Energy Economics 122 (2023) 106711

Contents lists available at ScienceDirect

Energy Economics

journal homepage: www.elsevier.com/locate/eneeco

Oil price volatility and new evidence from news and Twitter

Hooman Abdollahi

School of Business and Economics, UiT-The Arctic University of Norway, Tromsø, Norway

ARTICLE INFO

JEL classification:
C53
G12
G17

Keywords:
Oil price volatility
News sentiment
Twitter sentiment
Forecasting
Media and market

ABSTRACT

In this paper, we develop semantic-based sentiment indices through relevant news and Twitter feeds for oil market using a state-of-the-art natural language processing technique. We investigate the predictability of crude oil price volatility using the novel sentiment indices through a hybrid structure consisting of generalized autoregressive conditional heteroskedasticity and bidirectional long short-term memory models. Findings show that media sentiment considerably enhances forecasting quality and the proposed framework outperforms existing benchmark models. More importantly, we compare the predictive power of news stories with Twitter feeds and document the superiority of the news sentiment index over the counterpart. This is an important contribution as this paper is the first study that compares the impact of regular press with that of social media, as an alternative informative medium, on oil market dynamics.

1. Introduction

Playing a significant role in global financial markets, oil is presently known as an alternative investment (Adams et al., 2020; Cui et al., 2021). It is also considered a strategic commodity for economic development across nations (An et al., 2020) such that central banks in many countries regularly update their forecasting of oil prices to remain prepared for possible shocks in the future (Safari and Davallou, 2018). Until recently, the literature has focused on studying oil price volatility using economic fundamentals and statistical data (e. g. Ji et al., 2018; Zhang et al., 2019; Liu and Lee, 2021). Yet, the origin of a part of volatility witnessed in oil prices remains unidentified. The oil market is known as a financial market, and finance theories should be considered in exploring it. According to the finance literature, investors' sensitivity to new information (sentiment) drives a wedge between prices and fundamental values and causes volatility (Ross, 1989). Sentiment can be influenced by various factors such as variations in public opinions, geopolitical conditions, economic news announcements, natural disasters, terroristic attacks, and other exogenous factors (Bomfim, 2003; Brenner et al., 2009; Möbert, 2009; Lucca and Moench, 2015; Birz and Dutta, 2016; Qadan and Nama, 2018). Although these factors are spread via both press and social media, sentiment is not directly measurable. However, recent advancements in natural language processing have provided a tool to quantify the sentiment of textual media. Therefore, investigating the impact of sentiment on forecasting accuracy is a new and growing line of research. Motivated by this issue, we conduct this study to answer the following questions: (i) To what extent does using media sentiment increase the accuracy of oil price volatility forecasting? And (ii) which sentiment proxy is more informative for oil price volatility predictability? Thus, the aim of this study is to shed light on the role of media in oil market predictability.

To answer the questions above, we use weekly data for Brent crude oil prices from August 2014 to December 2020. One of the crucial factors in forecasting oil price volatility is interrelation between financial markets (Abdollahi and Ebrahimi, 2020). Hence, this study also uses historical data for Gold and The Standard and Poor's 500 (S&P 500) index with which the oil market has a high connectedness (Hung, 2022). We also scrape news headlines regarding the oil market as well as tweets mentioning hashtags relevant to the oil market over the same period. The textual materials are then processed using advanced natural language processing to generate two distinct sentiment indices for news and Twitter.

Forecasting oil market volatility is complex because of the various characteristics of oil price time series such as sensitivity to non-fundamental factors (sentiment), nonlinearity, lags, and time-varying volatility (Chen et al., 2016; Abdollahi and Ebrahimi, 2020). To effectively capture these impactful factors, we design a hybrid structure considering the particular ability of each constituent model to capture the aforementioned characteristics. We use the Bidirectional Encoder Representations from Transformers (BERT) to extract sentiment from news headlines and Twitter feeds.[1] Then, we employ a Generalized

E-mail address: hooman.abdollahi@uit.no.

[1] BERT is a state-of-the-art machine learning model that reportedly produces the best results in natural language processing and sentiment classification (González-Carvajal and Garrido-Merchán, 2020).

https://doi.org/10.1016/j.eneco.2023.106711
Received 9 February 2023; Received in revised form 16 April 2023; Accepted 25 April 2023
Available online 5 May 2023
0140-9883/© 2023 The Author. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Nomenclature

ARCH	Autoregressive conditional heteroscedasticity
BERT	Bidirectional encoder representations from transformers
BiLSTM	Bidirectional long-short term memory
DM	Diebold-Mariano
GARCH	Generalized autoregressive conditional heteroskedasticity
LSTM	Long short-term memory
MAE	Mean absolute error
RMSE	Root mean square error
S&P 500	The Standard and Poor's 500

Autoregressive Conditional Heteroskedasticity (GARCH) model to obtain an initial estimation of volatility. Using GARCH efficiently captures time-varying volatility and leverage effects (Crawford and Fratantoni, 2003), which are specific characteristics of the oil price time series. In the final step, we use all the parameters as input features for a Bidirectional Long-Short Term Memory (BiLSTM) model that generates the final forecast. Neural networks generally produce the best results for nonlinear settings like oil price volatility. Moreover, there is a time step dimension in the input array of the BiLSTM model. Therefore, a function of the model is to find patterns in the time step dimension that effectively responds to the need for lags inclusion. We also evaluate forecasting quality using various measures; namely root mean square error (RMSE), mean absolute error (MAE), and Diebold-Mariano (DM) test.

The main empirical finding of this paper indicates that news sentiment generates the paramount results for oil price volatility forecasting relative to the obtained forecast using Twitter sentiment. We also compare the proposed framework with well-established benchmark models fed with both statistical and sentiment data. Adding the Twitter sentiment index to the forecasting setting reduces the error (MAE) by 27%, while replacing that with the news sentiment index reduces the error by 23% further. This error reduction indicates that regular media provides more refined information for the oil market. This is somewhat inconsistent with the idea that the strength of evidence receives more attention than the weight of evidence (Griffin and Tversky, 1992). In this context, we consider the strength as the frequency or reportage of the information and the weight as the source of information. The reportage is higher on social media, but press is more credible. This finding shows that at a market level, and in aggregate, this is the weight that plays the more important role concerning the oil price volatility. Overall, using exclusive semantic sentiment indices enhances forecasting accuracy compared to the results using historical data alone.

From these findings, we conclude that social media such as Twitter is becoming important as an alternative platform for information circulation about the oil market. However, the traditional press and news agencies are more important than social media with regards to spreading more impactful information. This finding suggests that market watchers should place more weight on regular news than Twitter feeds for information updates. This finding also implies that since people can freely express their opinions and anticipations regarding oil market on social media; therefore, Twitter sentiment contains inauthentic interpretations as to news and events, which reduces the degree of sentiment reliability.

This paper adds three important contributions to the existing literature. Firstly, we document that investigating Twitter feeds significantly improves oil price volatility predictability, and this can be traced to incremental information that Twitter offers. To the best of our knowledge, this is the first study that develops a semantic Twitter sentiment index exclusively for the oil market. Previous studies use emotional sentiment proxy (Lehrer et al., 2021) or Twitter general economic uncertainty index (Lang et al., 2021), while those indices do not include non-economic factors or are just a general index. Secondly, we also develop a semantic sentiment index for news headlines relevant to the oil market and compare the predictive power of regular press with social media in forecasting oil price volatility for the first time in the literature. Using these new indices, we document the superiority of news sentiment over Twitter sentiment. Although both indices bring incremental information to the forecasting setting, our results indicate the quality of news information is higher. Thirdly, not only do we produce sentiment indices based on the meaning of the text which reveals the significance of information more refinedly (Shiller, 2020), but also fine-tune the BERT on financial language using Financial PhraseBank proposed by Malo et al. (2014). This practice helps capture the impact of textual information on market prices. We also prove the superiority of the proposed framework over various benchmark models, including individual constituent models and the framework proposed by Verma (2021). A good understanding of various factors impacting the prediction of oil price volatility is vital for policymakers, central banks, and investors across financial markets.

