Hybrid ARIMAX-LSTM Modeling for Oil ETF Price Forecasting

Chenming Dai
Department of Economics
University of Michigan
Ann Arbor, United States
cmdai@umich.edu

***Abstract*—This paper evaluates the efficacy of time series forecasting models for predicting the daily closing price of the Invesco DB Oil Fund (DBO) ETF. We implement and compare three approaches: Autoregressive Integrated Moving Average with Exogenous Regressors (ARIMAX), Long Short-Term Memory (LSTM) neural networks with technical indicators, and a novel hybrid ARIMAX-LSTM model designed to capture both linear and nonlinear patterns. Using historical DBO price data alongside relevant economic indicators, we evaluate model performance based on standard error metrics (MAE, MSE, RMSE) on a held-out test set. Results demonstrate that the hybrid model achieves superior predictive accuracy compared to standalone models, highlighting the benefits of combining traditional statistical methods with deep learning techniques for financial time-series forecasting.**

***Index Terms*—Time Series Forecasting, Financial Modeling, Oil ETF, DBO, ARIMAX, LSTM.**

I. Introduction

Financial market forecasting for commodity-based assets such as oil ETFs presents significant challenges due to inherent volatility and complex market dependencies. The Invesco DB Oil Fund (DBO) ETF tracks the DBIQ Optimum Yield Crude Oil Index Excess Return, making it a relevant proxy for crude oil market trends. Financial time series exhibit non-stationarity, nonlinearity, and dependence on various external factors [1].

Traditional statistical models like ARIMA have limitations in capturing nonlinear patterns [2], while deep learning approaches, particularly Long Short-Term Memory (LSTM) networks, show promise in modeling sequential data with long-range dependencies [3]. Hybrid approaches combining statistical and machine learning models have emerged as a potential avenue for improved accuracy [5].

This study aims to develop and evaluate forecasting models for DBO ETF price, comparing the predictive performance of ARIMAX, LSTM, and a hybrid ARIMAX-LSTM model. We formulate the problem as a supervised learning task where:

- **Input**: Historical DBO prices, exogenous economic variables (WTI price, USD index), and technical indicators.
- **Output**: Forecasted DBO closing price for future time steps.

The dataset comprises daily data spanning January 2010 to April 2025, chronologically split into training and testing sets.

II. Method

Our methodology comprises data preprocessing followed by the implementation and evaluation of three forecasting models using Python's data science and machine learning libraries.

A. Data Preprocessing

Raw data for DBO prices and exogenous variables underwent cleaning with forward-fill, backward-fill, and interpolation methods for missing values. Technical indicators were computed from historical prices. Features were normalized using MinMaxScaler to [0,1] range. The dataset was chronologically split into training (1511 days) and test sets (378 days). For the LSTM model, input sequences of 30 days were created.

B. ARIMAX Model

The ARIMAX(p,d,q) model extends ARIMA by incorporating exogenous variables to capture linear relationships and autocorrelation. Optimal orders (p=1, d=1, q=1) were determined using ACF/PACF plots. The model was fitted with two exogenous variables: WTI Crude Oil price and DXY Dollar Index.

C. LSTM Model

An LSTM network architecture with two stacked layers was implemented to capture non-linear patterns and temporal dependencies. The model configuration included:

- Architecture: Two LSTM layers (64, 32 units) followed by two Dense layers
- Sequence Length: 60
- Dropout Rate: 0.3
- Optimizer: Adam with MSE loss function
- Training: 100 epochs with batch size 32

Technical indicators served as input features for the model.

D. Hybrid ARIMAX-LSTM Model

This two-stage approach combines both models:

1) ARIMAX model fitted to capture linear components
2) LSTM network trained on ARIMAX residuals ($e_t = y_t - \hat{y}_{ARIMAX,t}$) to model non-linear patterns
3) Final forecast calculated as $\hat{y}_{Hybrid,t} = \hat{y}_{ARIMAX,t} + \hat{e}_{LSTM,t}$

This architecture allows ARIMAX to model linear structure while LSTM addresses remaining non-linear components.

III. Results

Model performance was evaluated on the unseen test dataset using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE).

A. Model Performance Metrics

Table I summarizes the error metrics for each model on the test set.

