

Aplikasi Machine Learning pada GEE untuk Prediksi Produktivitas Tanaman Teh Menggunakan EVI dari Citra Sentinel-2A

SATRIA AJI PRADANA, DEWI KANIA SARI

Institut Teknologi Nasional Bandung

Email: aji64739@gmail.com

ABSTRAK

Teh merupakan komoditas perkebunan penting di Indonesia sehingga pemantauan produktivitasnya diperlukan untuk mendukung pengelolaan yang efisien. Penelitian ini menguji potensi citra Sentinel-2 dan algoritma Random Forest regression pada platform Google Earth Engine (GEE) untuk mengestimasi produktivitas teh di Perkebunan Kertamanah, Pangalengan. Indeks vegetasi Enhanced Vegetation Index (EVI) diekstraksi dari citra Sentinel-2 periode 2023–2025, kemudian digunakan sebagai variabel prediktor tunggal. Model dievaluasi dengan metrik RMSE, MAE, dan R^2 . Hasil menunjukkan akurasi rendah dengan nilai RMSE 436 kg/ha, MAE 268,173 kg/ha, dan R^2 hanya 9,2%. Temuan ini menegaskan bahwa EVI saja belum cukup untuk merepresentasikan variasi produktivitas teh, yang juga dipengaruhi faktor agronomis lain. Penelitian ini menekankan pentingnya integrasi variabel tambahan seperti umur tanaman dan kondisi lingkungan untuk meningkatkan kinerja model estimasi produktivitas pada perkebunan teh.

Kata Kunci: Produktivitas Teh, Google Earth Engine, Sentinel-2, Random Forest, EVI

1. PENDAHULUAN

Teh merupakan salah satu komoditas perkebunan utama di Indonesia yang berperan penting bagi perekonomian, ekspor, serta kesejahteraan masyarakat sekitar perkebunan. Produksi teh Indonesia menunjukkan tren positif, meningkat dari 97,6 ribu ton pada 2023 menjadi 103,5 ribu ton pada 2024 (Badan Pusat Statistik, 2025). Jawa Barat, khususnya Kabupaten Bandung, menjadi salah satu sentra utama penghasil teh dengan kontribusi signifikan terhadap produksi nasional. Pemantauan produktivitas secara akurat dan berkelanjutan diperlukan untuk mendukung pengelolaan kebun, perencanaan panen, serta peningkatan efisiensi produksi.

Pendekatan konvensional dalam penilaian produktivitas masih mengandalkan data panen manual yang bersifat terbatas, memerlukan waktu lama, dan tidak selalu mencerminkan kondisi aktual di lapangan. Perkembangan teknologi penginderaan jauh menawarkan alternatif yang lebih efisien melalui pemanfaatan citra satelit, di antaranya Sentinel-2, yang memiliki resolusi spasial menengah, cakupan luas, dan ketersediaan data bebas biaya (Kawamuna dkk.,

2017). Dengan dukungan *cloud computing* pada Google Earth Engine (GEE), analisis spasial dapat dilakukan secara lebih cepat dan terintegrasi. GEE adalah platform yang sangat efisien (Wang dkk., 2023) karena menyediakan akses mudah ke berbagai sumber data citra satelit, dilengkapi dengan algoritma pengolahan citra, pemetaan interaktif, dan visualisasi data.

Di sisi lain, algoritma *machine learning* sebagai cabang kecerdasan buatan, memungkinkan sistem untuk belajar mandiri dari data pelatihan atau pengalaman masa lalu tanpa pemrograman eksplisit (Primartha, 2018). *Machine learning* banyak digunakan dalam sektor pertanian untuk memodelkan hubungan kompleks antara indikator vegetasi dan hasil tanaman. Indeks *Enhanced Vegetation Index* (EVI) memiliki sensitivitas tinggi terhadap kondisi biomassa dan mampu mengurangi gangguan atmosfer dibandingkan NDVI (Huete dkk., 1997). Beberapa studi (Aslan dkk., 2024; Shabrina dkk., 2020) telah menunjukkan potensi EVI dalam estimasi hasil tanaman pangan, namun penerapannya pada komoditas teh di Indonesia masih sangat terbatas.

