

## Article

# Forecasting Natural Gas Spot Prices with Machine Learning

Dimitrios Mouchtaris <sup>1</sup>, Emmanouil Sofianos <sup>2,\*</sup> , Periklis Gogas <sup>2</sup>  and Theophilos Papadimitriou <sup>2</sup><sup>1</sup> Faculty of Sciences, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece; dimmoucht@gmail.com<sup>2</sup> Department of Economics, Democritus University of Thrace, 69100 Komotini, Greece; pgkogkas@econ.duth.gr (P.G.); papadimi@econ.duth.gr (T.P.)

\* Correspondence: esofiano@econ.duth.gr

**Abstract:** The ability to accurately forecast the spot price of natural gas benefits stakeholders and is a valuable tool for all market participants in the competitive gas market. In this paper, we attempt to forecast the natural gas spot price 1, 3, 5, and 10 days ahead using machine learning methods: support vector machines (SVM), regression trees, linear regression, Gaussian process regression (GPR), and ensemble of trees. These models are trained with a set of 21 explanatory variables in a 5-fold cross-validation scheme with 90% of the dataset used for training and the remaining 10% used for testing the out-of-sample generalization ability. The results show that these machine learning methods all have different forecasting accuracy for every time frame when it comes to forecasting natural gas spot prices. However, the bagged trees (belonging to the ensemble of trees method) and the linear SVM models have superior forecasting performance compared to the rest of the models.

**Keywords:** natural gas; spot price; machine learning; forecasting



**Citation:** Mouchtaris, D.; Sofianos, E.; Gogas, P.; Papadimitriou, T. Forecasting Natural Gas Spot Prices with Machine Learning. *Energies* **2021**, *14*, 5782. <https://doi.org/10.3390/en14185782>

Academic Editor: Javier Reneses

Received: 28 July 2021

Accepted: 9 September 2021

Published: 14 September 2021

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



**Copyright:** © 2021 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

Natural gas has been proposed as a solution to increase the security of the energy supply and to reduce environmental pollution around the world. It is the second most widely used energy commodity after oil [1]. With the replacement of coal and the widespread use of natural gas, gas spot price forecasting has become one of the most critical issues in many sectors. The accurate forecasting of natural gas spot prices is of high importance, as these forecasts are used in the energy market, in power system planning and in regulatory decision making, covering both supply and demand in the natural gas market.

Due to the significant economic results obtained from forecasting, many techniques have been explored and studied, especially in electric load forecasting, such as artificial neural networks (ANN), as seen in [2] and SVM, as seen in [3] and many other works. The current studies on energy market forecasting mainly focus on crude oil prices [4]. Thus, publications in the field of natural gas price forecasting are relatively rare [1].

One of the few studies that has tried to directionally forecast natural gas price movements for the U.S. market is that of [5], which analyzed trader positions published on a weekly basis. [6] forecasted gas prices one day ahead, but they relied on monthly forward products and futures instead of focusing on current prices. They combined wavelet transform (WT) with fixed and adaptive machine learning/time series models: multi-layer perceptron (MLP), radial basis functions (RBF), linear regression, and GARCH (Generalized Autoregressive Conditional Heteroskedasticity). According to their results, the best models for electricity demand/gas price forecasting are the adaptive MLP/GARCH.

Another study analyzing gas prices is that of [7]. They trained several nonlinear models with the aid of a Gamma test: local linear regression (LLR), dynamic local linear regression (DLLR), and artificial neural networks (ANN). They used daily, weekly, and monthly Henry Hub spot prices from 1997 to 2012. They concluded that the forecasting model of daily spot prices using ANN can provide an accurate view. Moreover, ANN models have superior performance compared to LLR and DLLR models.

Ref. [8] tried to determine whether natural gas future prices can predict natural gas spot prices. They used daily observations for the spot and futures prices for natural gas for all trading days between 1 January 1997 and 3 March 2014 collected from the U.S. Energy Information Administration (EIA) for a total of 4294 observations. According to their results, gas futures prices are not superior in forecasting natural gas spot prices when compared to a random walk (RW) model.

Ref. [9] compared the long-horizon forecasting performance of traditional econometric models with machine learning methods (neural networks and random forests) for the main energy commodities in the world: oil, coal and gas. Their results showed that machine learning methods outperform traditional econometric methods and that they present an additional advantage, which is the ability to predict turning points.

Ref. [10] combined machine learning methodologies (XGboost, SVM, logistic regression, random forests, and neural networks) with dynamic moving windows and expanded windows to forecast crises in the U.S. natural gas market for a period spanning from 1994 to 2019. According to their results, the best forecasting accuracy was achieved with the XGboost combined with the dynamic moving window, reaching 49% accuracy and a false alarm of no more than 25%.

Ref. [11] presented a literature survey of the published papers forecasting natural gas prices, amongst others. According to their survey, predicting the exact future evolution of natural gas price is impossible.

According to the literature review, it can be observed that machine learning methodologies produce higher prediction accuracy compared to standard econometric methods. Therefore, in this paper we trained models that have the potential to successfully predict gas prices. The models trained in this paper are the support vector machines (SVM), regression trees, linear regression, Gaussian process regression (GPR), and ensemble of trees models. We focus on the short-term forecasting of the natural gas spot price 1, 3, 5, and 10 days ahead, and we compare the effectiveness of the machine learning models in natural gas price forecasting with a random walk model.

For the training of the models, we used the lags of the natural gas spot prices and a set of 21 explanatory variables that were selected based on the relevant literature (for instance, [1,8,12]) and determined their ability to enhance the predictive ability of natural gas price forecasting. The selected variables were then fed into the forecasting models through a training–testing learning process, resulting in the most efficient and least error-prone models for natural gas price forecasting.

The paper is organized as follows: in Section 2, we will briefly discuss the methodologies and the data used in our study, while in Section 3, we describe our empirical results. Finally, Section 4 will conclude the paper.

## 2. Methodology

### 2.1. Support Vector Machines

Support vector machines (SVM) are a set of methods for data classification and regression based on the maximization of the interclass distance: the basic concept of the SVM is to define the optimal (optimal in the sense of the model's generalization to unknown data) linear separator that separates the data points into two classes. To facilitate this, the algorithm employs the “kernel trick”: the initial data space is projected through a kernel function to a higher dimensional space (feature space) where the dataset may be linearly separable [13]. In this paper, we use four kernels, the linear, the quadratic, the cubic, and three different Gaussian kernels: fine, medium and coarse, following a different structure in the data each time.

