

BAB II

TINJAUAN PUSTAKA

2.1 Kajian Relevan

Penelitian mengenai peramalan konsumsi energi dengan pendekatan *time series* telah banyak dilakukan baik di tingkat internasional maupun nasional. Salah satu metode yang paling banyak digunakan adalah *Seasonal Autoregressive Integrated Moving Average with Exogenous Variables* (SARIMAX) karena kemampuannya menangkap pola musiman sekaligus melibatkan variabel eksternal. Beberapa studi, seperti yang dilakukan oleh Tarsitano & Amerise (2017), Elamin & Fukushige (2018), serta Alharbi & Csala (2022), menegaskan bahwa penggunaan SARIMAX dengan variabel eksogen mampu meningkatkan akurasi peramalan dibandingkan model klasik ARIMA. Hasil ini juga diperkuat oleh penelitian Elshewey et al. (2023) yang memperkenalkan varian WD-SARIMAX untuk prediksi suhu harian, menunjukkan relevansi model ini tidak hanya pada sektor energi tetapi juga pada prediksi iklim.

Seiring berkembangnya teknologi komputasi, penelitian terbaru cenderung mengarah pada integrasi SARIMAX dengan metode kecerdasan buatan. Kachalla et al. (2025) mengembangkan kerangka kerja *hybrid* SARIMAX-MLP yang terbukti meningkatkan akurasi prediksi konsumsi energi mikrogrid residensial. Hasil serupa ditunjukkan oleh Ashtar et al. (2025) yang membandingkan SARIMAX, SARIMAX-LSTM, dan model *sequence-to-sequence deep learning* untuk proyeksi kebutuhan listrik di Belanda. Penelitian Maleki et al. (2024) dan Sultana (2025) juga menekankan bahwa meskipun *deep learning* unggul dalam akurasi, model berbasis *time series* klasik seperti SARIMAX tetap lebih efisien secara komputasi, sehingga masih relevan dalam konteks tertentu. Temuan-temuan ini memperlihatkan bahwa pendekatan *hybrid* dan integrasi dengan *machine learning* menjadi tren utama dalam penelitian *forecasting* energi kontemporer.

Selain fokus pada aspek metodologi, sejumlah penelitian juga menekankan pentingnya variabel eksternal dalam model prediksi energi. Penelitian Zulfialda et al. (2025) menunjukkan bagaimana kondisi cuaca, seperti suhu dan curah hujan,

memengaruhi produksi listrik pada PLTA dan dapat dimodelkan secara efektif menggunakan SARIMAX. Sementara itu, Fauzan et al. (2024) berhasil mengimplementasikan SARIMAX dalam sistem monitoring penggunaan daya harian, menegaskan penerapan praktis metode ini di tingkat operasional. Di sisi lain, penelitian Rao et al. (2025) serta Jain et al. (2024) mengintegrasikan peramalan konsumsi energi rumah tangga dengan optimisasi biaya dan energi terbarukan, memperlihatkan relevansi model prediksi dalam mendukung transisi energi berkelanjutan.

Meskipun berbagai penelitian tersebut menunjukkan efektivitas SARIMAX dan variannya, sebagian besar kajian masih berfokus pada tingkat makro seperti beban nasional, perkotaan, atau sistem pembangkit besar. Kajian yang secara khusus menyoroti prediksi konsumsi listrik rumah tangga dengan daya tertentu, seperti 1300 VA, masih sangat terbatas. Padahal, segmen ini memiliki karakteristik konsumsi yang unik dan berfluktuasi tinggi akibat pengaruh sosial-ekonomi maupun kondisi lingkungan. Oleh karena itu, penelitian ini berusaha mengisi celah tersebut dengan menerapkan metode SARIMAX untuk membangun model prediksi konsumsi listrik rumah tangga di wilayah ULP Balai Karangan, dengan mempertimbangkan variabel eksogen yang relevan.

Tabel 2. 1 Matriks Penelitian Relevan

No	1
Judul Jurnal	Data-driven hybrid SARIMAX-MLP framework for energy consumption prediction in residential micro-grid
Penulis & Tahun	Ibrahim Ali Kachalla, Christian Ghiaus, Adeniran Ademuwagun, Olufemi Babajide Odeyinde, Majid Baseer (2025)
Metode	Metode yang digunakan adalah pendekatan <i>hybrid</i> , yaitu menggabungkan model SARIMAX dengan MLP (<i>Multi-Layer Perceptron</i>).

Dataset	Data yang digunakan adalah konsumsi energi listrik dan air panas dari <i>electric water boilers</i> (EWB) di 67 apartemen di Prancis, dengan rentang waktu sekitar 18 bulan.
Hasil Penelitian	Model <i>hybrid</i> SARIMAX-MLP menunjukkan performa terbaik dengan peningkatan akurasi dibandingkan model tunggal dan <i>hybrid</i> lainnya, serta stabil pada skenario musiman dan harian. Analisis SHAP menunjukkan konsumsi air panas sebagai faktor dominan (59%), diikuti oleh jam dalam sehari (20%), sementara hari dalam seminggu berpengaruh minimal.
Keterkaitan Penelitian	Persamaan: Keduanya menggunakan metode SARIMAX untuk menangkap pola musiman dan pengaruh variabel eksogen dalam prediksi konsumsi energi rumah tangga. Perbedaan: Penelitian ini menerapkan model <i>hybrid</i> SARIMAX-MLP untuk meningkatkan akurasi prediksi <i>electric water boilers</i> di apartemen Prancis, sedangkan penelitian penulis menggunakan SARIMAX sebagai model tunggal untuk konsumsi listrik rumah tangga dengan daya 1300 VA di ULP Balai Karangan.

No	2
Judul Jurnal	Short-term load forecasting using a two-stage SARIMAX model
Penulis & Tahun	Agostino Tarsitano & Ilaria L. Amerise, 2017
Metode	Penelitian ini mengusulkan metode <i>two-stage</i> SARIMAX untuk prediksi beban listrik jangka pendek. Tahap pertama menggunakan regresi linear dengan <i>lagged load</i> dan <i>calendar effect</i> untuk memodelkan variabel eksogen. Tahap kedua menerapkan model SARIMA pada <i>residual</i> regresi, yang kemudian diintegrasikan melalui SARIMAX untuk menangani autokorelasi dan meningkatkan akurasi prediksi.

