandungan C-organik tanah di Kabupaten Sampang Madura

No	Kode	Kecamatan	Desa	Koordinat X (m)	Koordinat Y (m)	C-Org (%)
1	SPG 002	Jrengik	Jrengik	735477	9212097	0,36
2	SPG 003a	Sampang	Polagan	746985	9202635	0,25
3	SPG 003b	Sampang	Pasean	749316	9205052	0,55
4	SPG 007	Sampang	Banyu Anyar	749908	9202428	1,46
5	SPG 009a	Sampang	Tanggumong	747246	9206219	0,62
6	SPG 009b	Robatal	Sawah Tengah	751529	9221540	0,61
7	SPG012	Tembalangan	Birem	740151	9226973	1
8	SPG 019	Kedundung	Ombul	743950	9219329	1,08
9	SPG 030	Robatal	Jlegung	752984	9226817	1,38
10	SPG 031	Robatal	Pandiyangan	750761	9227490	0,44
11	SPG 037	Kedundung	Ombul	744023	9218732	0,8
12	SPG 045	Pangarengan	Gullbung	742146	9201876	1,14
13	SPG 046	Sampang	Aeng Sareh	745003	9202700	0,71
14	SPG 047	Robatal	Sawah Tengah	750832	9222562	0,43
15	SPG 053	Kedundung	Kedundung	747331	9219823	0,81
16	SPG 059	Robatal	Sawah Tengah	752200	9222200	0,3
17	SPG 061	Robatal	Gunung Rancak	753678	9229296	1,07
18	SPG 063	Robatal	Jlegung	751604	9225900	0,94
19	SPG 065	Tembalangan	Birem	740292	9226698	0,73
20	SPG 068	Ketapang	Bunten Barat	751256	9232812	0,13
21	SPG 069	Banyu Ates	Tapa'an	737942	9232172	0,49
22	SPG 077	Banyu Ates	Planggaran Timur	740319	9229966	0,24
23	SPG 079	Ketapang	Bunten Barat	752155	9230636	0,76
24	SPG 081	Tembalangan	Birem	740079	9224790	0,54
25	SPG 082	Torjun	Tanah Merah	742524	9213505	1,34
26	SPG 083	Sokobanah	Bira Timur	764284	9233452	0,62
27	SPG 087	Robatal	Pandiyangan	752525	9230250	0,51
28	SPG 089	Tembalangan	Tembalangan	740071	9220330	0,54

Lampiran 1. (lanjutan)

No	Kode	Kecamatan	Desa	Koordinat X (m)	Koordinat Y (m)	C-Org (%)
29	SPG 090	Banyu Ates	Planggaran Timur	740659	9228951	0,43
30	SPG 091	Robatal	Tragih	751141	9225198	0,49
31	SPG 092	Kedundung	Banjar	741199	9217846	0,49
32	SPG 093	Tembalangan	Baturasang	734251	9219135	0,85
33	SPG 095	Sampang	Gunung Maddah	750504	9205188	0,37
34	SPG 096	Robatal	Jlegung	753216	9227971	1,12
35	SPG 099	Tembalangan	Karang Anyar	738280	9215879	0,6
36	SPG 100	Kedundung	Rahayu	749379	9214777	0,18
37	SPG 102	Sokobanah	Bira Tengah	764151	9235620	0,48
38	SPG 103	Sokobanah	Tobai Tengah	762912	9229845	0,18
39	SPG 104	Banyu Ates	Morbato	738357	9234093	0,56
40	SPG 105	Ketapang	Ketapang Timur	754155	9237354	1,41
41	SPG 110	Omben	Kamondung	760288	9212724	0,74
42	SPG 111	Omben	Tambak	760049	9214341	0,8
43	SPG 114	Banyu Ates	Trapang	735943	9237138	0,06
44	SPG 115	Banyu Ates	Masaran	739890	9237119	0,6
45	SPG 116	Pangarengan	Apaan	741088	9202610	0,71
46	SPG 117	Torjun	Ragung	738082	9206776	0,61
47	SPG 122	Banyu Ates	Morbato	738238	9235380	0,84
48	SPG 123	Sampang	Banyu Anyar	749823	9202948	1,3
49	SPG 126	Sokobanah	Temberu Barat	773761	9236855	0,55
50	SPG 128	Sampang	Aeng Sareh	745537	9202481	0,82
51	SPG 131	Sampang	Taman Sareh	750845	9209566	0,49
52	SPG 135	Torjun	Dulang	740972	9205729	0,7
53	SPG 138	Sreseh	Plasah	728518	9205617	0,38
54	SPG 139	Sreseh	Sreseh	734917	9201486	0,59
55	SPG 141	Sreseh	Bundeh	726670	9205945	0,69
56	SPG 142	Pangarengan	Pangarengan	739320	9204652	0,57

Lampiran 2. Pendeteksian Spasial *Outlier* Berdasarkan Tradisional *Outlier* (Pendeteksian Pertama).

