

Pengembangan Model Prediksi Daya Surya Jangka Pendek Menggunakan Ensemble Random Forest dengan Rekayasa Fitur Temporal

Christio Revano Mege^{1*}, Ferizandi Qauzar Gani¹, Septia Eka Marsha Putra¹, Friska Hasugian²

¹ Fakultas Teknologi Industri, Teknik Fisika, Institut Teknologi Sumatera, Lampung Selatan, Indonesia

² Fakultas Teknologi Industri, Rekayasa Minyak dan Gas, Institut Teknologi Sumatera, Lampung Selatan, Indonesia

INFORMASI ARTIKEL

Diterima Redaksi: 26 Agustus 2026
Revisi Akhir: 24 September 2026
Diterbitkan Online: 28 September 2026

KATA KUNCI

Mikrogrid
Prediksi Daya
Random Forest
Rekayasa Fitur
Peramalan Jangka Pendek

KORESPONDENSI (*)

Phone: +62 813-8685-6037
E-mail: christio.mege@tf.itera.ac.id

A B S T R A K

Ketersediaan listrik di wilayah terpencil Indonesia masih menghadapi tantangan akibat ketergantungan pada generator diesel dan sifat intermiten energi surya. Penelitian ini bertujuan mengembangkan model prediksi daya panel surya jangka pendek berbasis Random Forest *multi-output* dengan memanfaatkan data historis tiga jam untuk meramalkan keluaran daya pada horizon 30 menit ke depan. Data diperoleh dari sistem panel surya 50 Wp yang dilengkapi sensor arus, tegangan, dan suhu, direkam setiap 10 detik selama periode 1 bulan. Setelah pra-pemrosesan dan *resampling* ke resolusi satu menit, dilakukan rekayasa fitur meliputi komponen waktu siklik, *lag features*, statistik bergerak, serta diferensiasi. *Dataset supervised* dibentuk melalui pendekatan *sliding window*, di mana model memanfaatkan data historis selama 180 menit sebagai *input* untuk meramalkan keluaran daya pada horizon 30 menit ke depan. Kemudian model dilatih menggunakan MultiOutputRegressor dengan 100 pohon keputusan. Hasil evaluasi menunjukkan bahwa model mampu mencapai nilai R^2 hingga 0,94 pada sampel terbaik, dengan *MAE* dan *RMSE* yang masih berada pada rentang dapat diterima untuk aplikasi mikrogrid skala laboratorium. Prediksi pada horizon menit ke-5 hingga ke-15 cenderung lebih akurat dibandingkan horizon ke-30. Model ini menawarkan keseimbangan antara akurasi, interpretabilitas, dan efisiensi komputasi, sehingga berpotensi mendukung pengelolaan energi proaktif pada sistem mikrogrid berbasis surya di wilayah terpencil. Pengembangan lebih lanjut disarankan melalui penambahan variabel meteorologi eksternal dan pendekatan hibrida untuk meningkatkan kinerja pada kondisi cuaca dinamis.

PENDAHULUAN

Penyediaan listrik di wilayah kepulauan dan daerah terpencil Indonesia masih menghadapi berbagai hambatan struktural. Biaya pembangunan jaringan transmisi yang tinggi serta kondisi geografis yang menantang membuat banyak komunitas mengandalkan generator diesel sebagai sumber utama. Ketergantungan ini tidak hanya meningkatkan beban operasional, tetapi juga menimbulkan emisi yang bertentangan dengan target transisi energi nasional [1], [2]. Dalam situasi tersebut, mikrogrid berbasis panel surya menjadi pilihan strategis karena bersifat modular, memanfaatkan sumber daya lokal, dan dapat dioperasikan secara mandiri maupun terhubung dengan jaringan utama [3], [4].

Namun, karakteristik intermiten energi surya menjadi kendala utama. Variasi radiasi matahari akibat tutupan awan, perubahan suhu, dan kondisi cuaca lokal menyebabkan output daya panel fotovoltaik (PV) fluktuatif dan sulit diperkirakan secara akurat [5], [6]. Ketidakpastian ini mempersulit pengelolaan penyimpanan energi, penjadwalan beban, serta penentuan kapasitas sistem yang optimal. Tanpa prediksi yang andal, operator cenderung melakukan *oversizing* komponen atau tetap mempertahankan cadangan berbasis fosil, sehingga menurunkan efisiensi ekonomi dan lingkungan

sistem [7], [8]. Dengan demikian, permasalahan makro berupa ketimpangan akses listrik di wilayah terpencil secara langsung menuntut solusi komputasi mikro berupa prediksi fluktuasi daya PV yang akurat dan responsif.

Pendekatan yang terbukti efektif untuk peramalan daya PV adalah metode *ensemble* berbasis pohon keputusan, khususnya Random Forest. Algoritma ini mampu menangani hubungan nonlinier antar fitur, *robust* terhadap *noise*, serta memberikan interpretasi melalui tingkat kepentingan fitur [9], [10]. Berbagai penelitian terkini menunjukkan bahwa Random Forest dan variannya termasuk hibrida dengan teknik residual atau *stacking* memberikan kinerja yang kompetitif untuk prediksi jangka pendek hingga menengah, terutama ketika data sensor lokal seperti arus, tegangan, dan suhu tersedia [11], [12], [13]. Keunggulan lain terletak pada kemudahan implementasi, kebutuhan komputasi yang relatif ringan, serta kemampuan menangani data berdimensi tinggi hasil rekayasa fitur seperti *lag*, *rolling statistics*, dan *encoding* waktu [14], [15].

Meskipun Random Forest telah banyak diaplikasikan pada peramalan energi terbarukan, integrasinya dengan sistem akuisisi data *real-time* berskala laboratorium di Indonesia masih terbatas. Sebagian besar studi sebelumnya menggunakan data meteorologi sekunder atau resolusi temporal yang kasar. Penelitian yang secara khusus memanfaatkan data pengukuran lapangan beresolusi tinggi dengan interval 10 detik untuk prediksi daya panel surya di Indonesia masih sangat jarang ditemukan. Sebagian besar studi sebelumnya cenderung menggunakan data meteorologi sekunder atau resolusi temporal yang lebih kasar, sehingga belum mengoptimalkan potensi data sensor lokal berfrekuensi tinggi melalui rekayasa fitur temporal dan model *ensemble* Random Forest. Kesenjangan inilah yang menjadi fokus penelitian ini, yaitu mengembangkan model prediksi daya panel surya berbasis Random Forest yang dilatih langsung dengan data aktual beresolusi tinggi guna mendukung pengelolaan energi mikrogrid yang lebih adaptif dan efisien. Oleh karena itu, pengembangan model prediksi daya panel surya menggunakan Random Forest yang dilatih dengan data aktual menjadi langkah penting untuk mendukung pengelolaan energi mikrogrid yang lebih adaptif dan efisien.