Dataset	Dataset yang digunakan berupa data beban listrik jangka pendek berbasis <i>time series</i> dari enam makro-wilayah di Italia, dengan resolusi data per jam serta mempertimbangkan faktor kalender dan periode waktu khusus.
Hasil Penelitian	Metode <i>two-stage</i> SARIMAX menghasilkan prediksi dengan tingkat kesalahan rendah, dengan RMSPE sebesar 1,32%–2,80% untuk prediksi 24 jam dan 2,27%–3,54% untuk prediksi 216 jam. Metode ini menunjukkan kinerja yang lebih baik dibandingkan regresi linear dan SARIMAX tunggal karena mampu mengurangi autokorelasi <i>residual</i> .
Keterkaitan Penelitian	Persamaan: Sama-sama menggunakan SARIMAX untuk prediksi konsumsi listrik jangka pendek dengan variabel eksogen. Perbedaan: Penelitian ini berskala makro dan menggunakan <i>two-stage</i> SARIMAX, sedangkan penelitian penulis berfokus pada skala rumah tangga di wilayah spesifik menggunakan SARIMAX tunggal.

No	3
Judul Jurnal	Modeling and forecasting hourly electricity demand by SARIMAX with interactions
Penulis & Tahun	Niematallah Elamin & Mototsugu Fukushige, 2018
Metode	Model <i>interaction</i> -SARIMAX digunakan untuk memprediksi permintaan listrik per jam dengan membandingkan SARIMAX standar dan SARIMAX yang diperluas melalui <i>cross effects</i> variabel eksogen.
Dataset	Data konsumsi listrik per jam dari TEPCO selama periode 2012–2015 yang digunakan untuk pemodelan dan pengujian prediksi.

Hasil Penelitian	Model <i>interaction</i> -SARIMAX menunjukkan performa yang lebih baik dibandingkan SARIMAX standar dan model pembandingan lainnya, dengan penurunan MAPE sebesar 22,2%, MAE 21,3%, dan RMSE 21,8%. Peningkatan akurasi konsisten pada berbagai kondisi musiman dan tipe hari, menunjukkan pentingnya interaksi antar variabel eksogen dalam prediksi beban listrik jangka pendek.
Keterkaitan Penelitian	Persamaan: Sama-sama menggunakan SARIMAX untuk prediksi konsumsi listrik jangka pendek dengan variabel eksogen. Perbedaan: Penelitian ini menerapkan <i>interaction</i>-SARIMAX pada data makro Jepang, sedangkan penelitian penulis menggunakan SARIMAX standar pada konsumsi listrik rumah tangga di Indonesia.

No	4
Judul Jurnal	Hybrid Forecasting for Sustainable Electricity Demand in The Netherlands Using SARIMAX, SARIMAX-LSTM, and Sequence-to-Sequence Deep Learning Models
Penulis & Tahun	Ashtar, Duaa; Mohammadi Ziabari, Seyed Sahand; Alsahag, Ali Mohammed Mansoor, 2025
Metode	Pendekatan <i>hybrid</i> diterapkan dengan membandingkan SARIMAX, SARIMAX-LSTM, dan model Seq2Seq berbasis <i>deep learning</i> untuk prediksi permintaan listrik jangka pendek dan menengah.
Dataset	Data konsumsi listrik nasional Belanda berbasis <i>time series</i> dengan mempertimbangkan faktor musiman dan variabel cuaca.

Hasil Penelitian	Model <i>hybrid</i> SARIMAX–LSTM dan Seq2Seq menunjukkan akurasi prediksi yang lebih tinggi dibandingkan SARIMAX tunggal. SARIMAX–LSTM efektif menangkap pola non-linear sekaligus mempertahankan kemampuan SARIMAX dalam memodelkan variabel eksogen, sedangkan Seq2Seq menunjukkan kinerja yang baik untuk prediksi jangka menengah. Hasil ini menegaskan keunggulan pendekatan <i>hybrid</i> dalam meningkatkan presisi prediksi permintaan listrik.
Keterkaitan Penelitian	Persamaan: Sama-sama menggunakan SARIMAX untuk prediksi konsumsi listrik. Perbedaan: Penelitian ini menerapkan model <i>hybrid</i> SARIMAX–LSTM dan Seq2Seq pada data konsumsi listrik nasional Belanda, sedangkan penelitian penulis menggunakan SARIMAX tunggal pada konsumsi listrik rumah tangga di Indonesia.

No	5
Judul Jurnal	Prediction of Hydropower Plant Electricity Production Dependence on Weather Conditions Using the SARIMAX Model
Penulis & Tahun	Dennis Hasnan Zulfiyaldi, Catur Arie Nugroho, Hakim Luthfi Malasan, 2025
Metode	Model SARIMAX digunakan untuk memprediksi produksi listrik PLTA dengan memasukkan variabel cuaca sebagai variabel eksogen berdasarkan data historis produksi dan kondisi bendungan.
Dataset	Data produksi listrik PLTA berbasis <i>time series</i> dari salah satu pembangkit di Sulawesi Selatan selama periode 2020–2023, disertai data cuaca sebagai variabel pendukung.

Hasil Penelitian	Model SARIMAX dengan parameter optimal menunjukkan performa prediksi yang baik, dengan nilai MSE 0,00101, MAE 0,0274, dan RMSE 0,0318. Dan menunjukkan bahwa variabel eksogen terkait <i>El Niño</i> berperan penting dalam prediksi produksi listrik PLTA.
Keterkaitan Penelitian	Persamaan: Sama-sama menggunakan SARIMAX dengan variabel eksogen. Perbedaan: Penelitian ini mengkaji produksi listrik PLTA skala makro, sedangkan penelitian penulis berfokus pada konsumsi listrik rumah tangga skala mikro.