### 2.2. Gaussian Process Regression

Gaussian processes are a flexible class of non-parametric machine learning models that are primarily used for modeling spatial and time series data. Gaussian models are commonly used to solve difficult machine learning problems. They are particularly

useful and attractive due to their flexible non-parametric nature and their computational simplicity. A common application of Gaussian processes is regression. Gaussian process regression (GPR) is based on the determination of an appropriate kernel function or a measure of similarity between data points whose locations are known. Compared to other machine learning methods, the advantages of GPR lie in its ability to seamlessly integrate multiple machine learning tasks, such as parameter estimation. Moreover, it has excellent performance and needs a relatively small training dataset to perform predictions. However, a known problem that arises is that due to the computational complexity of the predictions, according to [12], it becomes infeasible for GPR to be effective for large datasets. In this paper we trained four different GPR models coupled with the most important kernel functions with same length scale for each predictor:

- (1) Rational Quadratic GPR: a Gaussian process model that uses the rational quadratic kernel;
- (2) Squared Exponential GPR: a Gaussian process model that uses the squared exponential kernel;
- (3) Matern 5/2 GPR: a Gaussian process model that uses the matern 5/2 kernel;
- (4) Exponential GPR: a Gaussian process model that uses the exponential kernel.

### 2.3. Decision Trees

Ref. [14] proposed decision trees as a forecasting modeling technique in statistics, data mining, and machine learning. It employs a decision tree (as a forecasting model) to shift from observations of an item (represented by the branches) to inferences about the object's target value (represented in the leaves). Regression trees are decision trees in which the target variable can take continuous values (typically real numbers). In this paper we use three different tree models:

- (1) Fine Tree where the minimum leaf size is 4;
- (2) Medium Tree where the minimum leaf size is 12;
- (3) Coarse Tree: where the minimum leaf size is 36.

### 2.4. Ensemble of Trees

An ensemble of trees is formed by several individual trees that are added together. Although decision trees are one of the most efficient and interpretable classification algorithms, they suffer from low generalization ability nonetheless. Thus, they provide a low bias in-sample but a high variance out-of-sample. Ensemble techniques have been shown to solve this problem. They combine several decision trees to produce better prediction performance, as opposed to using a single decision tree. The basic principle underlying the ensemble model is that a group of weak learners is combined to form a strong learner. The main techniques for training ensemble decision tree models are bagging and boosting [15].

#### 2.4.1. Bagging

Bagging (bootstrap aggregation) is used when our goal is to reduce the variance of a decision tree. In this process, the basic idea is to generate several subsets of data from the training sample, which is selected randomly by replacement. Each subset of data is used to train the corresponding decision tree model. As a result, we end up with a set of different models. Finally, the average of all predictions obtained from different trees is used, which is more powerful and accurate than a single decision tree (Figure 1).

#### 2.4.2. Boosting

Boosting is another ensemble technique that aims to improve the accuracy of predictions generated by one or many models. This technique starts by fitting an initial model (e.g., a tree or linear regression) to the data. Then, a second model is constructed that focuses on accurately predicting cases where the first model does not perform well by using a weighted data sample. The combination of these two models is better than either individual model separately. The boosting process is then repeated several times. Each

successive model attempts to correct the weaknesses and errors of the combined boosted set of all of the previous models (Figure 2). Combining the entire set at the end converts the weak learners into a better performing model.

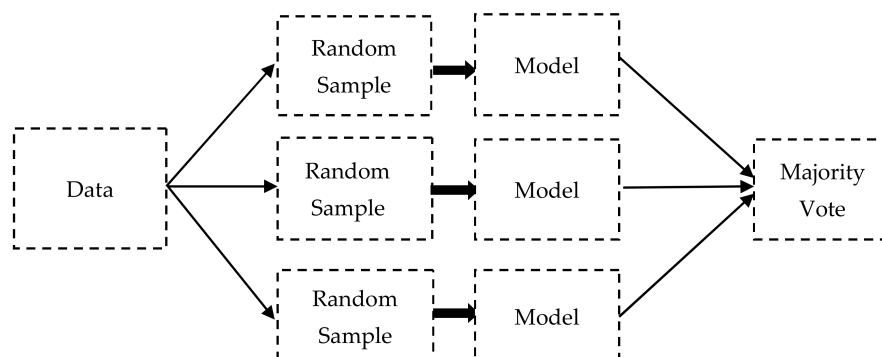


Figure 1. Bagging [16].

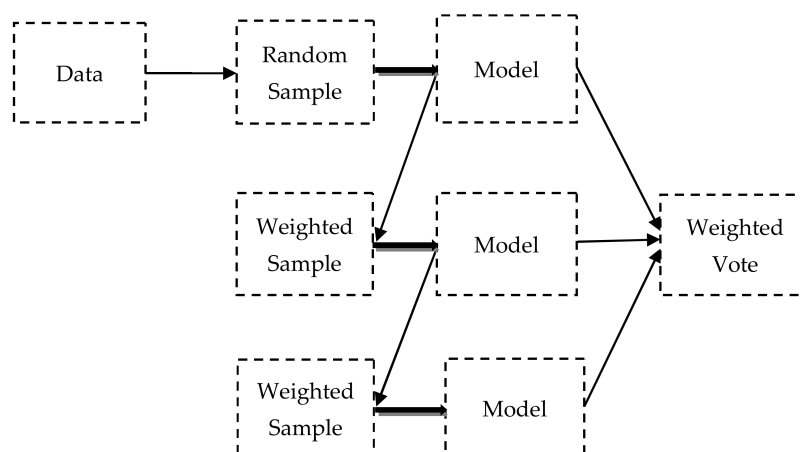


Figure 2. Boosting [16].

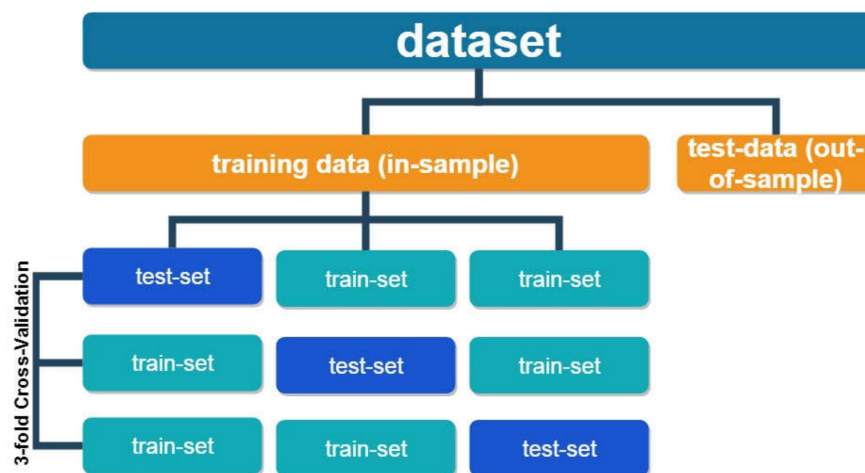
## 2.5. Cross-Validation

A common issue in this area of work is the problem of overfitting. A model can theoretically be conditioned to precisely fit the training data, hence exhibiting very high accuracy in-sample. Nonetheless, such a model would be useless in forecasting, as it will likely exhibit a low fit in the test (out-of-sample) data. In such cases, the model is trained to only fit the training data and not the underlying phenomenon. To avoid this, in the empirical part of the study, we employed a  $k$ -fold cross validation procedure. The in-sample data, which are used to train the model, are divided into  $k$  parts (folds) of equal size. Then, in each of the  $k$  iterations, one fold is used as the testing set, while the remaining  $k-1$  folds are used as the training set. This is repeated for all  $k$  folds. In this scheme, the model's accuracy is evaluated by the average performance over all of the  $k$  folds for each set of the model's parameters. Figure 3 provides a graphical representation of a 3-fold cross validation procedure.

