

Proof of Concept Digital Twin Berbasis Machine Learning untuk Prediksi dan Optimasi Kinerja Panel Surya

Ade Zulkarnain Hasibuan^{1*}, Irwanda Syahputra²

^{1*,2} Program Studi Informatika Universitas Samudra, Langsa, Indonesia

^{1*}adezulhsb@unsam.ac.id, ²irwanda.syahputra@unsam.ac.id

^{*)}adezulhsb@unsam.ac.id

Abstrak–Penelitian ini mengembangkan dan menguji proof of concept Digital Twin berbasis Machine Learning pada panel surya fisik berdaya nominal 100 W. Data lingkungan dan kelistrikan diperoleh melalui sensor yang terhubung dengan ESP32, kemudian dikirim ke server untuk prediksi daya menggunakan Random Forest Regression. Exhaustive Scenario Search digunakan untuk membandingkan sudut 25°, 35°, dan 45°, sedangkan dashboard menyajikan daya terukur, daya prediksi, dan rekomendasi sudut. Analisis menggunakan 9.720 rekaman hasil pengukuran yang dibagi secara kronologis menjadi data pelatihan, validasi, dan pengujian. Validasi pada panel fisik dilakukan dengan membandingkan prediksi terhadap daya terukur. Pada 1.944 rekaman pengujian, model menghasilkan MAE 2,769 W, RMSE 4,145 W, dan R^2 0,9446. Evaluasi skenario sudut menghasilkan energi 2.492,97 Wh, lebih tinggi 3,06% dibandingkan 2.418,91 Wh pada sudut tetap 35°. Hasil ini mendukung penerapan proof of concept untuk pemantauan kondisi panel, prediksi daya, dan penyusunan rekomendasi sudut pada kondisi pengujian yang diteliti.

Kata Kunci: Digital Twin, Machine Learning, Fotovoltaik, Random Forest Regression, Optimasi Sudut

Abstract–This study developed and evaluated a machine learning–based Digital Twin proof of concept for a physical solar panel with a nominal capacity of 100 W. Environmental and electrical data were collected using sensors connected to an ESP32 and transmitted to a server for power prediction using Random Forest Regression. Exhaustive Scenario Search was used to compare panel angles of 25°, 35°, and 45°, while the dashboard presented measured power, predicted power, and angle recommendations. The analysis used 9,720 measurement records divided chronologically into training, validation, and testing datasets. Physical-panel validation was performed by comparing the predicted values with the measured power. For the 1,944 testing records, the model achieved an MAE of 2.769 W, an RMSE of 4.145 W, and an R^2 of 0.9446. The angle scenario evaluation produced 2,492.97 Wh of energy, which was 3.06% higher than the 2,418.91 Wh obtained with a fixed angle of 35°. These findings support the implementation of a Digital Twin proof of concept for panel-condition monitoring, power prediction, and angle recommendation under the tested conditions.

Keywords: Digital Twin, Machine Learning, Photovoltaic, Random Forest Regression, Angle Optimization

1. PENDAHULUAN

Pengembangan Digital Twin panel surya memerlukan lebih dari sekadar pencatatan sensor. Representasi digital perlu menerima data kondisi aset, memperkirakan daya, membandingkan alternatif operasi, dan menghasilkan rekomendasi yang dapat diuji pada panel fisik. Hubungan antara entitas fisik, model virtual, data, dan komunikasi merupakan unsur utama Digital Twin [1]. Penelitian ini mengembangkan hubungan tersebut melalui akuisisi data panel surya, prediksi daya berbasis Machine Learning, evaluasi skenario sudut, dan validasi hasilnya.

Tingkat otomatisasi aliran data membedakan Digital Model, Digital Shadow, dan Digital Twin[2], koneksi physical-to-virtual, virtual-to-physical, dan twinning rate merupakan karakteristik penting Digital Twin [3]. Perbedaan ini menjadi dasar untuk menjelaskan peran akuisisi sensor, model prediksi, evaluasi skenario, dashboard, dan umpan balik dalam penelitian. Komponen-komponen tersebut disusun untuk menghubungkan kondisi fisik panel dengan representasi digital dan layanan analitik.

Pada sistem fotovoltaik, Digital Twin dapat digunakan untuk pemantauan, prediksi, diagnosis, simulasi, dan optimasi operasi. Kajian terdahulu menunjukkan bahwa pendekatan berbasis data banyak digunakan, meskipun tingkat integrasi fisik-digital, pembaruan model, dan umpan baliknya masih beragam [4], [5]. Keluaran panel sendiri berubah mengikuti cahaya, suhu panel, suhu dan kelembapan lingkungan, serta sudut pemasangan. Data kondisi panel yang konsisten memungkinkan model virtual memperkirakan daya dan membandingkan skenario sudut sebelum memberikan rekomendasi.

Penggunaan machine learning sebagai model virtual telah diteliti pada sistem energi. Natgunanathan et al. mengevaluasi model kecerdasan buatan untuk memprediksi pembangkitan daya pada digital twin microgrid [6].

Pendekatan berbasis transformer pada digital twin fotovoltaik untuk meningkatkan kemampuan prediksi adaptif [7]. Penelitian ini berfokus pada integrasi model prediksi dan evaluasi skenario dalam arsitektur Digital Twin panel surya, dengan data lingkungan, sudut panel, dan daya aktual sebagai variabel penelitian. Kontribusi tersebut terletak pada integrasi sistem, bukan pada pengembangan algoritma machine learning baru.

Random Forest Regression dipilih karena dapat memodelkan hubungan nonlinier pada data tabular dan telah digunakan untuk prediksi daya fotovoltaik [8], [9]. Model berbasis pohon juga sesuai untuk data yang memuat variabel lingkungan dan waktu [10]. Formulasi regresi serta evaluasinya mengikuti prinsip supervised learning [11]. Dalam penelitian ini, Random Forest memprediksi daya berdasarkan kondisi lingkungan dan sudut panel. Exhaustive Scenario Search kemudian membandingkan prediksi pada sudut 25°, 35°, dan 45° untuk memilih nilai tertinggi. Dengan tiga kandidat diskrit, optimasi dilakukan melalui evaluasi seluruh alternatif yang tersedia. Pemisahan data berdasarkan waktu diperlukan untuk mengurangi bias akibat kebocoran informasi. Evaluasi data temporal harus memperhatikan ketergantungan antarobservasi dan batas waktu antara data pelatihan dan pengujian [12], [13]. Kebocoran juga dapat terjadi ketika fitur berhubungan langsung dengan target [14]. Oleh karena itu, data dibagi secara kronologis dan rekaman dengan kondisi yang sama ditempatkan dalam satu partisi. Tegangan dan arus tidak digunakan sebagai prediktor karena kedua variabel tersebut membentuk target daya. Model selanjutnya diuji pada periode yang tidak digunakan dalam pelatihan.

