

Article

Forecasting Volatility of Energy Commodities: Comparison of GARCH Models with Support Vector Regression

Marcin Faldziński^{1,2,*}, Piotr Fiszeder^{1,2,*}  and Witold Orzeszko³

¹ Department of Econometrics and Statistics, Faculty of Economic Sciences and Management, Nicolaus Copernicus University in Torun, ul. Gagarina 13a, 87-100 Torun, Poland; marcin.faldzinski@umk.pl

² Faculty of Finance and Accounting, Prague University of Economics and Business, W. Churchill Sq. 1938/4, Žižkov, 130 67 Prague, Czech Republic

³ Department of Applied Informatics and Mathematics in Economics, Faculty of Economic Sciences and Management, Nicolaus Copernicus University in Torun, ul. Gagarina 13a, 87-100 Torun, Poland; witold.orzeszko@umk.pl

* Correspondence: piotr.fiszeder@umk.pl

Abstract: We compare the forecasting performance of the generalized autoregressive conditional heteroscedasticity (GARCH) -type models with support vector regression (SVR) for futures contracts of selected energy commodities: Crude oil, natural gas, heating oil, gasoil and gasoline. The GARCH models are commonly used in volatility analysis, while SVR is one of machine learning methods, which have gained attention and interest in recent years. We show that the accuracy of volatility forecasts depends substantially on the applied proxy of volatility. Our study confirms that SVR with properly determined hyperparameters can lead to lower forecasting errors than the GARCH models when the squared daily return is used as the proxy of volatility in an evaluation. Meanwhile, if we apply the Parkinson estimator which is a more accurate approximation of volatility, the results usually favor the GARCH models. Moreover, it is difficult to choose the best model among the GARCH models for all analyzed commodities, however, forecasts based on the asymmetric GARCH models are often the most accurate. While, in the class of the SVR models, the results indicate the forecasting superiority of the SVR model with the linear kernel and 15 lags, which has the lowest mean square error (MSE) and mean absolute error (MAE) among the SVR models in 92% cases.

Keywords: energy commodities; futures contracts; volatility; forecasting; GARCH models; support vector regression; machine learning



Citation: Faldziński, M.; Fiszeder, P.; Orzeszko, W. Forecasting Volatility of Energy Commodities: Comparison of GARCH Models with Support Vector Regression. *Energies* **2021**, *14*, 6. <https://dx.doi.org/10.3390/en14010006>

Received: 29 October 2020

Accepted: 18 December 2020

Published: 22 December 2020

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Energy risk has always been one of the major risk factors for most firms involved in key industrial sectors in both developed and developing countries. Risk management of energy commodities is a crucial issue for majority industrial firms, as it can seriously affect its competitiveness, viability and future profitability. Global economic developments, emerging technological advances and economic, geopolitical and environmental events have caused a significant increase in volatility of energy commodities prices in the last 20 years (cf. [1]). For these reasons, the ability to predict volatility of energy commodities has been gaining more and more importance.

Commodity market participants, like the large corporate producers, manufacturers, large energy consuming firms, alongside a number of investment banks, specialist funds, investors and traders, need volatility forecasts for effective investment, hedging risk and arbitrage strategies. There are many methods for volatility forecasting but the most popular in the literature are the generalized autoregressive conditional heteroscedasticity (GARCH) models. In particular, they have already been applied in plenty of studies for energy commodities, including: Crude oil (e.g., [2–11]), natural gas (e.g., [4,6,12]), heating oil (e.g., [2,3,6]), gasoline (e.g., [3,6]) and gasoil (e.g., [2]).

In order to forecast volatility, machine learning techniques can also be applied. One such technique, which has been gaining high popularity in recent years, is support vector regression (SVR). In particular, SVR and SVR-based hybrid models have been applied in many studies to forecast prices of energy commodities, including: crude oil (e.g., [13–18]) and natural gas (e.g., [19,20]). However, they have been used only by Zhang and Zhang [8] to forecast volatility of energy commodities, specifically crude oil.

In this paper, we compare the volatility forecasting performance of the GARCH-type models with support vector regression for futures contracts of selected energy commodities, namely, crude oil, gasoil, gasoline, heating oil and natural gas, quoted at New York Mercantile Exchange and International Exchange. It should be noted that, nowadays, derivative markets are very important research areas since they play an essential role in trading energy commodities, and turnover on such markets has become significantly higher in comparison to spot markets.

This study has two main contributions. Firstly, it is the first comparison of the GARCH-type models and SVR for futures contracts for energy commodities. As we mentioned before, according to our knowledge, SVR has been used for volatility forecasting of energy commodities only by Zhang, and Zhang [8] who applied the least-squares version of a SVR model—least squares support vector machines (LSSVM) [21] for the West Texas Intermediate (WTI) and Brent crude oil spot prices. However, they used this model only as a part of the hybrid forecasting method—to forecast the residual series from the exponential generalized autoregressive conditional heteroscedasticity (EGARCH) model.

Secondly, we show that the accuracy of volatility forecasts depends substantially on the applied proxy of volatility. It has been shown in the literature that volatility forecasts of financial assets based on SVR models can be more precise than those from GARCH-type models (e.g., [22–33]). However, in most of these papers, the squared daily returns have been taken as the ex-post measure of volatility. Our study confirms that SVR with properly determined hyperparameters can lead to lower forecasting errors than the GARCH models when the squared daily return is used as the proxy of volatility in an evaluation. Meanwhile, if we apply the Parkinson estimator, which is a more accurate approximation of volatility, the results are different since they usually favor the GARCH models over SVR.

The rest of the paper is organized in the following way. Section 2 provides a description of applied models and methods. In Section 3 we introduce and describe data, explain the forecasting procedure and present the results of the research for energy commodities. Finally, in the last section, we give our concluding remarks.

2. Description of Models

2.1. GARCH-Type Models

The GARCH models are a standard tool applied in volatility research. The primary model is the standard GARCH model introduced by Bollerslev [34], which describes time-varying variance. Let us assume that ε_t is the innovation process which can be presented as:

$$\varepsilon_t | \psi_{t-1} \sim N(0, h_t), \quad (1)$$

where h_t is the conditional variance, ψ_{t-1} is the set of all information available at time $t - 1$, N is the conditional normal distribution. Then the GARCH(p, q) model (denoted as GARCH-n) is given as,

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}, \quad (2)$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$ for $i = 1, 2, \dots, q$; $j = 1, 2, \dots, p$, however Nelson, Cao [35] gave weaker conditions for non-negativity of the conditional variance.

Instead of the conditional normal distribution in the Equation (1) the Student's t -distribution can be applied (the model denoted as GARCH-t) in order to better describe fatter tails and leptokurtosis of unconditional distributions of many empirical financial time series [36].

The conditional variance function of the standard GARCH model is symmetric in the lagged values of ε_t . Such function may be inappropriate for modelling the volatility of returns because it cannot represent a phenomenon, known as the leverage effect, i.e., the negative correlation between volatility and past returns. The first model describing asymmetric responses of the conditional variance to positive and negative errors is the exponential GARCH (EGARCH) model proposed by Nelson [37]. The EGARCH(p, q) model is specified as,

$$\ln h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \{ \theta z_{t-i} + \gamma [|z_{t-i}| - E(|z_{t-i}|)] \} + \sum_{j=1}^p \beta_j \ln h_{t-j}, \quad (3)$$

where $\alpha_1 \equiv 1$, $z_t = \varepsilon_t / h_t^{1/2}$, E is the expected value. The logarithmic form of the conditional variance means that it is not necessary to introduce any restrictions on parameters to ensure the positivity of the conditional variance.

The second most popular asymmetric GARCH model is the GJR model introduced by Glosten et al. [38]. The GJR(p, q) model is given as,

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^q \gamma_i I_{t-i} \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}, \quad (4)$$

where I_{t-i} is a dummy variable which satisfies $I_{t-i} = 1$ when $\varepsilon_{t-i} \leq 0$ and $I_{t-i} = 0$ when $\varepsilon_{t-i} > 0$. To ensure the positivity of the conditional variance parameters should meet the following requirements: $\alpha_0 > 0$, $\alpha_i \geq 0$, $\alpha_i + \gamma_i \geq 0$ for $i = 1, 2, \dots, q$, $\beta_j \geq 0$ for $j = 1, 2, \dots, p$.

Another model which captures asymmetry in volatility is the asymmetric power autoregressive conditional heteroscedasticity (APARCH) model proposed by Ding et al. [39]. The APARCH(p, q) model is defined as,

$$(h_t^{1/2})^d = \alpha_0 + \sum_{i=1}^q \alpha_i (|\varepsilon_{t-i}| + \gamma_i \varepsilon_{t-i})^d + \sum_{j=1}^p \beta_j (h_{t-j}^{1/2})^d, \quad (5)$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$ for $i = 1, 2, \dots, q$; $j = 1, 2, \dots, p$. The exponent d allows for more flexibility in the description of volatility. The class of APARCH models includes several other GARCH extensions as special cases, like the GJR model, the threshold autoregressive conditional heteroscedasticity (TARCH) model of Zakoian [40], the Taylor ([41])/Schwert ([42]) GARCH model, the nonlinear autoregressive conditional heteroscedasticity (NARCH) of Higgins and Bera [43], the Log-ARCH of Geweke [44] and Pentula [45].

In many empirical studies, the sum of parameters estimates (except α_0) in the standard GARCH model is close to 1 which makes variance highly persistent. That is why Engle and Bollerslev [46] proposed the integrated GARCH (IGARCH) model in the form of Equation (2), where $\alpha_1 + \dots + \alpha_q + \beta_1 + \dots + \beta_p = 1$. A shock to the conditional variance in the IGARCH model is persistent in the sense that it remains significant for forecasts of all horizons.

