Prediksi Produksi Minyak Mentah *Gulf Coast* (PADD 3) Menggunakan Algoritma *Prophet*

Muhammad Fattahul Falah Ehzra Candra[1*], Diyah Rosiani[1], Arya Dwi Candra[1]
[1]Teknik Produksi Minyak dan Gas, Politeknik Energi dan Mineral Akamigas, Jl. Gajah Mada No. 38, Mentul, Karangboyo, Kec. Cepu, Kabupaten Blora, Jawa Tengah 58315, Indonesia

ABSTRAK

Gulf Coast merupakan salah satu penyumbang terbesar produksi minyak mentah Amerika Serikat sehingga peramalan produksinya penting bagi perencanaan energi dan keputusan investasi. Penelitian ini menerapkan algoritma *Prophet* untuk memprediksi produksi minyak mentah bulanan *Gulf Coast* periode Januari 1981 hingga Mei 2026 dari *U.S. Energy Information Administration*, dengan pembagian 85% data latih dan 15% data uji. Model menggunakan musiman tahunan bermode multiplikatif dan dievaluasi dengan *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), *Mean Absolute Percentage Error* (MAPE), dan koefisien determinasi (R^2). Pada data uji diperoleh MAE 13.811,90 ribu barel/bulan, RMSE 16.987,10 ribu barel/bulan, MAPE 5,35%, dan R^2 0,6702. *Prophet* mengungguli ARIMA dan SARIMA yang menghasilkan R^2 negatif, serta menghasilkan MAPE 8,33% pada validasi *rolling cross-validation*. Proyeksi Juni 2026 hingga Mei 2027 menunjukkan tren meningkat dengan rata-rata 327.895,38 ribu barel/bulan. Analisis residual menunjukkan kecenderungan prediksi berlebih dan *overfitting* ringan, sehingga penelitian selanjutnya disarankan menambahkan variabel eksogen.

Kata Kunci: Deret waktu, *Gulf Coast*, Peramalan, Produksi minyak mentah, *Prophet*

alipcandra29@gmail.com

1. Pendahuluan

Dalam sektor minyak dan gas, produksi minyak mentah menjadi salah satu indikator utama yang memengaruhi perencanaan energi, pengembangan infrastruktur, serta proses pengambilan keputusan. Salah satu kawasan penghasil minyak yang memegang peranan penting di Amerika Serikat adalah *Gulf Coast*, atau *Petroleum Administration for Defense District* 3 (PADD 3), yang meliputi wilayah Alabama, Arkansas, Louisiana, Mississippi, New Mexico, dan Texas [1]. Tingkat produksi di kawasan ini terus mengalami perubahan seiring berjalannya waktu, sehingga dibutuhkan pendekatan peramalan yang dapat mengenali pola historis sekaligus kecenderungan produksi ke depannya.

Pendekatan peramalan (*forecasting*) pada data deret waktu telah luas digunakan dalam berbagai analisis di sektor energi. Metode statistik konvensional seperti ARIMA umum dipakai untuk memodelkan pola historis, sementara kemunculan metode *machine learning* menawarkan alternatif dalam menangani data dengan pola yang lebih kompleks. Sebagai contoh, Aziz dan Barawi menggunakan kombinasi ARIMA dan *Facebook Prophet* untuk meramalkan harga minyak mentah [2], sedangkan Ma menerapkan *Fbprophet* [3] dan Bontemps dan Ohnemus menggunakan *machine learning* [4] untuk peramalan harga minyak mentah, yang memperlihatkan bahwa pendekatan berbasis data relevan diterapkan dalam *forecasting* komoditas minyak. *Prophet*

sendiri merupakan model peramalan deret waktu hasil pengembangan Taylor dan Letham, yang bekerja dengan menggabungkan tiga komponen utama, yaitu tren, musiman, dan perubahan tren (*changepoint*) [5]. Karakteristik ini menjadikan *Prophet* cocok digunakan pada deret waktu yang memiliki pola musiman sekaligus perubahan tren, termasuk pada data dengan struktur historis yang kompleks. Pada penelitian ini, *Prophet* diterapkan untuk memodelkan data produksi minyak mentah bulanan *Gulf Coast* (PADD 3) sekaligus menghasilkan proyeksi produksi untuk periode-periode berikutnya.

Prophet telah diterapkan pada peramalan produksi hidrokarbon. Ning dkk. [6] membandingkan *Prophet*, ARIMA, dan LSTM pada produksi minyak satu sumur dan 65 sumur di *Denver-Julesburg Basin*. Pada satu sumur, *Prophet* menghasilkan RMSE terendah dan mampu menangkap fluktuasi musiman, tetapi pada 65 sumur ARIMA dan LSTM umumnya lebih unggul. Pada produksi gas sebuah lapangan lepas pantai, Satria dkk. [7] melaporkan *Prophet* lebih akurat daripada ARIMA. Kedua penelitian tersebut dilakukan pada level sumur atau lapangan dengan rentang data beberapa tahun. Sepengetahuan penulis, penelitian pada level regional dengan data bulanan jangka panjang (45 tahun) yang mencakup perubahan struktural, seperti *booming* minyak serpih (*shale*) dan pandemi COVID-19, masih terbatas. Penelitian ini mengisi celah tersebut dengan menerapkan *Prophet* pada produksi minyak mentah PADD 3, membandingkannya dengan ARIMA dan SARIMA sebagai *baseline* pada horizon uji 82 bulan, memvalidasinya dengan *rolling cross-validation*, serta mendiagnosis keterbatasan model melalui analisis residual dan komponen musiman.

2. Metodologi

Penelitian ini dirancang sebagai studi non-eksperimental dengan pendekatan *forecast-ing* deret waktu univariat. Variabel yang di-kaji adalah produksi minyak mentah bulanan di kawasan regional *Gulf Coast* (PADD 3), dengan satuan ribuan barel per bulan. Sifat observasional pada desain ini tercermin dari tidak digunakannya data teknis seperti data produksi rinci, *reservoir*, temperatur, maupun tekanan; penelitian ini murni memanfaatkan data historis yang tersedia dengan menerapkan algoritma *forecasting*. Pendekatan tersebut dinilai tepat mengingat produksi minyak mentah secara historis dapat dipandang sebagai suatu proses deret waktu yang memiliki komponen tren dan musiman [5].

