

RESEARCH

Open Access



# A hybrid econometrics and machine learning based modeling of realized volatility of natural gas

Werner Kristjanpoller<sup>1\*</sup>

\*Correspondence:  
werner.kristjanpoller@usm.cl

<sup>1</sup> Departamento de Industrias,  
Universidad Técnica Federico  
Santa María, Av. España 1680,  
Valparaíso, Chile

## Abstract

Determining which variables affect price realized volatility has always been challenging. This paper proposes to explain how financial assets influence realized volatility by developing an optimal day-to-day forecast. The methodological proposal is based on using the best econometric and machine learning models to forecast realized volatility. In particular, the best forecasting from heterogeneous autoregressive and long short-term memory models are used to determine the influence of the Standard and Poor's 500 index, euro-US dollar exchange rate, price of gold, and price of Brent crude oil on the realized volatility of natural gas. These financial assets influenced the realized volatility of natural gas in 87.4% of the days analyzed; the euro-US dollar exchange rate was the primary financial asset and explained 40.1% of the influence. The results of the proposed daily analysis differed from those of the methodology used to study the entire period. The traditional model, which studies the entire period, cannot determine temporal effects, whereas the proposed methodology can. The proposed methodology allows us to distinguish the effects for each day, week, or month rather than averages for entire periods, with the flexibility to analyze different frequencies and periods. This methodological capability is key to analyzing influences and making decisions about realized volatility.

**Keywords:** Deep learning, Heterogeneous autoregressive model, Long short-term memory model, Realized volatility, Volatility forecasting framework

## Introduction

Energy has always been a focus of study worldwide, given its importance in people's quality of life, global production, and environmental effects. Environmental effects and their consequent climate change are one of the main concerns in the sustainable development of human society (Wang et al. 2017), and this requires commitments by various countries, governments, organizations, and companies to reduce energy consumption or pollutant emissions into the environment, mainly by reducing the use of energy sources that emit excessive CO<sub>2</sub>.

In this context, renewable energies have gained relevance given that they are more environmentally friendly. Even so, fossil fuels represent more than 80% of the total

energy supply. Of the fossil fuels, natural gas (NG) has had greater importance in transitioning the world energy system and less influence on global climate change (Vidic et al. 2013). In fact, in recent decades, NG has been the fastest-growing primary energy source worldwide.

Price and its volatility are key factors influencing NG consumption and production. The importance of financial asset volatility stands out in studies by Anderson et al. (2005), Herrera et al. (2018) and Wang et al. (2020a), mainly because it is a recognized proxy for risk. Following the subprime crisis, commodity prices have been especially volatile. Knowing which variables influence NG price volatility supports better decisions. For example, we know that SPX volatility influences NG volatility, and financial reports indicate higher VIX forecasting. We can wait for an increase in NG price volatility and then decide whether to hedge or invest according to those forecasts. Furthermore, with the ability to handle and analyze high-frequency data, realized volatility has been defined and incorporates the intraday behavior of asset prices, which serves to better characterize risk. The consensus is that high-frequency models have outperformed low-frequency models in forecasting volatility (Lyócsa et al. 2021).

When assessing commodities associated with energy, most studies have focused on the volatility of crude oil. Given the importance that NG has acquired in recent times, studies that determine which variables or drivers influence NG price volatility are lacking. This study fills that gap.

The main objective of this study is to determine the key variables that influence NG price volatility by adjusting models to find the best forecasting model for each period or frequency analyzed. For each day of the out-of-sample period, the model and variables with the highest precision will be determined. The models to be analyzed are linear (econometric) and nonlinear (machine learning). In this way, it will be possible to conclude the linearity or nonlinearity of the volatility of the realized behavior and the influencing variables for the periods analyzed.

Following Degiannakis and Filis (2017), we use three information channels (stock, forex, and commodity) to determine the effects on the realized price futures volatility of NG. The financial assets used are the gold spot price (XAU), crude oil price (BRENT), Dow Jones Industrial Average, and the euro–US dollar exchange rate (EURO). Todorova et al. (2014) determined that volatility series of industrial metals contain useful incremental information to improve the accuracy of forecasting nonferrous metal futures price volatility, which determines that the volatility of commodities explains the behavior of the volatility of other commodities. These findings support the hypothesis that the volatility of the NG price can be explained by the volatilities of other commodities and of financial market indexes.

The classical econometric model, heterogeneous autoregressive (HAR), and the best machine learning model for forecasting time series, long short-term memory (LSTM), are used in the prediction analysis for volatility conducted on the price of NG. Then, mixing the best of econometrics and machine learning to predict realized volatility, we determine the best prediction per day. We use a nonlinear model to capture the complex patterns hidden in the linear models commonly used to forecast realized volatility.

The main innovation of this study is that it determines the influence of variables and nonlinearity of the realized volatility of the NG price through the most accurate model

for predicting realized volatility daily, weekly, or monthly. Moreover, the volatility prediction is out-of-sample rather than in-sample, as is typical to determine variable influences. This is important because we are interested in predicting the future and not the past. The methodology's focus relates to the forecasting power of the variables and models, leaving out the classical overfitting problem because the best model for each day, week, or month is assessed based on predictions, not training performance. Additionally, this methodology allows influencer variables to be found and accounts for the generalization ability of the models more than simply explanations during adjustment or training periods.

The contribution made by this work is supported in two dimensions, theoretical and empirical. The first relates to the methodology. In particular, a new method is proposed for determining the variables that influence volatility. This methodology is based on dynamically determining the variables that affect volatility (through time) and observing their effects according to their forecasting capacity for the out-of-sample period. This proposed methodology also allows for analysis of the accuracy of the models in each period (daily, weekly, monthly, or another required study period). During each period, the goal is to determine the best model and determine the relevant variables that influence volatility and the degree of nonlinear volatility behavior. The second dimension of the contribution made by this study is in applying the proposed methodology to better understand the behavior of NG price volatility. This allows one to identify which variables influence NG price volatility and how the influence and intensity of variables change over time. It also allows one to understand the period in which volatility behavior is nonlinear and when it is more linear.

The results and conclusions of this study are relevant for NG exporting and importing countries, electricity-generating companies, government policymakers and authorities, commodities investors, NG producers, traders, and consumers, and portfolio and risk managers.

The rest of the paper is structured as follows. Sect. "Literature review" reviews the existing literature. Sect. "Methodology and data" details the proposed methodology and the descriptive statistics of the variables. In Sect. "Analysis of results", the obtained results are analyzed.

### Literature review

Much evidence exists about the relationships between energy commodity prices, returns, and volatilities and financial indices, stock markets indices, and commodities. Primarily, a spillover, connectedness, or contagion effect is defined to characterize how prices, returns, and volatilities are affected. Regarding volatility and relationships among volatilities in commodity prices, Bouri et al. (2021) analyzed dynamic connectedness among the realized volatilities of 15 commodities. The results showed strong and moderate levels of volatility connectedness among commodities, and in some cases, evidence of a high proportion of realized volatility explained by the realized volatility of another commodity. Kumar et al. (2021) analyzed the relationships among NG, crude oil, gold prices, exchange rates, and a stock market index. They obtained evidence of a short-term, dynamic relationship among prices in energy, stock, and exchange rate markets. Asadi et al. (2022) studied volatility spillovers

within crude oil, NG, coal, stock, and currency markets in the USA and China, applying methodologies based on time and frequency. The results showed low total connectedness among energy, stock, and currency markets.

A key relationship studied within energy markets is spillover between oil and NG markets. Lovcha and Perez-Laborda (2020) demonstrated the dynamic volatility connectedness between oil and NG and how it changed over time. In addition, some studies have extended the variables to stock and financial markets. Stoupos and Kiohos (2021) found strong long-term dynamics between energy and developed stock markets. Liang et al. (2021) used GARCH-MIDAS to forecast US NG futures, finding that equity market volatility and geopolitical risk are two critical variables for forecasting. Caporin et al. (2019) analyzed relationships between the SPX and futures prices. The results showed relationships between the stock market and oil, NG, and ethanol price futures. Lin and Li (2015) found that gas prices are cointegrated with Brent crude oil prices, showing a long-term relationship and a price spillover from crude oil to NG markets.

For realized volatility forecasting, Lyócsa and Molnár (2018) found that HAR was the best econometric model for forecasting the realized volatility of NG. Recently, Alfeus and Nikitopoulos (2022) demonstrated that the HAR model can forecast the realized volatility of NG for daily and weekly frequencies. On the machine learning model side, Srivastava et al. (2023) showed the success of machine learning, particularly the LSTM-GJR-GARCH(1,1) with hyperparameter tuning, in forecasting the price volatilities of significant natural resources. Herrera et al. (2019) compared the forecasting performance of traditional econometric models with machine learning methods for the main energy commodities; the results showed that machine learning methods outperform traditional econometric methods, including for NG. Čeperić et al. (2017) forecast NG prices through classical time series models and machine learning methods, comparing the performance of time series and machine learning methods; they found that machine learning models performed better. Lyócsa et al. (2021) found that high-frequency volatility models outperformed low-frequency volatility models for short-term forecasts. Baruník and Křehlík (2016) used machine learning models and high-frequency data to improve the accuracy of forecasting energy market volatility. In the case of oil price volatility forecasting, in recent years, some studies have addressed the challenge and shown the importance of the issue (Jin et al. 2022; Song et al. 2022; Li et al. 2021; Luo et al. 2022a; Alfeus and Nikitopoulos 2022; Qu and Li 2023).

Rachal (2021) and Gao et al. (2021) addressed the challenge of forecasting NG prices by analyzing different models and giving flexibility to time-varying parameters but keeping the variable set fixed. There is no dynamic influence analysis of the variables. The conclusion is that the time-varying model is better for forecasting the price of NG. These conclusions support the idea of having hyperparameter flexibility over time. Our proposed methodology allows both the model and the parameters to change over time. Nevertheless, according to Armstrong (2001), no forecast method is always better than others; each has advantages and disadvantages. Time series dynamics require different models for different times. For this reason, our proposed methodology provides the flexibility to choose the best model for each day.

Econometric models require that the time series studied are stationary, regular, and linear; however, realized volatility time series are usually complex, irregular, or non-linear, or they have dynamic and random characteristics (Wang et al. 2018).

The proposed methodology for this study consists of two stages, training and forecasting. The data are separated into two groups, the first for training (in-sample) and the second (out-of-sample) for forecasting and thus analyzing the effects of the selected variables on the realized volatility of NG. The training is conducted for the two primary realized volatility prediction models, HAR (econometric) and LSTM (machine learning). For each model, a grid of parameters and variables is defined that includes independent variables in the econometric model and input variables in the machine learning model. The characteristics of the parameters of each model configuration allow us to infer and interpret variable influences on the realized volatility of NG and aspects of long memory and nonlinear behavior. With the model optimized during training for each configuration defined within the grid, its out-of-sample forecasting capacity is analyzed. The main difference from previous studies is that the effects of the studied variables on realized volatility can be determined dynamically, with the ability to discriminate the effects of influences having different time frequencies. With traditional methodologies, variable influences are usually obtained by the best in-sample configuration or the model that best predicts the complete out-of-sample period. With the proposed methodology, it is possible to obtain a dynamic analysis of influences and characterizations of the time series of realized volatility daily, weekly, and monthly. This capacity is fundamental since it is based on the fact that variables can change their influence on realized volatility over time, a condition that traditional methodologies assume is constant. Additionally, it has the flexibility to change the model to forecast over time, its parameter or hyperparameter configuration, and different sets of independent or input variables.

Most previous studies used econometric models to analyze spillover and relationships, because with machine learning, it is not easy to determine influences. With the proposed methodology, the influence of variables can be determined by applying machine learning models.