## 2.6. The Dataset

For the training and the testing of our models, we compiled a dataset consisting of 2423 daily natural gas spot price values from the Energy Information Administration (EIA) database and 21 related economic variables from the Federal Reserve Bank of Saint Louis and Yahoo Finance databases. They span the period from 3 December 2010 to 18 September 2020 (Table 1). In addition, the momentum of the last 5 and 10 days (Momentum 5 and 10 are defined as the sum of the times that natural gas spot price increases

in the last 5 and 10 days, respectively) as well as the 5- and 10-day moving average was calculated and added to the independent variable set. With the exception of interest rates, all of the variables were converted to natural logarithms.



**Figure 3.** A three-fold cross validation for a given set of model parameter values. Each fold serves as a test sample, while the remaining folds are used to train the model. The average prediction accuracy for each set of parameters over the k-folds is used to assess the model [17].

**Table 1.** List of explanatory variables with mean, standard deviation, skewness, kurtosis, and variance.

#	Name	Mean	Standard Deviation	Skewness	Kurtosis	Variance
Panel A: Stock Indices						
1	NASDAQ Composite Index	5289.44	2132.48	0.61	−0.4	4,549,055
2	S&P 500 Index	2112.97	607.77	0.18	−0.92	369,500
3	Dow Jones Industrial Average Index	18,797.36	5195.66	0.3	−1.04	27,003,372
Panel B: Exchange Rates						
4	USD/EUR	0.19	0.09	0.3	−1.29	0.008
5	JPY/USD	4.62	0.14	−0.82	−0.61	0.019
6	USD/GBP	0.37	0.1	0.19	−1.5	0.010
Panel C: WTI Spot Price						
7	Cushing, OK WTI Spot Price FOB	4.17	0.37	−0.57	0.34	0.1401
Panel D: Interest Rates						
8	Effective Federal Funds Rate	0.638	0.77	1.17	−0.15	0.5974
9	5-Year Breakeven Inflation Rate	1.7	0.32	−0.84	1.57	0.107
10	10-Year Breakeven Inflation Rate	1.94	0.33	−0.44	0.36	0.1142
11	1-Year Treasury Constant Maturity Rate	0.75	0.82	1.1	−0.23	0.6757
12	10-Year Treasury Constant Maturity Rate	2.22	0.61	−0.4	0.51	0.3736
13	Bank Prime Loan Rate	3.76	0.75	1.18	−0.12	0.5732
Panel E: Future Contracts						
14	Natural Gas Futures Contract 1	1.1	0.26	−0.171	−0.51	0.0686
15	Natural Gas Futures Contract 2	1.11	0.25	−0.174	−0.62	0.0626
16	Natural Gas Futures Contract 3	1.13	0.24	−0.163	−0.69	0.0573
17	Natural Gas Futures Contract 4	1.15	0.23	−0.096	−0.75	0.0519
18	OK Crude Oil Future Contract 1	4.181	0.371	−0.53	0.15	0.138
19	OK Crude Oil Future Contract 2	4.189	0.358	−0.37	−0.45	0.1283
20	OK Crude Oil Future Contract 3	4.195	0.348	−0.27	−0.81	0.1213
21	OK Crude Oil Future Contract 4	4.199	0.341	−0.22	−0.95	0.1165

In order to test the generalization ability of the trained models, the dataset was divided into two parts: the first 90% was used as the training data set (in-sample, consisting of 2180 observations), and the remaining 10% of the most recent observations was the test data set (out-of-sample, consisting of 243 observations).

### 3. Empirical Results

The prediction accuracy of each model for both the out-of-sample and in-sample data was measured using the Root Mean Square Error (RMSE) metric. Thus, the optimal model was selected as the one that minimizes the RMSE:

$$\text{RMSE} = \frac{\sqrt{\sum_{t=1}^T (\hat{y}_t - y_t)^2}}{T} \quad (1)$$

where  $\hat{y}$  = the forecasted value,  $y$  = the actual value, and  $T$  = the number of observations.

Our forecasts were produced for several alternative forecasting horizons, i.e.,  $t + 1$ ,  $t + 3$ ,  $t + 5$ , and  $t + 10$ . We completed the same task with a random walk model in order to compare our machine learning results to a naïve prediction model.

Before moving to structural models (The ones that include the independent variables of our data set.), we first tried to identify the best autoregressive representation, i.e., to produce the best AR( $q$ ) model (autoregressive model). The AR( $q$ ) model is a simple model that uses past (lagged) values of natural gas spot prices to forecast the future natural gas spot price.

$$X_t = c + \sum_{i=1}^q \varphi_i X_{t-i} + \varepsilon_t \quad (2)$$

where  $X$  is the natural gas spot price,  $q$  is the maximum number of lags, and  $\varphi_i$  the parameter vector of the lags to be estimated.

In order to identify the optimal number of lags, we train several linear SVM models by varying the number of lags we used each time, starting with an AR(1) up to an AR(15).

We concluded that by using the first 14 lags, we minimize the in-sample RMSE (0.04196). These results are presented in Figure 4.

