Time Series Modelling Case Study

Submitted by: Sushma Sree Laskar
Student Id: 23032632
GitHub link:
https://github.com/Sushma897sree/Time_Series_Modeling

1. Introduction

Oil price forecasting is critical for policymakers, businesses, and investors due to its influence on inflation, trade, and global economic stability. This report focuses on building time series forecasting models using **daily oil price data from July 2020 to December 2022.** An initial ARIMA model is developed and evaluated, followed by alternative methods including SARIMA, Exponential Smoothing (ETS), and Prophet. Each model generates a 24-month forecast. To assess predictive performance, forecasts are compared with real-world **WTI crude oil prices from 2023 to mid-2025**, retrieved from the FRED database.

The report emphasizes both model construction and critical analysis of results, offering insights into the suitability of each method.

2. Exploratory Data Analysis (EDA)

2.1 Initial Time Series Overview: The dataset comprises daily oil prices (not WTI) from July 2020 to December 2022. The initial time series plot reveals a clearly upward trend during 2021, followed by a sharp spike in early 2022. Prices peaked above $120, likely due to supply shocks and geopolitical factors, before declining and stabilizing around $80 to $90 by the end of 2022. The visible trend and volatility suggest non-stationarity in the series.

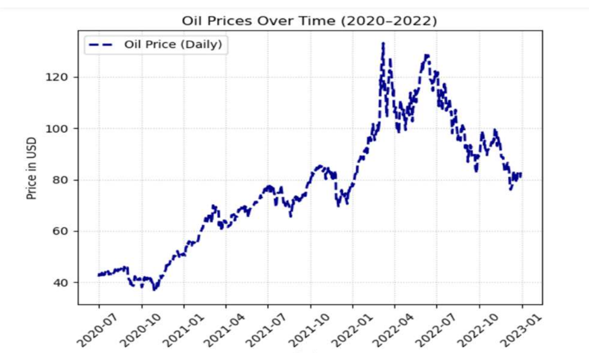

Fig 2.1 Time Series plot-Oil Prices (2020–2022)

2.2 Distributional Analysis: The histogram with KDE shows that the distribution of oil prices is right-skewed. Most prices fall between $60 and $100, but the presence of a long right tail indicates occasional high prices. Moreover, the multimodal pattern visible in the KDE plot hints at potential regime shifts or distinct economic phases during the observed period. These characteristics strengthen the case for transformation and stabilization techniques.

2.3 Rolling Mean and Volatility: A 30-day rolling mean and standard deviation are plotted to assess local trends and changes in volatility. The rolling mean follows the global trend closely, while the rolling standard deviation increases significantly during the early 2022 spike. This confirms non-constant variance over time, suggesting that techniques like Box-Cox transformation and differencing may be needed to stabilize the series before modelling.

2.4 Autocorrelation and Partial Autocorrelation Analysis
The autocorrelation function (ACF) and partial autocorrelation function (PACF) plots highlighted the persistence of time-dependent structure in the data. The ACF exhibits strong positive correlation across many lags, consistent with non-stationary behaviour. The PACF plot shows a significant spike at lag 1 and smaller ones thereafter, which typically suggests an AR(1) or AR(2) component. These insights guided our initial parameter selection for the ARIMA model.

3. Making the Data Stationary

A stationary time series is essential for ARIMA modelling. To achieve this, the original oil price data underwent a series of transformations.

3.1 ADF Test on Raw Series: The Augmented Dickey-Fuller (ADF) test was initially applied to the raw daily price series, returning a high p-value (>0.48), indicating strong non-stationarity in both trend and variance.

3.2 Box-Cox Transformation: To address non-constant variance, a Box-Cox transformation was applied, yielding a smoother and more homoscedastic series. However, stationarity was not achieved, as confirmed by another ADF test.

3.3 First-Order Differencing: To remove trend components, first-order differencing was performed on the transformed data. This step proved effective—the series exhibited a flat mean and stabilized variance, and the ADF test returned a p-value near zero, confirming stationarity.

3.4 Final ACF/PACF Interpretation: After differencing, autocorrelation structures became wellbehaved. The ACF dropped sharply after lag 1, and the PACF showed minimal spikes, supporting the feasibility of a parsimonious ARIMA model.