Penelitian ini menggunakan panel surya nominal 100 W pada sudut 25°, 35°, dan 45° di Kota Langsa. Akuisisi data dilakukan dengan ESP32 v4, DHT22 untuk suhu dan kelembapan lingkungan, BH1750 untuk iluminansi, PT100 tiga kabel melalui MAX31865 untuk suhu panel, serta INA228 untuk tegangan dan arus. Daya dihitung dari hasil kali tegangan dan arus. Setiap rekaman dilengkapi penanda waktu (*timestamp*) dan dikirim ke server sebagai representasi kondisi panel (*state*). Data ini menjadi masukan bagi model prediksi, evaluasi skenario, dashboard, dan penyusunan rekomendasi sudut.

Data penelitian diperoleh dari pengukuran pada panel surya fisik dan digunakan dalam validasi model terhadap daya terukur. Rekaman lingkungan dan kelistrikan menjadi dasar prapemrosesan, pelatihan, pengujian prediksi, serta evaluasi skenario sudut. Asal data pengukuran dibedakan dari perhitungan skenario: data berasal dari panel, sedangkan respons pada alternatif sudut diperkirakan oleh model untuk menyusun rekomendasi. Pemodelan daya perlu tetap mengacu pada hubungan tegangan, arus, dan daya dalam sistem fotovoltaik [15]. Kualitas rekomendasi sudut juga perlu dinilai terpisah dari akurasi prediksi karena prediksi yang akurat belum tentu menghasilkan keputusan operasi terbaik [16].

Pada tingkat implementasi, hubungan antara panel fisik dan model digital perlu diperiksa melalui data yang diperoleh selama pengoperasian. Digital Twin fotovoltaik dengan menghubungkan sensor, layanan IoT, dan model MATLAB, kemudian membandingkan keluaran model dengan hasil pengukuran [17]. Pengintegrasian Digital Twin dan IoT dalam lingkungan kendali sistem energi hibrida serta mengujinya pada demonstrator fotovoltaik [18]. Kedua penelitian tersebut memberi dasar bagi evaluasi yang mencakup aliran data dari perangkat, pembaruan model, dan kesesuaian keluaran digital dengan kondisi fisik. Dalam penelitian ini, hubungan tersebut diperiksa melalui rekaman sensor dan perbandingan daya prediksi terhadap daya terukur pada panel.

Kebutuhan perangkat berbiaya rendah juga mendapat perhatian dalam pengembangan Digital Twin fotovoltaik. Menyusun sistem pemantauan waktu nyata yang menghubungkan pengukuran berbasis IoT dengan model panel untuk memperoleh karakteristik arus–tegangan dan daya–tegangan [19]. Pendekatan ini memperlihatkan bagaimana representasi perilaku panel dapat dipadukan dengan perangkat akuisisi data. Penelitian yang dilakukan di sini menggunakan model berbasis data untuk memperkirakan daya pada kondisi lingkungan dan sudut tertentu. Karena itu, ketepatan prediksi perlu diperiksa pada rekaman pengujian yang terpisah, sedangkan hasil pemantauan digunakan untuk menelusuri kondisi saat penyimpangan prediksi terjadi.

Pemilihan Random Forest perlu mempertimbangkan karakter data dan tujuan prediksi. Sebuah penelitian dengan membandingkan sejumlah algoritma pada beberapa ukuran dataset dan rentang waktu prediksi; hasilnya menunjukkan bahwa keunggulan suatu model bergantung pada kondisi evaluasi [20]. Random Forest dengan pemilihan fitur cuaca untuk meningkatkan prediksi daya fotovoltaik [21], juga membandingkan Random Forest, jaringan saraf dalam, dan LSTM menggunakan variabel cuaca yang mudah diperoleh [22]. Temuan-temuan tersebut mendukung penggunaan Random Forest sebagai model yang layak diuji pada data panel, sekaligus menegaskan perlunya pembandingan dengan model acuan pada data yang sama. Oleh sebab itu, penelitian ini menyertakan baseline rata-rata pelatihan dan persistence dalam evaluasi kinerja.

Pengolahan data sebelum pelatihan turut memengaruhi hasil prediksi. Penggabungan pengelompokan kondisi data dengan Random Forest untuk prediksi daya surya [23]. Penggabungan Random Forest pada tahap pemilihan fitur, kemudian menggabungkannya dengan dekomposisi sinyal dan model prediksi [24]. Kedua pendekatan tersebut menempatkan pemilihan masukan dan pengolahan variasi temporal sebagai bagian dari proses pemodelan. Pada penelitian ini, perhatian diberikan pada kesesuaian penanda waktu, pemeriksaan mutu rekaman,

serta pemisahan data secara kronologis. Tegangan dan arus digunakan untuk menghitung daya terukur, sehingga keduanya tidak dimasukkan sebagai prediktor agar informasi pembentuk target tidak langsung masuk ke model.

Untuk pengaturan sudut, memadukan simulasi PVLib dan machine learning guna menentukan sudut kemiringan bulanan pada beberapa zona iklim [25]. Kajian tersebut menunjukkan bahwa pemilihan sudut berkaitan dengan kondisi lokasi dan periode evaluasi. Ruang pencarian dalam penelitian ini dibatasi pada tiga sudut, yaitu 25°, 35°, dan 45°, sesuai konfigurasi pengujian panel. Exhaustive Scenario Search mengevaluasi ketiga alternatif menggunakan prediksi daya pada kondisi masukan yang sama. Dengan batas tersebut, rekomendasi yang dihasilkan merupakan pilihan terbaik di antara tiga sudut yang diperiksa. Penilaian manfaatnya tetap perlu dibedakan dari pengujian akurasi model, karena perbandingan skenario dan validasi daya menjawab pertanyaan yang berbeda.

Kajian tersebut menempatkan pemantauan kondisi panel, prediksi daya, dan pengaturan sudut sebagai bagian yang saling berkaitan. Persoalan yang dikaji dalam penelitian ini adalah bagaimana hasil prediksi dapat digunakan untuk memilih sudut panel, kemudian dinilai berdasarkan respons daya pada kondisi pengujian yang sama. Persoalan ini memerlukan keterhubungan antara rekaman sensor, model prediksi, dan evaluasi keputusan, sehingga manfaat model dapat diperiksa hingga tahap rekomendasi operasi.

