Transactions on Quantitative Finance and Beyond

www.tqfb.reapress.com

Trans. Quant. Fin. Bey. Vol. 2, No. 1 (2025) 30–42.

Paper Type: Original Article

Evaluating and Forecasting Conventional Gasoline Price Fluctuations Using Garch Models with Two Distributions and Machine Learning Methods

Reza Roshanpour[1,*], Mohammadreza Parsanejad[1], Sorya Asgari[2], Fatemeh Hashtroodi Mahmoodi[2]

[1] Departmant of Management, Economics and Progress Engineering, Iran University Science and Technology, Tehran, Iran; roshanpour_reza@alumni.iust.ac.ir; parsanejad@iust.ac.ir.

[2] Department of Accounting, Economic and Accounting Faculty, Central Tehran Branch, Islamic Azad University, Tehran, Iran; soryaasgari@yahoo.com; f.h.mahmoodi@gmail.com.

Citation:

Received: 02 June 2024 *Revised: 15 July 2024* *Accepted: 13 October 2024*	Roshanpour, R., Parsanejad, M., Asgari, S., & Hashtroodi Mahmoodi, F. (2025). Evaluating and forecasting conventional gasoline price fluctuations using garch models with two distributions and machine learning methods. *Transactions on quantitative finance and beyond, 2*(1), 30-42.

Abstract

Conventional gasoline price can affect the government and society as a strategic commodity in the community. Conventional gasoline price fluctuations have economic, political, social, cultural, and environmental effects. Thus, the prediction of its volatility is essential but there is not any study to examine the price fluctuations. This study aims to hybridize and propose different Garch models based on two distributions and various algorithms in machine learning, such as random forest, ridge regression, Support Vector Regression (SVR), and elastic-net for predicting weekly gasoline price volatility. The results depict Garch and GJRgarch models based on t-student distribution can predict volatility. The combination of ridge regression and GJRgarch model can better predict volatility for the seven-step-ahead. The RMSE scale has been used to compare results that the scale value is 0.01475 in the hybrid method. In fact, combining the ridge regression with t-student-GJRgarch model has the slightest error prediction or the most accuracy among different Garch models and machine learning algorithms.

Keywords: Garch models, Machine learning, Gasoline price, Volatility, Distribution.

1 | Introduction

Fuel prices are highly volatile, creating significant challenges and uncertainties for the global economy. gasoline is a strategic product in the household goods basket. Whether households like it or not, it affects them [1]. The fluctuation in gasoline prices effectively influences society's welfare level [2], [3]. In other words, the volatility of gasoline price has two effects; the first effect is to affect the price of transportation, and the

Corresponding Author: roshanpour_reza@alumni.iust.ac.ir
10.22105/tqfb.v2i1.48

second effect is to increase the price of other goods, and therefore, inflation enlarges [4]. Also, gasoline price shocks can change one-year household inflation expectations [5]. It can affect the consumption of society and even the policies of policymakers and investors [6].

In this context, research has been performed on the long-run relation between gasoline price and other factors, including taxation, exchange rate, and crude price variation by the NARDL model [7], [8]. Also, Error Correction Models (ECM) have been utilized to predict retailed gasoline prices in China from 2019 to 2050 based on international crude oil prices [9]. Another study examined price transmission from crude oil to gasoline [10]. The hybrid method introduced by Escribano and Wang [11] contains random forest and cointegrated with international oil prices and exchange rates to weekly predicted gasoline prices. Moreover, considering asymmetry with structural beaks was examined in the Russian gasoline market, and the results revealed no asymmetry between the gasoline prices and the crude oil prices in the long run [12].

Regarding causality and volatility spillovers, Hammoudeh et al. [13] have investigated petroleum prices, gasoline, and heating oil in different locations and obtained results demonstrating the cointegration tests of each of them offer that spot and futures contracts suggest little benefits for long-run commodity portfolio diversification. In this regard, it is vital to forecast conventional gasoline prices accurately. However, no studies examine the volatility of traditional gasoline prices. Most studies have investigated the volatility of other commodities, including crude oil price and gas. Several articles have addressed just modeling and forecasting oil prices and tried to increase the prediction accuracy by various methods [14–16]. Regarding the new method of predicting, a comparison has also been performed between AdaBoost-LSTM and AdaBoost-GRU to make better forecasting performance for oil price prediction, which the AdaBoost-GRU is superior to predict [17]. Another research utilized the MRN networks to predict, and results showed that this method could be flexible for forecasting at various horizons [18].

Regarding optimization algorithms, the LSTM approach based on the Salp Swarm Algorithm (SSA) was utilized to predict [19]. Also, a dynamic ensemble learning based on the nondominated sorting genetic algorithm II was used to calculate a prediction for crude oil [20]. In summary, by reviewing energy prices, including natural gas, crude oil, electricity, and carbon in the systematic decade review, the researches are as follows: to hybridize basic model, data cleaning method, and optimizer [21]. Garch models have performed different studies under different distributions, including the volatility of the S&P 500 index and the evaluation of daily returns in the Istanbul stock exchange by Garch et.al methods with ANN [22], [23].

