

Kombinasi Algoritma *Spatial Autocorrelation G** dan Algoritma C5.0 untuk Deteksi Daerah Rawan Longsor di Pulau Jawa

Yeremia Alfa Susetyo

Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana
Jl. Dr. O. Notohamidjojo 1-10, Salatiga 50715, Jawa Tengah, Indonesia
Email: yeremia.alfa@uksw.edu

Masuk: 6 Agustus 2018 ; Direvisi: 10 September 2018 ; Diterima: 28 September 2018

Abstract. *Java Island is also an island with a high frequency of landslide natural disaster. Various efforts have been made to minimize the disaster risk, including the model compilation of early disaster detection at the areas with landslide potential. This research aims to develop an early warning model for landslide potential areas using Spatial Autocorrelation combined with Machine Learning Algorithm based on attributes of landslide causes. The first step is to classify areas in Java Island into a landslide hotspot and landslide coldspot using spatial autocorrelation G* algorithm. This algorithm revealed 124 polygons of sub-district in Java as landslide hotspots. The next step is building machine learning model using C5.0 method for hotspot and coldspot areas. In this research, we utilize landslide-causing attributes i.e. rainfall, land cover, area slope, soil type, and land movement. The hotspot model showed that landslide distribution focuses on land cover attributes. Meanwhile, for the coldspot area model, there is no focus of landslide distribution on one attribute. Furthermore, the accuracy level of hotspot model are 84.61% and 71.66% for coldspot model.*

Keywords: *Landslide, Machine Learning, Spatial Autocorrelation, C5.0, G**

Abstrak. *Pulau Jawa merupakan pulau dengan frekuensi bencana alam tanah longsor yang tinggi. Untuk meminimalkan resiko bencana, dilakukan penyusunan model komputasi deteksi dini daerah potensi longsor. Tujuan dari penelitian ini adalah membangun model deteksi daerah potensi longsor menggunakan Spatial Autocorrelation yang dikombinasikan dengan algoritma Machine Learning berbasis data variabel-variabel pemicu longsor. Tahap pertama yang dilakukan adalah mengklasifikasikan daerah-daerah di Pulau Jawa sebagai daerah hotspot longsor dan coldspot longsor menggunakan algoritma spatial autocorrelation G*. Dhasilkan 124 poligon kecamatan di pulau Jawa sebagai daerah hotspot longsor. Setelah diklasifikasikan, dibangun model machine learning menggunakan metode C5.0. Atribut-atribut pemicu longsor yang digunakan untuk membangun machine learning adalah curah hujan, tutupan lahan, kelerengan, jenis tanah, dan gerakan tanah. Hasil yang diperoleh dari model hotspot terlihat distribusi longsor memusat pada atribut tutupan lahan dan menghasilkan akurasi model 84.61%. Untuk model coldspot area tidak terlihat adanya pemusatan pada satu atribut pemicu longsor, akurasi untuk model ini adalah 71.66%.*

Kata Kunci: *Longsor, Machine Learning, Spatial Autocorrelation, C5.0, G**

1. Pendahuluan

Indonesia adalah negara kepulauan terbesar di dunia, yang secara geografis berada pada kawasan rawan bencana alam. Salah satu bencana alam yang sering terjadi adalah bencana longsor. Data dari otoritas penanggulangan bencana Indonesia yaitu Badan Nasional Penanggulangan Bencana (BNPB) [1] sampai bulan Juli 2017, jumlah kejadian longsor tahun 2017 telah terjadi sebanyak 458 kejadian, dengan jumlah korban meninggal dunia 106 jiwa. Daerah di Indonesia dengan frekuensi terjadinya longsor paling banyak adalah Provinsi Jawa

Tengah, Provinsi Jawa Barat, dan Provinsi Jawa Timur [2]. Sementara itu, ketiga Provinsi tersebut juga merupakan daerah dengan jumlah penduduk terbanyak di Indonesia [3]. Korelasi daerah paling rawan longsor sekaligus daerah paling banyak penduduknya, membuat besarnya resiko bencana menjadi turut berbanding lurus, yaitu dengan banyaknya jumlah korban jiwa maupun kerusakan infrastruktur [2].

Salah satu cara untuk mengurangi resiko bencana longsor adalah dengan dibangunnya sistem yang mampu mendeteksi terlebih dahulu titik potensi longsor yang kemudian menjadi dasar peringatan dini terjadinya longsor kepada masyarakat. Titik potensi longsor didefinisikan sebagai daerah yang memiliki kecenderungan untuk terjadi longsor. Kecenderungan terjadinya tanah longsor di suatu wilayah dapat dihubungkan dengan adanya kesamaan karakteristik lahan dan iklim di wilayah lain pada kejadian longsor masa lalu. Dengan adanya sistem deteksi dini daerah rawan longsor, maka dapat membantu untuk mengidentifikasi lokasi lain yang memiliki kesamaan ciri-ciri fisik tanah maupun iklim, sehingga daerah tersebut dapat dicurigai sebagai daerah yang berpotensi longsor [4].

Salah satu model yang dapat digunakan dalam mendeteksi daerah rawan longsor adalah *Machine Learning* [5]. Konsep dari metode ini adalah terbentuknya aturan-aturan yang berdasar pada data masa lalu. Data masa lalu diidentifikasi dalam variabel-variabel yang dipetakan dalam bentuk *decision tree* sehingga menghasilkan aturan-aturan sebagai *tools* yang tepat untuk memprediksi suatu kondisi berdasarkan beberapa variabel [6]. Metode ini memiliki kemampuan untuk memodelkan hubungan antar variabel tanpa harus terpaku pada aturan-aturan distribusi atau pembobotan data. Selain itu, tidak diperlukan aturan yang spesifik untuk format data [7]. Tidak bergantungnya metode *Machine Learning* pada pembobotan statistik merupakan kelebihan dari metode ini. Mesin akan secara otonom mempelajari ciri, korelasi, dan tingkat pengaruh variabel dari data masa lampau kemudian akan disimpulkan sebagai suatu *knowledge* yang dapat menggantikan aturan statistik. Dasar dari algoritma *Machine Learning* adalah pemodelan hubungan antar variabel [7]. Variabel-variabel pemicu longsor yang berbasis deret waktu (data historis) tersebut menjadi satu-satunya dasar pembentuk *Machine Learning*.

Sementara itu, faktor lain di samping variabel-variabel pemicu longsor untuk *Machine Learning*, perlu juga memperhatikan pola panel atau pola spasial antar lokasi kejadian longsor. Hal ini dikarenakan pengamatan terhadap kejadian longsor tidak hanya dilihat dari data masing-masing variabel di satu lokasi saja, namun juga dilihat pada beberapa lokasi [8]. Pada dasarnya, data panel atau spasial merupakan gabungan antara data lintas lokasi dan data deret waktu [8].