Kebaruan penelitian ini terletak pada rancangan integrasi Random Forest Regression dan Exhaustive Scenario Search dalam proof of concept Digital Twin panel surya 100 W. Rancangan tersebut menggunakan representasi kondisi panel yang sama untuk memprediksi daya dan membandingkan sudut 25°, 35°, dan 45°. Pada setiap kondisi lingkungan dan waktu, model memperkirakan daya ketiga kandidat, lalu hasilnya digunakan untuk menyusun rekomendasi sudut. Ketiga rekaman sudut dalam satu interval analisis dipertahankan dalam partisi kronologis yang sama agar evaluasi prediksi dan keputusan menggunakan susunan data yang konsisten. Kontribusi rancangan ini juga mencakup pemisahan penilaian galat prediksi, kesesuaian pemilihan sudut, dan hasil evaluasi energi. Dengan cara tersebut, hubungan antara ketepatan model dan manfaat rekomendasi dapat diperiksa melalui indikator yang berbeda.

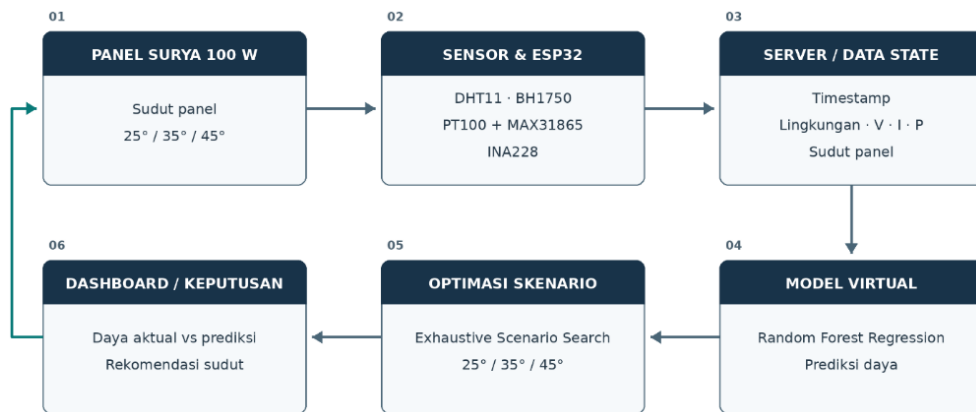
Penelitian ini bertujuan menguji rancangan tersebut melalui akuisisi data lingkungan dan kelistrikan pada panel fisik di Kota Langsa, pemodelan daya dengan Random Forest Regression, serta evaluasi tiga kandidat sudut menggunakan Exhaustive Scenario Search. Dashboard menghubungkan rekaman kondisi panel dengan daya terukur, daya prediksi, dan rekomendasi yang dihasilkan. Pengujian difokuskan pada kesesuaian prediksi terhadap pengukuran panel dan kualitas rekomendasi sudut dalam cakupan proof of concept.

2. METODOLOGI PENELITIAN

2.1 Desain Penelitian dan Arsitektur Digital Twin

Sistem diterapkan pada panel PV nominal 100 W yang diuji pada sudut 25°, 35°, dan 45°. Sensor membaca suhu dan kelembapan lingkungan, iluminansi, suhu panel, tegangan, serta arus. ESP32 v4 mengumpulkan pembacaan tersebut dan mengirimkannya ke server. Rekaman beserta penanda waktunya disimpan sebagai state panel, lalu digunakan oleh Random Forest Regression untuk memprediksi daya. Exhaustive Scenario Search membandingkan prediksi pada tiga kandidat sudut. Dashboard menampilkan daya terukur, prediksi, riwayat pengukuran, dan rekomendasi. Validasi pada panel fisik dilakukan dengan menggunakan daya terukur sebagai pembandingan keluaran model.

Penelitian dimulai dengan perancangan arsitektur dan penentuan variabel state, dilanjutkan dengan akuisisi serta prapemrosesan data lingkungan dan kelistrikan. Dataset berpenanda waktu disusun untuk sudut 25°, 35°, dan 45°, lalu digunakan dalam pelatihan dan pengujian Random Forest Regression. Model diintegrasikan dengan representasi digital dan dashboard. Seluruh kandidat sudut dievaluasi menggunakan Exhaustive Scenario Search, kemudian prediksi dan rekomendasinya divalidasi terhadap daya panel. Hubungan antarkomponen ditampilkan pada Gambar 1.



Gambar 1. Arsitektur Penelitian Digital Twin Panel Surya

Tabel 1 menjelaskan pembagian peran komponen serta aliran data dalam sistem yang dibangun. Panel surya merupakan objek fisik yang diamati, sedangkan sensor dan ESP32 v4 mengakuisisi kondisi lingkungan, suhu panel, tegangan, dan arus beserta penanda waktunya. Server menyimpan rekaman tersebut sebagai representasi kondisi panel dan menyediakan masukan bagi Random Forest Regression. Keluaran prediksi selanjutnya digunakan oleh Exhaustive Scenario Search untuk membandingkan tiga kandidat sudut. Dashboard menyajikan hasil pengukuran, prediksi, dan rekomendasi agar hubungan antara kondisi panel dan keluaran model dapat ditelusuri. Umpan balik digunakan untuk memeriksa kesesuaian hasil digital dengan pengukuran pada panel fisik.

Tabel 1. Pemetaan komponen penelitian Digital Twin

Komponen	Peran Dalam Penelitian	Data/Fungsi Utama	Implementasi Dalam Arsitektur
Aset fisik	Panel PV nominal 100 W; sudut 25°, 35°, 45°	Daya dan konfigurasi sudut	Sumber state fisik panel
Akuisisi sensor	DHT22, BH1750, PT100+MAX31865, INA228, ESP32 v4	Suhu, RH, iluminansi, suhu panel, tegangan, arus, timestamp	Akuisisi dan pengiriman state
Data state	Penyimpanan state panel bertimestamp pada server	Lingkungan, sudut, V, I, P, quality flag	Representasi data untuk model dan dashboard
Model prediksi	Random Forest Regression	Prediksi daya berdasarkan state dan sudut	Predictive service
Optimasi	Exhaustive Scenario Search untuk 25°, 35°, 45°	Tiga prediksi dan rekomendasi sudut	Decision service
Dashboard Digital Twin	Target-vs-prediksi, histori, rekomendasi	Representasi state dan layanan keputusan	Visualisasi dan monitoring
Feedback & validasi	Mengirim rekomendasi/aksi dan membaca respons baru	Sinkronisasi fisik-digital dan validasi respons	Closed-loop evaluation

2.2 Akuisisi dan Struktur Data Penelitian

Dataset berasal dari pengukuran panel fisik dan memuat penanda waktu, sudut panel, suhu lingkungan, kelembapan relatif, iluminansi, suhu panel, tegangan, arus, dan daya. Rekaman disusun untuk sudut 25°, 35°, dan 45° agar respons ketiga konfigurasi dapat dibandingkan. Kondisi panel direpresentasikan dengan interval lima menit untuk mengikuti perubahan lingkungan dan daya secara berurutan. Prapemrosesan meliputi pemeriksaan data hilang, konsistensi penanda waktu, rentang variabel, pencilan, serta kesesuaian daya dengan hasil kali tegangan dan arus.