2.1 Bahan

Data yang digunakan dalam penelitian ini berupa deret waktu bulanan produksi minyak mentah di kawasan *Gulf Coast* (PADD 3), bersumber dari *U.S. Energy Information Administration* (EIA) [1], dengan cakupan periode Januari 1981 hingga Mei 2026 (total 545 observasi bulanan). *Dataset* tersebut diunduh dalam format CSV, kemudian diimpor ke dalam *Google Colab* menggunakan pustaka *pandas*. Seluruh proses eksperimen dilaksanakan menggunakan *Google Colab*, sebuah *platform* komputasi berbasis *cloud* yang memungkinkan pengguna menjalankan kode *Python* beserta akses ke berbagai pustaka statistik dan *machine learning*. Pemanfaatan *Google Colab* memberikan sejumlah keuntungan, di antaranya integrasi langsung dengan penyimpa-nan *Google Drive* serta kemudahan dalam memvisualisasikan hasil dari kode yang dijalankan [8].

2.2 Metode

2.2.1 Pengolahan dan Pembagian Data

Proses pengolahan data dilakukan melalui tahapan berikut:

1. *Dataset* dibaca menggunakan fungsi *read_csv()* dengan parameter *skiprows=4* untuk melewati bagian metadata pada awal *file*.
2. Nama kolom disesuaikan menjadi *ds* untuk tanggal dan *y* untuk nilai produksi, mengikuti format data yang disyaratkan oleh *Prophet*.
3. Kolom tanggal dikonversi ke format *datetime* menggunakan pola *%b-%y*.
4. Data diurutkan secara kronologis, dari periode paling lama hingga paling baru.
5. Pemeriksaan *missing values* dilakukan untuk memastikan tidak ada data yang hilang sebelum tahap pemodelan.

6. Data kemudian dibagi secara kronologis menjadi 85% data latih (*training set*) dan 15% data uji (*testing set*). Dari total 545 observasi bulanan, sebanyak 463 data digunakan sebagai *training set* hingga Juli 2019, sementara 82 data sisanya menjadi *testing set* mulai Agustus 2019 hingga Mei 2026.

Pembagian data secara kronologis ini dilakukan untuk menjaga urutan temporalnya. Periode data yang lebih awal digunakan untuk membangun (melatih) model, sementara periode setelahnya berfungsi untuk menguji sejauh mana model mampu menghasilkan prediksi terhadap data baru yang belum pernah dilihat selama proses pelatihan.

Rasio 85% data latih dan 15% data uji dipilih karena data deret waktu membutuhkan riwayat latih yang panjang agar komponen tren, perubahan tren (*changepoint*), dan musiman dapat diestimasi dengan baik, sementara data uji tetap cukup panjang (82 bulan) untuk mencakup berbagai kondisi, termasuk periode pandemi COVID-19 dan pemulihannya. Rasio ini mendekati praktik umum pembagian data latih-uji sekitar 70/30 hingga 80/20 [9]. Secara teoretis, Gholamy dkk. [9] menunjukkan bahwa proporsi data latih sekitar 80% atau lebih menjaga estimasi akurasi tidak melebihi akurasi sebenarnya, sehingga proporsi 85% masih berada pada rentang tersebut. Proporsi data latih yang sedikit lebih besar ini juga memberi riwayat yang lebih panjang untuk mengestimasi komponen tren dan perubahan tren, sejalan dengan rekomendasi data uji sekitar 20% dari total sampel pada evaluasi peramalan [10]. Selain itu, *rolling cross-validation* digunakan untuk mengurangi ketergantungan hasil pada satu skema pembagian [11].

2.2.2 Model dan Algoritma *Prophet*

Prophet adalah model *forecasting* deret waktu hasil pengembangan Taylor dan Letham, yang dirancang khusus untuk menangani data dengan pola tren, musiman, serta perubahan tren (*changepoint*) [5]. Model ini merepresentasikan deret waktu ke dalam tiga komponen utama, yaitu tren, musiman, dan komponen kesalahan (*error*), sehingga cukup sesuai diterapkan pada data produksi yang karakteristiknya terus berubah dari waktu ke waktu:

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t \quad (1)$$

Pada penelitian ini, model *Prophet* dikonfigurasi dengan parameter sebagai berikut:

1. *yearly_seasonality = True*
2. *weekly_seasonality = False*
3. *daily_seasonality = False*
4. *seasonality_mode = 'multiplicative'*

Parameter *yearly_seasonality* diaktifkan mengingat data yang digunakan berfrekuensi bulanan, sehingga berpotensi menunjukkan pola musiman tahunan. Sebaliknya, *weekly_seasonality* dan *daily_seasonality* tidak diaktifkan karena unit observasi penelitian ini bersifat bulanan, bukan mingguan maupun harian. Adapun mode musiman *multiplicative* dipilih dengan asumsi bahwa variasi musiman turut berubah mengikuti tingkat produksi, sehingga amplitudo fluktuasi musiman dapat membesar atau mengecil seiring perubahan level deret waktu. Adapun *order Fourier* untuk komponen musiman tahunan menggunakan nilai bawaan *Prophet* [10], yang berpotensi memengaruhi kehalusan kurva musiman.

Parameter *changepoint_prior_scale* tidak diubah dan tetap menggunakan nilai bawaan (*default*) *Prophet*, yaitu 0,05. Selanjutnya, model dilatih menggunakan *training set* melalui fungsi *fit()*. Setelah tahap pelatihan selesai, prediksi terhadap periode *testing set* dilakukan dengan memanfaatkan fungsi *make_future_dataframe()* dan *predict()*. Selanjutnya, model akhir dilatih ulang dengan memanfaatkan seluruh data historis, yaitu periode Januari 1981 hingga Mei 2026. Model tersebut kemudian digunakan untuk menghasilkan proyeksi produksi selama 12 bulan berikutnya, mencakup periode Juni 2026 hingga Mei 2027.