Penelitian ini bertujuan membangun dan mengevaluasi model prediksi daya panel surya berbasis Random Forest. Dengan memanfaatkan data pengukuran lapangan yang diolah melalui teknik rekayasa fitur dan *sliding window*, model diharapkan mampu menghasilkan prediksi akurat untuk horizon 30 menit ke depan berdasarkan data historis 3 jam sebelumnya. Hasil prediksi tersebut diharapkan dapat mendukung pengambilan keputusan operasional secara proaktif, mengurangi ketergantungan pada generator diesel, serta mendorong penerapan sistem energi terbarukan yang lebih andal di wilayah terpencil Indonesia.

TINJAUAN PUSTAKA

State of the art dari penelitian ini dirangkum dalam Tabel 1 berikut. Tabel tersebut membandingkan beberapa penelitian terdahulu yang menggunakan pendekatan Random Forest dan model terkait untuk prediksi daya atau iradiasi surya, mencakup kontribusi masing-masing studi beserta keterbatasannya. Melalui perbandingan ini, posisi kebaruan penelitian yang diusulkan yaitu pemanfaatan data lapangan beresolusi tinggi, rekayasa fitur temporal, serta kerangka *sliding window multi-output* untuk prediksi horizon 5, 10, 15, dan 30 menit dapat terlihat secara lebih jelas.

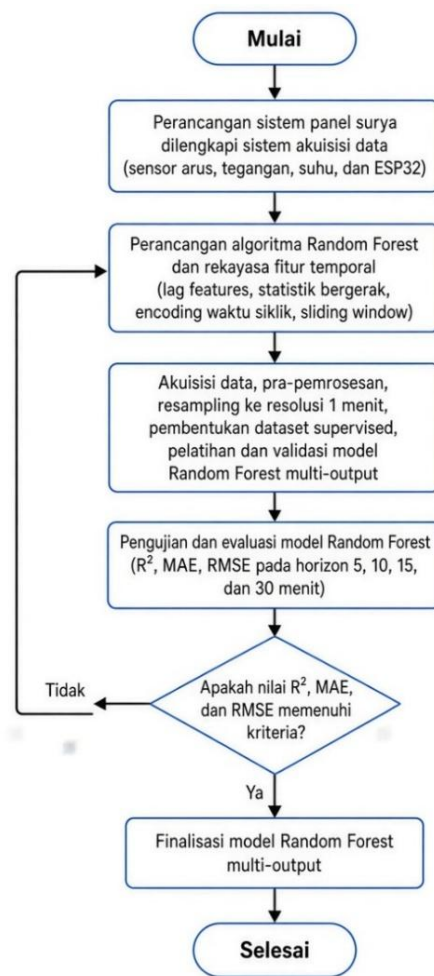
Tabel 1. Penelitian Terdahulu

No.	ID Artikel	Kontribusi
1	[16]	Menggunakan Random Forest untuk peramalan produksi daya surya dengan optimasi hyperparameter jumlah pohon dan kedalaman pohon. Model mencapai nilai R^2 yang tinggi. Namun kelemahan dari penelitian ini adalah data yang digunakan berasal dari sumber cuaca sekunder dan tidak memanfaatkan data sensor lapangan beresolusi tinggi. Selain itu, penelitian ini tidak menerapkan kerangka <i>sliding window multi-output</i> untuk prediksi beberapa horizon secara simultan.
2	[17]	Menerapkan Random Forest untuk prediksi multi-step jangka pendek yaitu horizon 15 dan 30 menit pada sistem PV 20 kW dengan mempertimbangkan variasi musiman. Hasil menunjukkan peningkatan akurasi hingga 50% dibanding model referensi. Kelemahan penelitian ini yaitu resolusi data masih relatif kasar dan belum mengeksplorasi pemanfaatan data pengukuran lapangan interval 10 detik yang diolah melalui rekayasa fitur temporal yang komprehensif.
3	[18]	Mengembangkan Random Forest yang dioptimasi menggunakan <i>Brown Bear Optimization Algorithm</i> untuk prediksi daya surya jangka pendek. Model berhasil

		menurunkan nilai kesalahan dibanding Decision Tree dan metode lain. Kelemahan penelitian ini adalah ukuran dataset terbatas dan belum mengintegrasikan pendekatan sliding window 180 menit serta prediksi multi-output pada beberapa horizon penting secara bersamaan.
4	[19]	Mengombinasikan Random Forest dengan Quantile Regression untuk peramalan iradiasi matahari jangka pendek, sehingga mampu memberikan informasi ketidakpastian prediksi. Kelemahan penelitian ini yaitu fokus utama masih pada prediksi iradiasi, bukan langsung pada daya keluaran panel surya. Selain itu, belum menggunakan data sensor lokal berfrekuensi tinggi dan kerangka <i>multi-output</i> .
5	[20]	Mengusulkan model hibrida XGBoost–LightGBM–Random Forest dengan optimasi bobot menggunakan algoritma L-BFGS-B. Model mencapai akurasi tinggi pada data pembangkit skala besar. Kelemahannya yaitu tidak memanfaatkan data lapangan beresolusi 10 detik dan tidak merancang prediksi berbasis sliding window multi-horizon seperti yang dilakukan pada penelitian ini.
6	Penelitian ini	Mengembangkan model Random Forest multi-output dengan rekayasa fitur temporal yang lengkap seperti <i>lag</i> , <i>rolling statistics</i> , <i>encoding</i> waktu siklik, dan diferensiasi serta pendekatan <i>sliding window</i> 180 menit sebagai <i>input</i> untuk memprediksi daya panel surya pada horizon 5, 10, 15, dan 30 menit. Data yang digunakan merupakan pengukuran lapangan beresolusi tinggi pada interval 10 detik dari sistem panel surya 50 Wp berskala laboratorium.