Berdasarkan hal tersebut, penelitian ini bertujuan menguji kemampuan algoritma Random Forest berbasis EVI dari citra Sentinel-2 dalam mengestimasi produktivitas teh di Perkebunan Kertamanah, Pangalengan. Studi ini diharapkan memberikan gambaran awal mengenai sejauh mana informasi satelit dapat dimanfaatkan untuk mendukung sistem monitoring produktivitas teh, sekaligus mengidentifikasi keterbatasan model berbasis indeks vegetasi tunggal.

2. METODOLOGI PENELITIAN

2.1 Data dan lokasi penelitian

Penelitian dilakukan di Perkebunan Teh Kertamanah, Kecamatan Pangalengan, Kabupaten Bandung, dengan luas ± 550 ha. Lokasi ini dipilih karena menjadi salah satu penghasil teh terbesar di Jawa Barat sekaligus representatif untuk uji coba pemodelan produktivitas berbasis citra satelit.

Data yang digunakan terdiri atas:

- 1) Citra Sentinel-2A Level-2A (periode 2023–2025) melalui Google Earth Engine;
- 2) Data panen teh per blok serta umur tanaman dari PTPN I Regional 2;
- 3) Batas administrasi dan batas kebun dari Ina-Geoportal dan PTPN I Regional 2.

2.2 Metodologi penelitian

Alur kerja metodologi penelitian digambarkan melalui diagram pada **Gambar 1**.



Gambar 1. Metodologi Penelitian

Tahapan utama penelitian dijelaskan di bawah ini.

- **Pengumpulan Data**

Pada tahapan ini dilakukan kompilasi data citra, data panen, dan batas kebun yang diperoleh dari beberapa sumber (Tabel 1).

- **Prapengolahan Data**

Pada tahap prapengolahan data dilakukan pembuatan dataset yang terbagi menjadi 2 buah dataset dalam bentuk file *Comma Separated Values* (CSV). Dataset tersebut adalah data latih dan data uji yang berisikan blok perkebunan yang panen setiap bulannya dalam kurun waktu 3 tahun, yaitu tahun 2023–2025. Setelah itu, dilakukan pengolahan pembagian data *shapfile* (SHP) per bulan pada *software* ArcGIS dengan menyesuaikan pada dataset yang telah dibuat. Setelah itu dilanjutkan pada prapengolahan dengan memanfaatkan platform GEE. Tahapan pertama pada platform tersebut adalah dengan melakukan proses *input* data citra sentinel-2A ke dalam platform, lalu melanjutkan dengan *composite* citra untuk memilih *band-band* yang akan digunakan dalam perhitungan indeks vegetasi, dan *masking cloud* dengan tutupan awan kurang dari 30%.

- **Pengolahan Data**

Pada tahap pengolahan data dilanjutkan pada platform Google Earth Engine. Pengolahan pertama dilakukan untuk menghitung nilai indeks vegetasi dengan menggunakan fungsi penentuan nilai rata-rata untuk setiap blok perkebunan. Setelah dilakukan proses transformasi indeks vegetasi, dilanjutkan mengekstrak nilai indeks vegetasi dengan cara dihitung rata-rata untuk setiap bloknya. Kemudian hasilnya akan diinput ke dalam dataset yang sebelumnya sudah dibuat. Data EVI dijadikan sebagai variabel tunggal untuk membuat model estimasi produktivitas. Pemodelan berbasis Random Forest regression dilakukan pada platform GEE.