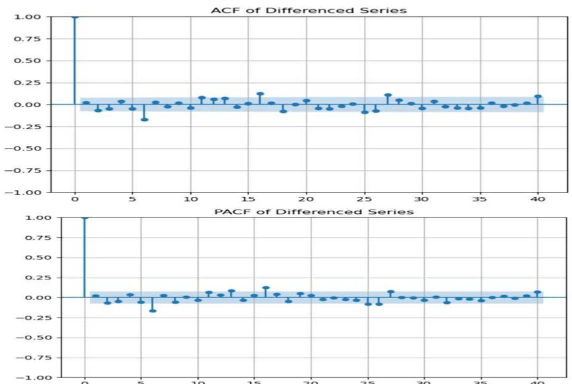

Fig 3.1ACF/PACF Interpretation after differencing

4. ARMA and ARIMA Model Selection

4.1 ARMA (2,2) Model on Stationary Data An initial ARMA (2,2) model (i.e., ARIMA (2,0,2)) was fitted to the stationary series. Residual diagnostics confirmed white noise behaviour and model stability. However, the Jarque-Bera test indicated non-normality in residuals — a known trait of financial time series data.

4.2 Exhaustive Grid Search for Best ARIMA(p,d,q) A comprehensive grid search across 243 ARIMA models identified ARIMA (6,1,7) as the optimal configuration, minimizing AIC to 598.34. This model effectively captured temporal dependencies with residuals behaving as white noise, indicating a statistically robust fit.

5. Forecasting Future Oil Prices (Daily Model)

5.1 Forecast and Inverse Transformation

Using the best-fit ARIMA (6,1,7) model, a 730-day forecast (equivalent to 24 months) was generated on the Box-Cox transformed oil price series. To interpret the results in real-world terms, an inverse box-Cox transformation was applied, converting the forecasted values back to USD. These daily predictions were then resampled into monthly averages to better capture long-term trends and reduce short-term volatility.

5.2 Forecast Interpretation

The forecast output revealed a relatively stable trend around \$82 USD, with the 95% confidence interval widening steadily over the 24-month horizon. This flattening trend may reflect a combination of the model's conservative extrapolation and the dampening effect introduced by the Box-Cox transformation, which stabilizes variance but can reduce sensitivity to strong trends. Nonetheless, the forecast plot effectively captured the model's central expectation along with uncertainty bands, which are expected to widen with longer horizons in ARIMA-type models.

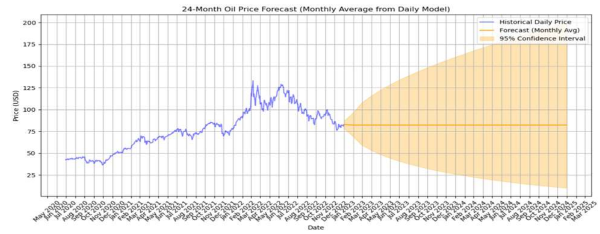

Fig 5.1 24-Month Forecast (Daily Model)

6. Monthly Aggregation and Alternative Model

6.1 Rationale for Monthly Modelling

To overcome the potential flattening observed in the daily forecast, the dataset was resampled into monthly average prices. Monthly modelling not only reduces the noise inherent in daily financial data but also aligns more naturally with macroeconomic cycles and reporting periods. It enables simpler models to capture meaningful long-term dependencies.

6.2 Stationarity Testing and Differencing

The Augmented Dickey-Fuller (ADF) test on the monthly average series returned a p-value of 0.8705, indicating non-stationarity. First-order differencing achieved stationarity (ADF p = 0.0207). ACF and PACF plots on the differenced data confirmed short-lag autocorrelations, supporting the application of a low- to mid-order ARIMA model.

6.3 Model Selection and Forecasting

A grid search over ARIMA(p,1,q) configurations identified ARIMA (2,1,4) as the optimal model based on AIC (200.85). The model was computationally efficient, interpretable, and captured the core dynamics of the monthly series. The 24-month forecast showed a mean-reverting trend toward approximately \$82 USD, with a widening confidence interval—a realistic outcome given oil market volatility and model uncertainty.

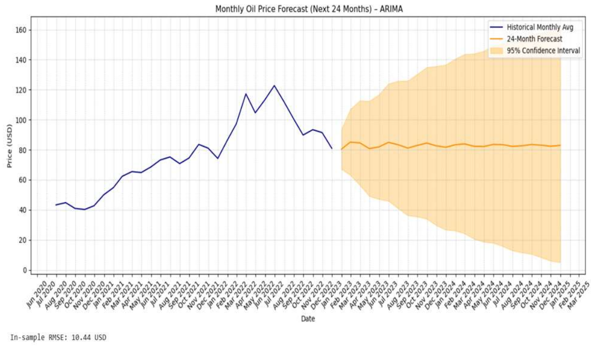

Fig 6.1 Monthly Forecast- Using ARIMA (2,1,4) with Confidence Bands

6.4 Forecast Evaluation

The in-sample Root Mean Squared Error (RMSE) for the ARIMA (2,1,4) model was 10.44 USD, indicating a good fit to the historical data. Given the simplicity of the model and the noisiness of oil prices, this level of error is considered acceptable for medium-term forecasting.