METODOLOGI

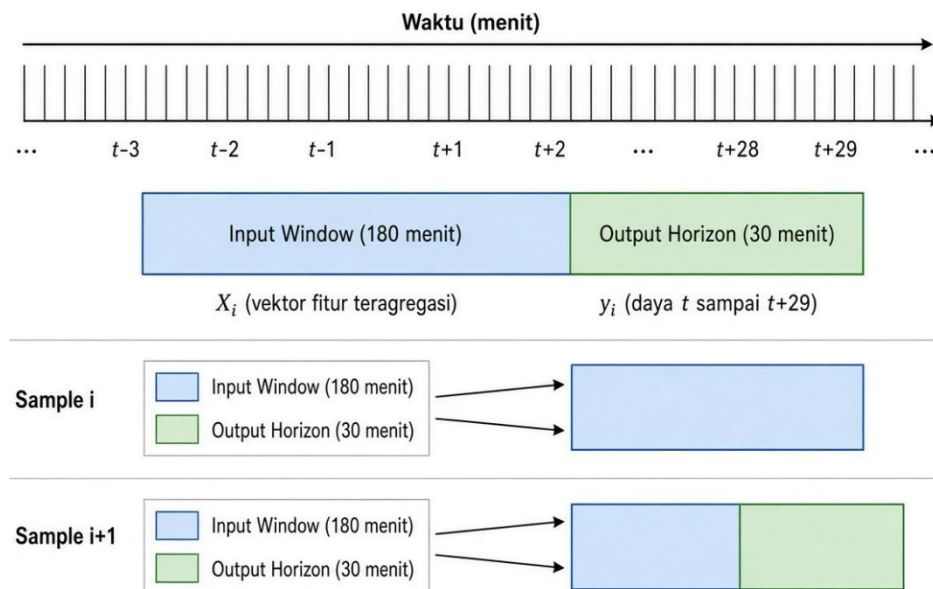
Penelitian ini dilaksanakan melalui pendekatan eksperimental-komputasional yang mengintegrasikan akuisisi data lapangan dengan pengembangan model prediksi berbasis *machine learning*. Data dikumpulkan dari sistem panel surya berkapasitas 50 Wp yang dilengkapi sensor arus INA260, sensor tegangan, sensor suhu DHT-22, dan mikrokontroler ESP32. Pengukuran dilakukan secara kontinu selama periode sekitar satu bulan, dimulai pada tanggal 6 Oktober 2025 hingga 6 November 2025 dengan interval 10 detik sehingga menghasilkan rekaman *timestamp*, arus panel, tegangan panel, daya, dan suhu udara. Data mentah kemudian dibersihkan dari nilai yang hilang dan diurutkan berdasarkan waktu. Selanjutnya, data disampel ulang menjadi resolusi satu menit menggunakan agregasi rata-rata untuk mengurangi noise sekaligus mempertahankan pola temporal yang relevan bagi prediksi jangka pendek. Diagram alir penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Diagram Alir Penelitian

Setelah tahap pra-pemrosesan, dilakukan rekayasa fitur guna memperkaya informasi yang dapat dipelajari oleh model. Fitur waktu yang diekstrak meliputi jam, menit, hari dalam seminggu, serta representasi siklik berupa *sine* dan *cosine* dari jam untuk menangkap periodisitas harian. Selain itu, dibuat sejumlah fitur *lag* dari daya, arus, dan suhu pada interval 1, 2, 3, 5, 10, 15, 30, dan 60 menit sebelumnya. Statistik bergerak berupa rata-rata, standar deviasi, nilai maksimum, dan nilai minimum dihitung dalam jendela 5, 15, 30, dan 60 menit. Fitur diferensiasi orde pertama juga ditambahkan untuk merepresentasikan perubahan jangka pendek pada daya dan suhu.

Dataset supervised dibentuk menggunakan pendekatan *sliding window*. Setiap sampel input dibangun dari data historis selama 180 menit terakhir atau tiga jam, sedangkan target prediksi berupa nilai daya pada 30 menit berikutnya. Dari setiap jendela historis, fitur diagregasi menjadi satu vektor yang terdiri atas nilai terakhir, rata-rata, standar deviasi, nilai maksimum, dan nilai minimum dari seluruh fitur yang tersedia. Proses pembentukan data ini diilustrasikan pada Gambar 2, di mana jendela input sepanjang 180 menit digeser secara bertahap untuk menghasilkan pasangan data *input-output*. Proses ini menghasilkan pasangan data *input-output* yang siap digunakan untuk pelatihan model *multi-output*.



Gambar 2 Proses Pembentukan Vektor *Input* (X) dan *Output* (y) berdasarkan *Sliding Window*

Model prediksi dikembangkan menggunakan algoritma Random Forest Regressor dalam kerangka MultiOutputRegressor sehingga mampu menghasilkan prediksi daya secara bersamaan. Parameter yang digunakan meliputi 100 pohon keputusan, kedalaman maksimum 12, minimum sampel untuk pemisahan sebanyak 5, dan minimum sampel pada daun sebanyak 3. Pelatihan dilakukan dengan memanfaatkan pemrosesan paralel. Pembagian data dilakukan secara kronologis dengan proporsi 80 persen untuk pelatihan dan 20 persen untuk pengujian guna menjaga urutan temporal dan menghindari kebocoran informasi.

Kinerja model dievaluasi menggunakan koefisien determinasi R^2 , *Mean Absolute Error* (MAE), *Root Mean Squared Error* ($RMSE$) pada keseluruhan horizon prediksi. Selain itu, dilakukan evaluasi secara terpisah pada menit ke-1, 5, 10, 15, dan 30 untuk menilai akurasi pada berbagai titik waktu dalam horizon 30 menit. Visualisasi perbandingan antara nilai aktual dan nilai prediksi juga dilakukan untuk menilai kemampuan model dalam mengikuti pola diurnal maupun fluktuasi jangka pendek daya panel surya.