### Methodology and data

In this study, we want to test models from both econometric and machine learning approaches, as both have achieved great results in forecasting problems, especially those that are financial in nature. The volatility of financial series is important for investment and hedging decisions; therefore, determining the drivers that generate or explain volatility is essential. A primary characteristic of financial time series volatility is its heterogeneity, the influence of lagged short-, medium-, and long-term volatility, and volatility clusters. These characteristics suggest that a single volatility forecasting model does not exist for forecasting volatility over an entire analysis period but that the best model may vary over time. For this reason, many studies, as analyzed in the literature review, address financial series volatility prediction with econometric models, machine learning, and hybrid models.

According to recent literature (Lehrer et al. 2021; Liu and Lee 2021; Luo et al. 2022a, b; Qiu et al. 2019; Audrino et al. 2019), the HAR model is one of the best econometric approaches for forecasting realized volatility, as it offers excellent performance and is

easy to interpret. One drawback of using the HAR model, however, is that it is a linear approach, which could lead to biased results if the analyzed period has nonlinear characteristics. For this reason, we also analyze a machine learning approach, the LSTM network, to forecast volatility. LSTM is NN with a recurrent state, which allows its to identify time patterns in the data. The LSTM neural network model is a widely used approach in various sequence tasks and achieves excellent performance in vision, text, and time series. Although both models provide excellent results, our methodology could also be used with other approaches, which could lead to future studies.

Realized volatility (RV) is a volatility measurement based on intraday data with a particular frequency that could be in minutes or hours (Andersen and Bollerslev 1998; Patton 2011). Although estimates for RV differ depending on the problem, our study considers RV as presented in Eq. 1:

$$RV_t = \sum_{t=1}^T r_t^2 \quad (1)$$

where  $r_t$  is the high-frequency return—in this case, every 5 min—and T is the number of high-frequency periods in a day.

The first econometric models developed to model and forecast volatility were the ARCH models developed by Engle (1982). These models were later generalized through the GARCH model by Bollerslev (1986). With the greater availability of high-frequency data and the definition of volatility carried out, the HAR model by Corsi (2009) was developed. The superiority of HAR models over other econometric models for forecasting volatility has been demonstrated in various studies of realized volatility (Andersen et al. 2007; Busch et al. 2011; Fernandes et al. 2014; Corsi and Reno 2012; Bekaert and Hoerova 2014; Santos and Ziegelmüller 2014; Celik and Ergin 2014; Seo and Kim 2015; Vortelinos 2017; Bergsli et al. 2022).

The HAR model can be interpreted economically and is given by Eq. 2:

$$RV_{t+1} = \beta_0 + \beta_d RV_t^{(d)} + \beta_w RV_t^{(w)} + \beta_m RV_t^{(m)} + \varepsilon_t \quad (2)$$

where  $RV_t^{(f)}$  is the realized volatility computed by the mean of the last  $f$  days of the previous  $t$ ; ( $d$ ) refers to the daily frequency, ( $w$ ) to the weekly frequency, and ( $m$ ) to the monthly frequency, or they may refer to realized lagged volatility at 1 day, 5 days, and 22 days. This model is usually estimated using ordinary least squares.

Machine learning models have broken into volatility forecasting studies. In the first stage, artificial neural networks were used to forecast volatility (Donaldson and Kamstra 1996, 1997; Hamid and Iqbal 2004; Kristjanpoller et al. 2014; Monfared and Enke 2014; Kristjanpoller and Minutolo 2016; Seo and Kim 2020; Bucci 2020, among others). Given the importance of forecast feedback and persistence in periods of volatility, the recurrent neural network (RNN) was used to forecast volatility (Tino et al. 2001; Bekiros and Georgoutsos 2008; Hsieh et al. 2011; Maknickiene et al. 2018; Petneházi and Gál 2019; Bucci 2020).

Finally, LSTM is a widely used neural network model for analyzing time-dependent data. Hochreiter and Schmidhuber (1997) introduced LSTM, corresponding to the RNN type. The LSTM explicitly avoids long-term dependency by remembering information over long

periods. Various studies have demonstrated the superiority of LSTM models for forecasting volatility. Jiang et al. (2019) demonstrated the superiority of LSTM models over traditional RNNs. Rodikov and Antulov-Fantulin (2022) demonstrated that LSTM models are superior to econometric models. Mücher (2022) determined the superiority of LSTM models over HAR and ARIMA in improving forecasts of the realized volatility of IBM stock. Yuyan et al. (2023) showed that LSTM is better than other artificial intelligence models at predicting realized volatility and that hybrid LSTM and HAR models have the best forecast accuracy. Li et al. (2023) demonstrated that the forecasting performance of the LSTM-based model is better than the GRU-based model for predicting realized volatility in the energy stock market.

LSTM is a specific type of RNN structure that addresses the learning of long-term relationships resulting from exponentially increasing or vanishing gradients (Bengio et al. 1994). The main operation of LSTM is through gates that regulate the flow of information and thus maintain the necessary history to carry out accurate temporal modeling of the analyzed data. Let  $X_t$  be an array of explanatory variables indexed by time and  $Y_t$  be the variable of interest. Figure 1 shows the general operation of LSTM (Greff et al. 2016). The main feature of LSTM is that it works with internal modules to perform calculations that allow time trends to be recognized. Mathematically,

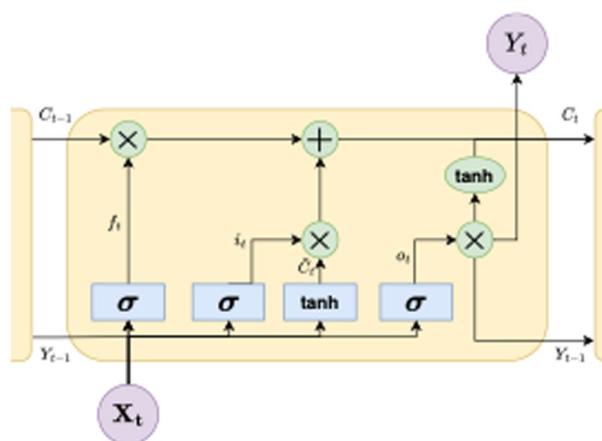
$$f_t = (W_f[Y_{t-1}, X_t] + b_f) \quad (3)$$

$$i_t = (W_i[Y_{t-1}, X_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C[Y_{t-1}, X_t] + b_C)$$

$$C_t = f_t \cdot C_{t-1} + i_t \tilde{C}_t$$

$$o_t = (W_o[Y_{t-1}, X_t] + b_o)$$



**Fig. 1** Diagram of long short-term memory (Greff et al. 2016)

$$\hat{Y}_t = o_t \bullet \tanh(C_t)$$

where  $\sigma$  is the sigmoid function and  $\tanh$  is the hyperbolic tangent function,  $f_t$  is the value of the forget gate,  $i_t$  is the input data for the next cell, and  $o_t$  is the value of the output gate.

The central idea is to pass information from period to period. A gate controls this handover with weights that must be learned. LSTM cell information is stored as a variable ( $C_t$ ).

The interaction of LSTM cells enables LSTM to model time series with a high degree of nonlinearity and capture long-term dependencies in the data.

To determine the accuracy of model predictions and for performance analysis, three loss functions are used: mean squared error (MSE), root mean squared error, and mean absolute error (MAE). The formula for each function is detailed below:

$$MSE = \frac{1}{N} \sum_1^N (Y_t - \hat{Y}_t)^2 \quad (4)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_1^N (Y_t - \hat{Y}_t)^2} \quad (5)$$

$$MAE = \frac{1}{N} \sum_1^N |Y_t - \hat{Y}_t| \quad (6)$$

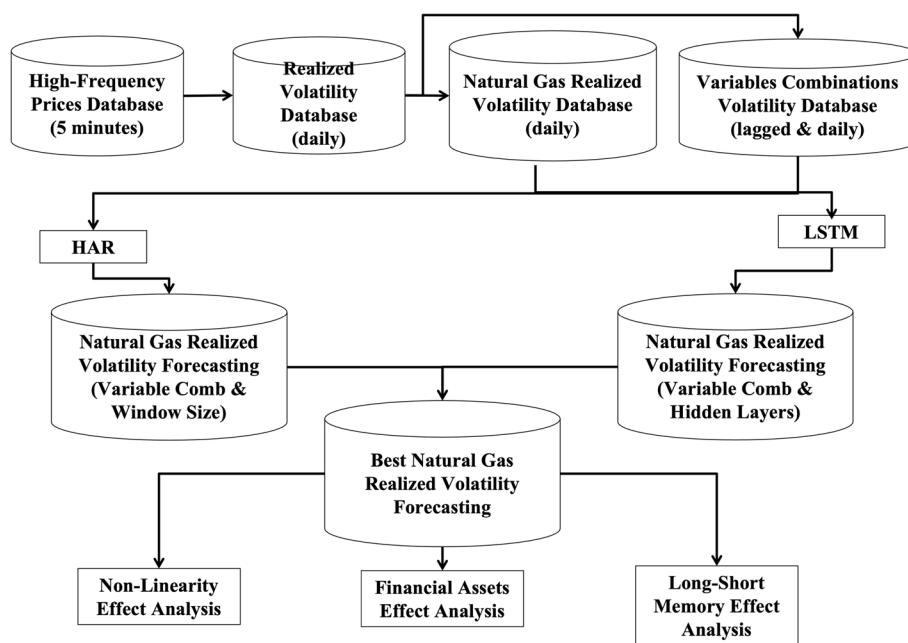
### Proposed analysis model

The proposed methodology to determine the effect of the main financial assets on the realized volatility of NG consists of predicting the realized volatility with the widely used forecasting models of the econometric and machine learning approaches (see Fig. 2).

Specifically, we fit different models with various configurations, and each day, we analyze the models to determine which had the best forecasting performance. This is done for each type of model, described as econometric and machine learning models. We analyze the HAR and LSTM models for each category, as they seem to be the best-performing approaches for forecasting realized volatility, as indicated by the literature review. The main idea of this work is to analyze the daily, weekly, and monthly impacts of models and variables. For the first part, we set various configurations for each model considered for testing. For the second part, we analyze input variable combinations that could influence the final forecast, leading to more models (Fig. 2).

The steps of the proposed methodology are.

- (1) Select the main and explanatory variables according to the literature review and available frequency of data.
- (2) Extract the high-frequency price data for the selected explanatory variables for the analyzed period.



**Fig. 2** Diagram of proposed methodology

- (3) Compute the log return for each period (for example, every 5 minutes). For each day, compute the realized volatility according to Eq. 1.
- (4) Extract natural gas high-frequency price data for the analyzed period.
- (5) Compute the log return for each period (for example, every 5 minutes). For each day, compute the realized volatility according to Eq. 1.
- (6) The HAR model is adjusted for different configurations for the training period.
- (7) The HAR model with explanatory variables is adjusted for different configurations for the training period.
- (8) The LSTM model is adjusted for different configurations for the training period.
- (9) The LSTM model with explanatory variables is adjusted for different configurations for the training period.
- (10) All the models estimated in steps 5, 6, 7, and 8 are used to forecast the out-of-sample realized volatility.
- (11) The model with the lowest error for the realized volatility forecast is selected for each day of the out-of-sample.
- (12) Given the best model for each out-of-sample day, analyses are conducted on the linearity of realized volatility behavior, the effects of explanatory variables, and the effects of long and short memory.
- (13) To conduct the weekly and monthly analyses, the best model for each day of the week or month, respectively, is taken as a basis. A linearity analysis is then conducted on realized volatility behavior, the effects of explanatory variables, and the effects of long and short memory.

Four possible rolling window sizes are used to configure the HAR model: 700, 1000, 1400, and 2100 days. The various LSTM models are optimized according to the grid of the number of layers, neurons per layer, number of lags, and rolling window lengths (Table 1).