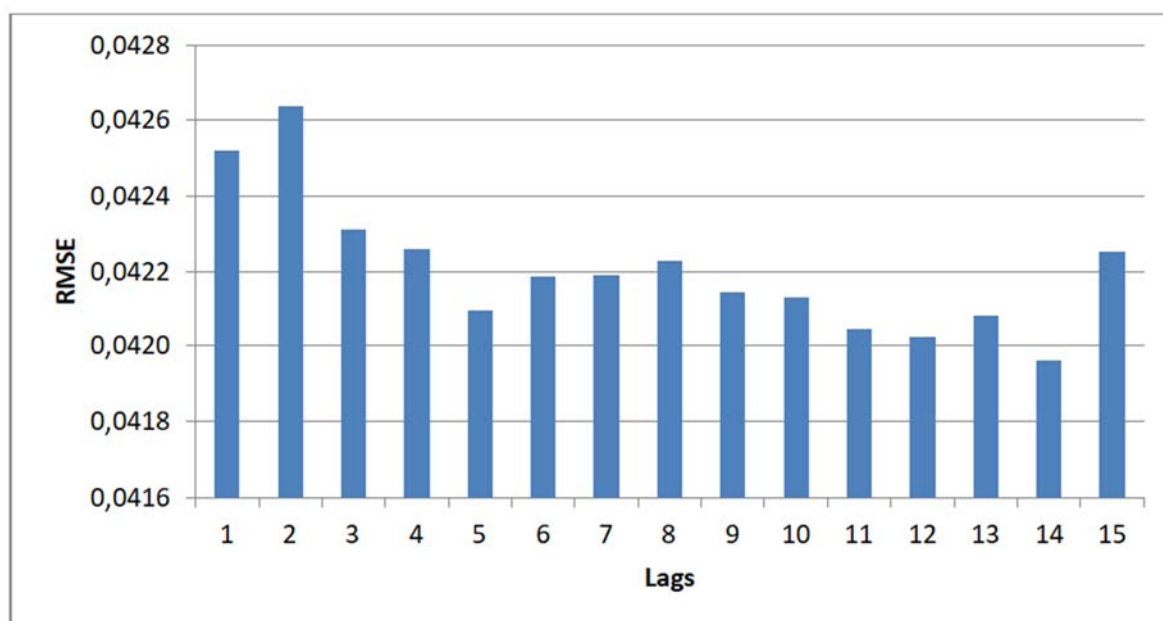


Figure 4. RMSE for AR models.



After identifying the best autoregressive representation, we built structural models. These include the 14 lags and all of the explanatory variables described earlier as independent variables to produce forecasts one day ahead. For this, we trained several alternative machine learning models and also produced the results for the random walk model. The in-sample and out-of-sample RMSEs of these models are presented in Table 2. An important issue in such forecasting models is to avoid overfitting in the in-sample or out-of-sample datasets. In the literature, this is known as the bias–variance trade-off. An efficient forecasting model is one that provides a balanced performance both in-sample and out-of-sample, i.e., the bias and variance are comparable. For this reason, we rejected all of models that provided evidence of overfitting and continued our empirical analysis with the rest. In the last column of Table 2, we note the models that overfit and are not used in the rest of our analysis. Interestingly, the tree models (plus bagged and boosted trees) do not overfit, and all of the GPR models overfit alongside most of the SVM models (with the exception of the linear SVM).

**Table 2.** In-sample and out-of-sample (OOS) RMSE of all models.

Models	In-Sample RMSE	OOS RMSE	Overfitting
RW	0.042643	0.057435	no
Linear Regression	0.038421	0.062872	no
Interactions Linear	0.1009	1.560244	yes
Robust Linear	0.039067	0.05736	no
Fine Tree	0.04992	0.071181	no
Medium Tree	0.045707	0.083954	no
Coarse Tree	0.047189	0.083388	no
Linear SVM	0.038581	0.056694	no
Quadratic SVM	0.044703	0.161214	yes
Cubic SVM	0.058968	0.456275	yes
Fine Gaussian SVM	0.098278	0.461634	yes
Medium Gaussian SVM	0.042352	0.243223	yes
Coarse Gaussian SVM	0.046224	0.079625	yes
Boosted Trees	0.065151	0.06169	no
Bagged Trees	0.041597	0.061089	no
Squared Exponential GPR	0.039915	0.216353	yes
Matern 5/2 GPR	0.039915	0.164098	yes
Exponential GPR	0.039989	0.098513	yes
Rational Quadratic GPR	0.040069	0.120337	yes

### 3.1. Time Frame $t + 1$

According to the results presented in Figure 5, we observed that for the time horizon  $t + 1$ , the optimal in-sample model was the linear regression model with RMSE = 0.038421 and that the best out-of-sample forecasting model was the linear SVM model with RMSE = 0.056694. The robust linear model also showed very good results, as it had the second lowest RMSE in the out-of-sample data and the third lowest in the in-sample data. Finally, the random walk model seemed to adequately predict the out-of-sample data. Therefore, we can generally conclude that linear models are able to predict the natural gas spot prices one day ahead with high accuracy and that the best model (linear regression) has good generalization ability (Figure 6).

### 3.2. Time Frame $t + 3$

The results for the forecasting window  $t + 3$  are presented in Figure 7. We observed that for time horizon  $t + 3$ , the optimal in-sample model was the bagged trees model with RMSE = 0.057793 and that the best out-of-sample forecasting model was the boosted trees model with RMSE = 0.077136. According to the above, it is clear that the best models at time horizon  $t + 3$  are tree based models. The out-of-sample performance of the bagged trees model is presented in Figure 8.

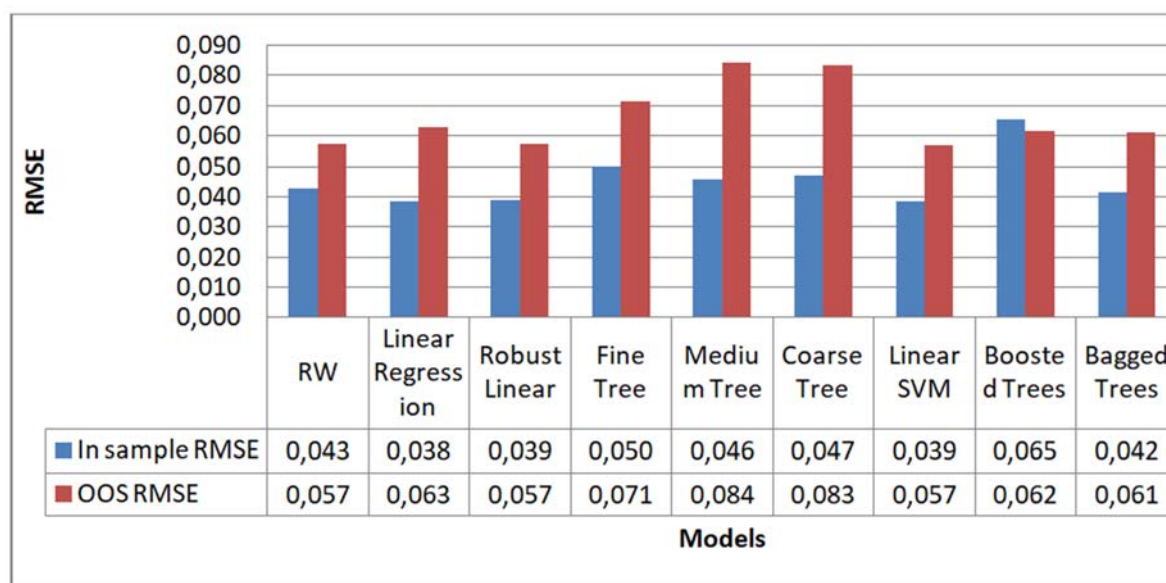


Figure 5. RMSE for  $t + 1$  forecasting.

