Intermarket Analysis: Influence of Brent Technical Indicators Using ARIMAX Modeling for Improving WTI Crude Oil Price Forecasting

Suhad A. Yousif[1], Nadia M. G. Al-Saidi[2], Suzan J. Obaiys[3], Redha A. Redha[4], and Nur Amalina Binti Jamaludin[5(✉)]

[1] Department of Computer Science, Al-Nahrain University, Baghdad, Iraq
suhad.a.yousif@nahrainuniv.edu.iq
[2] Department of Applied Sciences, University of Technology, 10066 Baghdad, Iraq
nadia.m.ghanim@uotechnology.edu.iq
[3] Department of Computer System & Technology, Faculty of Computer Science and Information Technology, Universiti Malaya, 50603 Kuala Lumpur, Malaysia
[4] State Organization for Marketing of Oil (SOMO), Bagdad, Iraq
redhaali18@gmail.com
[5] Department of Mathematics, Centre of Foundation Studies, National Defence University of Malaysia, Kuala Lumpur, Malaysia
suzan@um.edu.my

Abstract. The task of predicting crude oil prices is of great importance but is also highly demanding due to the economic ramifications associated with these prices and their intricate, non-linear characteristics. We introduce a new method which utilizes technical indicators from various oil benchmarks, specifically Brent crude, to forecast daily prices of WTI crude oil, acknowledging the potential of this integration. Through the utilization of the ARIMAX model, which integrates exogenous factors such as Brent crude technical indicators, In this study, we introduce a new model for price prediction that exhibits superior performance compared to conventional methods such as Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) models. The performance measures, namely Root Mean Square Error (RMSE) with a value of 2.9148, Mean Absolute Error (MAE) with a value of 2.3166, and Mean Absolute Percentage Error (MAPE) with a value of 3.971, not only serve as a validation for our model but also indicate its superiority over other models frequently mentioned in the existing literature. These models will be thoroughly examined and debated. This study not only introduces a new forecasting methodology but also demonstrates its potential applicability in several economic sectors that are impacted by fluctuations in oil prices.

Keywords: ARIMAX · Auto ARIMA · Technical Indicators · Intermarket Analysis · WTI · Brent · Forecasting

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16758, pp. 393–411, 2027.
https://doi.org/10.1007/978-3-032-30497-1_24

1 Introduction

Oil wells extract crude oil, a combination of petroleum liquids and gases, a crucial energy source for manufacturing a wide range of products [8]. Because oil is one of the world's key energy sources, crude oil prices significantly impact the global economy and security. Furthermore, since crude oil accounts for a substantial proportion of the exports of certain nations, a sharp price change can have serious economic ramifications, with lower crude oil prices resulting in lower economic activity [14].

Price is influenced by Many factors, such as supply and demand dynamics, stock levels, weather conditions, political influences, and other relevant variables, which have a significant role in affecting outcomes. Unexpected events; these factors cause significant oil price fluctuation, which causes problems for business leaders and policymakers. As a result, crude oil price forecasting is an interesting field [21]. Two of the world's most important crude oil types are the West Texas Intermediate (WTI) and Brent Crude oil. WTI crude oil is refined primarily in the Midwest and Gulf Coast regions of the United States of America. It is mainly supplied to the North American market, while Brent crude oil is sourced from the North Sea and is used mainly by Europe [12].

Several studies on crude oil time-series forecasting have been conducted [1,2]. These studies assume that future data values are proportionate to present and past data values. However, it is important to note that numerous real-world time series datasets exhibit intricate nonlinear, seasonal, and non-stationary patterns, which pose significant challenges for conventional machine learning methods in terms of capturing and modelling these patterns. One instance of a conventional forecasting technique is the utilization of autoregressive integrated moving average (ARIMA) models to generate precise forecasts in the event of price stability [28].

Time-series data, such as stock market statistics, do not display seasonal patterns, although they may demonstrate fluctuating levels of volatility. During the COVID-19 period, there has been significant volatility in crude oil prices [17]. Numerous empirical studies have recently exposed the nonlinear character of economic and financial data, demonstrating that standard methodologies cannot successfully analyze the complex nonlinear dynamics at hand [10].

Artificial Neural Networks (ANNs) are one of the most extensively utilized nonlinear time series modelling techniques to capture the complex nonlinear properties of crude oil price movements, one example of which is long short-term memory (LSTM) [9]. The disadvantage of utilizing LSTM is that it needs much tuning to acquire the optimal set of parameters. It is sometimes used with one of the optimizer methods, such as Adam optimizers. Adam calculates multiple learning rates in an adaptive learning system depending on various criteria [18]. Thus, if the time-series data are extensive, they may require more CPU time. Additionally, ANNs models frequently suffer from local minima and overfitting issues [14].

This research aims to get accurate results using a light model by exploiting strong interrelationships between two oil market benchmarks [8], and Also to

illustrate the impact of technical indicators extracted from the Brent dataset on forecasts WTI dataset. By exploiting strong interrelationships between two oil market benchmarks, this research contributes by transforming the traditional time series model from capturing simple transform patterns to capturing volatility patterns. The utilization of a specific collection of technical indicators derived from the Brent crude oil price as an exogenous variable was used to achieve this outcome within the ARIMA model. Over ANN, ARIMAX, with technical indicators, can be less complicated since it does not require the dataset to be normalized, the number of hidden variables to be determined, or the parameters to be hyperfine-tuned to capture nonlinearity. The result is similar to the ANN prediction.

The first section of this paper introduces the subject. Then, the related works are discussed in Sect. 2. Then, the methodology and the central concept of the proposed algorithm are described in Sect. 3. Next, in Sect. 4, the results of the experimental setup for three case studies are illustrated, followed by the conclusion Sect. 5.