2.2.3 Evaluasi Model dan Ukuran Akurasi

Kinerja model dievaluasi melalui empat metrik, yaitu *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), *Mean Absolute*

Percentage Error (MAPE), dan koefisien determinasi (R^2). Penggunaan beberapa metrik sekaligus dimaksudkan agar diperoleh gambaran yang lebih menyeluruh mengenai tingkat kesalahan model sekaligus kemampuannya dalam menjelaskan variasi data yang ada:

- *Mean Absolute Error* (MAE)

digunakan untuk mengukur besarnya rata-rata penyimpangan absolut antara nilai aktual dan nilai hasil prediksi [12]:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (2)$$

- *Root Mean Squared Error* (RMSE)

memberikan bobot yang lebih besar terhadap kesalahan (*error*) berukuran besar, sehingga metrik ini berguna untuk menilai tingkat stabilitas model:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (3)$$

- *Mean Absolute Percentage Error* (MAPE)

merepresentasikan besarnya kesalahan relatif dalam bentuk persentase, sehingga relatif mudah diinterpretasikan, meskipun metrik ini cukup sensitif terhadap nilai nol pada data aktual:

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (4)$$

- Koefisien Determinasi (R^2)

merepresentasikan seberapa besar proporsi variabilitas data yang mampu dijelaskan oleh model, dibandingkan dengan total variansi data secara keseluruhan [10]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (5)$$

Pada Persamaan (2) hingga Persamaan (5), n menyatakan jumlah total data pengamatan, y_i menyatakan nilai aktual pada data ke-i, $\hat{y}_i$ menyatakan nilai prediksi pada data ke-i, $\bar{y}$ menyatakan rata-rata nilai aktual, $|y_i - \hat{y}_i|$ menyatakan selisih absolut antara nilai aktual dan prediksi, $(y_i - \hat{y}_i)^2$ menyatakan kuadrat selisih antara nilai aktual dan prediksi, serta $|(y_i - \hat{y}_i)/y_i|$ menyatakan persentase kesalahan absolut pada setiap titik data.

2.2.4 Prosedur *Forecasting*

Secara keseluruhan, tahapan penelitian ini terdiri atas beberapa proses utama, meliputi pengumpulan *dataset* dari EIA, tahap *preprocessing*, pembagian data menjadi *training set* dan *testing set*, pembangunan model *Prophet*, prediksi terhadap *testing set*, evaluasi kinerja model menggunakan MAE, RMSE, MAPE, dan R^2, hingga pembangunan model akhir dengan memanfaatkan seluruh data historis yang tersedia. Model akhir tersebut kemudian digunakan untuk menghasilkan proyeksi produksi selama 12 bulan setelah periode observasi terakhir. Hasil prediksi ditampilkan dalam bentuk nilai estimasi produksi (*yhat*), yang dilengkapi dengan batas bawah dan batas atas interval prediksi (*yhat_lower* dan *yhat_upper*).

2.2.5 Perbandingan dengan Model *Baseline*

Untuk memperkuat validitas hasil, performa model *Prophet* dibandingkan dengan dua model peramalan konvensional, yaitu ARIMA dan SARIMA, yang diuji pada data latih dan data uji yang sama. Pemilihan *order* ARIMA dilakukan dengan mengevaluasi beberapa kombinasi parameter (*p, d, q*) dan memilih konfigurasi dengan nilai *Akaike Information Criterion* (AIC) terendah, yaitu ARIMA(1,1,0). Sebagai pembanding tambahan yang turut memperhitungkan komponen musiman, digunakan pula model SARIMA(1,1,1)(1,1,1,12) dengan periode musiman 12 bulan. Berbeda dengan ARIMA, *order* SARIMA ditetapkan secara tetap berdasarkan konfigurasi umum untuk data bulanan dan tidak dioptimasi melalui AIC, sehingga berfungsi sebagai pembanding yang mewakili spesifikasi musiman standar. Kedua model dilatih menggunakan pustaka *statsmodels* pada *Python* dan dievaluasi menggunakan metrik yang sama dengan model *Prophet* (MAE, RMSE, MAPE, dan R^2). ARIMA dan SARIMA dipilih sebagai model pembanding (*baseline*) karena merupakan metode statistik standar dan paling umum digunakan dalam peramalan deret waktu [13].

Kinerja yang lebih rendah dari kedua model tersebut merupakan hasil penelitian ini, bukan asumsi awal. Pembandingan dengan *baseline* konvensional diperlukan untuk menunjukkan apakah model yang lebih kompleks seperti *Prophet* benar-benar memberikan peningkatan kinerja.

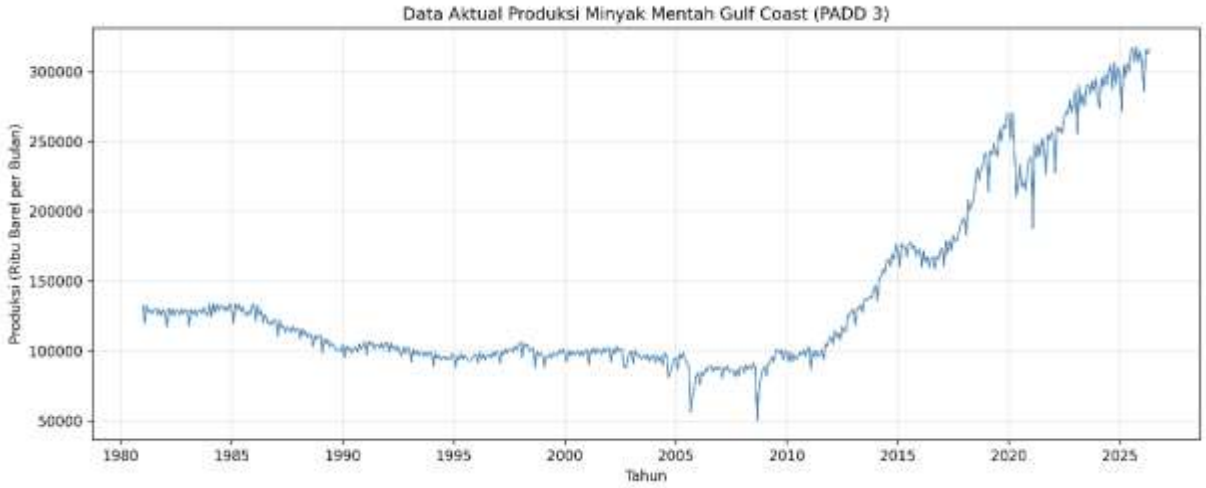
2.2.6 Validasi dengan *Rolling Cross-Validation*

Selain evaluasi menggunakan skema pembagian data tunggal (*single split*), dilakukan pula validasi tambahan menggunakan *rolling cross-validation* bawaan *Prophet* melalui fungsi *cross_validation()* dari modul *prophet.diagnostics*. Skema validasi ini menggunakan periode pelatihan awal (*initial*) selama 30 tahun, dengan horizon prediksi masing-masing 12 bulan dan pergeseran (*period*) setiap 12 bulan, menghasilkan total 15 titik potong (*cutoff*) yang tersebar dari tahun 2011 hingga 2025. Pendekatan ini memungkinkan evaluasi performa model pada berbagai periode waktu yang berbeda sehingga memberikan gambaran yang lebih representatif dibandingkan dengan evaluasi pada satu periode pengujian saja.