Dataset analitik mencakup 30 hari, dengan 108 state per hari pada interval lima menit dari pukul 08.00 hingga 16.55. Rekaman dikelompokkan menurut interval analisis dan tiga sudut panel sehingga diperoleh 3.240 kelompok state dan 9.720 rekaman. Pengelompokan ini digunakan untuk membandingkan kandidat sudut dalam pengujian prediksi dan evaluasi skenario. Interval analisis merupakan dasar pengelompokan data, bukan pernyataan bahwa ketiga sudut diukur serentak.

State disusun dari waktu pengukuran, suhu dan kelembapan lingkungan, iluminansi, suhu panel, serta daya terukur. Perubahan variabel tersebut menggambarkan kondisi operasi selama pengujian. Rekaman ketiga sudut digunakan untuk mempelajari hubungan antara kondisi lingkungan, waktu, sudut, dan daya sebelum model memilih kandidat dengan prediksi tertinggi.

Iluminansi dicatat hingga 65.535 lux. Suhu dan kelembapan menggambarkan kondisi lingkungan, sementara suhu panel mewakili respons termal modul. Pemeriksaan kualitas dilakukan pada setiap variabel untuk menandai nilai di luar rentang, data hilang, dan anomali sebelum data diproses oleh model. Penanda kualitas (*quality flag*) disimpan bersama data agar integritas state dapat dipantau pada server.

Daya pada setiap state dihitung sebagai hasil kali tegangan dan arus. Kedua besaran ini juga disimpan untuk memeriksa konsistensi perhitungan dan menggambarkan kondisi kelistrikan panel. Pada sistem fisik, INA228 membaca tegangan dan arus sesuai konfigurasi beban yang digunakan. Nilai daya menjadi target model prediksi sekaligus pembanding keluaran Digital Twin.

Rekaman hasil pengukuran digunakan sejak tahap prapemrosesan hingga evaluasi prediksi dan rekomendasi sudut. Struktur data yang sama digunakan pada model dan dashboard agar hasil analisis dapat ditelusuri ke kondisi panel yang mendasarinya. Pengujian model terhadap daya terukur menjadi bagian dari validasi proof of concept pada panel fisik.

2.3 Model Prediksi Random Forest Regression

Prapemrosesan menghasilkan satu state terstruktur untuk setiap penanda waktu. Fitur utama model terdiri atas iluminansi, suhu lingkungan, kelembapan relatif, sudut kemiringan, dan representasi waktu. Suhu panel dapat digunakan untuk diagnosis atau sebagai fitur apabila nilainya tersedia sebelum keputusan sudut dibuat. Tegangan dan arus tidak dimasukkan sebagai prediktor karena keduanya membentuk target daya secara langsung. Data dibagi secara kronologis agar informasi dari periode pengujian tidak masuk ke proses pelatihan.

Data D01-D18 digunakan untuk pelatihan, D19-D24 untuk validasi, dan D25-D30 untuk pengujian. Tiga rekaman sudut dalam kelompok interval analisis yang sama selalu berada dalam satu partisi. Random Forest Regression dikonfigurasi dengan 300 pohon, bootstrap sampling, squared-error splitting, dan kedalaman maksimum yang tidak dibatasi. Jumlah minimum observasi ditetapkan sebanyak dua untuk membagi node dan dua pada setiap leaf, dengan `max_features 1.0`. Konfigurasi ini dipertahankan pada seluruh evaluasi.

Kinerja prediksi diukur dengan mean absolute error (MAE), root mean squared error (RMSE), koefisien determinasi (R^2), dan mean bias error (MBE). Rata-rata daya pada data pelatihan serta persistence digunakan sebagai model pembanding (*baseline*). Metrik dihitung pada data holdout yang dipisahkan secara kronologis, sehingga penilaian dilakukan pada periode di luar pelatihan.

2.4 Optimasi Skenario Sudut dengan Exhaustive Scenario Search

Optimasi dilakukan dengan membandingkan seluruh kandidat sudut, yaitu 25°, 35°, dan 45°. Untuk setiap state, Random Forest dijalankan tiga kali dengan mengganti sudut, sementara variabel lingkungan dan waktu tetap. Exhaustive Scenario Search memilih sudut dengan prediksi daya tertinggi dari ketiga hasil tersebut. Secara komputasional, proses ini merupakan enumerasi seluruh alternatif yang diikuti operasi `argmax`. Metode ini digunakan sebagai prosedur optimasi dalam penelitian, bukan sebagai algoritma baru.

Akurasi prediksi dan kualitas keputusan dievaluasi secara terpisah. Layanan prediksi dinilai dari selisih antara daya prediksi dan target, sedangkan layanan keputusan dinilai dari ketepatan pemilihan sudut dengan daya tertinggi. Pemisahan ini diperlukan karena galat prediksi yang rendah tidak selalu menjamin urutan kandidat yang benar. Evaluasi Digital Twin dengan demikian mencakup ketepatan estimasi daya dan rekomendasi sudut.

Data pengukuran, model prediksi, evaluasi skenario, dan rekomendasi ditempatkan pada server serta dashboard yang sama. Pembacaan sensor memperbarui representasi digital, lalu model memprediksi daya dari kondisi tersebut. Layanan skenario membandingkan alternatif sudut dan menampilkan rekomendasi. Hubungan antara data fisik dan keluaran model diperiksa melalui perbandingan prediksi dengan daya terukur. Alur umpan balik pada Gambar 1 menjelaskan hubungan rekomendasi dengan respons panel dalam arsitektur sistem.

3. HASIL DAN PEMBAHASAN

3.1 Implementai Arsitektur Digital Twin

Sistem memproses rekaman pengukuran panel fisik menjadi prediksi daya, evaluasi tiga sudut, dan rekomendasi. Sebanyak 9.720 rekaman dikelompokkan ke dalam 3.240 state, masing-masing memuat sudut 25°, 35°, dan 45°. Seluruh kandidat dalam satu kelompok tetap berada pada partisi yang sama. Susunan ini menjaga konsistensi data bagi model virtual dan layanan keputusan serta memungkinkan keluaran model dibandingkan dengan daya hasil pengukuran.