**Table 1** Hyperparameters of long short-term memory (LSTM) configuration

Hyperparameter	Values
Number of layers	1, 2, 3, 4
Neurons per layer	20, 40, 60, 100, 150
Number of lags (days)	3, 6, 10
Window length (days)	700, 1000, 1400, 2100

The HAR model has 20 variable combinations for each configuration. One combination is solely the lagged values of the realized volatility of NG, four are separate combinations of the realized volatility of each financial asset, and the remaining fifteen are all possible combinations of the grouping of the four financial assets. Thus, the HAR model has 80 forecasts for each forecast day.

Each forecast is treated as an individual model for the variable importance analysis.

The LSTM model uses 16 variable combinations: a combination of only the lagged values of the realized volatility of NG and 15 combinations of all possible combinations of the four financial assets. Therefore, the LSTM model has 3840 forecasts each day.

Unlike the analysis of the HAR models, here we preprocess the 3840 models and determine the best according to the variable combinations and the number of layers, resulting in only 64 valid models for comparison. Thus, adding these to the possible HAR combinations, we will analyze 144 models to determine which is best on each forecast day.

Finally, we extract the financial assets associated with the best model for each day to determine the most influential variables in a disaggregated manner. This helps to better understand the key drivers in forecasting NG, as we hypothesize that no variable persistently provides the best explanations over the entire forecasting period. Moreover, a primary contribution is to determine the influential dynamics of financial assets on NG given the best model between the econometric and machine learning approaches. The day-to-day analysis clarifies the most influential variables not based directly on the model used but rather on the best-performing approach to forecasting.

### Data description

The NG price is the main variable in this study. This variable is characterized by the NG continuous contract future, and the price is expressed in US dollars per MMBtu. The criteria for selecting explanatory variables are the economic relationship and high-frequency data availability. Following Degiannakis and Filis (2017), Kumar et al. (2021), and Asadi et al. (2022), the explanatory variables were selected from the stock, forex, and commodities markets. The explanatory variables used are the XAU in US dollars, the BRENT futures price, the Standard and Poor's 500 (SPX), and the EURO. The XAU was selected because gold is used as a refuge in crisis periods and is a predictor of poor economic performance. The SPX was chosen because it is a good predictor of US and world economic performance. The EURO can serve as a buffer against or dampen the effects of inflation when energy prices rise. BRENT is an energy alternative to NG for two reasons: substitution and comovement in economic trends.

All the high-frequency data of these variables were extracted from [www.dukascopy.com](http://www.dukascopy.com). These variables were sampled at 5-min intervals to compute the daily realized volatility. For each variable, the realized volatility was calculated according to Eq. 1.

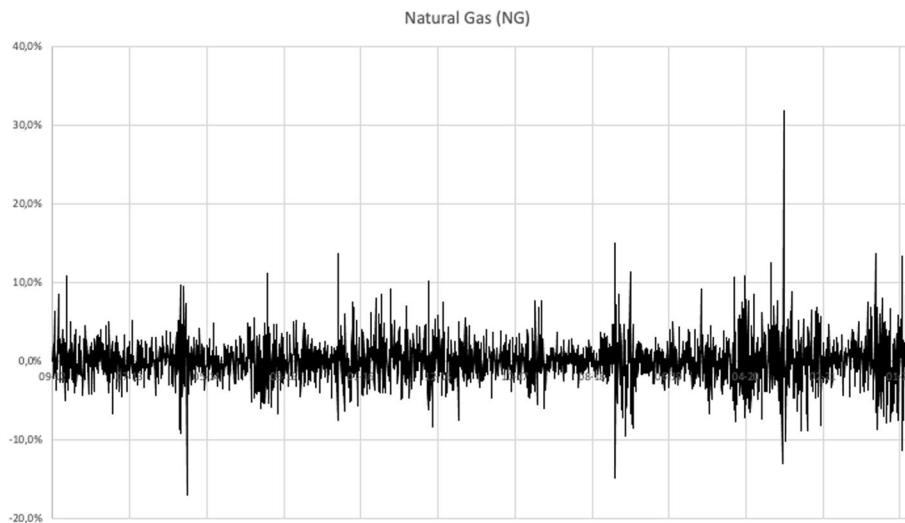
The period analyzed is from September 3rd, 2012, to January 31st, 2022 (977,497 intraday observations and 2724 daily observations, excluding nonwork days). Of the total analysis period, the last 2 years will be the prediction period to determine the variables that influence NG price volatility. This period is from February 4, 2020, to January 31, 2022, which contains 619 days of transactions.

Figure 3 shows NG price evolution for the entire period. In February 2014, a price peak was observed, exceeding \$6.40 (USD) per MMBtu. Then the price began to decline, reaching a valley of lower prices between the end of 2015 and the beginning of 2016. Subsequently, the price rose and remained around \$3 (USD) per MMBtu for some years. Toward the end of 2018, the price peaked above \$4.50 (USD) per MMBtu; it later fell into a valley low in early 2020. Afterward, an upward trend began, peaking above \$6 (USD) per MMBtu toward the end of 2021 and then falling. Brent oil had the most similar price behavior relative to NG price evolution. All the other variables had different price behaviors during the studied period (Appendix 2).

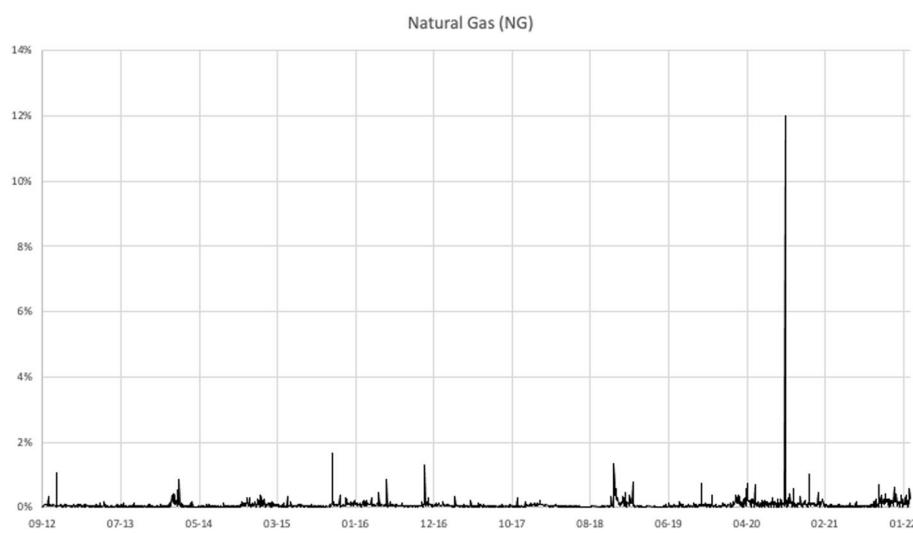
EURO had a period of high prices until mid-2014, when it fell into a channel until the end of 2017, where it exceeded 1.20 USD. It then returned to the same channel range until the end of 2020 and the beginning of 2021, when it again exceeded a value of 1.20 USD. The SPX had an upward trend throughout the analyzed period, increasing its value by approximately three times. A sharp drop can be observed in March 2020 due to the start of the COVID-19 pandemic. The price of Brent oil started the analysis period at a high value until mid-2014, at more than 100 USD per barrel. Then it experienced a strong downward trend until the end of 2015; subsequently, its price recovered, reaching a peak at the end of 2018 before beginning to fall, reaching its minimum value in that period when the pandemic began. The price of gold began the study period in decline until mid-2013, and remained relatively stable until the third



**Fig. 3** Natural gas (NG) price evolution



**Fig. 4** Natural gas (NG) price return evolution



**Fig. 5** Natural gas (NG) price realized volatility evolution

quarter of 2018. At that point, an upward trend began, reaching a price peak in mid-2020 and then maintaining prices of around 1800 USD per ounce.

Figures 4 and 5 show NG volatility clusters in 2016, at the end of 2018, and in the last 2 years of the study. The COVID-19 pandemic influenced this last period. All the explanatory variables had a volatility peak at the beginning of COVID-19 (03-2020), but they showed different volatility behaviors during the rest of the period.

The average NG price during the study period was 3.11 USD per MMBtu, while the realized volatility was 0.078%. During the entire period, the price fell to nearly half the average, and its maximum doubled the average price (Table 2). The standard deviation of return and volatility are much higher than the average, showing high price

**Table 2** Descriptive statistics of natural gas (NG) price, return, and realized volatility

Natural gas (NG)			
	Price	Return (%)	Realized volatility (%)
Mean	3.1077	0.0266	0.0783
St. deviation	0.8413	2.7582	0.2480
Median	2.9211	-0.0441	0.0459
Min	1.5280	-16.9456	0.0000
Max	6.3993	31.7281	11.9986

**Table 3** Results of best forecasting with the long short-term memory (LSTM) model

Number of layers			Neurons per layer			Number of lags			Window size		
Q	Obs	Perc (%)	Q	Obs	Perc (%)	Q	Obs	Perc (%)	L	Obs	Perc (%)
1	66	26.2	20	23	9.1	28	51	20.2	700	109	43.3
2	111	44.0	40	36	14.3	30	138	54.8	1000	42	16.7
3	36	14.3	60	13	5.2	37	63	25.0	1400	23	9.1
4	39	15.5	100	27	10.7				2100	78	31.0
			150	153	60.7						

Q is the number of layers, neurons per layer, and lags. L is the length of the windows to train. Obs is the days when the LSTM was the best model with this configuration

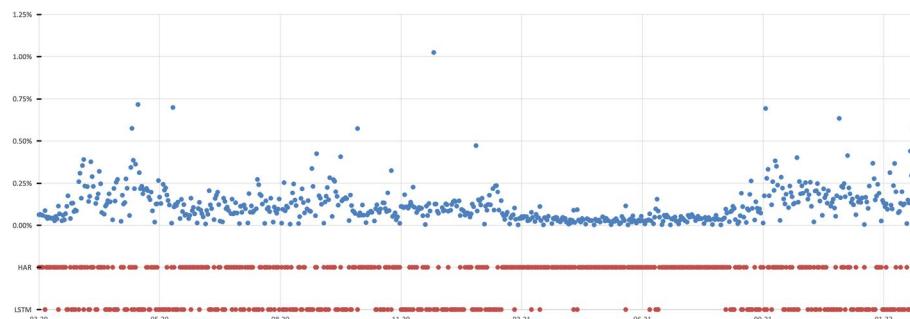
volatility. For its part, EURO is the variable with the least volatility during the period, while the Brent oil variable has the highest volatility (Appendix 1).

### Analysis of results

The influence of EURO, gold, Brent crude, and SPX prices on NG price volatility was analyzed for the February 4, 2020, to January 31, 2022, period; the linear relationship (HAR) explained more days than the nonlinear one (LSTM). For the 619 days analyzed, the HAR model provided the best forecast (least daily error) for 367 days (59%), whereas the LSTM model provided the best forecast for the other 252 days (41%). In the case of the best forecasts predicted by HAR, 41% corresponded to the model with the shortest training window (700 days) and 31% to the second-shortest adjustment window (1000 days). In 72% of cases, the best forecast was made with a short memory. In the case of the best forecasts made by LSTM in 109 days, the best training window was 700 days, and in 42 days, it was 1000 days, with 59.9% being small windows (Table 3). Therefore, independent of the model, on 67% of the days, the model was fitted with a small training window showing a short memory for training and forecasting.

In the case of nonlinearity, better forecasting with LSTM occurred in 70.2% of cases but with one or two layers showing low nonlinearity. However, in 71.4% of the cases, a high number of neurons per layer showed a high information-processing requirement, which generates higher nonlinearity.