HASIL DAN PEMBAHASAN

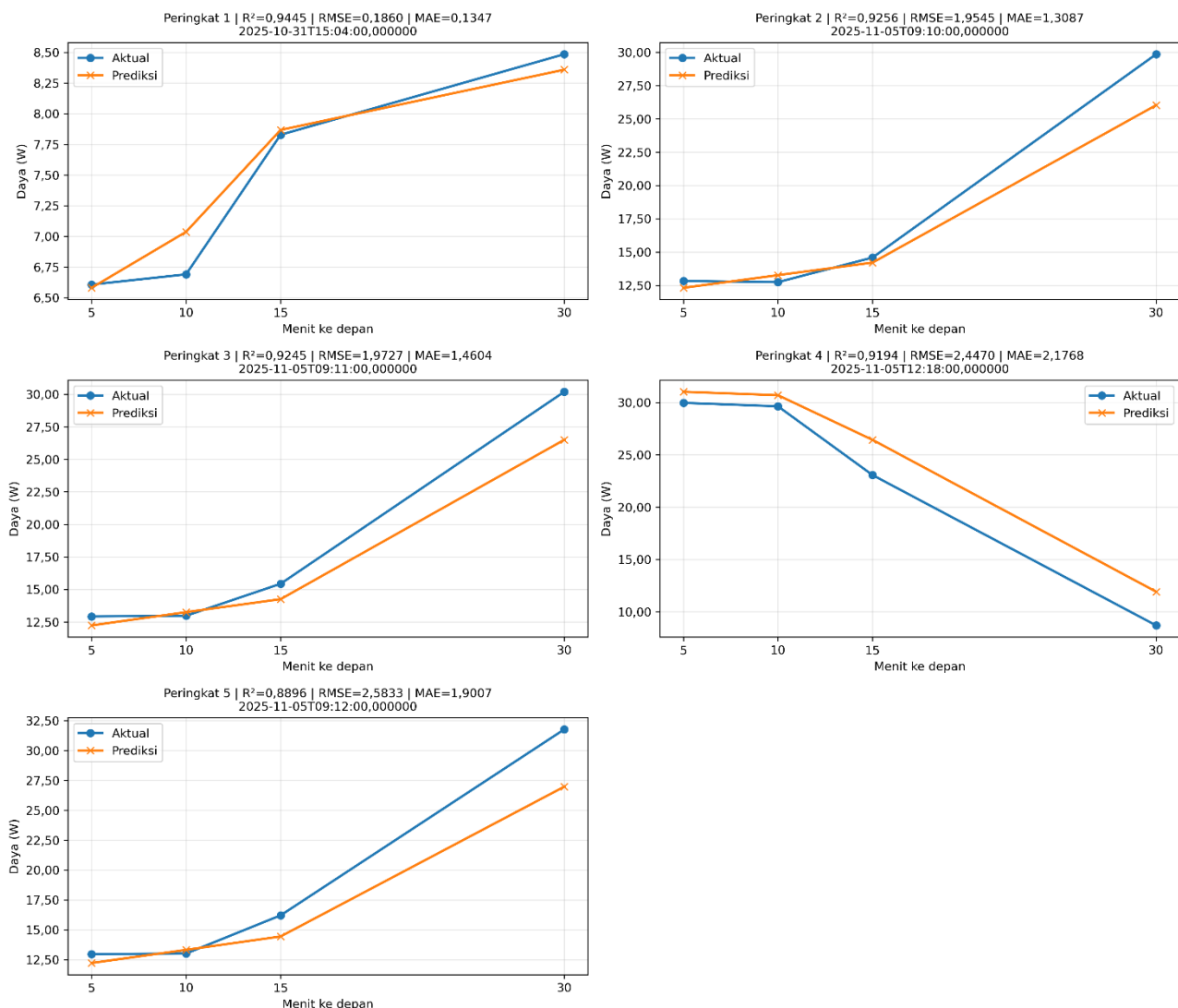
Hasil evaluasi kuantitatif seperti terlihat pada tabel 1 menunjukkan bahwa model mampu menangkap sebagian besar variasi daya aktual. Pada sampel uji terbaik, nilai koefisien determinasi mencapai 0,94. Beberapa sampel unggulan lainnya berada pada kisaran 0,89 hingga 0,92. Capaian ini mengindikasikan bahwa pada kondisi tertentu model dapat menjelaskan lebih dari 94% keragaman data daya yang diamati. Nilai MAE dan $RMSE$ secara keseluruhan berada pada rentang yang masih dapat diterima untuk keperluan prediksi jangka pendek pada sistem mikrogrid skala laboratorium. Evaluasi per horizon menunjukkan pola yang konsisten dengan karakteristik peramalan *multi-step*, di mana kesalahan prediksi cenderung meningkat seiring dengan semakin jauhnya menit yang diramalkan. Horizon menit kelima hingga kelima belas umumnya menunjukkan akurasi yang lebih tinggi dibandingkan horizon menit ketiga puluh.

Tabel 2. Hasil Evaluasi Sampel Terbaik

Peringkat Model	Indeks Data	R^2	$RMSE$	MAE
1	205	0,94	0,18	0,13
2	2856	0,92	1,95	1,30
3	2857	0,92	1,97	1,46
4	3044	0,92	2,44	2,17
5	2858	0,89	2,58	1,90

Visualisasi perbandingan antara nilai aktual dan nilai prediksi pada sampel dengan kinerja terbaik pada Gambar 1 memperlihatkan bahwa model mampu mengikuti pola umum kenaikan dan penurunan daya sepanjang hari. Pada periode perubahan yang relatif gradual, kurva prediksi berada sangat dekat dengan kurva aktual. Namun, pada titik-titik di mana terjadi fluktuasi tajam, umumnya terkait dengan pergerakan awan yang mengurangi radiasi matahari secara mendadak, terdapat sedikit penyimpangan berupa *lag* atau *underestimation*. Meskipun demikian, model tetap mampu mempertahankan arah tren secara keseluruhan. Analisis lebih lanjut terhadap distribusi kesalahan menunjukkan bahwa residual cenderung terpusat di sekitar nol pada sebagian besar sampel, dengan beberapa outlier yang muncul pada kondisi cuaca yang sangat dinamis.