Tabel 2. Struktur Partisi Kronologis Dataset Penelitian

Partisi	Hari	Kelompok state	Rekaman	Fungsi
Pelatihan	D01-D18	1.944	5.832	Fit model referensi
Validasi	D19-D24	648	1.944	Pemeriksaan holdout
Pengujian	D25-D30	648	1.944	Evaluasi final pipeline
Total	30 hari	3.240	9.720	Total dataset

Seperti ditunjukkan pada Tabel 2, pembagian data didasarkan pada hari, bukan pengacakan baris. Cara ini menjaga pemisahan waktu antara data pelatihan dan pengujian serta mempertahankan tiga kandidat sudut pada kondisi yang sama dalam satu kelompok. Model prediksi dan layanan keputusan kemudian dievaluasi dengan struktur data yang konsisten.

Distribusi variabel diperiksa untuk memastikan bahwa state masukan berada dalam rentang rancangan eksperimen. Pemeriksaan ini juga menilai kesesuaian skema data dengan variabel penelitian dan kemampuan perangkat lunak memprosesnya. Kualitas representasi kondisi panel menjadi dasar bagi penggunaan model virtual dan layanan keputusan dalam Digital Twin.

3.2 Kinerja Model Prediksi Random Forest Regression

Validasi terhadap daya terukur pada panel fisik menggunakan 1.944 rekaman pengujian yang dipisahkan dari data pelatihan. Random Forest menghasilkan MAE 2,769 W, RMSE 4,145 W, R^2 0,9446, dan MBE 0,283 W. Hasil ini menunjukkan kesesuaian prediksi dengan daya panel pada kondisi pengujian. Selisih keduanya digunakan sebagai residual untuk memeriksa penyimpangan model sebelum hasil prediksi digunakan dalam evaluasi sudut.

Tabel 3 menyajikan kinerja Random Forest pada data pelatihan, validasi, dan pengujian, disertai dua model pembandingan. Kolom n menunjukkan jumlah rekaman yang dievaluasi; MAE dan RMSE menyatakan besar galat prediksi, R^2 menunjukkan kemampuan model menjelaskan variasi daya terukur, sedangkan MBE menggambarkan bias rata-rata prediksi. RMSE meningkat dari 2,079 W pada pelatihan menjadi 4,492 W pada validasi dan 4,145 W pada pengujian. Pada 1.944 rekaman pengujian, RMSE Random Forest lebih rendah daripada baseline rata-rata pelatihan yang mencapai 18,029 W. Perbandingan dengan persistence lima menit menggunakan 1.926 rekaman yang dipadankan: RMSE Random Forest sebesar 4,158 W, sedangkan persistence sebesar 4,412 W. Dengan demikian, penurunan galat terhadap persistence lebih kecil daripada terhadap baseline konstan. Jumlah rekaman yang dipadankan perlu diperhatikan agar kedua model dibandingkan pada pengamatan yang sama.

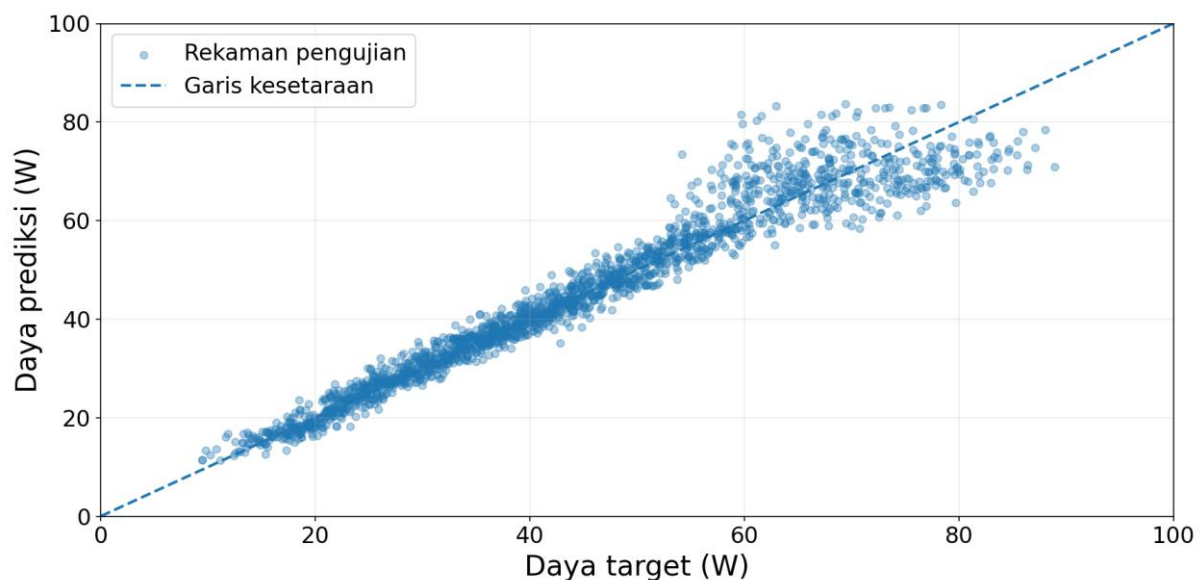
Tabel 3. Kinerja Random Forest pada Data Pelatihan, Validasi, Dan Pengujian

Model dan partisi	n	MAE (W)	RMSE (W)	R^2	MBE (W)
RF: pelatihan	5.832	1,312	2,079	0,9847	0,001
RF: validasi	1.944	2,995	4,492	0,9323	1,350
RF: pengujian penuh	1.944	2,769	4,145	0,9446	0,283
Mean pelatihan: pengujian	1.944	15,292	18,029	-0,0485	3,876
Persistence: pengujian dipadankan	1.926	3,481	4,412	0,9370	0,094
RF: pengujian dipadankan	1.926	2,779	4,158	0,9441	0,277

RMSE model mencapai 2,079 W pada pelatihan, 4,492 W pada validasi, dan 4,145 W pada pengujian. Baseline rata-rata pelatihan menghasilkan RMSE 18,029 W. Pada data pengujian yang dipadankan, RMSE persistence lima menit sebesar 4,412 W, sedangkan Random Forest sebesar 4,158 W. Random Forest memberikan galat lebih rendah daripada kedua pembanding tersebut, meskipun perbaikannya terhadap persistence lebih kecil daripada terhadap baseline konstan.

Gambar 2 membandingkan daya terukur pada panel dengan prediksi pada data pengujian. Sebagian besar titik mengikuti garis kesetaraan. Jarak titik dari garis tersebut menunjukkan residual prediksi dan membantu mengenali rentang daya dengan penyimpangan yang lebih besar. Pola ini digunakan untuk menilai kemampuan model virtual dalam merepresentasikan respons daya panel fisik.

Penilaian prediksi daya tidak cukup hanya berdasarkan R^2 . Residual perlu diperiksa terhadap waktu, iluminansi, suhu panel, kelembapan, dan sudut untuk mengenali kondisi dengan galat yang lebih besar. Pemeriksaan ini diperlukan karena kesalahan prediksi dapat memengaruhi hasil optimasi dan rekomendasi sudut.



