Analyzing the Impact of Global Events on Oil Prices

Literature Review, Data Research and Technology Review

Group 18:

no	Group members
1	Esmael Aliyi Uta
2	Furtuna Yimer Tilahun
3	Ermias Brhane Berhe

Literature Review

1. Introduction

Forecasting crude oil prices is essential in global energy economics due to the commodity's strong influence on financial markets, production planning, and policy decisions. However, the inherent volatility of oil prices—driven by geopolitical events, supply-demand dynamics, and macroeconomic shifts—poses significant modeling challenges. To address these, researchers have employed a wide range of time series forecasting models.

This literature review critically examines two widely used forecasting methods: the AutoRegressive Integrated Moving Average (ARIMA) model, a traditional statistical approach, and Long Short-Term Memory (LSTM) networks, a deep learning technique. By analyzing their theoretical underpinnings, applications, performance, and limitations, this review provides a foundation for comparative analysis in crude oil price prediction.

2. ARIMA: The Classical Benchmark

2.1 Theoretical Foundations

ARIMA, formulated by Box and Jenkins (1976), is a univariate time series model that captures linear temporal dependencies. It integrates three components:

Autoregression (AR): Models the relationship between current and past values.

Integration (I): Ensures stationary by differencing the series.

Moving Average (MA): Accounts for serial correlation in residuals.

The model is valued for its mathematical rigor and ease of interpretation, making it a strong baseline in many forecasting tasks.

2.2 Applications in Oil Price Forecasting

ARIMA has been widely applied in oil price forecasting. For example, Yu et al. (2008) reported Mean Absolute Percentage Errors (MAPE) of 3–5% for short-term predictions of West Texas Intermediate (WTI) prices. Similarly, Abramson and Finizza (1991) demonstrated ARIMA's capability to model oil price trends under relatively stable market conditions.

However, ARIMA assumes linearity and stationarity, which limit its ability to handle abrupt shocks or complex dynamics common in energy markets.

2.3 Limitations

Linearity Assumption: ARIMA performs poorly when dealing with nonlinear patterns, such as sudden geopolitical disruptions or demand shocks.

Stationarity Requirement: Differencing may obscure long-term trends or cyclical behavior.

Univariate Nature: Although ARIMAX extends the model to include exogenous variables, it introduces complexity and remains limited in handling high-dimensional inputs.

3. LSTM: A Deep Learning-Based Alternative

3.1 Architecture and Modeling Strengths

Long Short-Term Memory (LSTM) networks, introduced by Hochreiter and Schmidhuber (1997), are a specialized form of Recurrent Neural Networks (RNNs) designed to learn long-term dependencies in sequential data. Unlike ARIMA, LSTMs can model both linear and nonlinear relationships.

Key architectural features include:

Memory Cells: Preserve information over long sequences.

Gating Mechanisms: Regulate information flow (input, forget, and output gates), enabling the model to retain or discard data dynamically.

3.2 Applications in Oil Price Forecasting

Sagheer and Kotb (2019) applied LSTM networks to oil production and pricing data, achieving 20–30% lower Root Mean Square Error (RMSE) compared to ARIMA. Their study highlights LSTM's effectiveness in learning from complex, noisy, and nonlinear time series typical of commodity markets.

3.3 Challenges

Despite their high accuracy, LSTMs present practical challenges:

Data Requirements: LSTMs require large datasets (typically over 5,000 samples) to achieve stable and generalizable performance.

Model Complexity: Training involves careful tuning of multiple hyperparameters, such as the number of layers, units, and sequence length.

Interpretability: Unlike ARIMA's transparent structure, LSTMs operate as black-box models, which may limit their use in policy-sensitive contexts.

4. Research Gaps in ARIMA vs. LSTM Comparisons

Despite the extensive use of both models, direct comparisons in oil price forecasting are limited. The literature often focuses on hybrid or deep ensemble models, leaving a gap in understanding how ARIMA and LSTM perform under equivalent conditions.

4.1 Methodological Gaps

Few studies apply both models to the same dataset with identical preprocessing, making comparisons inconsistent.

Most LSTM studies focus on predictive accuracy, while interpretability and computational efficiency are underexplored.

4.2 Contextual Suitability

ARIMA may be more suitable for short-term, low-variance series with limited data.

LSTM is better suited for long-term forecasting in complex, high-variance, or multivariate environments.

5. Our Contribution

This project aims to directly address these gaps through a controlled comparative study of ARIMA and LSTM:

Data Sources: We will use daily and monthly datasets of WTI and Brent crude oil prices from 2000 to 2024.

Experimental Design: Both models will be trained on the same dataset, with comparisons made across univariate and multivariate settings.

Evaluation Metrics: Forecast accuracy will be assessed using MAPE, RMSE, and Diebold-Mariano tests to ensure statistical robustness.

6. Conclusion

ARIMA and LSTM each bring distinct advantages to crude oil price forecasting. ARIMA offers simplicity, interpretability, and efficiency in stable market conditions, while LSTM provides superior performance in dynamic and nonlinear environments. This review establishes a strong foundation for systematically comparing these models and informing model selection based on forecasting objectives and data characteristics.

References

Abramson, B., & Finizza, A. J. (1991). Using time series analysis for oil price forecasting. Energy Economics, 13(1), 31–35.

