

Analisis Kesesuaian Lahan Pertanian di Provinsi Jawa Timur Menggunakan Pendekatan Random Forest dan Sistem Informasi Geografis

Isna Kumala Sari^{1*}, Ardiana Fatma Dewi¹

¹Universitas Islam Negeri Sayyid Ali Rahmatullah Tulungagung, Tulungagung, Indonesia

isnakumalasari@uinsatu.ac.id*

| Received: 08/06/2026 | Revised: 29/06/2026 | Accepted: 13/07/2026 |

Copyright©2026 by authors. Authors agree that this article remains permanently open access under the
terms of the Creative Commons

Abstrak

Provinsi Jawa Timur merupakan salah satu sentra produksi pertanian utama di Indonesia dengan karakteristik biofisik yang beragam. Variasi kondisi topografi, iklim, dan tutupan lahan menyebabkan tingkat kesesuaian lahan pertanian berbeda antarwilayah. Penelitian ini bertujuan untuk menganalisis kesesuaian lahan pertanian di Provinsi Jawa Timur menggunakan algoritma *Random Forest* berbasis data geospasial multisumber. Variabel yang digunakan meliputi elevasi, kemiringan lereng, aspek lereng, curah hujan, suhu permukaan, pH tanah, kandungan karbon organik tanah, *Normalized Difference Vegetation Index* (NDVI), *Enhanced Vegetation Index* (EVI), dan *Normalized Difference Water Index* (NDWI). Data diperoleh dari Landsat 8, SRTM, CHIRPS, SoilGrids, dan ESA WorldCover. Model *Random Forest* dibangun menggunakan 80% data pelatihan dan 20% data pengujian. Hasil evaluasi menunjukkan bahwa model memiliki akurasi sebesar 80,12%, nilai Kappa sebesar 0,4212, serta ROC-AUC sebesar 0,8596 yang menunjukkan kemampuan klasifikasi yang baik. Analisis *feature importance* menunjukkan bahwa kemiringan lereng merupakan faktor paling berpengaruh dengan kontribusi sebesar 18,43%, diikuti oleh NDWI (14,69%), elevasi (12,24%), dan NDVI (11,63%). Hasil pemetaan menunjukkan bahwa kelas kesesuaian sedang mendominasi wilayah Jawa Timur dengan luas sekitar 3,28 juta ha (48,71%), diikuti oleh kelas kesesuaian rendah sebesar 2,00 juta ha (29,75%) dan kelas kesesuaian tinggi sebesar 1,45 juta ha (21,54%). Penelitian ini menunjukkan bahwa pendekatan *machine learning* dan data geospasial multisumber dapat digunakan secara efektif untuk mendukung perencanaan pertanian berkelanjutan dan pengambilan keputusan berbasis spasial di Provinsi Jawa Timur.

Kata kunci: Jawa Timur, kesesuaian lahan, pertanian, *Random Forest*, Sistem Informasi Geografis

Abstract

East Java Province is one of the major agricultural production centers in Indonesia with diverse biophysical characteristics. Variations in topography, climate, soil properties, and vegetation conditions contribute to differences in agricultural land suitability across the region. This study aimed to analyze agricultural land suitability in East Java Province using a *Random Forest* algorithm integrated with Geographic Information Systems (GIS) and multisource geospatial data. The variables employed included elevation, slope, aspect, rainfall, temperature, soil pH, soil organic carbon (SOC), *Normalized Difference Vegetation*

Index (NDVI), Enhanced Vegetation Index (EVI), and Normalized Difference Water Index (NDWI). Data were obtained from SRTM, CHIRPS, TerraClimate, SoilGrids, Landsat 8, and ESA WorldCover datasets. The Random Forest model was developed using 80% of the samples for training and 20% for testing. Model evaluation showed an overall accuracy of 80.12%, a Cohen's Kappa coefficient of 0.4212, and an ROC-AUC value of 0.8596, indicating good classification performance. Feature importance analysis revealed that slope was the most influential variable (18.43%), followed by NDWI (14.69%), elevation (12.24%), and NDVI (11.63%). Spatial analysis indicated that moderately suitable land dominated East Java, covering approximately 3.28 million ha (48.71%), followed by land of low suitability, covering 2.00 million ha (29.75%), and highly suitable land, covering 1.45 million ha (21.54%). The results demonstrate that integrating Random Forest with GIS provides an effective approach for assessing agricultural land suitability and can support sustainable agricultural planning and spatial decision-making in East Java Province.

Keywords: agricultural land suitability, East Java, Geographic Information System, Random Forest, remote sensing

Pendahuluan

Sektor pertanian memiliki peran strategis dalam mendukung ketahanan pangan dan pembangunan ekonomi nasional. Provinsi Jawa Timur merupakan salah satu sentra produksi tanaman pangan utama di Indonesia, dengan total luas panen padi mencapai lebih dari 1,7 juta hektare dan produksi sekitar 9,7 juta ton berdasarkan data BPS Jawa Timur (2023), sehingga menjadikannya salah satu lumbung pangan terbesar di tingkat nasional. Perbedaan kondisi biofisik seperti topografi, iklim, dan karakteristik tanah menyebabkan tingkat produktivitas dan potensi pengembangan pertanian tanaman pangan bervariasi antarwilayah, sehingga diperlukan informasi kesesuaian lahan yang komprehensif sebagai dasar perencanaan pemanfaatan lahan yang lebih efektif dan berkelanjutan (FAO, 1976; Rossiter, 2023).

