

BAB I

PENDAHULUAN

1.1 Latar Belakang

Tanah longsor merupakan salah satu bencana alam yang sering terjadi di berbagai wilayah dunia, terutama di daerah yang memiliki topografi curam dan curah hujan tinggi. Bencana ini tidak hanya menimbulkan kerugian ekonomi yang signifikan, tetapi juga mengancam keselamatan jiwa serta menimbulkan kerusakan infrastruktur. Menurut data dari Badan Nasional Penanggulangan Bencana (BNPB), setiap tahun tanah longsor mengakibatkan ratusan korban jiwa serta menimbulkan kerugian material yang mencapai triliunan rupiah. Di Indonesia, tanah longsor menjadi salah satu bencana alam yang paling sering terjadi, dikarenakan intensitas curah hujan yang tinggi di pegunungan dan lereng perbukitan, serta aktivitas tektonik yang intensif (Sulistio et al. (2020) dan Prawiradisastra (2014) menegaskan bahwa Indonesia merupakan negara dengan tingkat kerawanan bencana yang sangat rawan, dengan hampir 40% wilayah Indonesia merupakan daerah dengan risiko longsor tinggi.

Faktor-faktor yang mempengaruhi terjadinya tanah longsor sangat kompleks, karena melibatkan interaksi berbagai unsur fisik seperti topografi, jenis tanah, serta curah hujan, yang dipengaruhi pula oleh faktor antropogenik seperti perubahan tata guna lahan dan deforestasi. Sebagaimana diungkapkan oleh Varnes (1984), kerentanan tanah longsor merupakan fungsi dari kemiringan lereng, jenis batuan, struktur geologi, penutupan lahan, curah hujan, dan aktivitas manusia. Kompleksitas interaksi faktor-faktor ini menyebabkan prediksi tanah longsor menjadi tantangan serius dalam upaya mitigasi bencana. Oleh sebab itu, untuk meminimalkan dampak bencana tersebut, diperlukan langkah mitigasi yang efektif melalui identifikasi dini terhadap wilayah yang rawan longsor. Identifikasi yang akurat terhadap daerah rawan tanah longsor akan sangat membantu pemerintah dan masyarakat dalam menyusun strategi pencegahan serta penanggulangan bencana yang lebih tepat sasaran (Fadilah et al., 2019).

Berbagai teknik untuk memetakan dan memprediksi kerawanan tanah longsor telah dikembangkan karena pentingnya upaya untuk mengurangi bencana tanah longsor. Sebagaimana dinyatakan oleh Hong et al. (2018), keberhasilan upaya mitigasi tanah longsor sangat bergantung pada ketersediaan peta kerawanan yang akurat. Penelitian yang dilakukan oleh Arabameri et al. (2018) menunjukkan bahwa model prediksi kerawanan longsor yang akurat dapat mengurangi korban jiwa hingga 45% dan kerugian material hingga 35% melalui implementasi sistem peringatan dini dan perencanaan tata ruang yang tepat.

Seiring dengan kemajuan teknologi, metode pemodelan berbasis data semakin banyak digunakan dalam berbagai bidang, termasuk dalam analisis kerawanan bencana. Salah satu metode yang mulai banyak diterapkan dalam studi lingkungan adalah metode *random forest*. *Random forest* merupakan algoritma

pembelajaran mesin (*machine learning*) yang memiliki keunggulan dalam menangani data yang kompleks dan menghasilkan prediksi yang akurat. Metode ini mampu mengolah berbagai variabel yang mempengaruhi kerawanan tanah longsor, misalnya topografi, jenis tanah, curah hujan, dan penggunaan lahan, serta memberikan hasil yang lebih handal dibandingkan dengan metode konvensional (N. K. Dewi et al., 2011).

Pemilihan metode *random forest* dalam penelitian ini didasarkan pada beberapa keunggulan yang dimilikinya. Pertama, sebagaimana dikemukakan oleh Breiman (2001), *random forest* memiliki kemampuan untuk mengatasi masalah *overfitting*, yang sering menjadi kendala dalam model prediktif. Kedua, algoritma ini mampu menangani data non-linear dan interaksi kompleks antar variabel, yang sangat relevan dengan karakteristik faktor-faktor penyebab tanah longsor Chen et al. (2017). Ketiga, *random forest* dapat memberikan informasi mengenai pentingnya masing-masing variabel dalam model, sehingga dapat mengidentifikasi faktor-faktor yang paling berpengaruh terhadap kejadian tanah longsor Pourghasemi et al. (2019). Studi yang dilakukan oleh Y. Zhang et al. (2020) menunjukkan bahwa *random forest* memiliki akurasi 15-20% lebih tinggi dibandingkan metode tradisional seperti analisis statistik bivariate dan multivariate dalam memprediksi kerawanan tanah longsor. Hal ini menjadikan *random forest* sebagai pilihan yang tepat untuk menghasilkan peta kerawanan longsor dengan tingkat reliabilitas tinggi.