The remainder of the paper is organized as follows. Section 2 presents a review of the pertinent literature and hypothesis development. Section 3 describes the methodologies used for this study. In Section 4, we present the data. Section 5 reports empirical results, while Section 6 concludes the paper.

2. Literature review and hypothesis development

Studying oil price forecasting begins with Amano (1987) by proposing a small-scale econometric model for global oil market. Later, Sharma (1998) compares the quality of GARCH class of models for oil price volatility prediction using data from 1986 to 1997. Tang and Hammoudeh (2002) predict OPEC oil prices using a regression model and find that the prices are also affected by the expectations of market participants. The early studies mainly employ statistical models which are generally able to capture linearity and time-varying volatility in a time series (Bhar and Hamori, 2005; Yu et al., 2008). Another strand of the oil price forecasting literature encompasses machine learning agents mainly due to their ability to capture nonlinear patterns embedded in the time series. Yu et al. (2008) use a neural network based on an empirical mode decomposition to forecast crude oil spot prices. He et al. (2012) propose a wavelet decomposed ensemble model to enhance the prediction accuracy of oil prices with a closer look at the market microstructure, where the proposed model follows the heterogeneous market hypothesis. They document the superiority of their proposed framework over benchmark models in the literature. Mostafa and El-Masry (2016) use an artificial neural network coupled with an evolutionary algorithm to forecast oil prices and recommend using computational approaches like neural networks or fuzzy settings instead of traditional statistical models. On the one hand, statistical models are weak in capturing nonlinearity and non-stationarity, and this weakness decreases their forecasting quality. On the other, Makridakis et al. (2018) document that machine learning agents underperform statistical models in some cases. This fact elucidates the necessity of a third framework in the forecasting literature based on hybrid models consisting of both statistical and machine learning models (Fazelabdolabadi, 2019). The reason is that the structure of hybrid models makes it possible to capture different characteristics in the time series.

Numerous studies on hybrid models document the superiority of such models over single models (see Abdollahi, 2020; Prado et al., 2020). Yet, scholars seek new features or data to increase the prediction quality. This practice gives rise to data construction in forecasting process. Shiller (2020) argues that narratives and news stories can affect prices. This idea suggests that media provide incremental information which are useful for forecasting models. In recent years, scholars have attempted to cover this aspect. Wu et al. (2021) develop a text classification of online news headlines during the COVID-19 pandemic, which produces a binary output for fluctuation of each month, and combine the news proxy with statistical and machine learning models to increase the accuracy of oil price prediction. Li et al. (2021) use news on the oil market and find that shocks in news sentiment result in volatility across the future prices of oil. Liu et al. (2022) combine the Google search volume index as a sentiment proxy with a high-frequency heterogeneous autoregressive model to elevate the forecasting quality in the volatility of oil futures prices. Zhe et al. (2022) scrape the comments posted on Eastmoney forum, an online financial forum, and document the predictive power of sentiment as to the price of China's crude oil. Herrera et al. (2022) employ a lexicon-based method for sentiment extraction from Twitter feeds and combine that with a deep learning model to forecast the returns and volatility of selected renewable energy stocks. Using sentiment as an informational proxy is not limited to volatility forecasting, Zhao et al. (2023) use Google trends (Google search volume index) to forecast crude oil inventory. However, a notable fact about previous studies is the use of less-sophisticated methods based on counting words or pre-weighted lexicons for textual analysis. Those methods fail to provide a sentiment analysis based on the meaning of text. However, using semantic sentiment is recommended for investigating narratives across financial markets as it can show the psychological significance of sentiment (Shiller, 2017). This paper also attempts to bridge this gap in the literature by producing semantic sentiment for news and Twitter feeds.

de Medeiros et al. (2022) argue that data construction and curation can play a crucial role in improving prediction accuracy for the oil market dynamics. On the other hand, the literature suggests various factors influencing oil price volatility such as geopolitical escalations (Tahmassebi, 1986), macroeconomic news (Meng and Liu, 2019), shifts in supply and demand (Hosseini et al., 2021), etc. de Oliveira et al. (2018) argue that press and social media are a source for sentiment circulation. Lehkonen et al. (2022) note that media can be a genuine risk factor in financial markets. Abdollahi et al. (2023) document a long-lasting connectedness between media sentiment and financial market volatility. Therefore, based on this literature, we expect that constructing sentiment data from news and social media improves the forecasting accuracy of oil prices. We formalize this as the first hypothesis.

H1. Media sentiment increases the accuracy of oil price volatility forecasting.

Junttila et al. (2005) use analysts' perceptions published in a well-known magazine and show that such contents can be a value driver in the market. Birz and Lott Jr (2011) selects newspaper articles that provide an interpretation of statistical releases as a measure of news and notes that public interpretations of news affect stock returns. Kim et al. (2019) note that >60% of adults get news from social media by 2016 and the proportion is still increasing; even though the problem of fake news is a case. Shiller (2020) proposes that viral news stories, whether fact-based or not, can potentially influence prices across financial markets. Based on these studies, we expect that materials posted on Twitter, as an amalgamation of news and the concomitant public interpretations, can be a better proxy than regular news media for the oil price volatility forecasting. We formalize this in the second hypothesis.

H2. Twitter sentiment outperforms news sentiment in oil price volatility forecasting.

3. Methodology

3.1. BERT

BERT is a pre-trained natural language processing technique developed by Google (Devlin et al., 2018). The model is trained on English Wikipedia and BookCorpus (Zhu et al., 2015), containing >3500 million words. The principal feature of the BERT is the ability to understand the meaning of sophisticated language. We employ the BERT-base version to analyze text data by predicting the likelihood of positivity, neutrality, or negativity for a given input. We obtain the sentiment score for a given text as the difference between positive and negative likelihoods or logits (Lin and Luo, 2020):

Table 1
The proportion of different classes in Financial PhraseBank.

Class	Proportion
Positive	28.2%
Neutral	59.4%
Negative	12.4%

$$Sentiment\ Score = logit_{positive} - logit_{negative}. \tag{1}$$

The sentiment score for a given text is between 1 (extreme positive) and − 1 (extreme negative). We fine-tune the BERT for finance domain to further improve the model's ability to generate quality results. For this purpose, we use the Financial PhraseBank proposed by Malo et al. (2014). The Financial PhraseBank contains 4845 sentences from financial news classified by finance professionals. The classes show how the professionals perceive the impact of information embedded in a sentence on prices in the market (positive impact, neutral impact, or negative impact). Table 1 presents the proportion of different classes in the dataset. The PhraseBank also categorizes the sentences based on agreement levels. We use the entire PhraseBank including all agreement levels for training the BERT.[2] Fine-tuning the model is advantageous as it elevates the quality of sentiment extraction.

3.2. GARCH (1,1)

Engle (1982) proposes autoregressive conditional heteroscedasticity (ARCH) to model the conditional variance using error terms and to capture volatility clustering of a series as follows.

$$\epsilon_t = \sigma_t X_t, \sim N\left(0, \sigma_t^2\right) \tag{2}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \ldots + \alpha_q \epsilon_{t-q}^2 \tag{3}$$

Where ϵ_t denotes the error term which follows a normal distribution, and σ_t^2 is the conditional variance dependent on past squared residuals. Bollerslev (1986) extends the ARCH(q) model to build GARCH (p,q) model such that it contains both the past squared residuals and the values of conditional variances. The model is formulated as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \gamma_{k-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{k-j}^2. \tag{4}$$

GARCH estimates the conditional variance (σ_t^2) of series γ_t. The model parameters are conditioned on $\omega > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$, and $\alpha_i + \beta_i < 1$, which are the essential conditions for stationarity in the series. We employ a GARCH (1,1) model to estimate the conditional variance of the oil prices and use it as an input for the next model in the proposed structure.

3.3. BiLSTM

Long short-term memory (LSTM) is an artificial neural network with the capability to learn sequential patterns. This model was developed to tackle the problem of long-range dependency in backpropagation (Hochreiter and Schmidhuber, 1997). The standard LSTM follows Eqs. (5) to (10).

[2] There are different agreement levels regarding the impact of sentences among professionals.