2 Related Work

Crude oil price forecasting has evolved from classical time-series modeling toward hybrid learning frameworks that incorporate nonlinear relationships and engineered market descriptors. Early work relied on Box–Jenkins style modeling to capture linear temporal dependence and seasonality, typically using ARIMA-family models and their variants for short-horizon prediction [1,23]. Subsequent studies extended these baselines by improving preprocessing, stationarity handling, and error evaluation for energy price series [10]. While these approaches are computationally efficient and interpretable, their performance can degrade when market dynamics shift, or when the series exhibits nonlinear behavior and volatility clustering.

To address these limitations, researchers increasingly adopted machine learning methods that can model nonlinear patterns from historical sequences. Support vector machines and support vector regression (SVR) have been widely explored for commodity forecasting because they can handle nonlinear regression with controlled model capacity [27,33]. In parallel, deep learning gained attention for learning representations directly from price history; for example, convolutional and recurrent architectures have been evaluated for crude oil forecasting and often report improvements under certain regimes [14,16]. However, these models typically require extensive tuning and computational resources, which can limit rapid retraining in operational environments.

Another important direction is feature engineering and the use of exogenous drivers. Several studies report that integrating external variables (macroeconomic factors, correlated markets, or engineered technical indicators) improves forecast stability and generalization [7,15,26]. Review and survey studies similarly emphasize that hybridization and feature design are central determinants of forecasting performance across energy markets [5,20].

Building on this trajectory, the authors demonstrated that SVR combined with technical indicators can improve crude oil price forecasting by leveraging indicator-based features as informative predictors [29]. The present paper extends this idea by adopting an intermarket perspective: instead of extracting indicators from the target series alone, technical indicators are generated from a related benchmark (Brent) and used as exogenous variables to strengthen WTI forecasting within ARIMAX and SVR settings. This intermarket feature-transfer design differentiates our work from (i) pure ARIMA/SVR approaches that use only WTI history, and (ii) deep learning approaches that replace interpretability with complex architectures. Our contribution is to show that carefully selected Brent-derived indicators (via correlation filtering) can deliver strong forecasting improvements with low computational cost and clear interpretability.

3 Proposed Methodology

The proposed model is illustrated in Fig. 1. It consists of four main stages, starting from data preprocessing and ending with forecasting evaluation.

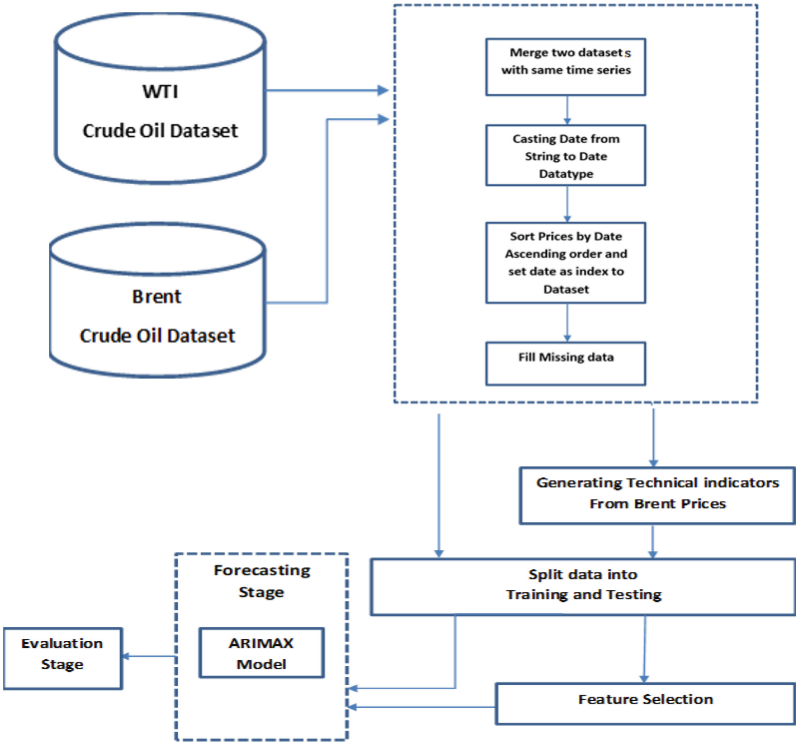

Fig. 1. Steps of the proposed methodology.

3.1 Dataset and Preprocessing

There are several crude oil markets around the world. The dataset in this study originates from the US Energy Information Administration, a major online data repository for energy statistics. The daily historical WTI and Brent Oil prices from January 1, 1988, through December 31, 2019, were downloaded as a CSV file containing only dates and pricing, as shown in Table 1 and Table 2.

Table 1. The daily time series of WTI crude oil prices

Date	Price
4-Jan-16	36.28
5-Jan-16	35.56
6-Jan-16	33.89
7-Jan-16	33.57

Table 2. The daily time series of Brent oil prices

Date	Price
4-Jan-16	33.57
5-Jan-16	31.67
6-Jan-16	30.14
7-Jan-16	29.14

The WTI and Brent datasets are imported into the system. The `Date` column is converted from string format to a proper date datatype and assigned as the index of the dataset. The time series is sorted in ascending hronological order. Since crude oil markets are closed on weekends, Saturday and Sunday entries are missing. To ensure continuity, missing weekend prices are filled using the last available Friday price as shown in Table 3 and Table 4.