3. Hasil

3.1 Karakteristik Data Produksi

Gambar 1 memperlihatkan perkembangan produksi minyak mentah bulanan di kawasan *Gulf Coast* (PADD 3) sepanjang periode Januari 1981 hingga Mei 2026. Terlihat bahwa tingkat produksi terus mengalami perubahan dari waktu ke waktu, ditandai dengan pola tren jangka panjang serta fluktuasi pada beberapa periode tertentu. Karakteristik semacam ini mengindikasikan bahwa produksi minyak mentah bersifat dinamis, bukan konstan, sehingga diperlukan pendekatan *time series forecasting* untuk mengidentifikasi kecenderungan produksi berdasarkan pola historisnya.



Gambar 1. Data aktual produksi minyak mentah *Gulf Coast* (PADD 3)

3.2 Evaluasi Kinerja Model *Prophet*

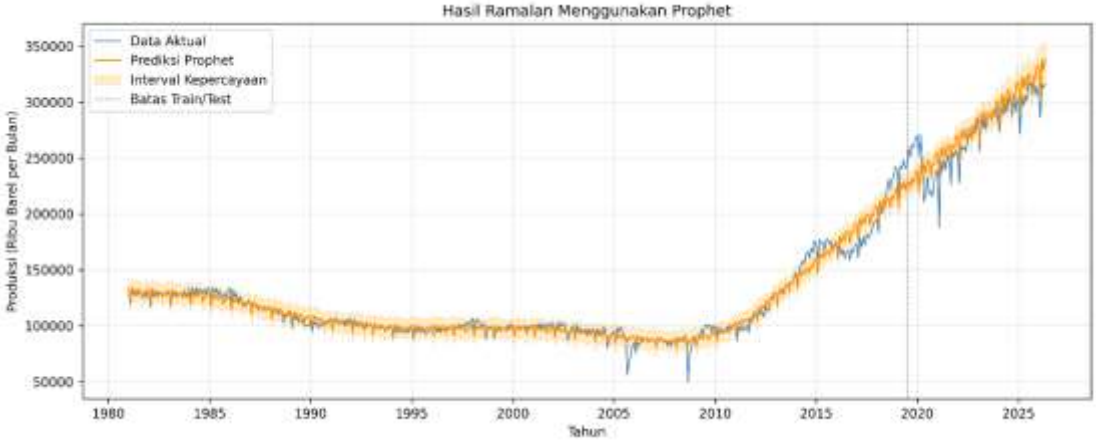
Pengujian model *Prophet* pada periode *testing set* (Agustus 2019 hingga Mei 2026) menghasilkan nilai evaluasi kuantitatif sebagaimana disajikan pada Tabel 1. Nilai MAE tercatat sebesar 13.811,90 ribu barel/bulan, RMSE sebesar 16.987,10 ribu barel/bulan, MAPE sebesar 5,35%, dan koefisien determinasi (R^2) sebesar 0,6702.

Tabel 1. Hasil evaluasi model *Prophet*

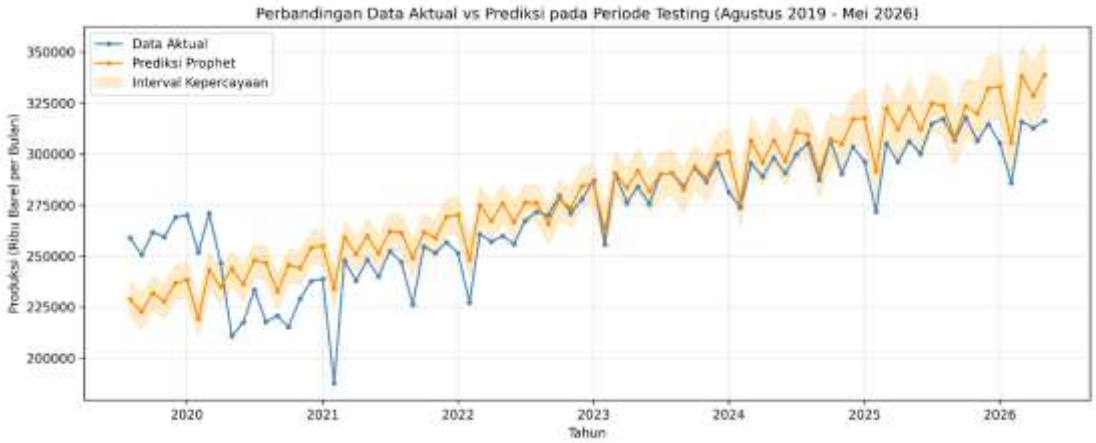
Metrik	Nilai
MAE	13.811,90 ribu barel/bulan
RMSE	16.987,10 ribu barel/bulan
MAPE	5,35%
R^2	0,6702

3.3 Hasil Prediksi terhadap Data Aktual

Gambar 2(a) menampilkan perbandingan antara data aktual dan hasil prediksi yang dihasilkan model *Prophet* pada rentang waktu penuh (Januari 1981–Mei 2026). Secara keseluruhan, hasil prediksi mampu mengikuti kecenderungan pola data aktual, khususnya dalam menggambarkan arah perubahan produksi secara umum. Untuk melihat lebih detail performa model pada periode pengujian, Gambar 2(b) menampilkan perbesaran (*zoom*) khusus pada periode *testing* (Agustus 2019–Mei 2026) beserta pita interval kepercayaan prediksi.



(a)



(b)

Gambar 2. Hasil prediksi terhadap data aktual: (a) perbandingan pada periode penuh (Januari 1981–Mei 2026); (b) perbandingan pada periode *testing* (Agustus 2019–Mei 2026)

3.4 Proyeksi Produksi 12 Bulan

Berdasarkan model akhir yang dilatih ulang menggunakan seluruh data historis (Januari 1981–Mei 2026), dihasilkan proyeksi produksi minyak mentah *Gulf Coast* (PADD 3) selama 12 bulan ke depan (Juni 2026 hingga Mei 2027) sebagaimana dirangkum pada Tabel 2. Hasil proyeksi memperlihatkan kecenderungan peningkatan produksi dengan rata-rata prediksi mencapai 327.895,38 ribu barel per bulan, bergerak dari 315.496,97 ribu barel pada Juni 2026 hingga mencapai 341.106,55 ribu barel pada Mei 2027.