Analisis keterhubungan antar lokasi dengan lingkungan geografi dapat dieksplorasi menggunakan model *Spatial Autocorrelation* [9]. *Spatial Autocorrelation* juga mendefinisikan hubungan spasial antar fitur yang berdekatan atau bertetangga atau properti yang menentukan hubungan relatif antar elemen spasial [10]. Dengan menggunakan metode *Spatial Autocorrelation* maka frekuensi kejadian longsor antar lokasi dapat diketahui pola distribusi dan hubungannya. Pola hubungan antar lokasi (konektivitas spasial) titik longsor yang dikombinasikan dengan variabel-variabel pemicu longsor dapat menjadi dasar pembangunan *Machine Learning* sistem deteksi dini longsor yang lebih akurat dan informatif, jika dibandingkan dengan *Machine Learning* yang hanya berbasis data deret waktu (data historis) variabel pemicu longsor saja [8].

Tujuan dari penelitian ini adalah membangun model deteksi daerah potensi longsor menggunakan *Spatial Autocorrelation* atau hubungan antar lokasi terjadinya longsor yang dikombinasikan dengan algoritma *Machine Learning* berbasis data variabel-variabel pemicu longsor.

2. Tinjauan Pustaka

Spatial Autocorrelation adalah salah satu model yang digunakan dalam penelitian ini untuk mencari hubungan atau keterkaitan antar lokasi terjadinya longsor. Hal ini didukung oleh penelitian yang dilakukan oleh Rudnick [9], disebutkan bahwa *Spatial Autocorrelation* bertujuan untuk mengeksplorasi hubungan spasial antara fitur yang berdekatan atau bertetangga yang dinyatakan dalam bentuk titik, garis, dan wilayah dalam keruangan. Selain itu, dalam

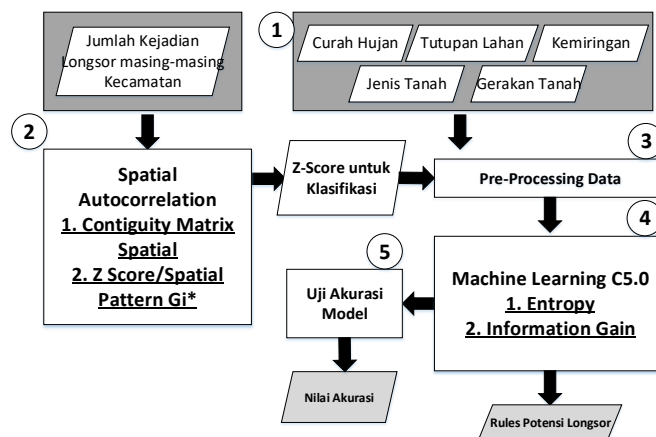
konektivitas spasial di dalamnya terdapat dua kategori: (1) konektivitas struktural yang merepresentasikan karakteristik fisik dari *landscape* yang memungkinkan terjadinya gerakan, termasuk topografi, hidrologi, vegetasi, dan pola penggunaan lahan. (2) Konektivitas fungsional yang merepresentasikan pergerakan populasi bergerak pada *landscape*.

Sementara itu, untuk membangun basis pengetahuan deteksi dini terjadinya longsor digunakan algoritma *Machine Learning*. *Machine Learning* adalah suatu algoritma yang bekerja berdasarkan kejadian masa lalu. Perbedaan karakteristik wilayah juga dapat diselesaikan dengan algoritma ini. Penelitian terdahulu mengenai deteksi daerah potensi longsor menggunakan algoritma *Machine Learning* menyebutkan bahwa kejadian longsor di masa lalu, dapat dijadikan sebagai peringatan untuk mewaspadaai terjadinya longsor di masa yang akan datang. Terjadinya longsor di setiap wilayah memiliki pemicu utama yang berbeda-beda. Salah satu faktor penyebab terjadinya perbedaan adalah kondisi geografi di setiap wilayah berbeda-beda. Algoritma ID3 yang berbasis pada *local-learning* dari kejadian masa lalu, dinyatakan mampu mengakomodasi perbedaan kondisi geografi di berbagai lokal wilayah [11].

Pada tahun 1980, Quinlan mengenalkan algoritma C4.5 sebagai pengembangan algoritma ID3. Kemudian Quinlan terus mengembangkan pohon klasifikasi dan model berbasis aturan menjadi algoritma C5.0 [12]. Algoritma C5.0 adalah salah satu jenis algoritma *Machine Learning* yang dibangun berdasarkan pohon keputusan. Kelebihan algoritma ini adalah mampu mengakomodasi variabel yang bersifat *continue* dan variabel yang bersifat diskrit. Untuk variabel yang bersifat diskrit, akan diperlakukan nilai variabel tersebut sebagai *splitter* [13].

3. Metodologi Penelitian

Pembangunan model deteksi daerah rawan longsor menggunakan Algoritma Gi* dan C5.0 ini terbentuk dari lima tahap (seperti yang ditunjukkan pada Gambar 1) yaitu : (1) tahap pertama adalah pengumpulan data spasial kejadian longsor dan data atribut-atribut pemicu longsor; (2) tahap kedua adalah mencari jenis pola spasial dengan algoritma Gi*; (3) tahap ketiga adalah tahap data preprocessing untuk membentuk data dari data spasial ke dalam bentuk data diskrit (*discretization*); (4) tahap keempat adalah membangun dua *Machine Learning* (untuk daerah *hotspot* dan daerah *coldspot*) deteksi potensi longsor dengan algoritma C5.0; (5) tahap terakhir adalah pengujian akurasi model menggunakan bantuan *Confusion Matrix*.



Gambar 1. Metode Penelitian Deteksi Daerah Potensi Longsor yang Dibangun

3.1. Pengumpulan Data

Untuk membangun pola spasial atau keterhubungan terjadinya longsor antar wilayah, maka digunakan data jumlah kejadian longsor setiap kecamatan (*sub-district*) di tiga provinsi di Indonesia, yaitu: Provinsi Jawa Tengah, Provinsi Daerah Istimewa Yogyakarta, dan Provinsi Jawa Barat. Selanjutnya, data yang digunakan untuk membentuk *Machine Learning* adalah data variabel pemicu longsor, antara lain: kelerengan, curah hujan, tutupan lahan, jenis tanah, dan gerakan tanah [14]. Pemilihan atribut ini, mengacu pada peraturan Menteri Pekerjaan Umum

Republik Indonesia Nomor 22/PRT/M/2007 tentang pedoman penataan ruang kawasan rawan bencana longsor. Data kejadian longsor diperoleh dari Badan Nasional Penanggulangan Bencana (BNPB) Republik Indonesia.

3.2. *Spatial Autocorrelation* dengan G_i^*

Spatial Autocorrelation didefinisikan sebagai korelasi antar wilayah observasi dalam bentuk pola spasial. Luaran dari *Spatial Autocorrelation* dapat bernilai positif jika nilai sekeliling memiliki kemiripan dengan nilai target dan bernilai negatif jika nilai sekeliling memiliki derajat kemiripan yang rendah [15]. Salah satu tahap yang paling penting dalam analisis *Spatial Autocorrelation* adalah membangun matriks bobot (*Weight Matrix*) objek spasial. *Weight Matrix* adalah fungsi untuk memberikan penilaian terhadap hasil dua pengamatan data asli yang berdekatan.