Evaluasi kesesuaian lahan merupakan proses penilaian tingkat kecocokan suatu lahan untuk penggunaan tertentu berdasarkan karakteristik lingkungan dan kebutuhan penggunaan lahan. Perkembangan teknologi penginderaan jauh dan Sistem Informasi Geografis (SIG) memungkinkan evaluasi kesesuaian lahan dilakukan secara lebih cepat, efisien, dan mencakup wilayah yang luas melalui integrasi berbagai data biofisik seperti topografi, vegetasi, iklim, dan karakteristik tanah (Malczewski, 2004). Penelitian terkini menunjukkan bahwa pemanfaatan data geospasial dan penginderaan jauh mampu meningkatkan akurasi identifikasi potensi lahan serta mendukung pengambilan keputusan dalam perencanaan penggunaan lahan pertanian tanaman pangan secara berkelanjutan (AL-Taani et al., 2021; Nguyen et al., 2023).

Selain SIG, pendekatan machine learning semakin banyak diterapkan dalam penelitian sumber daya lahan. Random Forest merupakan salah satu algoritma yang banyak digunakan karena mampu menangani data multidimensi, mengurangi risiko overfitting, serta menghasilkan akurasi klasifikasi yang tinggi pada berbagai aplikasi penginderaan jauh dan pemetaan lingkungan (Breiman, 2001; Belgiu & Drăgut, 2016). Beberapa penelitian terbaru menunjukkan bahwa Random Forest memiliki kinerja yang baik dalam analisis kesesuaian lahan pertanian tanaman pangan karena mampu mengakomodasi hubungan nonlinier antarvariabel lingkungan serta menghasilkan interpretasi tingkat kepentingan variabel yang lebih baik dibandingkan metode klasifikasi konvensional (Giannarakis et al., 2022; Shevchenko et al., 2023; Li et al., 2024). Keunggulan ini menjadikan Random Forest sebagai pilihan yang relevan untuk analisis

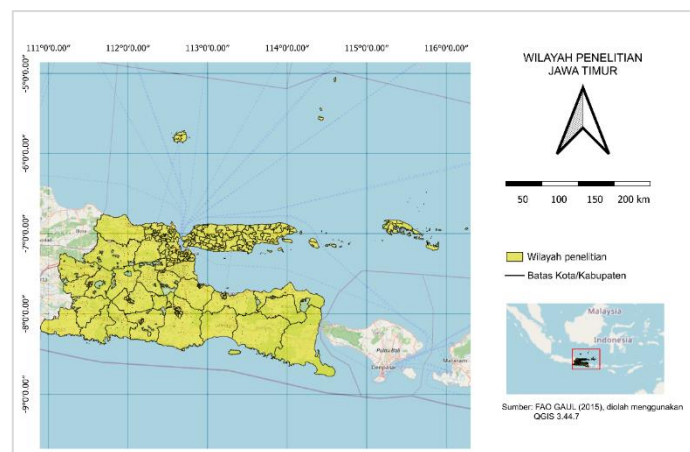
kesesuaian lahan pada skala regional yang melibatkan banyak variabel biofisik dengan pola hubungan yang kompleks.

Meskipun berbagai penelitian mengenai kesesuaian lahan telah dilakukan, kajian yang mengintegrasikan pendekatan Random Forest dan SIG untuk menganalisis kesesuaian lahan pertanian tanaman pangan pada skala Provinsi Jawa Timur masih relatif terbatas. Oleh karena itu, penelitian ini bertujuan untuk menganalisis kesesuaian lahan pertanian tanaman pangan di Provinsi Jawa Timur menggunakan pendekatan Random Forest dan SIG serta mengidentifikasi faktor-faktor lingkungan yang berpengaruh terhadap tingkat kesesuaian lahan. Hasil penelitian diharapkan dapat menjadi informasi pendukung dalam perencanaan dan pengelolaan sumber daya lahan pertanian tanaman pangan secara berkelanjutan di Jawa Timur.