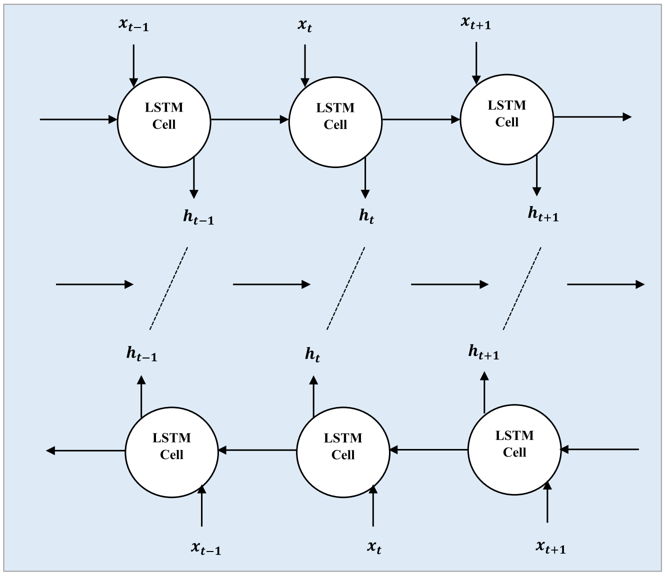


Fig. 1. The structure of BiLSTM.
Graphical representation of the BiLSTM model consisting of two LSTMs in which one takes inputs in a forward direction and the other in a backward direction leading to an increased amount of information absorbed by the model.

$$g_t = \sigma\left(b_g + U_g x_t + V_g h_{t-1}\right), \tag{5}$$

$$i_t = \sigma(b_i + U_i x_t + V_i h_{t-1}), \tag{6}$$

$$o_t = \sigma(b_o + U_o x_t + V_o h_{t-1}), \tag{7}$$

$$f_t = tanh\left(b_f + U_f x_t + V_f h_{t-1}\right), \tag{8}$$

$$c_t = g_t \odot c_{t-1} + i_t \odot f_t, \tag{9}$$

$$h_t = o_t \odot tanh(c_t). \tag{10}$$

The LSTM structure is composed of blocks consisting of one memory cell (c_t) and three gates, namely the input gate (i_t), forget gate (g_t), and output gate (o_t). In Eqs. (5)–(10), h_t and x_t denote the hidden state of the memory cell and input at time t, respectively. σ represents the activation function (sigmoid) and b is the bias term. U_i, U_g, and U_o denote the weight matrices of input, forget, and output gates, respectively. V_i, V_g, and V_o denote the recurrent weight matrices of input, forget, and output gates, respectively. f_t is an input modulate gate, which is a value showing the amount of new information received in the memory cell. Symbol $\odot$ represents element-wise multiplication. In the LSTM structure, forget gate g_t, as shown in Eq. (5), generates a value between 0 and 1, where 0 means that no input information passes through the gate and a value of 1 implies all input information is passed. The input gate, as shown in Eq. (6), ascertains the amount of information stored in the memory cell (c_t). Eq. (8) determines the new information at time t. Eq. (9) calculates past information and new information of the memory cell, which are controlled by input and forget gates, at time t. Finally, the hidden state information h_t is determined using the output gate (o_t).

BiLSTM algorithm is an update on LSTM, incorporating the bidirectional recurrent network structure into LSTM cells. This structure adds the ability to take advantage of feedback layers to the BiLSTM model. The model structure includes hidden layers that run in opposite directions. Therefore, the model acquires both forward and backward sequential information through the past and future of a given dataset. Fig. 1 illustrates the structure of the BiLSTM model.

3.4. Performance evaluation

The forecasts of the model must be further assessed to verify the findings. Also, a formal test is needed to examine the hypotheses. Hence, we employ the following measures and test:

I. RMSE: Shows the standard deviations of differences between actual and forecasted values. It is used as an accuracy indicator for comparing prediction errors of different models.

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Y_t - \widehat{Y}_t\right)^2} \tag{11}$$

II. MAE: Shows the average of absolute errors when comparing the actual and forecasted values.

$$MAE = \frac{1}{T}\sum_{t=1}^{T}\left|Y_t - \widehat{Y}_t\right| \tag{12}$$

In Eq. (11) and Eq. (12), Y_t and $\widehat{Y}_t$ are the actual and forecasted values at time t, respectively. T represents the sample size. The smaller values for RMSE and MAE indicate higher forecasting accuracy.

III. DM test: Determines if the difference between two forecasting models is statistically significant (Diebold and Inoue, 2001). The null hypothesis implies no significant difference between errors. The DM statistic is defined as follows:

$$DM = \frac{\overline{d}}{\left(\frac{2\pi \widehat{f}_d}{T}\right)^{½}} \sim N(0,1), \tag{13}$$

where T is the size of sample. $\overline{d}$ denotes the mean of loss differential between two forecasts, and $\widehat{f}_d$ represents spectral density of the loss differential.

3.5. Benchmark forecasting models

Benchmark models are needed to check the effectiveness of the proposed forecasting model. For this aim, we use the constituent forecasting models, namely the GARCH and the BiLSTM, along with their hybridization in which the outcome of the GARCH model is used as one of the features to feed the BiLSTM model. The hybridization of GARCH- and LSTM-based models in forecasting oil price volatility has been reliably documented in the literature (Verma, 2021). Therefore, these models can be employed as reliable benchmarks. Another interesting hybridization is to use the outcome of the BiLSTM together with sentiment series as regressors in the GARCH model. This practice makes it possible to compare the suitability of statistical and machine learning models for oil price volatility predictability in the presence of sentiment. Therefore, we use the GARCH model, BiLSTM model, GARCH-BiLSTM model, and BERT-BiLSTM-GARCH model as benchmarks for our proposed model.

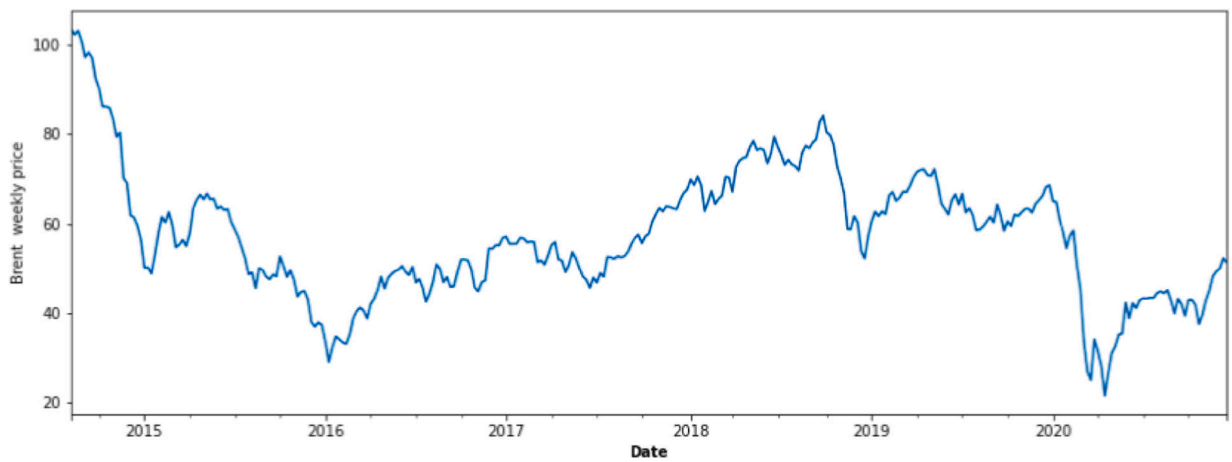
4. Data

4.1. Historical data

Weekly data of Brent oil were sourced from Refinitiv Eikon from 4 August 2014, to 20 December 2020. Brent oil historically plays a crucial role in the global oil pricing mechanism as it is considered a benchmark for around 55% of the international oil trade (Cheng et al., 2019). Therefore, Brent is an appropriate proxy for the global oil market. We also use weekly data for Gold and S&P 500 over the same period as financial indicators which have a high volatility connectedness with the oil market.