No	Kode	Kecamatan	Desa	Koordinat X (m)	Koordinat Y (m)	C-Org (%)
1	SPG 002	Jrengik	Jrengik	735477	9212097	0,36
2	SPG 003a	Sampang	Polagan	746985	9202635	0,25
3	SPG 003b	Sampang	Pasean	749316	9205052	0,55
4	SPG 007	Sampang	Banyu Anyar	749908	9202428	1,46
5	SPG 009a	Sampang	Tanggumong	747246	9206219	0,62
6	SPG 009b	Robatal	Sawah Tengah	751529	9221540	0,61
7	SPG012	Tembalangan	Birem	740151	9226973	1
8	SPG 019	Kedundung	Ombul	743950	9219329	1,08
9	SPG 030	Robatal	Jlegung	752984	9226817	1,38
10	SPG 031	Robatal	Pandiyangan	750761	9227490	0,44
11	SPG 037	Kedundung	Ombul	744023	9218732	0,8
12	SPG 045	Pangarengan	Gullbung	742146	9201876	1,14
13	SPG 046	Sampang	Aeng Sareh	745003	9202700	0,71
14	SPG 047	Robatal	Sawah Tengah	750832	9222562	0,43
15	SPG 053	Kedundung	Kedundung	747331	9219823	0,81
16	SPG 059	Robatal	Sawah Tengah	752200	9222200	0,3
17	SPG 061	Robatal	Gunung Rancak	753678	9229296	1,07
18	SPG 063	Robatal	Jlegung	751604	9225900	0,94
19	SPG 065	Tembalangan	Birem	740292	9226698	0,73
20	SPG 068	Ketapang	Bunten Barat	751256	9232812	0,13
21	SPG 069	Banyu Ates	Tapa'an	737942	9232172	0,49
22	SPG 077	Banyu Ates	Planggaran Timur	740319	9229966	0,24
23	SPG 079	Ketapang	Bunten Barat	752155	9230636	0,76
24	SPG 081	Tembalangan	Birem	740079	9224790	0,54
25	SPG 082	Torjun	Tanah Merah	742524	9213505	1,34
26	SPG 083	Sokobanah	Bira Timur	764284	9233452	0,62
27	SPG 087	Robatal	Pandiyangan	752525	9230250	0,51
28	SPG 089	Tembalangan	Tembalangan	740071	9220330	0,54

Lampiran 2. (lanjutan)

No	Kode	Kecamatan	Desa	Koordinat X (m)	Koordinat Y (m)	C-Org (%)
29	SPG 090	Banyu Ates	Planggaran Timur	740659	9228951	0,43
30	SPG 091	Robatal	Tragih	751141	9225198	0,49
31	SPG 092	Kedundung	Banjar	741199	9217846	0,49
32	SPG 093	Tembalangan	Baturasang	734251	9219135	0,85
33	SPG 095	Sampang	Gunung Maddah	750504	9205188	0,37
34	SPG 096	Robatal	Jlegung	753216	9227971	1,12
35	SPG 099	Tembalangan	Karang Anyar	738280	9215879	0,6
36	SPG 100	Kedundung	Rahayu	749379	9214777	0,18
37	SPG 102	Sokobanah	Bira Tengah	764151	9235620	0,48
38	SPG 103	Sokobanah	Tobai Tengah	762912	9229845	0,18
39	SPG 104	Banyu Ates	Morbato	738357	9234093	0,56
40	SPG 105	Ketapang	Ketapang Timur	754155	9237354	1,41
41	SPG 110	Omben	Kamondung	760288	9212724	0,74
42	SPG 111	Omben	Tambak	760049	9214341	0,8
43	SPG 114	Banyu Ates	Trapang	735943	9237138	0,06
44	SPG 115	Banyu Ates	Masaran	739890	9237119	0,6
45	SPG 116	Pangarengan	Apaan	741088	9202610	0,71
46	SPG 117	Torjun	Ragung	738082	9206776	0,61
47	SPG 122	Banyu Ates	Morbato	738238	9235380	0,84
48	SPG 123	Sampang	Banyu Anyar	749823	9202948	1,3
49	SPG 126	Sokobanah	Temberu Barat	773761	9236855	0,55
50	SPG 128	Sampang	Aeng Sareh	745537	9202481	0,82
51	SPG 131	Sampang	Taman Sareh	750845	9209566	0,49
52	SPG 135	Torjun	Dulang	740972	9205729	0,7
53	SPG 138	Sreseh	Plasah	728518	9205617	0,38
54	SPG 139	Sreseh	Sreseh	734917	9201486	0,59
55	SPG 141	Sreseh	Bundeh	726670	9205945	0,69
56	SPG 142	Pangarengan	Pangarengan	739320	9204652	0,57