The last considered parameterization is the GARCH-in-mean (GARCH-M) model introduced by Engle et al. [47]. The GARCH-M(p, q) model is described as,

$$r_t = \gamma_0 + \delta h_t + \varepsilon_t, \quad (6)$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}, \quad (7)$$

where $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$ for $i = 1, 2, \dots, q$; $j = 1, 2, \dots, p$. The GARCH-M model is able to describe the fundamental trade-off relationship between return and risk.

The parameters of the above GARCH-type models can be estimated by maximum likelihood or quasi-maximum likelihood methods. The log-likelihood function is described as,

$$L(\zeta) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^n \left(\ln h_t + \frac{\varepsilon_t^2}{h_t} \right), \quad (8)$$

where ζ is a vector containing unknown parameters of the model, n is the number of observations used in estimation.

2.2. SVR Model

Let y be the dependent variable and x the vector of regressors. Based on a training data set $\{(x_t, y_t)\}_{t=1, \dots, n}$, we want to estimate the regression function that has, at most, ε deviation from the outputs y_t and that, at the same time, is as flat as possible [48]. The idea of SVR is to map the vectors x onto a high-dimensional feature space using some fixed (nonlinear) transformation and then to estimate the linear model,

$$f(x) = \sum_{i=1}^d \omega_i \varphi_i(x) + b, \quad (9)$$

where d is the dimension of the space, $\varphi_i(x)$ denote transformations, ω_i are the coefficients and b is the bias term [49,50]. To assess the estimated model Vapnik [51] proposed the ε -insensitive loss function,

$$L_\varepsilon(y, f(x)) = \begin{cases} 0, & |y - f(x)| \leq \varepsilon, \\ |y - f(x)| - \varepsilon, & \text{otherwise,} \end{cases} \quad (10)$$

which does not penalize errors below ε . To estimate the coefficients of the SVR model (9) the ε -insensitive loss function is used, however at the same time, the postulate of the model complexity reduction is taken into account by minimizing the expression $\|\omega\|^2 = \omega^T \omega$, where $\omega = (\omega_1, \omega_2, \dots, \omega_d)^T$. In practice, it is not always possible to approximate all data of the training set with an error below ε (cf. [52]). In order to allow errors to be greater, the SVR model incorporates nonnegative slack variables ζ_t and ζ_t^* representing the upper and lower constraints, s.t.,

$$y_t - f(x_t) \leq \varepsilon + \zeta_t^*, \quad (11)$$

$$f(x_t) - y_t \leq \varepsilon + \zeta_t, \quad (12)$$

for all $t = 1, 2, \dots, n$. Finally, the regression function $f(x)$ is obtained as the minimum of the functional,

$$\Phi(\omega, \xi) = \frac{1}{2} \|\omega\|^2 + C \sum_{t=1}^n (\zeta_t + \zeta_t^*), \quad (13)$$

where C is a pre-specified positive value. The first term of the Functional (13) is used to penalize large weights and to maintain regression function flatness, whereas the second term penalizes training errors by using the ε -insensitive loss function [53]. Where, C is the hyperparameter to trade off these two terms. It controls the penalty imposed on observations that lie outside the ε -margin and, in consequence, helps to prevent overfitting. Both the ε and C hyperparameters must be determined by the user.

The optimization problem described above can be transformed into a corresponding dual problem using the Karush-Kuhn-Tucker conditions, with the solution,

$$f(x) = \sum_{t=1}^{n_{SV}} (\alpha_t - \alpha_t^*) K(x_t, x), \text{ s.t. } 0 \leq \alpha_t \leq C, 0 \leq \alpha_t^* \leq C, \quad (14)$$

where α_t and α_t^* are the Lagrange multipliers, n_{SV} is the number of support vectors and K is the kernel function of the form,

$$K(\mathbf{x}_t, \mathbf{x}) = \sum_{i=1}^d \varphi_i(\mathbf{x}) \varphi_i(\mathbf{x}_t). \quad (15)$$

Any function satisfying the Mercer's condition ([51]) can be used as the kernel. In practice, the most commonly used kernel functions include (cf. [54]):

- Linear (dot product): $K(\mathbf{x}_t, \mathbf{x}) = \mathbf{x}_t^T \mathbf{x}$,
- Radial basis function (RBF): $K(\mathbf{x}_t, \mathbf{x}) = \exp(-\gamma \|\mathbf{x}_t - \mathbf{x}\|^2)$,
- Polynomial: $K(\mathbf{x}_t, \mathbf{x}) = (1 + \mathbf{x}_t^T \mathbf{x})^p$; $p = 2, 3, \dots$

2.3. Ex-Post Volatility Measures

Volatility is not directly observable, even ex-post, therefore it has to be estimated. A popular proxy of the daily variance is the squared daily return,

$$\sigma_{sqr,t}^2 = r_t^2, \quad (16)$$

where r_t is the daily logarithmic return at day t . This proxy can be identified as the classical variance estimator.

Andersen and Bollerslev [55] showed that although the squared daily return is an unbiased estimator of the variance of return, it is also extremely noisy. A significantly more accurate measure of volatility is the realized variance (RV) calculated from intraday prices,

$$RV_t = \sum_{k=1}^K r_{k,t}^2, \quad (17)$$

where $r_{k,t}$ is the intraday return (e.g., the 5-min return), K is the number of intraday observations during a day. The realized variance is a significantly more efficient estimator of variance than the daily squared return, but high-frequency data are not commonly available and are significantly more expensive in comparison to daily data.

An alternative way is to calculate range-based variance estimators, based on daily opening, low and high prices. In practice, these values are easily available alongside with daily closing prices. The range-based estimators have already been used as the proxy of volatility in many studies (e.g., [56–59]). Furthermore, many volatility models have been proposed based on these estimators (e.g., [60–64]).

The Parkinson [65] estimator, the simplest of this class, is given as,

$$\sigma_{P,t}^2 = [\ln(H_t/L_t)]^2 / (4 \ln 2), \quad (18)$$

where H_t and L_t are high and low prices over a day t . This estimator assumes a zero drift process and is more than 4.9 times efficient than the classical variance estimator based on closing prices [65]. It has been shown that the accuracy of the Parkinson estimator is similar to the accuracy of the realized variance calculated from six observations during the day (see [59,66]).

Two other most popular range-based variance estimators are Garman-Klass [67] and Rogers-Satchell [68] estimators. The former can be presented as,

$$\sigma_{GK,t}^2 = 0.5[\ln(H_t/L_t)]^2 - (2 \ln 2 - 1)[\ln(C_t/O_t)]^2, \quad (19)$$

where C_t and O_t are closing and opening prices at day t , respectively. The Garman-Klass estimator assumes a zero drift process and is more than 7.4 times efficient than the classical variance estimator based on closing prices [67].

The Rogers-Satchell estimator is defined as:

$$\sigma_{RS,t}^2 = \ln(H_t/O_t) \ln(H_t/C_t) + \ln(L_t/O_t) \ln(L_t/C_t). \quad (20)$$

This estimator is independent of the drift. For a zero drift it is more than 6.0 times efficient than the classical variance estimator based on closing prices [68].

The realized variance is a more accurate proxy of volatility than the range-based estimators provided that intraday data are of good quality and the analyzed market is liquid (see [69]). However, the application of very high frequency data suffers from a large computational burden. Moreover, the range-based estimators are more robust than the realized variance to some microstructure effects like the bid-ask spread (e.g., [69,70]).

3. Forecasting Volatility of Energy Commodities

3.1. Data

In our study we investigate futures contracts of the selected energy commodities: WTI crude oil, gasoil, gasoline, heating oil and natural gas. Gasoil is quoted at the International Exchange (ICE) and the rest of considered contracts are listed at the New York Mercantile Exchange (NYMEX).

We analyze the data for the five-year period from 2 January 2015 to 31 December 2019. We apply two proxies of ex-post volatility in the forecast evaluation: the daily squared return (Equation (16)) and the Parkinson estimator (Equation (18)). In order to avoid the noise induced by measuring the overnight volatility we analyze open-to-close percentage returns $r_t = 100 \ln(C_t/O_t)$ instead of close-to-close returns. Additionally, this concept allows for better comparability with the Parkinson estimator which, by definition, measures volatility only during the exchange session. Since we analyze percentage returns, we compute the Parkinson estimator multiplied by 100^2 . Daily closing prices and open-to-close returns for the all analyzed future contracts are presented in Figure 1.