Metodologi Penelitian

Lokasi Penelitian

Penelitian dilakukan di Provinsi Jawa Timur yang terletak pada kisaran 111°–115° BT dan 5°–9° LS. Wilayah penelitian mencakup seluruh kabupaten dan kota di Provinsi Jawa Timur beserta pulau-pulau kecil yang berada dalam administrasi provinsi tersebut. Berdasarkan topografinya, wilayah Jawa Timur terbagi menjadi dataran rendah (0–100 mdpl) meliputi pesisir utara dan selatan, dataran menengah (100–500 mdpl) yang mendominasi bagian tengah, serta dataran tinggi (>500 mdpl) meliputi kawasan pegunungan seperti Bromo-Tengger, Arjuno-Welirang, dan Ijen. Keberagaman topografi ini sangat berpengaruh terhadap distribusi dan potensi pengembangan lahan pertanian tanaman pangan di Jawa Timur



Gambar 1. Peta Lokasi Penelitian Provinsi Jawa Timur

Data dan Sumber Data

Data yang digunakan dalam penelitian ini merupakan data geospasial yang diperoleh dari platform Google Earth Engine, dengan rincian seperti pada Tabel 1. Seluruh data bersumber dari dataset global yang tersedia secara terbuka dan telah divalidasi untuk digunakan pada skala regional, sehingga mampu merepresentasikan kondisi biofisik seluruh wilayah Provinsi Jawa Timur secara spasial.

Tabel 1. Data yang digunakan dalam penelitian

Variabel	Sumber Data	Resolusi	Keterangan
Elevasi	SRTM	30 m	Ketinggian permukaan (mdpl)
Slope	Turunan SRTM	30 m	Kemiringan lereng (°)
Aspect	Turunan SRTM	30 m	Arah lereng (°)
NDVI	Landsat 8	30 m	Kondisi vegetasi (indeks)
EVI	Landsat 8	30 m	Kepadatan vegetasi (indeks)
NDWI	Landsat 8	30 m	Kelembapan lahan (indeks)
Rainfall	CHIRPS	5 m	Curah hujan (mm/tahun)
Temperature	TerraClimate	± 4 km	Suhu rata-rata tahunan (°C)
pH Tanah	SoilGrids	250 m	Tingkat keasaman tanah
SOC	SoilGrids	250 m	Karbon organik tanah (g/kg)
Referensi Kesesuaian	ESA WorldCover	10 m	Lahan pertanian sebagai data referensi

Variabel yang digunakan dipilih berdasarkan faktor-faktor yang secara umum memengaruhi kesesuaian lahan pertanian tanaman pangan, meliputi kondisi topografi, iklim, sifat tanah, dan karakteristik vegetasi (FAO, 1976; Sys et al., 1991). Parameter elevasi, *slope*, dan *aspect* digunakan untuk menggambarkan kondisi morfologi lahan. Curah hujan dan suhu digunakan sebagai indikator kondisi iklim. Data pH dan SOC bersumber dari SoilGrids, yaitu platform pemetaan tanah global berbasis machine learning dengan resolusi 250 m yang telah terbukti andal dalam merepresentasikan kualitas tanah pada skala regional (Hengl et al., 2017). Parameter vegetasi berupa NDVI, EVI, dan NDWI dihitung dari citra Landsat 8 untuk menggambarkan kondisi kehijauan vegetasi, produktivitas biomassa, dan kelembapan lahan (Huete et al., 2002), sedangkan data ESA WorldCover digunakan sebagai referensi dalam pembentukan sampel pelatihan dan pengujian model (Zanaga et al., 2021).

Tahapan Pengolahan Data

Penelitian ini dilakukan melalui beberapa tahapan yang sistematis. Pertama, seluruh data geospasial dipotong sesuai batas administrasi Provinsi Jawa Timur menggunakan Google Earth Engine dan diseragamkan ke dalam sistem koordinat WGS84 (EPSG:4326) sebelum dilakukan analisis. Kedua, citra Landsat 8 melalui proses masking awan dan komposit median untuk menghasilkan representasi kondisi permukaan yang lebih stabil selama periode pengamatan (2018–2022). Ketiga, dilakukan perhitungan indeks vegetasi NDVI, EVI, dan NDWI. Keempat, seluruh data topografi, iklim, tanah, dan indeks vegetasi digabungkan menjadi satu kumpulan variabel prediktor. Tahapan terakhir adalah pemodelan Random Forest, evaluasi kinerja model, dan pemetaan distribusi spasial kesesuaian lahan pertanian tanaman pangan di seluruh Provinsi Jawa Timur.

Pemodelan Random Forest

Analisis kesesuaian lahan dilakukan menggunakan algoritma *Random Forest* yang diperkenalkan oleh Breiman (2001). *Random Forest* merupakan metode klasifikasi berbasis *ensemble learning* yang bekerja dengan membangun sejumlah pohon keputusan secara bersamaan menggunakan subset data dan variabel yang dipilih secara acak, kemudian menggabungkan hasil prediksi masing-masing pohon melalui mekanisme voting untuk menghasilkan klasifikasi akhir. Metode ini banyak digunakan dalam penelitian penginderaan jauh dan pemetaan sumber daya lahan karena memiliki kemampuan menangani data multidimensi serta menghasilkan akurasi klasifikasi yang tinggi (Belgiu & Drăgut, 2016).