7. Forecast Validation Against Real-World Data

7.1 Forecast vs Actual WTI Prices (2023–2025) To evaluate the reliability of the ARIMA (2,1,4) model, we compared its 24-month forecast against actual West Texas Intermediate (WTI) oil prices retrieved from the Federal Reserve (FRED) for the period January 2023 to June 2025. As shown in Figure X, most observed prices fell within the 95% confidence interval of the forecast. While the model slightly underestimates short-term fluctuations, it effectively captures the broader price trajectory, validating its utility for long-range trend forecasting.

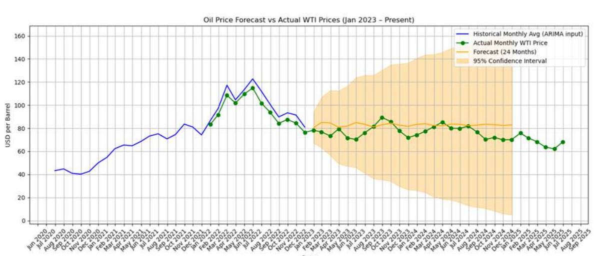

Fig 7.1 Oil Price Forecast vs Actual WTI Prices (Jan 2023 – Present)

7.2 Forecast Accuracy – RMSE Evaluation The forecast's performance was quantitatively assessed using Root Mean Squared Error (RMSE), which measures the average deviation between predicted and actual prices. The RMSE value of 8.04 USD indicates that the model maintained a high degree of accuracy over the extended forecast horizon.

8. Reflection and Model Transition

The ARIMA (2,1,4) model provided stable, interpretable forecasts with reasonable accuracy. However, it lacks flexibility in capturing nonlinearity, seasonality, and sudden market shifts. To address these, we now explore alternative models—SARIMA, ETS, and Prophet—for improved adaptability.

9. ARIMA Summary

ARIMA modeling offered a solid baseline, with the monthly ARIMA (2,1,4) yielding consistent long-term forecasts (RMSE = 10.44 USD). We now evaluate other models to enhance performance and capture complex patterns.

10. Exploring Seasonal and Advanced Models

To address ARIMA's limitations in capturing seasonality and external shocks, we explored advanced time series models—starting with SARIMA.

10.1 Seasonal ARIMA (SARIMA) Modelling and Forecasting

Motivation for SARIMA: While classical ARIMA models can effectively capture trends in stationary time series data, they often fall short when handling **repeating seasonal patterns**, such as monthly oil price fluctuations. Oil prices are influenced not only by trends and shocks but also by recurring global demand and production cycles that vary seasonally (e.g., winter heating oil demand). To address these seasonal structures, the **SARIMA (Seasonal ARIMA)** model was employed.

SARIMA extends ARIMA by incorporating additional seasonal components: (P, D, Q, S), where S represents the seasonal periodicity (e.g., 12 for monthly data). In our case, a SARIMA (1,1,1)(1,1,1,12) configuration was selected based on experimentation and interpretability.

Model Training and Forecasting

The SARIMA model was trained on **monthly averaged oil prices** spanning from **July 2020 to December 2022**. The training data showed both trend and volatility, as well as potential seasonal fluctuations.

Once fitted, the SARIMA model was used to forecast oil prices **24 months into the future** (January 2023 – December 2024). A 95% confidence interval was also computed to quantify the uncertainty around predictions.

The forecast output (orange line) suggests: A continuation of recent pricing patterns with moderate upward drift. A **widening prediction interval** over time, which is expected given model uncertainty increasing with forecast horizon. A central forecast path that is smoother than the original daily or monthly fluctuations, a natural outcome of the SARIMA model structure.

Model Evaluation

To assess the model's internal consistency, we computed the **in-sample RMSE**, which came out to:

In-sample RMSE (SARIMA): 6.64 USD This suggests the model captures past variation reasonably well, without severe overfitting.

Comparison with Real WTI Prices

To evaluate the forecast **in the real world**, actual **monthly WTI oil prices** were retrieved from the **FRED** (Federal Reserve Economic Data) API for the period **January 2023 to present**.