Lampiran 2. (lanjutan)

Pendeteksian Kedua

No	Kode	Kecamatan	Desa	Koordinat X (m)	Koordinat Y (m)	C-Org (%)
1	SPG 002	Jrengik	Jrengik	735477	9212097	0,36
2	SPG 003a	Sampang	Polagan	746985	9202635	0,25
3	SPG 003b	Sampang	Pasean	749316	9205052	0,55
4	SPG 007	Sampang	Banyu Anyar	749908	9202428	Dihapus
5	SPG 009a	Sampang	Tanggumong	747246	9206219	0,62
6	SPG 009b	Robatal	Sawah Tengah	751529	9221540	0,61
7	SPG012	Tembalangan	Birem	740151	9226973	1
8	SPG 019	Kedundung	Ombul	743950	9219329	1,08
9	SPG 030	Robatal	Jlegung	752984	9226817	1,38
10	SPG 031	Robatal	Pandiyangan	750761	9227490	0,44
11	SPG 037	Kedundung	Ombul	744023	9218732	0,8
12	SPG 045	Pangarengan	Gullbung	742146	9201876	1,14
13	SPG 046	Sampang	Aeng Sareh	745003	9202700	0,71
14	SPG 047	Robatal	Sawah Tengah	750832	9222562	0,43
15	SPG 053	Kedundung	Kedundung	747331	9219823	0,81
16	SPG 059	Robatal	Sawah Tengah	752200	9222200	0,3
17	SPG 061	Robatal	Gunung Rancak	753678	9229296	1,07
18	SPG 063	Robatal	Jlegung	751604	9225900	0,94
19	SPG 065	Tembalangan	Birem	740292	9226698	0,73
20	SPG 068	Ketapang	Bunten Barat	751256	9232812	0,13
21	SPG 069	Banyu Ates	Tapa'an	737942	9232172	0,49
22	SPG 077	Banyu Ates	Planggaran Timur	740319	9229966	0,24
23	SPG 079	Ketapang	Bunten Barat	752155	9230636	0,76
24	SPG 081	Tembalangan	Birem	740079	9224790	0,54
25	SPG 082	Torjun	Tanah Merah	742524	9213505	1,34
26	SPG 083	Sokobanah	Bira Timur	764284	9233452	0,62
27	SPG 087	Robatal	Pandiyangan	752525	9230250	0,51
28	SPG 089	Tembalangan	Tembalangan	740071	9220330	0,54

Lampiran 2. (lanjutan)

No	Kode	Kecamatan	Desa	Koordinat X (m)	Koordinat Y (m)	C-Org (%)
29	SPG 090	Banyu Ates	Planggaran Timur	740659	9228951	0,43
30	SPG 091	Robatal	Tragih	751141	9225198	0,49
31	SPG 092	Kedundung	Banjar	741199	9217846	0,49
32	SPG 093	Tembalangan	Baturasang	734251	9219135	0,85
33	SPG 095	Sampang	Gunung Maddah	750504	9205188	0,37
34	SPG 096	Robatal	Jlegung	753216	9227971	1,12
35	SPG 099	Tembalangan	Karang Anyar	738280	9215879	0,6
36	SPG 100	Kedundung	Rahayu	749379	9214777	0,18
37	SPG 102	Sokobanah	Bira Tengah	764151	9235620	0,48
38	SPG 103	Sokobanah	Tobai Tengah	762912	9229845	0,18
39	SPG 104	Banyu Ates	Morbato	738357	9234093	0,56
40	SPG 105	Ketapang	Ketapang Timur	754155	9237354	1,41
41	SPG 110	Omben	Kamondung	760288	9212724	0,74
42	SPG 111	Omben	Tambak	760049	9214341	0,8
43	SPG 114	Banyu Ates	Trapang	735943	9237138	0,06
44	SPG 115	Banyu Ates	Masaran	739890	9237119	0,6
45	SPG 116	Pangarengan	Apaan	741088	9202610	0,71
46	SPG 117	Torjun	Ragung	738082	9206776	0,61
47	SPG 122	Banyu Ates	Morbato	738238	9235380	0,84
48	SPG 123	Sampang	Banyu Anyar	749823	9202948	1,3
49	SPG 126	Sokobanah	Temberu Barat	773761	9236855	0,55
50	SPG 128	Sampang	Aeng Sareh	745537	9202481	0,82
51	SPG 131	Sampang	Taman Sareh	750845	9209566	0,49
52	SPG 135	Torjun	Dulang	740972	9205729	0,7
53	SPG 138	Sreseh	Plasah	728518	9205617	0,38
54	SPG 139	Sreseh	Sreseh	734917	9201486	0,59
55	SPG 141	Sreseh	Bundeh	726670	9205945	0,69
56	SPG 142	Pangarengan	Pangarengan	739320	9204652	0,57

Lampiran 3. Pendeteksian Spasial *Outlier* Berdasarkan Metode *Weighted z Algorithm*.