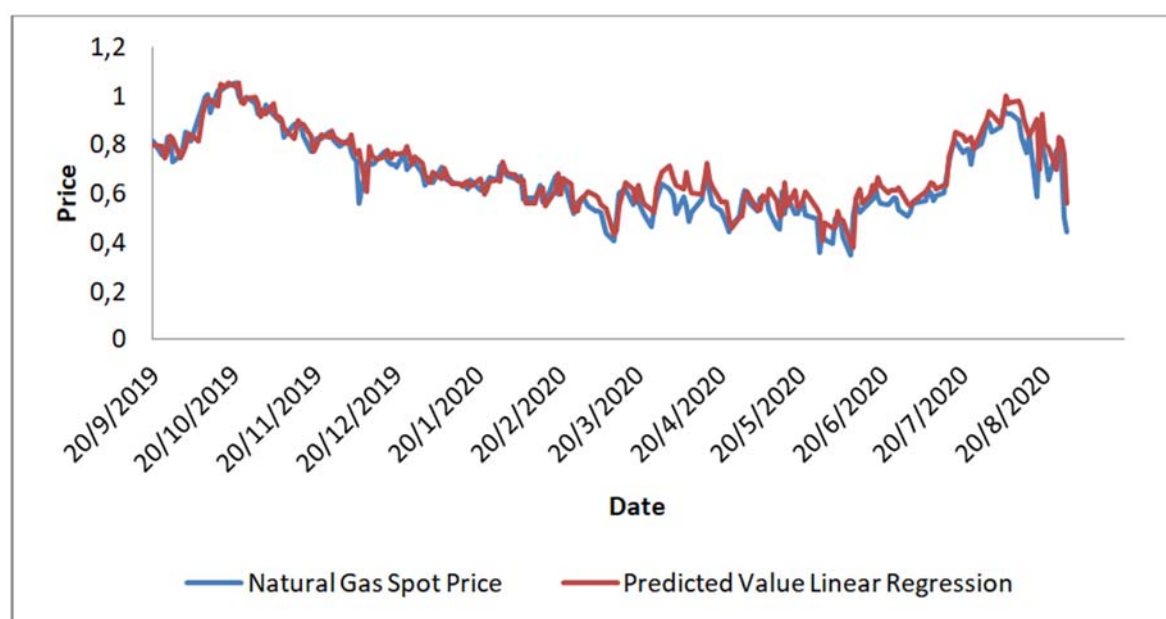


Figure 6. Comparison of the actual natural gas spot prices and the predicted prices with the linear regression model for  $t + 1$  in the out-of-sample part of the dataset.

### 3.3. Time Frame $t + 5$

The results for the forecasting window at  $t + 5$  are presented in Figure 9. In this window, we found that the optimal in-sample model was the bagged trees model with  $RMSE = 0.061787$  and that the best out-of-sample forecasting model was the linear SVM model with  $RMSE = 0.083687$ . The bagged trees model also shows good generalization ability (Figure 10). It is worth noting that the random walk model also showed good performance, as it holds the second lowest out-of-sample  $RMSE = 0.087654$ .



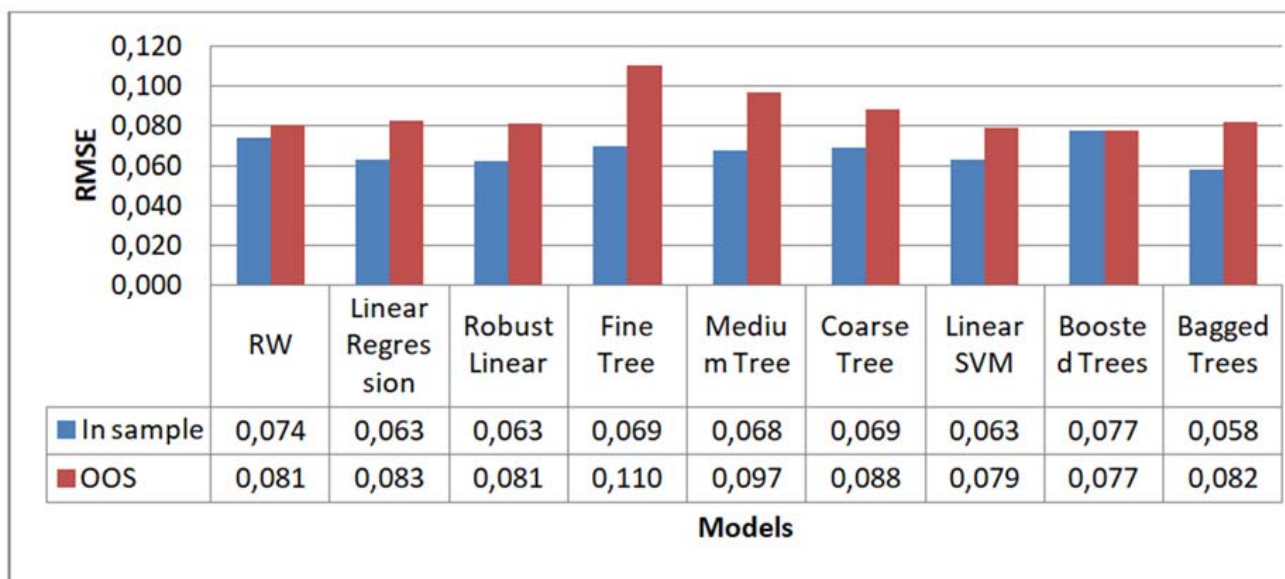


Figure 7. RMSE for  $t + 3$  forecasting.