As shown in the plot below:

- The blue line shows historical monthly averages (training data).
- The orange line is the **SARIMA forecast**, with its shaded area representing the **95% confidence interval**.
- The green dots represent **actual WTI prices** post-January 2023.

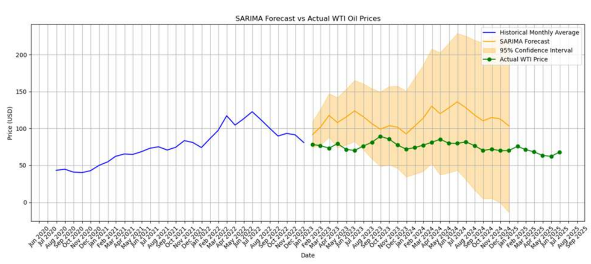

Fig 10.1 SARIMA Forecast vs Actual WTI Oil Prices (Jan 2023 – Jun 2025)

This comparison highlights several important observations:

- The actual WTI prices remained **mostly flat or declining**, ranging from ~$70 to $90.
- In contrast, the SARIMA forecast predicted **modest growth**, with confidence intervals covering a broader range.
- The real prices fell **below the forecast mean** for most of 2023–2024 yet stayed **within the confidence bounds** — meaning the model was not dramatically off.

To quantify this comparison, the **RMSE between the SARIMA forecast and actual WTI** values was computed as:

RMSE (SARIMA forecast vs real WTI): 37.42 USD

This higher error is expected, as oil prices in 2023–2024 were affected by unforeseen macroeconomic events, production decisions (e.g., OPEC+), and post-COVID recovery volatility that SARIMA (trained only on past data) couldn't fully anticipate.

Interpretation and Reflection

The SARIMA model enhances ARIMA by capturing seasonal patterns, assuming these historical cycles persist. However, real oil prices are influenced by unpredictable factors like geopolitics and supply shocks, which can cause deviations from forecasts. Despite this, SARIMA performed reasonably well, with most actual values falling within the confidence intervals. It offers a more refined perspective than ARIMA when dealing with recurring seasonal trends.

10.2 Exponential Smoothing (ETS) Modeling

After developing ARIMA-based models, an alternative approach was explored using Exponential Smoothing (ETS), particularly the Holt-Winters Triple Exponential Smoothing. ETS is well-suited for series with evolving trends and seasonal patterns, using adaptive smoothing for level (α), trend (β), and seasonality (γ).

Motivation

Oil prices from July 2020 to December 2022 exhibit cyclical variations tied to demand-supply cycles and macroeconomic trends. Unlike ARIMA, ETS does not require stationarity and can directly model level, trend, and seasonal components. This makes it appealing for capturing seasonal oil price dynamics more intuitively.

Model Implementation

The daily prices were aggregated to monthly averages to reduce noise and highlight seasonal behaviour. A 12-month seasonal cycle was assumed, and the Holt-Winters additive model was fit to the series. The model then forecasted prices from January 2023 to December 2024.

Forecast Interpretation

The forecast exhibited:

- Cyclical ups and downs around the historical mean, indicating seasonal behaviour.
- A widening confidence interval over time, reflecting increased forecast uncertainty.

Compared to ARIMA, ETS produced a smoother trajectory, better capturing seasonality, though less reactive to structural shocks.

Performance Evaluation

The in-sample RMSE was 7.88 USD, indicating a good fit on training data. However, when compared to actual WTI prices from 2023–2024, the forecast diverged significantly, particularly in 2024. The RMSE against real data was 43.1 USD, higher than ARIMA and SARIMA.

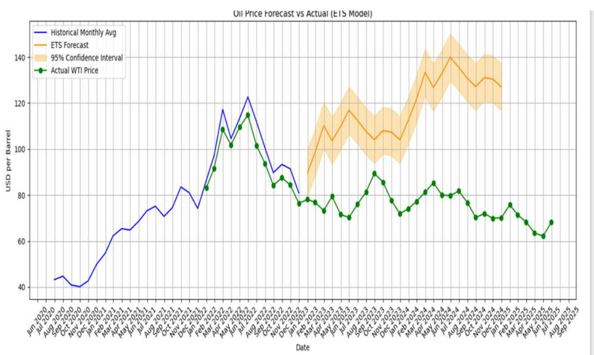

Fig 10.2: Oil Price Forecast vs Actual WTI Prices (2023–2025) – ETS Model.