3.2 Technical Indicators

Technical indicators are mathematical techniques that may detect whether a stock is under- or over-priced or whether a trend is weak or strong, hence estimating implied volatility. Indicators can also provide information on previous patterns and the present state of the market [24]. Technical indicators are divided into three groups based on their functions: lagging, leading, and volatility. Lagging indicators, often known as trend indicators, track prior price behavior. The main examples of lagging indicators are MA and MACD. Leading or momentum-based indicators are designed to forecast future price trends and reflect price

Table 3. The price of crude oil before gap-filling

Date	Price
4-Jan-19	53.3
5-Jan-19	51.57
6-Jan-19	52.29
7-Jan-19	53.95
10-Jan-19	53.33
11-Jan-19	53.03
12-Jan-19	51.13

Table 4. The price of crude oil after gap-filling

Date	Price
4-Jan-19	53.3
5-Jan-19	51.57
6-Jan-19	52.29
7-Jan-19	53.95
8-Jan-19	53.95
9-Jan-19	53.95
10-Jan-19	53.33
11-Jan-19	53.03
12-Jan-19	51.13

change rates. Leading indicators such as ROC and RSI are the most well-known. Finally, price volatility is measured using volatility-based indicators, the most commonly used of which is BB [32]. The lagging technical indicators used in this study are Standard Deviation (SD), Simple Moving Average (SMA), Exponential Moving Average (EMA), and Moving Average Convergence/Divergence (MACD). The standard deviation is a statistical measure of how volatile something is. In general, this measures the degree to which the observations in a dataset are dispersed around an average. SD is calculated using Eq. (1).

$$SD(t) = \sigma = \sqrt{\frac{\sum_{r}^{t} = t - n\left(A_{x_r} - SMA_{A_x}(n)\right)^2}{n}} \tag{1}$$

where SMA is the simple moving average, and n refers to the number of samples. The moving average (MA (is a trend-following indicator that reveals what is happening in a given market. Moving averages highlight a trend's direction while also smoothing out market volatility. Different forms of moving averages, such as the simple moving average (SMA) and the exponential moving average (EMA), have different ways of determining moving average prices, depending on how

historical values are weighted. The SMA is an essential mean value that uses the same weights as previous prices, while EMA gives more weight to the most recent prices. The SMA and EMA are calculated in Eqs. (2 and 3) below, respectively [15]:

$$SMA_n(t) = \frac{\sum_{k=t-n+1}^{t} P(k)}{n} \tag{2}$$

where $P(k)$ refers to the price at time period k.

$$EMA_n(t) = P(t) * a + (1 - a) * EMA_n(t - 1) \tag{3}$$

EMAn represents the EMAn(t-1) of today's price, represents the EMA of day ago (yesterday) t, and a = 2/ (n + 1). $MACD$ and $MACD$ signals are two indicators provided by the $MACD$. $MACD$ shows the difference between a fast and a slow EMA of closing prices. A short-period average is called "fast," whereas a long-period average is called "slow" [13]. $MACD$ is the difference between the 12-period and 26-period $EMAs$ based on the preset parameters as shown in Eqs. (4 and 5):

$$MACD = EMA(12) - EMA(26) \tag{4}$$

$$MACDSignal = EMA(MACD, 9) \tag{5}$$

The normal values for the parameters of MACD (12, 26, and 9) can be changed to suit the traders' demands. This study used the default MACD parameter settings because this value set is commonly recognized and used worldwide [12]. In our research with ARIMAX, the four lagging technical indicators listed above are employed initially as exogenous or affected factors. The rest of the types of technical indicators will used later in future work.

Technical Indicator Selection via Correlation Filtering. Not all generated indicators that are shown in Table 5 are suitable for the forecasting task; including weak or redundant indicators can reduce the efficiency of the forecasting method and increase running time. Therefore, a linear correlation coefficient (R) is employed as a filter-based feature selection technique to measure the relationship between each Brent-derived indicator and WTI prices, then rank indicators by significance and retain only the most informative set.

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum_i (x_i - \bar{x})^2 (y_i - \bar{y})^2} \tag{6}$$

Final Indicator Set After Correlation Filtering. To justify the feature selection process without overloading the manuscript, Table 6 summarizes the correlation between WTI and the most relevant Brent-derived indicators.

The SMA and EMA indicators exhibit extremely strong correlation ($R >$ 0.96), whereas volatility (STD) and MACD-related variables show weak association. Therefore, the filter method prioritized SMA/EMA indicators, while

Table 5. Correlation matrix between WTI and Brent-derived technical indicators

	wti	signal_ macd_ brent	macd_brent	brent_sma_30	brent_ema_30	brent_sma_7	brent_ema_7	brent_sma_3	brent_ema_3	brent_std_7	brent_std_3	brent_std_30
wti	1.000000	0.146324	0.135390	0.968040	0.969993	0.974563	0.974990	0.975574	0.975744	0.120260	0.085329	0.105543
signal_ macd_ brent	0.146324	1.000000	0.958017	0.052689	0.064069	0.139496	0.136605	0.146178	0.145539	−0.111843	−0.037486	−0.230733
macd_brent	0.135390	0.958017	1.000000	0.023347	0.038977	0.119302	0.117445	0.133060	0.132127	−0.133674	−0.056420	−0.204900
brent_sma_30	0.968040	0.052689	0.023347	1.000000	0.999605	0.994936	0.995183	0.993035	0.993176	0.173542	0.116827	0.170535
brent_ema_30	0.969993	0.064069	0.038977	0.999605	1.000000	0.996679	0.996860	0.995051	0.995168	0.172288	0.116284	0.171470
brent_sma_7	0.974563	0.139496	0.119302	0.994936	0.996679	1.000000	0.999910	0.999361	0.999390	0.160401	0.111287	0.154314
brent_ema_7	0.974990	0.136605	0.117445	0.995183	0.996860	0.999910	1.000000	0.999621	0.999658	0.160062	0.110509	0.154994
brent_sma_3	0.975574	0.146178	0.133060	0.993035	0.995051	0.999361	0.999621	1.000000	0.999956	0.156045	0.107786	0.153245
brent_ema_3	0.975744	0.145539	0.132127	0.993176	0.995168	0.999390	0.999658	0.999956	1.000000	0.156287	0.107853	0.153434
brent_std_7	0.120260	−0.111843	−0.133674	0.173542	0.172288	0.160401	0.160062	0.156045	0.156287	1.000000	0.420930	0.338171
brent_std_3	0.085329	−0.037486	−0.056420	0.116827	0.116284	0.111287	0.110509	0.107786	0.107853	0.420930	1.000000	0.166809
brent_std_30	0.105543	−0.230733	−0.204900	0.170535	0.171470	0.154314	0.154994	0.153245	0.153434	0.338171	0.166809	1.000000