- **Evaluasi Model**

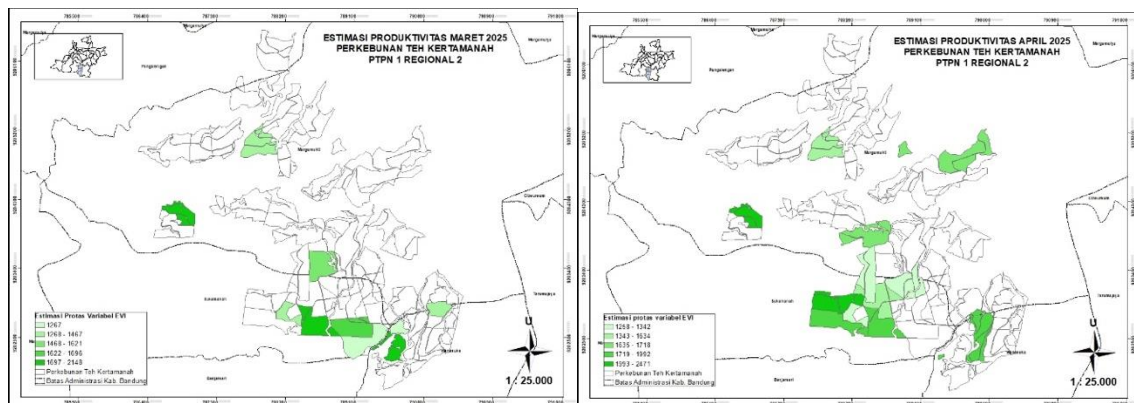
Untuk mengevaluasi model yang dihasilkan, dilakukan asesmen akurasi dengan menggunakan metrik RMSE, MAE, dan R^2 .

- **Analisis dan Visualisasi Hasil**

Tahapan ini merupakan proses akhir dari penelitian ini. Hasil pemodelan dianalisis berdasarkan kinerja yang dihasilkan dengan memperhatikan ketiga metrik akurasi. Untuk menyajikan hasil pemodelan dibuat visualisasi dalam bentuk peta yang dilakukan dengan bantuan SIG.

3. HASIL DAN PEMBAHASAN

Di bawah ini merupakan hasil pemodelan produktivitas teh berdasarkan indeks vegetasi EVI (**Gambar 2**). Pemodelan produktivitas teh berbasis indeks vegetasi EVI menghasilkan estimasi yang belum konsisten dengan data aktual. Beberapa blok menunjukkan *overestimate*, seperti Cibanteng I (Maret 2025) dan Ranca Gede (April 2025) dengan selisih masing-masing 1.046 kg/ha dan 1.004 kg/ha. Sebaliknya, sebagian blok lain menunjukkan *underestimate*, misalnya Cikijang (Maret 2025) dengan selisih 38,5 kg/ha. Hasil evaluasi menunjukkan RMSE 436 kg/ha, MAE 268,173 kg/ha, dan R^2 hanya 9,2%. Nilai R^2 yang rendah mengindikasikan bahwa sebagian besar variasi produktivitas tidak dapat dijelaskan oleh EVI saja.



Gambar 2. Hasil Pemodelan Produktivitas Teh

Hal ini sejalan dengan sifat EVI yang peka terhadap biomassa dan tutupan kanopi, namun kurang merepresentasikan faktor agronomis lain seperti umur tanaman, pemangkasan, dan kondisi iklim. Temuan ini sejalan dengan penelitian Shabrina dkk. (2020) dan Aslan dkk. (2024) yang menegaskan perlunya integrasi variabel tambahan agar model lebih reliabel. Studi terbaru pada perkebunan teh di China menunjukkan bahwa integrasi fitur spektral Sentinel-2 dengan UAV mampu menghasilkan estimasi hasil teh yang lebih akurat dibandingkan hanya mengandalkan satu indeks vegetasi (Wang dkk., 2024). Demikian pula, penelitian pada padi dan gandum menemukan bahwa variasi hasil lebih baik dijelaskan oleh kombinasi indeks

vegetasi multitemporal daripada satu indeks tunggal (Tiruneh dkk., 2023; Shabrina dkk., 2020).