Tabel 2. Proyeksi produksi minyak mentah *Gulf Coast* (PADD 3) periode Juni 2026–Mei 2027

No.	Periode	Prediksi (ribu barel/bulan)	Batas Bawah	Batas Atas
1	Juni 2026	315.496,97	305.540,66	324.916,57
2	Juli 2026	327.409,70	316.723,18	338.171,13
3	Agustus 2026	330.172,27	320.275,56	340.621,69
4	September 2026	317.768,86	306.995,88	328.983,13
5	Oktober 2026	331.319,44	320.320,06	341.589,77
6	November 2026	324.820,22	314.726,31	335.944,98
7	Desember 2026	336.357,20	325.690,95	346.226,34
8	Januari 2027	335.179,70	324.369,00	345.314,42
9	Februari 2027	309.373,85	299.106,92	320.095,06
10	Maret 2027	335.906,25	325.335,08	345.947,55
11	April 2027	329.833,54	318.915,64	340.629,56
12	Mei 2027	341.106,55	329.929,19	352.063,86
Rata-rata		**327.895,38**	**-**	**-**

4. Analisa & Pembahasan

4.1 Interpretasi Kinerja Prediksi dan Faktor Eksternal

Pada beberapa periode di Gambar 1 dan Gambar 2, terlihat perubahan tingkat produksi yang cukup signifikan. Kondisi ini diduga dipengaruhi oleh berbagai faktor eksternal yang tidak secara langsung tercakup dalam model, seperti pergeseran kondisi pasar minyak, kebijakan energi, maupun gangguan ekonomi. Salah satu periode dengan perubahan paling mencolok terjadi sekitar tahun 2020, bertepatan dengan masa pandemi COVID-19 [14] serta perubahan aktivitas ekonomi global secara luas. Faktor-faktor tersebut menjadi salah satu sumber variasi yang tidak sepenuhnya mampu dijelaskan oleh model, mengingat model hanya mengandalkan data historis produksi sebagai input.

Berdasarkan Tabel 1, nilai MAE sebesar 13.811,90 ribu barel/bulan mengindikasikan bahwa rata-rata selisih absolut antara hasil prediksi dan nilai aktual berada di kisaran 13.811,90 ribu barel per bulan. Adapun nilai RMSE (16.987,10 ribu barel/bulan) yang lebih besar dibandingkan MAE menandakan adanya sebagian kesalahan prediksi dengan magnitudo yang cukup besar, mengingat RMSE secara matematis memberikan penalti lebih tinggi terhadap kesalahan berskala besar [15], [12]. Sementara itu, nilai MAPE yang diperoleh sebesar 5,35% menunjukkan bahwa rata-rata kesalahan prediksi, secara relatif, berada pada kisaran 5,35% terhadap nilai aktual. Karena disajikan dalam bentuk persentase, MAPE relatif lebih mudah dipahami untuk menggambarkan tingkat kesalahan relatif dari suatu model [16].

Nilai R^2 sebesar 0,6702 mengindikasikan bahwa model mampu menjelaskan sekitar 67,02% variasi nilai produksi pada data pengujian, sedangkan sisanya, sekitar 32,98%, belum dapat dijelaskan oleh model. Pada Gambar 2(b), tampak adanya beberapa periode di mana nilai aktual dan prediksi saling berselisih, yang mengindikasikan bahwa model belum sepenuhnya mampu menangkap se-

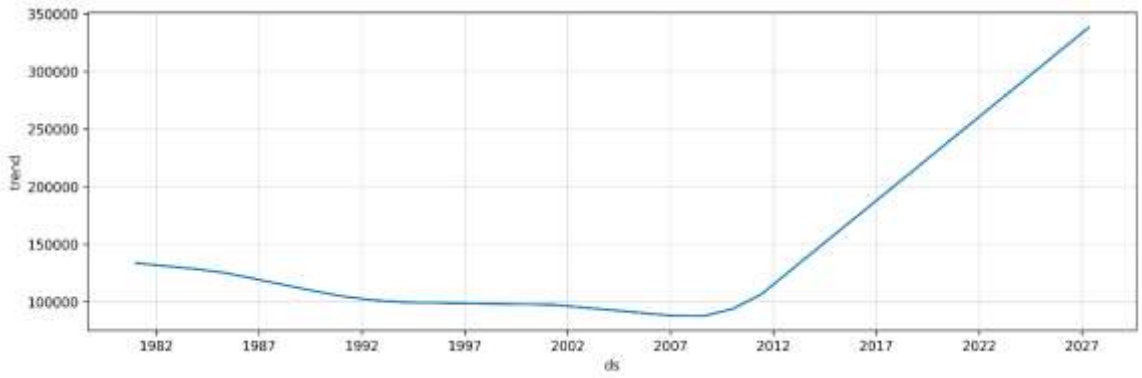
luruh fluktuasi jangka pendek pada data produksi. Kondisi tersebut berkaitan erat dengan karakteristik dasar *Prophet*, yang mengandalkan pola historis berupa tren dan musiman sebagai landasan peramalan. Ketika terjadi perubahan produksi akibat faktor eksternal atau perubahan struktural yang tidak tercakup dalam variabel input, model berpotensi menghasilkan deviasi terhadap nilai aktual. Hal inilah yang menjadi salah satu penyebab nilai R^2 belum mendekati angka 1, sehingga masih terdapat *error* pada hasil prediksi.