Table 6. Correlation Summary (WTI vs Selected Indicators)

Indicator	Correlation with WTI
brent_sma_3	0.9756
brent_ema_3	0.9757
brent_sma_7	0.9746
brent_ema_7	0.9750
brent_sma_30	0.9680
brent_ema_30	0.9700
brent_std_7	0.1203
brent_std_3	0.0853
brent_std_30	0.1055
macd_brent	0.1354
signal_macd_brent	0.1463

discarding weak predictors to avoid unnecessary complexity and performance degradation.

Based on this empirical evidence, a filter-based selection strategy is adopted to retain only highly informative indicators while discarding weak predictors that may introduce noise and unnecessary model complexity. Consequently, the final exogenous feature set prioritizes SMA and EMA indicators with multiple lag configurations (3, 7, and 30 days). These selected indicators are then incrementally incorporated into the ARIMAX model to evaluate their individual and cumulative contribution to forecasting accuracy. This strategy ensures improved predictive performance while maintaining interpretability and computational efficiency.

3.3 Forecasting Models

This section provides an overview of the mathematical foundations of the time series techniques utilized and researched in this article. Before delving into the ARIMAX model, it is necessary to have an explanation of the background of Autoregressive (AR), Moving Average (MA), Autoregressive Moving Average (ARMA), and Autoregressive Integrated Moving Average (ARIMA) modelling techniques.

1. Autoregressive (AR): The AR is a lagging linear time series forecasting methodology. Equation (7) is used in an AR model with p lags to determine the value for period t in a time series [25].

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots \phi_p y_{t-p} + \varepsilon_t \tag{7}$$

2. Moving Average (MA) The MA model only relies on the lagged forecast errors, which are the AR model errors of the relevant lags. Equation (8) is for the

MA model [6].

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \tag{8}$$

where y_t is the value of observation in period t, c is a constant, θ is a numeric constant by which we multiply the residuals, ε is the residual of time t.

3. Autoregressive Moving Average (ARMA) The fundamental approach for assessing a stationary time series is Auto-Regressive Moving Average (ARMA). The ARMA model is a mixture of AR and MA models. The AR model explains the movements and means reversion effects, whereas the shock effects in the white noise terms are captured by the MA model [30]. The ARMA model is expressed in Eq. (9):

$$y_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi y_{t-1} + \sum_{j=1}^{q} \theta_{j\varepsilon_{t-1}} \tag{9}$$

where y_t is the value of prediction in period t, c is a constant, φ is a coefficient that multiplies the residuals and ε_t represents an error in period t.

4. Autoregressive Integrated Moving Average (ARIMA): An ARIMA model works with non-stationary data by reducing it using the differencing technique [25]. The model employs three parameters: p, d, and q. The p parameter determines the number of lag periods [19,30]. When $p = 2$ is used in the autoregression component of the equation, for example, two previous periods of the time series are utilized. The number of differencing transformations used to remove trends and/or seasonality, resulting in a stationary time series, is represented by parameter d. The lag of the error component of the ARIMA model is represented by the parameter q [34].
5. Autoregressive Integrated Moving Average with Exogenous Variable (ARIMAX): The ARIMAX model is a developed or enhanced version of the ARIMA model that includes factors and extra variables. It may be thought of as a multiple regression model with one or more autoregressive (AR) and/or moving average (MA) variables [26]. This method may be used to forecast stationary or non-stationary data and any data time series patterns such as trend, seasonality, or cyclist [6]. The ARIMAX model requires independent variables that perform as extra variables. The variables used in this inquiry, technical indicators of another crude oil price benchmark Brent are employed.
6. Auto ARIMA: Automated Machine Learning (AutoML) frameworks have been developed to reduce manual effort in model selection and hyperparameter optimization. For instance, H2O AutoML [35] employs ensemble learning and stacked models, TPOT uses genetic programming to evolve optimal machine learning pipelines [3], Auto-WEKA applies Bayesian optimization for algorithm and parameter selection [31], and AutoKeras focuses on automated deep learning through neural architecture search (NAS) and automated model tuning [22]. While these frameworks are primarily designed for classification and regression tasks, specialized AutoML techniques have also been proposed for time-series forecasting. One such approach is AutoARIMA, which automatically identifies the optimal (p, d, q) parameters of the ARIMA model for

temporal data analysis. In the context of time-series forecasting, AutoARIMA represents a specialized AutoML approach that automatically determines the optimal ARIMA (p, d, q) parameters, making it particularly suitable for modeling temporal dependencies in sequential data. Even though ARIMA is a powerful model for forecasting time series data, the data preparation and parameter tuning operations are quite time-consuming. Before adopting ARIMA, the series must be made stationary, and plots must be used to establish the values of p and q (ACF and PACF) [4]. This work is simplified by the Auto ARIMA function, which is frequently used to find the best model by automatically producing a set of optimum parameters by evaluating all possible (p, d, q) combinations and returning the model with the lowest Akaike information criterion (AIC) and Bayesian Information Criterion (BIC) values [4,34].