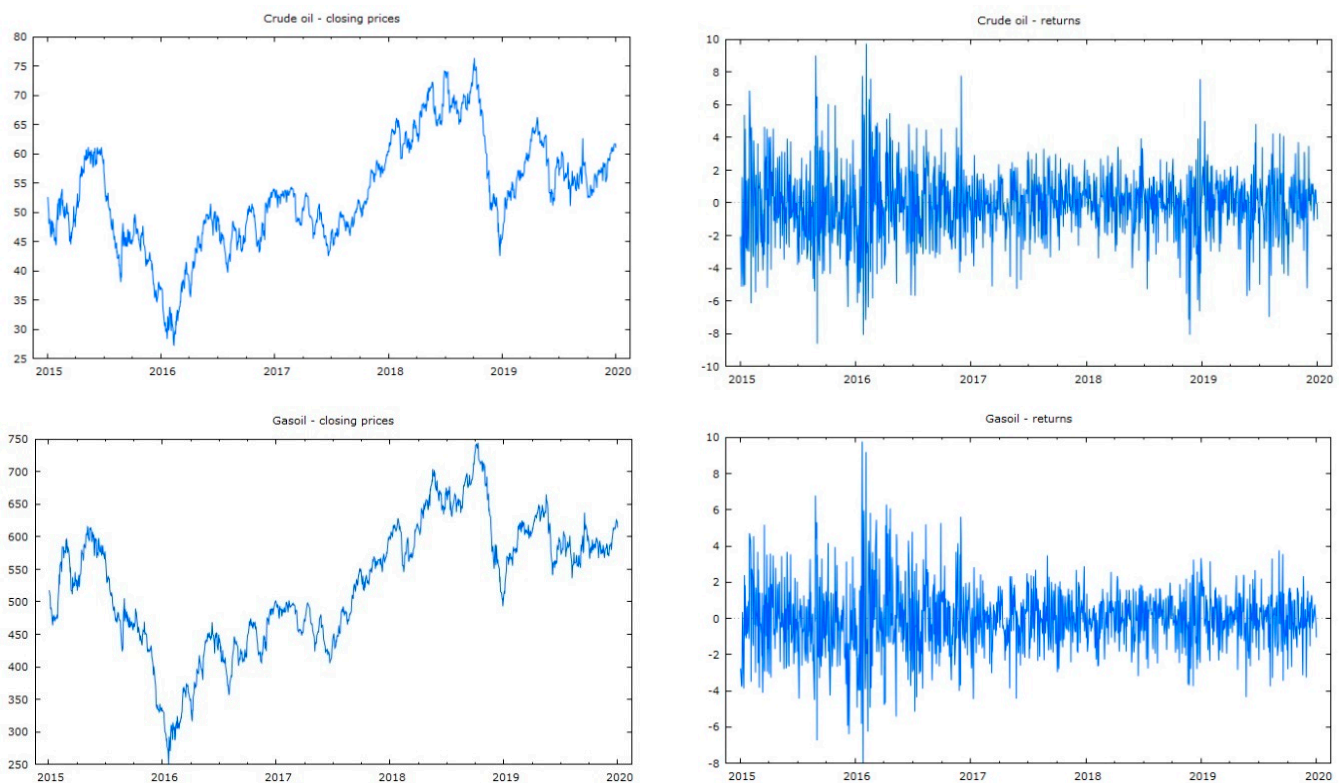


Figure 1. Cont.

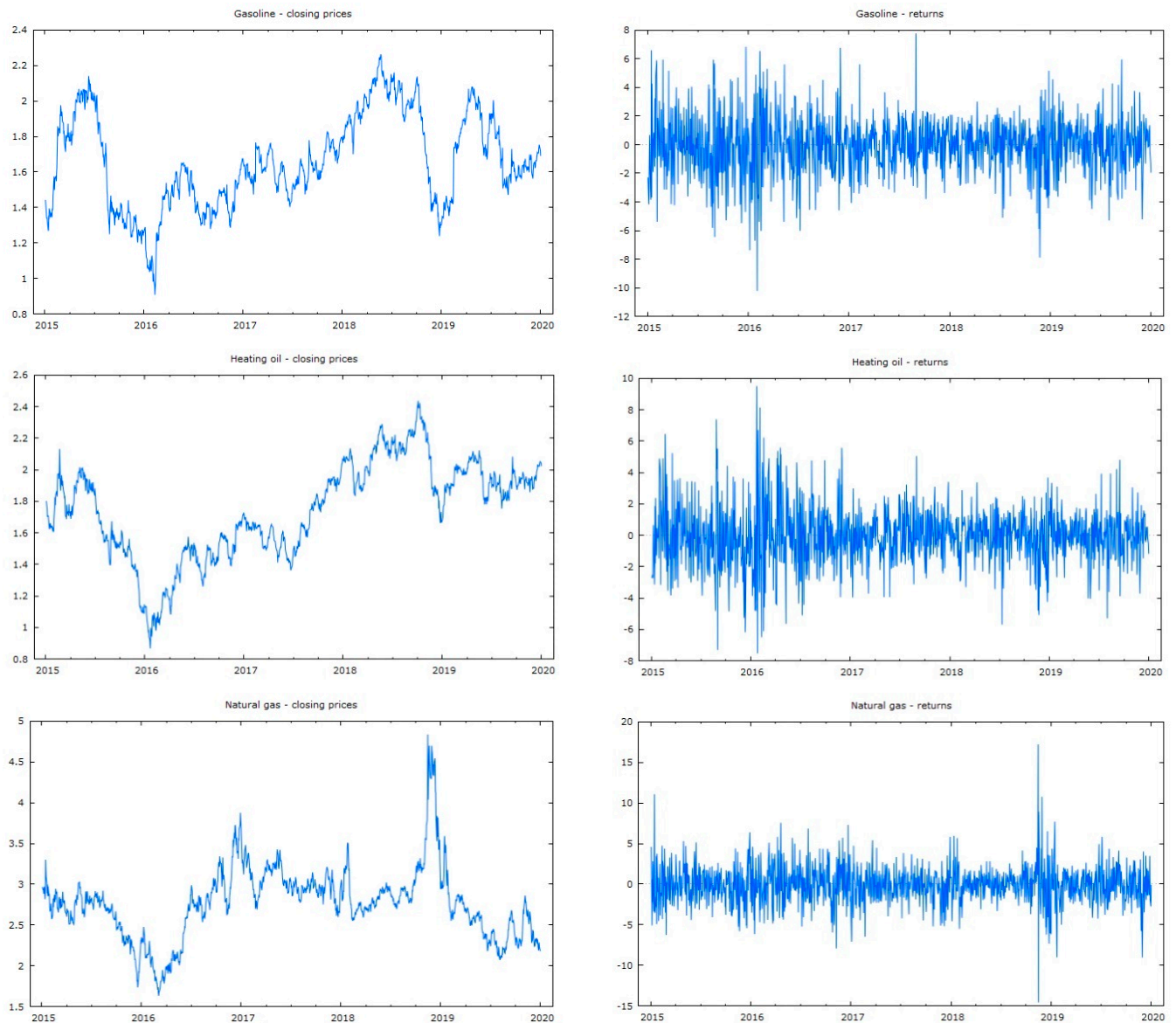


Figure 1. Daily closing prices and open-to-close returns of the investigated energy commodities.

The descriptive statistics for daily returns and two considered proxies of volatility are presented in Table 1. The calculated means of returns are negative, despite the prices of all commodities (except natural gas) increased during the analyzed period. It stems from the fact that we analyze open-to-close returns instead of close-to-close ones. The absolute value of the minimum and maximum returns are relatively high. The highest volatility of returns (see standard deviations for returns) but also variation of volatility (see the standard deviations for both the squared returns and the Parkinson estimator) can be seen for natural gas, which is a well-known fact for this energy asset (see e.g., [71]). Crude oil is the second most volatile contract. Meanwhile, the lowest volatility is for gasoil and heating oil.

Table 1. Summary statistics of daily returns, squared returns and values of the Parkinson estimator.

Commodities	Mean	Min	Max	SD	Skew	Kurt	LB
Returns							
Crude oil	−0.053	−8.624	9.742	2.190	−0.042	4.509	0.187
Gasoil	−0.019	−7.994	7.775	1.797	0.224	5.386	0.096
Gasoline	−0.034	−8.624	9.742	2.004	−0.13	4.492	0.205
Heating oil	−0.022	−10.226	7.776	1.852	0.182	4.994	0.103
Natural gas	−0.091	−14.484	17.216	2.364	0.260	7.242	0.147
Squared returns							
Crude oil	4.799	0.000	94.905	8.993	4.212	27.429	0.000
Gasoil	3.231	0.000	95.543	6.759	5.886	57.243	0.000
Gasoline	4.018	0.000	104.575	7.517	4.751	40.042	0.000
Heating oil	3.432	0.000	90.421	6.850	5.018	41.009	0.000
Natural gas	5.597	0.000	296.392	13.926	11.445	197.959	0.000
Parkinson estimator							
Crude oil	4.763	0.253	54.245	5.618	3.276	17.581	0.000
Gasoil	3.432	4.245	43.255	4.245	3.848	23.695	0.000
Gasoline	4.146	0.191	49.761	4.579	3.488	21.647	0.000
Heating oil	3.355	0.204	36.649	3.887	3.549	20.646	0.000
Natural gas	5.581	0.262	162.507	8.354	9.635	154.035	0.000

Note: Mean is arithmetic mean, Min is minimum, Max is maximum, SD is standard deviation, Skew is skewness, Kurt is excess kurtosis, LB denotes the p -value of the Ljung-Box test for 10 lags. The sample period is 2 January 2015 to 31 December 2019.

The distributions of both proxies of volatility exhibit strong skewness and high kurtosis. Significantly higher variability can be seen for the squared returns than for the Parkinson estimator which indicates much stronger noise in the former. The autocorrelation of returns is weak and mostly insignificant. The autocorrelation of the squared returns and the Parkinson estimator is very high and significant, much stronger for the latter measure of volatility. These results shows that the Parkinson estimator can be useful as the volatility measure.

3.2. Forecasting Procedure

We compared the forecasting performance of the GARCH models with SVR. We have not detected any significant dependencies in the conditional mean, which is the reason for modelling only the conditional variance. To estimate both the GARCH and SVR models, we used a rolling window and apply the following procedure. For the starting sample (i.e., 2 January 2015 to 30 December 2016) we estimate models and obtain one-day-ahead forecasts. Consecutively, we added one new observation to the estimation sample, while at the same time removing the oldest observation. Then, based on each estimation sample, we re-estimated the models and made forecasts. We repeat this procedure until we obtain forecasts for the three-year period from 3 January 2017 to 31 December 2019. In our procedure we consider a small estimation sample, because the persistence of the conditional volatility in large samples could be exaggerated by the existence of structural breaks in the GARCH parameters (see [72]).

The considered GARCH models are GARCH-n, GARCH-t, EGARCH, GJR, APARCH, IGARCH and GARCH-M. The parameters of the models are estimated using the quasi-maximum likelihood method, except for the GARCH-t model, whereby the maximum likelihood method was applied.