Kabupaten Barru menjadi batas administrasi yang mendominasi di hulu Daerah Aliran Sungai (DAS) Segeri. Karakteristik hulu DAS umumnya curam dengan kemiringan lereng yang tinggi, elevasi yang bervariasi, serta keanekaragaman jenis tanah yang umumnya mudah lapuk dan berpotensi labil ketika jenuh air selain itu, area hulu idealnya tertutup oleh tutupan vegetasi yang rapat seperti hutan primer atau sekunder yang berfungsi menahan erosi dan menjaga stabilitas lereng. Namun, tekanan antropogenik seperti deforestasi, alih fungsi lahan untuk pertanian atau pemukiman menyebabkan penurunan tutupan hutan yang berakibat pada peningkatan risiko longsor dan erosi, serta perubahan pola hidrologi yang mempercepat laju aliran permukaan (Fitri, 2020). Berdasarkan penelitian Arsyad & Hikmah (2023), Kabupaten Barru memiliki potensi yang cukup tinggi terhadap bencana tanah longsor, yaitu sebesar 61.786 ha dari seluruh total luas wilayah atau sekitar 52,5%.

Tanah longsor dapat menyebabkan kerusakan serius terhadap bentang lahan, sumber daya alam, dan lingkungan, serta berpotensi memicu bencana alam lainnya. Berdasarkan data yang tercatat oleh BNPB, kejadian tanah longsor di Desa Parenreng, Kecamatan Segeri pada tanggal 21 Desember 2020 dan Desa Bulobulo, Kecamatan Pujananting, Kabupaten Barru pada tanggal 21 Februari 2023 menjadi salah satu riwayat kejadian tanah longsor di DAS Segeri yang menyebabkan kerusakan serius terhadap 3 rumah dan akses jalan. Menurut Ulfa et al. (2016), peristiwa tanah longsor dapat menyebabkan kerugian jangka panjang berupa degradasi kualitas tanah dan produktivitas lahan yang dapat berlangsung

hingga puluhan tahun pasca kejadian. Hal ini menjadikan upaya pencegahan dan mitigasi tanah longsor sebagai prioritas yang mendesak dalam pengelolaan risiko bencana di wilayah tersebut.

Hingga saat ini, belum ada pemetaan komprehensif mengenai kerawanan tanah longsor di DAS Segeri yang mengintegrasikan berbagai faktor pemicu secara holistik dengan menggunakan pendekatan pembelajaran mesin. Sebagaimana ditekankan oleh Boyd et al. (2018), pendekatan tradisional dalam pemetaan kerawanan longsor sering kali gagal menangkap interaksi kompleks antara faktor pemicu, sehingga menghasilkan peta kerawanan dengan tingkat akurasi yang terbatas. Oleh karena itu, penelitian ini bertujuan untuk mengidentifikasi distribusi potensi tanah longsor di Daerah Aliran Sungai (DAS) Segeri dengan memanfaatkan algoritma pembelajaran mesin *random forest*. Melalui penerapan teknik pembelajaran mesin ini, diharapkan mampu menghasilkan peta kerawanan tanah longsor yang lebih akurat dan komprehensif, sehingga dapat menjadi acuan ilmiah dalam upaya mitigasi risiko tanah longsor serta pengambilan kebijakan yang tepat terkait tata ruang di wilayah DAS Segeri. Sebagaimana ditekankan oleh Xing et al. (2021), menekankan bahwa pemetaan kerawanan longsor dapat menunjukkan daerah mana yang mungkin mengalami pergerakan lereng di masa depan. Pemetaan kerawanan ini lebih akurat dengan bantuan teknik statistik yang canggih dan model pembelajaran mesin terintegrasi. Ini sangat penting untuk perencanaan wilayah dan mitigasi risiko bencana.

1.2 Tujuan dan Kegunaan Penelitian

Adapun tujuan dari penelitian ini yaitu:

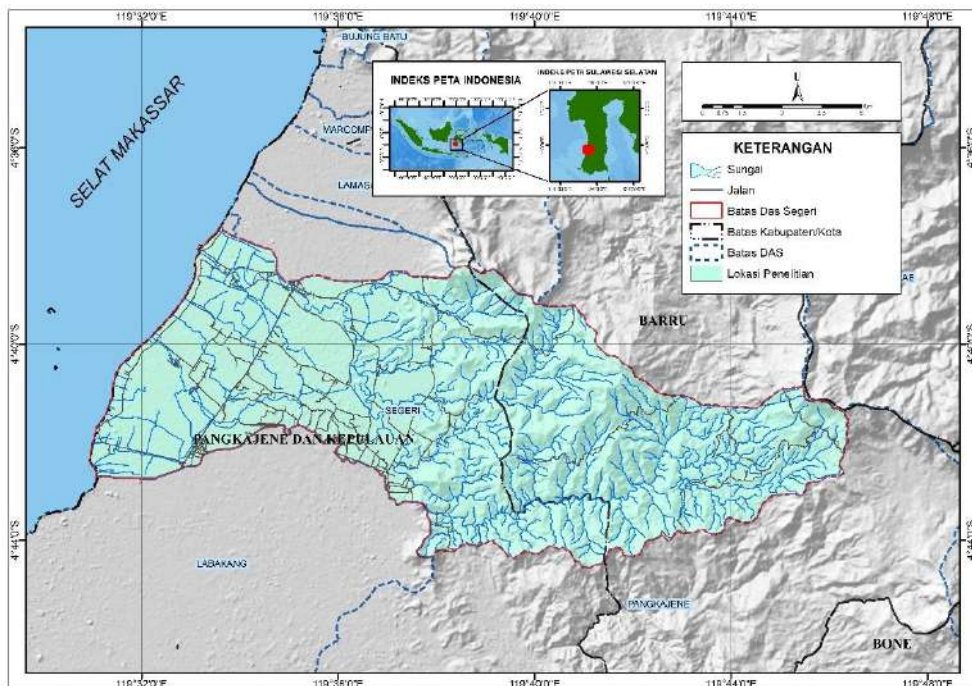
1. Mengidentifikasi tanah longsor di DAS Segeri.
2. Menganalisis faktor yang berpengaruh terhadap terjadinya tanahlongsor di DAS Segeri.
3. Membuat peta kerawanan longsor di DAS Segeri .