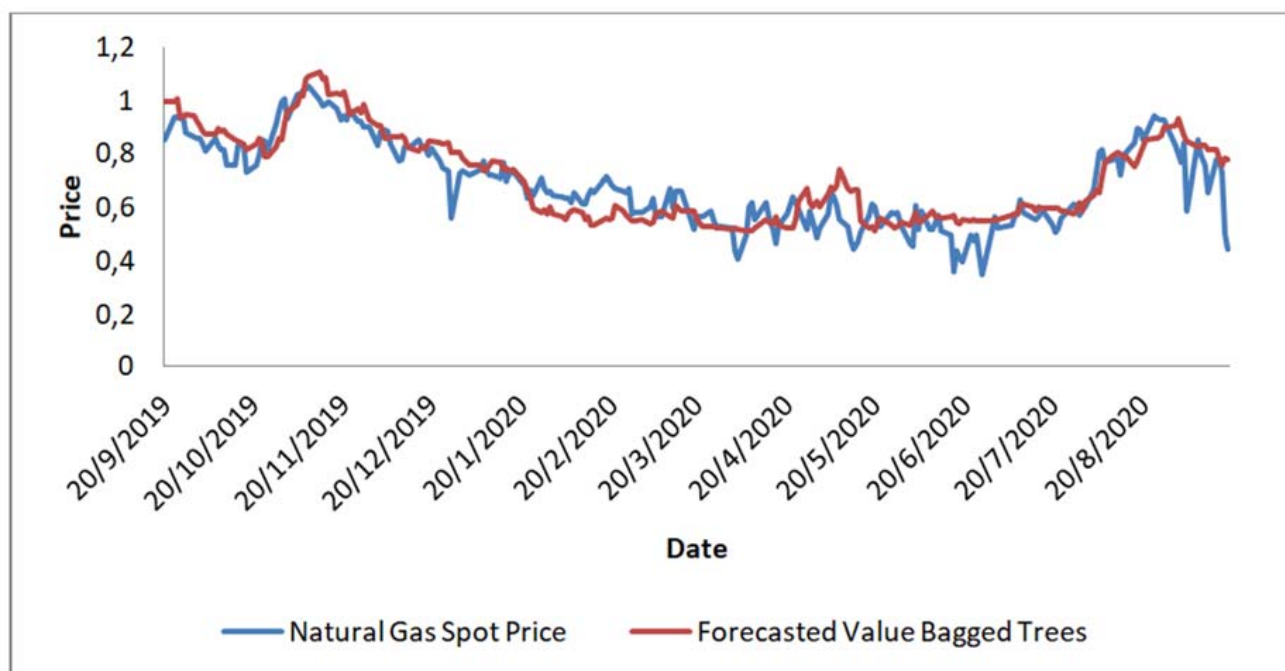


Figure 8. Comparison of the actual natural gas spot prices and the predicted prices with the bagged trees model for  $t + 3$  in the out-of-sample part of the dataset.

### 3.4. Time Frame $t + 10$

Finally, for the  $t + 10$  forecasting window the results are presented in Figure 11. We observed that the optimal in-sample model was the bagged trees model with  $\text{RMSE} = 0.064968$  and that the best out-of-sample forecasting model was the linear SVM model with  $\text{RMSE} = 0.102711$ . Additionally, the random walk model showed good results, as it achieved the second lowest out-of-sample  $\text{RMSE} = 0.109871$ . The best model for time horizon  $t + 10$  (bagged trees) has also good generalization ability (Figure 12).

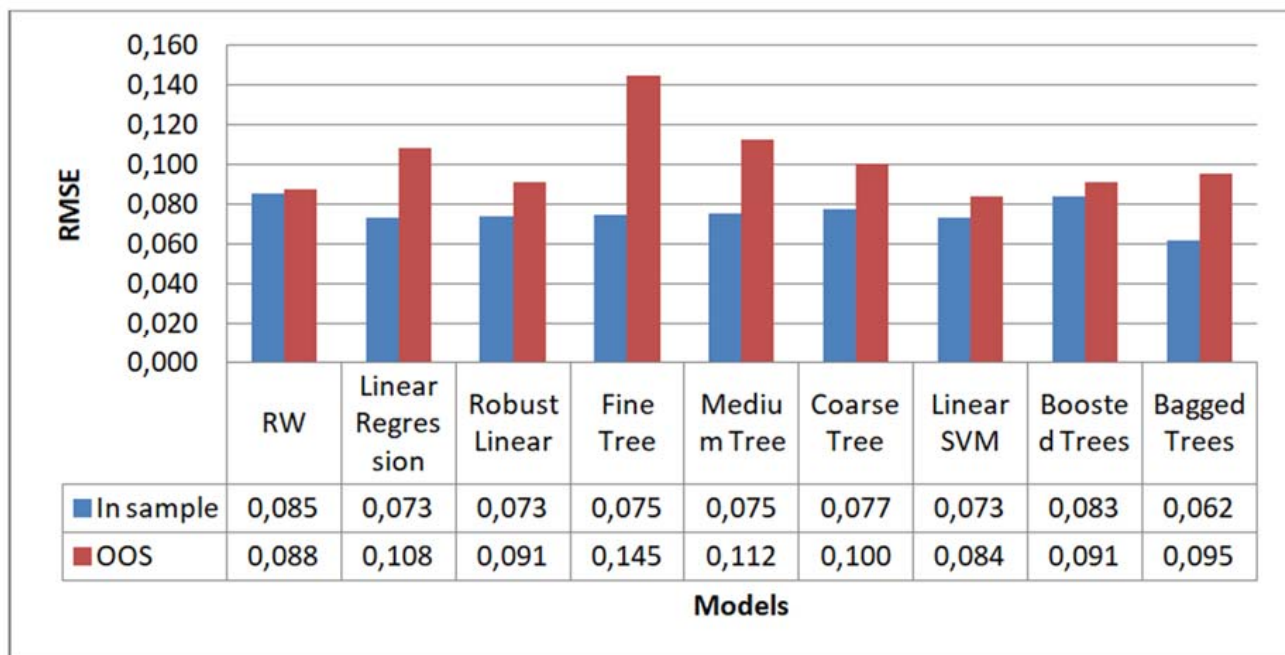


Figure 9. RMSE for  $t + 5$  forecasting.

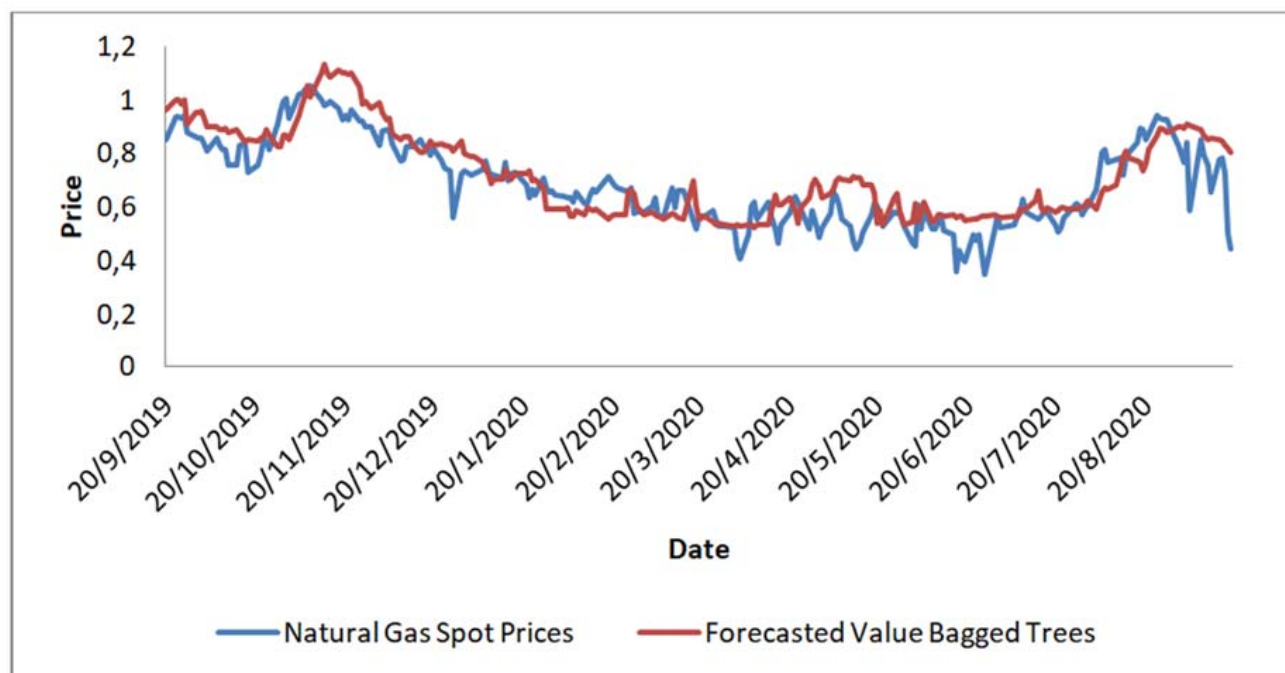


Figure 10. Comparison of the actual natural gas spot prices and the predicted prices with the bagged trees model for  $t + 5$  in the out-of-sample part of the dataset.