Gambar 2. Perbandingan Prediksi Random Forest Dengan Daya Terukur Pada Data Pengujian

3.3 Hasil Optimasi Sudut Panel

Pada 648 state lingkungan pengujian, layanan skenario menghasilkan tiga prediksi untuk setiap state dan memilih salah satu sudut 25°, 35°, atau 45°. Sudut 25° dipilih pada 273 state, 35° pada 70 state, dan 45° pada 305 state. Referensi pembanding memilih ketiga sudut tersebut masing-masing pada 284, 85, dan 279 state. Kesesuaian keputusan mencapai 88,43%, yang menunjukkan bahwa pilihan model sejalan dengan referensi pada sebagian besar kondisi pengujian.

Tabel 4. Hasil Evaluasi Tiga Skenario Sudut

Kebijakan	Energi (Wh)	Perubahan terhadap tetap 35° (%)
Tetap 25°	2.409,56	-0,39
Tetap 35° (baseline)	2.418,91	0,00
Tetap 45°	2.359,15	-2,47
RF + Exhaustive Scenario Search	2.492,97	+3,06
Referensi maksimum tiga kandidat	2.495,84	+3,18

Pada Tabel 4, kombinasi Random Forest dan Exhaustive Scenario Search menghasilkan energi 2.492,97 Wh. Nilai ini meningkat 3,06% dibandingkan energi pada sudut tetap 35° sebesar 2.418,91 Wh. Sudut tetap 25° dan 45° masing-masing menghasilkan 2.409,56 Wh dan 2.359,15 Wh. Dalam periode pengujian ini, pemilihan sudut berdasarkan evaluasi skenario menghasilkan energi lebih besar daripada penggunaan satu sudut tetap.

Proses optimasi dimulai dari penerimaan state lingkungan dan pembentukan tiga kandidat sudut. Random Forest memprediksi daya setiap kandidat, lalu Exhaustive Scenario Search memilih prediksi tertinggi. Hasil prediksi diterjemahkan menjadi rekomendasi operasi dan ditampilkan pada dashboard Digital Twin.

3.4 Pembahasan Kinerja dan Integrasi Digital Twin

Random Forest memodelkan hubungan nonlinier antara kondisi lingkungan, waktu, sudut, dan daya hasil pengukuran. Validasi pada panel fisik menghasilkan R^2 0,9446 dan RMSE 4,145 W pada data pengujian. Selisih kinerja terhadap persistence lebih kecil dibandingkan terhadap baseline konstan. Hasil tersebut menunjukkan perlunya beberapa pembanding agar kemampuan model dinilai secara proporsional. Kesesuaian ini berlaku pada kondisi pengujian yang dilaporkan dan tidak dengan sendirinya menunjukkan kinerja yang sama pada panel atau lokasi lain.

Illuminansi berkaitan langsung dengan perubahan daya panel. Karena rentang pencatatan mencapai 65.535 lux, data pada intensitas tinggi perlu disertai penanda kualitas agar dapat ditelusuri. Validasi sensor, pemeriksaan rentang, deteksi kejenuhan, dan pemeriksaan drift ditempatkan sebelum tahap prediksi untuk memantau kualitas masukan model.

Tabel 5. Interpretasi Hasil dan Implikasinya Terhadap Komponen Digital Twin

Aspek	Hasil utama	Implikasi pada Digital Twin
Random Forest	Pipeline training, holdout, dan prediksi berjalan konsisten	Predictive service menyediakan estimasi daya untuk setiap state
Skenario 25°/35°/45°	Pembentukan kandidat dan pemilihan sudut berjalan	Decision service membandingkan seluruh alternatif sudut
Sensor dan telemetri	Variabel state dan quality control ditetapkan	State fisik dapat ditelusuri melalui timestamp dan quality flag
Keluaran listrik	Tegangan, arus, dan daya tersusun konsisten	Target daya dapat dibandingkan dengan keluaran model
Digital Twin	Model virtual dan decision service terintegrasi	State, prediksi, rekomendasi, dashboard, dan feedback membentuk alur fisik-digital

Kontribusi penelitian terletak pada penggabungan akuisisi data panel, representasi state, prediksi daya, evaluasi sudut, dashboard, dan umpan balik dalam satu sistem Digital Twin. Random Forest Regression dan Exhaustive Scenario Search digunakan sebagai metode analitik utama. Integrasi keduanya memungkinkan prediksi dan rekomendasi sudut dikaitkan dengan kondisi operasi panel.

Perangkat keras menyediakan data kondisi fisik bagi Digital Twin. ESP32, DHT22, BH1750, PT100+MAX31865, dan INA228 mengirimkan data suhu, kelembapan, iluminansi, suhu panel, tegangan, arus, dan penanda waktu ke server. Kualitas telemetri dinilai dari kelengkapan data, konsistensi waktu, kehilangan data, latensi, dan kestabilan pembacaan. Pemeriksaan tersebut diperlukan agar representasi digital tetap sesuai dengan kondisi panel.

Validasi pada panel fisik dilakukan dengan membandingkan daya hasil pengukuran dan daya prediksi pada periode holdout yang tidak digunakan untuk pelatihan. Evaluasi sudut 25°, 35°, dan 45° menggunakan respons daya sebagai pembanding keputusan model. Penanda waktu dan pengelompokan interval digunakan untuk menelusuri kondisi yang mendasari perbandingan. Galat prediksi, kesesuaian keputusan sudut, dan hasil evaluasi energi dilaporkan sebagai indikator yang berbeda, sehingga akurasi prediksi tidak disamakan dengan keberhasilan pengendalian panel secara otomatis.

Konsistensi kondisi fisik dan digital juga menjadi bagian dari validasi. Setiap paket data yang diterima server harus dapat ditelusuri ke waktu pengukuran, sensor, dan sudut panel. Data terlambat atau hilang diberi penanda agar tidak diproses sebagai state yang lengkap. Jika residual meningkat, pemeriksaan perlu membedakan perubahan karakter panel, drift sensor, kondisi di luar distribusi data pelatihan, dan gangguan komunikasi.

Penerapan rekomendasi sudut perlu memperhitungkan waktu dan biaya perpindahan. Exhaustive Scenario Search memilih kandidat dengan prediksi daya tertinggi, tetapi perubahan sudut dalam operasi dapat dibatasi dengan ambang atau hysteresis. Perubahan dilakukan apabila selisih prediksi melampaui biaya aktuasi dan variasi pengukuran. Pendekatan ini memungkinkan pengembangan kebijakan operasi yang lebih stabil daripada pemilihan maksimum sesaat.