Kegunaan dari penelitian ini yaitu memberikan pengetahuan tentang sistem informasi geografis dan *machine learning* dalam perannya melihat area kerawanan Tanah Longsor serta, data hasil penelitian dapat dimanfaatkan sebagai sumber informasi dasar bagi pemerintah setempat dalam pembangunan tata ruang wilayah dan dasar mitigasi bencana kedepannya di DAS Segeri.

BAB III METODE PENELITIAN

2.1 Waktu dan Tempat

Penelitian ini dilakukan dari bulan Maret hingga Oktober 2025 yang dilakukan di wilayah DAS Segeri, yang secara administratif melintasi dua kabupaten. Kabupaten Pangkajene dan Kepulauan dan Kabupaten Barru, yang terletak di Provinsi Sulawesi Selatan. Gambar 1 menunjukkan lokasi penelitian dan seluruh data diolah dan dianalisis di Laboratorium Pengelolaan Daerah Aliran Sungai di Fakultas Kehutanan Universitas Hasanuddin.



Gambar 1. Peta Lokasi Penelitian

2.2 Alat dan Bahan

2.2.1 Alat

Alat yang digunakan dalam penelitian terkait analisis longsor di DAS Segeri dapat dilihat pada Tabel 1.

Tabel 1. Alat yang digunakan dalam proses penelitian.

No	Alat	Jenis/Tipe	Kegunaan
1	Perangkat lunak GIS	ArcGIS	Menganalisis spasial dan

No	Alat	Jenis/Tipe	Kegunaan
		10.8	inventarisasi kejadian longsor
		Google Earth Pro	
		R Studio	Pemodelan dan analisis uji akurasi
2	Perangkat lunak statistik	Microsoft Office	Menganalisis pengaruh <i>variable</i> pada pemodelan
3	Global Position System (GPS)	Avenza Maps	Melakukan <i>ground check</i> kejadian longsor
4	Personal Computer (PC)/Laptop		Mengelola data (analisis dan pemodelan)
5	Handphone		Mendokumentasikan penelitian di lapangan
6	Alat tulis menulis		Mencatat alur penelitian (teknis dan nonteknis)

2.2.2 Bahan

Bahan yang digunakan pada penelitian terkait analisis kerawanan longsor di DAS Segeri dapat dilihat pada Tabel 2.

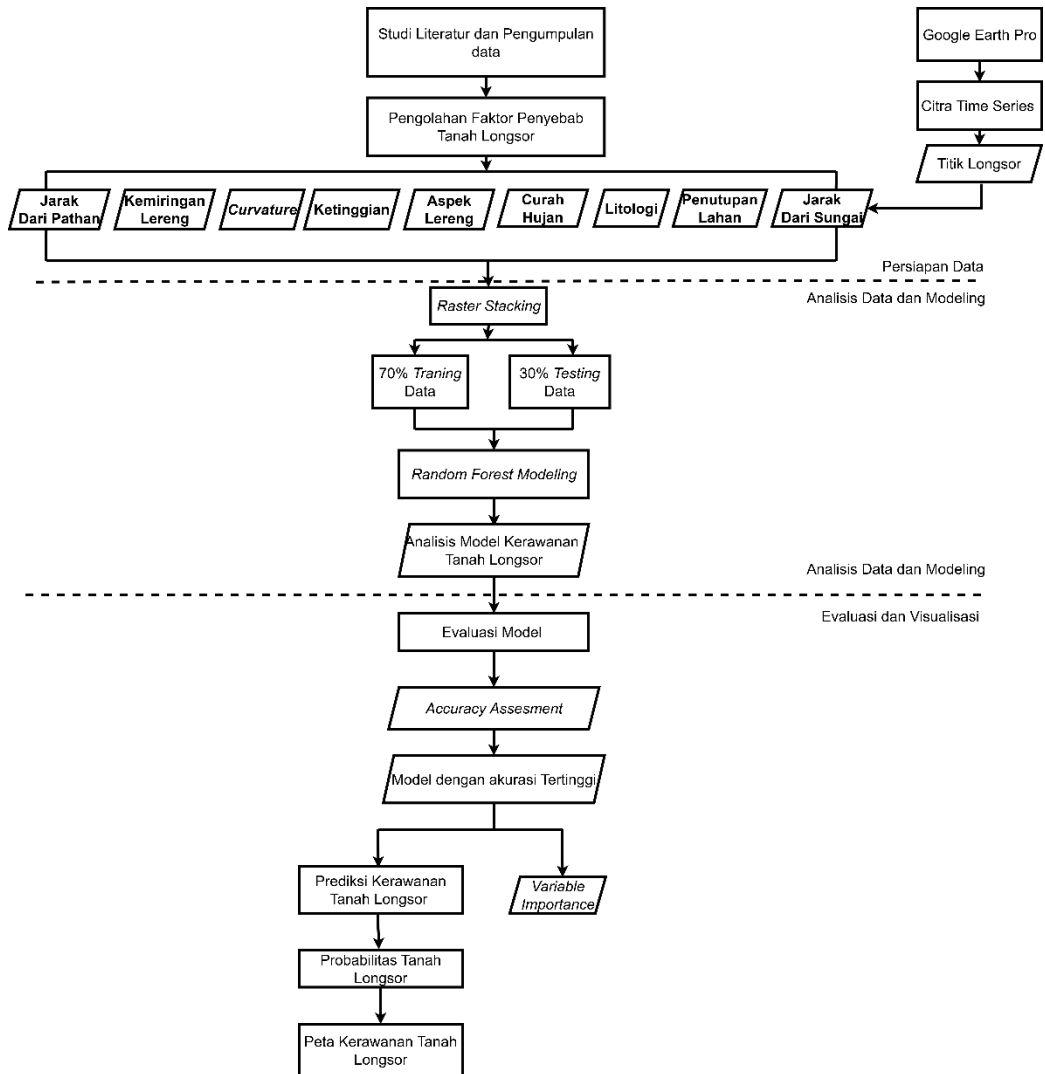
Tabel 2. Bahan yang digunakan dalam proses penelitian.