Sebelum membentuk matriks bobot objek spasial harus dilakukan perhitungan matriks kedekatan spasial (*Spatial Contiguity Matrix*) menggunakan fungsi yang ditunjukkan pada Persamaan 1 [16]. Dalam penelitian ini satuan spasial adalah berupa poligon wilayah kecamatan. Selanjutnya poligon kecamatan akan dianalisis ketetanggaannya dengan matriks bobot.

$$C = \begin{bmatrix} C_{11} & \dots & C_{1n} \\ \vdots & \ddots & \vdots \\ C_{n1} & \dots & C_{nn} \end{bmatrix} \quad (1)$$

Dimana C_{ij} adalah ukuran perbandingan nilai kedekatan wilayah ke i dan ke j , sehingga matriks bobot spasial dapat didefinisikan seperti Persamaan 2.

$$W = \begin{bmatrix} W_{11} & \dots & W_{1n} \\ \vdots & \ddots & \vdots \\ C_{n1} & \dots & C_{nn} \end{bmatrix} \quad (2)$$

Dimana nilai W_{ij} ditentukan berdasarkan ketetangaan antar poligon i dan poligon j seperti yang ditunjukkan dengan Persamaan 3.

$$W_{i,j} = \begin{cases} 1, & \text{Jika } i \text{ dan } j \text{ berdekatan} \\ 0, & \text{Jika } i \text{ dan } j \text{ tidak berdekatan} \end{cases} \quad (3)$$

Untuk standar statistika G_i^* satu hal yang menjadi ciri utamanya adalah ketetangaan terhadap dirinya sendiri bernilai 1, seperti yang ditunjukkan pada Persamaan 4 [17]. Maka hal ini menuntut *Spatial Contiguity Matrix* harus melalui tahap normalisasi matriks.

$$W_{i,i} \neq 0 \quad (4)$$

Setelah *Spatial Contiguity Matrix* yang dinormalisasi diperoleh, tahap selanjutnya adalah mendapatkan model pola spasial dengan G_i^* . Pada tahap ini, akan mendapatkan luaran berupa *Z Score* [17], dimana *Z Score* yang diperoleh merepresentasikan pengelompokan poligon kecamatan [18]. Adapun karakteristik pola spasial menurut *Getis-Ord* terdiri dari (1) *Hotspot*, yaitu pola atribut data yang memusat, jika *Z Score* > 1.96 dan (2) *Coldspot*, yaitu pola atribut data menyebar, jika *Z Score* < 1.96 [15]. Persamaan G_i^* ditunjukkan pada Persamaan 5 [18].

$$G_i^* = \frac{\sum_{j=1}^n W_{i,j} X_j - \bar{X} \sum_{j=1}^n W_{i,j}}{S \sqrt{\frac{n \sum_{j=1}^n w_{i,j}^2 - (\sum_{j=1}^n W_{i,j})^2}{n-1}}} \quad (5)$$

Dimana X_j adalah nilai dari atribut j (dalam penelitian ini, adalah nilai jumlah kejadian longsor di suatu kecamatan). W_{ij} adalah bobot spasial antara poligon i dan j . n adalah jumlah poligon kecamatan. Sementara itu S dan \bar{X} diperoleh menggunakan Persamaan 6 dan Persamaan 7 [18].

$$\bar{X} = \frac{\sum_{j=1}^n X_j}{n} \tag{6}$$

$$S = \sqrt{\frac{\sum_{j=1}^n X_j^2}{n} - (\bar{X})^2} \tag{7}$$

Hasil perhitungan statistik G_i^* adalah sama dengan Z Score, sehingga nilai Z Score tidak perlu dihitung lagi [18].

3.3. Data Pre-Processing

Tahap data *pre-processing* adalah tahap penyaringan data dari karakteristik-karakteristik data yang tidak diinginkan, sehingga data menjadi lebih tepat ketika diimplementasikan dalam pemrosesan *data mining* [19]. Pada penelitian ini, data yang diperoleh adalah data spasial, oleh karena itu perlu ada aturan data *pre-processing* yang jelas, agar data dapat diimplementasikan secara tepat dalam algoritma *Machine Learning*. Kebutuhan pertama untuk membangun *Machine Learning* menggunakan algoritma C5.0 adalah pembentukan struktur tabel yang mencakup pemilihan atribut dan tipe datanya serta *role* yang diberikan pada masing-masing atribut. Selanjutnya nilai atribut numerik yang diperoleh dari data spasial diubah dalam bentuk interval label atau diskrit (*discretization*). Aturan diskritisasi ditunjukkan pada Tabel 1.

Dasar dari diskritisasi atribut curah hujan adalah pengelompokan dengan skala 1000, atribut kelerengan dikelompokkan dengan skala 15, atribut tutupan lahan dikelompokkan berdasarkan jenis penggunaan lahan di lokasi, jenis tanah dikelompokkan berdasarkan tingkat potensi kerusakan tanah menurut aturan Departemen Lingkungan Hidup Republik Indonesia [20].

Tabel 1. Nilai Diskrit untuk Atribut Pemicu Longsor

Atribut	Nilai Numerik	Nilai Diskrit
Curah Hujan	< 2000 mm/tahun	Rendah
	2000 – 3000 mm/tahun	Sedang
	> 3000 mm/tahun	Tinggi
Kelerengan	< 15%	Datar
	15% - 30%	Berombak
	> 30%	Curam
Jenis Tanah	Vertisol, Oxisol	Ringan
	Alfisol, Mollisol, Ultisols	Sedang
	Inceptisols, Entisols, Histosols, Spodosol, Andisol	Tinggi
Tutupan	Hutan	Hutan
Lahan	Sawah / Lahan Basah	Sawah
	Kebun / Perkebunan	Kebun
	Pertanian Lahan Kering	Pertanian Kering
	Pemukiman	Pemukiman
Gerakan Tanah	Rendah	Rendah
	Menengah	Menengah
	Tinggi	Tinggi

Pola spasial adalah atribut yang diperoleh dari hasil *Spatial Autocorrelation* jumlah kejadian longsor menggunakan algoritma G_i^* . Hasil yang diperoleh adalah interpretasi nilai z (z score), dimana nilai $z > 1.96$ menunjukkan daerah *hotspot* bencana longsor, sementara nilai $z < 1.96$ menunjukkan daerah *coldspot* bencana longsor [15]. Selanjutnya data dikelompokkan berdasarkan variabel pola spasial, yaitu data yang memiliki nilai pola spasial *hotspot* dan data yang memiliki nilai pola spasial *coldspot*. Dua kelompok data tersebut yang akan digunakan sebagai dasar pembangunan dua *Machine Learning*.

Tahap terakhir dari data *pre-processing* ini adalah masing-masing kelompok data, dikelompokkan lagi menjadi dua kelompok, yaitu data *training* dan data *testing*. Data *training* yang digunakan untuk membentuk model sebanyak 90% dari keseluruhan data, dan data *testing* untuk menguji akurasi model sebanyak 10% dari keseluruhan data.