We consider autoregressive SVR models, which means that we calculate the variance forecasts $\sigma_{f,t+1}^2$ using the lagged squared returns as predictor variables:

$$\sigma_{f,t+1}^2 = f(r_t^2, r_{t-1}^2, \dots, r_{t-l+1}^2), \quad (21)$$

where l is the lag length. In order to construct the SVR models the regressors $r_t^2, r_{t-1}^2, \dots, r_{t-l+1}^2$ were first standardized, i.e., the lagged squared returns were centered by subtracting their mean and divided by standard deviation. After applying the Model (21) we use reverse standardization of $\sigma_{f,t+1}^2$ to calculate the final variance forecast.

We applied two kernels in the SVR models: The linear and RBF ones and four values of lags: $l = 1, l = 5, l = 10$ and $l = 15$, however, we present only the results for lags $l = 1$ and $l = 15$. The lag $l = 1$ led to the simplest specification of the model, in which only the previous lagged squared return is used as the predictor variable. Obviously, higher lags led to a more general form of the model, which could potentially generate more accurate forecasts. On the other hand, high lags could introduce spurious information to the model, and additionally, substantially increase computation time. Our calculations show that $l = 15$ leads to slightly more accurate forecasts than $l = 5$ and $l = 10$ and it seems to be optimal when considering the accuracy of the forecasts and computation time.

Finally, we consider four specifications of the SVR models:

- (1) SVR with the linear kernel and $l = 1$ (hereafter SVR-lin-1),
- (2) SVR with the linear kernel and $l = 15$ (SVR-lin-15),
- (3) SVR with the RBF kernel and $l = 1$ (SVR-rbf-1),
- (4) SVR with the RBF kernel and $l = 15$ (SVR-rbf-15).

As stated before, the values of the ε and C hyperparameters (and additionally γ in the case of the RBF kernel) must be determined. For this aim, we applied the grid search technique. This method consisted in constructing many models for different values of the hyperparameters and selecting the optimal model on the basis of a validation set (we use the function `fitsvm` in MATLAB R2015b to train the SVR models). We performed the grid search by considering consecutive values of $C = 0.5, 1, 1.5, \dots, 10$, $\varepsilon = 0.5, 0.6, 0.7, \dots, 2.5$ and $\gamma = 2^{-3}, 2^{-2}, \dots, 2^2, 2^3$. To evaluate the model for each combination of hyperparameters, a 10-fold cross-validation procedure is applied. According to this approach, the investigated sample was randomly partitioned into 10 equal-sized subsamples. Nine of them were used to construct the SVR model, while the remaining one was used to validate the model. To this end, the mean square error (MSE) was computed on the observations in the validation subsample. This procedure is repeated 10 times (for each of the 10 subsamples used as a test set), and the average of 10 values of the obtained MSEs was calculated. Finally, the hyperparameters that led to the smallest MSE were considered to be optimal. It is worth noting that the optimal hyperparameters were determined for each forecast separately.

It should be emphasized that the computation of forecasts based on the SVR model (Equation (21)) does not ensure that it will always output non-negative values. In such cases we propose to take the previous squared return as the forecast, i.e., $\sigma_{f,t+1}^2 = r_t^2$. However, situations where our models led to negative outputs are very exceptional. Such problem occurs only for natural gas, for which SVR-lin-15 leads to 9 (out of 759) negative forecasts.

In the evaluation of forecasts we consider two proxies of volatility: the squared daily return and the Parkinson estimator. In Section 3.3 we present results for the squared daily return and Section 3.4 shows the results for the Parkinson estimator. Despite shortcomings of the squared daily return, we apply it due to its popularity in previous studies in which GARCH models have been compared with SVR models (see Introduction).

The forecasts are evaluated based on two primary measures, namely, the mean squared error and the mean absolute error.

The MSE is the most frequently used criterion in forecasting studies. It is written as,

$$\text{MSE} = \frac{1}{T} \sum_{t=1}^T (\sigma_{R,t}^2 - \sigma_{f,t}^2)^2, \quad (22)$$

where $\sigma_{R,t}^2$ is the proxy of volatility of returns and $\sigma_{f,t}^2$ is the variance forecasts at time t , T is the number of forecasts.

The MSE is robust to the use of a noisy volatility proxy (it yields the same ranking of competing forecasts using an unbiased volatility proxy, see e.g., [59]).

The mean absolute error (MAE) is less sensitive to outliers, which is very important when evaluating extraordinary events. It is given as:

$$\text{MAE} = \frac{1}{T} \sum_{t=1}^T \left| \sigma_{R,t}^2 - \sigma_{f,t}^2 \right|. \quad (23)$$

In order to assess whether the loss differentials between competing models are statistically significant two different tests are applied: the test of superior predictive ability (SPA) of Hansen [73] and the model confidence set (MCS) of Hansen et al. [74]. In the first test, it is checked whether each of the models considered is outperformed significantly by any of the alternatives. In this regard, the performance of the benchmark model relative to model k can be described as,

$$d_{k,t} = L(\sigma_{R,t}^2, \sigma_{B,t}^2) - L(\sigma_{R,t}^2, \sigma_{k,t}^2), \quad k = 1, \dots, m, \quad t = 1, \dots, T, \quad (24)$$

where $\sigma_{B,t}^2$ and $\sigma_{k,t}^2$ are the volatility forecasts from the benchmark model and model k , respectively, $L(\sigma_{R,t}^2, \sigma_{B,t}^2)$, $L(\sigma_{R,t}^2, \sigma_{k,t}^2)$ denote the loss functions, and m is the number of competing models (excluding the benchmark model). In this study, we applied two measures, namely MSE and MAE, to calculate $L(\sigma_{R,t}^2, \sigma_{B,t}^2)$. The null hypothesis of the SPA test is formulated as,

$$H_0 : E[d_{k,t}] \leq 0, \text{ for all } k = 1, \dots, m, \quad (25)$$

meaning that the benchmark model is the best forecasting model compared to any of the models $k = 1, \dots, m$. The test statistic can be expressed as,

$$\text{SPA} = \max_k \frac{\sqrt{T} \bar{d}_k}{\bar{\omega}_k}, \quad (26)$$

where \bar{d}_k is the mean of $d_{k,t}$ and $\bar{\omega}_k$ is a consistent estimator of the asymptotic variance.

The objective of the MCS procedure is to determine the set of best models, denoted as M_{best} , from a given collection of models, M . The set of the best models is defined as,

$$M_{best} \equiv \{i \in M : E[d_{ij}] \leq 0\} \text{ for all } j \in M, \quad (27)$$

where $d_{ij} = L(\sigma_{R,t}^2, \sigma_{i,t}^2) - L(\sigma_{R,t}^2, \sigma_{j,t}^2)$ is the loss differential for $i, j \in M$.

The null hypothesis is as follows,

$$H_0 : E[d_{ij,t}] = 0, \text{ for all } i, j \in M_s, \quad (28)$$

where $M_s \subset M$. The testing procedure begins with initially setting $M_s = M$. Then the null hypothesis is tested at a given significance level. If the null is not rejected then the $\hat{M}_{best} = M_s$, otherwise the model that contributes most to the test statistic is removed from M_s and the whole procedure is repeated until there is no more models to be removed. The \hat{M}_{best} is then referred to as the model confidence set (MCS). The best models are selected with a given level of confidence in terms of a criterion for the loss function that is user-specified. In our case we use the MSE and MAE as such criteria.

3.3. Results for the Squared Daily Return Used as a Proxy of Volatility

In this Section we evaluate the forecasts by applying the squared daily return as a proxy of volatility. The results are given in Tables 2 and 3, for the MSE, and MAE measures, respectively.

Table 2. Evaluation of the variance forecasts in terms of the MSE measure for the squared daily returns used as a proxy of volatility.

Model	Crude Oil			Gasoil			Gasoline			Heating Oil			Natural Gas		
	MSE	SPA	MCS	MSE	SPA	MCS	MSE	SPA	MCS	MSE	SPA	MCS	MSE	SPA	MCS
GARCH-n	3.862	0.155	0.294 *	0.727	0.869	0.851 *	2.854	0.589	0.614 *	1.372	0.800	0.951 *	24.575	0.695	0.651 *
GARCH-t	3.860	0.210	0.325 *	0.727	0.843	0.851 *	2.889	0.337	0.568 *	1.379	0.351	0.839 *	24.306	0.681	0.651 *
EGARCH	3.777	0.799	0.656 *	0.726	0.723	0.851 *	2.849	0.397	0.614 *	1.360	0.886	1.000 *	26.036	0.206	0.651 *
GJR	3.754	0.945	1.000 *	0.739	0.197	0.773 *	2.807	0.747	0.750 *	1.379	0.334	0.893 *	24.231	0.501	0.651 *
APARCH	3.819	0.453	0.615 *	0.720	0.960	1.000 *	2.790	0.896	1.000 *	1.416	0.130	0.428 *	24.428	0.221	0.651 *
IGARCH	3.915	0.005	0.133 *	0.731	0.519	0.851 *	2.911	0.119	0.540 *	1.402	0.032	0.343 *	24.708	0.422	0.651 *
GARCH-M	3.896	0.171	0.133 *	0.734	0.554	0.851 *	2.865	0.180	0.577 *	1.381	0.048	0.789 *	24.656	0.372	0.651 *
SVR_lin_1	3.968	0.154	0.185 *	0.757	0.000	0.170 *	2.884	0.450	0.600 *	1.366	0.770	0.954 *	26.128	0.051	0.651 *
SVR_lin-15	3.918	0.387	0.350 *	0.729	0.629	0.851 *	2.891	0.401	0.577 *	1.362	0.902	0.954 *	25.680	0.844	1.000 *
SVR-rbf-1	3.926	0.169	0.129 *	0.762	0.065	0.577 *	2.893	0.306	0.568 *	1.390	0.168	0.647 *	26.073	0.064	0.651 *
SVR-rbf-15	3.957	0.173	0.128 *	0.757	0.051	0.275 *	3.013	0.007	0.444 *	1.410	0.010	0.193 *	25.543	0.380	0.651 *

Note: The values of MSE are multiplied by 10^{-1} , the lowest values of MSE for each energy commodity are marked in bold, * indicates that models belong to MCS with a confidence level of 0.90. SPA and MCS denote the p -value of the SPA and MCS tests, respectively. The evaluation period is from 3 January 2017 to 31 December 2019.