Data sampel dibagi menjadi data pelatihan sebesar 80% dan data pengujian sebesar 20% menggunakan metode stratified random sampling. Pembagian ini bertujuan agar model dapat belajar dari sebagian besar data (data pelatihan) dan kemudian diuji secara independen menggunakan data yang belum pernah digunakan sebelumnya (data pengujian), sehingga performa yang dilaporkan mencerminkan kemampuan generalisasi model yang sesungguhnya. Model dibangun menggunakan 500 pohon keputusan ($n_{\text{estimators}} = 500$) dengan parameter max_features menggunakan akar kuadrat dari jumlah variabel prediktor. Selain itu, dilakukan validasi silang (10-fold cross-validation) untuk mengevaluasi stabilitas model. Kinerja model dievaluasi menggunakan nilai Overall Accuracy, Precision, Recall, F1-score, Cohen's Kappa, dan Area Under the Curve (AUC) (Cohen, 1960; Powers, 2011). Ketiga kelas kesesuaian lahan yang digunakan dalam penelitian ini adalah: kelas kesesuaian tinggi (lahan dengan karakteristik biofisik yang optimal untuk pertanian tanaman pangan), kelas kesesuaian sedang (lahan dengan faktor pembatas yang masih dapat dikelola), dan kelas kesesuaian rendah (lahan dengan kendala biofisik yang signifikan sehingga membatasi potensi pertanian tanaman pangan).

Hasil dan Pembahasan

Karakteristik Variabel Lingkungan

Karakteristik biofisik wilayah Jawa Timur menunjukkan variasi yang cukup tinggi pada hampir seluruh variabel yang digunakan dalam penelitian ini. Variasi tersebut mengindikasikan adanya heterogenitas kondisi lingkungan yang berpotensi memengaruhi tingkat kesesuaian lahan pertanian untuk tanaman pangan. Karakteristik biofisik wilayah Jawa Timur disajikan pada Tabel 2.

Tabel 2. Statistik deskriptif variabel penelitian

Variabel	Mean	Std. Dev	Minimum	Maximum
Aspect (°)	163,38	105,93	0,00	358,89
EVI	0,39	0,13	-0,16	1,07
Elevation (m)	300,80	415,71	-6,00	3.649,00
NDVI	0,62	0,17	-0,83	0,95
NDWI	-0,61	0,13	-0,89	0,88
Rainfall (mm)	2.216,62	363,45	1.410,39	4.492,30

SOC (%)	5,21	2,74	2,00	64,00
Slope (°)	7,57	8,64	0,00	61,36
Temperature (°C)	24,81	2,67	6,70	27,60
pH	5,60	0,32	4,40	7,10

Elevasi wilayah berkisar antara -6 hingga 3.649 mdpl dengan nilai rata-rata sebesar 300,80 mdpl. Berdasarkan klasifikasi topografi, wilayah ini dapat dikategorikan menjadi: dataran rendah (0–100 mdpl) yang meliputi kawasan pesisir dan lahan sawah irigasi di wilayah utara; dataran menengah (100–500 mdpl) yang merupakan wilayah transisi dengan potensi lahan kering produktif; serta dataran tinggi (>500 mdpl) yang mencakup kawasan pegunungan dengan keterbatasan untuk pertanian tanaman pangan intensif. Keberagaman topografi ini berdampak langsung terhadap suhu, drainase, tingkat erosi, dan kemampuan pengelolaan lahan pertanian tanaman pangan di Jawa Timur. Secara umum, wilayah dataran rendah hingga menengah memiliki potensi yang lebih tinggi untuk pengembangan tanaman pangan seperti padi, jagung, dan kedelai (Sys et al., 1991).

Variabel iklim juga menunjukkan variasi yang cukup besar. Curah hujan tahunan berkisar antara 1.410–4.492 mm dengan rata-rata 2.216 mm, mencerminkan perbedaan distribusi hujan yang signifikan antara wilayah utara dan selatan Jawa Timur. Curah hujan tertinggi umumnya terdapat di wilayah pegunungan bagian selatan seperti kawasan Banyuwangi dan Malang selatan, sedangkan wilayah pantai utara cenderung lebih kering. Dari perspektif budidaya tanaman pangan, curah hujan optimal untuk padi sawah berkisar antara 1.500–2.500 mm/tahun, sehingga sebagian besar wilayah Jawa Timur berada dalam rentang yang mendukung produksi padi. Namun, wilayah dengan curah hujan di atas 3.000 mm dapat meningkatkan risiko genangan dan serangan hama dan penyakit (FAO, 1976). Suhu rata-rata tahunan sebesar 24,81°C dengan variasi 6,70–27,60°C mencerminkan perbedaan kondisi iklim antara dataran rendah dan dataran tinggi yang berpengaruh terhadap pemilihan jenis komoditas tanaman pangan yang sesuai.