This paper's innovation is forecasting the volatility of weekly gasoline prices using the family of Autoregressive Conditional Heteroscedasticity (ARCH) models based on two distributions and various machine learning algorithms. In other words, this paper aims to fill the gap in the research on gasoline price volatility.

Since fluctuation is an unobservable variable, it has to be estimated, and then conditional variance or fluctuation is combined into machine learning algorithms as an input. To address the volatility of gasoline prices, different models of Garch have been utilized, such as GARCH, GJRGARCH, and EGARCH [24–26]. The remainder of this paper is organized as follows. Section 2 presents the description of data and the statistical characteristics of sample data. Section 3 discusses a general description of GARCH classes and explains various algorithms in machine learning. Section 4 contains the results, and the final section, Section 5, includes some concluding remarks.

2 | Data

The data in this research consists of 1932 weekly price observations of weekly U.S. Gulf coast conventional gasoline regular spot price FOB (Dollars per gallon) from 1986/6/6 to 2023/9/6. The data has been downloaded from the U.S. energy information admiration. Descriptive statistics for the conventional gasoline price have been summarized in *Table 1*. It shows a significant difference between the maximum and minimum values, which depict high volatility in the series. The mean and the corresponding data variance are small, but the kurtosis in the table is so high that the price distribution is not normally distributed. Also, the Jarque-Bera test is rejected because the probability of the null hypothesis of normality in the data distribution is less than 0.05.

Table 1. Descriptive statistics for conventional gasoline price.

Descriptive Statistics	Value
Mean	1.332370
Median	1.104500
Maximum	4.253000
Minimum	0.277000
Std. Dev.	0.844510
Skewness	0.686776
Kurtosis	2.346002
Jarque-Bera	186.3060
Probability	0.000000

We also assessed the data graph, and the skewness is positive in *Fig. 1*. It demonstrates a red curve when the data are not distributed symmetrically to the left and right sides of the mean on a bell curve. In other words, the distribution is positively skewed when its tail is more pronounced on the right side than on the left.

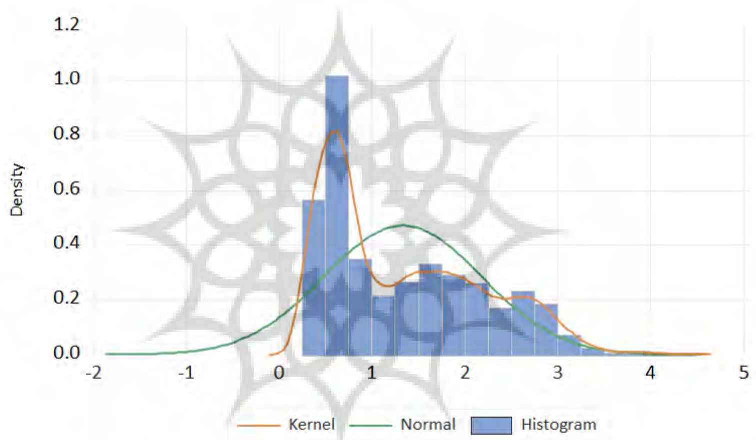


Fig. 1. The graph of descriptive statistics.

We applied the Augment Dickey-Fuller (ADF) test to examine the unit root test and study stationary data. According to the ADF test, ProbProb. > 0.05, and the null hypothesis of stationary is rejected *Table 2*. In other words, the ADF test's p-value is more than the significant level (0.05), which the null hypothesis will reject. Therefore, outputs infer that the time series is not stationary, and it has to take the first difference to stationary the data.

Table 2. Stationary test.

		T-Statistic	Prob.*
Augmented diskey-fuller test statistic		-0.057134	0.9521
Test critical values	1% level	-3.433542	
	5% level	-2.862836	
	10% level	-2.567507	

Fig. 2 shows that the raw data are not stationary, and *Fig. 3* represents the raw data after the first difference.

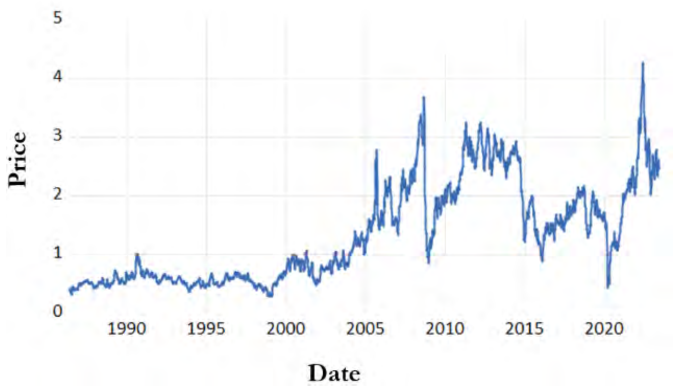


Fig. 2. The raw data of gasoline price (None-stationary).