3.5 Integrasi Fisik-Digital pada Digital Twin

Implementasi penelitian menghubungkan pembacaan panel fisik dengan representasi digital dan layanan prediksi. Pengiriman data sensor melalui ESP32 ke server menyediakan aliran physical-to-virtual. Dalam klasifikasi Digital

Twin, hubungan virtual-to-physical merujuk pada penerapan keluaran digital kembali ke aset fisik [2], [3]. Penelitian ini disajikan sebagai proof of concept yang telah diuji pada panel fisik; validasi prediksi dibedakan dari klaim pengendalian dua arah otomatis.

Data lingkungan dan kelistrikan hasil pengukuran menyediakan state bagi model. Random Forest memprediksi daya, lalu Exhaustive Scenario Search mengevaluasi alternatif sudut. Dashboard menampilkan kondisi panel, prediksi, dan rekomendasi. Perbandingan keluaran model dengan daya terukur menghubungkan evaluasi komputasi dengan respons panel fisik. Peran setiap komponen dan cakupan evaluasinya dirangkum pada Tabel 6.

Tabel 6. Implementasi dan Cakupan Evaluasi *Proof Of Concept* Pada Panel Fisik

Komponen	Implementasi pada penelitian	Cakupan evaluasi
Akuisisi data fisik	Sensor dan ESP32 membaca kondisi panel dan mengirim data ke server	Rekaman pengukuran menjadi sumber data analisis
Model prediksi	Random Forest memprediksi daya dari kondisi panel	Prediksi dibandingkan dengan daya terukur pada data holdout
Evaluasi sudut	Exhaustive Scenario Search membandingkan 25°, 35°, dan 45°	Kesesuaian keputusan dan energi hasil evaluasi skenario
Proof of concept	Data fisik, model, rekomendasi, dan dashboard diintegrasikan	Validasi pada panel fisik sesuai cakupan pengujian yang dilaporkan

Validasi yang dilaporkan mencakup kesesuaian prediksi terhadap daya terukur pada data holdout serta evaluasi pemilihan sudut. Pengukuran panel menjadi acuan bagi penilaian model, sementara rekomendasi merupakan keluaran dari perbandingan tiga kandidat. Cakupan ini mendukung pengujian proof of concept pada panel fisik tanpa menyamakan hasilnya dengan validasi seluruh fungsi Digital Twin yang bekerja secara otonom.

Setiap prediksi dan rekomendasi perlu dapat ditelusuri ke kondisi yang mendasarinya. Karena itu, sistem menyimpan penanda waktu, sudut, variabel lingkungan, tegangan, arus, daya, prediksi, dan penanda kualitas. Rekaman tersebut memungkinkan penelusuran dari sensor hingga layanan keputusan serta mendukung pembaruan model ketika residual mengindikasikan perubahan karakteristik sistem.

Dashboard menampilkan pembacaan sensor, riwayat daya, suhu panel, iluminansi, suhu dan kelembapan lingkungan, serta status konektivitas. Model menggunakan data tersebut untuk menghasilkan prediksi yang dapat dibandingkan dengan daya terukur. Residual yang kecil menunjukkan kesesuaian pada kondisi yang diuji, sedangkan residual yang meningkat menjadi petunjuk perlunya pemeriksaan sensor, kalibrasi ulang, atau penambahan data pelatihan.

Rekomendasi sudut merupakan keluaran layanan keputusan untuk mendukung pengoperasian panel. Penilaian rekomendasi mengacu pada respons daya dalam data pengujian. Pengendalian melalui aktuator, penerapan hysteresis, dan pembaruan model secara otomatis merupakan fungsi tersendiri; keberhasilannya tidak disimpulkan hanya dari nilai galat prediksi atau hasil evaluasi skenario.

Ruang lingkup sistem mencakup telemetry lingkungan dan kelistrikan, prediksi daya dengan Random Forest Regression, serta optimasi sudut menggunakan Exhaustive Scenario Search. Data lingkungan, tegangan, arus, daya, suhu panel, dan sudut menjadi bagian dari representasi panel. Keluaran sistem berupa prediksi, rekomendasi sudut, dan dashboard digunakan dalam proof of concept yang menghubungkan model digital dengan panel surya.

4. KESIMPULAN

Proof of concept Digital Twin berbasis Machine Learning telah diterapkan dan divalidasi pada panel surya fisik berdaya nominal 100 W. Data diperoleh melalui sensor lingkungan dan kelistrikan, dikirim oleh ESP32 ke server, lalu digunakan oleh Random Forest Regression untuk memprediksi daya. Exhaustive Scenario Search membandingkan sudut 25°, 35°, dan 45°, sedangkan dashboard menyajikan kondisi panel, daya terukur, prediksi, dan rekomendasi. Pada 1.944 rekaman pengujian, Random Forest menghasilkan MAE 2,769 W, RMSE 4,145 W, dan R^2 0,9446. Evaluasi skenario menghasilkan energi 2.492,97 Wh, meningkat 3,06% dibandingkan 2.418,91 Wh pada sudut tetap 35°. Kesesuaian keputusan dengan referensi tiga kandidat mencapai 88,43%. Hasil ini menunjukkan bahwa prediksi daya dan evaluasi skenario dapat dijalankan secara terintegrasi untuk mendukung



rekomendasi sudut dalam Digital Twin. Validasi pada panel fisik menunjukkan kesesuaian model dengan daya hasil pengukuran pada kondisi yang diteliti. Cakupan hasil meliputi akuisisi data panel, prediksi daya, dan evaluasi rekomendasi sudut dalam satu proof of concept. Hasil ini belum menjadi dasar untuk menyimpulkan kemampuan pengendalian otonom, pembaruan model otomatis, atau kinerja pada kondisi di luar pengujian. Pengembangan fungsi tersebut memerlukan evaluasi tersendiri.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM) Universitas Samudra atas pendanaan melalui program hibah penelitian internal yang mendukung pelaksanaan penelitian ini, serta penulis juga mengucapkan terima kasih kepada seluruh pihak yang membantu dalam penyelesaian penelitian ini.