Indeks vegetasi NDVI dan EVI menunjukkan nilai rata-rata masing-masing 0,62 dan 0,39, yang mengindikasikan tutupan vegetasi yang relatif baik di sebagian besar wilayah penelitian. Nilai NDWI yang cenderung negatif (rata-rata -0,61) menunjukkan bahwa sebagian besar wilayah berada dalam kondisi ketersediaan air permukaan yang terbatas, yang menjadi faktor pembatas penting bagi pertanian tanaman pangan. Nilai pH rata-rata 5,60 (kisaran 4,40–7,10) menunjukkan dominasi tanah agak masam hingga netral, sedangkan SOC rata-rata 5,21 g/kg mengindikasikan variasi kandungan bahan organik antarwilayah. Berdasarkan keseluruhan variasi biofisik tersebut, perencanaan pengembangan lahan pertanian tanaman pangan di Jawa Timur sebaiknya diprioritaskan pada wilayah dataran rendah hingga menengah dengan curah hujan 1.500–2.500 mm/tahun, pH 5,5–7,0, dan nilai NDWI yang lebih tinggi sebagai zona potensial untuk intensifikasi pertanian tanaman pangan.

Secara keseluruhan, hasil statistik deskriptif menunjukkan bahwa Provinsi Jawa Timur memiliki karakteristik lingkungan yang beragam baik dari aspek topografi, iklim, vegetasi, maupun tanah. Keragaman tersebut menjadi dasar penting dalam pembangunan model kesesuaian lahan karena setiap variabel memiliki kontribusi yang berbeda terhadap potensi pengembangan lahan pertanian.

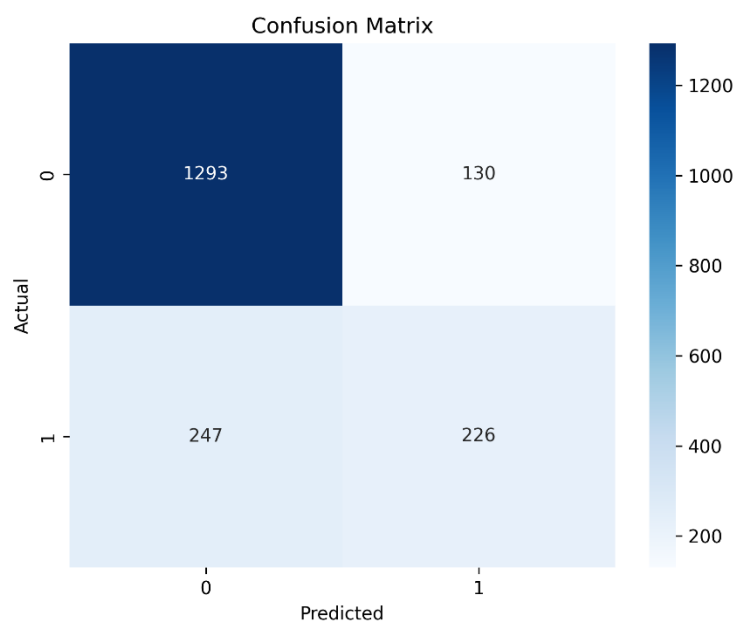
Evaluasi Kinerja Model Random Forest

Kinerja model Random Forest dievaluasi menggunakan beberapa metrik statistik untuk mengetahui kemampuan model dalam mengklasifikasikan kesesuaian lahan pertanian tanaman pangan di Provinsi Jawa Timur. Hasil evaluasi model disajikan pada Tabel 3.

Tabel 3. Hasil evaluasi model *Random Forest*

Metrik Evaluasi	Nilai
Accuracy	0,8012
Precision	0,6348
Recall	0,4778
F1-score	0,5452
Kappa	0,4212
ROC-AUC	0,8596

Berdasarkan Tabel 3, model Random Forest menghasilkan nilai akurasi sebesar 80,12%, menunjukkan bahwa sebagian besar sampel pengujian berhasil diklasifikasikan dengan benar. Nilai ROC-AUC sebesar 0,8596 mengindikasikan kemampuan model yang sangat baik dalam membedakan kelas kesesuaian lahan pertanian tanaman pangan. Sementara itu, nilai Cohen's Kappa sebesar 0,4212 menunjukkan tingkat kesepakatan klasifikasi pada kategori moderat, yang mengindikasikan performa model berada di atas peluang acak namun masih memiliki ruang untuk peningkatan, khususnya dalam membedakan kelas-kelas kesesuaian yang memiliki karakteristik biofisik yang saling tumpang tindih.



Gambar 2. *Confusion Matrix Random Forest*

Nilai precision sebesar 63,48% menunjukkan bahwa sebagian besar lokasi yang diprediksi sesuai memang merupakan lahan pertanian tanaman pangan yang sesuai. Nilai recall sebesar 47,78% mengindikasikan bahwa masih terdapat sejumlah lokasi yang sesuai yang belum berhasil teridentifikasi oleh model, sehingga nilai F1-score berada pada tingkat sedang, yaitu 54,52%. Berdasarkan confusion matrix pada Gambar 2, sebanyak 1.293 sampel lahan tidak sesuai berhasil diidentifikasi dengan benar (true negative), 226 sampel lahan sesuai berhasil diprediksi dengan benar (true positive), 130 sampel false positive, dan 247 sampel false negative.