Objek	Kecamatan	Koordinat X (m)	Koordinat Y (m)	C-Org (%)	Pendeteksian pertama		
					M_w	$h(x_i)$	$z(x_i)$
1	Jrengik	735477	9212097	0,36	0,98	-0,62	-1,7238715
2	Sampang	746985	9202635	0,25	0,87	-0,62	-1,7238715
3	Sampang	749316	9205052	0,55	1,89	-1,34	-1,9447472
4	Sampang	749908	9202428	1,46	0,97	0,49	3,24164528
5	Sampang	747246	9206219	0,62	0,56	0,06	1,31806671
6	Robatal	751529	9221540	0,61	1,81	-1,2	-4,3184658
7	Tembalangan	740151	9226973	1	1,87	-0,87	-0,8422311
8	Kedundung	743950	9219329	1,08	0,78	0,3	1,39169196
9	Robatal	752984	9226817	1,38	1,23	0,15	1,72067618
10	Robatal	750761	9227490	0,44	0,86	-0,42	-0,8291838
11	Kedundung	744023	9218732	0,8	0,77	0,03	1,18386355
12	Pangarengan	742146	9201876	1,14	0,76	0,38	2,74956704
13	Sampang	745003	9202700	0,71	1,45	-0,74	-1,2606841
14	Robatal	750832	9222562	0,43	0,65	-0,22	0,06550392
15	Kedundung	747331	9219823	0,81	1,12	-0,31	-0,3371055
16	Robatal	752200	9222200	0,3	0,45	-0,15	0,37864462
17	Robatal	753678	9229296	1,07	1,56	-0,49	-1,1423245
18	Robatal	751604	9225900	0,94	0,53	0,41	2,88377019
19	Tembalangan	740292	9226698	0,73	0,61	0,12	1,58647302
20	Ketapang	751256	9232812	0,13	0,84	-0,71	-1,126481
21	Banyu Ates	737942	9232172	0,49	1,3	-0,81	-1,5738248
22	Banyu Ates	740319	9229966	0,24	0,55	-0,31	-0,3371055
23	Ketapang	752155	9230636	0,76	0,82	-0,06	0,78125409
24	Tembalangan	740079	9224790	0,54	0,49	0,05	1,27333232
25	Torjun	742524	9213505	1,34	0,7	0,64	0,91266106
26	Sokobanah	764284	9233452	0,62	0,8	-0,18	0,24444146
27	Robatal	752525	9230250	0,51	1,14	-0,63	-1,7686059
28	Tembalangan	740071	9220330	0,54	0,71	-0,17	0,28917585

Lampiran 3. (lanjutan)

Objek	Kecamatan	Koordinat X (m)	Koordinat Y (m)	C-Org (%)	Pendeteksian pertama		
					M_w	$h(x_i)$	$z(x_i)$
29	Banyu Ates	740659	9228951	0,43	0,87	-0,44	-0,9186526
30	Robatal	751141	9225198	0,49	0,81	-0,32	-0,3818399
31	Kedundung	741199	9217846	0,49	0,3	0,19	1,89961372
32	Tembalangan	734251	9219135	0,85	1,07	-0,22	0,06550392
33	Sampang	750504	9205188	0,37	0,94	-0,57	-1,5001996
34	Robatal	753216	9227971	1,12	0,56	0,56	1,55478597
35	Tembalangan	738280	9215879	0,6	1,41	-0,81	-1,5738248
36	Kedundung	749379	9214777	0,18	0,74	-0,56	-1,4554652
37	Sokobanah	764151	9235620	0,48	0,8	-0,32	-0,3818399
38	Sokobanah	762912	9229845	0,18	0,06	0,12	1,58647302
39	Banyu Ates	738357	9234093	0,56	0,6	-0,04	0,87072286
40	Ketapang	754155	9237354	1,41	0,71	0,7	4,18106737
41	Omben	760288	9212724	0,74	0,61	0,13	1,63120741
42	Omben	760049	9214341	0,8	0,84	-0,04	0,87072286
43	Banyu Ates	735943	9237138	0,06	0,53	-0,47	-1,0528557
44	Banyu Ates	739890	9237119	0,6	0,61	-0,01	1,00492601
45	Pangarengan	741088	9202610	0,71	0,84	-0,13	0,46811339
46	Torjun	738082	9206776	0,61	1,3	-0,69	-1,0370122
47	Banyu Ates	738238	9235380	0,84	0,55	0,29	1,34695757
48	Sampang	749823	9202948	1,3	0,82	0,48	1,19691089
49	Sokobanah	773761	9236855	0,55	0,49	0,06	1,31806671
50	Sampang	745537	9202481	0,82	1,87	-1,05	-1,6474501
51	Sampang	750845	9209566	0,49	0,82	-0,33	-0,4265743
52	Torjun	740972	9205729	0,7	1,24	-0,54	-1,3659964
53	Sreseh	728518	9205617	0,38	0,55	-0,17	0,28917585
54	Sreseh	734917	9201486	0,59	1,13	-0,54	-1,3659964
55	Sreseh	726670	9205945	0,69	0,8	-0,11	0,55758216
56	Pangarengan	739320	9204652	0,57	1,66	-1,09	-1,8263876