When analyzing the influence of the variables, the results show that in 78 of the 619 days analyzed (12.6%), the best model only used lagged NG volatility values. However, in 87.4% of the observations, the variables used as exogenous improved the forecast



**Fig. 6** Evolution of the best model in the study period daily analysis. Note Red filled circle indicates the best model for each day; dark blue filled circle represents realized volatility of the Natural Gas

**Table 4** Effect of financial assets on natural gas volatility

	Days effect	Standardized effect	
BRENT	206	38.1%	111
EUR	299	55.3%	217
XAU	159	29.4%	87
SPX	196	36.2%	126

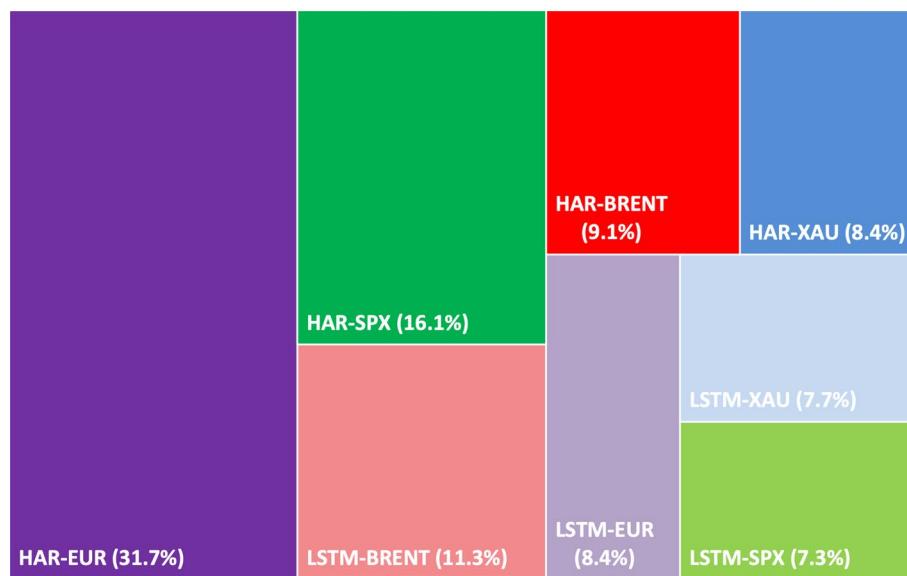
Days correspond to the number explained by the Natural Gas volatility. Standardized refers to days standardized by the number of variables explained each day

of the best model built only based on the autoregressive terms of the volatility of NG. In the case of the best models with only the autoregressive terms of the volatility of NG in 82.1%, it was through an LSTM model. Figure 6 shows that in the period of low volatility between February and September 2021, the best model for predicting volatility is HAR.

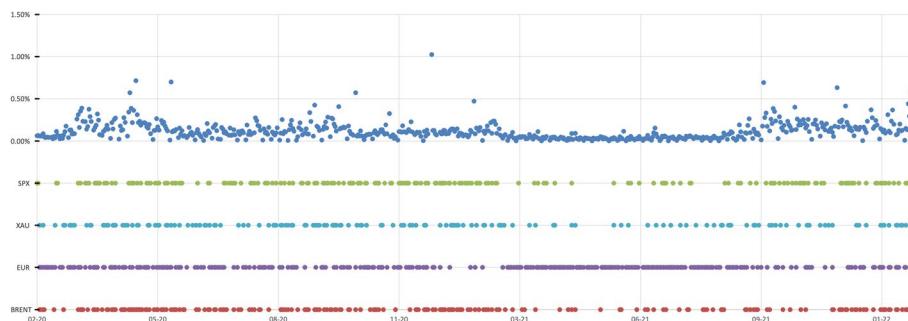
The results of the effect of financial assets on NG volatility are presented in Table 4. EURO explained NG volatility on 55.3% of the days where there was interference from exogenous variables. In comparison, Brent oil and the SPX explained volatility for 38.1% and 36.2%, respectively. When the daily effect was standardized by dividing by the number of variables that influenced each day, EURO continued to maintain a high capacity for explaining NG volatility. This result showed that for several days, EURO was the only explanatory variable that explained realized volatility, unlike the other variables. For this reason, the standardized percentage of the other variables dropped proportionally more.

In Fig. 7, the effects of EURO and SPX are established mainly with the HAR model, with nearly 50% of the standardized effects. In the case of the LSTM model, the most significant standardized effect is with Brent oil. Figure 8 shows the high concentration of the EURO effect at the beginning of 2020 and March–September 2021. This last period is related to a cluster of low volatility in the NG price. Both Brent oil and the SPX had a significant effect in 2020, but by 2021 their effects began to get smaller. In the case of gold, its effects are clustered and spaced over time.

As seen in Table 5, for periods of low realized volatility, the best predictor was the HAR model (98.7%). The main variable to explain the realized volatility was the EURO with 134.0 standardized effects (87.0%). In the second quartile, the HAR models have a



**Fig. 7** Financial asset effects by models (standardized). Note The red area indicates the Brent effect, the purple area indicates the Euro effect, the blue area indicates the Gold effect, and the green area indicates the Standard and Poor's 500 effect. The dark colors are the cases for HAR models, and the light colors for LSTM



**Fig. 8** Evolution of the financial asset effect on natural gas volatility in the study period daily analysis. Note Red filled circle indicates the Brent effect, purple filled circle indicates the Euro effect, light blue filled circle indicate the Gold effect, and light green filled circle the Standard and Poor's 500 effects; dark blue filled circle represents the realized volatility of the Natural Gas

**Table 5** Model and financial asset effects on natural gas volatility by realized volatility quartile

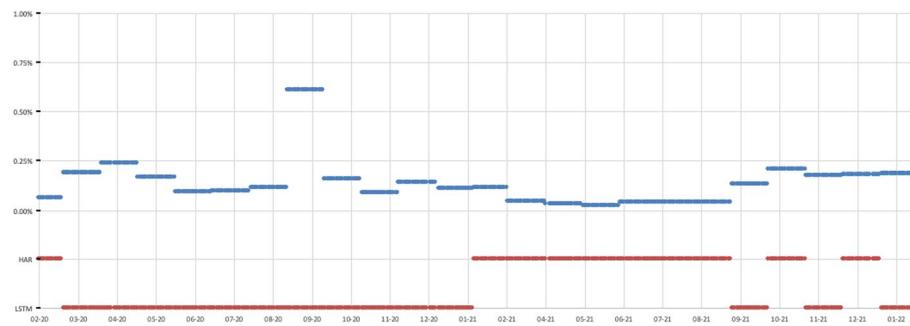
RVQ	N	HAR	LSTM	AR	BRENT	EUR	XAU	SPX
1	154	152	2	7	6.2	134.0	4.5	2.3
2	155	125	30	2	39.5	28.2	35.7	49.7
3	155	24	131	18	36.8	33.6	32.1	34.6
4	155	66	89	51	28.1	21.1	15.1	39.8

RVQ is the Realized Volatility Quartile, and the effects are the standardized effects. AR refers to the lagged Natural Gas realized volatility

**Table 6** Financial asset effects on natural gas volatility by realized volatility quartile and model

RVQ	N	AR	BRENT	EUR	XAU	SPX	AR	BRENT	EUR	XAU	SPX
1	154	7	5.5	133.3	3.8	2.3	0.0	0.7	0.7	0.7	0.0
2	155	2	28.0	24.2	29.8	41.0	0.0	11.5	4.0	5.8	8.7
3	155	3	4.9	4.4	3.6	8.1	15.0	31.8	29.2	28.5	26.5
4	155	2	10.8	9.3	8.3	35.5	49.0	17.3	11.8	6.8	4.3

RVQ is the Realized Volatility Quartile, and the effects are the standardized effects. AR refers to the lagged Natural Gas realized volatility

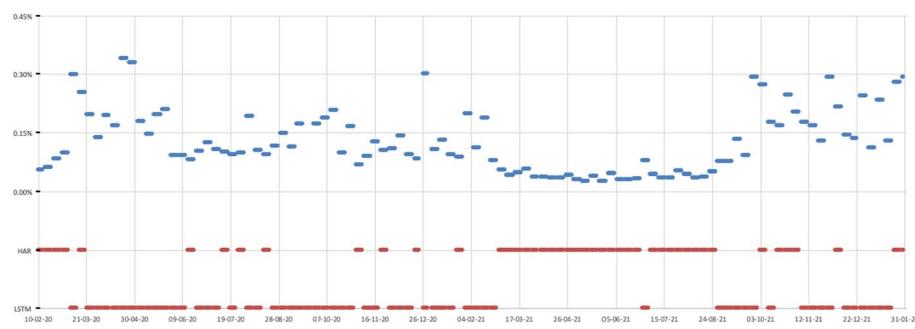


**Fig. 9** Evolution of the best model in the study period monthly analysis. Note Red filled circle indicates the best model for each month; dark blue filled circle represents realized volatility of the Natural Gas

high capability to forecast the NG realized volatility (80.6%). The financial assets had a similar standardized effect in the third quartile, but in the highest volatility quartile, the SPX had a stronger influence. For the two quartiles of high volatility, the LSTM model showed the best accuracy, indicating that nonlinearity is recommended to predict high volatility.

Finally, the effects are classified by model and realized volatility quartile (Table 6). The combination of the HAR model and EURO in the lowest volatility quartile best explains the NG volatility. In the second volatility quartile, EURO had a lower effect, and SPX had a stronger influence on explaining the realized volatility of NG. LSTM dominates the third quartile, and the influence of the financial assets is similar. In the highest volatility quartile, two combinations in high proportion explain the realized volatility of NG: the LSTM-AR (31.6%) and the HAR-SPX (22.9%).

The analysis was carried out monthly and weekly to complement the results obtained in the daily analysis and demonstrate the flexibility of the proposed methodology. It can be seen in Fig. 9 that from 2020 to September 2021, in months of low realized volatility, the HAR model is the best, whereas in months of high realized volatility, it is the LSTM model. Subsequently, from October 2021 to January 2022, there is an intercalation between the models, with no explicit high- or low-volatility behavior in this period. These findings are confirmed in the weekly analysis (see Fig. 10). In some weeks within months of high volatility, the HAR model was the best, coinciding with weeks of low realized volatility. In the monthly analysis for those months, the best model was LSTM, but when seeing 1 or 2 weeks of low volatility weekly, the best

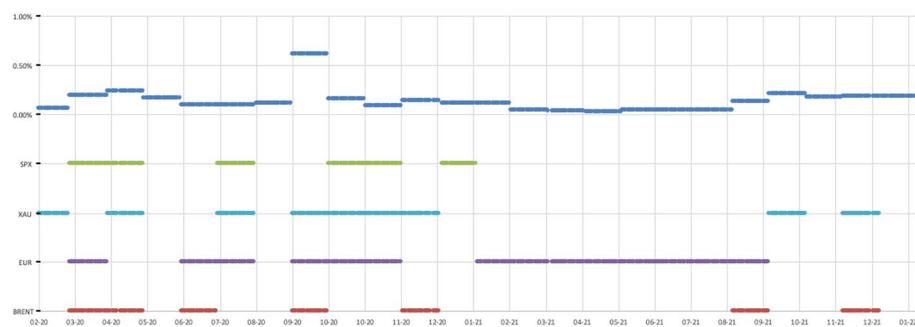


**Fig. 10** Evolution of the best model in the study period weekly analysis. Note Red filled circle indicates the best model for each week; dark blue filled circle represents realized volatility of the Natural Gas