Interestingly the random walk model showed very good results for the out-of-sample part of the dataset at all time instances, while at the same time, we can conclude that all of the linear models have the ability to predict natural gas prices with high accuracy and showed very good performance with small RMSE values. The bagged trees models also showed very good predictive ability, having the lowest in-sample RMSE error at all time instances except for  $t + 1$ .

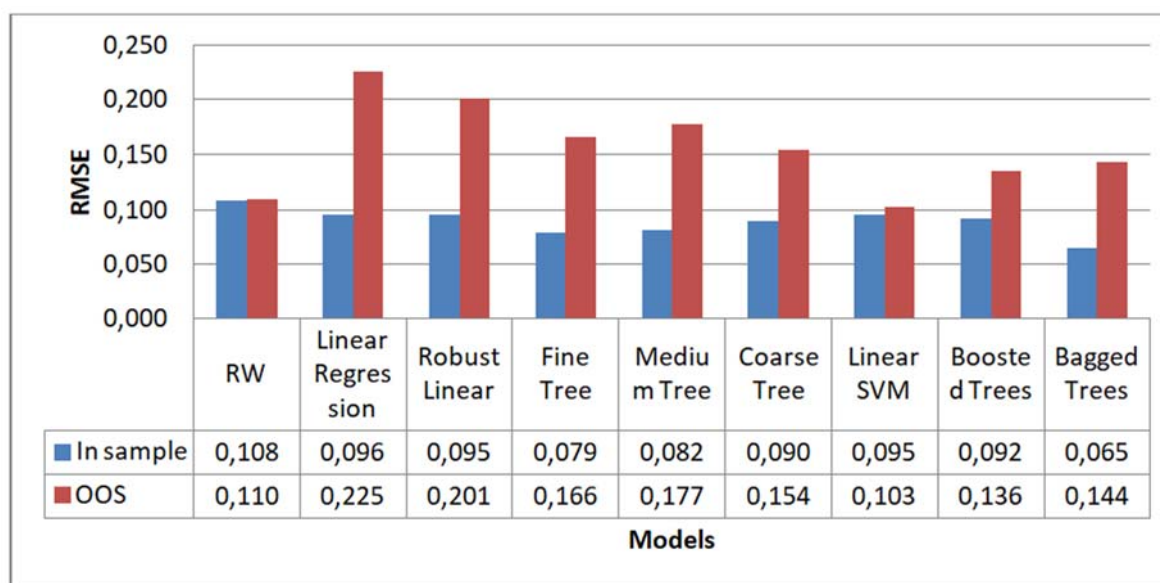


Figure 11. RMSE for  $t + 10$  forecasting.

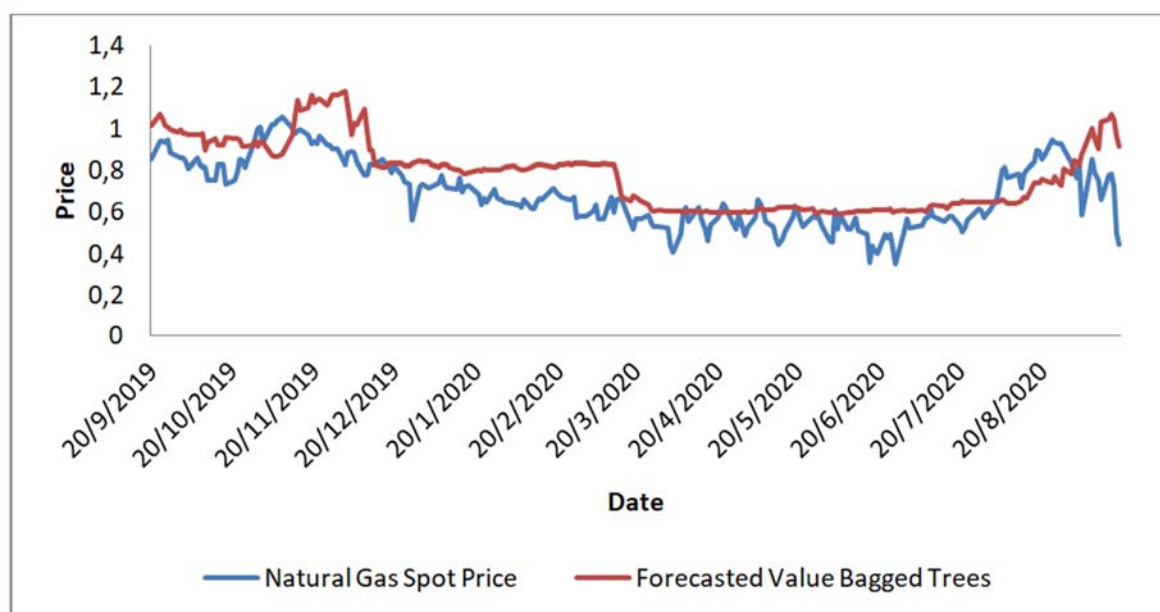


Figure 12. Comparison of the actual natural gas spot prices and the predicted prices with the bagged trees model for  $t + 10$  in the out-of-sample part of the dataset.

#### 4. Conclusions

The accurate forecasting of any asset has obvious practical implications. It can help individuals on both the supply and demand sides to reduce associated risk by better anticipating future changes in prices and by being prepared and acting on time to optimize their participation and behavior in the relevant market via positive or negative storage, substitution from and to this market, and the alteration of budget plans and in general decreases uncertainty, which has adverse effects on both suppliers and consumers. Moreover, government officials can use such information for larger scale planning, as they can anticipate prices swings.

The effective forecasting of natural gas prices is obviously important for all market participants: suppliers, distributors, consumers, investors, and regulatory agencies. It is also a powerful and important tool that has become increasingly important for various

stakeholders in the natural gas market, helping them to make better decisions for risk management, reducing the demand-supply gap, and optimizing resource utilization based on accurate predictions. For investors trading in the U.S. energy equity markets, the current boom around green energy investing offers significant hurdles. If and how these investors learn to cope with various information frictions, navigate through broad fluctuations in market risk appetite and uncertainty, and deal with unexpected changes in energy laws and regulations will be crucial to their investment decisions [18].