Fig. 2 shows oil price time series over the period under consideration. As can be observed, high production levels together with low demand pushed the oil prices to a nadir in early 2016, when it hit a 13-year low of

A) Brent closing prices



A) Brent price volatility

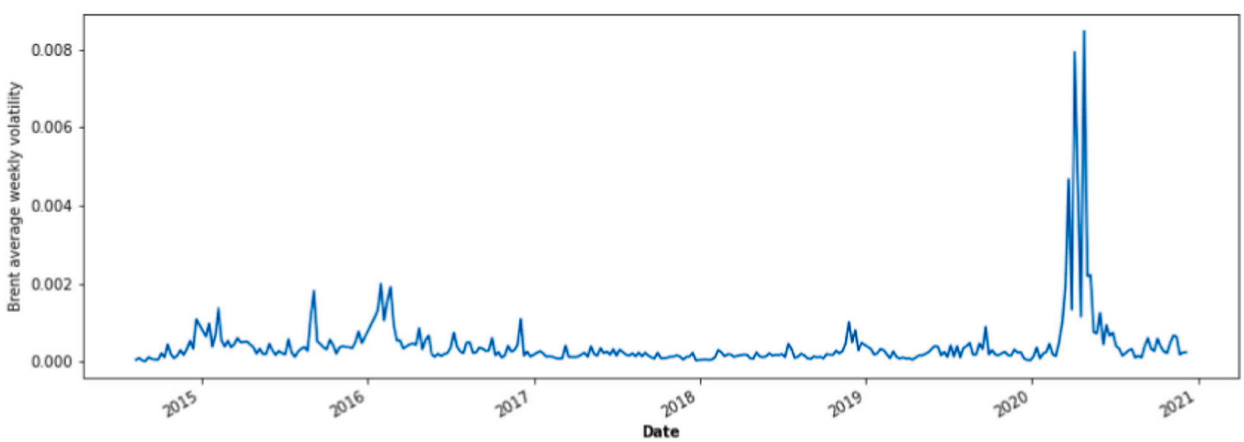


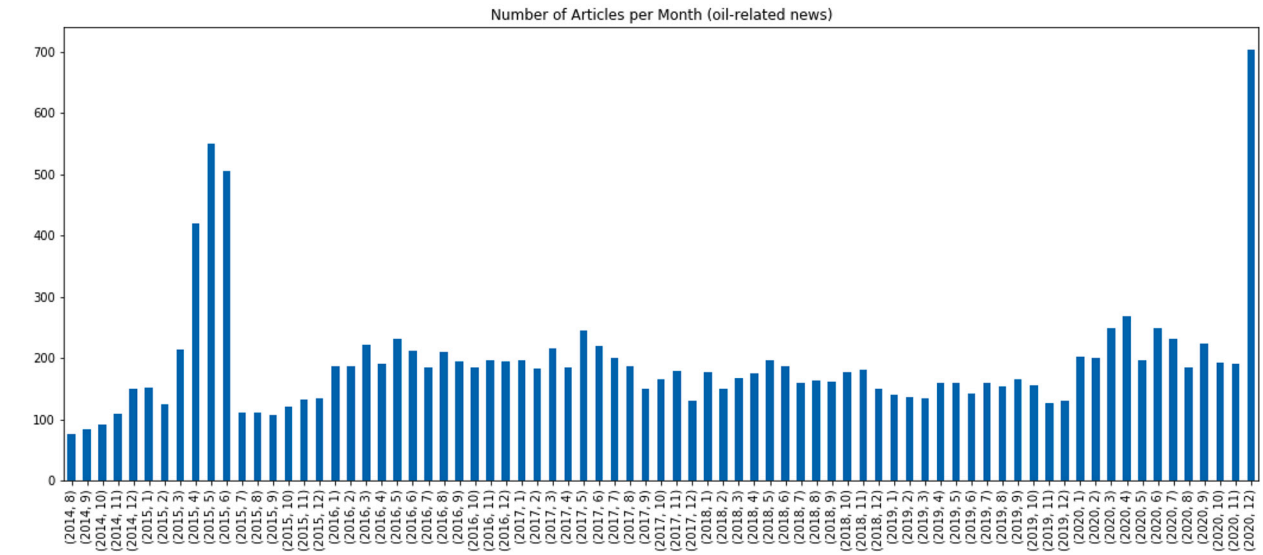
Fig. 2. Brent weekly price and volatility from August 2014 to December 2020.
Weekly closing prices and range-based volatility of Brent crude oil from August 2014 to December 2020. Panel A shows the historical trajectories in closing prices and Panel B shows the volatility for Brent crude oil, respectively.

Table 2
Descriptive statistics.

	Mean	Minimum	Maximum	SD	Skewness	ADF
Panel A. Financial series						
Brent	0.000422	0.000024	0.008468	0.000753	7.8625	−4.1483 (0.0008)
Gold	0.000054	0	0.0007	0.000077	4.5268	−7.3323 (0)
S&P 500	0.000066	0.000003	0.001784	0.00015	7.8192	−7.7669 (0)
Panel B. Sentiment series						
News sentiment index	0.179640	−0.997356	0.984978	0.463667	−0.3259	−10.0038 (0)
Twitter sentiment index	0.325211	−0.107918	0.749671	0.14302	0.2199	−14.5428 (0)

This table presents descriptive statistics for Brent crude oil prices, gold prices, and S&P 500 index together with news sentiment index and Twitter sentiment index for oil market over the period from August 2014 to December 2020 in panel A and B, respectively. Values in the first column shows weekly mean for each series followed by the minimum and maximum values in the next two columns. SD stands for standard deviations. Number in parenthesis in the last columns show the *p*-value at 5% significance level.

A) News frequency



B) Twitter feeds frequency

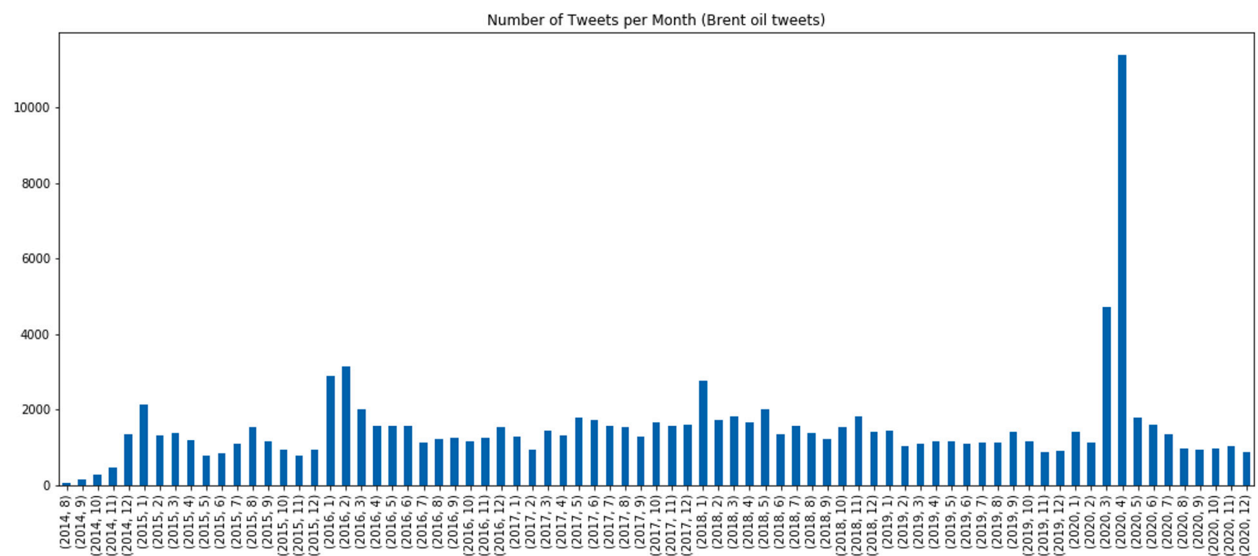


Fig. 3. Number of news and tweets relevant to the oil market from August 2014 to December 2020.
Panel A shows the frequency of online news about Brent crude oil across all news services on Investing.com from August 2014 to December 2020. Panel B sows the frequency of Twitter feeds using one of the hashtags: #oilmarket, #brentoil, #oilprice, #wtioilprice from August 2014 to December 2020.