3.4. *Machine Learning* dengan Algoritma C5.0

Algoritma C5.0 adalah suatu algoritma yang secara khusus digunakan dalam *Machine Learning*. Teknik algoritma ini adalah dengan membangun suatu model pohon klasifikasi. Algoritma C5.0 merupakan algoritma klasifikasi yang dibangun berdasarkan nilai *entropy*, yaitu dengan mengevaluasi semua atribut yang ada untuk mendapatkan tingkat pengaruh suatu atribut dalam mengklasifikasikan sampel data menggunakan suatu ukuran yang dikenal dengan istilah *information gain* [12].

Entropy adalah suatu parameter untuk mengukur heterogenitas atau keberagaman dari suatu kumpulan sampel data. Jika kumpulan sampel data semakin heterogen, maka nilai *entropy* semakin besar [12]. Secara matematis, *entropy* dirumuskan seperti pada Persamaan 8.

$$Entropy(S) = \sum_i^c -p_i \log_2 p_i \quad (8)$$

Dimana c adalah jumlah nilai dari atribut, dan p_i adalah jumlah sampel data dari kelas i . Setelah nilai *entropy* untuk suatu kumpulan sampel data didapatkan, maka tingkat pengaruh atau tingkat efektivitas suatu atribut dalam mengklasifikasikan data dapat diukur. Ukuran efektivitas ini disebut sebagai *information gain* [12]. Secara matematis, *information gain* dari suatu atribut A , dituliskan sebagai Persamaan 9.

$$Gain(S, A) = Entropy(S) - \sum_{v \in \text{values}(A)} \frac{|S_v|}{|S|} Entropy(S_v) \quad (9)$$

Dimana:

- A : atribut
- V : satu nilai yang mungkin dari A
- $\text{Values}(A)$: satu set nilai untuk atribut A
- $|S_v|$: jumlah sampel nilai v
- $|S|$: jumlah keseluruhan sampel data
- S : *entropy* sampel data

3.5. Uji Akurasi dengan *Confusion Matrix*

Untuk menguji tingkat akurasi klasifikasi dan prediksi pada model deteksi daerah potensi longsor dengan algoritma C5.0, maka langkah pertama yang digunakan adalah membangun *Confusion Matrix* [21]. *Confusion Matrix* digunakan untuk mengukur performa dari dua model keputusan yang dihasilkan. Pada Tabel 2 tersebut, yang dimaksud dengan TP adalah jumlah data yang diprediksi YA dan kenyataannya YA, FN adalah jumlah data yang diprediksi YA namun kenyataannya TIDAK, FP adalah jumlah data yang diprediksi TIDAK, namun kenyataannya YA, dan TN adalah jumlah data yang diprediksi TIDAK dan kenyataannya TIDAK.

Tabel 2. *Confusion Matrix*

Prediksi	Aktual	
	YA	TIDAK
YA	TP	FP
TIDAK	FN	TN

Akurasi didefinisikan sebagai rasio data yang diklasifikasikan secara tepat terhadap total keseluruhan data. Adapun persamaan matematika untuk menghitung tingkat akurasi adalah seperti yang ditunjukkan pada Persamaan 10 [21].

$$Akurasi = \frac{TP + TN}{TP + FP + TN + FN} \tag{10}$$

Dimana TP, FP, TN, dan FN diperoleh dari *confusion matrix*.

4. Hasil dan Diskusi

Untuk membangun model deteksi dini daerah rawan longsor ini, digunakan titik-titik observasi. Titik lokasi penelitian yang diambil dikelompokkan menjadi dua, yaitu daerah terjadinya longsor dan daerah yang tidak terjadi longsor pada tahun 2011 sampai 2015 di tiga provinsi di Indonesia, antara lain: Provinsi Jawa Tengah, Daerah Istimewa Yogyakarta, dan Jawa Barat. Diperoleh 590 titik lokasi, dengan statistik 429 titik terjadi longsor dan 161 titik tidak terjadi longsor. Selanjutnya data dikelompokkan menjadi dua kelompok, yaitu data *training* dan data *testing*. Data *training* untuk pembentukan model sebanyak 531 data, dan data *testing* untuk pengujian model sebanyak 10% dari keseluruhan sampel data, yaitu 59 data. Data *training* yang digunakan adalah data kejadian longsor pada bulan Januari 2011 sampai Februari 2015 sedangkan untuk data *testing* adalah kejadian longsor pada bulan Maret 2015 sampai Juni 2015.

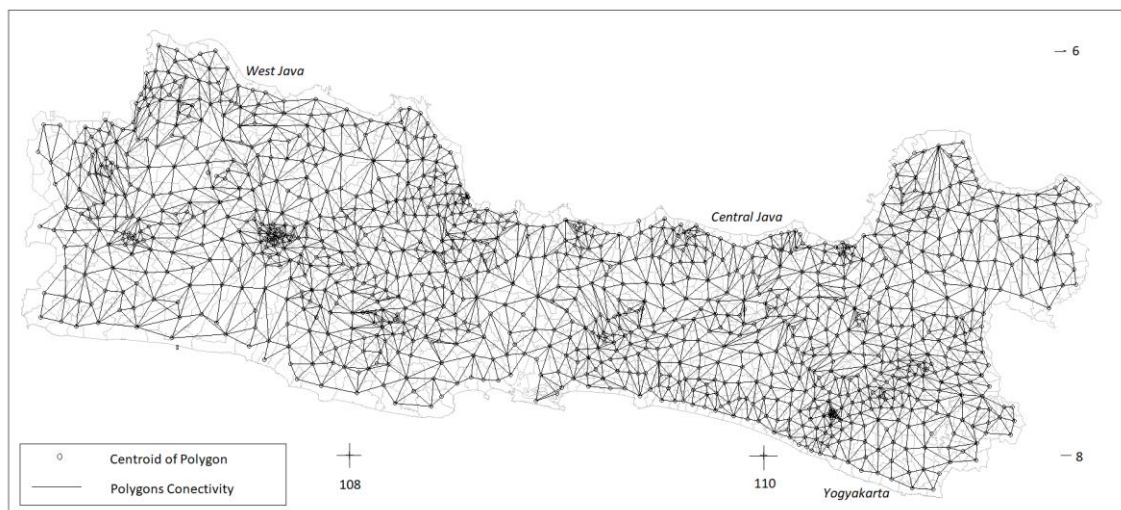
4.1. Spatial Autocorrelation Daerah Potensi Longsor