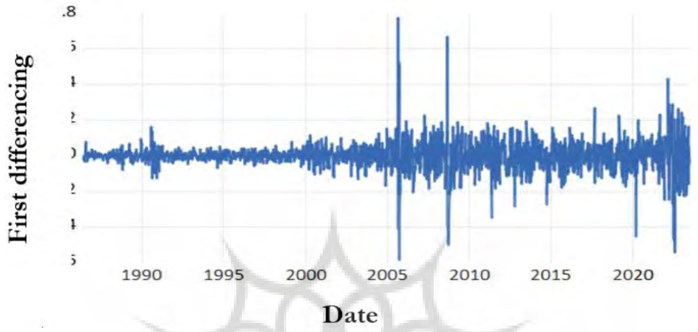


Fig. 3. The raw data of gasoline price (Stationary with taking the first difference).

3 | Methodology

3.1 | Garch Models

The purpose of the garch model is to model the variance when the variance of the residuals is not constant in a sequence of values. The fundamental assumption of a linear regression model is equal variance, but this assumption is rejected in different models. The behavior of values is non-linear, chaotic, and dynamic. The conditional variance equations of the three models have been described in *Eqs. (1)-(3)*. In each equation, $\widehat{\sigma}_t^2$ refers to the conditional variance, ε_{t-1} is the term of the error of the first lag, and $\alpha_0, \alpha_i, \beta_j, \gamma_i$ are the parameters to be estimated from the maximization of the sample log-likelihood function.

3.1.1 | GARCH (p,q)

$$\widehat{\sigma}_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i\, \varepsilon_{t-1}^2 + \sum_{j=1}^{q} \beta_j\, \sigma_{j-1}^2 \tag{1}$$

$\widehat{\sigma}_t^2$ is the conditional variance to which a squared ARCH (ε_{t-1}^2) is added. In other words, the conditional variance of the variable depends on the previous lags and the squared lag residuals. The Garch model has some limitations because the parameters must be positive in the conditional equation, and $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j$ are expected to be less than one [27].

3.1.2 | GJR-garch (p,q)

$$\widehat{\sigma}_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \gamma_i \varepsilon_{t-i}^2 d_{t-1} + \sum_{j=1}^{q} \beta_j\, \widehat{\sigma}_{t-j}^2 \tag{2}$$

Where d_{t-1} is a dummy variable:

$$d_{t-1} = \begin{cases} 1, & \text{if } \varepsilon_{t-1}^2 < 0, \\ 0, & \text{if } \varepsilon_{t-1}^2 \geq 0 \end{cases}$$

And γ refers to the coefficient that measures the impact of news. The rest of the parameters in the equation remain the same as those of the GARCH model.

3.1.3 | EGARCH (p,q)

$$\log \hat{\sigma}_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| + \sum_{k=1}^{r} \gamma_k \frac{u_{t-k}}{\sigma_{t-k}} + \sum_{j=1}^{q} \beta_j \log \hat{\sigma}_{t-j}^2 \tag{3}$$

ε_{t-1} is the term of the error of the first lag. Unlike the Garch model, the EGarch model has no restrictions [28]. In the equation, γ_k shows the leverage effects that depict the asymmetry of the model. If $\gamma_k \neq 0$, the effect is asymmetric, and if $\gamma_k = 0$, the effect is symmetric. In this regard, if $\gamma_k < 0$, it shows the existence of a leverage effect. It means that bad news reinforces the volatility. Conversely, if $\gamma_k > 0$, it indicates the presence of leverage effect, and it denotes that negative shocks at time $t-1$ have a weaker impact on the variance at time t than positive shocks.

3.2 | Gaussian Distribution

When estimating Garch models, the Gaussian or normal distribution is used. One must maximize the log-likelihood function to estimate the conditional mean, variance, and density function [29].

3.3 | Student's Distribution

Student or T-student distribution is utilized when financial time series have fat tails. The T-student distribution corresponds to the normal distribution when the degree of freedom is infinite [30].

To define the orders of the model, this process is made by considering the AC and the PAC for each case *Table 3*. Regarding the AC and the PAC for the data,

Table 3. The correlogram of date.

Auto Correlation	Partial Correlation		AC	PAC	Q-State	Prop
		1	0.130	0.130	32.606	0.000
		2	-0.016	-0.033	33.071	0.000
		3	0.083	0.091	46.393	0.000
		4	0.068	0.046	55.402	0.000
		5	-0.029	-0.040	57.030	0.000
		6	-0.025	-0.020	58.247	0.000
		7	-0.030	-0.036	59.959	0.000
		8	0.011	0.021	60.178	0.000

The best model will be chosen based on the outputs of this analysis. However, to increase the accuracy of the analysis, the information criteria have been considered. Table 4 depicts each case, which presents the lowest value of information criteria as the best one.

Table 4. Information criteria.

Model	LOGL	AIC*	BIC	HQ
(4,3)(0,0)	2160.007319	-2.227869	-2.201928	-2.218327
(0,4)(0,0)	2155.349385	-2.226152	-2.208858	-2.219790
(1,4)(0,0)	2156.269186	-2.226069	-2.205892	-2.218647

Table 4. Continued.