Area interval prediksi yang ditampilkan pada Gambar 2(a) dan Gambar 2(b) merepresentasikan rentang ketidakpastian dari hasil peramalan yang dihasilkan. Semakin lebar rentang interval tersebut, semakin besar pula tingkat ketidakpastian prediksi model. Informasi ini penting untuk diperhatikan, sebab hasil *forecasting* tidak hanya menghasilkan satu nilai estimasi tunggal, melainkan juga memberikan gambaran mengenai rentang kemungkinan nilai produksi yang dapat terjadi. Demikian pula pada hasil proyeksi 12 bulan ke depan (Tabel 2), kecenderungan peningkatan produksi menandakan bahwa model tetap mempertahankan pola tren yang telah dipelajari dari data historis hingga memasuki periode proyeksi. Meskipun demikian, hasil proyeksi tersebut sebaiknya dimaknai sebagai estimasi yang didasarkan pada pola historis, bukan sebagai kepastian mutlak tentang realisasi produksi di masa mendatang. Perubahan kondisi pasar, kebijakan energi, dinamika aktivitas ekonomi, maupun faktor-faktor lain yang tidak dimasukkan sebagai variabel input, berpotensi menyebabkan realisasi produksi berbeda dari hasil proyeksi yang telah dihasilkan.

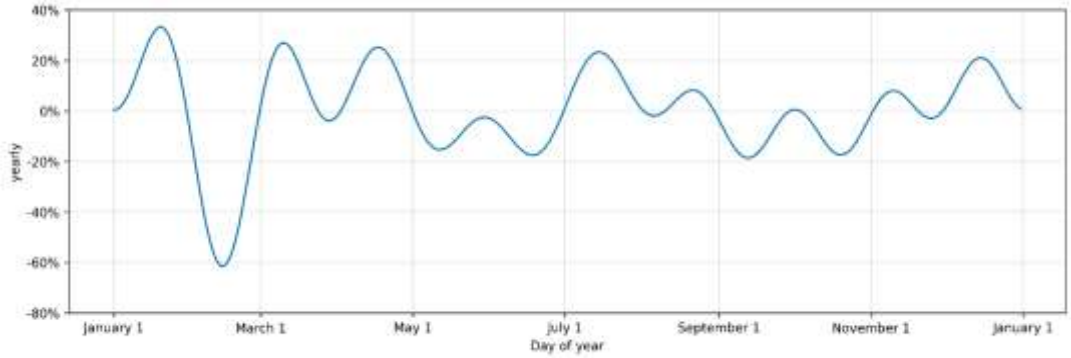
4.2 Analisis Komponen Model *Prophet*

Gambar 3(a) memperlihatkan komponen tren yang dihasilkan oleh model *Prophet*, menunjukkan adanya penurunan produksi hingga sekitar tahun 2007–2008, yang kemudian diikuti oleh peningkatan signifikan hingga akhir periode pengamatan. Pola inilah yang mendasari kecenderungan peningkatan pada hasil proyeksi 12 bulan, dengan rata-rata prediksi sebesar 327.895,38 ribu barel per bulan. Sementara itu, Gambar 3(b) menampilkan komponen *yearly seasonality*, yang memperlihatkan pola produksi bervariasi antarbulan dengan fluktuasi positif dan negatif yang cukup kuat.

Rentang amplitudo kurva komponen musiman tahunan pada Gambar 3(b) tergolong ekstrem, yaitu −62,1% (14 Februari) hingga +33,5% (21 Januari). Namun, pada titik-titik bulanan yang benar-benar diobservasi, rentang yang relevan terhadap hasil prediksi jauh lebih sempit, yakni −10,6% hingga +3,9%. Kesenjangan ini mengindikasikan *order* musiman tahunan (*yearly_seasonality*) bawaan *Prophet* (10) kemungkinan terlalu besar untuk data bulanan yang hanya memiliki 12 observasi per tahun, sehingga kurva musiman berosilasi tajam di sela-sela titik data aktual akibat *overfitting*. Meski demikian, arah polanya tetap stabil: Februari secara berulang menjadi titik terendah sepanjang 1981–2026, bukan hanya pada 2020, sehingga bukan sekadar artefak dari pandemi COVID-19. Yang bersifat artefak hanyalah besarnya osilasi pada kurva kontinu, bukan arah polanya. Ini menjadi keterbatasan yang perlu diperhatikan pada penelitian berikutnya, misalnya dengan mengecilkan *yearly_seasonality* ke kisaran 3–6 agar kurva musiman lebih representatif terhadap resolusi data bulanan. Penerapan *seasonality mode multiplicative* memungkinkan pengaruh musiman turut berubah mengikuti tingkat produksi, sehingga variasi musiman secara absolut dapat semakin membesar seiring meningkatnya level produksi. Temuan ini menunjukkan bahwa *Prophet* mampu menangkap pola produksi melalui kombinasi tren jangka panjang dan pola musiman tahunan, meskipun perubahan yang dipengaruhi oleh faktor eksternal belum dapat direpresentasikan secara langsung dalam model [5].



(a)



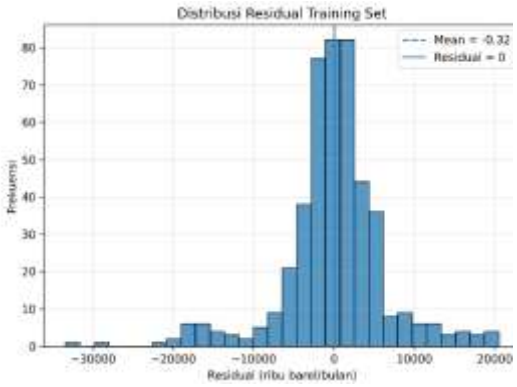
(b)

Gambar 3. Dekomposisi komponen model *Prophet*: (a) komponen tren; (b) komponen musiman tahunan

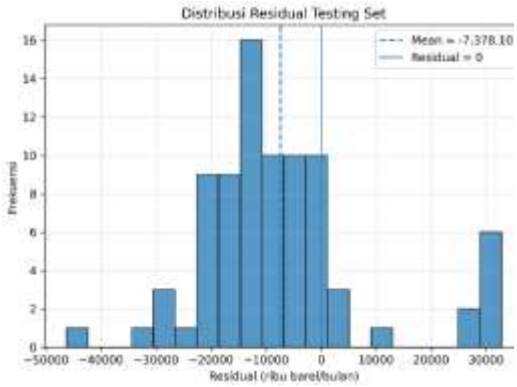
4.3 Analisis Residual dan *Actual vs Predicted*

Gambar 4(a) dan Gambar 4(b) menampilkan distribusi residual pada *training set* dan *testing set*. Pada *training set*, residual memiliki rata-rata sebesar −0,32 ribu barel/bulan dengan sebaran yang relatif sempit di sekitar nol. Sebaliknya, pada *testing set*, residual memiliki rata-rata −7.378,10 ribu barel/bulan dengan sebaran yang jauh lebih lebar, mengindikasikan adanya kecenderungan *overprediction*. Kondisi ini juga tampak pada Gambar 4(c) dan Gambar 4(d), yang menampilkan hubungan *Actual vs Predicted* untuk *training set* dan *testing set*.