Table 3. Evaluation of the variance forecasts in terms of the MAE measure for the squared daily returns used as a proxy of volatility.

Model	Crude Oil			Gasoil			Gasoline			Heating Oil			Natural Gas		
	MAE	SPA	MCS	MAE	SPA	MCS	MAE	SPA	MCS	MAE	SPA	MCS	MAE	SPA	MCS
GARCH-n	0.343	0.000	0.000	0.184	0.000	0.000	0.302	0.000	0.000	0.229	0.139	0.082	0.568	0.000	0.010
GARCH-t	0.341	0.000	0.000	0.185	0.000	0.000	0.300	0.000	0.001	0.229	0.163	0.083	0.554	0.001	0.004
EGARCH	0.330	0.023	0.000	0.175	0.721	1.000 *	0.302	0.024	0.003	0.224	0.968	1.000 *	0.590	0.000	0.000
GJR	0.334	0.003	0.000	0.180	0.001	0.022	0.303	0.007	0.001	0.228	0.061	0.083	0.577	0.000	0.002
APARCH	0.332	0.048	0.000	0.176	0.280	0.564 *	0.301	0.018	0.003	0.227	0.338	0.326 *	0.601	0.000	0.001
IGARCH	0.353	0.000	0.000	0.183	0.000	0.002	0.308	0.001	0.000	0.234	0.007	0.020	0.579	0.000	0.001
GARCH-M	0.342	0.000	0.000	0.202	0.000	0.000	0.303	0.000	0.000	0.234	0.000	0.009	0.569	0.001	0.005
SVR_lin_1	0.326	0.000	0.000	0.198	0.000	0.000	0.290	0.096	0.197 *	0.234	0.000	0.042	0.486	0.512	0.679 *
SVR_lin-15	0.317	0.568	1.000 *	0.190	0.000	0.000	0.288	0.915	1.000 *	0.229	0.184	0.142 *	0.464	0.720	1.000 *
SVR-rbf-1	0.329	0.000	0.000	0.197	0.000	0.000	0.292	0.046	0.117 *	0.236	0.001	0.010	0.488	0.279	0.679 *
SVR-rbf-15	0.339	0.000	0.000	0.195	0.000	0.000	0.314	0.000	0.000	0.238	0.001	0.005	0.531	0.000	0.110 *

Note: The values of MAE are multiplied by 10^{-1} , the lowest values of MAE for each energy commodity are marked in bold, * indicates that models belong to MCS with a confidence level of 0.90. SPA and MCS denote the p -values of the SPA and MCS tests, respectively. The evaluation period is from 3 January 2017 to 31 December 2019.

Both for the MSE and MAE measures, the highest errors of the forecasts are for natural gas, while the lowest are for gasoil. It corresponds with the fact that these commodities have the highest, and the lowest volatility of daily returns, respectively (see Table 1). The values of MSE are considerably higher than of MAE because the former measure is more sensitive to outliers.

Figure 2 depicts the one-day-ahead volatility forecasts for one selected commodity, namely crude oil. For greater clarity, we present the forecasts only for the two best models in their classes chosen according to the MSE criterion, i.e., SVR-lin-15 and GARCH-t. We found that the forecasts from the GARCH models are greater than those from SVR in times of high volatility. It relates to the observation that the GARCH models react more quickly and strongly to the past huge changes in volatility than the SVR models. That is the reason squared forecasting errors are higher for the SVR models than for the GARCH ones. These general conclusions are also valid for other analyzed energy commodities.

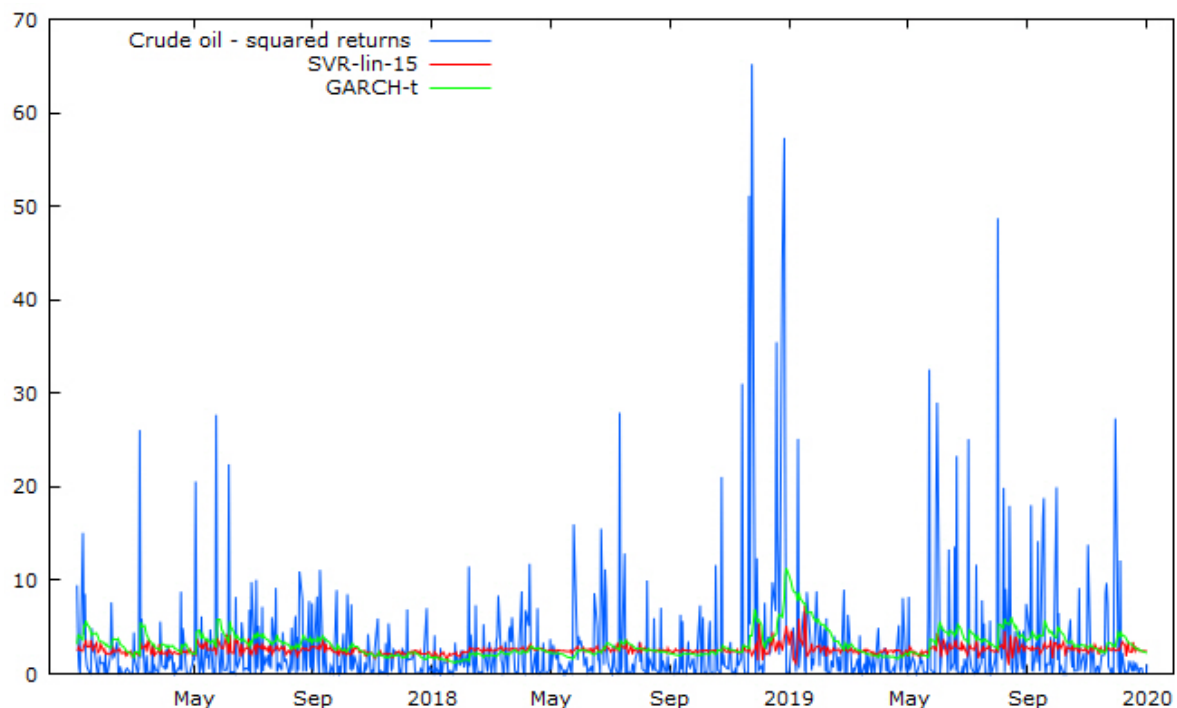


Figure 2. One-day-ahead volatility forecasts and the squared daily returns for crude oil.

Generally, the GARCH models have lower MSE values than the SVR models, but when it comes to MAE we cannot derive such general conclusion. However, according to the MAE measure the SVR-lin-15 model is often preferable. In order to assess whether these findings are statistically significant we apply the SPA and MCS tests (calculated p -values are given in Tables 2 and 3, for the MSE, and MAE measures, respectively).

The results of the SPA test for the MSE measure indicate that only the following models are outperformed significantly (at the 10% significance level) by any of the alternatives: IGARCH for crude oil, SVR-lin-1, SVR-rbf-1, SVR-rbf-15 for gasoil, SVR-rbf-15 for gasoline, IGARCH, GARCH-M, SVR-rbf-15 for heating oil and SVR-lin-1, SVR-rbf-1 for natural gas. According to the results of the MCS test, all models for all commodities belong to the model confidence set and there is no evidence to reject the null hypothesis of equal predictive ability. It means that the MCS test with the MSE criterion does not differentiate the examined models.

In contrast to the results for MSE, the SPA test for the MAE measure rejects the null hypothesis for most cases, indicating that most models are outperformed significantly by any of the alternatives. The only exceptions are: SVR-lin-15 for crude oil, EGARCH, APARCH for gasoil, SVR-lin-15 for gasoline, GARCH-n, GARCH-t, EGARCH, APARCH,

SVR-lin-15 for heating oil and SVR-lin-1, SVR-lin-15, SVR-rbf-1 for natural gas. Similar conclusions come from the results of the MCS test and indicate that the most accurate forecast of volatility are obtained from SVR-lin-15 for crude oil, EGARCH, APARCH for gasoil, SVR-lin-1, SVR-rbf-1 for gasoline, EGARCH, APARCH, SVR-lin-15 for heating oil and SVR-lin-1, SVR-lin-15, SVR-rbf-1, SVR-rbf-15 for natural gas. It is worth noting that SVR-lin-15 is among the best models for all commodities except gasoil. It is difficult to choose the best model among the GARCH models for all analyzed commodities. However, forecasts based on the asymmetric GARCH models are often the most accurate.

3.4. Results for the Parkinson Estimator Used as a Proxy of Volatility

In this Section we evaluate the forecasting performance of the analyzed models by applying the Parkinson estimator as a proxy of volatility instead of the squared daily return. As it was discussed, this estimator is significantly more efficient than the classical variance estimator based on closing prices. For this reason we expect that the results of an evaluation of models may change significantly. The corresponding results are given in Tables 4 and 5 for the MSE, and MAE measures, respectively.

Both the MSE and MAE forecasting errors are significantly lower when the Parkinson estimator is used for the forecasts evaluation. The values of these measures are sometimes even more than three times lower than those obtained for the squared daily returns (compare Tables 2 and 3). The highest errors of forecasts are for natural gas, while the lowest for heating oil and gasoil.