Lampiran 3. (lanjutan)

Objek	Kecamatan	Koordinat X (m)	Koordinat Y (m)	C-Org (%)	Pendeteksian kedua		
					M_w	$h(x_i)$	$z(x_i)$
1	Jrengik	735477	9212097	0,36	1,12	-0,76	-1,426872
2	Sampang	746985	9202635	0,25	0,83	-0,58	-1,931212
3	Sampang	749316	9205052	0,55	1,74	-1,19	-0,555739
4	Sampang	749908	9202428	1,46	0,95	0,51	3,6165297
5	Sampang	747246	9206219	0,62	0,62	0	-0,234795
6	Robatal	751529	9221540	0,61	1,82	-1,21	-4,280644
7	Tembalangan	740151	9226973	1	1,8	-0,8	1,5074707
8	Kedundung	743950	9219329	1,08	0,8	0,28	1,8742636
9	Robatal	752984	9226817	1,38	1,2	0,18	1,6497368
10	Robatal	750761	9227490	0,44	0,83	-0,39	-1,060079
11	Kedundung	744023	9218732	0,8	0,72	0,08	0,5904886
12	Pangarengan	742146	9201876	1,14	0,69	0,45	2,1493582
13	Sampang	745003	9202700	0,71	1,39	-0,68	0,1778466
14	Robatal	750832	9222562	0,43	0,65	-0,22	-1,105928
15	Kedundung	747331	9219823	0,81	1,1	-0,29	0,6363377
16	Robatal	752200	9222200	0,3	0,5	-0,2	-1,701967
17	Robatal	753678	9229296	1,07	1,66	-0,59	1,8284145
18	Robatal	751604	9225900	0,94	0,57	0,37	2,2323761
19	Tembalangan	740292	9226698	0,73	0,63	0,1	0,2695448
20	Ketapang	751256	9232812	0,13	0,9	-0,77	-1,481402
21	Banyu Ates	737942	9232172	0,49	1,24	-0,75	-0,830834
22	Banyu Ates	740319	9229966	0,24	0,58	-0,34	-1,977061
23	Ketapang	752155	9230636	0,76	0,82	-0,06	0,4070921
24	Tembalangan	740079	9224790	0,54	0,48	0,06	-0,601588
25	Torjun	742524	9213505	1,34	0,73	0,61	0,9663404
26	Sokobanah	764284	9233452	0,62	0,88	-0,26	-0,234795
27	Robatal	752525	9230250	0,51	1,15	-0,64	-0,739136
28	Tembalangan	740071	9220330	0,54	0,8	-0,26	-0,601588

Lampiran 3. (lanjutan)

Objek	Kecamatan	Koordinat X (m)	Koordinat Y (m)	C-Org (%)	Pendeteksian kedua		
					M_w	$h(x_i)$	$z(x_i)$
29	Banyu Ates	740659	9228951	0,43	0,78	-0,35	-1,105928
30	Robatal	751141	9225198	0,49	0,83	-0,34	-0,830834
31	Kedundung	741199	9217846	0,49	0,4	0,09	-0,830834
32	Tembalangan	734251	9219135	0,85	1,1	-0,25	0,8197341
33	Sampang	750504	9205188	0,37	0,98	-0,61	-1,381023
34	Robatal	753216	9227971	1,12	0,59	0,53	1,55766
35	Tembalangan	738280	9215879	0,6	1,46	-0,86	-0,326494
36	Kedundung	749379	9214777	0,18	0,77	-0,59	-1,252156
37	Sokobanah	764151	9235620	0,48	0,87	-0,39	-0,876683
38	Sokobanah	762912	9229845	0,18	0,09	0,09	-2,252156
39	Banyu Ates	738357	9234093	0,56	0,76	-0,2	-0,50989
40	Ketapang	754155	9237354	1,41	Dihapus		
41	Omben	760288	9212724	0,74	0,65	0,09	0,3153939
42	Omben	760049	9214341	0,8	0,88	-0,08	0,5904886
43	Banyu Ates	735943	9237138	0,06	0,58	-0,52	-2,802345
44	Banyu Ates	739890	9237119	0,6	0,64	-0,04	-0,326494
45	Pangarengan	741088	9202610	0,71	0,88	-0,17	0,1778466
46	Torjun	738082	9206776	0,61	1,43	-0,82	-0,280644
47	Banyu Ates	738238	9235380	0,84	0,67	0,17	0,773885
48	Sampang	749823	9202948	1,3	0,77	0,53	1,182944
49	Sokobanah	773761	9236855	0,55	0,45	0,1	-0,555739
50	Sampang	745537	9202481	0,82	1,82	-1	0,6821868
51	Sampang	750845	9209566	0,49	0,86	-0,37	-0,830834
52	Torjun	740972	9205729	0,7	1,3	-0,6	0,1319975
53	Sreseh	728518	9205617	0,38	0,6	-0,22	-1,335174
54	Sreseh	734917	9201486	0,59	1,2	-0,61	-0,372343
55	Sreseh	726670	9205945	0,69	0,92	-0,23	0,0861484
56	Pangarengan	739320	9204652	0,57	1,78	-1,21	-0,464041