Pada Gambar 4(c), titik-titik data relatif berhimpit dengan garis *Perfect Prediction* dengan nilai R^2 mencapai 0,9617 — mengindikasikan bahwa model mampu merepresentasikan pola historis dengan sangat baik. Sementara itu, pada Gambar 4(d), banyak titik prediksi berada di atas garis *Perfect Prediction*, dan nilai R^2 menurun menjadi 0,6702 pada data pengujian, menunjukkan bahwa kemampuan generalisasi model tidak sebaik kemampuannya dalam menyesuaikan data *training*. Kesenjangan yang cukup besar antara R^2 *training* (0,9617) dan R^2 *testing* (0,6702) ini meru-pakan indikasi terjadinya *overfitting* ringan, di mana model cenderung terlalu menyesuaikan diri dengan pola spesifik pada data *training* sehingga kurang optimal dalam menggeneralisasi ke data baru. Kendati demi-kian, model tetap mampu menjelaskan seki-tar 67,02% variasi produksi pada data pengujian. Perbedaan performa antara *training* dan *testing* ini mengindikasikan bahwa *Prophet* cukup baik dalam menangkap pola utama produksi, namun belum sepenuhnya mampu merepresentasikan fluktuasi jangka pendek maupun perubahan-perubahan yang tidak tercermin dalam pola historis.



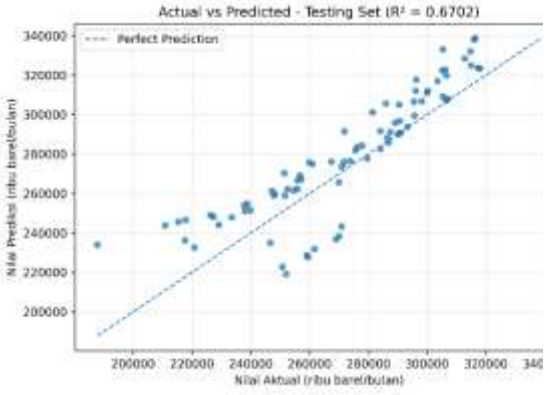
(a) Distribusi residual *training set*



(b) Distribusi residual *testing set*



(c) *Actual vs predicted* pada *training set*



(d) *Actual vs predicted* pada *testing set*

Gambar 4. Analisis residual dan hubungan nilai aktual dengan nilai prediksi pada *training set* dan *testing set*

4.4 Perbandingan dengan Model *Baseline*

Untuk memperkuat validitas performa model yang digunakan, penelitian ini juga membandingkan hasil *Prophet* dengan dua model *baseline* konvensional, yaitu ARIMA (1,1,0) dan SARIMA (1,1,1)(1,1,1,12), yang diuji pada data dan periode pengujian yang sama. Sebagaimana dirangkum pada Tabel 3, hasil perbandingan menunjukkan bahwa kedua model ARIMA memiliki performa yang jauh lebih rendah dibandingkan *Prophet*, dengan ARIMA menghasilkan MAPE sebesar 13,07% (R^2 = −1,1499) dan SARIMA menghasilkan MAPE sebesar 20,83% (R^2 = −3,3074), sedangkan *Prophet* menghasilkan MAPE sebesar 5,35% (R^2 = 0,6702).

Nilai R^2 negatif pada kedua model ARIMA mengindikasikan performa yang lebih buruk dibandingkan sekadar memprediksi nilai rata-rata data historis. Kondisi ini dapat dijelaskan oleh sifat model ARIMA dengan *differencing* (d = 1) tanpa suku *drift*, yang prediksi jangka panjangnya cenderung konvergen ke nilai konstan sehingga tidak mengikuti tren, sementara lebar interval prediksinya terus membesar seiring bertambahnya horizon [10], [13]. Pada horizon uji 82 bulan yang mencakup perubahan tren besar, kelemahan ini menyebabkan galat yang besar. Sebaliknya, *Prophet* memodelkan tren secara eksplisit melalui mekanisme *changepoint* sehingga lebih adaptif terhadap perubahan tren jangka panjang [5].

SARIMA (1,1,1)(1,1,1,12) menghasilkan kinerja terendah (MAPE 20,83%; R^2 = −3,3074), bahkan lebih buruk daripada ARIMA(1,1,0). Ada beberapa kemungkinan penyebab. Pertama, pola musiman pada data bulanan ini relatif lemah, yaitu hanya −10,6% hingga +3,9% pada komponen musiman *Prophet*, dan ke-

mungkinan besar mencerminkan perbedaan jumlah hari per bulan (misalnya Februari terendah), bukan siklus produksi yang kuat. Kedua, kombinasi *differencing* biasa (d = 1) dan musiman (D = 1) berisiko *over-differencing*, yang memperbesar derau dan mempercepat akumulasi galat pada horizon panjang [10]. Ketiga, *order* SARIMA ditetapkan tetap dan tidak dipilih melalui AIC seperti pada ARIMA, sehingga belum tentu optimal. Keempat, perubahan struktural (*booming shale* pasca-2008 dan penurunan 2020) menyulitkan asumsi model [15]. Hasil ini menunjukkan bahwa penambahan komponen musiman tidak otomatis meningkatkan akurasi bila pola musimannya lemah.

Tabel 3. Perbandingan performa model *Prophet* dengan ARIMA dan SARIMA

Model	MAE (ribu barel/bulan)	RMSE (ribu barel/bulan)	MAPE	R^2
ARIMA(1,1,0)	37.072,43	43.371,81	13,07%	−1,1499
SARIMA(1,1,1)(1,1,1,12)	57.111,47	61.390,67	20,83%	−3,3074
Prophet (penelitian ini)	13.811,90	16.987,10	5,35%	0,6702

4.5 Perbandingan Hasil Evaluasi *Single Split* dan *Rolling Cross-Validation*

Untuk memperkuat validitas evaluasi, dilakukan pula *rolling cross-validation* menggunakan fungsi bawaan *Prophet* dengan 15 titik potong (*cutoff*) yang tersebar dari tahun 2011 hingga 2025, masing-masing dengan horizon prediksi 12 bulan. Sebagaimana ditunjukkan pada Tabel 4, hasil rata-rata *cross-validation* menghasilkan MAE sebesar 15.457,95 ribu barel/bulan, RMSE sebesar 18.828,30 ribu barel/bulan, dan MAPE sebesar 8,33%, sedikit lebih tinggi dibandingkan hasil evaluasi *single split* (MAPE 5,35%). Perbedaan ini mengindikasikan bahwa performa model pada evaluasi *single split* kemungkinan sedikit dipengaruhi oleh karakteristik spesifik pada periode pengujian yang dipilih, sehingga hasil *cross-validation* memberikan estimasi performa model yang lebih representatif dan *robust* terhadap variasi periode waktu.