**Table 7** Best model and effects of financial assets on natural gas volatility

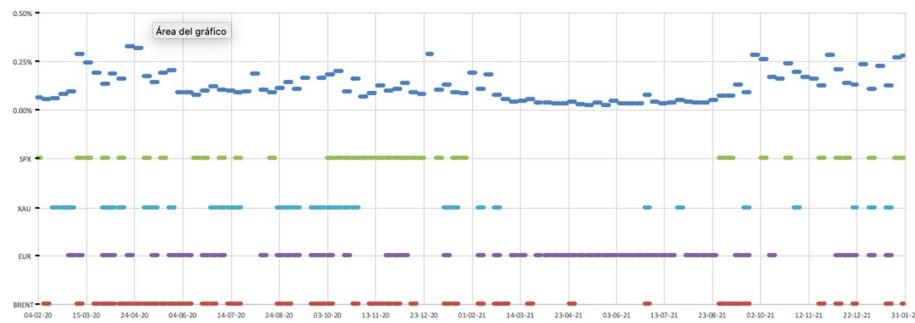
	Daily (%)	Weekly (%)	Monthly (%)
<i>Model</i>			
HAR	59.3	43.8	41.7
LSTM	40.7	56.2	58.3
<i>Variable</i>			
BRENT	38.1	50.0	35.0
EURO	55.3	62.0	70.0
XAU	29.4	35.9	45.0
SPX	36.2	40.2	30.0

Percentage of days, weeks and months as the best model to forecast the Natural Gas realized volatility and effect of financial asset which explained by the Natural Gas volatility



**Fig. 11** Evolution of the financial asset effect on natural gas volatility in the study period monthly analysis. Note: Red filled circle indicates the Brent effect, purple filled circle indicates the Euro effect, light blue filled circle indicate the Gold effect, and light green filled circle the Standard and Poor's 500 effects; dark blue filled circle repersents the realized volatility of the Natural Gas

model was HAR. The same happens in the period of low monthly volatility, which, when broken down into weekly volatility, shows a week of high volatility at the end of June 2021, for which the best model was LSTM. Since the first week of October 2021, there has yet to be an apparent behavior of the best model concerning weekly



**Fig. 12** Evolution of the financial asset effect on natural gas volatility in the study period weekly analysis. Note Red filled circle indicates the Brent effect, purple filled circle indicates the Euro effect, light blue filled circle indicate the Gold effect, and light green filled circle the Standard and Poor's 500 effects; dark blue filled circle repersents the realized volatility of the Natural Gas

volatility. Interestingly, the lower the frequency, the greater the number of months; the best model is LSTM (monthly); at a daily level, almost 60% of the days, the best model is HAR (Table 7). The nonlinear behavior of the days or weeks of high realized volatility that makes up a particular month characterizes its monthly behavior.

The weekly and monthly analyses (Figs. 11 and 12) show that the variable with the most significant influence in periods of low realized volatility is the euro, while in periods of high realized volatility, the most influential variable is Brent oil. Since October 2021, the variables that had the most influence at the monthly level were gold and Brent oil, while at a weekly level, it can be observed that all four variables studied influenced the realized volatility of NG.

As can be seen in Table 7, EURO, independent of frequency, is the financial variable with the most significant influence on NG price volatility. The other variables studied generally were included in between 30 and 40% of the best prediction models. At a weekly level, the 50% influence of Brent oil is striking. The explainability of EURO and XAU increases as the frequency becomes lower. The characteristics of these explanatory variables show that they are better long-term predictors than BRENT and SPX. The maximum explanatory power that SPX and BRENT have is at the monthly level, which indicates that their volatility adjusts more quickly to news and economic events. However, over the longer term, EURO and XAU are more efficient in projecting the realized volatility of NG.

## Discussion

The traditional methodology is applied first to understand the innovation of the proposed methodology to determine which financial assets influence the realized volatility of NG. The traditional methodology addresses determining the influential variables for the entire out-of-sample period, selecting the model that generates the lowest MSE in the whole period. The influence of the variables studied is concluded based on the best model obtained, its configuration, and the variables used.

**Table 8** Model configuration for the best model in the entire period for classical analysis

Ranking	Model	Financial assets	NoL	NpL	WS	MSE	RMSE	MAE
1	LSTM	BRENT-SPX	1	150	1000	<b>2,38E-01</b>	<b>4,88E-01</b>	8,85E-04
2	LSTM	BRENT-XAU	1	150	1000	2,38E-01	4,88E-01	8,64E-04
3	LSTM	BRENT-XAU-SPX	1	150	700	2,38E-01	4,88E-01	8,27E-04
4	LSTM	BRENT-EURO-XAU	2	150	700	2,38E-01	4,88E-01	9,31E-04
5	LSTM	BRENT	1	150	700	2,38E-01	4,88E-01	<b>8,18E-04</b>
	HAR	SPX			700	2,43E-01	4,93E-01	8,35E-04

In bold is the best model by loss function criteria.

NoL Number of layers, NpL Neurons per layer, WS Window size.

**Table 9** Model configuration for the best model for the entire period using classical analysis (daily analysis)

Model	Type	MSE	RMSE	MAE
LSTM	Only	2,38E-05	4,88E-03	8,18E-04
HAR	Only	2,43E-05	4,93E-03	8,35E-04
LSTM	Combined	2,32E-05	4,82E-03	5,54E-04
HAR	Combined	2,35E-05	4,84E-03	4,92E-04
Proposed model	Hybrid	<b>2,31E-05</b>	<b>4,81E-03</b>	<b>3,84E-04</b>

Only: the best configuration for the whole period. Combined: the best forecasting for each day of the model. Hybrid: the best forecasting for each day of independent of the model (HAR or LSTM). In bold the best model by loss function criteria

Applying the traditional methodology for all the possible configurations of the HAR and LSTM models, given the hyperparameter grid and the variables studied, only one of the two models would be chosen, with a single hyperparameter configuration and a single set of variables of all the variables studied. By performing the analysis with the proposed methodology, it is expected to improve the accuracy of the forecast. By improving the accuracy of the forecast, the conclusions that can be drawn about the influence of the variables on the volatility of the price of NG will be more robust.

The results presented in Table 8 correspond to the application of the traditional methodology. The model whose forecasts are the most accurate is chosen, that is, the model that obtained the lowest out-of-sample loss function. Under this traditional analysis methodology, the conclusion would be that the best way to predict the realized volatility of NG is by using an LSTM model with one hidden layer by MSE criteria. With this model, we can conclude that the NG realized volatility time series behavior is nonlinear but with low nonlinearity since the optimal model only has one hidden layer. The training window was 1000 days, which shows that it has a short memory to learn, whereas 150 neurons implies a high information-processing requirement.

The influencing variables are Brent oil and the SPX, leaving out gold and EURO. Overall, the best HAR model has a 2.248% higher MSE than the best LSTM. The traditional methodology is based on the lowest MSE for the entire period methodology, the best model, and the influencing variables in general of the total period, and there is no more refined analysis day by day, losing helpful information to make conclusions and to be the basis for decision-making. By MAE criteria, the conclusions are similar, defining the

**Table 10** Model configuration for the best model in the entire period for classical analysis (weekly analysis)

Model	Type	MSE	RMSE	MAE
LSTM	Only	1,40E-04	1,18E-02	3,18E-03
HAR	Only	1,45E-04	1,20E-02	3,28E-03
LSTM	Combined	1,38E-04	1,17E-02	2,73E-03
HAR	Combined	1,41E-04	1,19E-02	2,73E-03
Proposed Model	Hybrid	<b>1,37E-04</b>	<b>1,17E-02</b>	<b>2,52E-03</b>

Only: the best configuration for the whole period. Combined: the best forecasting for each day of the model. Hybrid: the best forecasting for each day of independent of the model (HAR or LSTM). In bold the best model by loss function criteria

**Table 11** Model configuration for the best model in the entire period for classical analysis (monthly analysis)

Model	Type	MSE	RMSE	MAE
LSTM	Only	6,14E-04	2,48E-02	9,54E-03
HAR	Only	6,28E-04	2,51E-02	9,87E-03
LSTM	Combined	6,05E-04	2,46E-02	8,95E-03
HAR	Combined	6,17E-04	2,48E-02	8,86E-03
Proposed Model	Hybrid	<b>6,04E-04</b>	<b>2,46E-02</b>	<b>8,61E-03</b>

Only: the best configuration for the whole period. Combined: the best forecasting for each day of the model. Hybrid: the best forecasting for each day of independent of the model (HAR or LSTM). In bold the best model by loss function criteria

BRENT as the only variable, and the hyperparameters are the same, but the window size is shorter (700 days). The accuracy difference is lower measured by MAE than MSE, supporting the best performance of LSTM models in high volatility periods.

When applying the proposed methodology daily, the HAR model better predicts realized volatility on more days than the LSTM model. Given the results, we can conclude that, on average, when the best model is the LSTM, it has much less error than the HAR model, while when the HAR model is the best day model, the LSTM model has a slightly higher error.

Suppose we do a deeper analysis on the possibility of finding the best forecast for each day for each model, HAR and LSTM. In that case, we can observe that the hybrid characteristic of the proposed methodology improves the precision. This accuracy improvement means that although each is the best model for forecasting volatility made from the econometric and machine learning models, mixing them generates a better model (Table 9). For all the loss functions, the proposed methodology had a better performance. Then the interpretations of the results and the conclusions are better supported because the methodology had better accuracy than the classical methodology.

The advantages of the proposed hybrid methodology can also be observed in the weekly and monthly analyses of Tables 10 and 11, respectively. Therefore, the influences of the variables on NG price volatility and the conclusions that can be obtained are determined from a more accurate model.

Like Degiannakis and Filis (2017) found evidence that information channels (stocks, forex, commodities, and macro) improve the predictive accuracy of oil prices' realized volatility. Our study found effects from stocks, forex, and commodities on NG price volatility. Macro information was not included because the frequency analyzed did not match the proposed methodology. In particular, the volatility of EURO had more information about the future behavior of NG price volatility than the other variables. It is the only explanatory variable, even in a low monthly NG price volatility cluster. As recently discovered by Garzón and Hierro (2022) for the price of oil, EURO variations partially dampen the impact of changes in the dollar price of oil on inflation; we can conclude that the same effect occurs with the NG price.

Oil price volatility is the second most important explanatory variable of NG price volatility. This influence has two primary drivers: real demand and substitution. First, since both are energy commodities, they tend to behave similarly in the face of growth or decline in the world economy. Second, when one of the two commodities becomes more expensive than the other, it influences the quantity demanded of the cheaper, which affects the price in the short run.

The influence found on the part of the stock market index indicates its recognized ability to take precedence over changes in economic growth. Thus, it influences the volatility of the NG price, anticipating the demand for NG given the market projections on the gross domestic product of the primary NG-consuming countries. Finally, gold is a refuge commodity; its volatility indicates uncertainty about the world economy, which affects projections of real demand for NG. In fact, in a longer-term view, gold was the second explanatory variable for the monthly analysis of the NG price volatility.

Regarding the relationship that the explanatory variables have with the volatility of the NG price, the results show that when using the proposed methodology, it is possible to distinguish a linear behavior in periods of low volatility and a nonlinear behavior in periods of high volatility. Thus, for periods of high volatility, the prediction is more complex, and therefore the best model to forecast volatility is the LSTM model.

Our work differs from the cited literature mainly in the analysis of variable influence for forecasting. Although works like Caporin et al. (2019), Lovcha and Perez-Laborda (2020), and Liu and Chen (2022) analyze variable influence, our present study does so daily, demonstrating that no variable consistently explains NG volatility over the whole forecast period. In other words, here we demonstrate the importance of understanding the influences in the out-of-sample period. Otherwise, an overfitting problem could lead to biased influence conclusions. This also depends on the model used, as works like Liu and Lee (2021) used purely econometric models, and in the work of Wang et al. (2020b), only machine learning approaches were used. Here we use both and select the best one each day, which reduces the bias of assuming the same characteristics for the entire out-of-sample period. These are the two main contributions of the paper and make it quite distinct from the existing studies presented in the literature review.