Jumlah false negative yang lebih tinggi dibandingkan dengan false positive menunjukkan bahwa model cenderung lebih konservatif dalam mengidentifikasi lahan yang sesuai untuk pertanian tanaman pangan. Dalam konteks perencanaan wilayah pertanian, sifat konservatif ini memiliki implikasi praktis yang positif: pemerintah dan petani dapat terhindar dari keputusan investasi pada lahan yang sebenarnya marginal atau kurang produktif, sehingga meminimalkan risiko kegagalan pengembangan lahan pertanian tanaman pangan. Nilai spesifisitas sebesar 90,86% memperkuat hal ini, menunjukkan bahwa model sangat andal dalam mengenali lahan yang tidak sesuai dan dapat digunakan sebagai instrumen seleksi awal yang efektif dalam program perluasan atau intensifikasi pertanian tanaman pangan di Jawa Timur.

Pengaruh Variabel terhadap Kesesuaian Lahan

Untuk mengetahui faktor lingkungan yang paling berpengaruh terhadap kesesuaian lahan pertanian, dilakukan analisis *feature importance* pada model Random Forest. Hasil analisis disajikan pada Tabel 4.

Tabel 4. Tingkat kepentingan variabel pada model *Random Forest*

Variabel	Importance (%)
Slope	18,43
NDWI	14,69
Elevation	12,24
NDVI	11,63
EVI	9,45
Rainfall	9,37
Temperature	8,12
Aspect	6,82
pH	5,20
SOC	4,06

Hasil analisis menunjukkan bahwa slope merupakan variabel yang paling berpengaruh terhadap kesesuaian lahan pertanian tanaman pangan dengan kontribusi sebesar 18,43%. Hal ini menunjukkan bahwa kondisi topografi masih menjadi faktor utama penentu kesesuaian lahan di Provinsi Jawa Timur. Kemiringan lereng berpengaruh terhadap risiko erosi, kemampuan pengelolaan lahan, drainase, serta efisiensi penggunaan alat dan mesin pertanian; semakin curam

suatu wilayah, semakin besar potensi kehilangan tanah dan degradasi lahan yang dapat menurunkan produktivitas tanaman pangan (Li et al., 2024). Hasil ini selaras dengan berbagai penelitian kesesuaian lahan serupa yang juga menempatkan topografi sebagai faktor dominan (Giannarakis et al., 2022; Shevchenko et al., 2023).

Variabel penting berikutnya adalah NDWI dengan kontribusi 14,69%, diikuti oleh elevasi 12,24% dan NDVI 11,63%. Tingginya kontribusi NDWI menunjukkan bahwa kondisi kelembapan lahan berperan penting dalam menentukan kesesuaian lahan pertanian untuk tanaman pangan. Ketersediaan air merupakan salah satu faktor pembatas utama dalam sistem produksi tanaman pangan—terutama padi sawah—sehingga wilayah dengan kondisi kelembapan yang lebih baik cenderung memiliki tingkat kesesuaian yang lebih tinggi (Gao, 1996). Besarnya kontribusi NDVI dan EVI menunjukkan bahwa kondisi vegetasi eksisting juga berperan penting dalam membedakan wilayah yang sesuai dan kurang sesuai untuk pertanian tanaman pangan, karena indeks vegetasi yang tinggi umumnya mencerminkan kondisi lahan yang lebih produktif dengan kapasitas untuk mendukung pertumbuhan tanaman yang lebih baik (Huete et al., 2002).

Sebaliknya, variabel pH dan SOC menunjukkan kontribusi yang relatif lebih rendah, masing-masing sebesar 5,20% dan 4,06%. Rendahnya kontribusi kedua variabel tanah ini di tingkat regional dapat dijelaskan oleh resolusi spasial data SoilGrids sebesar 250 m, sehingga variasi kimia tanah pada skala sub-piksel tidak terdeteksi secara optimal. Pada skala regional seperti Jawa Timur, data tanah dari sumber global cenderung lebih homogen secara spasial dibandingkan data satelit (NDVI, EVI, NDWI) yang bervariasi pada resolusi 30 m dari piksel ke piksel, sehingga variabel tanah memiliki daya diskriminatif yang lebih rendah dalam model Random Forest (Hengl et al., 2017). Meskipun demikian, faktor tanah tetap berperan penting dalam menentukan produktivitas lahan pertanian tanaman pangan pada tingkat tapak atau skala yang lebih detail, dan perlu dipertimbangkan dalam perencanaan pengembangan pertanian pada tingkat kabupaten atau lahan spesifik (FAO, 1976; Sys et al., 1991).