No	6
Judul Jurnal	Sistem Forecasting Penggunaan Daya Harian dengan Metode SARIMAX
Penulis & Tahun	Ahmad Fauzan, Muhammad Erpandi Dalimunthe, Parlin Siagian, 2024
Metode	Metode SARIMAX diterapkan untuk membangun sistem peramalan penggunaan daya listrik harian dengan mempertimbangkan pola musiman dan variabel eksogen.
Dataset	Data penggunaan daya listrik harian berbasis <i>time series</i> yang diperoleh dari pemantauan konsumsi listrik rumah tangga, disertai variabel pendukung terkait kondisi lingkungan dan waktu.
Hasil Penelitian	Model SARIMAX menunjukkan performa yang memadai dalam memprediksi konsumsi daya harian dengan variabilitas tinggi, dengan hasil evaluasi MAE, RMSE, dan MAPE yang baik. Namun, model masih memiliki keterbatasan pada signifikansi parameter dan asumsi normalitas <i>residual</i> .

Keterkaitan Penelitian	Persamaan: Sama-sama menggunakan SARIMAX untuk peramalan konsumsi listrik rumah tangga. Perbedaan: Penelitian ini berorientasi pada implementasi perangkat keras, sedangkan penelitian penulis berfokus pada pemodelan statistik.
------------------------	---

No	7
Judul Jurnal	A Seasonal Autoregressive Integrated Moving average with Exogenous Factors (SARIMAX) Forecasting Model-Based Time series Approach
Penulis & Tahun	Fahad Radhi Alharbi, Denes Csala, 2022
Metode	Model SARIMAX diterapkan untuk meramalkan kinerja sektor kelistrikan jangka panjang dengan mempertimbangkan variabel eksogen dan evaluasi performa model.
Dataset	Data sektor kelistrikan Arab Saudi berbasis <i>time series</i> kuartalan periode 1980–2020.
Hasil Penelitian	Model SARIMAX menunjukkan performa sangat baik dengan nilai R^2 sekitar 99% dan tingkat kesalahan rendah (MAPE $\pm 0,3\%$ dan RMSE konsumsi listrik sekitar 1 TWh). Proyeksi jangka panjang menunjukkan konsumsi listrik Arab Saudi mencapai sekitar 644 TWh pada 2050, dengan kapasitas terpasang 160 GW dan beban puncak 105 GW,
Keterkaitan Penelitian	Persamaan: Sama-sama menggunakan SARIMAX untuk peramalan konsumsi listrik. Perbedaan: Penelitian ini berskala makro jangka panjang, sedangkan penelitian penulis berskala mikro jangka menengah.

No	8
Judul Jurnal	A Novel WD-SARIMAX Model for Temperature Forecasting Using Daily Delhi Climate Dataset

Penulis & Tahun	Ahmed M. Elshewey, Mahmoud Y. Shams, Abdelghafar M. Elhady, Samaa M. Shohieb, Abdelaziz A. Abdelhamid, Abdelhameed Ibrahim, Zahraa Tarek, 2023
Metode	Model <i>hybrid</i> WD-SARIMAX diterapkan dengan menggabungkan <i>wavelet decomposition</i> dan SARIMAX untuk meningkatkan akurasi peramalan suhu harian.
Dataset	Data iklim harian di Delhi, India periode 2013–2017.
Hasil Penelitian	Model WD-SARIMAX menunjukkan performa terbaik dibandingkan model pembanding, dengan akurasi tinggi (MSE 2,8; MAE 1,13; RMSE 1,67; MAPE 4,9%; dan R^2 0,91). Proyeksi hingga 2025 menunjukkan tren suhu yang relatif stabil.
Keterkaitan Penelitian	Persamaan: Sama-sama menggunakan SARIMAX. Perbedaan: Penelitian ini menerapkan WD-SARIMAX untuk prediksi suhu, sedangkan penelitian penulis menggunakan SARIMAX untuk konsumsi listrik rumah tangga.

No	9
Judul Jurnal	Automated Energy Management for Sustainable and Cost-Efficient Households: A Forecasting and Optimization Approach
Penulis & Tahun	Abhishek S. Rao, Apoorva Nayak B, Anand Bhat B, Anusha V. Shenoy, Bakthi S. Shetty, Ramaprasad Poojary, dan Anoushka Acharya, 2025
Metode	Model SARIMAX digunakan untuk peramalan konsumsi energi rumah tangga, yang kemudian dikombinasikan dengan XGBoost dan algoritma <i>Greedy</i> untuk optimasi penggunaan energi.

Dataset	Data konsumsi energi harian rumah tangga berbasis <i>time series</i> yang mencakup penggunaan peralatan listrik utama dan komponen biaya listrik.
Hasil Penelitian	Model SARIMAX menghasilkan prediksi konsumsi dengan MAE 1,74, RMSE 2,22, dan MAPE 23,81%. Integrasi dengan metode optimasi meningkatkan kinerja secara signifikan (MAE 0,24 dan R^2 0,97) serta mampu menekan tagihan listrik sesuai batas anggaran melalui rekomendasi pengelolaan penggunaan energi.
Keterkaitan Penelitian	Persamaan: Sama-sama menggunakan SARIMAX untuk prediksi konsumsi energi rumah tangga. Perbedaan: Penelitian ini menambahkan optimasi biaya dan sistem IoT, sedangkan penelitian penulis berfokus pada prediksi statistik

No	10
Judul Jurnal	Balancing Accuracy and Efficiency in Multi-Step Building-Level Electric Load Forecasting: Deep Learning vs. Classical Time-Series Models
Penulis & Tahun	Nahid Sultana, 2025
Metode	Pendekatan klasik SARIMAX dibandingkan dengan model <i>hybrid</i> BOA-NARX dan BOA-LSTM berbasis <i>deep learning</i> untuk peramalan beban listrik jangka pendek.
Dataset	Data beban listrik per jam berbasis <i>time series</i> pada tingkat bangunan, disertai variabel cuaca pendukung.
Hasil Penelitian	BOA-LSTM memberikan performa terbaik ($R^2 > 0,99$; MAPE $< 3\%$) di semua musim, SARIMAX kompetitif pada volatilitas rendah dengan biaya komputasi lebih rendah, sedangkan BOA-NARX memiliki akurasi terendah.