The proposed methodology also provides the ability to determine the effects of each variable over time—daily, weekly, monthly, or some other frequency—distinguishing periods with a more significant influence from one or more specific variables. The

flexibility of analyzing at different periods allows us to determine the effects of the variables analyzed. On the contrary, traditionally, if the study is only carried out for part of the period, it is impossible to discern the different temporary effects of the variables. On the other hand, the analysis focused on predictive capacity leaves aside overfitting problems, and the conclusions of the effects of the analyzed variables imply their generalization capacity.

## Conclusions

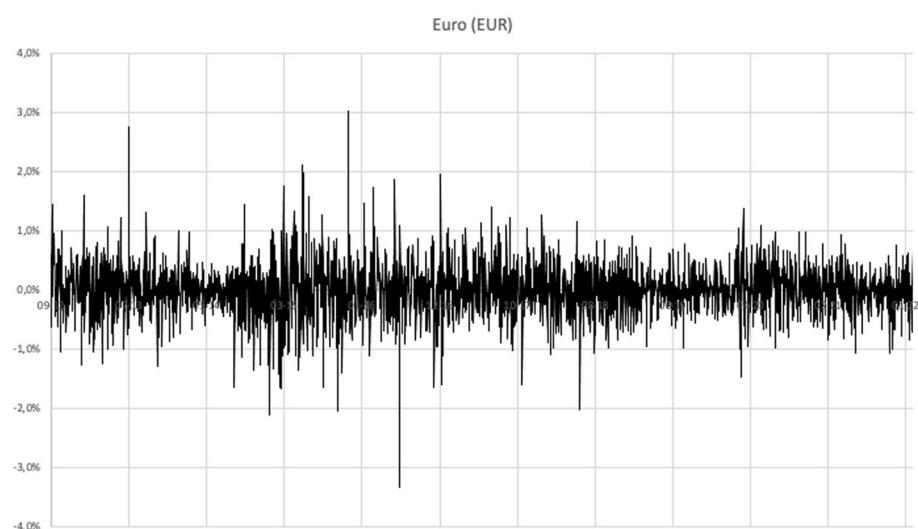
The proposed methodology made it possible to distinguish the effects of the main financial assets on the realized volatility of NG. By performing the day-to-day analysis, it was possible to have greater detail on the influence of the different financial assets on the realized volatility, unlike the usual way of finding the best model for the entire study period. This allows for better forecasting as we do not assume a particular model for all the out-of-sample periods. We see that in periods of higher volatility, a combination of linear (HAR) and nonlinear (LSTM) models provides the best forecast. On the contrary, in periods of lower volatility, the HAR model predominates as the best forecasting approach. This could indicate that, generally, the HAR model is more suited for this kind of problem. With the proposed methodology, it was possible to detect changes over time in the influences and influences according to the level of volatility of the period. We see that no variable can consistently explain NG volatility, regardless of the period analyzed. In the case of the realized volatility of NG, it was possible to conclude that in periods of low volatility, the HAR model with EURO provides the best forecast. Therefore, EURO is the variable that best explains the behavior of the realized volatility of NG for periods of low volatility in a linear manner. However, for periods of high volatility, there is an explanation for the realized volatility of NG lagged nonlinearly and the SPX linearly.

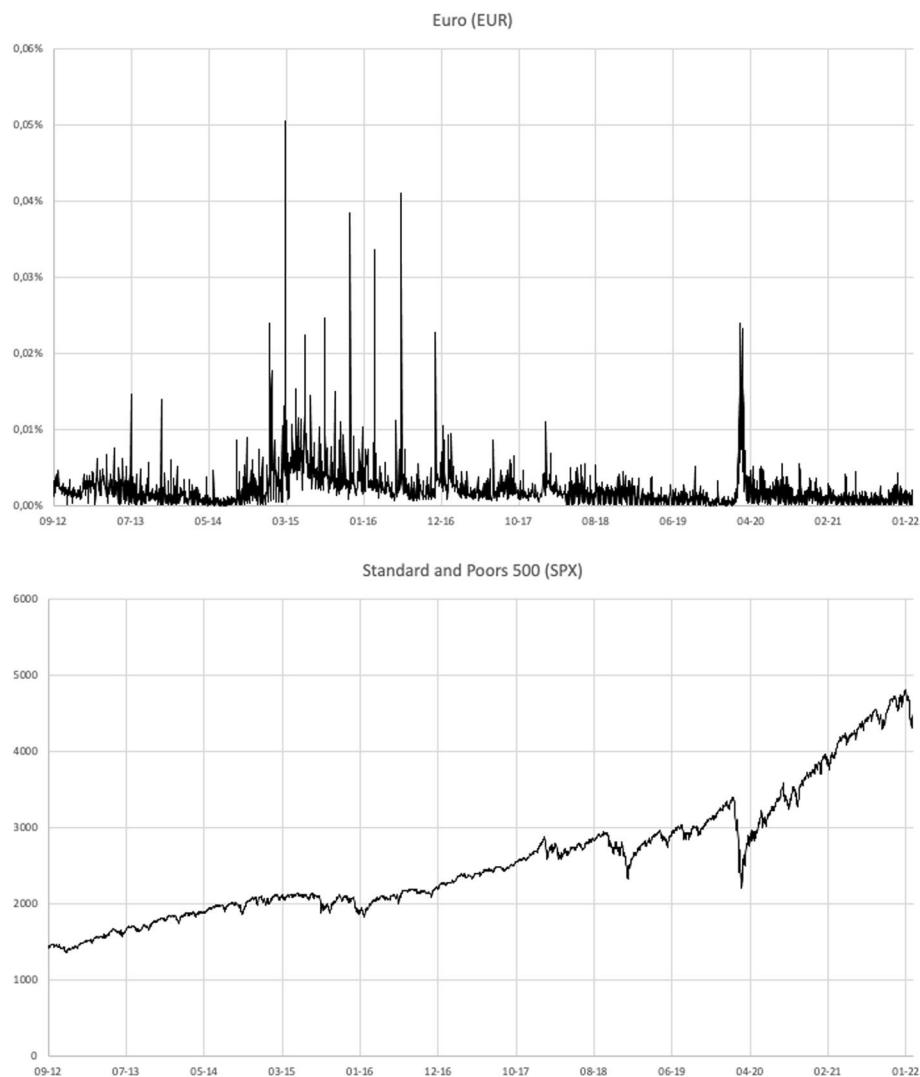
The traditional model is binary on the influence of the variables; that is, the variable influences volatility for the entire period, or it does not influence it at all. The traditional model does not allow the ability to identify subperiods or periods of characteristic volatility where some variables have an influence and other periods where they do not have an influence. The proposed methodology demonstrated that it could determine the temporal effects of the variables analyzed throughout the study period. This approach opens the opportunity to consider more complex settings, including more variables or models to better explain the behavior of NG prices. Moreover, this framework could be used in other problems where it is important to deeply understand the variable influence dynamics in a predictive way.

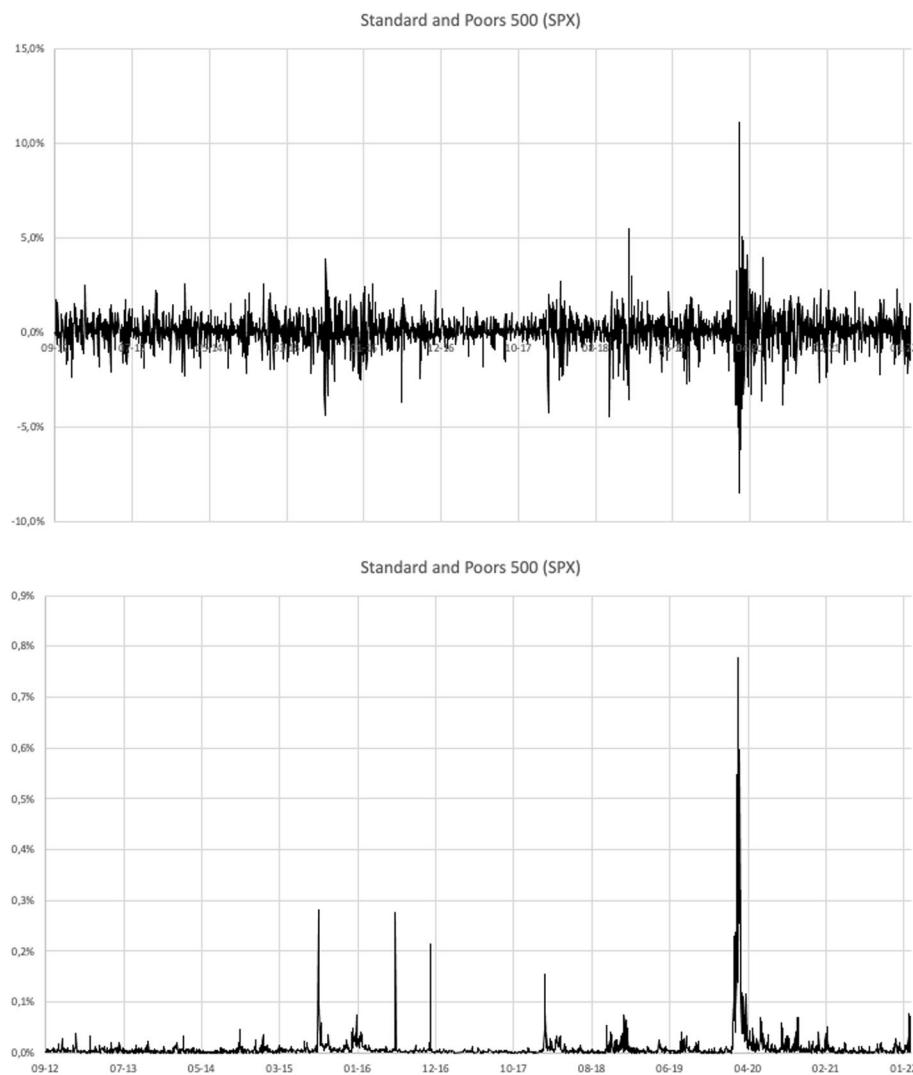
In future work, we propose raising awareness of the methodology by not only analyzing the best model for each frequency but also including a statistical test of superiority to determine the best models. Other future work could include characterizing and relating economic facts with the findings of results obtained through the proposed methodology. In addition, as a complement to the proposed methodology, we propose that studies be conducted that incorporate an analysis of the influence of variables over time with economic and political moments experienced globally.