Longsor merupakan gejala alami yakni suatu proses perpindahan *massa* tanah atau batuan pembentuk lereng dengan arah miring dari kedudukan semula, sehingga terpisah dari *massa* yang mantap karena pengaruh gravitasi, dengan jenis gerakan berbentuk translasi dan/atau rotasi [14]. Salah satu manfaat analisis *Spatial Autocorrelation* adalah melihat adanya konektivitas struktural, yang di dalamnya termasuk representasi karakteristik fisik dari *landscape* yang memungkinkan terjadinya gerakan, termasuk topografi, hidrologi, vegetasi, dan pola penggunaan lahan [9]. Untuk menganalisis pola gerakan longsor beserta konektivitas longsor dengan daerah lain atau poligon lain, maka terlebih dahulu model ini akan membagi wilayah penelitian dalam bentuk poligon-poligon. Poligon yang digunakan dalam penelitian ini adalah poligon wilayah kecamatan yang ada di tiga provinsi yang diobservasi. Jumlah poligon kecamatan adalah 1,277 poligon yang tersebar di tiga provinsi. Selanjutnya, *Contiguity Matrix* akan menganalisis nilai ketetanggaan antar poligon tersebut. Setiap poligon yang berdekatan dengan poligon lain diberi bobot satu, yang artinya antar poligon tersebut terdapat konektivitas. Sebaliknya, setiap poligon yang tidak berdekatan dengan poligon lain diberi nilai nol, yang artinya antar poligon tersebut tidak terdapat konektivitas. Hasil konektivitas antar poligon kecamatan, digambarkan dalam jaringan ketetanggaan seperti yang ditunjukkan pada Gambar 2.

Setelah jaringan konektivitas terbentuk, maka tahap selanjutnya adalah normalisasi masing-masing matriks. Normalisasi matriks mengacu pada Persamaan 4, hal ini dilakukan karena untuk standar statistika G_i^* satu hal yang menjadi ciri utamanya adalah ketetanggaan terhadap dirinya sendiri bernilai satu [19]. Sebagai contoh, salah satu bentuk normalisasi *Contiguity Matrix* untuk poligon Kecamatan Selo, Kabupaten Boyolali, Provinsi Jawa Tengah adalah sebagai berikut:

$$W = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix} \quad \|W\| = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Notasi W adalah matriks sebelum dinormalisasi, $\|W\|$ adalah matriks setelah dinormalisasi. Kecamatan Selo adalah kecamatan yang memiliki konektivitas ketetangaan dengan Kecamatan Getasan, Kecamatan Ampel, Kecamatan Cepogo, Kecamatan Musuk, Kecamatan Srumbung, Kecamatan Dukun, Kecamatan Sawangan, dan Kecamatan Ngablak. Bobot ketetangaan Kecamatan Selo dengan delapan kecamatan lainnya adalah satu, bobot ketetangaan Kecamatan Selo dengan dirinya sendiri pada mulanya adalah nol, namun setelah dinormalisasi menjadi satu.



Gambar 2. Peta Konektivitas antar Poligon Kecamatan Hasil *Contiguity Matrix*

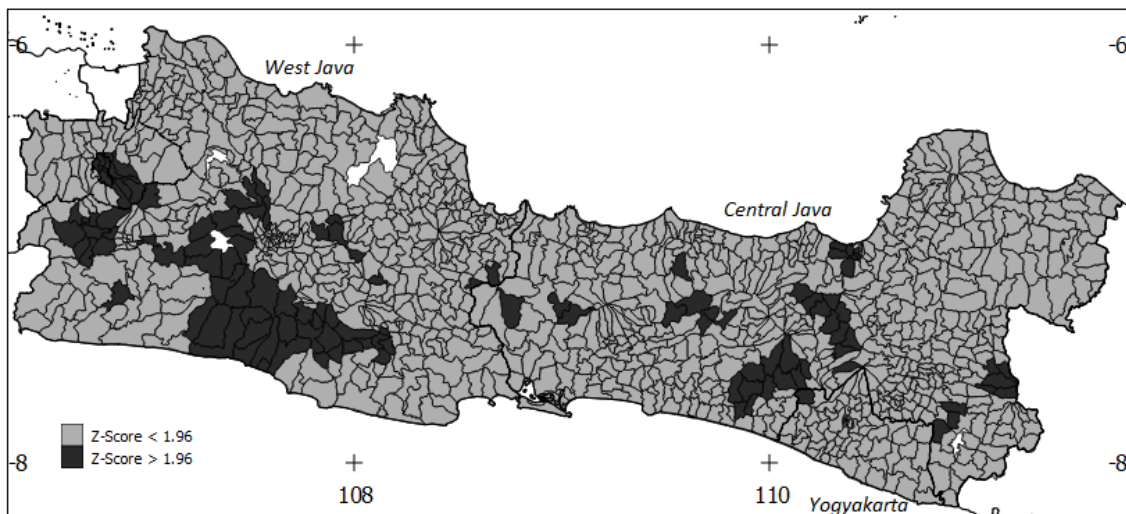
Setelah pola ketetangaan terbentuk lewat normalisasi *Contiguity Matrix*, selanjutnya akan dihasilkan pola distribusi spasial untuk kejadian longsor. Pola distribusi spasial yang dihasilkan mengacu pada Persamaan 5. Mengacu kepada Persamaan 5, variabel X dalam penelitian ini adalah jumlah kejadian longsor di masing-masing poligon kecamatan.

Tahap selanjutnya adalah nilai dari variabel X yang telah diperoleh dikombinasikan dengan variabel W . Hal ini sesuai dengan tujuan konsep *Spatial Autocorrelation* yang digunakan dalam penelitian ini. Lokasi longsor yang menyebar di berbagai kecamatan di Pulau Jawa perlu diperhitungkan pola spasialnya [8][9]. Kombinasi variabel X dan W adalah representasi dari ordo ruang dan ordo waktu, yaitu kejadian longsor dan konektivitasnya dalam hubungan ketetangaan antar poligon (ordo ruang bernilai satu, diperoleh dengan *Contiguity Matrix*) [8]. Algoritma G_i^* bekerja dengan cara melihat setiap fitur dengan bobot konteks ketetanggaannya [20]. Maka mengacu pada Persamaan 5, nilai G_i^* masing-masing poligon kecamatan dapat diperoleh. Nilai G_i^* yang diperoleh adalah representasi dari *Z-Score* masing-masing poligon. Pemetaan nilai G_i^* atau *Z-Score* kejadian longsor ditunjukkan pada Gambar 3.

Gambar 3 menunjukkan daerah yang berwarna gelap pada peta adalah daerah dengan *Z-Score* di atas 1.96, sementara daerah yang berwarna terang, nilai *Z-score* di bawah 1.96. Secara

statistik, interpretasi nilai Z adalah semakin besar *Z-Score*, maka semakin cenderung terbentuk pengelompokan nilai tinggi (*hotspot*). Semakin kecil *Z-Score*, maka semakin cenderung terbentuk pengelompokan nilai rendah (*coldspot*) [20]. Menurut Getis-Ord, jika *Z-Score* bernilai di atas 1.96 ($Z > 1.96$) maka daerah tersebut merupakan daerah *hotspot*, dengan pola distribusi spasial memusat. Sementara *Z-Score* di bawah 1.96 ($Z < 1.96$) maka daerah tersebut merupakan daerah *coldspot*, dengan pola distribusi spasial menyebar [17].