No	Bahan	Kegunaan	Sumber
1	Citra Time Series Tahun 2019-2025	Untuk menginventarisasi kejadian tanah longsor	Google EarthPro
2	Citra DEMNAS (<i>Digital Elevation Nasional</i>)	Untuk data pembangun faktor kemiringan lereng, ketinggian, arah lereng, kelengkungan bumi, dan jarak dari sungai	INAGEOPORTAL https://tanahair.indonesia.go.id/

No	Bahan	Kegunaan	Sumber
3	Citra Sentinel-2 Tahun 2024	L2A Untuk analisis tutupan lahan	ESA (<i>European Space Agency</i>) https://scihub.copernicus.eu/
4	Peta Rupa Bumi Indonesia Kabupaten Pangkajene dan Kepulauan Kabupaten Barru	(RBI) Untuk dijadikan data pendukung analisis dan administrasi	INAGEOPORTAL https://tanahair.indonesia.go.id/
5	Peta Geologi skala 1:250.000 tahun 2016	Untuk data pembangunan faktor litologi dan jarak dari patahan	Pusat Penelitian dan Pengembangan Geologi Indonesia
6	Data spasial curah hujan PERSIANN-CCS tahun 2019 - 2023	Pembangun faktor curah hujan	Analisis Stasiun merra PERSIANN-CCS https://chrsdata.eng.uci.edu/
7	Peta Batas Daerah Aliran Sungai Segeri	Untuk peta dasar lokasi penelitian	SK/304/MenLHK/PDAS HL/DAS.0/7/2018 Tentang Penetapan Peta Daerah Aliran Sungai

2.3 Alur Penelitian

Metode ini dilakukan melalui beberapa tahapan kerja, yaitu persiapan data, pengolahan data, analisis data, dan evaluasi model, serta hasil penelitian yang secara umum termuat di dalam bagan alur penelitian pada Gambar 2.



Gambar 2. Bagan alur penelitian

2.3.1 Persiapan dan Pengumpulan Data

Data primer dan sekunder digunakan dalam penelitian ini untuk memberikan gambaran kondisi lingkungan dan karakteristik wilayah penelitian. Perangkat lunak R (v4.5.1) dan ArcGIS 10.8.2 digunakan untuk menganalisis semua data. Jenis data sekunder yang digunakan dalam penelitian ini disajikan dalam Tabel 2. Variabel prediktor yang berkontribusi pada pemodelan kerentanan tanah longsor kemudian diolah.

Inventarisasi data primer dilakukan melalui interpretasi citra satelit resolusi tinggi periode 2019 - 2024 menggunakan Google Earth Pro. Setiap kejadian longsor disimpan dalam bentuk poligon untuk merepresentasikan daerah terdampak secara

spasial, dengan titik (point) sebagai penanda lokasi pusat longsor. Validasi lapangan diterapkan pada 30% sampel acak titik longsor teridentifikasi menggunakan Avenza Maps. Kesesuaian antara interpretasi kejadian longsor dengan kondisi aktual. Kesalahan identifikasi objek menjadi dasar penghapusan kejadian longsor dari basis data (Guzzetti et al., 2012). Seluruh data spasial yang digunakan dalam penelitian ini diproyeksikan dalam sistem koordinat UTM-Zone 50S (WGS 1984).

2.3.2 Analisis Data Faktor Penyebab Tanah Longsor

Dalam memprediksi kerentanan kejadian longsor tidak ada pedoman yang tetap untuk pemilihan faktor penyebab yang akan digunakan dalam analisis sehingga, faktor yang dipilih sangat bervariasi antar penelitian (Soma & Kubota, 2017). Pemilihan faktor penyebab tanah longsor penelitian ini mengacu dua prinsip yang dikemukakan (Wang et al., 2021): (1) Prediktor harus memiliki hubungan yang erat dengan mekanisme pemicu serta faktor-faktor yang menyebabkan tanah longsor, atau dengan karakteristik tanah di lokasi longsor; (2) prediktor harus dapat diukur dan tersedia. Pengolahan data faktor yang digunakan pada penelitian ini diolah menggunakan *ArcGIS*, Berikut adalah Langkah kerja dari setiap faktor :

Kemiringan Lereng

Setiap daerah mempunyai kemiringan yang berbeda-beda. Data kemiringan lereng ini diperoleh dari data geoportal DEMNAS. Data mentah ini diproses di *ArcGIS* menggunakan *tools slope* dan diklasifikasikan menggunakan *tools* yang tersedia. Hasil pengelolaan data diklasifikasikan dalam lima kelas yaitu 0-8% (datar), 8%-15% (landai), 15%-25% (agak curam), 25%-45% (curam), dan >45% (sangat curam). Hal inilah yang menjadi faktor penyebab terjadinya tanah longsor. Menurut (Fadilah et al., 2019) mendeskripsikan bahwa semakin curam lereng suatu lahan maka akan semakin besar gaya pendorong dan pengaruh gaya gravitasi terhadap material yang terdapat pada lereng tersebut.