TABLE I
Model Performance Comparison on Test Set

Model	MAE	MAPE (%)	RMSE
ARIMAX	2.57	18.27	2.62
LSTM	0.50	3.63	0.62
Hybrid	0.27	1.93	0.34

The hybrid model achieved the lowest error metrics across all measures, demonstrating superior ability to capture DBO price dynamics.

B. Forecast Visualization

Figure 1 illustrates model predictions against actual DBO prices on the test set. The hybrid model demonstrates closer alignment with actual price movements, particularly during periods of higher volatility.

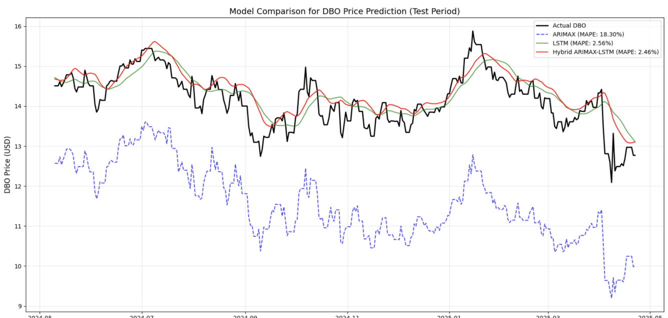

Fig. 1. Model Comparison for DBO Price Prediction (Test Period). The figure shows the actual DBO price (solid black) compared with predictions from ARIMAX (dashed blue, MAPE: 18.30%), LSTM (solid green, MAPE: 2.56%), and Hybrid ARIMAX-LSTM (solid red, MAPE: 2.46%) models over the test period from May 2024 to April 2025.

C. Future Forecast Analysis

The models generated 30-day forecasts beyond the test set. Table II presents statistical properties of these forecasts.

TABLE II
Statistical Properties of 30-Day Forecasts

Model	Min	Max	Mean	Std Dev
ARIMAX	9.98	9.99	9.98	0.0002
LSTM	12.83	13.05	12.88	0.0461
Hybrid	12.94	13.14	13.02	0.0766

Given its superior accuracy on test data, the hybrid model's forecast (mean: 13.02, std: 0.08) likely provides the most reliable projection for future DBO prices.

IV. Conclusion

This study compared ARIMAX, LSTM, and a hybrid ARIMAX-LSTM model for DBO ETF price forecasting. Performance metrics demonstrate that the hybrid approach achieved superior accuracy, suggesting that integrating linear modeling capabilities of ARIMAX with non-linear pattern recognition of LSTM yields synergistic benefits for financial time series prediction.

Limitations include the assumption that historical patterns persist, vulnerability to unforeseen market shocks, and sensitivity to hyperparameter selection. Future research directions include incorporating sentiment analysis data, exploring transformer architectures, and implementing adaptive retraining strategies.

In conclusion, our results highlight the potential of hybrid modeling approaches for financial forecasting, offering a more robust analytical tool compared to standalone traditional or machine learning models.

References

[1] T.-Y. Tsai, C.-C. Wang, G.-W. Wu, G.-Y. Lin, and J.-S. Chiang, "Deep Learning for Financial Time Series Prediction: A State-of-the-Art Review of Standalone and Hybrid Models," Computer Modeling in Engineering & Sciences (CMES), vol. 139, no. 1, pp. 1–42, 2023.

[2] F. Alharbi, A. Alarifi, A. Alshdadi, A. Albeshri, R. Alsaqour, and M. Amoon, "Navigating the Financial Landscape: The Power and Limitations of the ARIMA Model," Humanities and Social Science Research, vol. 7, no. 1, pp. p10, 2024.

[3] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[4] R. A. Bakar, M. A. Abdullah, M. F. M. Fudzee, and K. A. Bakar, "A Systematic Literature Review on The Role of LSTM Networks in Capturing Temporal Dependencies in Data Mining Algorithms," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 15, no. 10, 2024.

[5] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," Neurocomputing, vol. 50, pp. 159–175, 2003.

[6] S. Alotaibi, M. Alrasheed, Y. Alghamdi, and O. Aljohani, "Prediction of the Stock Market Using LSTM, ARIMA, and Hybrid of LSTM-ARIMA Models," International Journal of Computing and Digital Systems, vol. 17, no. 1, pp. 1–15, 2024.