3.4 Evaluation Metrics

There are numerous different methods to measure how well a model works. The most popular metrics used for forecasting accuracy are MAE and RMSE, defined in Eqs. (10), (11), and (12) [11].

$$MAE = \frac{1}{n-m} \sum_{t=m+1}^{n} |y_t - \hat{y}_t| \tag{10}$$

$$RMSE = \sqrt{\frac{1}{n-m} \sum_{t=m+1}^{n} (y_t - \hat{y}_t)^2} \tag{11}$$

$$MAPE = 100.\frac{1}{N} \sum_{i=1}^{N} \frac{|y_t - \hat{y}|}{y_t} \tag{12}$$

R-squared is another metric used to measure the performance of a forecasting model. It is usually between 0 and 1. A model's fit is better if the R-squared is close to 1. Indicates how much of the volatility of the dependent variable can be predicted by the independent variables or features. Equation (13) is used to compute R-squared [14]. Where y_i is the outcome value from the forecasting model and $\hat{y}$ is the actual value.

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - \hat{y})^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2} \tag{13}$$

where N represents the number of samples, y_t is the forecasted value, and $\hat{y}$ is the actual value.

4 Results and Discussion

This section evaluates the impact of Brent-derived technical indicators on improving WTI crude oil forecasting accuracy.

The WTI dataset from January 1, 2010 to December 31, 2019 is used in the experiments. The data are split into 70% training and 30% testing sets. The training data are used to estimate model parameters, while the testing data evaluate predictive performance.

4.1 Technical Indicator Generation

In this section, we generate technical indicators from the time series of the Brent crude oil utilizing the prior equations clarified in the methodology section and then utilize them as an exogenous ARIMAX variable. Three sets of values from the past are used to create each technical indicator (SMA, SD, and EMA): one looking back on the previous three days, one from the previous seven days, and one from the previous 30 days. These values stand in for the last three days, weeks, and months. MACD is assigned 12 lag and 26 lag as default parameters, and MACD signals use nine as the default parameter, as illustrated in Table 7.

Table 7. The intervals of a daily technical indicator time

No.	Technical indicators	Selected period of lags
1	SMA	3 lags-7 lags-30 lags
2	SD	3 lags-7 lags-30 lags
3	EMA	3 lags-7 lags-30 lags
4	MACD	12 lags-26 lags-9 lags

4.2 Forecasting Model

We used Auto ARIMA and the filter method to add technical indicators with a high degree of significance as an exogenous variable to help the model perform well. As mentioned previously, we use the filter method for feature selection, begin running the model with a null feature, and then fit it with the indicators, one at a time, with the highest degree of correlation selected. As shown in Fig. 2 and 3, the forecasting model begins to improve and further enhancement is observed in Fig. 4 and Fig. 5.

The final improvement stage is illustrated in Fig. 6, 7 and Fig. 8. In all charts, a forecast model is represented in a green line, while the orange represents the actual value. Looking at the simple drawings above, it can be seen that technical indicators affect expectations. Each time a technical indicator is added, the model improves. In Table 8, we will calculate RMSE, MAE, MAPE, and AIC for each stage and compare them to see how much the forecasting model improves.

The final model achieves the lowest MAE, RMSE, MAPE, and AIC values, demonstrating that incorporating carefully selected Brent-derived indicators significantly enhances WTI forecasting accuracy.