$27.10. Later, the trend reversed with a simultaneous increase in global demand and a drop in output levels. Also, a series of geopolitical issues took place over 2017–18, like the imposition of the U.S. sanctions on Iran, or Russia and Saudi Arabia's decisions to curb production levels. All these events sent oil prices to a four-year high of more than $80 in late 2018. Throughout 2019, rise in the U.S. oil production put downward pressure on prices. Moreover, geopolitical issues such as the attack on Saudi Arabia's petroleum installations and production cut announcements by OPEC contributed to a lower average of oil prices. Finally, the fall in 2020 was related to a remarkable drop in demand coupled with weakened economic prospects during the Covid-19 pandemic even though it reversed later. The oil price is both volatile and trending at some points over this period in which media shocks play a role; therefore, using this period is beneficial to investigating the reliability of the proposed model.

Panel A in Table 2 presents descriptive statistics for the financial series. For all markets, the standard deviation values are greater than the average. This shows that Brent (0.000753) has the highest volatility level, while gold (0.000077) is of the lowest volatility. The skewness values indicate that all financial series are highly skewed. We check the stationarity of the series using the augmented Dickey-Fuller (ADF) test, where the null hypothesis implies the existence of a unit root in series. The last column of Table 2 presents the results of the ADF test, confirming the stationarity of all series at the 5% significance level.

4.2. Textual data

Using the archival news of *investing.com*[3] Website, a total of 14,833

[3] Investing.com is a financial markets platform providing exclusive and eclectic news and other facilities regarding multifarious exchanges and commodities around the world.

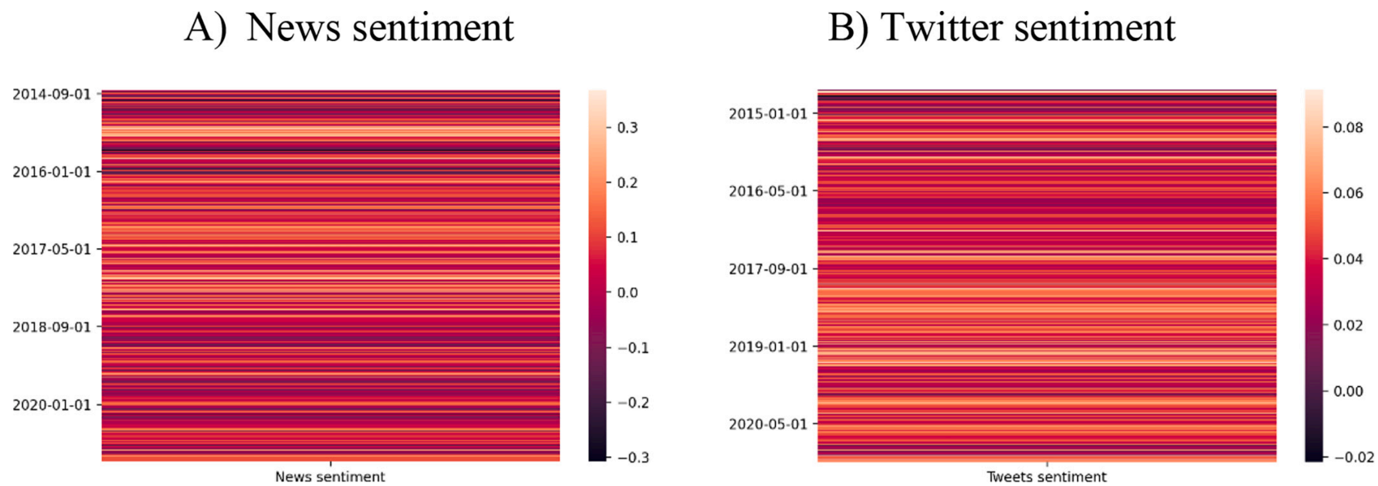


Fig. 4. Heatmaps of sentiment indices.
Heatmaps of the semantic indices through news and Twitter feeds exclusively for the oil market from August 2014 to December 2020. The indices rely on real-time events causing the sentiment to be varying. Time-varying characteristic is more pronounced in the case of news sentiment, while Twitter sentiment is a bit more positive.

news headlines about the oil market was collected from 4 August 2014, to 22 December 2020. We use news headlines as headlines efficiently convey the gist of the whole article (Li et al., 2019). Panel A in Fig. 3 shows the number of news articles per month. There is a sharp increase in the quantity of news in December 2020 which can relate to the OPEC meeting, OPEC and non-OPEC ministerial meeting, and announcement of Covid-19 vaccine production.

We also collected 116,463 tweets for different hashtags related to the oil market over the same period.[4] Panel B in Fig. 3 illustrates the number of tweets per month. As can be seen, there is a dramatic increase in March and April 2020 when the oil price unprecedently plunged and, simultaneously, there was a serious conflict between OPEC+ members where the U.S. president mediated between them and repeatedly updated the public via his Twitter account.

5. Empirical analysis

To check Hypothesis 1, which predicts media sentiment elevates the accuracy of oil price volatility forecasting, we commence the analysis by sentiment extraction using the BERT model trained on the Financial PhraseBank. We use 80% of the dataset for training and validation and 20% for test. This practice boosts the model accuracy to 0.83%. The weighted average of model's precision and recall equals 0.82, and model's loss is 0.38. We then process the textual contents through the fine-tuned model. Fig. 4 represents the heatmaps for sentiment indices. As can be seen, the sentiment indices are very dynamic and varying as they mostly depend on exogenous news like real-time events. Twitter sentiment also includes public interpretations, which seemingly add a tinge of pessimism to the sentiment as most of the index values fall within the positive area. Also, From Panel B in Table 2, we see that Twitter sentiment index is positively skewed, while the news sentiment index is negatively skewed. Standard deviations also indicate that the news sentiment index (0.463667) is more volatile than the Twitter sentiment index (0.14302).

We use the sentiment indices to feed the BiLSTM model in the last step of the proposed model. Before that, we estimate the GARCH model, whose parameters are presented in Table 3, to obtain an initial forecast of oil prices volatility. Not only do we assess this forecast against the actual volatility, but we also use it as an input feature for the final forecast.

Table 3
The estimated parameters of GARCH model.

	Estimate	SE	t-statistic	P-value
ω	13×10^{-9}	0	304.875	(0)
α	0.2	0.0922	2.169	(0.03006)
β	0.78	0.0331	23.568	(0)

Table presents the estimation of GARCH (1,1) parameters where ω is the constant coefficient of the variance equation, α represents the value of the autoregressive coefficient, and β denotes the value of variance coefficient. SE: standard error.

Finally, we feed the BiLSTM model with the sentiment indices, GARCH estimation, and financial data as input features. As for the sentiment data, we use the Twitter index and news index each at a time. The structure of the BiLSTM model consists of two layers, having 12 and 1 neurons, and the dense layer. We also use Adam optimizer (Kingma and Ba, 2014) to optimize the training process. The learning rate is set at 0.001. All BiLSTM settings in this study are trained for 60 epochs. The prediction horizon is 22 weeks, and we use 12-week window of explanatory variables as input.

Fig. 5 illustrates the final prediction for Brent oil volatility by the proposed hybrid model using the news sentiment index (News-GARCH-BiLSTM) in Panel A and using the Twitter sentiment index (Twitter-GARCH-BiLSTM) in Panel B. Fig. 5 also shows forecasting results obtained by benchmark models including News-BiLSTM-GARCH (Panel C), Twitter-BiLSTM-GARCH (Panel D), GARCH-BiLSTM using financial data (Panel E), BiLSTM using financial data (Panel F), and GARCH using oil historical data (Panel G).

Table 4 presents forecast errors for all the models employed in this study. We can see significant drops in forecast errors when the hybrid model is fed with sentiment indices. The interpretation is that refined information concerning the oil market in media increases the forecasting quality if properly converted into numerical values. The news sentiment index produces the paramount forecast as it reduces the RMSE measure derived by the Twitter sentiment index by 21%. The interpretation is that regular news is of more genuine information for the oil market.