Keterkaitan Penelitian	Persamaan: Sama-sama menggunakan SARIMAX dengan variabel eksogen cuaca. Perbedaan: Penelitian ini berskala bangunan dan membandingkan dengan <i>deep learning</i>, sedangkan penelitian penulis berskala rumah tangga.
------------------------	--

No	11
Judul Jurnal	Prediction and Analysis of Household Energy Consumption Integrated with Renewable Energy Sources Using Machine learning Algorithms in Energy Management
Penulis & Tahun	Nirbhi Jain, Shreya Sharma, Vijyant Thakur, Mounica Nutakki, dan Srihari Mandava, 2024
Metode	Pendekatan <i>machine learning</i> dan <i>time series</i> diterapkan dalam sistem manajemen energi rumah tangga berbasis IoT untuk prediksi konsumsi energi, dengan membandingkan beberapa model termasuk SARIMAX dan LSTM.
Dataset	Data konsumsi energi rumah tangga dari <i>smart meter</i> dengan variabel cuaca pendukung.
Hasil Penelitian	Model AdaBoost dan Random Forest menunjukkan performa terbaik pada pendekatan klasifikasi (RMSE 0,064–0,072; $R^2 > 0,69$). Pada peramalan deret waktu, LSTM multivariat memberikan akurasi tertinggi (RMSE 0,110; R^2 0,70), mengungguli ARIMA dan SARIMAX.
Keterkaitan Penelitian	Persamaan: Sama-sama memprediksi konsumsi listrik rumah tangga dengan variabel cuaca. Perbedaan: Penelitian ini mengkaji <i>machine learning</i> dan energi terbarukan, sedangkan penelitian penulis berfokus pada SARIMAX pada skala rumah tangga tertentu.

No	12
Judul Jurnal	Future Energy Insights: Time-Series and Deep Learning Models for City Load Forecasting
Penulis & Tahun	Neda Maleki, Oxana Lundström, Arslan Musaddiq, John Jeansson, Tobias Olsson, dan Fredrik Ahlgren, 2024
Metode	Penelitian ini membandingkan model <i>time series</i> klasik dan <i>deep learning</i> untuk peramalan konsumsi energi tingkat kota, dengan menggunakan SARIMAX, FB Prophet, Random Forest, XGBoost, dan CNN.
Dataset	Data konsumsi listrik per jam tingkat kota dengan variabel suhu dan harga listrik.
Hasil Penelitian	Model CNN menunjukkan performa terbaik dengan kesalahan prediksi terendah (RMSE 8,22 MWh; MAE 5,74 MWh; MAPE 4,35%), mengungguli SARIMAX, FB Prophet, XGBoost, dan Random Forest.
Keterkaitan Penelitian	Persamaan: Sama-sama menggunakan SARIMAX dengan variabel eksogen. Perbedaan: Penelitian ini berskala kota dan membandingkan <i>deep learning</i> , sedangkan penelitian penulis berskala rumah tangga.

Berdasarkan hasil tinjauan literatur, metode SARIMAX terbukti andal dalam memodelkan dan meramalkan konsumsi energi pada berbagai skala, mulai dari rumah tangga hingga tingkat kota dan sistem energi terbarukan. Sejumlah penelitian menunjukkan bahwa akurasi peramalan dapat ditingkatkan melalui pemanfaatan variabel eksogen maupun pengembangan pendekatan *hybrid* dengan metode *machine learning* dan *deep learning*. Namun demikian, penelitian terdahulu umumnya masih terbatas pada konteks wilayah atau objek tertentu, sehingga kajian yang lebih spesifik pada tingkat unit layanan pelanggan dan kelompok daya tertentu masih relatif terbatas.

Oleh karena itu, penelitian ini difokuskan pada prediksi konsumsi listrik rumah tangga daya 1300 VA di ULP Balai Karangan dengan menggunakan metode SARIMAX dan mempertimbangkan variabel eksogen yang relevan dengan kondisi lokal. Diharapkan, penelitian ini tidak hanya menghasilkan peramalan yang akurat, tetapi juga memberikan kontribusi praktis dalam mendukung pengelolaan energi yang lebih efektif dan berkelanjutan pada tingkat unit layanan pelanggan.

2.2 Landasan Teori

Landasan teori berfungsi sebagai dasar konseptual yang mencakup pembahasan berbagai teori yang akan digunakan dalam penelitian ini dan didapat dari berbagai buku, jurnal maupun artikel ilmiah. Berikut merupakan landasan teori yang digunakan :

2.2.1 Energi Listrik

Energi listrik merupakan bentuk energi sekunder yang dihasilkan melalui proses konversi berbagai sumber energi primer seperti batubara, gas, air, angin, maupun sinar matahari menjadi energi listrik yang dapat digunakan untuk memenuhi kebutuhan manusia. Dalam konteks sistem ketenagalistrikan, energi listrik berperan penting sebagai penggerak utama aktivitas industri, ekonomi, dan rumah tangga, karena sifatnya yang mudah dikonversi, didistribusikan, serta digunakan dalam berbagai sektor. Efisiensi dalam produksi dan distribusi energi listrik menjadi salah satu indikator penting dalam pembangunan energi berkelanjutan (Zulfialda et al., 2025).

Lebih lanjut, penelitian oleh Ashtar et al. (2025) menekankan bahwa peningkatan konsumsi listrik global menuntut adanya sistem pengelolaan dan peramalan energi yang lebih adaptif. Energi listrik tidak hanya dilihat sebagai komoditas teknis, tetapi juga sebagai faktor strategis dalam mendukung stabilitas energi nasional dan transisi menuju sistem energi yang lebih ramah lingkungan melalui integrasi teknologi cerdas dan sumber energi terbarukan.

2.2.2 Konsumsi Listrik

Konsumsi listrik rumah tangga merupakan jumlah energi listrik yang digunakan oleh rumah tangga dalam memenuhi kebutuhan aktivitas sehari-hari,

seperti penerangan, pendingin ruangan, pemanas air, peralatan elektronik, dan mesin rumah tangga lainnya. Pola konsumsi listrik rumah tangga sangat dipengaruhi oleh berbagai faktor, antara lain jumlah penghuni, tingkat pendapatan, kondisi iklim, dan efisiensi peralatan listrik yang digunakan (Jain et al., 2024; Ridha, 2025).

Selain itu, peningkatan penggunaan teknologi rumah tangga modern dan sistem otomasi turut mendorong naiknya kebutuhan listrik secara signifikan, terutama pada wilayah urban dan suburban. Penelitian oleh Rao et al. (2025) menjelaskan bahwa rumah tangga berdaya besar cenderung menunjukkan pola konsumsi yang fluktuatif akibat adanya variasi penggunaan alat berdaya tinggi seperti pendingin udara, pemanas, dan mesin cuci. Oleh karena itu, memahami karakteristik konsumsi listrik rumah tangga menjadi dasar penting dalam upaya pengelolaan energi serta perancangan sistem prediksi beban listrik yang lebih efisien dan berkelanjutan.