Ketinggian (*Elevasi*)

Peta ketinggian atau *elevasi* dianalisis dari data pengolahan DEMNAS yang dimana prosesnya hampir sama dengan kemiringan lereng tetapi berbeda pada proses klasifikasi yang dibagi menjadi 5 (lima) kelas dengan jarak atau interval 200 m yaitu 0-200 mdpl, 200-400 mdpl, 400-600 mdpl, 600-800 mdpl, dan >800 mdpl. Sesuai dengan pendapat (Tazik et al., 2014) menyatakan bahwa salah satu faktor yang menentukan kerawanan tanah longsor adalah ketinggian, karena kekuatan tanah yang jatuh lebih besar jika lokasinya lebih tinggi karena pengaruh gaya gravitasi.

Aspek lereng

Arah lereng didapatkan dari pengolahan data DEMNAS hasil klasifikasi peta aspek lereng menghasilkan 8 kelas arah lereng yaitu utara, timur, tenggara, selatan, selatan barat daya, barat, utara barat daya. Aspek lereng mempengaruhi longsor

karena secara langsung menerima sinar matahari, sehingga mengeringkan tanah yang membentuk lereng. Apabila musim hujan tiba, air akan masuk ke dalam rongga tanah yang terbuka dan mengakibatkan longsor dan erosi tanah (Hidayah et al., 2017).

Kelengkungan Bumi (*Curvature*)

Kelengkungan bumi diperoleh dari data geoportal DEMNAS. Dalam membuat peta digunakan *tools contour* di *ArcGIS*. Analisis *contour* akan menghasilkan 3(tiga) kelas morfologi topografi yaitu cembung, datar, dan cekung. Hasil morfologi topografi ini menunjukkan bahwa daerah penelitian mempunyai ciri topografi yang beragam. Menurut (Sujatha et al., 2012) kemungkinan terjadinya tanah longsor pada lereng yang cekung sangat tinggi karena lereng cekung memiliki kecenderungan dalam menahan kelembaban tanah, sehingga rentan terhadap aktivitas terjadinya tanah longsor.

Curah hujan

Curah hujan dalam penelitian ini diperoleh dari perhitungan rata-rata curah hujan harian selama periode 2019 hingga 2024. Data ini diambil dari PERSIANN-CCS (<https://chrdata.eng.uci.edu/>), kemudian diolah pada *software MS.Excel*. Data semula berupa data tabular diolah lebih lanjut menggunakan *software ArcGIS* dengan metode Isohyet pada *tools Inverse Distance Weighted* (IDW). Metode ini dipilih karena area penelitian memiliki topografi yang sebagian besar berupa perbukitan. Menurut (Nurhijriah et al., 2022), metode Isohyet dianggap sebagai metode akurat dalam menghitung rata-rata kedalaman hujan di suatu wilayah perbukitan.

Penutupan Lahan

Penutupan lahan menjadi salah satu indikator yang berpengaruh dalam terjadinya tanah longsor. Dalam pengelolaan data penutupan lahan didapatkan dari Citra Sentinel 2A tahun 2024 pada platform *Esa Copernicus* yang akan dilakukan digitasi *onscreen* dengan acuan klasifikasi dari Direktorat Perdirjen Planologi tahun 2020. Hasil digitasi tutupan lahan diuji untuk memverifikasi keakuratan sistem dengan membandingkan data prediksi dan data lapangan. Perhitungan akurasi klasifikasi citra dapat dilakukan menggunakan metode *confusion matrix* yang terdapat pada Tabel 3.

Tabel 3. *Confusion matrix* validasi tutupan lahan

		Data Acuan			Total Baris	<i>User's Accuracy</i>
		A	B	C		
Data Interpretasi Citra	A	Xii			Xi+	Xii/Xi+
	B		Xii			
	C			Xii		
Total Kolom		X+i			N	
<i>Producer's Accuracy</i>		Xii/X+i				Σ Xii

Sumber : Jaya, 2007

Pengujian ketelitian klasifikasi bertujuan untuk melihat kesuksesan suatu klasifikasi. Perhitungan akurasi suatu klasifikasi menggunakan *kappa accuracy* (KA) dengan persamaan sebagai berikut :

$$Kappa\ Accuracy = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^y (x_{ii} + x_{+i})}{N^2 - \sum_{i=1}^r (x_{ii} + x_{+i})} \times 100\% \quad (2.2)$$

Keterangan :