Having the sentiment indices involved in the forecasting models, we can also see that using BiLSTM as the ultimate forecasting model improves the prediction errors by 15% in comparison to having the GARCH as the final predictive model. Another interesting fact goes back to the propriety of predictive models and data in prediction improvement. For the Twitter sentiment index, we can see that using the GARCH as the

[4] The hashtags are: #brentoil, #oilprice, #wtioilprice, #oilmarket.

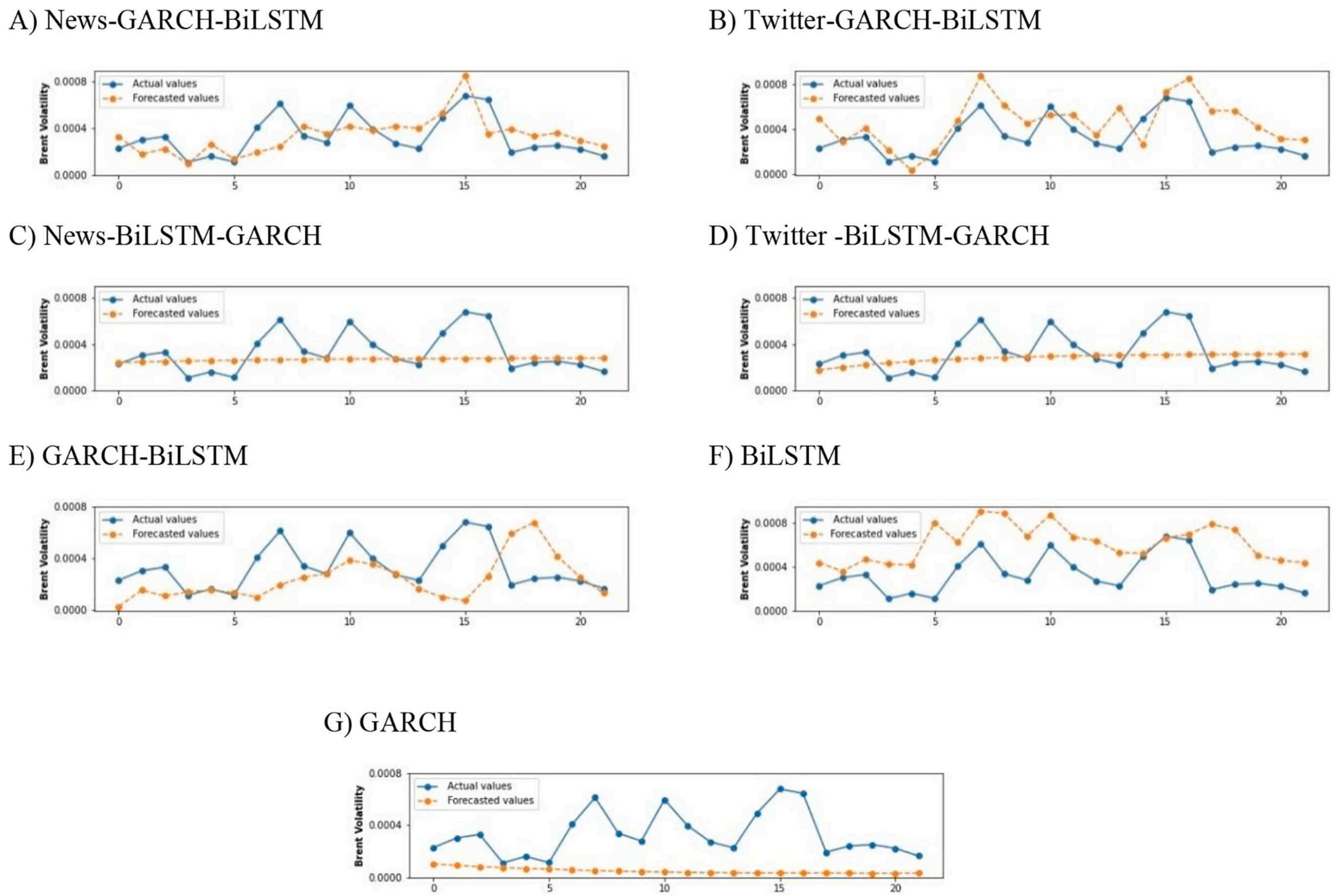


Fig. 5. Forecasting results of Brent oil volatility for 22 weeks ahead.
Forecasting 22-week ahead for Brent crude oil volatility. Panel A shows the predictions for News-GARCH-BiLSTM, Panel B illustrates the forecasting by Twitter-GARCH-BiLSTM, Panel C represents the predictions using News-BiLSTM-GARCH, panel D shows the forecasting using Twitter-BiLSTM-GARCH, panel E illustrate the prediction using GARCH-BiLSTM, panel F shows the forecasting using BiLSTM, and panel G illustrates the prediction using GARCH model.

Table 4
Results of oil price volatility forecasts with error functions.

Model	MAE ($\times 10^{-4}$)	RMSE ($\times 10^{-4}$)
GARCH	0.2819	0.3329
BiLSTM	0.2871	0.3365
GARCH-BiLSTM	0.2286	0.2948
Twitter-BiLSTM-GARCH	0.1393	0.1818
News-BiLSTM-GARCH	0.1331	0.1818
Twitter-GARCH-BiLSTM	0.1661	0.1951
News-GARCH-BiLSTM	0.1269	0.1531

Table shows the values of error measures for all the forecasting models using root mean square error (RMSE) and mean absolute error (MAE).

final forecasting model results in smaller prediction errors than the BiLSTM. The opposite is true with regard to the news sentiment index.

We can also see that the GARCH provides a slightly better forecast than that of the BiLSTM when we scrutinize the performance of single models. The interpretation is that the statistical models provide superior results than machine learning agents when we focus on using single models for forecasting oil price volatility under this setting. We also observe that the hybrid GARCH-BiLSTM model (using financial data) produces smaller forecast errors than single models. The interpretation is that the hybridization of statistical and machine learning models improves the forecasting accuracy, as the errors subsided. The reason is that this process enables the model to capture different characteristics existing in the time series more effectively.

Table 5
The results of DM test.

Pair of models	DM statistic	P-value
GARCH vs. BiLSTM	8.679	(0.0000)
GARCH-BiLSTM vs. GARCH	3.1692	(0.0046)
GARCH-BiLSTM vs. BiLSTM	−7.3704	(0.0000)
Twitter-BiLSTM-GARCH vs. GARCH	6.68	(0.0000)
News-BiLSTM-GARCH vs. GARCH	6.6767	(0.0000)
Twitter-BiLSTM-GARCH vs. BiLSTM	7.2509	(0.0000)
News-BiLSTM-GARCH vs. BiLSTM	7.2925	(0.0000)
Twitter-BiLSTM-GARCH vs. GARCH-BiLSTM	6.7	(0.0000)
News-BiLSTM-GARCH vs. GARCH-BiLSTM	6.6768	(0.0000)
Twitter-GARCH-BiLSTM vs. GARCH	5.5788	(0.0001)
News-GARCH-BiLSTM vs. GARCH	4.2697	(0.0003)
Twitter-GARCH-BiLSTM vs. BiLSTM	−3.8205	(0.0009)
News-GARCH-BiLSTM vs. BiLSTM	−5.1502	(0.0000)
Twitter-GARCH-BiLSTM vs. GARCH-BiLSTM	3.5132	(0.002)
News-GARCH-BiLSTM vs. GARCH-BiLSTM	10.1343	(0.0000)
Twitter-GARCH-BiLSTM vs. Twitter-BiLSTM-GARCH	3.9351	(0.0007)
News-GARCH-BiLSTM vs. Twitter-BiLSTM-GARCH	10.1344	(0.0000)
Twitter-GARCH-BiLSTM vs. News -BiLSTM-GARCH	2.132	(0.04497)
News-GARCH-BiLSTM vs. News -BiLSTM-GARCH	4.0629	(0.0005)
News-BiLSTM-GARCH vs. Twitter-BiLSTM-GARCH	2.6968	(0.0135)
News-GARCH-BiLSTM vs. Twitter-GARCH-BiLSTM	10.1337	(0.00)

Table presents the result of pairwise comparisons between forecasting errors using Diebold-Mariano test. The null hypothesis of the test implies no significant difference between the predictions. Significance level is 5%.