2.2.3 Prediksi

Prediksi (*forecasting*) merupakan suatu proses ilmiah untuk memperkirakan kejadian yang akan terjadi di masa mendatang berdasarkan pola dan data historis yang telah dikumpulkan. Proses ini tidak hanya berfungsi sebagai alat bantu analisis, tetapi juga sebagai dasar pengambilan keputusan strategis dalam berbagai bidang seperti industri, ekonomi, dan manajemen operasional (Puji Catur Siswipraptini et al., 2023). Dalam konteks sektor energi, khususnya pada penyediaan listrik rumah tangga, kemampuan untuk melakukan prediksi secara akurat menjadi hal yang sangat penting untuk mengoptimalkan produksi, distribusi, serta perencanaan kebutuhan daya di masa depan (Abdurrasyid et al., 2021).

Dalam menghasilkan prediksi yang efektif, terdapat dua elemen utama yang harus diperhatikan:

1. Pengumpulan data yang relevan, kualitas prediksi sangat bergantung pada data historis yang lengkap, akurat, dan representatif terhadap kondisi yang akan diprediksi.
2. Pemilihan metode prediksi yang tepat, teknik yang digunakan harus disesuaikan dengan karakteristik data dan tujuan peramalan, misalnya *statistical forecasting* untuk data terstruktur atau *machine learning forecasting* untuk pola kompleks.

Dalam hal ini, prediksi bukan hanya sekadar menebak tetapi melibatkan penggunaan teknik dan alat yang berbasis pada data historis untuk memperkirakan peristiwa yang akan datang. Oleh karena itu, prediksi yang baik dapat menjadi alat yang sangat berharga dalam proses pengambilan keputusan untuk mencapai tujuan organisasi maupun operasional. Secara umum, pendekatan dalam prediksi dibagi menjadi dua metode utama:

1. Metode Kualitatif bersifat subjektif dan biasanya digunakan ketika data historis tidak tersedia atau tidak relevan. Penilaian dilakukan berdasarkan intuisi, pengalaman, atau pendapat para ahli. Contohnya digunakan pada peramalan kebijakan energi baru di wilayah yang belum memiliki data konsumsi historis.
2. Metode Kuantitatif menggunakan data numerik atau statistik untuk membangun model matematis. Metode ini bersifat objektif karena hasilnya dapat diukur dan diuji. Model ARIMA, SARIMA, hingga SARIMAX termasuk dalam kategori ini dan banyak digunakan dalam peramalan energi serta sistem ekonomi modern.

Berdasarkan jangka waktunya, prediksi dapat dibedakan menjadi tiga kategori utama (Tarsitano & Amerise, 2017) :

1. Prediksi Jangka Panjang (*Long-Term Forecasting*) yaitu prediksi untuk kebutuhan dalam periode lebih dari 18 bulan, biasanya digunakan dalam perencanaan strategis dan pembangunan infrastruktur energi jangka panjang.
2. Prediksi Jangka Menengah (*Mid-Term Forecasting*) mencakup periode antara 3 hingga 18 bulan, sering digunakan dalam pengaturan pasokan energi musiman dan perencanaan operasional sistem kelistrikan.
3. Prediksi Jangka Pendek (*Short-Term Forecasting*) digunakan untuk memprediksi kebutuhan dalam jangka waktu kurang dari 3 bulan, seperti beban harian atau mingguan untuk menjaga kestabilan sistem distribusi listrik.

2.2.4 Machine Learning

Machine learning merupakan cabang dari kecerdasan buatan (*artificial intelligence*) yang berfokus pada pengembangan algoritma yang mampu belajar dari data untuk melakukan prediksi atau pengambilan keputusan tanpa perlu diprogram

secara eksplisit (Hendra Jatnika et al., 2025). Pendekatan ini memungkinkan sistem untuk mengenali pola, menganalisis hubungan antar variabel, serta meningkatkan akurasi hasil prediksi seiring bertambahnya data yang dipelajari. Dalam penelitian modern, *machine learning* banyak digunakan dalam berbagai bidang, termasuk analisis konsumsi energi dan sistem prediksi permintaan listrik karena kemampuannya dalam mengolah data besar dan kompleks secara efisien (Afandi et al., 2024).

CRISP-DM (*Cross Industry Standard Process for Data Mining*) adalah metodologi standar yang digunakan dalam pengembangan proyek data mining secara sistematis melalui enam tahapan utama, yaitu *business understanding*, *data understanding*, *data preparation*, *modelling*, *evaluation*, dan *deployment*. Setiap tahap saling berkaitan untuk memastikan proses analisis berjalan terarah, mulai dari identifikasi kebutuhan bisnis hingga penerapan hasil model secara praktis (Rifai et al., 2019).

Gambar 2. 1 CRISP-DM

Tahap ini berfokus pada pemahaman tujuan dan kebutuhan dari sudut pandang bisnis. Tujuannya adalah untuk merumuskan masalah utama yang akan diselesaikan melalui proses data mining serta menentukan sasaran analisis yang sejalan dengan kebutuhan organisasi atau penelitian.

2. *Data Understanding*

Pada tahap ini dilakukan pengumpulan data awal dan eksplorasi untuk memahami karakteristik data. Proses ini mencakup identifikasi sumber data, pengecekan kelengkapan data, serta analisis awal untuk menemukan pola atau anomali.

3. *Data Preparation*

Tahap ini mencakup pemrosesan dan pembersihan data agar siap digunakan untuk pemodelan. Kegiatan yang dilakukan meliputi seleksi variabel, penanganan nilai hilang, normalisasi, dan transformasi data sesuai kebutuhan model.

4. *Modeling*

Pada tahap ini dilakukan pemilihan dan penerapan teknik pemodelan yang sesuai, seperti algoritma statistik atau *machine learning*. Parameter model diatur dan diuji untuk memperoleh hasil yang paling akurat sesuai dengan tujuan analisis.