Model	LOGL	AIC*	BIC	HQ
(4,4)(0,0)	2158.627311	-2.225404	-2.196580	-2.214802
(3,2)(0,0)	2155.519179	-2.225292	-2.205115	-2.217870
(4,1)(0,0)	2155.347538	-2.225114	-2.204938	-2.217692
(2,4)(0,0)	2156.336811	-2.225103	-2.2002044	-2.216621
(4,0)(0,0)	2153.911419	-2.224662	-2.207368	-2.218301
(3,4)(0,0)	2156.677029	-2.224420	-2.198478	-2.214878
(2,3)(0,0)	2154.634665	-2.224375	-2.204199	-2.216954
(4,2)(0,0)	21.55.531476	-2.224269	-2.201210	-2.215787

The order of ACs is four, and PACs is three. In this regard, (4,3) has the lowest value of information criteria, and all the inverse roots lie inside the unit circle *Fig. 5*. The residuals, shown in *Table 5*, are white noise and a positive point in the ARIMA (4,1,3) because there is no information in the residuals. Therefore, the ARIMA (4,1,3) has satisfied the stability conditions, and the error terms are white noise.

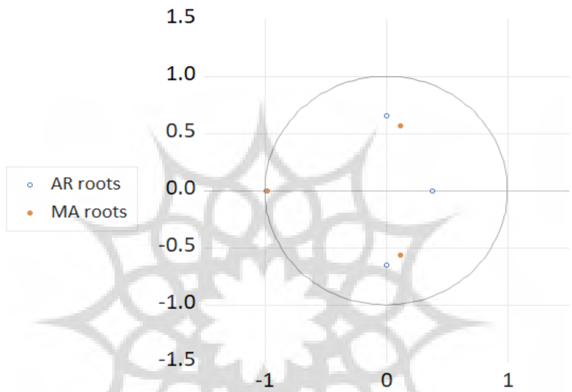
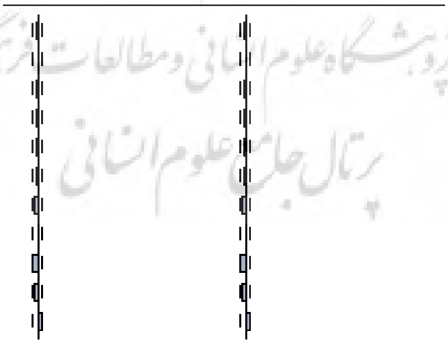


Fig. 4. Inverse roots of AR/MA polynomial(s).

Table 5. The correlogram of residuals in ARIMA (4,1,3).

Autocorrelation	Partial Correlation



To examine conditional heteroscedasticity, the null hypothesis of heteroskedasticity is rejected by ProbProb. <0.05. *Table 6* presents the data as not being homoscedastic. Thus, it has to model the volatility for the data.

Table 6. ARCH test results–Evidence of heteroskedasticity in the model.

Heteroskedasticity Test: ARCH			
F-statistic	422.5478	Prob. F(1,1928)	0.0000
Obs*R-squared	346.9477	Prob. chi-squre(1)	0.0000

The Garch models will be calculated to compare the results and examine fluctuations in *Tables 7-9*.

Table 7. GARCH Model Estimations – Estimated parameters of GARCH, GJR-GARCH, and EGARCH models under normal and t-student distributions.

Mean Equation	Distribution	Models					
		Garch	Prob	GJR-Garch	Prob	EGarch	Prob
AR(1)	Normal	-0.3628	(0.2755)	1.2355	(0.0013)	0.3984	(0.3808)
AR(2)		-0.5885	(0.0012)	-0.3869	(0.5668)	-0.5605	(0.2228)
AR(3)		-0.6744	(0.0477)	-0.3655	(0.4524)	-0.4607	(0.3751)
AR(4)		0.1759	(0.0316)	0.0511	(0.5382)	0.0931	(0.3505)
MA(1)		0.5750	(0.0801)	-1.0202	(0.0074)	-0.2014	(0.6546)
MA(2)		0.6885	(0.0000)	0.1376	(0.8139)	0.5079	(0.1611)
MA(3)		0.8216	(0.0069)	0.4650	(0.1871)	0.6013	(0.1788)
AR(1)	T-student	-0.4324	(0.0001)	-0.4318	(0.0001)	0.3363	(0.5516)
AR(2)		0.8763	(0.0000)	0.8730	(0.0000)	0.4889	(0.1149)
AR(3)		0.6114	(0.0000)	0.6123	(0.0000)	0.2033	(0.6067)
AR(4)		-0.1588	(0.0000)	-0.1578	(0.0000)	-0.0962	(0.2433)
MA(1)		0.6577	(0.0000)	0.6564	(0.0000)	-0.1186	(0.8340)
MA(2)		-0.7946	(0.0000)	-0.7936	(0.0000)	-0.5670	(0.0097)
MA(3)		-0.8090	(0.0000)	-0.8096	(0.0000)	-0.2815	(0.5230)

Table 8. Variance Equation Estimations–Estimated variance equations for GARCH, GJR-GARCH, and EGARCH models with different distributions.