Table 6
Results of oil price volatility forecasts for different prediction horizons.

Model	Prediction horizon = 16		Prediction horizon = 28	
	MAE ($\times 10^{-4}$)	RMSE ($\times 10^{-4}$)	MAE ($\times 10^{-4}$)	RMSE ($\times 10^{-4}$)
GARCH	0.3396	0.3793	0.3051	0.3627
BiLSTM	0.296	0.3446	0.2567	0.3042
GARCH-BiLSTM	0.2712	0.3558	0.784	1.1202
Twitter-BiLSTM-GARCH	0.1715	0.2063	0.1886	0.2394
News-BiLSTM-GARCH	0.1554	0.2052	0.1875	0.2349
Twitter-GARCH-BiLSTM	0.1698	0.2154	0.1899	0.2353
News-GARCH-BiLSTM	0.1326	0.1826	0.1868	0.2058

Table presents the robustness check for the predictions of Brent crude oil price volatility using different forecast horizons.

We perform the DM test as a formal test to check the hypotheses. This test provides pairwise comparisons between the forecasts generated by different models. The DM tests the assumption of no significant difference between two forecasts. Table 5 presents the results of the DM test. From rows 4–15 in Table 5, we see that there are statistically significant differences between the forecasts of hybrid models fed with sentiment indices and those of models without sentiment inclusion. The interpretation is that media sentiment provides novel and recent information regarding the oil market that boosts the quality of forecasting. A notable feature of sentiment is that it frequently varies with respect to the dynamic atmosphere of the real world. Given the rapidity of information incorporation within financial markets, take the effects of the Covid-19 pandemic, the US-China Tariff clash, and Brexit for instance, building an index that covers the behavioral aspects relevant to the oil market adds a contribution to the predictive power of models. Therefore, the results are consistent with hypothesis 1 that media sentiment elevates the accuracy of oil price volatility predictability. We conclude that media sentiment enhances the forecasting quality of oil price volatility because it includes new information which is absent in historical data.

Hypothesis 2 states that Twitter sentiment outperforms news sentiment in oil price volatility forecasting. From Table 4, we find that the opposite is true as forecasts using the news sentiment index generate smaller prediction errors than those of the Twitter sentiment index. To check if this finding is statistically significant, we perform the DM test for these forecasts. The last six rows in Table 5 show statistically significant differences between forecasts using the news sentiment index and the Twitter sentiment index under both forecasting models. Therefore, hypothesis 2 is rejected as news stories provide more quality prediction for oil price volatility than Twitter feeds. The interpretation is that although quantified information about the exogenous factors of the oil market increases the forecasting quality, the results are still sensitive to the quality and reliability of the information. In this analysis, news sentiment is built upon globally prestigious press, where those news agencies apply a high level of ethical and professional observations. Therefore, we are ensured that their contents usually are fact-checked and reliable. On the contrary, Twitter feeds are fraught with heterogeneous users causing the index contains dual information as one part of the index includes reliable information, while the other part is based on non-fact contents in the form of public interpretations or even false news. We conclude that news sentiment outperforms Twitter sentiment in forecasting oil price volatility, as the results indicate that oil market fluctuations are more sensitive to genuine news.

We further examine the robustness of our results using two different forecasting horizons of 16 and 28 weeks ahead. Table 6 presents the corresponding forecasting errors. From the second column in Table 6, we see that sentiment-based models produce smaller errors than those of other models over a shorter forecasting horizon. We also observe that the news sentiment index generates a better forecast than the Twitter sentiment index. From the third column in Table 6, we see that the findings remain unchanged over a longer forecasting horizon as well. These facts signify that changes in the forecasting horizon lead to no significant variations in the results. Therefore, we conclude that the obtained results are robust and contributory to the empirical literature.

6. Conclusion

In this paper, we investigate the contribution of news stories and Twitter feeds to the quality of predictive models for oil price volatility. Oil market fluctuations have always been influenced by exogenous factors such as geopolitical conflicts, natural disasters, viral narratives, etc. These factors are spread in media. Research is at an early stage as to incorporating media sentiment into forecasting settings. This is the first study that builds semantic sentiment indices for both news and Twitter feeds and compares their effectiveness.

We collect a total of 14,833 news headlines and 116,463 tweets regarding various aspects of the oil market from August 2014 to December 2020. We then process the textual contents using a financially fine-tuned BERT model to extract weekly sentiment indices for both news and Twitter feeds. We add the sentiment indices into a GARCH-BiLSTM model to check the impact of media sentiment. Findings indicate that media sentiment enhances forecasting accuracy relative to using only historical data. More importantly, we also find that news sentiment generates superior results for oil price volatility forecasting compared with Twitter-based forecasting. This improvement results from the inclusion of various features to cover certain aspects of oil price characteristics, specifically the sentiment.

Empirical implications suggest that policymakers, central banks, hedgers, and others who deal with the oil market should consider the crucial role of media sentiment in their anticipations. More weight should be put on regular press than social media as the sentiment derived from news agencies further reduces forecasting errors.

Results also present theoretical insights. We document that news and Twitter sentiments contribute to forecasting quality. One way to extend this line of research is to add a weighting function, which assigns optimized weight to each forecasting feature, to the current setting and use both indices simultaneously. As the paramount result was obtained by news sentiment, future models can also divide the news index into different categories, like economic news, political news, war news, etc., and compare the predictive power of each category.

Credit author statement

This is a single-authored paper (the author confirms sole responsibility for idea and design, data collection, analysis and interpretation of results, and manuscript preparation).

Acknowledgements

Special thanks are due to Espen Sirnes and Sturla L. Fjesme for their advice during the conduct of this study. Thanks also to Babak Ebrahimi for useful comments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eneco.2023.106711.

References

Abdollahi, H., 2020. A novel hybrid model for forecasting crude oil price based on time series decomposition. Appl. Energy 267, 115035.
Abdollahi, H., Ebrahimi, S.B., 2020. A new hybrid model for forecasting Brent crude oil price. Energy 200, 117520.