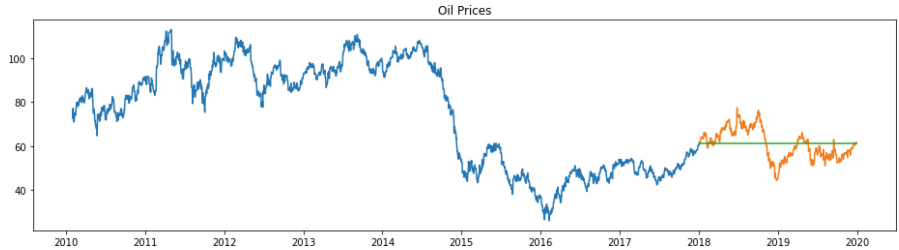

Fig. 2. ARIMAX without exogenous variable

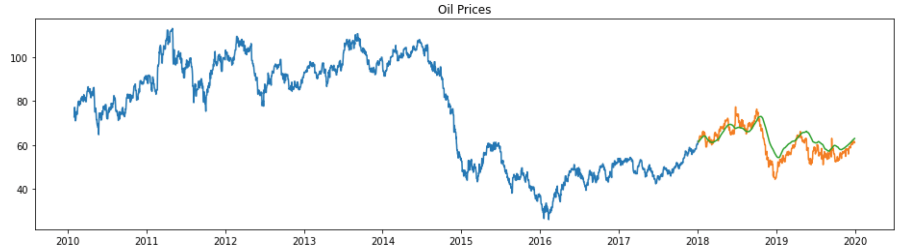

Fig. 3. ARIMAX with brent_sma_30

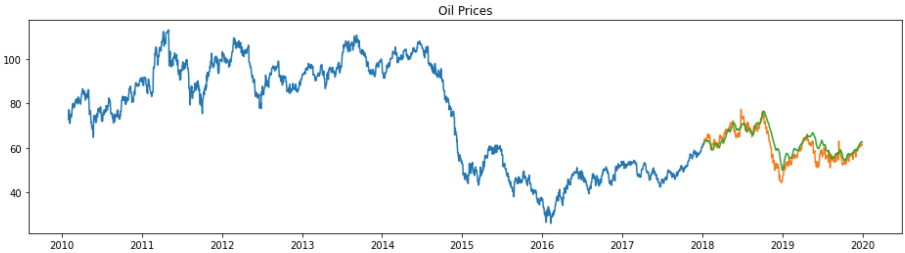

Fig. 4. ARIMAX with brent_sma_30 and brent_ema_30

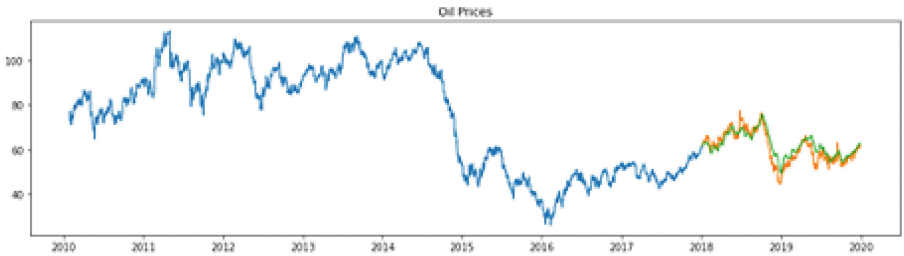

Fig. 5. ARIMAX with brent_sma_30, brent_ema_30, and brent_sma_7

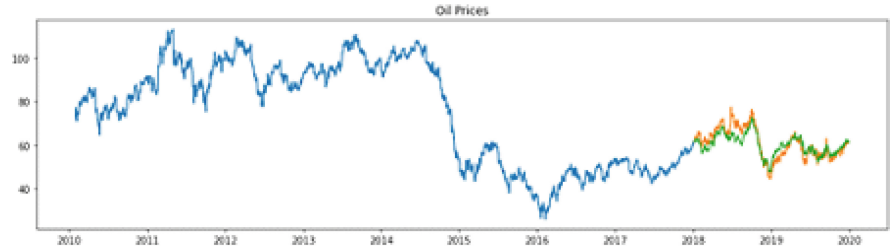

Fig. 6. ARIMAX with brent_sma_30, brent_ema_30, and brent_ema_7

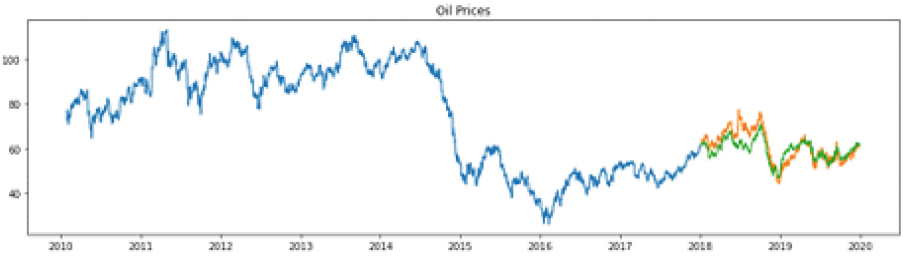

Fig. 7. ARIMAX with brent_sma_30, brent_ema_30, brent_ema_7, and brent_sma_3

4.3 Computational Efficiency and Deployment Considerations

A key motivation of this study is practicality: in many industrial forecasting pipelines, models must be retrained frequently and executed under strict computational budgets. Deep learning approaches such as LSTM/CNN hybrids can yield high accuracy, but they often require GPU acceleration, extensive hyperparameter tuning, and longer training time, especially when retraining is performed frequently to adapt to evolving market regimes.

In contrast, ARIMAX enhanced with a small, well-selected set of exogenous indicators offers three practical advantages: (i) it can be trained quickly using classical likelihood-based estimation and Auto ARIMA search, (ii) it does not require complex neural architecture selection, and (iii) it is easier to interpret and monitor in production settings. These characteristics make the proposed approach well-suited for online or near-real-time forecasting contexts where the model must be updated as new prices arrive.

4.4 Practical Implications for Energy Forecasting

Oil price forecasts have immediate consequences for budgeting, hedging, and risk management. Improving WTI forecasts using Brent-derived information can be valuable in multiple contexts, such as:

- planning and procurement in energy-intensive industries,
- hedging strategies for producers and refiners,

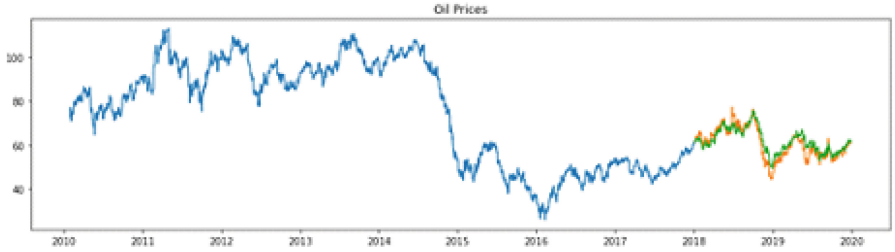

Fig. 8. ARIMAX with brent_sma_30, brent_sma_7, brent_ema_7, and brent_ema_3

Table 8. Accuracy measures of the ARIMAX model across incremental experiments

Figure	MAE	RMSE	MAPE	AIC
Fig. 1	5.7447	6.7965	9.6686	9058.027
Fig. 2	3.6121	4.6262	6.2933	9014.384
Fig. 3	2.7577	3.6896	4.7785	8791.368
Fig. 4	2.5368	3.2396	4.3549	8778.644
Fig. 5	2.5400	3.2143	4.3571	8098.878
Fig. 6	3.3100	4.1398	5.2880	7978.189
Fig. 7	2.3166	2.9148	3.9710	7566.847

- macroeconomic planning for oil-exporting and oil-importing economies,
- downstream forecasting tasks (inflation, exchange rates, and sector-level indices).