5. *Evaluation*

Model yang telah dibangun kemudian dievaluasi untuk menilai performa dan kesesuaiannya terhadap tujuan bisnis atau penelitian. Tahapan ini memastikan bahwa model tidak hanya akurat secara statistik, tetapi juga relevan secara praktis.

6. *Deployment*

Tahap terakhir adalah penerapan model yang telah tervalidasi ke dalam sistem atau proses bisnis yang nyata. Hasil model dapat digunakan untuk pengambilan keputusan, pembuatan laporan, atau implementasi sistem otomatisasi prediksi.

2.2.6 *Data Time series*

Data time series merupakan sekumpulan data yang dikumpulkan dan direkam secara berurutan berdasarkan waktu dengan interval yang teratur, seperti harian, mingguan, bulanan, atau tahunan. Data jenis ini digunakan untuk menganalisis pola dan kecenderungan (*trend*) yang terjadi dari waktu ke waktu sehingga dapat dilakukan peramalan terhadap nilai di masa mendatang. Analisis *time series* sangat penting dalam bidang energi listrik karena memungkinkan identifikasi pola musiman, tren jangka panjang, serta fluktuasi konsumsi daya yang terjadi secara periodik (Tarsitano & Amerise, 2017).

2.2.7 ARIMA (*Autoregressive Integrated Moving Average*)

Model ARIMA (*Autoregressive Integrated Moving Average*) merupakan salah satu metode peramalan deret waktu yang digunakan untuk menganalisis data non-stasioner dengan cara melakukan proses pembedaan (*differencing*) agar menjadi stasioner. Model ini menggabungkan tiga komponen utama, yaitu *autoregressive* (AR), *integrated* (I), dan *moving average* (MA). Komponen AR menggambarkan hubungan antara nilai saat ini dengan nilai masa lalunya, komponen I menunjukkan tingkat pembedaan data untuk mencapai kestasioneran, sedangkan komponen MA menggambarkan hubungan antara nilai saat ini dengan kesalahan (*error*) masa lalu. Model ARIMA banyak digunakan dalam analisis deret waktu karena kemampuannya dalam menangkap pola tren dan fluktuasi jangka pendek secara efektif (Alharbi & Csala, 2022). Secara umum, model ARIMA dilambangkan sebagai $ARIMA(p, d, q)$, dengan persamaan matematis:

$$Y_t = c + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (2.1)$$

Keterangan :

Y_t : nilai deret waktu pada periode ke- t

c : konstanta

ϕ_i : koefisien *autoregressive* (AR) yang menunjukkan pengaruh nilai masa lalu terhadap nilai saat ini

θ_j : koefisien *moving average* (MA) yang menggambarkan pengaruh *error* masa lalu terhadap nilai saat ini

ε_t : *error* acak pada periode ke- t

p : orde *autoregressive* (jumlah lag pada variabel dependen)

d : tingkat pembedaan (jumlah kali data di-*differencing* agar stasioner)

q : orde *moving average* (jumlah lag pada *error*)

2.2.8 SARIMA (*Seasonal Autoregressive Integrated Moving Average*)

Model SARIMA merupakan pengembangan dari model ARIMA yang digunakan untuk menganalisis dan memprediksi data deret waktu yang memiliki pola musiman (*seasonal*). Model ini menambahkan komponen musiman pada struktur ARIMA agar mampu menangkap fluktuasi data yang berulang secara periodik, seperti

pola bulanan, triwulanan, atau tahunan (Tarsitano & Amerise, 2017). SARIMA dilambangkan dengan notasi:

$$SARIMA(p, d, q)(P, D, Q)_s$$

dengan persamaan umum:

$$\Phi_P(B^s)\phi_p(B)(1-B)^d(1-B^s)^D Y_t = \Theta_Q(B^s)\theta_q(B)\varepsilon_t \quad (2.2)$$

Keterangan:

Y_t : nilai deret waktu pada periode ke-t

B : operator lag

s : panjang periode musiman (misalnya, $s = 12$ untuk data bulanan)

p, d, q : parameter non-musiman ARIMA

P, D, Q : parameter musiman

$\phi_p(B)$: komponen *autoregressive* non-musiman

$\theta_q(B)$: komponen *moving average* non-musiman

$\Phi_P(B^s)$: komponen *autoregressive* musiman

$\Theta_Q(B^s)$: komponen *moving average* musiman

ε_t : *error* acak pada periode ke-t

2.2.9 SARIMAX

Model SARIMAX (*Seasonal Autoregressive Integrated Moving Average with Exogenous Variables*) merupakan pengembangan dari model SARIMA yang menambahkan variabel eksogen (X) sebagai faktor eksternal yang dapat memengaruhi nilai deret waktu. Variabel eksogen ini dapat berupa data yang relevan dan memiliki hubungan signifikan dengan variabel utama, seperti suhu, kelembapan, atau data ekonomi dalam konteks peramalan energi listrik. Model ini sangat berguna ketika pola musiman dan tren tidak cukup menjelaskan variasi data, sehingga faktor eksternal perlu dipertimbangkan untuk meningkatkan akurasi peramalan. Hal ini menjadikan SARIMAX lebih adaptif dan akurat dalam memprediksi data kompleks seperti konsumsi energi listrik rumah tangga yang dipengaruhi oleh berbagai faktor luar (Alharbi & Csala, 2022). Secara umum, model SARIMAX dituliskan dalam bentuk:

$$Y_t = \beta X_t + \Phi_P(B^s)\phi_p(B)(1-B)^d(1-B^s)^D Y_t = \Theta_Q(B^s)\theta_q(B)\varepsilon_t \quad (2.3)$$

Keterangan:

Y_t : nilai deret waktu pada periode ke- t

X_t : variabel eksogen pada periode ke- t

β : koefisien pengaruh variabel eksogen

B : operator lag

s : panjang periode musiman

p, d, q : parameter non-musiman (AR, I, MA)

P, D, Q : parameter musiman

ε_t : *error* acak pada periode ke- t

2.2.10 RMSE (*Root Mean Square Error*)

Root Mean Square Error (RMSE) merupakan salah satu metrik evaluasi yang umum digunakan untuk mengukur tingkat kesalahan antara nilai hasil prediksi dengan nilai aktual. RMSE menunjukkan seberapa besar penyimpangan rata-rata prediksi terhadap data sebenarnya semakin kecil nilai RMSE, maka semakin baik kinerja model prediksi (Fauzan et al., 2024). Rumus RMSE dapat dituliskan sebagai berikut:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_t - \hat{y}_t)^2}{n}} \quad (2.4)$$