DAFTAR PUSTAKA

- [1] X. Liu *et al.*, “A systematic review of digital twin about physical entities, virtual models, twin data, and applications,” *Advanced Engineering Informatics*, vol. 55, p. 101876, Jan. 2023, doi: 10.1016/j.aei.2023.101876.
- [2] W. Kritzinger, M. Karner, G. Traar, J. Henjes, and W. Sihm, “Digital Twin in manufacturing: A categorical literature review and classification,” *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1016–1022, 2018, doi: 10.1016/j.ifacol.2018.08.474.
- [3] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, “Characterising the Digital Twin: A systematic literature review,” *CIRP J. Manuf. Sci. Technol.*, vol. 29, pp. 36–52, May 2020, doi: 10.1016/j.cirpj.2020.02.002.
- [4] A. Elnosh, M. Calais, and D. Parlevliet, “A systematic literature review of digital twin research for photovoltaic systems: Trends, challenges, and opportunities,” *Renewable and Sustainable Energy Reviews*, vol. 226, p. 116326, Jan. 2026, doi: 10.1016/j.rser.2025.116326.
- [5] K. T. Alao, S. I. U. H. Gilani, K. Sopian, and T. O. Alao, “A review on digital twin application in photovoltaic energy systems: challenges and opportunities,” *JMST Advances*, vol. 6, no. 3, pp. 257–282, Sep. 2024, doi: 10.1007/s42791-024-00083-z.
- [6] I. Natgunanathan, V. Mak-Hau, S. Rajasegarar, and A. Anwar, “Deakin microgrid digital twin and analysis of AI models for power generation prediction,” *Energy Conversion and Management: X*, vol. 18, p. 100370, Apr. 2023, doi: 10.1016/j.ecmx.2023.100370.
- [7] X. Zhao, “A novel digital-twin approach based on transformer for photovoltaic power prediction,” *Sci. Rep.*, vol. 14, no. 1, p. 26661, Nov. 2024, doi: 10.1038/s41598-024-76711-4.
- [8] K. Mahmud, S. Azam, A. Karim, S. Zobaed, B. Shanmugam, and D. Mathur, “Machine Learning Based PV Power Generation Forecasting in Alice Springs,” *IEEE Access*, vol. 9, pp. 46117–46128, 2021, doi: 10.1109/ACCESS.2021.3066494.
- [9] A. F. Amiri, A. Chouder, H. Oudira, S. Silvestre, and S. Kichou, “Improving Photovoltaic Power Prediction: Insights through Computational Modeling and Feature Selection,” *Energies (Basel)*, vol. 17, no. 13, p. 3078, Jun. 2024, doi: 10.3390/en17133078.
- [10] L. Grinsztajn, E. Oyallon, and G. Varoquaux, “Why Do Tree-Based Models Still Outperform Deep Learning on Typical Tabular Data?,” in *Advances in Neural Information Processing Systems 35*, San Diego, California, USA: Neural Information Processing Systems Foundation, Inc. (NeurIPS), 2022, pp. 507–520. doi: 10.52202/068431-0037.
- [11] G. James, D. Witten, T. Hastie, R. Tibshirani, and J. Taylor, *An Introduction to Statistical Learning*. Cham: Springer International Publishing, 2023. doi: 10.1007/978-3-031-38747-0.
- [12] D. R. Roberts *et al.*, “Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure,” *Ecography*, vol. 40, no. 8, pp. 913–929, Aug. 2017, doi: 10.1111/ecog.02881.
- [13] V. Cerqueira, L. Torgo, and I. Mozetič, “Evaluating time series forecasting models: an empirical study on performance estimation methods,” *Mach. Learn.*, vol. 109, no. 11, pp. 1997–2028, Nov. 2020, doi: 10.1007/s10994-020-05910-7.
- [14] S. Kapoor and A. Narayanan, “Leakage and the reproducibility crisis in machine-learning-based science,” *Patterns*, vol. 4, no. 9, p. 100804, Sep. 2023, doi: 10.1016/j.patter.2023.100804.





- [15] R. A. . Messenger and Amir. Abtahi, *Photovoltaic systems engineering*. CRC Press, Taylor & Francis Group, CRC Press is an imprint of the Taylor & Francis Group, an informa business, 2017.
- [16] P. Donti, B. Amos, and J. Z. Kolter, "Task-based End-to-end Model Learning in Stochastic Optimization," in *Advances in Neural Information Processing Systems*, 2017, pp. 5484–5494.
- [17] D. K. M. Idris, C. S. Martiş, R. Birasa, F. naziri Moghaddam, and D. R. Philippe, "From Architecture to Action: Implementing and Validating a Digital Twin for PV Systems Using Real-Time IoT Data," in *2025 International Conference on Future Energy Solutions (FES)*, IEEE, Sep. 2025, pp. 1–6. doi: 10.1109/FES65767.2025.11341797.
- [18] L. Chalal, A. Saadane, and A. Rachid, "Unified Environment for Real Time Control of Hybrid Energy System Using Digital Twin and IoT Approach," *Sensors*, vol. 23, no. 12, p. 5646, Jun. 2023, doi: 10.3390/s23125646.
- [19] P. J. Hueros-Barrios, F. J. R. Sánchez, M. Tradacete-ágreda, P. Martín, C. S. Pérez, and D. Perez-Saura, "A low-cost digital twin for real-time monitoring of photovoltaic panels," in *2023 IEEE 32nd International Symposium on Industrial Electronics (ISIE)*, IEEE, Jun. 2023, pp. 1–6. doi: 10.1109/ISIE51358.2023.10228172.
- [20] C. Scott, M. Ahsan, and A. Albarbar, "Machine learning for forecasting a photovoltaic (PV) generation system," *Energy*, vol. 278, p. 127807, Sep. 2023, doi: 10.1016/j.energy.2023.127807.
- [21] M. Massaoudi, I. Chihi, L. Sidhom, M. Trabelsi, S. S. Refaat, and F. S. Oueslati, "Enhanced Random Forest Model for Robust Short-Term Photovoltaic Power Forecasting Using Weather Measurements," *Energies (Basel)*, vol. 14, no. 13, p. 3992, Jul. 2021, doi: 10.3390/en14133992.
- [22] D. Rangelov, M. Boerger, N. Tcholtchev, P. Lämmel, and M. Hauswirth, "Design and Development of a Short-Term Photovoltaic Power Output Forecasting Method Based on Random Forest, Deep Neural Network and LSTM Using Readily Available Weather Features," *IEEE Access*, vol. 11, pp. 41578–41595, 2023, doi: 10.1109/ACCESS.2023.3270714.
- [23] D. Liu and K. Sun, "Random forest solar power forecast based on classification optimization," *Energy*, vol. 187, p. 115940, Nov. 2019, doi: 10.1016/j.energy.2019.115940.
- [24] D. Niu, K. Wang, L. Sun, J. Wu, and X. Xu, "Short-term photovoltaic power generation forecasting based on random forest feature selection and CEEMD: A case study," *Appl. Soft Comput.*, vol. 93, p. 106389, Aug. 2020, doi: 10.1016/j.asoc.2020.106389.
- [25] O. Z. Lin, L. Štěpanec, E. Koutroulis, D. Juchelkova, and H. Y. Aye, "Optimizing monthly solar PV tilt angles and energy yield across global climate zones: A hybrid machine learning and PVLib approach," *Renew. Energy*, vol. 260, p. 125163, Mar. 2026, doi: 10.1016/j.renene.2025.125163.