Distribusi Spasial Kesesuaian Lahan Pertanian

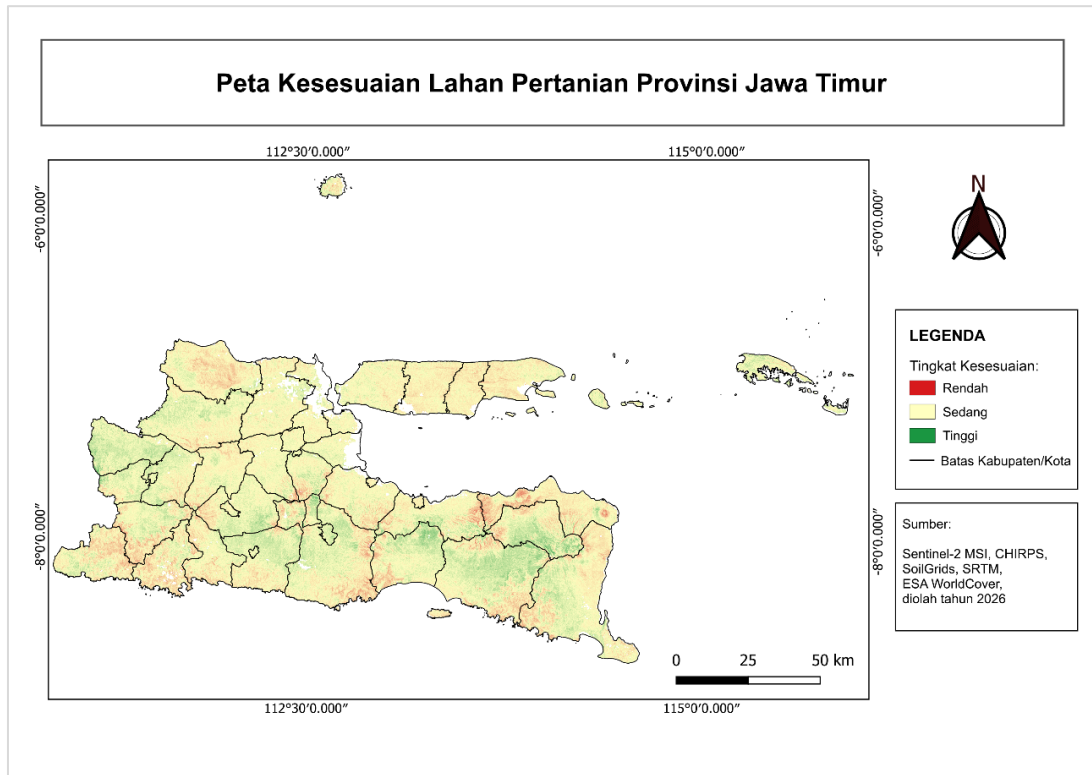
Hasil klasifikasi *Random Forest* selanjutnya digunakan untuk menghasilkan peta kesesuaian lahan pertanian di Provinsi Jawa Timur. Distribusi luas masing-masing kelas kesesuaian disajikan pada Tabel 5.

Tabel 5. Distribusi luas kelas kesesuaian lahan

Kelas	Luas (ha)	Persentase (%)
Rendah	2.002.079	29,75
Sedang	3.278.437	48,71
Tinggi	1.449.509	21,54

Berdasarkan hasil klasifikasi, kelas kesesuaian sedang mendominasi wilayah Jawa Timur dengan luas sekitar 3,28 juta ha atau 48,71% dari total wilayah penelitian. Dominasi kelas ini menunjukkan bahwa sebagian besar wilayah memiliki potensi yang cukup baik untuk mendukung kegiatan pertanian tanaman pangan, meskipun masih terdapat beberapa faktor pembatas seperti kemiringan lereng yang moderat, keterbatasan kelembapan, atau kondisi tanah yang memerlukan

perbaikan lebih lanjut. Kelas kesesuaian rendah mencakup sekitar 2,00 juta ha atau 29,75% dari total wilayah, umumnya berkaitan dengan kondisi topografi yang kurang mendukung, keterbatasan ketersediaan air, atau kombinasi beberapa faktor lingkungan yang kurang optimal bagi pertumbuhan tanaman pangan, termasuk keberadaan kawasan terbangun dan lahan non-pertanian.



Gambar 3. Peta Kesesuaian Lahan Pertanian Provinsi Jawa Timur

Sementara itu, kelas kesesuaian tinggi memiliki luas sekitar 1,45 juta ha atau 21,54% dari total wilayah penelitian. Secara spasial, wilayah dengan tingkat kesesuaian tinggi cenderung tersebar di bagian selatan dan tengah Jawa Timur yang memiliki kombinasi kondisi topografi datar hingga bergelombang ringan, nilai NDWI dan NDVI yang relatif tinggi, serta curah hujan yang memadai. Mengingat data geospasial yang digunakan bersumber dari platform global dengan cakupan spasial yang menyeluruh, hasil pemetaan ini secara umum dapat merepresentasikan kondisi biofisik Provinsi Jawa Timur untuk tujuan perencanaan pada skala regional, meskipun validasi lapangan tetap dianjurkan untuk verifikasi pada skala lokal. Berdasarkan distribusi spasial yang dihasilkan, pemerintah daerah disarankan untuk memprioritaskan pengembangan intensifikasi pertanian tanaman pangan pada zona kesesuaian tinggi, serta melakukan rehabilitasi lahan dan perbaikan sistem irigasi pada zona kesesuaian sedang agar potensi lahan pertanian tanaman pangan di Jawa Timur dapat dimanfaatkan secara optimal dan berkelanjutan.