4. KESIMPULAN

Penelitian ini menguji pemodelan produktivitas teh di Perkebunan Kertamanah menggunakan algoritma Random Forest regression berbasis EVI dari citra Sentinel-2 pada Google Earth Engine. Hasil menunjukkan akurasi rendah (RMSE 436 kg/ha, MAE 268,173 kg/ha, R^2 9,2%), dengan kecenderungan overestimasi pada blok bervegetasi rapat. Hal ini menegaskan bahwa EVI tunggal belum mampu merepresentasikan faktor kompleks yang memengaruhi produktivitas teh. Penelitian ini menekankan pentingnya penggunaan variabel tambahan, seperti umur tanaman, iklim, serta kombinasi indeks vegetasi, untuk meningkatkan performa model. Meskipun terbatas, studi ini memberikan gambaran awal potensi pemanfaatan citra satelit dan machine learning dalam pemantauan produktivitas teh di Indonesia.

UCAPAN TERIMA KASIH

Ucapan terima kasih penulis tujukan kepada PTPN 1 Regional 2, khususnya Perkebunan Teh Kertamanah, atas izin dan fasilitas yang diberikan untuk akses data lapangan.

DAFTAR RUJUKAN

- Aslan, M. F., Sabanci, K., & Aslan, B. (2024). Artificial intelligence techniques in crop yield estimation based on Sentinel-2 data: A comprehensive survey. *Sustainability*, 16(18), 8277. <https://doi.org/10.3390/su16188277>
- Badan Pusat Statistik. (2025). Statistik Indonesia 2025. Jakarta: BPS.
- Huete, A.R., Liu, H.Q., Batchily, K., dan Leeuwen, W. Van. (1997). A Comparison of Vegetation Indices over a Global Set of TM Images for EOS-MODIS. *Remote Sensing of Environment*, 4257.
- Huang, S., Tang, L., Hupy, J. P., Wang, Y., & Shao, G. (2021). A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *Journal of Forestry Research*, 32(1), 1–6. <https://doi.org/10.1007/s11676-020-01155-1>
- Wang, S., Yu, C., Ma, J., Ouyang, J., Zhao, Z., Xuan, Y., Fan, D., Yu, J., Wang, X., dan Zheng, X. (2024). Tea yield estimation using UAV images and deep learning, *Industrial Crops and Products*, 212, 118358, ISSN 0926-6690, <https://doi.org/10.1016/j.indcrop.2024.118358>.
- Kawamuna, A., Suprayogi, A., & Wijaya, A. P. (2017). Analisis kesehatan hutan mangrove berdasarkan metode klasifikasi NDVI pada citra Sentinel-2 (Studi kasus: Teluk Pangpang

- Kabupaten Banyuwangi). *Jurnal Geodesi Undip*, 6(1), 277–284.
- Primartha, R. (2018). *Belajar machine learning: Teori dan praktik*. Bandung: Informatika.
- Shabrina, N., Sukmono, A., & Subiyanto, S. (2020). Analisis identifikasi fase tumbuh padi untuk estimasi produksi padi dengan algoritma EVI dan NDRE multitemporal pada citra Sentinel-2 di Kabupaten Demak. *Jurnal Geodesi Undip*, 9(4), 59–70.
- Wang, Q., Moreno-Martínez, Á., Muñoz-Marí, J., Campos-Taberner, M., & Camps-Valls, G. (2023). Estimation of vegetation traits with kernel NDVI. *ISPRS Journal of Photogrammetry and Remote Sensing*, 195, 408–417. <https://doi.org/10.1016/j.isprsjprs.2022.12.018>
- Tiruneh, G. A., Meshesha, D. T., Adgo, E., Tsunekawa, A., Haregeweyn, N., Fenta, A. A., Alemayehu, T. Y., Mulualem, T., Fekadu, G., Demissie, S., & Reichert, J. M. (2023). Mapping crop yield spatial variability using Sentinel-2 vegetation indices in Ethiopia. *Arabian Journal of Geosciences*, 16(11).