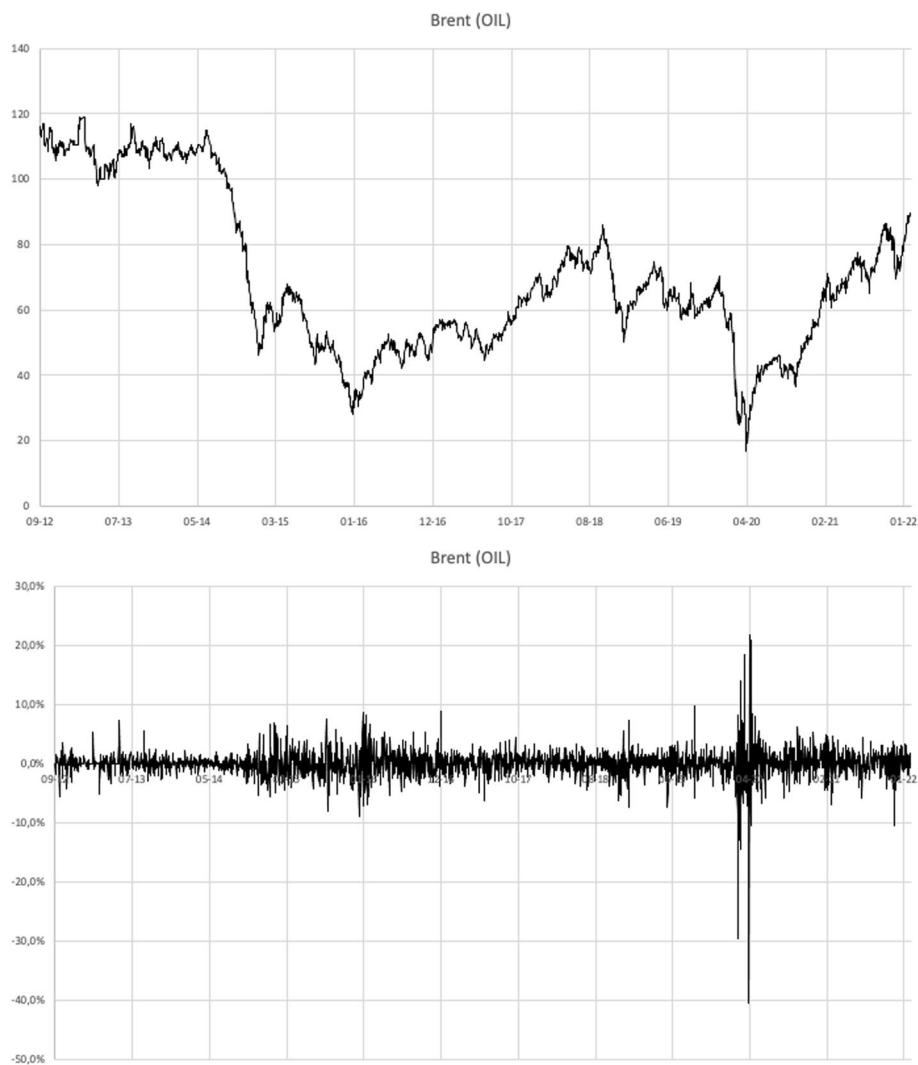
**Appendix 1. Descriptive statistics of price, return and variance of EURO, standard and Poor's 500, brent oil and gold**

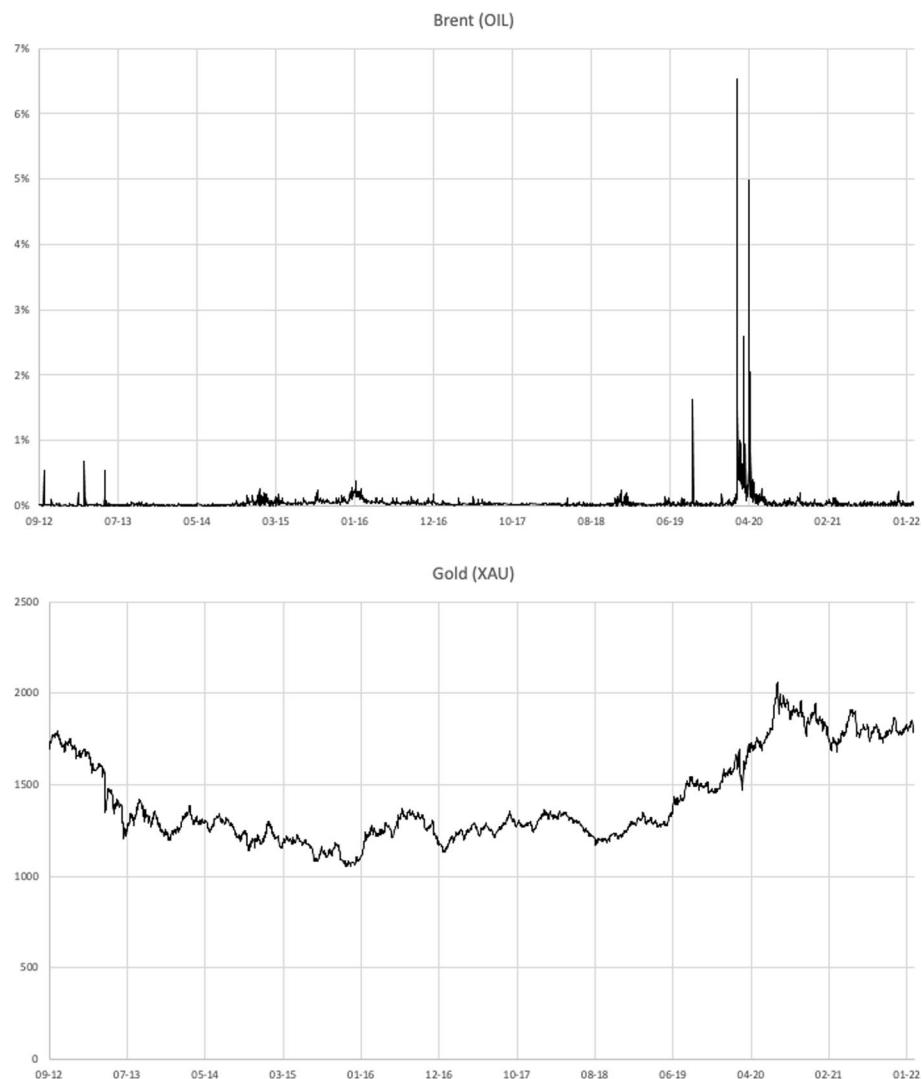
	EURO (EUR)			Standard and Poors 500 (SPX)			Brent (Oil)			Gold (XAU)		
	Price	Return (%)	Variance (%)				Price	Return (%)	Variance (%)			
				Price	Return (%)	Variance (%)				Price	Return (%)	Variance (%)
Mean	1.1862	-0.0033	0.0022	2602.4868	0.0425	0.0099	69.9809	-0.0064	0.0504	1421.1541	-0.0013	0.0076
St. Deviation	0.0912	0.4453	0.0028	829.5000	0.9197	0.0342	23.8278	2.3235	0.1987	243.8623	0.8605	0.0109
Median	1.1614	0.0036	0.0016	2442.9395	0.0689	0.0036	64.5530	0.0000	0.0235	1313.8030	0.0096	0.0050
Min	1.0406	-3.3421	0.0000	1353.3300	-8.4447	0.0000	16.9350	-40.3988	0.0000	1052.2400	-6.9245	0.0001
Max	1.3933	3.0197	0.0505	4797.6360	11.0757	0.7768	119.0000	21.7756	6.5258	2056.4180	5.5366	0.2210

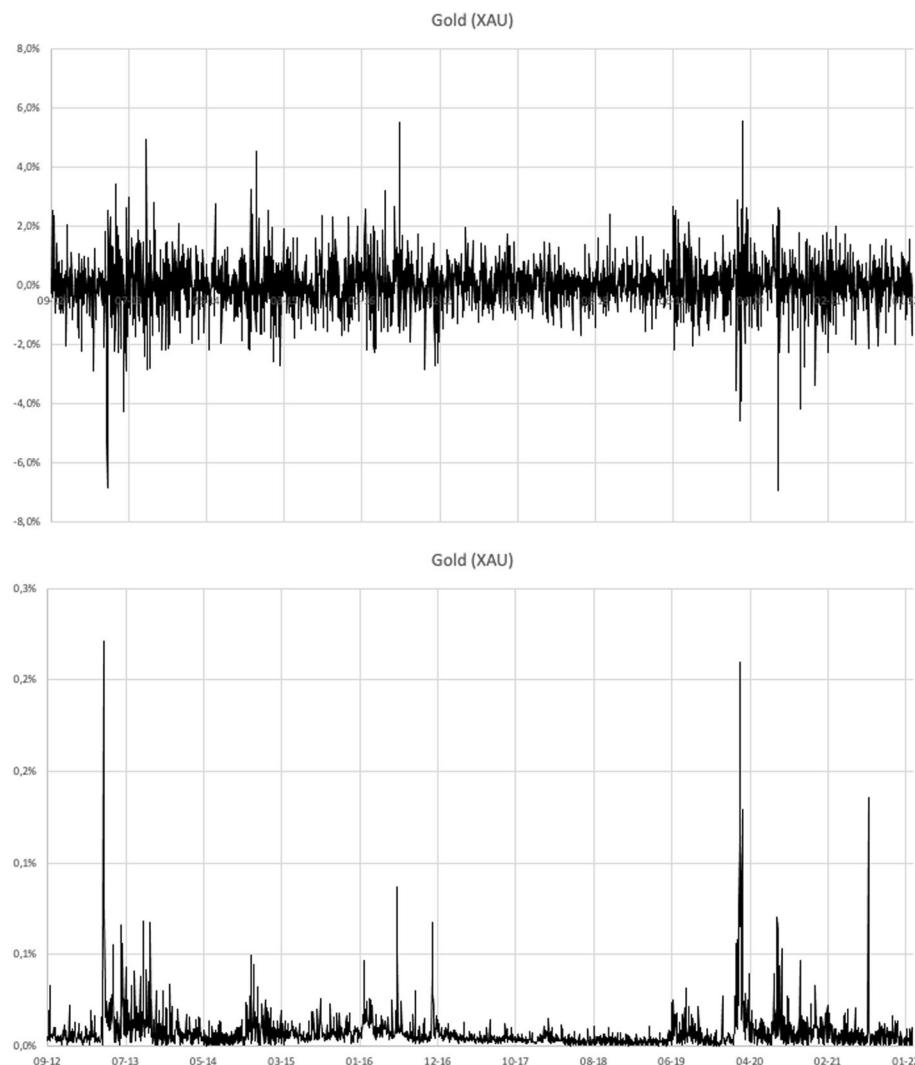
**Appendix 2 Evolution of price, return and variance of EURO, standard and poor's 500, brent oil and gold**











#### Abbreviations

NG	Natural gas
IEA	International Energy Agency
LNG	Liquefied natural gas
CNG	Natural gas compressed
XAU	Gold price
BRENT	Crude oil price
SPX	Standard and Poor's 500
EURO	Euro U.S. dollar exchange rate
HAR	Heterogeneous autoregressive
NN	Neural networks
LSTM	Long-short term memory
RV	Realized volatility
ENet	Elastic net
PCA	Principal component analysis
INE	Chine's Shanghai International Energy Exchange
MS	Markov switching
MIDAS	Mixed data sampling
NYMEX	New York Mercantile
SVM	Support vector machine
AANN	Auto-associative neural network

SDAE	Stacked denoising autoencoders
GRU	Gate recurrent units based NN
t	Time index (current)
f	Time index (previous days)
d	Daily frequency
w	Weekly frequency
m	Monthly frequency
$RV_t^{(\ell)}$	Realized volatility computed by the mean of the last $f$ days previous $t$
$\beta_*$	Linear coefficient depending on frequency
RNN	Recurrent neural network
$X_t$	Array of variables indexed by time
$Y_t$	Variable of interest indexed by time
$\sigma$	Sigmoid function
tanh	Hyperbolic tangent function
$W_*$	Weights matrices of each LSTM gate
$b_f$	Bias of each LSTM gate
$C_t$	Cell state
$f_t$	Forget gate
$i_t$	Input/update gate
$\tilde{C}_t$	Cell input activation
$o_t$	Output gate
Q	Number of layers, neurons and lags
L	Length of the window to train
Obs.	Days in which the LSTM was the best model with a particular configuration
RVQ	Realized volatility quartile
AR	Lagged natural gas realized volatility
MSE	Mean square error

#### Acknowledgements

Not applicable.

#### Author contributions

WK: Conceptualization, Methodology, Software, Data curation, Writing- Original, Validation.

#### Funding

No funding was received.

#### Availability of data and materials

The data that support the findings of this study are available from dukascopy, but restrictions apply to the availability of these data.

#### Declarations

##### Consent to publication

The authors agreed with the content and gave explicit consent to submit the manuscript.

##### Competing interests

The author has no relevant financial or non-financial interests to disclose.