Intermarket feature transfer is particularly beneficial when a benchmark incorporates information earlier than another market (e.g., due to liquidity, trading hours, or regional shocks). The findings of this study provide evidence that such transfer can be realized using light-weight statistical learning rather than computationally expensive deep learning.

4.5 Comparison with Related Work

The proposed ARIMAX model is compared with previous studies that utilize ARIMA, ANN, and LSTM-based approaches. Despite its lower computational complexity, the proposed model achieves competitive or superior accuracy as shown in

5 Conclusion and Future Work

The capacity of ARIMAX to model the WTI crude oil price using exogenous data from another time series was investigated in this research. The lagged technical indicators MACD, SMA, EMA, and SD, were extracted from a Brent crude time

Author	year	Title	Period	Model	RMSE	MAE	MAPE
Ting Yao (Yao and Zhang, 2017)	2020	Forecasting crude oil prices with the Google Index	From 2004-01-01 To 2016-06-01	ARIMA	4.501	3.790	82.993
				ARIMAX with google index	6.072	5.072	138.751
				ARMA-GARCH	4.265	3.830	97.329
				ARIMAX-GARCH	6.361	5.102	296.431
		Proposed Models		ARIMAX with a technical indicator	2.0189	4.0759	4.56

Author	Year	Title	Period	Model	R-Square
Guo, Junhui (Guo, 2019)	2019	Oil price forecast using deep learning and3 ARIMA	From 1986-01-01 To 2018-07-01	ARIMA	0.69
				LSTM	0.80
				GRU	0.81
				CNN	0.83
		Proposed Models		ARIMAX with a technical indicator	0.8679

Author	Year	Title	Period	Model	RMSE	R-Square
D. Makala and Z. Li (Makala and Li, 2019)	2019	Economic Forecasting with Deep Learning: Crude Oil	From 2009-03-01 To 2019-03-01	Moving Average	23.79	0.77
				Linear Regression	24.77	0.32
				ARIMA	106.03	0.50
				LSTM	1.18	0.97
		Proposed Models		ARIMAX with a technical indicator	2.5288	0.9425

Author	Year	Title	Period	Model	RMSE	MAPE
Herrera, Gabriel (Herrera et al., 2019)	2019	Long-term forecast of energy commodities price using machine learning	From 1980-01-01 To 2017-06-01	ANN	14.7208	16.5554
				Random forests	7.8539	8.9603
				Hybrid	23.3935	33.7099
		Proposed Models		ARIMAX with a technical indicator	8.8064	13.25

Fig. 9. Comparison of forecasting accuracy with related works.

series and employed as external factors in this study. The experimental results show that this prediction approach produces the best or closest results to those obtained from the Related Works. Thus, we can conclude that the technical indicators derived from a Brent oil time series substantially impact the forecasted WTI crude oil price. Future research on this topic could go in several interesting directions; for example, utilizing a different crude oil price, stock prices in the United States, major European stock markets, and the EUR/USD and USD/JPY exchange rates all provide valuable information for projecting crude oil prices. Although the results are strong, several limitations should be recog-

nized. First, the selected indicators are primarily lagging indicators; momentum and leading indicators (e.g., RSI, ROC) were not included in the present experiments. Second, correlation-based filtering captures only linear association; nonlinear dependence may exist and could justify alternative selection criteria in future work. Third, the evaluation is based on a specific historical period; performance under major regime shifts may require re-validation and adaptive retraining strategies. These limitations open clear directions for future extensions without changing the core intermarket idea demonstrated here.

Acknowledgements. This research was sponsored by the University of Malaya Translational Research Grant Scheme, TRG0019-2025 (UM.0002388/HRU.IP) and the Research Excellence Grant, UMREG 068 - 2024 (UM.0000285/HRU.RC).

Conflicts of Interest. The authors declare no conflict of interest.

References

1. Ahmed, R.A., Shabri, A.B.: Daily crude oil price forecasting model using Arima, generalized autoregressive conditional heteroscedastic and support vector machines. Am. J. Appl. Sci. **11**(3), 425 (2014)
2. Al-Qazzaz, R.A., Yousif, S.A.: High performance time series models using auto autoregressive integrated moving average. Indones. J. Electr. Eng. Comput. Sci **27**, 422–430 (2022)
3. Al-Shaikhli, H.I., Yousif, S.A., Abdulhameed, T.Z., Obaiys, S.J.: Riceaml: An auto machine learning model to identify abiotic stress-associated single nucleotide polymorphisms in rice. In: Gervasi, O., et al. (eds.) ICCSA 2025. LNCS, vol. 15649, pp. 81–96. Springer, Cham (2025). https://doi.org/10.1007/978-3-031-96997-3_6
4. Alabdulrazzaq, H., Alenezi, M.N., Rawajfih, Y., Alghannam, B.A., Al-Hassan, A.A., Al-Anzi, F.S.: On the accuracy of Arima based prediction of COVID-19 spread. Resu. Phys. **27**, 104509 (2021)
5. Alzubaidi, L., et al.: Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J. Big Data **8**(1), 1–74 (2021)
6. Anggraeni, W., Vinarti, R.A., Kurniawati, Y.D.: Performance comparisons between Arima and AriMax method in MOSLEM kids clothes demand forecasting: case study. Procedia Comput. Sci. **72**, 630–637 (2015)
7. Aziz, N., Abdullah, M.H.A., Zaidi, A.N.: Predictive analytics for crude oil price using RNN-LSTM neural network. In: 2020 International Conference on Computational Intelligence (ICCI), pp. 173–178. IEEE (2020)
8. Bollapragada, R., Mankude, A., Udayabhanu, V.: Forecasting the price of crude oil. Decision **48**(2), 207–231 (2021). https://doi.org/10.1007/s40622-021-00279-5
9. Büyükşahin, Ü.Ç., Ertekin, Ş: Improving forecasting accuracy of time series data using a new Arima-ANN hybrid method and empirical mode decomposition. Neurocomputing **361**, 151–163 (2019)
10. Chen, Y., He, K., Tso, G.K.: Forecasting crude oil prices: a deep learning based model. Procedia Comput. Sci. **122**, 300–307 (2017)
11. Ciulla, G., D'Amico, A.: Building energy performance forecasting: a multiple linear regression approach. Appl. Energy **253**, 113500 (2019)
12. Deng, S., Sakurai, A.: Crude oil spot price forecasting based on multiple crude oil markets and timeframes. Energies **7**(5), 2761–2779 (2014)