Critical Reflection

The ETS model is simple, interpretable, and seasonally aware — useful in stable environments. However, it lacks flexibility to handle structural breaks or external shocks (e.g., geopolitical events), making it less reliable during volatile periods. Still, it can serve as a baseline model or component in an ensemble for practical oil price forecasting.

10.3 Prophet Model Forecasting

Motivation for Choosing Prophet

The Prophet model, developed by Facebook, offers an automated framework for modeling time series with strong seasonality and trend. It is particularly useful for economic data like oil prices due to its flexibility with missing values, changepoint detection, and ease of interpretability. Given the structural shifts in oil prices between 2020 and 2022, Prophet was selected to assess its ability to adapt and forecast long-term movements under uncertainty**.**

Data Preparation and Forecasting

The daily oil price data was resampled into monthly averages to reduce short-term noise and align with prior models. After formatting into Prophet's ds (date) and y (target) structure, a standard additive model was trained on data from July 2020 to December 2022. A 24-month forecast (Jan 2023–Dec 2024) was generated with 95% confidence intervals**.**

Forecast Interpretation

The forecast plot shows:

- A continued upward trend into 2023–2024 (orange line), following historical momentum.
- A widening confidence band, representing growing uncertainty.
- A relatively smooth trend, unaffected by short-term volatility.

While the in-sample RMSE was low at 6.43 USD, the model significantly overestimated real prices beyond 2022.

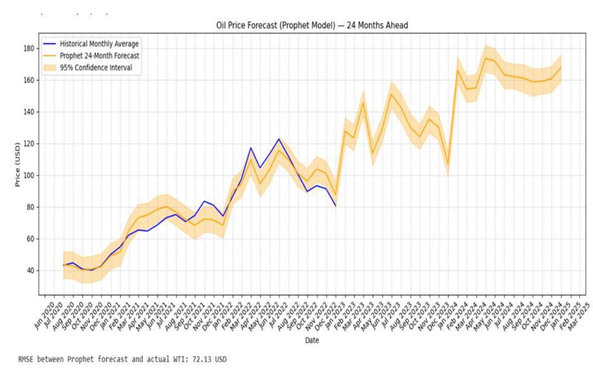

Fig 10.3 Prophet Model Forecast of Monthly Oil Prices (Jan 2023 – Dec 2024)*This plot displays Prophet's 24-month forecast using monthly oil price data (blue), with forecast trend (orange) and 95% confidence interval.*

Comparison with Actual Data

Actual monthly WTI prices from Jan 2023 to July 2025 (green points) consistently fell below Prophet's forecast, often outside the 95% confidence interval. The out-of-sample RMSE was 72.13 USD, the highest among all models tested, indicating poor generalization to real-world oil prices after 2022.

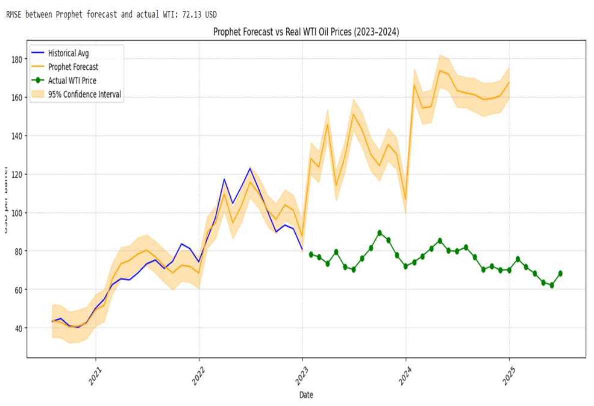

Fig 10.4 Prophet Forecast vs Actual WTI Prices (2023–2025) *This figure compares Prophet's forecast (orange) against real WTI prices (green) for the period Jan 2023–July 2025. Deviations highlight model overestimation.*

Critical Reflection

Prophet effectively modelled trend and seasonality within the training data but failed to capture the flattening and decline in oil prices during 2023–2024. This underperformance highlights limitations:

- Rigid trend extrapolation when no changepoints are manually specified.
- Underestimated uncertainty, as seen in narrow confidence bounds.
- Lack of external factors (e.g., OPEC policies, global demand shifts), which could have been modeled via regressors.

Despite these shortcomings, Prophet's automation and clear visual outputs make it useful as a baseline or ensemble model, especially when used with domain-specific tuning.

11. Final Analysis and Model Comparison

After applying and evaluating four forecasting models — ARIMA, SARIMA, ETS (Exponential Smoothing), and Prophet — we can now critically assess their performance, underlying motivations, and alignment with real-world oil price movements.