Figure 3 depicts the one-day-ahead volatility forecasts for the same commodity as shown in Figure 2, namely crude oil. We present the forecasts only for the two best models in their classes chosen according to the MSE criterion, i.e., SVR-lin-15 and GJR. The main difference between Figures 2 and 3 is that the values of the Parkinson estimator are considerably lower than the squared daily returns in the periods of high volatility. The findings derived from Figure 3 are similar to those from Figure 2. One can see that the forecasts from the GARCH models are greater than those from the SVR ones in times of high volatility. Moreover, the GARCH models fit even better to extreme changes in volatility than it is observed in the Figure 2. These general conclusions are also valid for the other investigated energy commodities.

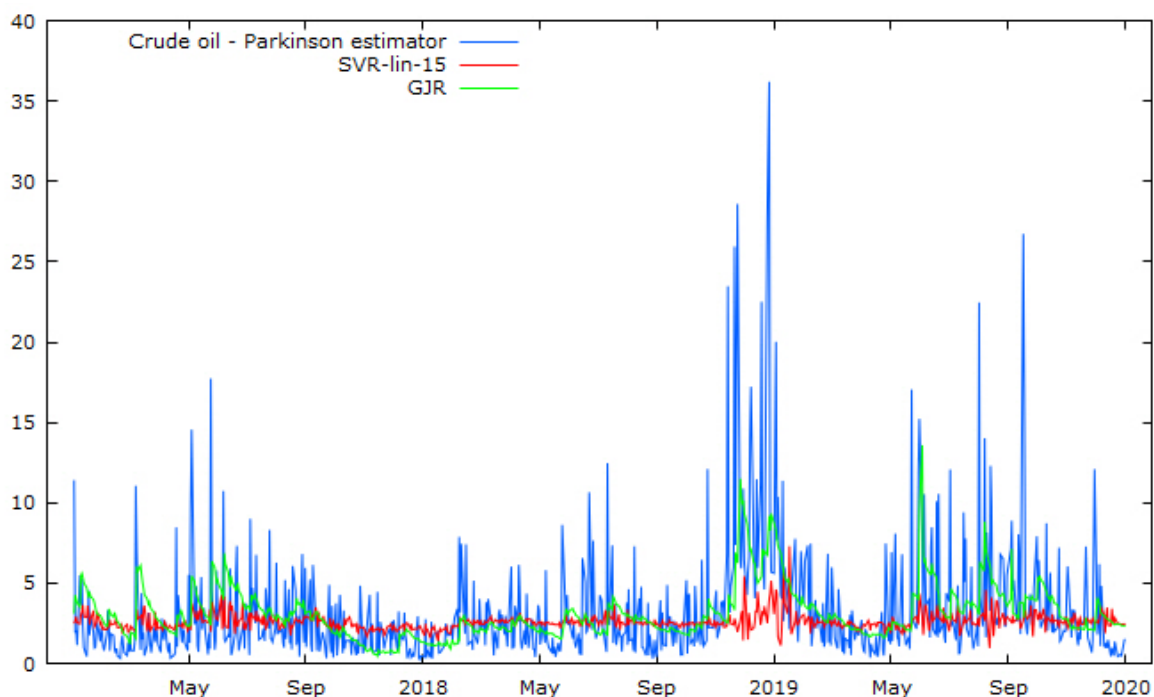


Figure 3. One-day-ahead volatility forecasts and values of the Parkinson estimator for crude oil.

Table 4. Evaluation of the variance forecasts in terms of the MSE measure for the Parkinson estimator used as a proxy of volatility.

Model	Crude Oil			Gasoil			Gasoline			Heating Oil			Natural Gas		
	MSE	SPA	MCS	MSE	SPA	MCS	MSE	SPA	MCS	MSE	SPA	MCS	MSE	SPA	MCS
GARCH-n	1.231	0.015	0.047	0.587	0.017	0.104 *	0.922	0.391	0.600 *	0.430	0.289	0.143 *	7.568	0.923	0.939 *
GARCH-t	1.244	0.025	0.031	0.582	0.128	0.116 *	0.961	0.056	0.341 *	0.434	0.125	0.113 *	7.428	0.906	0.948 *
EGARCH	1.188	0.265	0.306 *	0.618	0.030	0.063	0.943	0.183	0.449 *	0.411	0.994	1.000 *	8.539	0.185	0.154 *
GJR	1.147	0.935	1.000 *	0.621	0.034	0.029	0.910	0.543	0.668 *	0.430	0.110	0.113 *	7.359	0.985	1.000 *
APARCH	1.197	0.364	0.306 *	0.586	0.111	0.116 *	0.896	0.634	0.668 *	0.451	0.118	0.074	7.691	0.404	0.703 *
IGARCH	1.232	0.008	0.095	0.577	0.179	0.116 *	0.879	0.882	1.000 *	0.444	0.078	0.074	7.671	0.317	0.703 *
GARCH-M	1.267	0.026	0.026	0.553	0.968	1.000 *	0.929	0.158	0.490 *	0.434	0.223	0.113 *	7.609	0.381	0.703 *
SVR_lin_1	1.431	0.018	0.018	0.620	0.005	0.023	1.044	0.031	0.210 *	0.449	0.078	0.068	9.924	0.004	0.012
SVR-lin-15	1.378	0.006	0.022	0.601	0.049	0.085	1.027	0.020	0.160 *	0.441	0.148	0.113 *	8.726	0.029	0.045
SVR-rbf-1	1.405	0.002	0.002	0.631	0.018	0.034	1.052	0.027	0.041	0.464	0.022	0.049	9.865	0.007	0.058
SVR-rbf-15	1.385	0.022	0.006	0.633	0.000	0.011	1.078	0.000	0.007	0.478	0.003	0.030	9.120	0.106	0.154 *

Note: The values of MSE are multiplied by 10^{-1} , the lowest values of MSE for each energy commodity are marked in bold, * indicates that models belong to MCS with a confidence level of 0.90. SPA and MCS denote the p -values of the SPA and MCS tests, respectively. The evaluation period is from 3 January 2017 to 31 December 2019.

Table 5. Evaluation of the variance forecasts in terms of the MAE measure for the Parkinson estimator used as a proxy of volatility.

Model	Crude Oil			Gasoil			Gasoline			Heating Oil			Natural Gas		
	MAE	SPA	MCS	MAE	SPA	MCS	MAE	SPA	MCS	MAE	SPA	MCS	MAE	SPA	MCS
GARCH-n	0.212	0.003	0.021	0.135	0.012	0.001	0.186	0.978	1.000 *	0.139	0.011	0.005	0.342	0.085	0.143 *
GARCH-t	0.212	0.005	0.009	0.135	0.014	0.001	0.188	0.352	0.880 *	0.139	0.012	0.009	0.328	0.872	1.000 *
EGARCH	0.198	0.991	1.000 *	0.134	0.017	0.007	0.191	0.308	0.761 *	0.131	0.791	1.000 *	0.370	0.000	0.004
GJR	0.201	0.321	0.271 *	0.136	0.002	0.000	0.188	0.633	0.937 *	0.137	0.007	0.040	0.350	0.055	0.045
APARCH	0.203	0.162	0.108*	0.129	0.694	1.000 *	0.186	0.865	0.998 *	0.133	0.364	0.412 *	0.382	0.000	0.001
IGARCH	0.216	0.000	0.004	0.135	0.006	0.001	0.186	0.792	0.998 *	0.142	0.000	0.001	0.349	0.019	0.027
GARCH-M	0.213	0.004	0.001	0.143	0.000	0.000	0.187	0.259	0.974 *	0.141	0.000	0.001	0.342	0.082	0.072
SVR_lin_1	0.217	0.000	0.001	0.149	0.000	0.000	0.193	0.023	0.534 *	0.145	0.000	0.000	0.364	0.004	0.013
SVR-lin-15	0.204	0.237	0.108 *	0.140	0.000	0.000	0.188	0.569	0.938 *	0.141	0.004	0.002	0.347	0.279	0.364 *
SVR-rbf-1	0.219	0.000	0.000	0.149	0.000	0.000	0.195	0.003	0.330 *	0.147	0.000	0.000	0.359	0.024	0.018
SVR-rbf-15	0.222	0.000	0.000	0.148	0.000	0.000	0.206	0.000	0.045	0.148	0.000	0.000	0.362	0.034	0.013

Note: The values of MAE are multiplied by 10^{-1} , the lowest values of MAE for each energy commodity are marked in bold, * indicates that models belong to MCS with a confidence level of 0.90. SPA and MCS denote the p -values of the SPA and MCS tests, respectively. The evaluation period is from 3 January 2017, to 31 December 2019.

The results of the SPA test for the MSE measure show that the following models are not outperformed significantly by any of the alternatives: EGARCH, GJR, APARCH for crude oil, GARCH-t, APARCH, IGARCH, GARCH-M for gasoil, GARCH-n, EGARCH, GJR, APARCH, IGARCH, GARCH-M for gasoline, GARCH-n, GARCH-t, EGARCH, GJR, APARCH, GARCH-M, SVR-lin-15 for heating oil and all GARCH models with SVR-rbf-15 for natural gas. The results of the MCS test are similar and the following models belong to the model confidence set: EGARCH, GJR, APARCH for crude oil, GARCH-n, GARCH-t, APARCH, IGARCH, GARCH-M for gasoil, all GARCH models with SVR-lin-1, SVR-lin-15 for gasoline, GARCH-n, GARCH-t, EGARCH, GJR, GARCH-M, SVR-lin-15 for heating oil and all GARCH models with SVR-rbf-15 for natural gas. According to the MSE measure the forecasts of volatility from the SVR models are generally inferior to the forecasts based on the GARCH models.