Lampiran 4. Pendeteksian Spasial *Outlier* Berdasarkan Metode *Iterative r Algorithm*.

Objek	Kecamatan	Koordinat X (m)	Koordinat Y (m)	C-Org (%)	Pendeteksian pertama	
					Iterasi 1	Iterasi 2
1	Jrengik	735477	9212097	0,36	0,359999415	0,349999275
2	Sampang	746985	9202635	0,25	0,249999413	0,249999501
3	Sampang	749316	9205052	0,55	0,549999418	0,539999768
4	Sampang	749908	9202428	1,46	1,459999432	1,449999426
5	Sampang	747246	9206219	0,62	0,619999419	0,709999824
6	Robatal	751529	9221540	0,61	0,609999419	0,629999425
7	Tembalangan	740151	9226973	1	0,999999425	0,979999514
8	Kedundung	743950	9219329	1,08	1,079999426	1,199998144
9	Robatal	752984	9226817	1,38	1,379999431	1,279999681
10	Robatal	750761	9227490	0,44	0,439999416	0,449999211
11	Kedundung	744023	9218732	0,8	0,799999422	0,809999237
12	Pangarengan	742146	9201876	1,14	1,139999427	1,099999616
13	Sampang	745003	9202700	0,71	0,70999942	0,739999118
14	Robatal	750832	9222562	0,43	0,429999416	0,449999814
15	Kedundung	747331	9219823	0,81	0,809999422	0,819999625
16	Robatal	752200	9222200	0,3	0,299999414	0,299999726
17	Robatal	753678	9229296	1,07	0,699994261	0,499991928
18	Robatal	751604	9225900	0,94	0,939999424	0,919999154
19	Tembalangan	740292	9226698	0,73	0,729999421	0,749999881
20	Ketapang	751256	9232812	0,13	1,129999411	1,199999423
21	Banyu Ates	737942	9232172	0,49	0,489999417	0,509999832
22	Banyu Ates	740319	9229966	0,24	0,239999413	0,249999137
23	Ketapang	752155	9230636	0,76	0,759999421	0,779999892
24	Tembalangan	740079	9224790	0,54	0,539999418	0,559999187
25	Torjun	742524	9213505	1,34	1,33999943	1,43999987
26	Sokobanah	764284	9233452	0,62	0,619999419	0,639999116
27	Robatal	752525	9230250	0,51	0,509999417	0,549999872
28	Tembalangan	740071	9220330	0,54	0,539999418	0,569999224

Lampiran 4. (lanjutan)