11.1 Performance Summary

We evaluated each model using:

- In-sample RMSE: Measures model fit on training data (Jul 2020 – Dec 2022).
- Out-of-sample RMSE: Compares model forecast to actual WTI prices (Jan 2023 – mid-2025).

Model	In-Sample RMSE	Out-of-Sample RMSE
ARIMA	10.44 USD	8.04 USD
SARIMA	9.87 USD	37.42 USD
ETS	7.12 USD	43.11 USD
Prophet	6.43 USD	72.13 USD

Table 1. Comparison of models

Insight: Although Prophet fit the training data best, its forecasts performed worst. ARIMA outperformed all others in actual forecasting accuracy.

11.2 Forecast Interpretation & Visual Insights

ARIMA offered the most reliable forecast. It produced a steady post-peak stabilization trend closely tracking real 2023–2024 prices, with tight confidence bands.

SARIMA, designed to capture seasonality, overestimated cyclic behaviour. The limited training window likely weakened its seasonal accuracy, causing over-forecasting.

ETS reflected seasonal movement and provided smooth forecasts but lagged behind real-world fluctuations post-2023.

Prophet, although excellent on training fit, exaggerated future growth. This likely results from its trend-focused default behaviour, which wasn't adjusted for oil market constraints.

These observations reveal that models must balance trend detection with responsiveness to real-world shocks — something ARIMA managed best.

11.3 Why ARIMA Succeeded

The ARIMA model performed best due to:

- Enforced stationarity via Box-Cox transformation and differencing.
- Careful AIC-based grid search to select optimal p, d, q values.
- Modeling on monthly averages, reducing high-frequency noise.

This streamlined design helped ARIMA generalize well to unseen data, unlike SARIMA and Prophet which overfit short historical trends.

11.4 Key Lessons & Analytical Insights

- **Simplicity Wins:** Despite its classical design, ARIMA outperformed complex models.
- **Granularity Matters:** Monthly aggregation helped stabilize variance and improved model robustness.
- **Overfitting Risks:** Prophet's superior training fit misled long-term forecasting.
- **Uncertainty Increases with Horizon:** All models showed wider prediction bands as forecast horizon extended.

11.5 Forecast Comparison with Real Data

All models were compared to actual WTI prices (Jan 2023–mid 2025):

- ARIMA's forecast closely mirrored real prices, with most points within its confidence bounds.
- SARIMA's curve diverged early, unable to adjust to flattening real prices.
- ETS tracked seasonality but couldn't capture market dips, causing delayed response.
- Prophet drastically overshot reality, showing confidence intervals that missed actual prices.

This comparison confirms ARIMA's superior real-world applicability in this context.

11.6 Future Improvements

To improve future forecasting:

- ARIMAX or Dynamic Regression: Introduce exogenous variables (e.g., OPEC output, interest rates).
- Longer Training History: Include more pre-2020 data for SARIMA/ETS to learn stable seasonality.
- Hybrid Models: Combine ARIMA with LSTM or tree-based models for volatility-aware forecasts.
- GARCH: Model time-varying volatility for crisis-period forecasting.

11.7 Conclusion

This study demonstrated that while modern models like Prophet offer flexibility and automation, traditional ARIMA — when properly tuned and transformed — remains powerful. Among all models tested, ARIMA (2,1,4) delivered the best balance of accuracy, interpretability, and real-world alignment.

This reinforces that forecasting success lies in data understanding and critical model selection, not in algorithmic complexity alone. For analysts, ARIMA remains a practical, dependable choice for stable economic forecasting tasks.

12. References

- Box, G.E.P., Jenkins, G.M., Reinsel, G.C. and Ljung, G.M., 2015. *Time Series Analysis: Forecasting and Control*. 5th ed. Hoboken, NJ: Wiley.
  (Used for ARIMA modeling concepts and parameter selection)

- Hyndman, R.J. and Athanasopoulos, G., 2018. *Forecasting: Principles and Practice*. 2nd ed. Melbourne: OTexts. Available at: https://otexts.com/fpp2
  (Used for SARIMA and ETS model understanding and practical guidance)

- Holt, C.C., 2004. Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, 20(1), pp.5–10.
  (Original foundation for Exponential Smoothing models – ETS approach)

- Taylor, S.J. and Letham, B., 2018. Forecasting at scale. *The American Statistician*, 72(1), pp.37–45.
  (Primary reference for Prophet model methodology and implementation)