Received: 8 September 2022 Accepted: 6 December 2023

Published online: 29 January 2024

#### References

- Alfeus M, Nikitopoulos CS (2022) Forecasting volatility in commodity markets with long-memory models. *J Commodity Markets* 28:100248
- Andersen TG, Bollerslev T (1998) Answering the skeptics: yes, standard volatility models do provide accurate forecasts. *Int Econ Rev* 39:885–905
- Andersen TG, Bollerslev T, Christoffersen PF, Diebold FX (2005) Volatility forecasting (No. w11188). National Bureau of Economic Research, Cambridge
- Andersen TG, Bollerslev T, Diebold FX (2007) Roughing it up: including jump components in the measurement, modeling, and forecasting of return volatility. *Rev Econ Stat* 89(4):701–720
- Armstrong JS (ed) (2001) Principles of forecasting: a handbook for researchers and practitioners. Kluwer Academic, Boston
- Asadi M, Roubaud D, Tiwari AK (2022) Volatility spillovers amid crude oil, natural gas, coal, stock, and currency markets in the US and China based on time and frequency domain connectedness. *Energy Econ* 109:105961
- Audrino F, Huang C, Okhrin O (2019) Flexible HAR model for realized volatility. *Stud Nonlinear Dyn Econom* 23(3):20170080

- Baruník J, Křehlík T (2016) Combining high frequency data with non-linear models for forecasting energy market volatility. *Expert Syst Appl* 55:222–242
- Bekaert G, Hoerova M (2014) The VIX, the variance premium and stock market volatility. *J Econom* 183(2):181–192
- Bekiros SD, Georgoutsos DA (2008) Direction-of-change forecasting using a volatility-based recurrent neural network. *J Forecast* 27(5):407–417
- Bengio Y, Simard P, Frasconi P (1994) Learning long-term dependencies with gradient descent is difficult. *IEEE Trans Neural Netw* 5(2):157–166
- Bergsli LØ, Lind AF, Molnár P, Polaski M (2022) Forecasting volatility of Bitcoin. *Res Int Bus Financ* 59:101540
- Bollerslev T (1986) Generalized autoregressive conditional heteroskedasticity. *J Econom* 31(3):307–327
- Bouri E, Lucey B, Saeed T, Vo XV (2021) The realized volatility of commodity futures: interconnectedness and determinants. *Int Rev Econ Financ* 73:139–151
- Bucci A (2020) Realized volatility forecasting with neural networks. *J Financ Econom* 18(3):502–531
- Busch T, Christensen BJ, Nielsen MØ (2011) The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets. *J Econom* 160:48–57
- Caporin M, Chang CL, McAleer M (2019) Are the S&P 500 index and crude oil, natural gas and ethanol futures related for intra-day data? *Int Rev Econ Financ* 59:50–70
- Celik S, Ergin H (2014) Volatility forecasting using high frequency data: evidence from stock markets. *Econ Model* 36:176–190
- Čeperić E, Žiković S, Čeperić V (2017) Short-term forecasting of natural gas prices using machine learning and feature selection algorithms. *Energy* 140:893–900
- Corsi F (2009) A simple approximate long-memory model of realized volatility. *J Financ Econom* 7(2):174–196
- Corsi F, Renò R (2012) Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *J Bus Econ Stat* 30(3):368–380
- Degiannakis S, Filis G (2017) Forecasting oil price realized volatility using information channels from other asset classes. *J Int Money Financ* 76:28–49
- Donaldson RG, Kamstra M (1996) Forecast combining with neural networks. *J Forecast* 15(1):49–61
- Donaldson RG, Kamstra M (1997) An artificial neural network-GARCH model for international stock return volatility. *J Empir Financ* 4(1):17–46
- Engle RF (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econom J Econom Soc* 50:987–1007
- Fernandes M, Medeiros MC, Scharth M (2014) Modeling and predicting the CBOE market volatility index. *J Bank Finance* 40:1–10
- Gao S, Hou C, Nguyen BH (2021) Forecasting natural gas prices using highly flexible time-varying parameter models. *Econ Model* 105:105652
- Garzón A, Hierro L (2022) Inflation, oil prices and exchange rates. The Euro's dampening effect. *J Policy Model* 44(1):130–146
- Greff K, Srivastava RK, Koutrník J, Steunebrink BR, Schmidhuber J (2016) LSTM: a search space odyssey. *IEEE Trans Neural Netw Learn Syst* 28(10):2222–2232
- Hamid SA, Iqbal Z (2004) Using neural networks for forecasting volatility of S&P 500 Index futures prices. *J Bus Res* 57(10):1116–1125
- Herrera AM, Hu L, Pastor D (2018) Forecasting crude oil price volatility. *Int J Forecast* 34(4):622–635
- Herrera GP, Constantino M, Tabak BM, Pistori H, Su JJ, Naranpanawa A (2019) Long-term forecast of energy commodities price using machine learning. *Energy* 179:214–221
- Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780
- Hsieh TJ, Hsiao HF, Yeh WC (2011) Forecasting stock markets using wavelet transforms and recurrent neural networks: an integrated system based on artificial bee colony algorithm. *Appl Soft Comput* 11(2):2510–2525
- Jiang Q, Tang C, Chen C, Wang X, Huang Q (2019) Stock price forecast based on LSTM neural network. In: Proceedings of the twelfth international conference on management science and engineering management (pp. 393–408). Springer International Publishing.
- Jin D, He M, Xing L, Zhang Y (2022) Forecasting China's crude oil futures volatility: how to dig out the information of other energy futures volatilities? *Resour Policy* 78:102852
- Kristjanpoller W, Minutolo MC (2016) Forecasting volatility of oil price using an artificial neural network-GARCH model. *Expert Syst Appl* 65:233–241
- Kristjanpoller W, Fadic A, Minutolo MC (2014) Volatility forecast using hybrid neural network models. *Expert Syst Appl* 41(5):2437–2442
- Kumar S, Choudhary S, Singh G, Singhal S (2021) Crude oil, gold, natural gas, exchange rate and Indian stock market: evidence from the asymmetric nonlinear ARDL model. *Resour Policy* 73:102194
- 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 H, Zhou D, Hu J, Li J, Su M, Guo L (2023) Forecasting the realized volatility of energy stock market: a multimodel comparison. *N Am J Econ Financ* 66:101895
- Liang C, Ma F, Wang L, Zeng Q (2021) The information content of uncertainty indices for natural gas futures volatility forecasting. *J Forecast* 40(7):1310–1324
- Lin B, Li J (2015) The spillover effects across natural gas and oil markets: based on the VEC–MGARCH framework. *Appl Energy* 155:229–241
- Liu W, Chen X (2022) Natural resources commodity prices volatility and economic uncertainty: evaluating the role of oil and gas rents in COVID-19. *Resour Policy* 76:102581

- Liu M, Lee CC (2021) Capturing the dynamics of the China crude oil futures: Markov switching, co-movement, and volatility forecasting. *Energy Econ* 103:105622
- Lovcha Y, Perez-Laborda A (2020) Dynamic frequency connectedness between oil and natural gas volatilities. *Econ Model* 84:181–189
- Luo J, Demirer R, Gupta R, Ji Q (2022a) Forecasting oil and gold volatilities with sentiment indicators under structural breaks. *Energy Econ* 105:105751
- Luo J, Klein T, Ji Q, Hou C (2022b) Forecasting realized volatility of agricultural commodity futures with infinite hidden Markov HAR models. *Int J Forecast* 38(1):51–73
- Lyócsa Š, Molnár P (2018) Exploiting dependence: day-ahead volatility forecasting for crude oil and natural gas exchange-traded funds. *Energy* 155:462–473
- Lyócsa Š, Molnár P, Výrost T (2021) Stock market volatility forecasting: do we need high-frequency data? *Int J Forecast.* <https://doi.org/10.1016/j.ijforecast.2020.12.001>
- Maknickiene N, Lapinskaite I, Maknickas A (2018) Application of ensemble of recurrent neural networks for forecasting of stock market sentiments. *Equilib Q J Econ Econ Policy* 13(1):7–27
- Monfared SA, Enke D (2014) Volatility forecasting using a hybrid GJR-GARCH neural network model. *Proced Comput Sci* 36:246–253
- Mücher C (2022) Artificial neural network based non-linear transformation of high-frequency returns for volatility forecasting. *Front Artif Intell* 4:787534
- Patton AJ (2011) Data-based ranking of realised volatility estimators. *J Econom* 161(2):284–303
- Petneházi G, Gáll J (2019) Exploring the predictability of range-based volatility estimators using recurrent neural networks. *Intell Syst Account Financ Manag* 26(3):109–116
- Qiu Y, Zhang X, Xie T, Zhao S (2019) Versatile HAR model for realized volatility: a least square model averaging perspective. *J Manag Sci Eng* 4(1):55–73
- Qu H, Li G (2023) Multi-perspective investor attention and oil futures volatility forecasting. *Energy Econ* 119:106531
- Rodikov G, Antulov-Fantulin N (2022) Can LSTM outperform volatility-econometric models?. arXiv preprint [arXiv:2202.11581](https://arxiv.org/abs/2202.11581).
- Santos DG, Ziegelmann FA (2014) Volatility forecasting via MIDAS, HAR and their combination: an empirical comparative study for IBOVESPA. *J Forecast* 33(4):284–299
- Seo SW, Kim JS (2015) The information content of option-implied information for volatility forecasting with investor sentiment. *J Bank Financ* 50:106–120
- Seo M, Kim G (2020) Hybrid forecasting models based on the neural networks for the volatility of bitcoin. *Appl Sci* 10(14):4768
- Song Y, He M, Wang Y, Zhang Y (2022) Forecasting crude oil market volatility: a newspaper-based predictor regarding petroleum market volatility. *Resour Policy* 79:103093
- Srivastava M, Rao A, Parihar JS, Chavriya S, Singh S (2023) What do the AI methods tell us about predicting price volatility of key natural resources: evidence from hyperparameter tuning. *Resour Policy* 80:103249
- Stoupous N, Kiohos A (2021) Energy commodities and advanced stock markets: a post-crisis approach. *Resour Policy* 70:101887
- Tino P, Schittenkopf C, Dorffner G (2001) Financial volatility trading using recurrent neural networks. *IEEE Trans Neural Netw* 12(4):865–874
- Todorova N, Worthington A, Souček M (2014) Realized volatility spillovers in the non-ferrous metal futures market. *Resour Policy* 39:21–31
- Vidic RD, Brantley SL, Vandebossche JM, Yoxtheimer D, Abad JD (2013) Impact of shale gas development on regional water quality. *Science* 340(6134):1235009
- Vortelinos DI (2017) Forecasting realized volatility: HAR against principal components combining, neural networks and GARCH. *Res Int Bus Financ* 39:824–839
- Wang J, Feng L, Tang X, Bentley Y, Höök M (2017) The implications of fossil fuel supply constraints on climate change projections: a supply-side analysis. *Futures* 86:58–72
- Wang M, Zhao L, Du R, Wang C, Chen L, Tian L, Stanley HE (2018) A novel hybrid method of forecasting crude oil prices using complex network science and artificial intelligence algorithms. *Appl Energy* 220:480–495
- Wang J, Huang Y, Ma F, Chevallier J (2020a) Does high-frequency crude oil futures data contain useful information for predicting volatility in the US stock market? New evidence. *Energy Econ* 91:104897
- Wang J, Lei C, Guo M (2020b) Daily natural gas price forecasting by a weighted hybrid data-driven model. *J Petrol Sci Eng* 192:107240
- Yuyan G, di H, Yan M, Hongmin Z (2023) Realised volatility prediction of high-frequency data with jumps based on machine learning. *Connect Sci* 35(1):2210265

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.