Objek	Kecamatan	Koordinat X (m)	Koordinat Y (m)	C-Org (%)	Pendeteksian pertama	
					Iterasi 1	Iterasi 2
29	Banyu Ates	740659	9228951	0,43	0,429999416	0,439999116
30	Robatal	751141	9225198	0,49	0,489999417	0,479999314
31	Kedundung	741199	9217846	0,49	0,489999417	0,489999662
32	Tembalangan	734251	9219135	0,85	0,849999423	0,829999157
33	Sampang	750504	9205188	0,37	0,369999415	0,389999178
34	Robatal	753216	9227971	1,12	1,119999427	1,209999731
35	Tembalangan	738280	9215879	0,6	0,599999419	0,609999165
36	Kedundung	749379	9214777	0,18	0,179999412	0,209999982
37	Sokobanah	764151	9235620	0,48	0,479999417	0,469999478
38	Sokobanah	762912	9229845	0,18	0,179999412	0,199999337
39	Banyu Ates	738357	9234093	0,56	0,559999418	0,599999927
40	Ketapang	754155	9237354	1,41	1,409999431	1,429999417
41	Omben	760288	9212724	0,74	0,739999421	0,709999432
42	Omben	760049	9214341	0,8	0,799999422	0,809999378
43	Banyu Ates	735943	9237138	0,06	1,05999941	0,97999982
44	Banyu Ates	739890	9237119	0,6	0,599999419	0,569999314
45	Pangarengan	741088	9202610	0,71	0,70999942	0,699999762
46	Torjun	738082	9206776	0,61	0,609999419	0,629999187
47	Banyu Ates	738238	9235380	0,84	0,839999422	0,889999723
48	Sampang	749823	9202948	1,3	1,29999943	1,49999725
49	Sokobanah	773761	9236855	0,55	0,549999418	0,509999372
50	Sampang	745537	9202481	0,82	0,819999422	0,799999213
51	Sampang	750845	9209566	0,49	0,489999417	0,499999114
52	Torjun	740972	9205729	0,7	0,69999942	0,70999939
53	Sreseh	728518	9205617	0,38	0,379999415	0,359999942
54	Sreseh	734917	9201486	0,59	0,589999419	0,609999326
55	Sreseh	726670	9205945	0,69	0,68999942	0,699999882
56	Pangarengan	739320	9204652	0,57	0,569999418	0,549999163

Lampiran 4. (lanjutan)

Objek	Kecamatan	Koordinat X (m)	Koordinat Y (m)	C-Org (%)	Pendeteksian kedua	
					Iterasi 1	Iterasi 2
1	Jrengik	735477	9212097	0,36	0,359999438	0,359999214
2	Sampang	746985	9202635	0,25	0,249999436	0,259999281
3	Sampang	749316	9205052	0,55	0,549999441	0,519999285
4	Sampang	749908	9202428	1,46	0	0
5	Sampang	747246	9206219	0,62	0,619999442	0,63999927
6	Robatal	751529	9221540	0,61	0,609999442	0,609999227
7	Tembalangan	740151	9226973	1	0,999999448	0,899999261
8	Kedundung	743950	9219329	1,08	1,079999449	1,059999249
9	Robatal	752984	9226817	1,38	1,379999454	1,349999225
10	Robatal	750761	9227490	0,44	0,439999439	0,409999239
11	Kedundung	744023	9218732	0,8	0,799999445	0,789999235
12	Pangarengan	742146	9201876	1,14	1,13999945	1,16999925
13	Sampang	745003	9202700	0,71	0,709999443	0,739999234
14	Robatal	750832	9222562	0,43	0,429999439	0,409999236
15	Kedundung	747331	9219823	0,81	0,809999445	0,829999255
16	Robatal	752200	9222200	0,3	0,299999437	0,279999235
17	Robatal	753678	9229296	1,07	0,68999916	0,699999249
18	Robatal	751604	9225900	0,94	0,939999447	0,909999246
19	Tembalangan	740292	9226698	0,73	0,729999444	0,709999264
20	Ketapang	751256	9232812	0,13	1,129999434	1,159999232
21	Banyu Ates	737942	9232172	0,49	0,48999944	0,45999924
22	Banyu Ates	740319	9229966	0,24	0,239999436	0,279999246
23	Ketapang	752155	9230636	0,76	0,759999444	0,739999244
24	Tembalangan	740079	9224790	0,54	0,539999441	0,579999241
25	Torjun	742524	9213505	1,34	1,339999453	1,439999234
26	Sokobanah	764284	9233452	0,62	0,619999442	0,609999242
27	Robatal	752525	9230250	0,51	0,50999944	0,539999241
28	Tembalangan	740071	9220330	0,54	0,539999441	0,549999244

Lampiran 4. (lanjutan)