A,B,C = Data acuan

X_{ii} = Data yang diuji

X_{+i} = Jumlah masing – masing data acuan

X_{i+} = Jumlah masing – masing data klasifikasi citra

N = Total data yang diuji

Jarak dari sungai

Jarak dari sungai adalah ukuran jarak horizontal antara suatu tempat atau objek dan sungai terdekat. Ukuran ini umum digunakan dalam hidrologi, geomorfologi, dan perencanaan penggunaan lahan. Faktor penting dalam kerawanan tanah longsor adalah jarak dari sungai karena dapat mempengaruhi risiko terjadinya tanah longsor. Saat hujan, titik terendah air hujan berada di atas permukaan, sementara limpasan dari bawah permukaan mengisi dan mengalir ke sungai. Air yang mengalir lambat laun akan mengikis dinding sungai, jika posisi sungai dekat dengan lereng. Menurut Hidayah et al., (2017) menyatakan bahwa semakin dekat lereng dengan sungai maka semakin tinggi tingkat kejenuhan air pada lereng secara langsung dapat berpengaruh pada stabilitas lereng. Peta jarak dari sungai menggunakan data sungai dari RBI Tahun 2019. Setelah diunduh, selanjutnya diolah menggunakan software ArcGIS dengan tools *euclidean distance*. Klasifikasi data jarak dari sungai diantaranya 0-100m, 100-200m, 200-300m, serta >300 m.

Litologi

Litologi merupakan variabel standar yang mengendalikan bahaya longsor. Hal ini berkaitan dengan kekuatan material karena komposisi dan struktur litologi bervariasi untuk jenis batuan yang berbeda, menentukan bagaimana batuan menyerap air, merespon gaya gravitasi, serta seberapa besar ketahanan batuan terhadap erosi dan pelapukan (Soma & Kubota, 2017). Peta litologi yang digunakan dalam penelitian ini diperoleh melalui proses ekstraksi data dari peta geologi yang diterbitkan oleh Pusat Penelitian dan Pengembangan Geologi Indonesia. Data ini memberikan informasi terperinci mengenai distribusi jenis-jenis batuan di wilayah penelitian, yang menjadi dasar penting dalam analisis kerawanan tanah longsor.

Jarak dari patahan

Jarak dari patahan Salah satu faktor penting yang mempengaruhi kemungkinan terjadinya tanah longsor. Patahan yang aktif, karena pergerakan tanah dan tekanan tektonik yang terus menerus, dapat menyebabkan ketidakstabilan di sekitarnya. Dalam penelitian ini, *tools eucliden distance*. Digunakan untuk mengolah data dari patahan untuk membuat peta zonasi jarak dari patahan geologi. Jarak patahan dari 0 -500meter, 500-1000 meter, 1000-1500meter, da >1500meter adalah lima kelas jarak patahan yang dipilih berdasarkan hasil pengolahan data (Soma et al., 2019). Pembagian ini bertujuan untuk memeriksa hubungan antara kedekatan dengan patahan dan tingkat kerawanan longsor.

Setelah semua komponen yang menyebabkan tanah longsor diproses, langkah berikutnya adalah menyatukan mereka ke dalam satu *raster stack*. Dengan metode ini, sekumpulan nilai variabel prediktor yang lengkap dan konsisten dapat digunakan untuk menunjukkan setiap piksel dari wilayah penelitian. Dengan cara ini, setiap lokasi dalam penelitian dapat dianggap sebagai unit analisis yang secara bersamaan mengandung semua informasi tentang faktor penyebab. Selanjutnya, hasil stacking raster diekstraksi ke dalam bentuk tabular, yang berfungsi sebagai dasar untuk tahap pemodelan. Dengan demikian, sebelum dilakukan pemodelan menggunakan algoritma *Random forest*, telah tersedia satu dataset komposit yang siap digunakan untuk membangun model kerentanan tanah longsor.

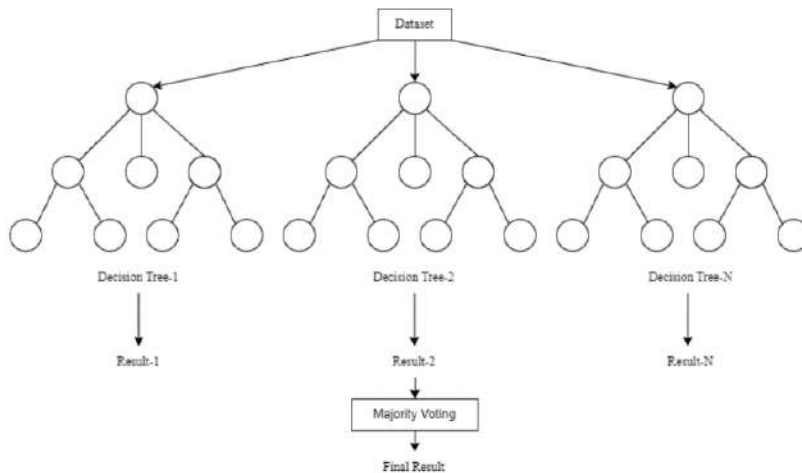
2.3.3 Analisis Model

Pemodelan kerawanan tanah longsor dalam penelitian ini dilakukan menggunakan algoritma *random forest* (RF) melalui paket ranger pada perangkat lunak R. Seluruh skrip yang digunakan dalam proses pemodelan ini dapat dilihat secara rinci pada Lampiran 5. Algoritma ini membangun sejumlah pohon keputusan dan menghasilkan prediksi akhir melalui mekanisme *majority voting*, yang efektif dalam mengurangi varians dan meningkatkan ketahanan model terhadap data spasial yang kompleks (Breiman, 2001).