Abdollahi, H., Fjesme, S.L., Sirnes, E., 2023. Measuring Connectedness between Media Sentiment and Market Volatility (Working paper).
Adams, Z., Collot, S., Kartsakli, M., 2020. Have commodities become a financial asset? Evidence from ten years of Financialization. Energy Econ. 89, 104769.
Amano, A., 1987. A small forecasting model of the world oil market. J. Policy Model 9 (4), 615–635.
An, S., Gao, X., An, H., An, F., Sun, Q., Liu, S., 2020. Windowed volatility spillover effects among crude oil prices. Energy 200, 117521.
Bhar, R., Hamori, S., 2005. Empirical Techniques in Finance. Springer Science & Business Media.
Birz, G., Dutta, S., 2016. Us macroeconomic news and international stock prices: evidence from newspaper coverage. Account. Fin. Res. 5 (1), 247.
Birz, G., Lott Jr., J.R., 2011. The effect of macroeconomic news on stock returns: new evidence from newspaper coverage. J. Bank. Financ. 35 (11), 2791–2800.
Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. J. Econ. 31 (3), 307–327.
Bomfim, A.N., 2003. Pre-announcement effects, news effects, and volatility: monetary policy and the stock market. J. Bank. Financ. 27 (1), 133–151.
Brenner, M., Pasquariello, P., Subrahmanyam, M., 2009. On the volatility and comovement of US financial markets around macroeconomic news announcements. J. Financ. Quant. Anal. 44 (6), 1265–1289.
Chen, H., Liao, H., Tang, B.J., Wei, Y.M., 2016. Impacts of OPEC's political risk on the international crude oil prices: An empirical analysis based on the SVAR models. Energy Econ. 57, 42–49.
Cheng, F., Li, T., Wei, Y.M., Fan, T., 2019. The VEC-NAR model for short-term forecasting of oil prices. Energy Econ. 78, 656–667.
Crawford, G.W., Fratantoni, M.C., 2003. Assessing the forecasting performance of regime-switching, ARIMA and GARCH models of house prices. Real Estate Econ. 31 (2), 223–243.
Cui, J., Goh, M., Zou, H., 2021. Coherence, extreme risk spillovers, and dynamic linkages between oil and China's commodity futures markets. Energy 225, 120190.
de Medeiros, R.K., da Nóbrega Besarria, C., de Jesus, D.P., de Albuquerquemello, V.P., 2022. Forecasting oil prices: new approaches. Energy 238, 121968.
de Oliveira, F.A., Maia, S.F., de Jesus, Besarria, C.D.N., 2018. Which information matters to market risk spreading in Brazil? Volatility transmission modelling using MGARCH-BEKK, DCC, t-Copulas. The North American Journal of Economics and Finance 45, 83–100.
Devlin, J., Chang, M.W., Lee, K., Toutanova, K., 2018. Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint. *arXiv:1810.04805*.
Diebold, F.X., Inoue, A., 2001. Long memory and regime switching. J. Econ. 105 (1), 131–159.
Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econ. J. Econ. soc. 987–1007.
Fazelabdolabadi, B., 2019. A hybrid Bayesian-network proposition for forecasting the crude oil price. Fin. Innovat. 5 (1), 1–21.
González-Carvajal, S., Garrido-Merchán, E.C., 2020. Comparing BERT against Traditional Machine Learning Text Classification. *arXiv preprint*. *arXiv:2005.13012*.
Griffin, D., Tversky, A., 1992. The weighing of evidence and the determinants of confidence. Cogn. Psychol. 24 (3), 411–435.
He, K., Yu, L., Lai, K.K., 2012. Crude oil price analysis and forecasting using wavelet decomposed ensemble model. Energy 46 (1), 564–574.
Herrera, G.P., Constantino, M., Su, J.J., Naranpanawa, A., 2022. Renewable energy stocks forecast using twitter investor sentiment and deep learning. Energy Econ. 114, 106285.
Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780.
Hosseini, S.H., Shakouri, H., Kazemi, A., 2021. Oil price future regarding unconventional oil production and its near-term deployment: a system dynamics approach. Energy 222, 119878.
Hung, N.T., 2022. Asymmetric connectedness among S&P 500, crude oil, gold and bitcoin. Manag. Financ. 48 (4), 587–610.
Ji, Q., Bouri, E., Roubaud, D., 2018. Dynamic network of implied volatility transmission among US equities, strategic commodities, and BRICS equities. Int. Rev. Financ. Anal. 57, 1–12.
Junttila, J., Kallunki, J.P., Kärja, A., Martikainen, M., 2005. Stock market response to analysts' perceptions and earnings in a technology-intensive environment. Int. Rev. Financ. Anal. 14 (1), 77–92.
Kim, A., Moravec, P.L., Dennis, A.R., 2019. Combating fake news on social media with source ratings: the effects of user and expert reputation ratings. J. Manag. Inf. Syst. 36 (3), 931–968.
Kingma, D.P., Ba, J., 2014. Adam: A Method for Stochastic Optimization. *arXiv preprint*. *arXiv:1412.6980*.
Lang, Q., Lu, X., Ma, F., Huang, D., 2021. Oil futures volatility predictability: evidence based on twitter-based uncertainty. Financ. Res. Lett. 102536.
Lehkonen, H., Heimonen, K., Pukthuanthong, K., 2022. Is media tone just a tone? Time-series and cross-sectional evidence from the currency market. In: Time-Series and Cross-Sectional Evidence from the Currency Market (November 22, 2022).
Lehrer, S., Xie, T., Zhang, X., 2021. Social media sentiment, model uncertainty, and volatility forecasting. Econ. Model. 102, 105556.
Li, Y., Jiang, S., Li, X., Wang, S., 2021. The role of news sentiment in oil futures returns and volatility forecasting: data-decomposition based deep learning approach. Energy Econ. 95, 105140.
Li, X., Shang, W., Wang, S., 2019. Text-based crude oil price forecasting: A deep learning approach. Int. J. Forecast. 35 (4), 1548–1560.
Lin, P., Luo, X., 2020, October. A survey of sentiment analysis based on machine learning. In: CCF International Conference on Natural Language Processing and Chinese Computing. Springer, Cham, pp. 372–387.
Liu, M., Lee, C.C., 2021. Capturing the dynamics of the China crude oil futures: Markov switching, co-movement, and volatility forecasting. Energy Econ. 103, 105622.
Liu, Y., Niu, Z., Suleman, M.T., Yin, L., Zhang, H., 2022. Forecasting the volatility of crude oil futures: the role of oil investor attention and its regime switching characteristics under a high-frequency framework. Energy 238, 121779.
Lucca, D.O., Moench, E., 2015. The pre-FOMC announcement drift. J. Financ. 70 (1), 329–371.
Makridakis, S., Spiliotis, E., Assimakopoulos, V., 2018. Statistical and machine learning forecasting methods: concerns and ways forward. PLoS One 13 (3), e0194889.
Malo, P., Sinha, A., Korhonen, P., Wallenius, J., Takala, P., 2014. Good debt or bad debt: detecting semantic orientations in economic texts. J. Assoc. Inf. Sci. Technol. 65 (4), 782–796.
Meng, F., Liu, L., 2019. Analyzing the economic sources of oil price volatility: An out-of-sample perspective. Energy 177, 476–486.
Möbert, J., 2009. Dispersion in Beliefs among Speculators as a Determinant of Crude Oil Prices. Deutsche Bank Research, Research Notes, p. 32.
Mostafa, M.M., El-Masry, A.A., 2016. Oil price forecasting using gene expression programming and artificial neural networks. Econ. Model. 54, 40–53.
Prado, F., Minutolo, M.C., Kristjanpoller, W., 2020. Forecasting based on an ensemble autoregressive moving average-adaptive neuro-fuzzy inference system–neural network-genetic algorithm framework. Energy 197, 117159.
Qadan, M., Nama, H., 2018. Investor sentiment and the price of oil. Energy Econ. 69, 42–58.
Ross, S.A., 1989. Information and volatility: the no-arbitrage martingale approach to timing and resolution irrelevancy. J. Financ. 44 (1), 1–17.
Safari, A., Davallou, M., 2018. Oil price forecasting using a hybrid model. Energy 148, 49–58.
Sharma, N., 1998. Forecasting Oil Price Volatility (Doctoral dissertation, Virginia Tech).
Shiller, R.J., 2017. Narrative economics. Am. Econ. Rev. 107 (4), 967–1004.
Shiller, R.J., 2020. Narrative Economics: How Stories Go Viral and Drive Major Economic Events. Princeton University Press.
Tahmassebi, H., 1986. The impact of the Iran-Iraq war on the world oil market. Energy 11 (4–5), 409–411.
Tang, L., Hammoudeh, S., 2002. An empirical exploration of the world oil price under the target zone model. Energy Econ. 24 (6), 577–596.
Verma, S., 2021. Forecasting volatility of crude oil futures using a GARCH–RNN hybrid approach. Intell. Syst. Account. Fin. Manag. 28 (2), 130–142.
Wu, B., Wang, L., Wang, S., Zeng, Y.R., 2021. Forecasting the US oil markets based on social media information during the COVID-19 pandemic. Energy 226, 120403.
Yu, L., Wang, S., Lai, K.K., 2008. Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. Energy Econ. 30 (5), 2623–2635.
Zhang, D., Ji, Q., Kutan, A.M., 2019. Dynamic transmission mechanisms in global crude oil prices: estimation and implications. Energy 175, 1181–1193.
Zhao, L.T., Zheng, Z.Y., Wei, Y.M., 2023. Forecasting oil inventory changes with Google trends: a hybrid wavelet decomposer and ARDL-SVR ensemble model. Energy Econ. 106603.
Zhe, J., Lin, Z., Lingling, Z., Bo, W., 2022. Investor sentiment and machine learning: predicting the price of China's crude oil futures market. Energy 123471.
Zhu, Y., Kiros, R., Zemel, R., Salakhutdinov, R., Urtasun, R., Torralba, A., Fidler, S., 2015. Aligning books and movies: towards story-like visual explanations by watching movies and reading books. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 19–27.