Objek	Kecamatan	Koordinat X (m)	Koordinat Y (m)	C-Org (%)	Pendeteksian kedua	
					Iterasi 1	Iterasi 2
29	Banyu Ates	740659	9228951	0,43	0,429999439	0,469999219
30	Robatal	751141	9225198	0,49	0,48999944	0,499999239
31	Kedundung	741199	9217846	0,49	0,48999944	0,479999236
32	Tembalangan	734251	9219135	0,85	0,849999446	0,829999255
33	Sampang	750504	9205188	0,37	0,369999438	0,349999231
34	Robatal	753216	9227971	1,12	1,11999945	1,219999248
35	Tembalangan	738280	9215879	0,6	0,599999442	0,609999246
36	Kedundung	749379	9214777	0,18	0,179999435	0,159999265
37	Sokobanah	764151	9235620	0,48	0,47999944	0,449999237
38	Sokobanah	762912	9229845	0,18	0,179999435	0,199999234
39	Banyu Ates	738357	9234093	0,56	0,559999441	0,529999214
40	Ketapang	754155	9237354	1,41	1,409999454	1,439999236
41	Omben	760288	9212724	0,74	0,739999444	0,709999236
42	Omben	760049	9214341	0,8	0,799999445	0,809999265
43	Banyu Ates	735943	9237138	0,06	1,059999433	1,079999243
44	Banyu Ates	739890	9237119	0,6	0,599999442	0,549999235
45	Pangarengan	741088	9202610	0,71	0,709999443	0,739999246
46	Torjun	738082	9206776	0,61	0,609999442	0,639999272
47	Banyu Ates	738238	9235380	0,84	0,839999445	0,819999248
48	Sampang	749823	9202948	1,3	1,299999453	1,309999244
49	Sokobanah	773761	9236855	0,55	0,549999441	0,579999261
50	Sampang	745537	9202481	0,82	0,819999445	0,849999279
51	Sampang	750845	9209566	0,49	0,48999944	0,459999245
52	Torjun	740972	9205729	0,7	0,699999443	0,719999263
53	Sreseh	728518	9205617	0,38	0,379999438	0,359999234
54	Sreseh	734917	9201486	0,59	0,589999442	0,549999218
55	Sreseh	726670	9205945	0,69	0,689999443	0,699999217
56	Pangarengan	739320	9204652	0,57	0,569999441	0,579999261

Lampiran 5. Peta Kabupaten Sampang Madura



Lampiran 6. Macro Minitab Untuk *Weighted z Algorithm*

```
macro
weighted data x y cluster nbravg1 diff1
stddiff1 outlier1
mcolumn data jcluster cluster bcluster data1 x
x1 y y1
titik11 titik12 jarak1 invjar1 tinvar1 nbravg1
diff1
stddiff1 outlier1
mconstant n i j d1 a k l cjarak weight kali

let n=count(data)
let jcluster[1]=cluster[1]
let j=1
let bcluster[1]=1
do i=2:n
if cluster[i]=cluster[i-1]
let bcluster[j]=bcluster[j]+1
elseif cluster[i]~cluster[i-1]
let j=j+1
let bcluster[j]=1
let jcluster[j]=cluster[i]
endif
enddo
let a=0
do k=1:d1
let tinvar1[k]=0
let nbravg1[k]=0
do l=1:d1
if k~l
let a=a+1
let titik11[a]=k
let titik12[a]=l
let jarak1[a]=sqrt((x1[k]-x1[l])**2)+(y1[k]-
y1[l])**2)
let invjar1[a]=1/jarak1[a]
let tinvar1[k]=tinvar1[k]+invjar1[a]
endif
enddo
do l=1:d1
if k~l
let tinvar1[k]=tinvar1[k]+invjar1[a]
```

Lampiran 6. (lanjutan)

```
endif
enddo
do l=1:d1
if k~=1
let cjarak=sqrt(((x1[k]-x1[l])**2)+(y1[k]-
y1[l])**2)
let cinvjar=1/cjarak
let weight=cinvgjar/tinvjar1[k]
let kali=data1[l]*weight
let nbravg1[k]=nbravg1[k]+kali
endif
enddo
let diff1[k]=data1[k]-nbravg1[k]
enddo

center diff1 stddiff1
do i=1:d1
if stddiff1[i]>2 or stddiff1[i]<-2
let outlier1[i]=1
else
let outlier1[i]=0
endif
enddo
endmacro
```

Lampiran 7. Macro Minitab Untuk *Iterative r Algorithm*

```
macro
iterative data x y cluster nbr1 nbravg1 diff1
stddiff1 outlier1
mcolumn data data1 jcluster cluster bcluster x y
x1 y1 nbr1
nbravg1 diff1 stddiff1 outlier1
mconstant n i j d1 k l

let n=count(data)
let jcluster [1] =cluster[1]
let j =1
let bcluster [1] =1
do i=2:n
    if cluster [i] = cluster [i-1]
        let bcluster [j] = bcluster [j]+1
    elseif cluster [i]~cluster [i-1]
        let j = j+1
        let bcluster [j]=1
        let jcluster [j]=cluster[i]
    endif
enddo

let d1=0
do i=1:n
    if cluster[i]=1
        let d1=d1+1
        let data1[d1]=data[i]
        let x1[d1]=x[i]
        let y1[d1]=y[i]
    endif
enddo
do k=1:d1
let nbr1[k]=0
do l=1:d1
if k~=l
let nbr1[k]=nbr1[k]+data1[l]
endif
enddo
```

Lampiran 7. (lanjutan)

```
let nbravg1[k]=nbr1[k]/(d1-1)
let diff[k]=data1[k]/nbravg1[k]
enddo
center diff1 stddiff1
do i=1:d1
if stddiff1[i]>1 or stddiff1[i]<-1
let outlier[i]=1
else
let outlier[i]=0
endif
enddo
endmacro
```