Poligon berwarna gelap yang ditunjukkan pada Gambar 3 adalah daerah *hotspot* lokasi longsor. Daerah-daerah tersebut memiliki konektivitas terpusat sebagai daerah *hotspot* lokasi longsor. Daerah *hotspot* merupakan daerah yang memiliki kecenderungan positif *Spatial Autocorrelation*, artinya memiliki pola yang menuju pemusatan terjadinya bencana longsor. Dinamika konektivitas terpusat terjadinya longsor dipengaruhi oleh banyak faktor, antara lain: (1) perubahan penggunaan guna lahan, (2) perubahan besarnya curah hujan, (3) kesamaan struktur kemiringan lahan, (3) kesamaan jenis tanah dan pola batuan, dan (4) berada dalam satu jenis rangkaian geologi yang sama [22]. Hasil dari penelitian ini adalah diperoleh 124 poligon kecamatan yang merupakan daerah *hotspot* lokasi longsor atau diprediksi sebagai pusat rangkaian terjadinya longsor. Jumlah poligon *hotspot* longsor paling banyak terdapat di Kabupaten Garut, Kabupaten Bandung, dan Kabupaten Sukabumi (Provinsi Jawa Barat). Setelah terbentuk klasifikasi daerah *hotspot* dan daerah *coldspot*, maka pada masing-masing klasifikasi akan dibentuk *Machine Learning* untuk deteksi daerah rawan longsor.



Gambar 3. Peta Gi* atau *Z-Score* Kejadian Longsor di Masing-Masing Poligon Kecamatan

4.2. *Machine Learning* untuk Daerah *Hotspot* Longsor

Daerah *hotspot* longsor adalah daerah dengan pola distribusi longsor memusat atau daerah pusat terjadinya longsor. Untuk membentuk *Machine Learning*, pertama harus menentukan nilai *entropy*. Nilai *entropy* dan *information gain* masing-masing atribut untuk menentukan *the best classifier* dari pohon keputusan. *Entropy* keseluruhan data *training*, diperoleh dengan mengacu pada Persamaan 8:

$$\begin{aligned} \text{Entropy (S)} &= - (169/172) \log_2(169/172) - (3/172) \log_2 (3/172) \\ &= 0.12682573 \end{aligned}$$

Dari lima nilai *information gain* yang dihasilkan, nilai *information gain* tutupan lahan adalah yang terbesar (0.066795096). Sehingga tutupan lahan adalah atribut yang merupakan *the best classifier* dan diletakkan sebagai *root* dari pohon keputusan yang dibentuk. Setelah seluruh

iterasi dilakukan sampai iterasi keempat dan struktur *decision tree* terbangun seluruhnya, maka dihasilkan aturan-aturan untuk mendeteksi longsor seperti yang ditunjukkan pada Tabel 3.

Tabel 3. Rules yang Dihasilkan Model Deteksi Longsor Daerah Hotspot Longsor di Pulau Jawa

Node	Rules untuk Daerah Hotspot Longsor
1	IF (Tutupan_Lahan='pertanian kering') THEN 'Potensi Longsor'
2	IF (Tutupan_Lahan='kebun' AND Lereng='curam') THEN 'Potensi Longsor'
3	IF (Tutupan_Lahan='kebun' AND Lereng='datar' AND gerakan_tanah='rendah' AND jenis_tanah='sedang') THEN 'Potensi Longsor'
4	IF (Tutupan_Lahan='kebun' AND Lereng='datar' AND gerakan_tanah='tinggi') THEN 'Potensi Longsor'
5	IF (Tutupan_Lahan='kebun' AND Lereng='berombak') THEN 'Potensi Longsor'
6	IF (Tutupan_Lahan='sawah') THEN 'Potensi Longsor'
7	IF (Tutupan_Lahan='hutan' AND jenis_tanah='tinggi') THEN 'Potensi Longsor'
8	IF (Tutupan_Lahan='pemukiman') THEN 'Potensi Longsor'

Pada aturan-aturan yang dihasilkan pada model *hotspot* longsor ini, terlihat atribut yang paling berpengaruh dalam terjadinya potensi longsor adalah tutupan lahan. Artinya dinamika konektivitas longsor pada area *hotspot* ini banyak dipengaruhi oleh faktor perubahan tata guna lahan. Daerah dengan tutupan lahan pertanian lahan kering, pemukiman penduduk, dan sawah adalah daerah yang berpotensi longsor cukup kuat. Tiga jenis tutupan lahan tersebut tanpa dikombinasikan dengan atribut lain sudah dapat disimpulkan sebagai daerah rawan longsor pada area *hotspot* longsor. Tutupan lahan adalah satu parameter yang paling sensitif mudah terpengaruh oleh perubahan lingkungan dan aktivitas manusia. Oleh karena itu, struktur tutupan lahan harus menjadi pertimbangan skenario manajemen daerah rawan longsor khususnya manajemen daerah pemukiman, penggunaan lahan untuk pertanian kering dan basah [23].

4.3. Machine Learning untuk Daerah Coldspot Longsor

Setelah nilai keberagaman data atau *entropy* untuk daerah *coldspot* longsor dihasilkan, selanjutnya dihasilkan nilai *information gain*, dimana nilai tertinggi adalah pada atribut tutupan lahan (0.221726911), maka atribut tutupan lahan akan menjadi *the best classifier* atau menempati posisi *root*. Rule yang dihasilkan ditunjukkan pada Tabel 4.

Tabel 4. Rules yang Dihasilkan Model Deteksi Longsor Daerah Coldspot Longsor di Pulau Jawa

No	Rules untuk Daerah Coldspot Longsor
1	IF (Vegetasi = Pemukiman AND Curah_Hujan = Sedang) THEN "Potensi Longsor"
2	IF (Vegetasi = Pemukiman AND Curah_Hujan = Tinggi) THEN "Potensi Longsor"
3	IF (Vegetasi = Pemukiman AND Curah_Hujan = Rendah AND Jenis_tanah = Tinggi AND Gerakan_tanah = Menengah) THEN "Potensi Longsor"
4	IF (Vegetasi = Pemukiman AND Curah_Hujan = Rendah AND Jenis_tanah = Sedang) THEN "Potensi Longsor"
5	IF (Vegetasi = Lahan Kering AND Curah_Hujan = Sedang) THEN "Potensi Longsor"
6	IF (Vegetasi = Lahan Kering AND Curah_Hujan = Tinggi AND Jenis_tanah = Tinggi) THEN "Potensi Longsor"
7	IF (Vegetasi = Lahan Kering AND Curah_Hujan = Tinggi AND Jenis_tanah = Sedang) THEN "Potensi Longsor"
8	IF (Vegetasi = Lahan Kering AND Curah_Hujan = Tinggi AND Jenis_tanah = Ringan) THEN "Potensi Longsor"
9	IF (Vegetasi = Lahan Kering AND Curah_Hujan = Rendah AND Gerakan_tanah = Rendah AND Jenis_tanah = Tinggi) THEN "Potensi Longsor"
10	IF (Vegetasi = Lahan Kering AND Curah_Hujan = Rendah AND Gerakan_tanah = Rendah AND Jenis_tanah = Ringan AND Lereng = Curam) THEN "Potensi Longsor"
11	IF (Vegetasi = Lahan Kering AND Curah_Hujan = Rendah AND Gerakan_tanah = Menengah) THEN "Potensi Longsor"
12	IF (Vegetasi = Lahan Kering AND Curah_Hujan = Rendah AND Gerakan_tanah = Tinggi) THEN "Potensi Longsor"
13	IF (Vegetasi = Sawah AND Lereng = Curam) THEN "Potensi Longsor"