Model	Distribution	Variance Equation
Garch	Normal	$\hat{\sigma}^2{}_t = 5.897 + 0.165*\varepsilon^2_{t-1} + 0.865*\sigma^2_{j-1}$
	Prob	(0.0317) (0.0000) (0.0000)
	T-student	$\hat{\sigma}^2{}_t = 8.335 + 0.128*\varepsilon^2_{t-1} + 0.885*\sigma^2_{j-1}$
	Prob	(0.0397) (0.0000) (0.0000)
GJRGarch	Normal	$\hat{\sigma}^2{}_t = 7.812 + 0.251*\varepsilon^2_{t-1} - 0.166*\varepsilon^2_{t-1}*d_{t-1} + 0.865*\sigma^2_{j-1}$
	Prob	(0.0036) (0.0000) (0.0000) (0.0000)
	T-student	$\hat{\sigma}^2{}_t = 9.399 + 0.174*\varepsilon^2_{t-1} - 0.099*\varepsilon^2_{t-1}*d_{t-1} + 0.887*\sigma^2_{j-1}$
	Prob	(0.0184) (0.0000) (0.0000) (0.0000)
EGarch	Normal	$\log \hat{\sigma}^2{}_t = -0.238 + 0.262*\frac{\lvert\varepsilon_{t-1}\rvert}{\sqrt{\sigma^2_{j-1}}} + 0.071*\frac{\varepsilon_{t-1}}{\sqrt{\sigma^2_{j-1}}} + 0.992*\mathrm{Log}(\sigma^2_{j-1})$
	Prob	(0.0000) (0.0000) (0.0000) (0.0000)
	T-student	$\log \hat{\sigma}^2{}_t = -0.233 + 0.232*\frac{\lvert\varepsilon_{t-1}\rvert}{\sqrt{\sigma^2_{j-1}}} + 0.061*\frac{\varepsilon_{t-1}}{\sqrt{\sigma^2_{j-1}}} + 0.990*\mathrm{Log}(\sigma^2_{j-1})$
	Prob	(0.0000) (0.0000) (0.0003) (0.0000)

Table 9. Model Selection Criteria – Comparison of GARCH models based on Akaike, Schwarz, and Hannan-Quinn information criteria.

Model	Distribution	Information Criteria			Maximum Likelihood
		Akaike	Schwarz	Hannan-Quinn	
Garch	Normal	-3.036588	-3.004828	-3.024905	2936.753
	T-student	-3.142173	-3.107526	-3.129427	3039.483
GJRGarch	Normal	-3.056684	-3.022038	-3.043939	2957.115
	T-student	-3.149621	-3.112087	-3.135813	3047.660
EGarch	Normal	-3.060373	-3.025726	-3.047628	2960.670
	T-student	-3.151873	-3.114339	-3.138065	3049.829

3.4 | Machine Learning Models

3.4.1 | Random forest

Random forest method based on decision trees is an ensemble learning method for classification and regression [31]. A random forest tree, which is based on a structure consisting of many decision trees, grows and aggregates regression trees for the predictions of each tree. This method is suitable for overfitting decision trees in the training set. Random forest has one stage more than bagging, which usually causes the performance of random forest to be better than that of the decision tree. The method utilizes bagging and random subspace to decrease the variance of the model [32]. The combination of bagging and random subspace helps to generate appropriate diversity in the model. In this regard, the random subspace method can increase diversity among trees by restricting trees to work on different random subsets of the predictor space, and bagging can examine additional diversity by constructing each tree in the forest from a bootstrap dataset sample.

3.4.2 | Ridge regression

The least square method aims to estimate variables when the relation between variables is linear. However, the least square estimates have high variance, which can become overfitting in the model. A small change in the training data can generate a large change in the coefficients [33]. Ridge regression estimates the coefficients in linear models where the independent variables are correlated [34]. This method can control the extent of coefficients by imposing a penalty for the error function. The penalty is a trade-off between the variance and bias that reduces the effects of multicollinearity and variance to increase prediction accuracy in the model.

3.4.3 | Support vector regression

Support Vector Regression (SVR) is a supervised machine-learning algorithm based on the Support Vector Machine [35]. This algorithm depicts sample data as points in space, and the points are separated using a hyperplane. Due to some characteristics of SVR, this method can be used for incomplete data and outliers [36]. The algorithm is a method to estimate a function that is mapped from an input to an output based on training data. The SVR considers kernel functions, which calculate the similarity among two observations in non-linear problems. SVR aims to map the vectors of regressor x onto a high-dimensional space using some fixed transformation.

3.4.4 | Elastic net

The elastic net method is a hybrid regression that utilizes penalties to improve the regularization of statistical models. Regularization contributes to solving the overfitting problem in the model. This procedure combines the lasso and ridge regression methods by learning to reduce loss function [37]. In this regard, two stages contain the lasso and regression algorithms to find the elastic net estimator. If first finds the ridge regression coefficients, perform the second step by utilizing a lasso sort of shrinkage of the coefficients. Also, the elastic net determines models that variable selection can be too dependent on data and unstable.

3.4.5 | Prediction performance metrics

To evaluate prediction accuracy, the Root Mean Square Error (RMSE) scale is utilized for examination, which is calculated as follows;

$$\mathrm{RMSE} = \sqrt{(1/n)\sum_{i=1}^{n}(y_i - \hat{y}_i)^2},$$

where $\hat{y}_i$ indicates the predicted value, and y_i depicts the actual value.