Kesimpulan

Penelitian ini berhasil memetakan tingkat kesesuaian lahan pertanian tanaman pangan di Provinsi Jawa Timur menggunakan algoritma Random Forest berbasis data geospasial

multisumber. Model yang dikembangkan menunjukkan kinerja yang andal dengan tingkat akurasi sebesar 80,12%, nilai Kappa sebesar 0,4212 (kategori moderat), serta nilai ROC-AUC yang tinggi sebesar 0,8596. Analisis feature importance menunjukkan bahwa kemiringan lereng, NDWI, elevasi, dan NDVI merupakan faktor yang paling berpengaruh terhadap tingkat kesesuaian lahan pertanian untuk tanaman pangan. Hasil pemetaan menunjukkan bahwa kelas kesesuaian sedang mendominasi wilayah penelitian dengan luas sekitar 3,28 juta ha (48,71%), diikuti oleh kelas rendah sebesar 2,00 juta ha (29,75%) dan kelas tinggi sebesar 1,45 juta ha (21,54%). Pendekatan machine learning dan data geospasial multisumber ini terbukti efektif dalam mendukung evaluasi kesesuaian lahan pertanian tanaman pangan pada skala regional, serta berpotensi digunakan sebagai dasar pengambilan keputusan dalam perencanaan pembangunan pertanian berkelanjutan di Jawa Timur. Penelitian selanjutnya disarankan untuk mengintegrasikan data tanah beresolusi tinggi skala lokal dan melakukan validasi lapangan guna meningkatkan akurasi pemetaan kesesuaian lahan pada tingkat kabupaten/kota.

Daftar Pustaka

- Al-Taani, A., Al-Husban, Y., & Farhan, I. (2021). Land suitability evaluation for agriculture use using GIS and remote sensing techniques: The case study of Ma'an Governorate Jordan. *Egyptian Journal of Remote Sensing and Space Science*, 24(1), 109–117. <https://doi.org/10.1016/j.ejrs.2020.01.001>
- Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
- Badan Pusat Statistik Provinsi Jawa Timur. (2023). *Jawa Timur dalam angka 2023*. Badan Pusat Statistik Provinsi Jawa Timur.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46. <https://doi.org/10.1177/001316446002000104>
- Food and Agriculture Organization of the United Nations. (1976). *A framework for land evaluation*. FAO.
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., & Alsdorf, D. (2007). The shuttle radar topography mission. *Reviews of Geophysics*, 45(2), RG2004. <https://doi.org/10.1029/2005RG000183>
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Scientific Data*, 2, Article 150066. <https://doi.org/10.1038/sdata.2015.66>
- Gao, B. C. (1996). NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, 58(3), 257–266. [https://doi.org/10.1016/S0034-4257\(96\)00067-3](https://doi.org/10.1016/S0034-4257(96)00067-3)

- Giannarakis, G., Kalivas, D., & Zervas, E. (2022). Machine learning approaches for agricultural land suitability assessment using geospatial data. *Sustainability*, 14(9), Article 5438. <https://doi.org/10.3390/su14095438>
- Hengl, T., Mendes de Jesus, J., Heuvelink, G. B. M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., & Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLOS ONE*, 12(2), e0169748. <https://doi.org/10.1371/journal.pone.0169748>
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1–2), 195–213. [https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2)
- Li, X., Zhang, Y., Wang, H., & Liu, J. (2024). Spatial suitability evaluation based on multisource geospatial data and random forest modeling for sustainable land management. *Frontiers in Environmental Science*, 12, Article 1338931. <https://doi.org/10.3389/fenvs.2024.1338931>
- Malczewski, J. (2004). GIS-based land-use suitability analysis: A critical overview. *Progress in Planning*, 62(1), 3–65. <https://doi.org/10.1016/j.progress.2003.09.002>
- Nguyen, T. T., Nguyen, L. T., Nguyen, T. H., & Pham, T. D. (2023). Agricultural land suitability mapping using multi-criteria evaluation and GIS: A case study in the Mekong Delta, Vietnam. *Land*, 12(4), Article 812. <https://doi.org/10.3390/land12040812>
- Powers, D. M. W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1), 37–63.
- Rossiter, D. G. (2023). Updating and extending the land suitability evaluation framework for the 21st century. *Geoderma*, 429, Article 116248. <https://doi.org/10.1016/j.geoderma.2022.116248>
- Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *Proceedings of the Third Earth Resources Technology Satellite Symposium*, 1, 309–317.
- Shevchenko, O., Terekhov, A., & Kussul, N. (2023). Application of machine learning algorithms for agricultural land suitability mapping using remote sensing data. *Remote Sensing Applications: Society and Environment*, 31, Article 100996. <https://doi.org/10.1016/j.rsase.2023.100996>
- Sys, C., Van Ranst, E., & Debaveye, J. (1991). *Land evaluation. Part I: Principles in land evaluation and crop production calculations*. Agricultural Publications.
- Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Wevers, J., Cartus, O., Santoro, M., Fritz, S., & Herold, M. (2021). *ESA WorldCover 10m 2020 product validation report*. European Space Agency.