14	IF (Vegetasi = Sawah AND Lereng = Berombak AND Curah_Hujan = Sedang AND Gerakan_tanah = Menengah) THEN "Potensi Longsor"
15	IF (Vegetasi = Sawah AND Lereng = Berombak AND Curah_Hujan = Tinggi) THEN "Potensi Longsor"
16	IF (Vegetasi = Sawah AND Lereng = Datar AND Gerakan_tanah = Tinggi) THEN "Potensi Longsor"
17	IF (Vegetasi = Kebun AND Gerakan_tanah = Rendah AND Lereng = Curam AND Curah_Hujan = Sedang) THEN "Potensi Longsor"
18	IF (Vegetasi = Kebun AND Gerakan_tanah = Menengah AND Lereng = Curam) THEN "Potensi Longsor"
19	IF (Vegetasi = Kebun AND Gerakan_tanah = Menengah AND Lereng = Berombak AND Curah_Hujan = Sedang) THEN "Potensi Longsor"
20	IF (Vegetasi = Kebun AND Gerakan_tanah = Menengah AND Lereng = Datar) THEN "Potensi Longsor"
21	IF (Vegetasi = Kebun AND Gerakan_tanah = Tinggi AND Lereng = Curam) THEN "Potensi Longsor"
22	IF (Vegetasi = Kebun AND Gerakan_tanah = Tinggi AND Lereng = Berombak AND Curah_Hujan = Sedang) THEN "Potensi Longsor"
23	IF (Vegetasi = Hutan AND Lereng = Curam AND Gerakan_tanah = Rendah AND Jenis_tanah = Tinggi AND Curah_Hujan = Rendah) THEN "Potensi Longsor"
24	IF (Vegetasi = Hutan AND Lereng = Curam AND Gerakan_tanah = Menengah AND Jenis_tanah = Tinggi AND Curah_Hujan = Sedang) THEN "Potensi Longsor"
25	IF (Vegetasi = Hutan AND Lereng = Curam AND Gerakan_tanah = Menengah AND Jenis_tanah = Tinggi AND Curah_Hujan = Tinggi) THEN "Potensi Longsor"
26	IF (Vegetasi = Hutan AND Lereng = Curam AND Gerakan_tanah = Tinggi) THEN "Potensi Longsor"

Kesamaan yang didapatkan untuk *Machine Learning* model *hotspot* dan *Machine Learning* model *coldspot* adalah sama-sama menempatkan atribut tutupan lahan sebagai atribut utama pemicu longsor. Hal ini disebabkan karena banyaknya alih fungsi tutupan lahan tanpa memperhatikan asas-asas pelestarian lingkungan [24]. Perbedaan dengan model *hotspot* adalah pada model *coldspot* ini tidak terdapat atribut tunggal pemicu potensi longsor, semua atribut secara merata saling mempengaruhi sebagai pemicu potensi longsor. Khusus wilayah dengan tutupan lahan pemukiman dan pertanian lahan kering akan berpotensi longsor ketika curah hujan tinggi. Sementara itu, area persawahan dan hutan akan berpotensi longsor bila itu berada pada lahan dengan kelerengan curam.

4.4. Uji Akurasi Machine Learning

Data *testing* untuk masing-masing model diterapkan di model yang telah terbentuk. Hasil *Confusion Matrix* ditunjukkan pada Tabel 5 dan Tabel 6 (Untuk model daerah *hotspot* dan *coldspot*). Pada Tabel 5 tersebut ditunjukkan bahwa dari 59 data *testing* yang diuji berdasarkan *rule* pohon keputusan yang telah dihasilkan, diperoleh 11 kejadian untuk *hotspot* area dan 20 untuk *coldspot* area diprediksi berpotensi longsor dan kenyataannya terjadi longsor (*True Positive* - TP), sementara yang diprediksi berpotensi longsor, namun tidak terjadi longsor ada dua kejadian di *hotspot* longsor dan 14 di *coldspot* longsor (*False Positive* - FP). Sebaliknya, ada tiga kejadian diprediksi tidak terjadi longsor di *coldspot*, namun justru terjadi longsor (*False Negative* - FN), dan ada 23 kejadian diprediksi tidak terjadi longsor dan kenyataannya tidak terjadi longsor untuk daerah *coldspot* (*True Negative* - TN). Performa hasil keputusan yang diperoleh untuk dua model ditunjukkan pada *Confusion Matrix* berikut:

Tabel 5. Tabel *Confusion Matrix* untuk Uji *Machine Learning* Daerah *Hotspot* Rawan Longsor

HOTSPOT	Aktual	
	YA	TIDAK
Prediksi		
YA	11	2
TIDAK	0	0

Tabel 6. Tabel Confusion Matrix untuk Uji Machine Learning Daerah Coldspot Rawan Longsor

COLDSPOT	Aktual	
	YA	TIDAK
Prediksi		
YA	20	14
TIDAK	3	23

Dari masing-masing *Confusion Matrix* tersebut, selanjutnya dapat dihitung akurasi masing-masing model, menggunakan Persamaan 10.

$$Akurasi = \frac{TP + TN}{TP + FP + TN + FN}$$

$$Akurasi Hotspot = \frac{11 + 0}{11 + 2 + 0 + 0} = 0.8461$$

$$Akurasi Coldspot = \frac{20 + 23}{20 + 14 + 23 + 3} = 0.7166$$

Tingkat akurasi untuk model *Machine Learning hotspot* longsor adalah 0.8461 atau 84.61%, sedangkan tingkat akurasi untuk model *Machine Learning coldspot* longsor adalah 0.7166 atau 71.66%.

5. Kesimpulan dan Saran

Penelitian ini telah menghasilkan dua model *Machine Learning* untuk deteksi dini daerah potensi longsor. Untuk model deteksi potensi longsor di daerah *hotspot* menghasilkan akurasi 84.61% dan untuk model deteksi potensi longsor di daerah *coldspot* menghasilkan akurasi 71.66%. Selain itu, dihasilkan klasifikasi daerah *hotspot* longsor dan daerah *coldspot* longsor. Terdapat 124 poligon atau wilayah kecamatan di Pulau Jawa yang merupakan daerah *hotspot* longsor. Pada model *Machine Learning* untuk daerah *hotspot* longsor, terdeteksi pusat distribusi longsor dipengaruhi oleh konektivitas kesamaan tutupan lahan atau konektivitas perubahan tata guna lahan. Khususnya daerah dengan tutupan lahan pertanian lahan kering, pemukiman penduduk, dan sawah. Ketiganya dapat menjadi pusat distribusi longsor, sehingga dapat dijadikan pertimbangan bagi pemerintah dalam merencanakan tata ruang wilayah khususnya di 124 kecamatan *hotspot* longsor di Pulau Jawa. Sementara itu, pada model *Machine Learning* daerah *coldspot* longsor tidak terlihat pemusatan faktor terjadinya longsor. Semua atribut menjadi faktor pemicu potensi longsor secara merata.