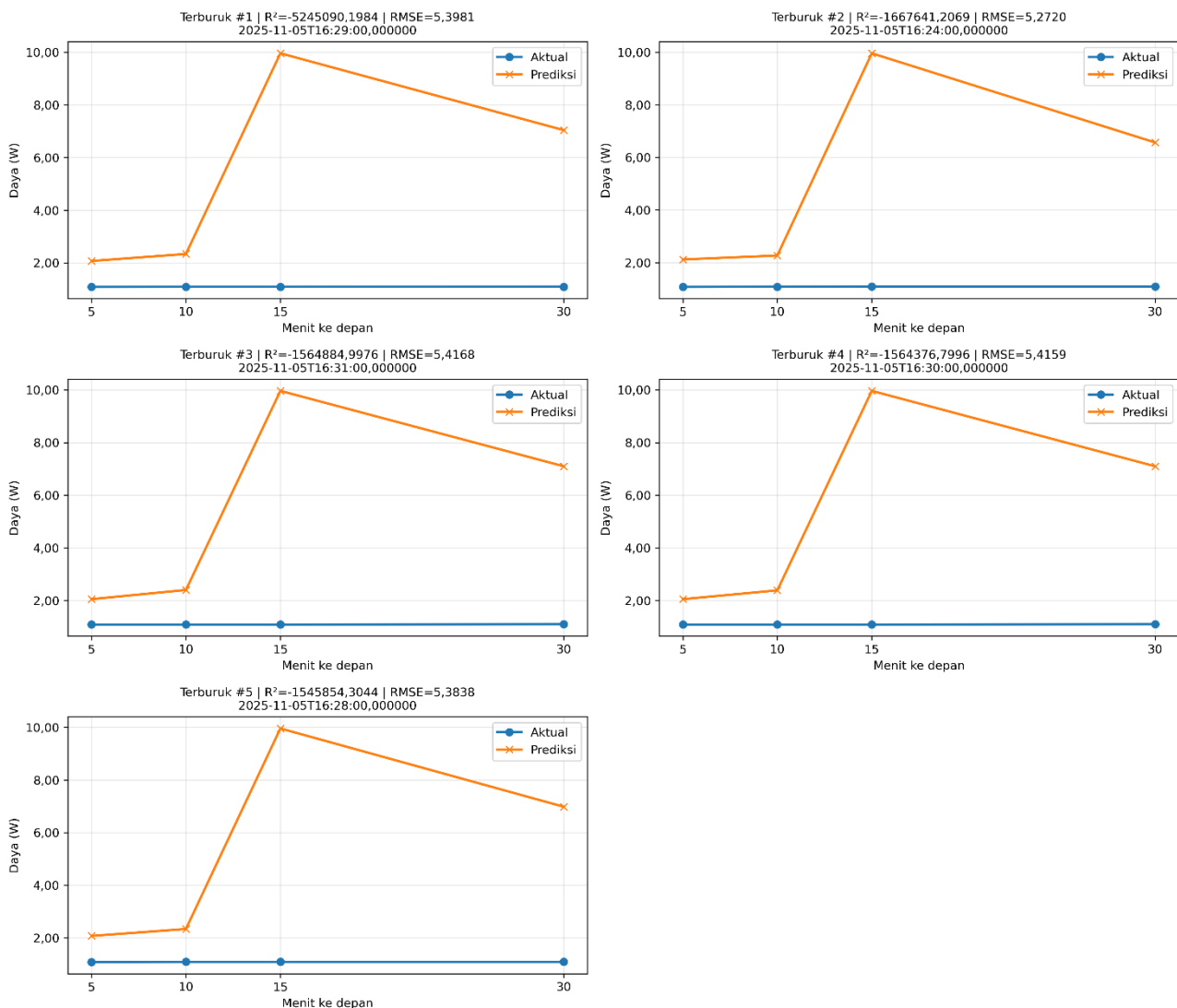
Evaluasi pada horizon-horizon penting secara terpisah, yaitu menit ke-5, 10, 15, dan 30 menunjukkan bahwa prediksi pada horizon yang lebih dekat memiliki kesalahan yang lebih rendah, sementara horizon tiga puluh menit masih mempertahankan tingkat akurasi yang memadai untuk mendukung pengambilan keputusan operasional. Secara keseluruhan, hasil penelitian ini menunjukkan bahwa pendekatan Random Forest dengan rekayasa fitur yang komprehensif mampu menghasilkan prediksi daya panel surya jangka pendek yang andal berdasarkan data sensor lokal.



Gambar 3. Grafik Daya Aktual Terhadap Daya Prediksi untuk 5 Sampel Terbaik

Meskipun model mampu mencapai R^2 hingga 0,94 pada beberapa sampel terbaik, evaluasi terhadap seluruh data uji sebanyak 3.616 sampel menghasilkan Mean R^2 sebesar -6729,34 dengan Mean RMSE 8,83 dan Mean MAE 7,56. Nilai R^2 yang sangat negatif ini menunjukkan bahwa secara keseluruhan model masih memiliki keterbatasan generalisasi, terutama pada periode sore dan malam hari ketika daya aktual rendah dan stabil, sementara prediksi model masih terpengaruh oleh pola diurnal dari jendela historis.

Gambar 4 menunjukkan sampel terburuk yang terjadi pada rentang waktu yang berdekatan, yaitu antara pukul 16:24 hingga 16:31 tanggal 5 November 2025. Pada seluruh sampel tersebut, nilai daya aktual berada pada kondisi hampir konstan dan sangat rendah sekitar 1,1–1,2 W. Sebaliknya, model Random Forest garis oranye memprediksi kenaikan daya yang tajam hingga mendekati 10 W pada horizon menit ke-15, kemudian sedikit menurun. Akibatnya, nilai R^2 menjadi sangat negatif dan RMSE relatif tinggi.



Gambar 4. Grafik Daya Aktual Terhadap Daya Prediksi untuk 5 Sampel Terburuk

Fenomena ini dapat dijelaskan bahwa pada sore hari menjelang matahari terbenam, produksi daya panel surya secara alami menurun drastis dan cenderung stabil di nilai rendah. Namun, jendela historis 180 menit yang digunakan sebagai input masih mengandung data dari periode siang hari di mana daya relatif tinggi. Selain itu, fitur waktu siklik dan statistik bergerak yang dibangun dari jendela tersebut masih mengingat pola kenaikan daya yang biasa terjadi pada pagi hingga siang hari. Akibatnya, model cenderung menggeneralisasi pola diurnal yang dipelajari selama pelatihan dan memprediksi kenaikan daya, padahal kondisi aktual sudah memasuki fase penurunan tajam atau tertutup awan tebal.