Dataset diperoleh dari ekstraksi nilai raster pada titik longsor dan titik non-longsor, yang dihasilkan secara acak dengan jarak buffer 1000 meter untuk menghindari overlap spasial, dengan rasio non-longsor:longsor sebesar 3:1. Sembilan variabel prediktor, yakni aspek lereng, curah hujan, kelengkungan, elevasi, jarak dari patahan, jarak dari sungai, kemiringan lereng, jenis litologi, dan penutupan lahan, diperoleh dari stack raster yang disusun dari file TIFF hasil klasifikasi geospasial. Semua variabel tersebut telah diolah sebelumnya menjadi data kategorikal melalui proses klasifikasi, sehingga model hanya menangani tipe data nominal atau ordinal tanpa komponen kontinu.

Data selanjutnya dibagi menjadi subset pelatihan (70%) dan pengujian (30%), dengan penerapan balancing pada data pelatihan melalui *downsampling* kelas mayoritas untuk mengatasi ketidakseimbangan kelas. *Hyperparameter tuning* dilakukan secara bertahap: pertama, nilai *mtry* (jumlah variabel kategorikal yang dipertimbangkan pada setiap split) diuji pada rentang 2–9 dengan jumlah pohon tetap 500, dan nilai optimal dipilih berdasarkan tingkat kesalahan out-of-bag (OOB)

terendah; kedua, model dilatih ulang dengan variasi jumlah pohon (*ntree*) dari 100 hingga 1000 menggunakan *mtry* optimal tersebut. Pendekatan ini sejalan dengan rekomendasi Trisasongko et al. (2017), yang menekankan penyesuaian *mtry* dan *ntree* untuk meningkatkan performa RF pada data spasial, serta memanfaatkan OOB error sebagai metrik evaluasi internal yang andal. Selanjutnya, Gambar berikut menyajikan ilustrasi konsep dasar algoritma *random forest* (Abdi et al., 2024).



Sumber: (Abdi et al., 2024)

Gambar 3. Ilustrasi *random forest*

2.3.4 Evaluasi Model

Setelah parameter optimal ditentukan dan model *random forest* selesai dibangun, tahap selanjutnya adalah melakukan evaluasi kinerja model. Proses evaluasi sangat penting untuk mengukur kemampuan model untuk mengklasifikasikan kejadian longsor dengan tepat dan memastikan bahwa hasil prediksi yang dihasilkan konsisten dan dapat dipertanggungjawabkan secara ilmiah. Evaluasi dilakukan dengan menggunakan sejumlah metrik kuantitatif yang umum digunakan dalam penelitian pembelajaran mesin. Tujuannya adalah untuk menilai tingkat akurasi secara keseluruhan, serta kemampuan model untuk membedakan kelas positif dan negatif, serta stabilitas prediksi pada data.

ROC AUC

Kurva ROC (*Receiver Operating Characteristic*) digunakan untuk validasi model penelitian ini untuk mengevaluasi dan memilih model *random forest* terbaik. *Sensitivity* (*true positive rate*) dan *1-specificity* (*false positive rate*) berhubungan satu sama lain pada berbagai ambang batas klasifikasi melalui kurva ROC. Nilai AUC (*Area Under the Curve*) tertinggi dari data tes, yang berkisar antara 0 dan 1, menunjukkan bahwa model dapat membedakan kelas longsor dan non-longsor

dengan sangat baik. Nilai AUC yang lebih tinggi dari 1 menunjukkan kemampuan model untuk membedakan antara kelas longsor dan non-longsor, seperti yang ditunjukkan dalam tabel 5 (Kleinbaum, 2010).

Tabel 4. Klasifikasi nilai interpretasi AUC

Nilai AUC	Interpretasi
$AUC \geq 0.9$	<i>Excellent</i>
$0.8 \leq AUC < 0.9$	<i>Considerable</i>
$0.7 \leq AUC < 0.8$	<i>Fair</i>
$0.6 \leq AUC < 0.7$	<i>Poor</i>
$0.7 \leq AUC < 0.6$	<i>Fail</i>

Confusion matrix

Hubungan antara kelas sebenarnya dan kelas prediksi digambarkan dalam matriks *confusion matrix*. Tabel 4 berikut menunjukkan isi matrix confusion dari hasil klasifikasi (Abdi et al., 2024).

Tabel 5. *Confusion matrix* validasi model

	Prediksi Negatif	Prediksi Positif
Aktual Positif	<i>False negative (FN)</i>	<i>True positive (TP)</i>
Aktual Negatif	<i>True negative (TN)</i>	<i>False Positive (FP)</i>

Keterangan:

TP: Kelas yang diprediksi adalah positif dan prediksi tersebut sesuai dengan kelas yang sebenarnya.

TN: Kelas yang diprediksi negatif dan prediksi sesuai dengan kelas yang sebenarnya.

FP: Kelas yang diprediksi positif tetapi prediksi tidak sesuai dengan kelas yang sebenarnya.

FN: Kelas yang diprediksi negatif tetapi prediksi tidak sesuai dengan kelas yang sebenarnya.