In this paper, we tested the effectiveness of various machine learning algorithms in forecasting natural gas spot prices. We trained multiple machine learning models and a naïve random walk model. In machine learning models, we used the optimal number of lagged natural gas spot prices and 21 other explanatory variables (regressors). These were selected based on economic theory and the relevant literature. Hence, these regressors included macroeconomic and stock market indicators, exchange rates, interest rates, the spot prices and future contracts of Oklahoma West Texas Intermediate Crude Oil, the corresponding future contracts of natural gas, the momentum of the last 5 and 10 days, and the 5- and 10-day moving average. The models were trained to forecast horizons one, three, five and ten days ahead ( $t + 1$ ,  $t + 3$ ,  $t + 5$  and  $t + 10$ ).

The dataset included 2423 daily observations for the time period from 3 December 2010 to 18 September 2020. This dataset was divided into two subsets, with the first part covering the range from 19 November 2010 to 19 September 2019, or 2180 observations that were used to train our models, and the second part spanning the period from 20 September 2019 to 18 September 2020, or the remaining 243 observations, which were used to test the generalization ability of the models to unknown data that were not used in the training process. In order to avoid the issue of overfitting, we employed a 5-fold cross validation method.

The optimal AR representation was found to be 14 lags using a linear SVM model. Next, we added all of the explanatory variables to train the 19 models. In 10 of these models, we detected overfitting; thus, they were not used in the subsequent analysis. These models were the interactions linear, SVM (quadratic, cubic, fine Gaussian, medium Gaussian, coarse Gaussian) and GPR (squared exponential, matern 5/2, exponential, and rational quadratic) models. The models that did not show overfitting were the random walk, linear regression, robust linear, fine tree, medium tree, coarse tree, linear SVM, boosted trees, and bagged trees models.

According to the results, the optimal model for in-sample data at  $t + 1$  is a linear regression model, and for  $t + 3$ ,  $t + 5$ , and  $t + 10$  bagged trees models are optimal. For the out-of-sample data, the best models are linear SVM models for  $t + 1$ ,  $t + 5$ , and  $t + 10$  and a boosted trees model for  $t + 3$ . The aforementioned models do not overfit since the RMSE's for the in-sample and out-of-sample data are comparable.

Therefore, from our research, we conclude that the most effective methods for natural gas spot price forecasting are the linear SVM and the bagged trees.

**Author Contributions:** D.M., E.S., P.G., and T.P. contributed equally to all parts of the paper, except for funding acquisition (E.S.). All authors have read and agreed to the published version of the manuscript.

**Funding:** This research is co-financed by Greece and the European Union (European Social Fund-ESF) through the Operational Programme “Human Resources Development, Education and Lifelong Learning” in the context of the project “Strengthening Human Resources Research Potential via Doctorate Research” (MIS-5000432), implemented by the State Scholarships Foundation (IKY).

**Data Availability Statement:** The data that support the findings of this study are available from the corresponding author upon reasonable request.

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

## References

1. Čeperić, E.; Žiković, S.; Čeperić, V. Short-term forecasting of natural gas prices using machine learning and feature selection algorithms. *Energy* **2017**, *140*, 893–900. [[CrossRef](#)]
2. Park, D.; El-Sharkawi, M.; Marks, R.; Atlas, L.; Damborg, M. Electric load forecasting using an artificial neural network. *IEEE Trans. Power Syst.* **1991**, *6*, 442–449. [[CrossRef](#)]
3. Chen, B.-J.; Chang, M.-W.; Lin, C.-J. Load Forecasting Using Support Vector Machines: A Study on EUNITE Competition 2001. *IEEE Trans. Power Syst.* **2004**, *19*, 1821–1830. [[CrossRef](#)]
4. Wang, J.; Cao, J.; Yuan, S.; Cheng, M. Short-term forecasting of natural gas prices by using a novel hybrid method based on a combination of the CEEMDAN-SE-and the PSO-ALS-optimized GRU network. *Energy* **2021**, *233*, 121082. [[CrossRef](#)]
5. Buchanan, W.; Hodges, P.; Theis, J. Which way the natural gas price: An attempt to predict the direction of natural gas spot price movements using trader positions. *Energy Econ.* **2001**, *23*, 279–293. [[CrossRef](#)]
6. Nguyen, H.T.; Nabney, I. Short-term electricity demand and gas price forecasts using wavelet transforms and adaptive models. *Energy* **2010**, *35*, 3674–3685. [[CrossRef](#)]
7. Salehnia, N.; Falahi, M.A.; Seifi, A.; Adeli, M.H.M. Forecasting natural gas spot prices with nonlinear modeling using Gamma test analysis. *J. Nat. Gas Sci. Eng.* **2013**, *14*, 238–249. [[CrossRef](#)]
8. Mishra, V.; Smyth, R. Are natural gas spot and futures prices predictable? *Econ. Model.* **2016**, *54*, 178–186. [[CrossRef](#)]
9. Herrera, G.P.; Constantino, M.; Tabak, B.M.; Pistori, H.; Su, J.-J.; Naranpanawa, A. Long-term forecast of energy commodities price using machine learning. *Energy* **2019**, *179*, 214–221. [[CrossRef](#)]
10. Zhang, W.; Hamori, S. Do Machine Learning Techniques and Dynamic Methods Help Forecast US Natural Gas Crises? *Energies* **2020**, *13*, 2371. [[CrossRef](#)]
11. Tamba, J.G.; Essiane, N.; Sapnken, E.F.; Koffi, F.D.; Nsouand, J.L.; Soldo, B.; Njomo, D. Forecasting natural gas: A literature survey. *Int. J. Energy Econ. Policy* **2018**, *8*, 216.
12. Su, M.; Zhang, Z.; Zhu, Y.; Zha, D.; Wen, W. Data Driven Natural Gas Spot Price Prediction Models Using Machine Learning Methods. *Energies* **2019**, *12*, 1680. [[CrossRef](#)]
13. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
14. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. *Classification and Regression Trees*; Wadsworth Inc.: Monterey, CA, USA, 1984.
15. Zimmermann, A. Ensemble-Trees: Leveraging Ensemble Power Inside Decision Trees. *Computer Vision* **2008**, *5255*, 76–87. [[CrossRef](#)]
16. Peterson, K. Resting Heart Rate Variability Can Predict Track and Field Sprint Performance. *OA Journal-Sports* **2018**, *1*, 1–9.
17. Gogas, P.; Papadimitriou, T.; Sofianos, E. Money Neutrality, Monetary Aggregates and Machine Learning. *Algorithms* **2019**, *12*, 137. [[CrossRef](#)]
18. Dragomirescu-Gaina, C.; Galariotis, E.; Philippas, D. Chasing the ‘green bandwagon’ in times of uncertainty. *Energy Policy* **2021**, *151*, 112190. [[CrossRef](#)]