4 | Results

Table 8 indicates all the parameters are statistically significant in Garch, GJRGarch, and EGarch models. In other words, the parameters significantly differ from zero, highlighting the models' high validity. The best conditional volatility model is EGarch because EGarch indicates significant parameters and has the smallest information criteria while having the biggest maximum likelihood *Table 8* and *Table 9*. Also, the γ coefficient in the EGarch indicates the leverage effect. It shows the news effect in the model, and $\gamma > 0$ depicts that positive news is more destabilizing than negative news. The EGarch is exponential Garch; the parameters can be negative and positive. However, the residuals of EGarch with normal and T-student distribution have an Arch effect; hence, this model is refused *Table 12* and *Table 13*. Also, all estimations in the mean equation are insignificant for EGarch *Table 7*.

The γ coefficient in the GJRGarch is not equal to zero ($\gamma \neq 0$) with the normal and the T-student distribution. This means that the impact of news on the series is asymmetric. This model is accepted because this model has no Arch effect in the residuals in both the normal distribution *Table 10* and the T-student distribution *Table 11*. However, the normal-GJRGarch is rejected for the mean equation because its coefficient is insignificant. All coefficients are insignificant except Garch with T-student distribution and GJRGarch with T-student distribution in *Table 7*. Therefore, the mean equation is statistically significant in T-student-Garch and T-student-GJRGarch.

In the Garch model, whether the normal distribution is normal or not, all parameters are significant and positive. Also, the Garch model has no Arch effect *Table 14* and *Table 15*. The T-student Garch model is accepted because the mean equation is just as significant in the T-student-Garch model.

Table 10. ARCH LM test for GJRgarch–normal distribution.

F-Statistic	0.884100	Prob. F (1,1924)	0.3472
Obs*Rsquard	0.884613	Prob.chi-square(1)	0.3469

Table 11. ARCH LM test for GJRgarch–t-student distribution.

F-Statistic	1.316068	Prob. F (1,1924)	0.2514
Obs*Rsquard	1.316535	Prob.chi-square(1)	0.2512

Table 12. ARCH LM test for egarch-normal distribution.

F-Statistic	2.877158	Prob. F (1,1924)	0.0900
Obs*Rsquard	2.875845	Prob.chi-square(1)	0.0899

Table 13. ARCH LM test for egarch – t-student distribution.

F-Statistic	3.629835	Prob. F (1,1924)	0.0569
Obs*Rsquard	3.626766	Prob.chi-square(1)	0.0569

Table 14. ARCH LM test for garch – normal distribution.

F-Statistic	0.853707	Prob. F (1,1924)	0.3556
Obs*Rsquard	0.854216	Prob.chi-square(1)	0.3554

Table 15. ARCH LM test for garch – t-student distribution.

F-Statistic	1.381600	Prob. F (1,1924)	0.2400
Obs*Rsquard	1.382044	Prob.chi-square(1)	0.2398

All coefficients of both T-student-Garch and T-student-GJRGarch in the mean equation are significant. Therefore, Garch(1,1) and GJRGarch(1,1) are the best way to model the conditional fluctuations of conventional gasoline in the case of the T-student distribution. *Table 16* depicts the hybridized machine learning with Garch and GJRGarch; the RMSE is considered a criterion. The RMSE criterion to predict the seven-step-ahead in the primary model (Garch and GJRGarch) is equal. This means that the results are identical when forecasting seven-ahead in two models. However, with machine learning, the results are changed, and the criterion can decrease. In other words, machine learning can better forecast seven-step-ahead. The criterion to evaluate random forest, ridge regression, SVR, and elastic-net is RMSE, the slightest RMSE of ridge regression in the GJRGarch model. The ridge regression is the best to forecast conditional variance than the primary model and other machine learning algorithms *(Table 16)*. In this regard; the T-student-GJRGarch-ridge regression model has the most accuracy among algorithms to predict seven-step-ahead.

Table 16. Hybrid ML with GARCH & GJR-GARCH.

RMSE		Model	Distribution	Garch	GJRGarch
	Machine learning models	Random forest	t-student	0.01751	0.01541
		Ridge regression	t-student	0.01695	0.01475
		SVR	t-student	0.08862	0.08901
		Elastic-net	t-student	0.01853	0.01612
	Main model		t-student	0.08052	0.08052

Table 17. Index of notations and abbreviations.

AC	Autocorrelation
ADF	Augmented-dickey-fuller
AIC	Akaike information
AR	Autoregressive
ARCH	Autoregressive conditional heteroskedasticity
MA	Moving average
PAC	Partial correlation
Prob	P-value