Kegagalan prediksi pada sampel-sampel ini merupakan ilustrasi klasik dari keterbatasan model berbasis pola historis dalam menghadapi perubahan rezim yang cepat dan jarang muncul, khususnya pada periode transisi sore hari. Temuan ini sejalan dengan keterbatasan yang telah diidentifikasi dalam literatur, di mana model *machine learning* sering kali kesulitan menangkap fluktuasi ekstrem atau perubahan mendadak yang tidak banyak terwakili dalam data pelatihan.

Temuan utama penelitian ini adalah kemampuan model Random Forest *multi-output* dalam menghasilkan prediksi daya panel surya yang cukup akurat untuk horizon tiga puluh menit berdasarkan data historis tiga jam. Capaian nilai R^2 hingga 0,94 pada sampel terbaik memberikan indikasi kuat bahwa model berhasil menangkap pola temporal utama yang terdapat

pada data pengukuran lapangan. Hasil ini sejalan dengan sejumlah penelitian sebelumnya yang menyoroti keunggulan algoritma *ensemble* berbasis pohon keputusan dalam peramalan energi terbarukan. Random Forest dikenal mampu menangani hubungan nonlinier antar variabel, *robust* terhadap *noise*, serta memberikan interpretasi melalui tingkat kepentingan fitur, sehingga cocok diterapkan pada data daya PV yang bersifat intermiten dan dipengaruhi banyak faktor [9], [10].

Keberhasilan model tidak dapat dilepaskan dari proses rekayasa fitur yang dilakukan secara sistematis. Penggunaan *lag features* pada berbagai interval memungkinkan model memanfaatkan informasi historis secara langsung. Statistik bergerak memberikan gambaran tentang tren lokal dan tingkat volatilitas dalam jendela waktu tertentu. *Encoding* waktu secara siklik melalui transformasi *sine* dan *cosine* berhasil merepresentasikan sifat periodik harian tanpa menciptakan diskontinuitas pada pergantian hari. Kombinasi fitur-fitur tersebut terbukti efektif dalam membedakan karakteristik daya pada pagi hari ketika radiasi masih rendah, siang hari ketika produksi mencapai puncak, dan sore hari ketika terjadi penurunan gradual. Temuan ini mendukung hasil studi terdahulu yang menekankan bahwa kualitas rekayasa fitur sering kali lebih menentukan kinerja model berbasis pohon dibandingkan dengan sekadar penyesuaian *hyperparameter* [11], [12].

Meskipun demikian, model masih menunjukkan keterbatasan yang perlu dibahas secara kritis. Pada kondisi fluktuasi ekstrem yang disebabkan oleh pergerakan awan, prediksi cenderung mengalami *lag* atau *underestimation*. Fenomena ini merupakan tantangan klasik dalam peramalan daya fotovoltaik dan telah dilaporkan secara luas dalam literatur. Model berbasis *machine learning*, termasuk Random Forest, pada dasarnya belajar dari pola historis. Ketika terjadi peristiwa yang jarang muncul dalam data pelatihan, seperti penurunan radiasi yang sangat tajam dalam waktu singkat, model kesulitan untuk merespons secara akurat [5], [6]. Penggunaan agregasi statistik dari jendela tiga jam, meskipun mengurangi dimensi data dan mencegah *overfitting*, berpotensi menghilangkan detail temporal beresolusi tinggi yang sebenarnya penting untuk menangkap perubahan cepat. *Trade-off* antara stabilitas model dan sensitivitas terhadap dinamika jangka pendek menjadi salah satu poin penting yang perlu dipertimbangkan dalam pengembangan selanjutnya. Perbandingan dengan pendekatan deep learning, khususnya Long Short-Term Memory (LSTM) yang mendominasi literatur prediksi daya surya dalam beberapa tahun terakhir, memberikan perspektif yang menarik. Sejumlah studi menunjukkan bahwa LSTM unggul dalam menangkap ketergantungan jangka panjang dan pola sekuensial yang kompleks [9], [13]. Namun, Random Forest menawarkan keunggulan praktis yang tidak dapat diabaikan. Proses pelatihan jauh lebih cepat, kebutuhan terhadap penyesuaian *hyperparameter* relatif lebih sederhana, dan model memiliki tingkat interpretabilitas yang lebih tinggi melalui *feature importance*. Dalam konteks sistem mikrogrid skala laboratorium atau penerapan di wilayah dengan keterbatasan sumber daya komputasi, keunggulan-keunggulan tersebut menjadikan Random Forest sebagai alternatif yang kompetitif. Beberapa penelitian terbaru bahkan menunjukkan bahwa pada dataset dengan ukuran sedang dan fitur yang telah direkayasa dengan baik, kinerja Random Forest dapat mendekati atau bahkan menyaingi arsitektur deep learning untuk horizon prediksi jangka pendek [10], [14].

Dari sudut pandang aplikasi, kemampuan model untuk memberikan prediksi tiga puluh menit ke depan memiliki implikasi praktis yang signifikan bagi pengelolaan mikrogrid berbasis energi surya. Prediksi jangka pendek ini dapat dimanfaatkan untuk mendukung berbagai keputusan operasional, antara lain penjadwalan pengisian dan pengosongan baterai secara lebih optimal, pengaturan prioritas beban, serta minimisasi penggunaan generator diesel atau impor daya dari jaringan utama. Dalam konteks wilayah kepulauan dan daerah terpencil di Indonesia, di mana ketergantungan pada bahan bakar fosil masih tinggi, peningkatan akurasi prediksi daya surya berkontribusi langsung pada pengurangan biaya operasional dan emisi. Temuan ini sejalan dengan penelitian-penelitian yang menekankan peran krusial peramalan dalam meningkatkan keandalan dan efisiensi sistem mikrogrid terisolasi [3], [7].