Tabel 4. Perbandingan hasil evaluasi *single split* dan *rolling cross-validation*

Metrik	*Single Split* (85/15)	*Cross-Validation* (15 *fold*)
MAE (ribu barel/bulan)	13.811,90	15.457,95
RMSE (ribu barel/bulan)	16.987,10	18.828,30
MAPE	5,35%	8,33%

5. Kesimpulan

Berdasarkan hasil penelitian, algoritma *Prophet* terbukti dapat digunakan untuk meramalkan produksi minyak mentah *Gulf Coast* (PADD 3) dengan memanfaatkan data historis bulanan periode Januari 1981 hingga Mei 2026. Model dengan konfigurasi *yearly seasonality* dan *multiplicative seasonality mode* mampu menangkap pola tren jangka panjang sekaligus pola musiman tahunan yang terdapat pada data produksi, dengan hasil evaluasi pada data uji menunjukkan MAE sebesar 13.811,90 ribu barel/bulan, RMSE sebesar 16.987,10 ribu barel/bulan, MAPE sebesar 5,35%, dan R^2 sebesar 0,6702. Untuk memperkuat validitas hasil, performa *Prophet* dibandingkan dengan model ARIMA dan SARIMA, yang menunjukkan keunggulan signifikan *Prophet*, ditandai dengan nilai R^2 negatif pada kedua model pembanding (ARIMA: −1,1499; SARIMA: −3,3074). Validasi tambahan melalui *rolling cross-validation* pada 15 titik potong turut mengonfirmasi kestabilan performa model, dengan MAPE rata-rata sebesar 8,33%. Meski demikian, analisis residual mengindikasikan adanya kecenderungan *overprediction* serta *overfitting* ringan pada data uji, sehingga model masih memiliki keterbatasan dalam menggeneralisasi perubahan produksi pada periode yang belum pernah diamati sebelumnya.

Proyeksi menggunakan model akhir untuk periode Juni 2026 hingga Mei 2027 menunjukkan kecenderungan peningkatan produksi, dengan rata-rata prediksi 327.895,38 ribu barel/bulan. Hasil ini tetap perlu dimaknai sebagai estimasi berbasis pola historis, bukan kepastian mutlak. Disarankan untuk mempertimbangkan penambahan variabel eksogen pada penelitian selanjutnya, seperti kebijakan energi dan kondisi ekonomi global, serta eksplorasi mekanisme pertumbuhan yang lebih realistis seperti *logistic growth*, guna meningkatkan akurasi dan generalisasi model.

Daftar Pustaka

[1] U.S. Energy Information Administration, "Gulf Coast (PADD 3) Field Production of Crude Oil (Thousand Barrels)," *EIA*, 2026. [Online]. Available: https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?f=M&n=PET&s=MCRFPP31

[2] M. I. A. Aziz and M. H. Barawi, "Forecasting crude oil price using ARIMA and Facebook Prophet with machine learning," in *AES 2021 Proceedings*, 2021.

[3] H. Ma, "Research on crude oil price forecasting technology based on Fbprophet," in *Proc. 2023 International Conference on Data Science, Advanced Algorithm and Intelligent Computing (DAI 2023)*, *Advances in Intelligent Systems Research*, vol. 180, pp. 261–269, 2024.

[4] A. Bontemps and S. Ohnemus, "Forecasting crude oil prices using machine learning techniques," *Journal of Petroleum Economics and Policy*, vol. 29, no. 1, pp. 103–114, 2020.

[5] S. J. Taylor and B. Letham, "Forecasting at scale," *The American Statistician*, vol. 72, no. 1, pp. 37–45, 2018.

[6] Y. Ning, H. Kazemi, and P. Tahmasebi, "A comparative machine learning study for time series oil production forecasting: ARIMA, LSTM, and Prophet," *Computers & Geosciences*, vol. 164, p. 105126, 2022, doi: 10.1016/j.cageo.2022.105126.

[7] M. R. E. Satria, D. Rosiani, A. D. Candra, and A. Setiyono, "Analisis forecasting produksi gas menggunakan algoritma Prophet dan ARIMA," *SNTEM*, vol. 5, pp. 562–572, Nov. 2025.

[8] E. Bisong, *Building Machine Learning and Deep Learning Models on Google Colab*. New York, NY, USA: Apress, 2019.

[9] A. Gholamy, V. Kreinovich, and O. Kosheleva, "Why 70/30 or 80/20 relation between training and testing sets: A pedagogical explanation," *Departmental Technical Reports (CS)*, Univ. of Texas at El Paso, Tech. Rep. UTEP-CS-18-09, no. 1209, Feb. 2018. [Online]. Available: https://scholarworks.utep.edu/cs_techrep/1209

[10] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 3rd ed. Melbourne, Australia: OTexts, 2021.

[11] L. J. Tashman, "Out-of-sample tests of forecasting accuracy: An analysis and review," *International Journal of Forecasting*, vol. 16, no. 4, pp. 437–450, 2000.

[12] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International Journal of Forecasting*, vol. 22, no. 4, pp. 679–688, 2006.

[13] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, 5th ed. Hoboken, NJ, USA: Wiley, 2015.

[14] E. Bouri, R. Gupta, and D. Roubaud, "Volatility spillovers and the time-varying correlation between oil and stock markets: Evidence from COVID-19," *Energy Economics*, vol. 93, p. 104989, 2021.

[15] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "Statistical and Machine Learning forecasting methods: Concerns and ways forward," *PLOS ONE*, vol. 13, no. 3, p. e0194889, 2018.

[16] S. Makridakis and M. Hibon, "The M3 competition: Results, conclusions and implications," *International Journal of Forecasting*, vol. 16, no. 4, pp. 451–476, 2000.