5 | Conclusion

The trend of this search is to evaluate gasoline price volatility based on different Garch models and machine learning methods. This study presents an effective model to predict volatility. Seven of the hybrid models were performed in several stages. In the first stage, comparative research was conducted on fluctuations based on different Garch models with two distributions, Garch, EGarch, and GJRGarch. All models were based on T-student and normal distribution. All the Garch models considered are evaluated to predict fluctuations. Two models were selected based on economic theories, which were significant to estimating all parameters, including Garch and GJRGarch, based on T-student distribution. These models can better forecast volatility than other models. In the second stage, the mentioned models were utilized to predict the seven-step-ahead. The rose scale to predict the seven-step-ahead was 0.08052, which means the scale value was equal in the two models. In the next stage, the machine learning algorithms were used to predict the horizon, which was seven steps more precisely. Conditional variance was entered as an input to machine learning algorithms. The purpose of entering input into algorithms is to analyze volatility. The RMSE scale was also examined in all algorithms, and the scale was the slightest in all algorithms except the support machine vector. There were more RMSE metrics in SVR than in the primary model and other algorithms. The RMSE metrics in ridge regression were less than other algorithms and the primary model with a close look. The scale of ridge regression in GJRGarh was 0.01475, which was the slightest in all. Therefore, the hybrid model ridge regression-GJRGarch based on T-student distribution can better decrease prediction error to forecast volatility for seven-step-ahead. In other words, the proposed model is more accurate than another model and algorithms to predict the seven-step-ahead. Other Garch models based on different distributions and algorithms in machine learning can be investigated to evaluate volatility in future research.

Ethical statement

I. This material is the authors' original work, which has not been previously published elsewhere.

II. The paper is not currently being considered for publication elsewhere.

III. The paper reflects the authors' research and analysis wholly and truthfully.

IV. The results are appropriately placed in prior and existing research context.

V. All sources used are correctly disclosed (Correct citation).

VI. All authors have been personally and actively involved in substantial work leading to the paper, and will take public responsibility for its content.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to affect the work reported in this research. Also, the authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Acknowledgments

The authors wish to convey their appreciation to everyone and every institution that played a role in completing this study. We extend special gratitude to our colleagues for their insightful contributions and engaging discussions, and to the reviewers for their valuable feedback and recommendations.

Author Contribution

Every author made a meaningful contribution to the conception, design, and implementation of this research. Reza Roshanpour was in charge of collecting and analyzing the data, Mohammadreza Parsanejad assisted with the theoretical framework, while Sorya Asgari and Fatemeh Hashtroodi Mahmoodi participated in writing and revising the manuscript. All authors have reviewed and approved the final manuscript version.

Data Availability

For access to the dataset, please reach out to the corresponding author.

References

[1] Setyawan, D. (2014). The impacts of the domestic fuel increases on prices of the Indonesian economic sectors. *Energy procedia*, *47*, 47–55. https://doi.org/10.1016/j.egypro.2014.01.195

[2] Kpodar, K., & Liu, B. (2022). The distributional implications of the impact of fuel price increases on inflation. *Energy economics*, *108*, 105909. https://doi.org/10.1016/j.eneco.2022.105909

[3] Norouzia, N. (2019). Short-term and long-term effects of fuel carrier's price rise on consumer price inflation in Iran. *Technology transfer and entrepreneurship*, 1–13. https://B2n.ir/ww5979

[4] Jansen, D. J., & Jonker, N. (2018). Fuel tourism in Dutch border regions: Are only salient price differentials relevant? *Energy economics*, *74*, 143–153. https://doi.org/10.1016/j.eneco.2018.05.036

[5] Kilian, L., & Zhou, X. (2020). *Oil prices, gasoline prices and inflation expectations: A new model and new facts*. https://hdl.handle.net/10419/223588

[6] Rodrigues, N., Losekann, L., & Silveira Filho, G. (2018). Demand of automotive fuels in Brazil: Underlying energy demand trend and asymmetric price response. *Energy economics*, *74*, 644–655. https://doi.org/10.1016/j.eneco.2018.07.005

[7] Bagnai, A., & Ospina, C. A. M. (2015). Long-and short-run price asymmetries and hysteresis in the Italian gasoline market. *Energy policy*, *78*, 41–50. https://doi.org/10.1016/j.enpol.2014.12.017

[8] Bagnai, A., & Ospina, C. A. M. (2016). "Asymmetric asymmetries" in Eurozone markets gasoline pricing. *The journal of economic asymmetries*, *13*, 89–99. https://doi.org/10.1016/j.jeca.2016.03.004

[9] Ou, S., Lin, Z., Xu, G., Hao, X., Li, H., Gao, Z., Bouchard, J. (2020). The retailed gasoline price in China: Time-series analysis and future trend projection. *Energy*, *191*, 116544. https://doi.org/10.1016/j.energy.2019.116544

[10] Bastianin, A., Galeotti, M., & Manera, M. (2014). Forecasting the oil-gasoline price relationship: Do asymmetries help? *Energy economics*, *46*, S44-S56. https://doi.org/10.1016/j.eneco.2014.08.006

[11] Escribano, Á., & Wang, D. (2021). Mixed random forest, cointegration, and forecasting gasoline prices. *International journal of forecasting*, *37*(4), 1442–1462. https://doi.org/10.1016/j.ijforecast.2020.12.008

[12] Kolesnikova, A., & Fantazzini, D. (2021). Asymmetry and hysteresis in the Russian gasoline market: The rationale for green energy exports. *Energy policy*, *157*, 112466. https://doi.org/10.1016/j.enpol.2021.112466

[13] Hammoudeh, S., Li, H., & Jeon, B. (2003). Causality and volatility spillovers among petroleum prices of WTI, gasoline and heating oil in different locations. *The north american journal of economics and finance*, *14*(1), 89–114. https://doi.org/10.1016/S1062-9408(02)00112-2