Box, G. E. P., & Jenkins, G. M. (1976). Time Series Analysis: Forecasting and Control. Holden-Day.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

Sagheer, A., & Kotb, M. (2019). Time series forecasting of petroleum production using deep LSTM recurrent networks. Neurocomputing, 323, 203–213.

Yu, L., Wang, S., & Lai, K. K. (2008). Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. Energy Economics, 30(5), 2623–2635.

2. Data Research

Introduction

This data research explores the relationship between global events and oil price volatility, a critical factor in energy market stability and sustainable economic planning. Investigating this connection is essential to answer how specific events drive price changes and why data-driven insights are vital for informed decision-making in energy policy.

Organization

The findings are organized chronologically based on the dataset's time span, with thematic insights drawn from event-price correlations.

Data Description

The data set contains historical Brent oil prices. It includes daily prices from May 20, 1987, to September 30, 2022. Total records: 9,011 rows.

Data fields

- Date: Represents the date of the recorded Brent oil price. Each entry is formatted as 'day-month-year' (e.g., 20-May-87). The dataset covers daily prices from May 20, 1987, to September 30, 2022.
- Price: This column represents the price of Brent oil on the corresponding date. The price is recorded in USD per barrel.
- Data Source: Historical Brent crude oil prices sourced from yahoo finance
- Data Size: Covers May 20, 1987, to September 30, 2022, and the size is 152 KB.

Here's a quick look at the first few entries:

Date	Price
20-May-87	18.63
21-May-87	18.45
22-May-87	18.55
25-May-87	18.60
26-May-87	18.63

Selection Rationale: Brent crude is a global benchmark, reflecting international market dynamics influenced by diverse events. This dataset's long time span and daily granularity enable precise event-impact analysis.

Data Analysis and Insights
Preliminary analysis reveals distinct price spikes tied to major events:
1990–1991 Gulf War: Prices surged from $17 to $36 per barrel due to supply fears.
2008 Financial Crisis: Prices dropped from $144 to $33 amid demand collapse.
-2014 OPEC Production Increase: Prices fell from $112 to $55 as supply flooded the market.
Descriptive statistics (e.g., mean price ~$58, standard deviation ~$32) indicate high volatility, while visualizations (e.g., time series plots) highlight event-driven peaks and troughs. These patterns suggest that geopolitical and policy events significantly disrupt price stability.

Conclusion
The data analysis confirms that global events like wars, economic crises, and OPEC decisions—drive substantial oil price fluctuations. The analysis of Brent Oil prices from 1987 to 2020 reveals substantial volatility, with global economic factors playing a significant role in price fluctuations. By observing the raw price data and the smoothed 1-year rolling average, it is clear that while there are periods of stability, prices are highly responsive to geopolitical and economic events.

3 Technology Review

Introduction
This technology review evaluates time series analysis tools likeARIMA, GARCH, and LSTM used to model the impact of global events on oil prices. Selecting appropriate technologies is crucial for accurate forecasting and volatility analysis, directly supporting the project's aim of delivering actionable energy market insights.

Technology Overview

Purpose

The technologies reviewed are designed to analyze historical oil price data and correlate it with significant global events. The primary purpose is to develop insights that help stakeholders understand how these events impact oil price fluctuations, allowing for better-informed decisions in a volatile market environment.

Key Features

1. Data Aggregation: These tools gather data from diverse sources, including historical oil prices and records of geopolitical events, to create a comprehensive dataset for analysis.

2. Predictive Modeling: Advanced algorithms utilize historical data to forecast future price changes based on identified global events.

3. Visualization Tools: User-friendly dashboards present data trends and relationships visually, making complex information accessible for decision-making.

4. Machine Learning Techniques: Models such as ARIMA, GARCH, and LSTM are employed to improve predictive accuracy and analyze price volatility.

1. ARIMA (AutoRegressive Integrated Moving Average):
   Purpose: Models time series data for trend analysis and short-term forecasting.
   Key Features: Incorporates autoregression, differencing, and moving averages.
   Common Use: Widely applied in economics for price and demand predictions.

2. GARCH (Generalized Autoregressive Conditional Heteroskedasticity):
   Purpose: Captures volatility clustering in financial time series.
   Key Features: Models time-varying variance, ideal for volatile datasets.
   Common Use: Used in finance and energy markets to assess price risk.

3. LSTM (Long Short-Term Memory) Networks:
   Purpose: model for capturing long-term dependencies in sequential data.
   Key Features: Memory cells retain past information, improving pattern recognition.
   Common Use: Applied in complex time series tasks, including stock and commodity forecasting.

Relevance

ARIMA: Suitable for identifying short-term price trends post-events (e.g., sanctions).
GARCH: Essential for modeling oil price volatility during turbulent periods (e.g., conflicts).
-LSTM: Captures long-term event impacts (e.g., multi-year OPEC policy shifts), enhancing predictive depth.
These tools address the challenge of linking diverse events to price movements, improving analytical precision.

Conclusion

The technologies reviewed are invaluable for analyzing the impact of global events on oil prices. Their capabilities in predictive analytics and data visualization significantly enhance our understanding of market dynamics. By utilizing these tools, our project can achieve greater accuracy and relevance, ultimately benefiting stakeholders involved in the oil market. The insights gained from this analysis will contribute to more informed decision-making, aligning with sustainable development goals for a stable energy future.

References

1. Hamilton, J. D. (2009). Causes and Consequences of the Oil Shock of 2007–08.
2. Kilian, L., & Zhou, X. (2020). Oil Prices and Stock Returns in a Time-Varying World.