Berdasarkan tabel *confusion matrix* seperti ditunjukkan di atas, maka nilai akurasi, *Sensitivity*, *Specifty*, dan *Balanced Accuracy* selanjutnya dapat dihitung dengan formulasi berikut:

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (2.3)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (2.4)$$

$$Specifty = \frac{TN}{TN + FP} \quad (2.5)$$

$$Precision = \frac{TP}{TP + FP} \quad (2.6)$$

$$Balanced Accuracy = \frac{sensitivity + Specifty}{2} \quad (2.7)$$

$$F1 Score = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \quad (2.8)$$

Akurasi menggambarkan proporsi prediksi yang benar secara keseluruhan, *sensitivity* mengukur kemampuan model mendeteksi kejadian longsor, sedangkan *specifty* mengukur kemampuan mengenali kelas *non*-longsor. Selain itu, penelitian ini juga menghitung *balanced accuracy* yang merupakan rata-rata dari *sensitivity* dan *specifty*, sehingga dapat memberikan gambaran kinerja model yang lebih seimbang terutama pada data yang memiliki distribusi kelas tidak seimbang. Sebagai pelengkap, penelitian ini juga mengevaluasi *precision*, dan *F1-score* untuk menilai performa model secara lebih komprehensif.

2.3.5 Analisis Faktor Penyebab Longsor

Dalam penelitian ini, kontribusi masing-masing komponen penyebab longsor diteliti dengan menggunakan ukuran *variable importance* pada algoritma *random forest*. Menurut (R. Dewi et al., 2016), Salah satu hasil penting dari *random forest* adalah variabel signifikansi, yang menunjukkan seberapa besar peran variabel penjelas dalam meningkatkan akurasi model. Analisis ini tidak hanya membantu menemukan faktor yang paling dominan yang mempengaruhi kejadian longsor, tetapi juga menunjukkan variabel yang kontribusinya sangat kecil sehingga relevansinya dalam model dapat dipertimbangkan kembali.

Dalam penelitian ini, dua metode digunakan untuk menghitung *variable penting*: *Mean Decrease Gini* (MDG) dan *Mean Decrease Accuracy* (MDA). Ketika suatu variabel digunakan sebagai pemisah dalam *node* pohon, MDG menghitung penurunan *impurity* rata-rata (menggunakan indeks Gini). Berikut ini adalah rumus umum *Millennium Development Goals* (MDG) (Louppe et al., 2013):

$$MDG(X_m) = \left(\frac{1}{N_T}\right) \sum_T \sum_{\{t \in T: v(s_t) = X_m\}} p(t) \Delta i(s_t, t) \quad (2.9)$$

N_T adalah jumlah pohon, $p(t)$ proporsi sampel yang mencapai node t , dan $\Delta i(s_t, t)$ adalah penurunan impurity pada node t akibat split s_t .

Sementara itu, MDA atau *permutation importance* mengukur seberapa besar akurasi model menurun ketika nilai suatu variabel diacak secara acak pada data *out-of-bag*. Rumus MDA dapat dirumuskan sebagai berikut (Altmann et al., 2010):

$$MDA(X_m) = \left(\frac{1}{N_T} \right) \sum_T (err_{oob}^T - err_{\{oob, perm(X_m)\}}^T) \quad (3.0)$$

err_{oob}^T adalah error OOB asli pada pohon T dan $err_{\{oob, perm(X_m)\}}^T$ adalah error setelah variabel X_m diacak. Nilai MDA yang lebih besar menunjukkan bahwa variabel tersebut berperan penting dalam klasifikasi.

Perlu dipahami bahwa peran setiap faktor dapat bervariasi antar lokasi, bergantung pada kondisi geo-lingkungan serta mekanisme longsor yang dominan di wilayah tersebut. Hal ini sejalan dengan (Y. Zhang et al., 2020) yang menyatakan bahwa faktor-faktor geo-lingkungan dapat memiliki pengaruh besar pada satu wilayah tetapi hanya memiliki pengaruh kecil di wilayah lain. Oleh karena itu, hasil dari *variable importance* bersifat kontekstual dan menunjukkan ciri-ciri khusus dari wilayah penelitian.

2.3.6 Peta Kerawanan Tanah Longsor

Analisis kerawanan tanah longsor diperoleh dari hasil analisis probabilitas kejadian longsor yang diprediksi oleh model terbaik *random forest* (RF). Probabilitas ini diperoleh dari proporsi pohon keputusan dalam RF yang mengklasifikasikan suatu piksel sebagai longsor, sehingga setiap piksel pada hasil raster *stacking* memiliki nilai peluang (0–1) yang mencerminkan tingkat kerawanan sesuai kombinasi faktor prediktornya. Dengan demikian, probabilitas yang dihasilkan RF tidak hanya memberikan informasi biner tentang apakah longsor atau tidak longsor, tetapi juga memberikan estimasi kuantitatif yang lebih akurat tentang kemungkinan longsor terjadi di setiap lokasi (Breiman, 2001).

Nilai probabilitas tersebut kemudian diklasifikasikan menjadi peta tematik menggunakan metode *natural breaks* (Jenks). Menurut (Fariza et al., 2022), teknik ini berhasil memisahkan distribusi data ke dalam kelas-kelas yang terbentuk secara alami. Ini memungkinkan untuk menunjukkan perbedaan nyata dalam tingkat kerawanan. Sangat rendah, rendah, sedang, tinggi, dan sangat tinggi adalah lima kategori kerawanan yang digunakan dalam penelitian ini. Untuk membuat pemetaan kerawanan berbasis *machine learning* lebih mudah dipahami dan digunakan untuk kepentingan kebijakan, skema ini mengubah hasil probabilitas algoritma klasifikasi ke dalam kategori spasial.