Analisis lebih lanjut terhadap distribusi kesalahan prediksi menunjukkan bahwa residual cenderung terpusat di sekitar nol pada sebagian besar sampel uji, dengan beberapa *outlier* yang muncul pada periode cuaca sangat dinamis. Pola ini mengindikasikan bahwa model tidak mengalami bias sistematis yang signifikan. Namun, keberadaan *outlier* pada kondisi ekstrem menegaskan perlunya pengembangan lebih lanjut, misalnya melalui teknik *residual modeling*, penambahan variabel eksogen seperti intensitas radiasi matahari atau data citra langit, atau penerapan pendekatan hibrida yang menggabungkan Random Forest dengan model statistik atau *deep learning*. Beberapa studi mutakhir telah mengeksplorasi hibrida residual di mana model statistik menangkap komponen linear dan tren, sementara Random Forest atau *gradient boosting* memodelkan residual nonlinier, dan hasilnya menunjukkan peningkatan akurasi yang konsisten [11], [12].

Keterbatasan lain yang perlu diakui adalah ruang lingkup data yang digunakan. Meskipun *dataset* mencakup periode sekitar satu bulan dengan resolusi tinggi, variasi musiman yang lebih panjang belum sepenuhnya terwakili. Perilaku daya panel surya di Indonesia dipengaruhi oleh pola musim hujan dan kemarau, serta variasi lokal yang cukup signifikan antar wilayah. Oleh karena itu, generalisasi model ke lokasi atau periode yang berbeda masih memerlukan validasi tambahan. Selain itu, penelitian ini belum memasukkan variabel meteorologi eksternal secara eksplisit, padahal intensitas radiasi, kelembaban, dan kecepatan angin diketahui memiliki korelasi kuat dengan keluaran daya fotovoltaik. Penambahan variabel-variabel tersebut pada penelitian lanjutan berpotensi meningkatkan kemampuan model dalam menghadapi kondisi cuaca yang sangat dinamis.

Secara keseluruhan, penelitian ini memberikan bukti empiris bahwa Random Forest dengan rekayasa fitur yang memadai merupakan pendekatan yang efektif dan praktis untuk prediksi daya panel surya jangka pendek. Model yang dihasilkan tidak hanya menunjukkan kinerja kuantitatif yang kompetitif, tetapi juga menawarkan keunggulan dari segi interpretabilitas dan efisiensi komputasi. Hasil penelitian ini memperkuat argumen bahwa metode *ensemble* berbasis pohon keputusan tetap relevan dan berharga di tengah dominasi *deep learning*, khususnya untuk aplikasi nyata pada sistem energi terbarukan skala kecil hingga menengah. Pengembangan lebih lanjut yang mengintegrasikan data multi-sumber, teknik hibrida, dan validasi lintas lokasi diharapkan dapat mendorong terciptanya sistem prediksi yang lebih *robust* dan siap diterapkan pada mikrogrid di wilayah terpencil Indonesia.

KESIMPULAN DAN SARAN

Penelitian ini berhasil mengembangkan model prediksi daya panel surya berbasis Random Forest *multi-output* yang memanfaatkan data historis selama tiga jam untuk meramalkan keluaran daya pada horizon 30 menit ke depan. Melalui tahapan pra-pemrosesan, rekayasa fitur yang komprehensif, serta pembentukan *dataset supervised learning*, model mampu menangkap pola temporal utama pada data pengukuran lapangan. Hasil evaluasi menunjukkan bahwa model mencapai nilai R^2 hingga 0,94 pada sampel terbaik, dengan kinerja yang konsisten pada horizon-horizon penting yaitu menit ke-5, 10, 15, dan 30. Temuan ini mengonfirmasi bahwa Random Forest merupakan pendekatan yang efektif dan praktis untuk prediksi daya surya jangka pendek, khususnya ketika didukung oleh rekayasa fitur yang memadai. Meskipun demikian, evaluasi terhadap seluruh data uji menghasilkan Mean R^2 yang sangat rendah. Model masih memiliki keterbatasan dalam menangani perubahan rezim yang cepat, sehingga diperlukan pengembangan lebih lanjut agar generalisasi model menjadi lebih *robust*.

Model yang dihasilkan menawarkan keseimbangan yang baik antara akurasi, interpretabilitas, dan efisiensi komputasi. Keunggulan tersebut menjadikannya relevan untuk diterapkan pada sistem mikrogrid berbasis energi terbarukan, di mana prediksi jangka pendek dibutuhkan untuk mendukung pengelolaan energi secara proaktif, seperti penjadwalan pengisian baterai dan pengaturan beban. Meskipun masih terdapat keterbatasan pada kondisi fluktuasi ekstrem akibat perubahan cuaca yang mendadak, secara keseluruhan model menunjukkan kemampuan yang memadai dalam mengikuti pola diurnal maupun dinamika jangka pendek daya panel surya.

Berdasarkan hasil penelitian, beberapa saran dapat diajukan untuk pengembangan lebih lanjut. Pertama, penambahan variabel eksogen seperti intensitas radiasi matahari, kelembaban udara, atau data citra langit berpotensi meningkatkan kemampuan model dalam menghadapi kondisi cuaca yang sangat dinamis. Kedua, eksplorasi pendekatan hibrida, misalnya *residual modeling* yang menggabungkan Random Forest dengan model statistik atau algoritma *boosting* lain, dapat dipertimbangkan untuk memperbaiki akurasi pada periode fluktuasi tajam. Ketiga, validasi model pada data lintas musim dan lintas lokasi perlu dilakukan agar generalisasi hasil penelitian menjadi lebih kuat. Keempat, penerapan teknik optimasi *hyperparameter* secara lebih sistematis serta evaluasi terhadap horizon prediksi yang lebih panjang dapat memperluas manfaat praktis model.

DAFTAR PUSTAKA

- [1] T. Mosetlhe, A. Yusuff, T. Ayodele, dan A. Ogunjuyigbe, "Sustainable rural electrification through micro-grids in developing nations — A review of recent development," *Energy Reports*, vol. 13, hlm. 1171–1177, Jun 2025, doi: 10.1016/J.EGYR.2024.11.040.