Untuk penelitian dan pengembangan lebih lanjut, dapat dikembangkan dengan menambahkan atribut-atribut lain pemicu longsor membentuk *Machine Learning* yang lebih detail dan luas. Selain itu, dapat dicoba juga menggunakan algoritma *Spatial Autocorrelation* yang lain, seperti *G* dan *Moran I*. Pengembangan untuk penelitian di wilayah yang lain juga perlu diujicobakan, karena setiap wilayah memiliki ciri-ciri geografi masing-masing.

Referensi

- [1] Badan Nasional Penanggulangan Bencana. (2017) Data Informasi Bencana Indonesia. [Online]. Tersedia: <http://dibi.bnpb.go.id/>
- [2] Badan Nasional Penanggulangan Bencana. (2017) Jumlah Kejadian Bencana Indonesia. [Online]. Tersedia: <http://dibi.bnpb.go.id/data-bencana/crosstab>
- [3] Badan Pusat Statistik. (2010) Data Sensus Penduduk Tahun 2010. [Online]. Tersedia: <http://sp2010.bps.go.id/>

- [4] M. S. Alkhasawneh, U. K. Ngah, L. T. Tay, N. A. M. Isa, and M. S. Al-Batah, "Modeling and Testing Landslide Hazard Using Decision Tree," *Journal of Applied Mathematics*, vol. 2014, pp. 1–9, 2014.
- [5] M. S. Tehrany, B. Pradhan, and M. N. Jebur, "Spatial prediction of flood susceptible areas using rule based decision tree (DT) and a novel ensemble bivariate and multivariate statistical models in GIS," *Journal of Hydrology*, vol. 504, pp. 69–79, 2013.
- [6] A. J. Myles, R. N. Feudale, Y. Liu, N. A. Woody, and S. D. Brown, "An introduction to decision tree modeling," *Journal of Chemometrics: A Journal of the Chemometrics Society*, vol. 18, no. 6, pp. 275–285, 2004.
- [7] R. B. Kheir, M. H. Greve, C. Abdallah, and T. Dalgaard, "Spatial soil zinc content distribution from terrain parameters: A GIS-based decision-tree model in Lebanon," *Environmental Pollution*, vol. 158, no. 2, pp. 520–528, 2010.
- [8] F. B. Wijaya, "Pendekatan Space Time Autoregressive (Star) dan Generalized Space Time Autoregressive (GStar) melalui Metode Autoregressive (AR) dan Vector Autoregressive (VAR)," Doctoral dissertation, Bogor Agricultural University (IPB), 2015.
- [9] D. A. Rudnick and S. J. Ryan, "The Role of Landscape Connectivity in Planning and Implementing Conservation and Restoration Priorities," *The Ecology Society of America*, No. XVI, 2012.
- [10] D. M. Theobald, "Topology revisited: representing spatial relations," *International Journal of Geographical Information Science*, vol. 15, no. 8, pp. 689–705, 2001.
- [11] Y. A. Susetyo, "Detection Model of Landslide-Potential Areas based on Local-Learning using Iterative Dichotomiser Three Algorithm," *International Journal of Computer Science and Information Security (IJCSIS)*, vol. 14 no. 9, September 2016.
- [12] G. P. Siknun and I. S. Sitanggang, "Web-based Classification Application for Forest Fire Data Using the Shiny Framework and the C5.0 Algorithm," *Procedia Environmental Sciences*, vol. 33, pp. 332–339, 2016.
- [13] W. Yogi Yusuf, "Perbandingan Performansi Algoritma Decision Tree C5.0, Cart, Chaid: Kasus Prediksi Status Resiko Kredit di Bank X", *Seminar Nasional Aplikasi Teknologi Informasi 2007*, pp. 59-62, 2007.
- [14] *Peraturan Menteri Pekerjaan Umum Republik Indonesia No. 22/PRT/M/2007 tentang Pedoman Tata Ruang Kawasan Rawan Bencana*, Departemen Pekerjaan Umum Republik Indonesia, 2007.
- [15] S. Y. J. Prasetyo, S. Subanar, E. Winarko, and B. S. Daryono, "ESSA: Exponential Smoothing and Spatial Autocorrelation, Methods for Prediction of Outbreaks Pest in Indonesia," *International Review on Computers and Software (IRECOS)*, vol. 10, no. 4, p. 362, 2015.
- [16] J. Chen and J. Jiang, "Analysis for spatial autocorrelation for point objects based on line buffer", *Proceeding of Accuracy 2010 Symposium*, Leicester, UK, 2010.
- [17] J. K. Ord and A. Getis, "Local Spatial Autocorrelation Statistics: Distributional Issues and an Application," *Geographical Analysis*, vol. 27, no. 4, pp. 286–306, Mar. 2010.
- [18] J. A. Lentz. (2009) Spatial Autocorrelation Statistic.[Online]. Tersedia: <http://jenniferlentz.info/Teaching/StudyGuides/SpatialAutocorrelation.pdf>
- [19] H. Bhalekar and S. Kumbhar, "Pre-processing data using ID3 classifier", *International Journal of Engineering and Techniques*, vol. 1, no. 3, pp. 68-73, June 2015.
- [20] *Pedoman Teknis Penyusunan Peta Status Kerusakan Tanah*, Deputi Konservasi Sumber Daya Mineral Departemen Lingkungan Hidup Republik Indonesia, 2009.
- [21] D. L. Gupta, A. K. Malviya, and S. Singh, "Performance Analysis of Classification Tree Learning Algorithms," *International Journal of Computer Applications*, vol. 55, no. 6, pp. 39–44, 2012.
- [22] M. S. Alkhasawneh, U. K. Ngah, L. T. Tay, N. A. M. Isa, and M. S. Al-Batah, "Modeling and Testing Landslide Hazard Using Decision Tree," *Journal of Applied Mathematics*, vol. 2014, pp. 1–9, 2014.

- [23] T. Kavzoglu, E. K. Sahin, and I. Colkesen, "Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression," *Landslides*, vol. 11, no. 3, pp. 425–439, 2013.
- [24] Maridi and A. Saputra, "Role of Vegetation for Water and Soil Conservation in Watershed: Case Study in 3 Sub-Watershed of Bengawan Solo (Keduang, Dengkeng, dan Samin)", *Seminar Nasional Konservasi dan Pemanfaatan SDA*, Surakarta, 2015.