When it comes to the MAE measure the conclusions are quite similar to those for the MSE criterion. Significantly more precise forecasts of volatility are based on: EGARCH, GJR, APARCH, SVR-lin-15 for crude oil, APARCH for gasoil, all GARCH models with SVR-lin-15 (according to the SPA test) and all considered models except SVR-rbf-15 (according to the MCS test) for gasoline, EGARCH, APARCH for heating oil, GARCH-t, SVR-lin-15 (according to the SPA test) and GARCH-n, GARCH-t, SVR-lin-15 (according to the MCS test) for natural gas. Therefore for three energy commodities the SVR-lin-15 model is not significantly inferior to the GARCH models. It is difficult to choose the best model among the GARCH models for all analyzed commodities, however, similarly to the results in Section 3.3, forecasts based on the asymmetric GARCH models are often the most accurate.

3.5. Discussion of the Results

In this study, we apply two proxies of volatility to evaluate forecasts: The squared return and the Parkinson estimator. The former is an extremely noisy variance estimator and its usage can lead to an unreliable evaluation of models. That is the reason why the Parkinson estimator, which is significantly more accurate measure of volatility is also adopted.

When the squared daily returns are used as an ex-post volatility measure the results are ambiguous. The MSE values are large (due to the existence of large outliers) and it is not possible to indicate significantly better models amongst analyzed ones. Meanwhile, the MAE criterion favors the SVR-lin-15 model for most commodities. On the other hand, when the Parkinson estimator is used as a proxy of volatility the forecasting errors are significantly lower, indicating more accurate predictions from the considered models and usually the obtained results favor the GARCH models over the SVR ones. Our findings indicate that the accuracy of volatility forecasts depends substantially on the applied proxy of volatility. This conclusion is important since in most papers concerning the application of SVR models to volatility forecasting, only the squared daily returns (or a moving average of the daily squared returns) have been analyzed (e.g., [22–33]). Our results, obtained for the squared daily returns, confirm the conclusion formulated in these studies that SVR can lead to lower forecasting errors than the GARCH models. However, we argue that this forecasting superiority of SVR models is not unequivocal, since it depends on the measure used to evaluate the forecasting errors and is valid only for models with properly determined hyperparameters.

4. Conclusions

Due to global economic developments, emerging technological advances and economic, geopolitical and environmental events a significant increase in volatility of energy commodities prices has occurred in the last 20 years. This highly volatile environment has become attractive to financial speculators, magnifying the risk on the energy commodities markets. That is reason there is a strong need to look for more accurate methods of volatility forecasting for such commodities.

In the paper we compare the forecasting performance of the GARCH-type models with support vector regression for futures contracts of selected energy commodities: Crude oil, natural gas, heating oil, gasoil and gasoline. The GARCH models are a standard tool applied in the volatility literature, while SVR is one of machine learning techniques which have been gaining huge popularity in recent years.

We show that the accuracy of volatility forecasts depends substantially on the applied proxy of volatility. Our study confirms that SVR with properly determined hyperparameters can lead to lower forecasting errors than the GARCH models when the squared daily return is used as the proxy of volatility in an evaluation. Meanwhile, if we apply the Parkinson estimator which is a more accurate approximation of volatility, the results are different since they usually favor the GARCH models over SVR.

Moreover, it is difficult to choose the best model among the GARCH models for all analyzed commodities, however, forecasts based on the asymmetric GARCH models are often the most accurate. While, in the class of the SVR models, the results indicate the forecasting superiority of the SVR model with the linear kernel and 15 lags. Precisely speaking, in 92% (i.e., 18 cases out of 20) the SVR model with the linear kernel and 15 lags has the lowest MSE or MAE among SVR models, for all analyzed time series and for both volatility proxies.

In the future, this study can be extended in several directions. Firstly, other machine learning methods like neural networks or hybrid models can be applied. Secondly, other proxies of volatility like the realized variance or the bi-power variation can be used for the evaluation of forecasts. Thirdly, the analysis can be done for the COVID-19 crisis that is, for a period with unheard-of volatility on the energy market. Fourthly, the comparison of the models can be performed for simulated data assuming different generating processes.

Author Contributions: W.O. described, estimated and applied the SVR models. M.F. and P.F. described, estimated and applied the GARCH-type models. All authors wrote, reviewed and commented on the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Science Centre under Grant 2019/35/B/HS4/00642.

Acknowledgments: The authors would like to thank four anonymous reviewers for helpful and constructive comments.

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

References

1. Halkos, G.E.; Tzirivis, A.S. Effective energy commodity risk management: Econometric modeling of price volatility. *Econ. Anal. Policy* **2019**, *63*, 234–250. [\[CrossRef\]](#)
2. Nomikos, N.K.; Pouliasis, P.K. Forecasting petroleum futures markets volatility: The role of regimes and market conditions. *Energy Econ.* **2011**, *33*, 321–337. [\[CrossRef\]](#)
3. Wang, Y.; Wu, C. Forecasting energy market volatility using GARCH models: Can multivariate models beat univariate models? *Energy Econ.* **2012**, *34*, 2167–2181. [\[CrossRef\]](#)
4. Chkili, W.; Hammoudeh, S.; Nguyen, D.K. Volatility forecasting and risk management for commodity markets in the presence of asymmetry and long memory. *Energy Econ.* **2014**, *41*, 1–18. [\[CrossRef\]](#)
5. Klein, T.; Walther, T. Oil price volatility forecast with mixture memory GARCH. *Energy Econ.* **2016**, *58*, 46–58. [\[CrossRef\]](#)
6. Kumar, D. Forecasting energy futures volatility based on the unbiased extreme value volatility estimator. *IIMB Manag. Rev.* **2017**, *29*, 294–310. [\[CrossRef\]](#)
7. Herrera, A.M.; Hu, L.; Pastor, D. Forecasting crude oil price volatility. *Int. J. Forecast.* **2018**, *34*, 622–635. [\[CrossRef\]](#)
8. Zhang, Y.-J.; Zhang, J.-L. Volatility forecasting of crude oil market: A new hybrid method. *J. Forecast.* **2018**, *37*, 781–789. [\[CrossRef\]](#)
9. Bildirici, M.; Bayazit, N.G.; Ucan, Y. Analyzing crude oil prices under the impact of COVID-19 by using LSTARGARCHLSTM. *Energies* **2020**, *13*, 2980. [\[CrossRef\]](#)
10. Lin, L.; Jiang, Y.; Xiao, H.; Zhou, Z. Crude oil price forecasting based on a novel hybrid long memory GARCH-M and wavelet analysis model. *Phys. A* **2020**, *543*, 123532. [\[CrossRef\]](#)
11. Lin, Y.; Xiao, Y.; Li, F. Forecasting crude oil price volatility via a HM-EGARCH model. *Energy Econ.* **2020**, *87*, 104693. [\[CrossRef\]](#)
12. Lv, X.; Shan, X. Modeling natural gas market volatility using GARCH with different distributions. *Phys. A* **2013**, *392*, 5685–5699. [\[CrossRef\]](#)