13. Fernández-Blanco, P., Bodas-Sagi, D.J., Soltero, F.J., Hidalgo, J.I.: Technical market indicators optimization using evolutionary algorithms. In: Proceedings of the 10th Annual Conference Companion on Genetic and Evolutionary Computation, pp. 1851–1858 (2008)
14. Guo, J.: Oil price forecast using deep learning and Arima. In: 2019 International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI), pp. 241–247. IEEE (2019)
15. Gupta, N., Nigam, S.: Crude oil price prediction using artificial neural network. Procedia Comput. Sci. **170**, 642–647 (2020)
16. Herrera, G.P., Constantino, M., Tabak, B.M., Pistori, H., Su, J.J., Naranpanawa, A.: Long-term forecast of energy commodities price using machine learning. Energy **179**, 214–221 (2019)
17. Kamdem, J.S., Essomba, R.B., Berinyuy, J.N.: Deep learning models for forecasting and analyzing the implications of COVID-19 spread on some commodities markets volatilities. Chaos Sol. Fract. **140**, 110215 (2020)
18. Kumar, S., Ningombam, D.: Short-term forecasting of stock prices using long short term memory. In: 2018 International Conference on Information Technology (ICIT), pp. 182–186. IEEE (2018)
19. Madan, R., Mangipudi, P.S.: Predicting computer network traffic: a time series forecasting approach using dwt, Arima and RNN. In: 2018 Eleventh International Conference on Contemporary Computing (IC3), pp. 1–5. IEEE (2018)
20. Makala, D., Li, Z.: Economic forecasting with deep learning: crude oil. MATTER: Int. J. Sci. Technol. **5**(2) (2019)
21. Malik, A.: The impact of oil price changes on inflation in Pakistan. Int. J. Energy Econ. Pol. **6**(4), 727–737 (2016)
22. Matti, R., Yousif, S.: Autokeras for fake news identification in Arabic: leveraging deep learning with an extensive dataset. Al-Nahrain J. Sci. **26**(3), 60–66 (2023)
23. Mensah, E.K.: Box-Jenkins modelling and forecasting of Brent crude oil price (2015)
24. Nguyen, H.N., Zhou, J., Kang, H.J.: A calibration method for enhancing robot accuracy through integration of an extended Kalman filter algorithm and an artificial neural network. Neurocomputing **151**, 996–1005 (2015)
25. Ohyver, M., Pudjihastuti, H.: Arima model for forecasting the price of medium quality rice to anticipate price fluctuations. Procedia Comput. Sci. **135**, 707–711 (2018)
26. Rahmayanti, I.A., Andreas, C., Ulyah, S.M.: Does us-china trade war affect the brent crude oil price? an arimax forecasting approach. In: AIP Conference Proceedings, vol. 2329, p. 060010. AIP Publishing LLC (2021)
27. Ramyar, S., Kianfar, F.: Forecasting crude oil prices: a comparison between artificial neural networks and vector autoregressive models. Comput. Econ. **53**, 743–761 (2019)
28. Redha, R.A., Yousif, S.A.: Crude oil price forecasts using support vector regression and technical indicators. Iraqi J. Sci. 2512–2520 (2023)
29. Redha, R.A., Yousif, S.A.: Crude oil price forecasts using support vector regression and technical indicators. Iraqi J. Sci. **64**(5), 2512–2520 (2023). https://doi.org/10.24996/ijs.2023.64.5.34
30. Şahinli, M.A.: Potato price forecasting with holt-winters and Arima methods: a case study. Am. J. Potato Res. **97**(4), 336–346 (2020)
31. Samawi, V.W., Yousif, S.A., Al-Saidi, N.M.: Intrusion detection system: an automatic machine learning algorithms using auto-Weka. In: 2022 IEEE 13th control and system graduate research colloquium (ICSGRC), pp. 42–46. IEEE (2022)

32. Shynkevich, Y., McGinnity, T.M., Coleman, S.A., Belatreche, A., Li, Y.: Forecasting price movements using technical indicators: investigating the impact of varying input window length. Neurocomputing **264**, 71–88 (2017)
33. Yao, T., Zhang, Y.J.: Forecasting crude oil prices with the google index. Energy Procedia **105**, 3772–3776 (2017)
34. Yermal, L., Balasubramanian, P.: Application of auto Arima model for forecasting returns on minute wise amalgamated data in NSE. In: 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), pp. 1–5. IEEE (2017)
35. Yousif, S.A., Samawi, V.W., Al-Saidi, N.M.: Automatic machine learning classification algorithms for stability detection of smart grid. In: 2022 IEEE 5th International Conference on Big Data and Artificial Intelligence (BDAI), pp. 34–39. IEEE (2022)