Keterangan:

y_t : nilai aktual

$\hat{y}_t$: nilai hasil prediksi

n : jumlah data

2.2.11 MAPE (*Mean Absolute Percentage Error*)

Mean Absolute Percentage Error (MAPE) merupakan salah satu metrik evaluasi yang paling umum digunakan untuk menilai akurasi suatu model peramalan. MAPE mengukur seberapa besar rata-rata kesalahan absolut antara nilai aktual dan nilai hasil prediksi dalam bentuk persentase terhadap nilai aktual. Keunggulan utama dari MAPE adalah interpretasinya yang intuitif, karena menunjukkan tingkat kesalahan dalam satuan persen, sehingga mudah dibandingkan antar model. Namun, MAPE dapat menjadi kurang akurat apabila terdapat nilai aktual yang mendekati nol, karena akan menyebabkan nilai *error* menjadi sangat besar. Semakin kecil nilai

MAPE yang dihasilkan, semakin baik tingkat akurasi model prediksi tersebut (Alharbi & Csala, 2022). Rumus MAPE dapat dituliskan sebagai berikut:

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (2.5)$$

Keterangan:

A_t : nilai aktual pada periode ke- t

F_t : nilai hasil prediksi pada periode ke- t

n : jumlah total data pengamatan

2.2.12 MAE (*Mean Absolute Error*)

Mean Absolute Error (MAE) merupakan salah satu metrik evaluasi yang digunakan untuk mengukur tingkat kesalahan rata-rata absolut antara nilai aktual dan nilai hasil prediksi. MAE menunjukkan seberapa besar rata-rata deviasi hasil prediksi terhadap nilai sebenarnya tanpa memperhatikan arah kesalahan (positif atau negatif). Nilai MAE yang lebih kecil menandakan bahwa hasil prediksi semakin mendekati nilai aktual, sehingga model dianggap memiliki tingkat akurasi yang baik (Elshewey et al., 2023). Rumus MAE dapat ditulis sebagai berikut:

$$MAE = \frac{1}{n} \sum_{t=1}^n |A_t - F_t| \quad (2.6)$$

Keterangan:

A_t : nilai aktual pada periode ke- t

F_t : nilai hasil prediksi pada periode ke- t

n : jumlah total data pengamatan

2.2.13 *Maximum Likelihood Estimation* (MLE)

Maximum Likelihood Estimation (MLE) merupakan metode statistik yang digunakan untuk mengestimasi parameter model dengan cara memaksimalkan nilai *likelihood function*, yaitu probabilitas munculnya data yang diamati berdasarkan parameter model tertentu. Dalam konteks pemodelan deret waktu seperti SARIMAX, MLE digunakan untuk mencari kombinasi parameter yang menghasilkan kemungkinan terbesar terhadap data aktual yang diamati.

MLE berperan penting dalam menentukan parameter terbaik model agar hasil prediksi lebih akurat dan stabil. Menurut Zhou (2021), penggunaan metode MLE pada model deret waktu membantu meminimalkan kesalahan estimasi dengan menyeimbangkan kompleksitas model dan kemampuan model dalam menangkap pola data yang sesungguhnya.

Dalam proses estimasi menggunakan MLE, kriteria penilaian model umumnya dilihat melalui tiga ukuran utama, yaitu:

1. *Log-Likelihood* menunjukkan tingkat kecocokan model terhadap data. Semakin tinggi nilainya, semakin baik model tersebut menggambarkan data.
2. *Akaike Information Criterion* (AIC) digunakan untuk menilai keseimbangan antara tingkat kecocokan dan kompleksitas model. Semakin kecil nilainya, semakin baik model tersebut.
3. *Bayesian Information Criterion* (BIC) mirip dengan AIC namun memberikan penalti lebih besar pada model dengan jumlah parameter yang banyak. Nilai BIC yang lebih kecil menunjukkan model yang lebih efisien dan stabil.

2.2.14 Flowchart

Flowchart merupakan diagram yang digunakan untuk menggambarkan alur atau langkah-langkah dalam suatu proses secara sistematis dan terstruktur. Setiap simbol pada flowchart mewakili jenis aktivitas tertentu, seperti proses, keputusan, input, atau output, sehingga membantu memahami logika dan urutan kerja suatu sistem. Dalam penelitian, flowchart berfungsi sebagai alat visual untuk memudahkan perancangan, analisis, dan komunikasi alur kerja dari tahap awal hingga akhir (Fauzan et al., 2024).

	Flow Direction symbol Yaitu simbol yang digunakan untuk menghubungkan antara simbol yang satu dengan simbol yang lain. Simbol ini disebut juga connecting line.		Simbol Manual Input Simbol untuk pemasukan data secara manual on-line keyboard
	Terminator Symbol Yaitu simbol untuk permulaan (start) atau akhir (stop) dari suatu kegiatan		Simbol Preparation Simbol untuk mempersiapkan penyimpanan yang akan digunakan sebagai tempat pengolahan di dalam storage.
	Connector Symbol Yaitu simbol untuk keluar - masuk atau penyambungan proses dalam lembar / halaman yang sama.		Simbol Predefine Proses Simbol untuk pelaksanaan suatu bagian (sub-program)/prosedure
	Connector Symbol Yaitu simbol untuk keluar - masuk atau penyambungan proses pada lembar / halaman yang berbeda.		Simbol Display Simbol yang menyatakan peralatan output yang digunakan yaitu layar, plotter, printer dan sebagainya.
	Processing Symbol Simbol yang menunjukkan pengolahan yang dilakukan oleh komputer		Simbol disk and On-line Storage Simbol yang menyatakan input yang berasal dari disk atau disimpan ke disk.
	Simbol Manual Operation Simbol yang menunjukkan pengolahan yang tidak dilakukan oleh computer		Simbol magnetik tape Unit Simbol yang menyatakan input berasal dari pita magnetik atau output disimpan ke pita magnetik.
	Simbol Decision Simbol pemilihan proses berdasarkan kondisi yang ada.		Simbol Punch Card Simbol yang menyatakan bahwa input berasal dari kartu atau output ditulis ke kartu
	Simbol Input-Output Simbol yang menyatakan proses input dan output tanpa tergantung dengan jenis peralatannya		Simbol Dokumen Simbol yang menyatakan input berasal dari dokumen dalam bentuk kertas atau output dicetak ke kertas.