13. Xie, W.; Yu, L.; Xu, S.; Wang, S. A New Method for Crude Oil Price Forecasting Based on Support Vector Machines. In Proceedings of the Computational Science–ICCS 2006, Reading, UK, 28–31 May 2006; Lecture Notes in Computer Science, 3994. Alexandrov, V.N., van Albada, G.D., Sloot, P.M.A., Dongarra, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2006.
14. Li, S.; Ge, Y. Crude Oil Price Prediction Based on a Dynamic Correcting Support Vector Regression Machine. *Abstr. Appl. Anal.* **2013**, *2013*, 528678.
15. Fan, L.; Pan, S.; Li, Z.; Li, H. An ICA-based support vector regression scheme for forecasting crude oil prices. *Technol. Forecast. Soc. Chang.* **2016**, *112*, 245–253. [[CrossRef](#)]
16. Li, T.; Zhou, M.; Guo, C.; Luo, M.; Wu, J.; Pan, F.; Tao, Q.; He, T. Forecasting Crude Oil Price Using EEMD and RVM with Adaptive PSO-Based Kernels. *Energies* **2016**, *9*, 1014. [[CrossRef](#)]
17. Yu, L.; Zhang, X.; Wang, S. Assessing Potentiality of Support Vector Machine Method in Crude Oil Price Forecasting. *EURASIA J. Math. Sci. Technol. Educ.* **2017**, *13*, 7893–7904. [[CrossRef](#)]
18. Li, T.; Zhou, Y.; Li, X.; Wu, J.; He, T. Forecasting Daily Crude Oil Prices Using Improved CEEMDAN and Ridge Regression-Based Predictors. *Energies* **2019**, *12*, 3603. [[CrossRef](#)]
19. Hu, Y.; Trafalis, T.B. New kernel methods for asset pricing: Application to natural gas price prediction. *Int. J. Financ. Mark. Deriv.* **2011**, *2*, 106–120. [[CrossRef](#)]
20. Su, M.; Zhang, Z.; Zhu, Y.; Zha, D. Data-Driven Natural Gas Spot Price Forecasting with Least Squares Regression Boosting Algorithm. *Energies* **2019**, *12*, 1094. [[CrossRef](#)]
21. Suykens, J.A.K.; Vandewalle, J. Least squares support vector machine classifiers. *Neural Process. Lett.* **1999**, *9*, 293–300. [[CrossRef](#)]
22. Gavrishchaka, V.V.; Ganguli, S.B. Volatility forecasting from multiscale and high-dimensional market data. *Neurocomputing* **2003**, *55*, 285–305. [[CrossRef](#)]
23. Perez-Cruz, F.; Afonso-Rodriguez, J.; Giner, J. Estimating GARCH Models Using Support Vector Machines. *Quant. Financ.* **2003**, *3*, 163–172. [[CrossRef](#)]
24. Gavrishchaka, V.V.; Banerjee, S. Support vector machine as an efficient framework for stock market volatility forecasting. *Comput. Manag. Sci.* **2006**, *3*, 147–160. [[CrossRef](#)]
25. Chen, S.; Härdle, W.K.; Jeong, K. Forecasting volatility with support vector machine-based GARCH model. *J. Forecast.* **2010**, *29*, 406–433. [[CrossRef](#)]
26. Ou, P.; Wang, H. Financial Volatility Forecasting by Least Square Support Vector Machine Based on GARCH, EGARCH and GJR Models: Evidence from ASEAN Stock Markets. *Int. J. Econ. Financ.* **2010**, *2*, 51–64. [[CrossRef](#)]
27. Bildirici, M.; Ersin, O.O. Support Vector Machine GARCH and Neural Network GARCH Models in Modeling Conditional Volatility: An Application to Turkish Financial Markets. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2227747 (accessed on 15 October 2020).
28. Geng, L.-Y.; Yu, F. Forecasting Stock Volatility using LSSVR-based GARCH Model Optimized by Siwpsso Algorithm. *J. Appl. Sci.* **2013**, *13*, 5132–5137. [[CrossRef](#)]
29. Santamaría-Bonfil, G.; Frausto-Solís, J.; Vázquez-Rodarte, I. Volatility forecasting using support vector regression and a hybrid genetic algorithm. *Comput. Econ.* **2015**, *45*, 111–133. [[CrossRef](#)]
30. Bezerra, P.C.S.; Albuquerque, P.H.M. Volatility forecasting via SVR–GARCH with mixture of Gaussian kernels. *Comput. Manag. Sci.* **2017**, *14*, 179–196. [[CrossRef](#)]
31. Bezerra, P.C.S.; Albuquerque, P.H.M. Volatility Forecasting: The Support Vector Regression Can Beat the Random Walk. *Econ. Comput. Econ. Cybern. Stud. Res.* **2019**, *4*, 115–126.
32. Peng, Y.; Albuquerque, P.H.; de Sá, J.M.C.; Padula, A.J.A.; Montenegro, M.R. The best of two worlds: Forecasting high frequency volatility for cryptocurrencies and traditional currencies with support vector regression. *Expert Syst. Appl.* **2018**, *97*, 177–192. [[CrossRef](#)]
33. Gong, X.-L.; Liu, X.-H.; Xiong, X.; Zhuang, X.-T. Forecasting stock volatility process using improved least square support vector machine approach. *Soft Comput.* **2019**, *23*, 11867–11881. [[CrossRef](#)]
34. Bollerslev, T. Generalised Autoregressive Conditional Heteroskedasticity. *J. Econom.* **1986**, *31*, 307–327. [[CrossRef](#)]
35. Nelson, D.B.; Cao, C.Q. Inequality Constraints in the Univariate GARCH Model. *J. Bus. Econ. Stat.* **1992**, *10*, 229–235.
36. Bollerslev, T. A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *Rev. Econ. Stat.* **1987**, *69*, 542–547. [[CrossRef](#)]
37. Nelson, D.B. Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica* **1991**, *59*, 347–370. [[CrossRef](#)]
38. Glosten, L.R.; Jagannathan, R.; Runkle, D.E. On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *J. Financ.* **1993**, *48*, 1779–1801. [[CrossRef](#)]
39. Ding, Z.; Granger, C.W.J.; Engle, R.F. A Long Memory Property of Stock Market Returns and a New Model. *J. Empir. Financ.* **1993**, *1*, 83–106. [[CrossRef](#)]
40. Zakoian, J.M. Threshold heteroskedastic models. *J. Econ. Dyn. Control* **1994**, *18*, 931–955. [[CrossRef](#)]
41. Taylor, S.J. *Modelling Financial Time Series*; Wiley: Chichester, UK, 1986.
42. Schwert, W. Stock volatility and the crash of '87. *Rev. Financ. Stud.* **1990**, *3*, 77–102. [[CrossRef](#)]
43. Higgins, M.; Bera, A. A class of nonlinear arch models. *Int. Econ. Rev.* **1992**, *33*, 137–158. [[CrossRef](#)]
44. Geweke, J. Modelling the Persistence of Conditional Variances: A Comment. *Econom. Rev.* **1986**, *5*, 57–61. [[CrossRef](#)]
45. Pentula, S. Modelling the Persistence of Conditional Variances: A Comment. *Econom. Rev.* **1986**, *5*, 71–74.

46. Engle, R.F.; Bollerslev, T. Modelling the Persistence of Conditional Variances. *Econom. Rev.* **1986**, *5*, 1–50. [[CrossRef](#)]
47. Engle, R.F.; Lilien, D.M.; Robins, R.P. Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model. *Econometrica* **1987**, *55*, 391–407. [[CrossRef](#)]
48. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222. [[CrossRef](#)]
49. Cherkassky, V.; Ma, Y. Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Netw.* **2004**, *17*, 113–126. [[CrossRef](#)]
50. Lee, S.; Kim, C.K.; Lee, S. Hybrid CUSUM Change Point Test for Time Series with Time-Varying Volatilities Based on Support Vector Regression. *Entropy* **2020**, *22*, 578. [[CrossRef](#)] [[PubMed](#)]
51. Vapnik, V.N. *The Nature of Statistical Learning Theory*; Springer: Berlin/Heidelberg, Germany, 1995.
52. Martínez-Álvarez, F.; Troncoso, A.; Asencio-Cortés, G.; Riquelme, J.C. A Survey on Data Mining Techniques Applied to Electricity-Related Time Series Forecasting. *Energies* **2015**, *8*, 13162–13193. [[CrossRef](#)]
53. Peng, L.-L.; Fan, G.-F.; Huang, M.-L.; Hong, W.-C. Hybridizing DEMD and Quantum PSO with SVR in Electric Load Forecasting. *Energies* **2016**, *9*, 221. [[CrossRef](#)]
54. Lux, M.; Härdle, W.K.; Lessmann, S. Data driven value-at-risk forecasting using a SVR-GARCH-KDE hybrid. *Comput. Stat.* **2020**, *35*, 947–981. [[CrossRef](#)]
55. Andersen, T.G.; Bollerslev, T. Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts. *Int. Econ. Rev.* **1998**, *39*, 885–905. [[CrossRef](#)]
56. Mapa, D. A Range-Based GARCH Model for Forecasting Volatility. *Philipp. Rev. Econ.* **2003**, *60*, 73–90.
57. Chou, R.Y. Forecasting Financial Volatilities with Extreme Values: The Conditional Autoregressive Range (CARR) Model. *J. Money Credit Bank.* **2005**, *37*, 561–582. [[CrossRef](#)]
58. Liu, H.C.; Hung, J.C. Forecasting Volatility and Capturing Downside Risk of the Taiwanese Futures Markets under the Financial Tsunami. *Manag. Financ.* **2010**, *36*, 860–875. [[CrossRef](#)]
59. Patton, A.J. Volatility Forecast Comparison using Imperfect Volatility Proxies. *J. Econom.* **2011**, *160*, 246–256. [[CrossRef](#)]
60. Molnár, P. High-low range in GARCH models of stock return volatility. *Appl. Econ.* **2016**, *48*, 4977–4991. [[CrossRef](#)]
61. Fiszeder, P. Low and high prices can improve covariance forecasts: The evidence based on currency rates. *J. Forecast.* **2018**, *37*, 641–649. [[CrossRef](#)]
62. Fiszeder, P.; Fałdziński, M. Improving Forecasts with the Co-Range Dynamic Conditional Correlation Model. *J. Econ. Dyn. Control* **2019**, *108*, 103736. [[CrossRef](#)]
63. Fiszeder, P.; Fałdziński, M.; Molnár, P. Range-Based DCC Models for Covariance and Value-at-Risk Forecasting. *J. Empir. Financ.* **2019**, *54*, 58–76. [[CrossRef](#)]
64. Wu, X.; Hou, X. Forecasting volatility with component conditional autoregressive range model. *N. Am. J. Econ. Financ.* **2020**, *51*, 101078. [[CrossRef](#)]
65. Parkinson, M. The Extreme Value Method for Estimating the Variance of the Rate of Return. *J. Bus.* **1980**, *53*, 61–65. [[CrossRef](#)]
66. Degiannakis, S.; Livada, A. Realized Volatility or Price Range: Evidence from a Discrete Simulation of the Continuous Time Diffusion Process. *Econ. Model.* **2013**, *30*, 212–216. [[CrossRef](#)]
67. Garman, M.B.; Klass, M.J. On the Estimation of Security Price Volatilities from Historical Data. *J. Bus.* **1980**, *53*, 67–78. [[CrossRef](#)]
68. Rogers, L.C.G.; Satchell, S.E. Estimating Variance From High, Low and Closing Prices. *Ann. Appl. Probab.* **1991**, *1*, 504–512. [[CrossRef](#)]
69. Shu, J.; Zhang, J.E. Testing Range Estimators of Historical Volatility. *J. Futures Mark.* **2006**, *26*, 297–313. [[CrossRef](#)]
70. Alizadeh, S.; Brandt, M.; Diebold, F.X. Range-Based Estimation of Stochastic Volatility Models. *J. Financ.* **2002**, *57*, 1047–1091. [[CrossRef](#)]
71. Alterman, S. *Natural Gas Price Volatility in the UK and North America*; NG 60; Oxford Institute for Energy Studies: Oxford, UK, 2012.
72. Hwang, S.; Valls Pereira, P.L. The effects of structural breaks in ARCH and GARCH parameters on persistence of GARCH models. *Commun. Stat.-Simul. Comput.* **2006**, *37*, 571–578. [[CrossRef](#)]
73. Hansen, P.R. A Test for Superior Predictive Ability. *J. Bus. Econ. Stat.* **2005**, *23*, 365–380. [[CrossRef](#)]
74. Hansen, P.R.; Lunde, A.; Nason, J.M. The Model Confidence Set. *Econometrica* **2011**, *79*, 453–497. [[CrossRef](#)]