[14] An, J., Mikhaylov, A., & Moiseev, N. (2019). Oil price predictors: machine learning approach. *International journal of energy economics and policy*, *9*(5), 1–6. https://www.econjournals.net.tr/index.php/ijeep/article/view/7597

[15] Gabralla, L. A., Jammazi, R., & Abraham, A. (2013). Oil price prediction using ensemble machine learning. In *2013 International conference on computing, electrical and electronic engineering (ICCEEE)* (pp. 674–679). IEEE. https://doi.org/10.1109/ICCEEE.2013.6634021

[16] Gumus, M., & Kiran, M. S. (2017). Crude oil price forecasting using XGBoost. In *2017 International conference on computer science and engineering (UBMK)* (pp. 1100-1103). IEEE. https://doi.org/10.1109/UBMK.2017.8093500

[17] Busari, G. A., & Lim, D. H. (2021). Crude oil price prediction: A comparison between AdaBoost-LSTM and AdaBoost-GRU for improving forecasting performance. *Computers & chemical engineering*, *155*, 107513. https://doi.org/10.1016/j.compchemeng.2021.107513

[18] Orojo, O., Tepper, J., McGinnity, T. M., & Mahmud, M. (2019). A multi-recurrent network for crude oil price prediction. In *2019 IEEE symposium series on computational intelligence (SSCI)* (pp. 2940-2945). IEEE. https://doi.org/10.1109/SSCI44817.2019.9002841

[19] Jovanovic, L., Jovanovic, D., Bacanin, N., Jovancai Stakic, A., Antonijevic, M., Magd, H., … & Zivkovic, M. (2022). Multi-step crude oil price prediction based on lstm approach tuned by salp swarm algorithm with disputation operator. *Sustainability*, *14*(21), 14616. https://doi.org/10.3390/su142114616

[20] Hao, J., Feng, Q., Yuan, J., Sun, X., & Li, J. (2022). A dynamic ensemble learning with multi-objective optimization for oil prices prediction. *Resources policy*, *79*, 102956. https://doi.org/10.1016/j.resourpol.2022.102956

[21] Lu, H., Ma, X., Ma, M., & Zhu, S. (2021). Energy price prediction using data-driven models: A decade review. *Computer science review*, *39*, 100356.https://doi.org/10.1016/j.cosrev.2020.100356

[22] Lahmiri, S., & Boukadoum, M. (2015). An ensemble system based on hybrid EGARCH-ANN with different distributional assumptions to predict S&P 500 intraday volatility. *Fluctuation and noise letters*, *14*(01), 1550001. https://doi.org/10.1142/S0219477515500017

[23] Bildirici, M., & Ersin, Ö. Ö. (2009). Improving forecasts of GARCH family models with the artificial neural networks: An application to the daily returns in Istanbul stock exchange. *Expert systems with applications*, *36*(4), 7355–7362. https://doi.org/10.1016/j.eswa.2008.09.051

[24] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the econometric society*, 987–1007. https://doi.org/10.2307/1912773

[25] Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, *48*(5), 1779–1801. https://doi.org/10.1111/j.1540-6261.1993.tb05128.x

[26] Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: journal of the econometric society*, 347–370. https://doi.org/10.2307/2938260

[27] Uğurlu, E. (2019). Research data analysis using EViews: An empirical example of modeling volatility. In *Research data access and management in modern libraries* (pp. 292–324). IGI Global. https://doi.org/10.4018/978-1-5225-8437-7.ch014

[28] Thomas, S., & Mitchell, H. (2005). GARCH modeling of high-frequency volatility in Australia's national electricity market. *System*, 1–39. https://B2n.ir/fy4358

[29] Cover, T. M. (1999). *Elements of information theory*. John Wiley & Sons. https://doi.org/10.1002/047174882X

[30] Hurst, S. (1995). *The characteristic function of the Student t distribution*. https://B2n.ir/fu2748

[31] Breiman, L. (2001). Random forests. machine learning. *Machine learning*, *45*, 5–32. https://doi.org/10.1023/A:1010933404324

[32] Dudek, G., Fiszeder, P., & Kobus Pawełand Orzeszko, W. (2024). Forecasting cryptocurrencies volatility using statistical and machine learning methods: A comparative study. *Applied soft computing*, *151*, 111132. https://doi.org/10.1016/j.asoc.2023.111132

[33] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112, No. 1). New York: Springer. https://doi.org/10.1007/978-3-031-38747-0

[34] Hilt, D. E., & Seegrist, D. W. (1977). *Ridge, a computer program for calculating ridge regression estimates* (Vol. 236). Department of agriculture, forest service, northeastern forest experiment. https://research.fs.usda.gov/treesearch/19260

[35] Vapnik, V., Golowich, S., & Smola, A. (1996). Support vector method for function approximation, regression estimation and signal processing. *Advances in neural information processing systems*, *9*. https://B2n.ir/rt8212

[36] Gavrishchaka, V. V, & Ganguli, S. B. (2003). Volatility forecasting from multiscale and high-dimensional market data. *Neurocomputing*, *55*(1–2), 285–305. https://doi.org/10.1016/S0925-2312(03)00381-3

[37] Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the royal statistical society series b: Statistical methodology*, *67*(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x