Gambar 2. 2 Simbol-Simbol Flowchart

2.2.15 Python

Python merupakan bahasa pemrograman tingkat tinggi yang banyak digunakan dalam pengembangan perangkat lunak, analisis data, kecerdasan buatan, serta pemodelan statistik. Python memiliki sintaks yang sederhana dan mudah dipahami, sehingga memudahkan pengguna dalam mengembangkan berbagai aplikasi, termasuk dalam bidang *machine learning* dan *forecasting*. Selain itu, Python dilengkapi dengan berbagai pustaka (*library*) seperti *NumPy*, *Pandas*, *Matplotlib*, dan *Statsmodels* yang sangat membantu dalam pengolahan data dan pembuatan model prediksi berbasis *time series* (Afandi et al., 2024).

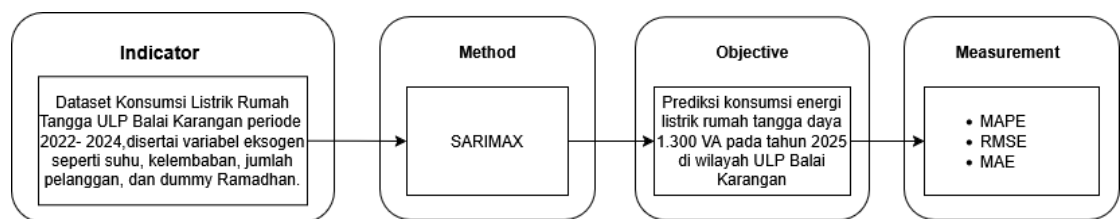
2.2.16 Google Colab

Google Colab merupakan platform berbasis cloud yang disediakan oleh Google untuk menjalankan kode Python secara daring tanpa perlu melakukan instalasi perangkat lunak tambahan. Platform ini sangat populer di kalangan peneliti dan praktisi data karena mendukung komputasi berbasis *machine learning* serta pemrosesan data dalam skala besar dengan dukungan GPU dan TPU. Selain itu, Google Colab memungkinkan kolaborasi *real-time* antar pengguna serta integrasi

langsung dengan Google Drive untuk penyimpanan dan pengelolaan data (Maleki et al., 2024).

2.3 Kerangka Pemikiran

Kerangka pemikiran merupakan bagian yang menjelaskan alur logis dan konseptual dari penelitian. Ibaratnya adalah “peta pikiran” yang menggambarkan keterkaitan antara teori, data, metode, dan tujuan penelitian dalam menjawab rumusan masalah. Bagian ini menjadi dasar untuk memahami arah penelitian secara sistematis dan terstruktur.



Gambar 2. 3 Kerangka Pemikiran

Gambar 2.3 memperlihatkan kerangka pemikiran dari penelitian ini, yang menggambarkan tahapan sistematis mulai dari pengumpulan data hingga evaluasi hasil prediksi. Setiap komponen saling terhubung dan berperan penting dalam mencapai tujuan penelitian. Berikut penjelasan dari masing-masing komponennya:

1. *Indicator*

Indikator merupakan elemen utama yang menjadi dasar pengolahan data dalam penelitian. Pada penelitian ini, indikator yang digunakan adalah data historis konsumsi listrik rumah tangga daya 1300 VA wilayah ULP Balai Karangan, yang dikumpulkan secara bulanan dari tahun 2022 hingga 2024. Selain itu, penelitian ini juga melibatkan variabel eksogen seperti suhu, kelembaban, jumlah pelanggan, dan *dummy* Ramadhan yang diharapkan dapat memberikan pengaruh terhadap fluktuasi konsumsi listrik masyarakat. Data historis ini menjadi landasan dalam membangun model prediksi yang lebih akurat dan kontekstual terhadap kondisi sosial masyarakat setempat.

2. *Method*

Metode yang digunakan dalam penelitian ini adalah SARIMAX (*Seasonal Autoregressive Integrated Moving Average with Exogenous Variables*), yaitu

pengembangan dari model ARIMA dan SARIMA yang memungkinkan penambahan variabel eksternal (eksogen). Model ini mampu menangkap pola musiman (*seasonal*), *trend*, serta pengaruh variabel eksternal yang memengaruhi konsumsi listrik. Proses pemodelan dilakukan melalui beberapa tahap utama: identifikasi pola data, penentuan parameter (p, d, q, P, D, Q, s) , pelatihan model, serta validasi hasil prediksi menggunakan data uji.

3. *Objective*

Tujuan utama dari penelitian ini adalah memprediksi konsumsi energi listrik rumah tangga daya 1300 VA pada tahun 2025 di wilayah ULP Balai Karangan. Model prediksi diharapkan dapat membantu PT PLN ULP Balai Karangan dalam menyusun strategi perencanaan kapasitas, distribusi energi, serta pengelolaan sumber daya listrik secara lebih efisien. Selain itu, model ini juga dapat menjadi dasar dalam pengambilan keputusan operasional, terutama untuk mengantisipasi periode puncak konsumsi energi seperti bulan Ramadhan.

4. *Measurement*

Evaluasi terhadap hasil prediksi dilakukan menggunakan beberapa metrik evaluasi akurasi, yaitu:

- a. RMSE (*Root Mean Square Error*) digunakan untuk mengukur rata-rata kesalahan kuadrat antara nilai aktual dan nilai prediksi. RMSE lebih sensitif terhadap kesalahan besar sehingga cocok untuk mengukur stabilitas model.
- b. MAPE (*Mean Absolute Percentage Error*) menghitung rata-rata kesalahan absolut dalam bentuk persentase terhadap nilai aktual, memberikan gambaran akurasi relatif model dalam konteks operasional.
- c. MAE (*Mean Absolute Error*) mengukur rata-rata selisih absolut antara nilai aktual dan nilai prediksi, dengan hasil yang mudah diinterpretasikan. Nilai MAE yang kecil menunjukkan model memiliki performa prediksi yang baik.