- [2] M. A. Shehu *dkk.*, "Control and energy management of standalone microgrids in remote areas: A review of recent advances, challenges, and opportunities for future research," *Engineering Science and Technology, an International Journal*, vol. 75, hlm. 102288, Mar 2026, doi: 10.1016/J.JESTCH.2026.102288.
- [3] W. M. Figueredo, O. R. Saavedra, P. B. Leite Neto, dan D. Q. Oliveira, "Design and operational challenges of renewable-powered isolated microgrids: a zero-carbon approach," *Scientific Reports 2025 15:1*, vol. 15, no. 1, hlm. 45296-, Nov 2025, doi: 10.1038/s41598-025-28882-x.
- [4] L. Vankudoth dan A. Q. H. Badar, "Development and analysis of scheduling strategies for utilizing shared energy storage system in networked microgrids," *J. Energy Storage*, vol. 97, hlm. 112691, Sep 2024, doi: 10.1016/J.EST.2024.112691.
- [5] M. M. Husein, E. J. Gago, B. Hasan, dan M. C. Pegalajar, "Towards energy efficiency: A comprehensive review of deep learning-based photovoltaic power forecasting strategies," *Heliyon*, vol. 10, no. 13, hlm. e33419, Jul 2024, doi: 10.1016/J.HELİYON.2024.E33419.
- [6] O. Khouili, M. Hanine, M. Louzazni, M. A. L. Flores, E. G. Villena, dan I. Ashraf, "Evaluating the impact of deep learning approaches on solar and photovoltaic power forecasting: A systematic review," *Energy Strategy Reviews*, vol. 59, hlm. 101735, Mei 2025, doi: 10.1016/J.ESR.2025.101735.
- [7] D. Hu, Y. Fan, dan W. Shao, "Optimal sizing and operation of microgrid considering renewable energy uncertainty based on scenario generation," *J. Energy Storage*, vol. 109, hlm. 115174, Feb 2025, doi: 10.1016/J.EST.2024.115174.
- [8] O. Motamedisedeh, S. Omrani, A. Karim, R. Drogemuller, dan G. Walker, "A comprehensive review of optimum integration of photovoltaic-based energy systems," *Renewable and Sustainable Energy Reviews*, vol. 207, hlm. 114935, Jan 2025, doi: 10.1016/J.RSER.2024.114935.
- [9] H. Z. Anonto *dkk.*, "Intelligent energy management of microgrids using machine learning: Leveraging random forest models for solar and wind power," *Results in Engineering*, vol. 27, hlm. 106539, Sep 2025, doi: 10.1016/J.RINENG.2025.106539.
- [10] M. U. Tariq, S. M. Saqib, T. Mazhar, M. A. Khan, T. Shahzad, dan H. Hamam, "PV forecasting-driven decision-support for solar microgrids with battery storage," *Appl. Energy*, vol. 412, hlm. 127685, Jun 2026, doi: 10.1016/J.APENERGY.2026.127685.
- [11] M. Abdelsattar, M. A. Azim, A. AbdelMoety, dan A. Emad-Eldeen, "Comparative analysis of deep learning architectures in solar power prediction," *Scientific Reports 2025 15:1*, vol. 15, no. 1, hlm. 31729-, Agu 2025, doi: 10.1038/s41598-025-14908-x.
- [12] D. Sharma, C. Kumar, dan M. Kumar, "A review of machine learning and deep learning models, features, and applications for solar PV forecasting," *Journal of Electrical Systems and Information Technology 2026 13:1*, vol. 13, no. 1, hlm. 74-, Jul 2026, doi: 10.1186/S43067-026-00382-6.
- [13] C. Rodriguez-Aburto *dkk.*, "Machine Learning for Photovoltaic Power Forecasting Integrated with Energy Storage Systems: A Scientometric Analysis, Systematic Review, and Meta-Analysis," *Energies 2025, Vol. 18, Page 6291*, vol. 18, no. 23, hlm. 6291, Nov 2025, doi: 10.3390/EN18236291.
- [14] F. Moazzen dan M. J. Hossain, "Multivariate Deep Learning Long Short-Term Memory-Based Forecasting for Microgrid Energy Management Systems," *Energies 2024, Vol. 17, Page 4360*, vol. 17, no. 17, hlm. 4360, Agu 2024, doi: 10.3390/EN17174360.
- [15] H. Xue, J. Ma, J. Zhang, P. Jin, J. Wu, dan F. Du, "Power Forecasting for Photovoltaic Microgrid Based on MultiScale CNN-LSTM Network Models," *Energies 2024, Vol. 17, Page 3877*, vol. 17, no. 16, hlm. 3877, Agu 2024, doi: 10.3390/EN17163877.
- [16] M. A. Azman, H. Jantan, U. F. M. Bahrin, dan E. A. Kadir, "Solar Power Production Forecasting Model Using Random Forest Algorithm," *Lecture Notes in Networks and Systems*, vol. 1050 LNNS, hlm. 135–144, 2024, doi: 10.1007/978-3-031-64847-2_12/SAVE-RESEARCH.
- [17] S. Jogunuri *dkk.*, "Random forest machine learning algorithm based seasonal multi-step ahead short-term solar photovoltaic power output forecasting," *IET Renewable Power Generation*, vol. 19, no. 1, hlm. e12921, Jan 2025, doi: 10.1049/RPG2.12921;PAGE:STRING:ARTICLE/CHAPTER.
- [18] R. S. Kumar, P. S. Meera, V. Lavanya, dan S. Hemamalini, "Brown bear optimized random forest model for short term solar power forecasting," *Results in Engineering*, vol. 25, hlm. 104583, Mar 2025, doi: 10.1016/J.RINENG.2025.104583.
- [19] A. Masache, P. Mdlongwa, D. Maposa, dan C. Sigauke, "Short-term forecasting of solar irradiance using decision tree-based models and non-parametric quantile regression," *PLoS One*, vol. 19, no. 12, hlm. e0312814, Des 2024, doi: 10.1371/JOURNAL.PONE.0312814.

- [20] T. A. Nguyen *dkk.*, “A Study on Novel Solar Power Forecasting using an XGB-LiGBM-RF Hybrid Model and the L-BFGS-B Optimization Algorithm,” *Engineering, Technology & Applied Science Research*, vol. 15, no. 4, hlm. 24516–24522, Agu 2025, doi: 10.48084/ETASR.11308.