**Using machine learning to predict clean energy stock prices: How important are volatility and uncertainty?**

Perry Sadorsky

Schulich School of Business

York University

4700 Keele Street

Toronto, Ontario, Canada M3J 1P3

[psadorsky@schulich.yorku.ca](mailto:psadorsky@schulich.yorku.ca)

October 2022

**Abstract**

Addressing climate change and transitioning to a low carbon economy requires investment in clean energy. Increases in clean energy usage are creating new opportunities for clean energy equity investing. The existing literature mostly focuses on the dynamic relationship between clean energy equities, oil prices, technology stock prices, and other important macroeconomic variables like market volatility and economic policy uncertainty. However, there is a shortage of literature on forecasting clean energy stock prices. Forecasting clean energy equity prices is important for making investment decisions. This paper uses machine learning methods to predict clean energy stock price direction. The analysis reveals that random forests, extremely randomized trees, stochastic gradient boosting, and support vector machine have higher prediction accuracy than Lasso or Naïve Bayes. For 10-day to 20-day forecasts random forests, extremely randomized trees, stochastic gradient boosting, and support vector machine achieve prediction accuracies greater than 85%. In some cases, prediction accuracy reaches 90%. Lasso prediction accuracy is higher than Naïve Bayes but never greater than 65%. Technical indicators like MA200, MA50, and WAD are, on average, the features most important for predicting clean energy stock price direction. Of the non-technical indicators, VIX and OVX are consistently ranked high in importance. Of the forecasting methods considered, extremely randomized trees are impressive due to high accuracy and less computational time.

**Keywords**: clean energy stock prices; forecasting; machine learning; random forests; VIX

JEL Classification: G17; Q42, Q47

**1. Introduction**

Addressing climate change and transitioning to a low carbon economy requires investment in clean energy. According to BloombergNEF (BloombergNEF, 2022) global energy transition investment in 2021 was $755 billion (US). This investment includes sustainable materials, carbon capture and storage (CCS), hydrogen, nuclear, electrified heat, electrified transport, electricity storage, and renewable energy. Over the period 2004 to 2021 the compound annual growth rate (CAGR) of energy transition investment was 20%. For the year 2021 the two largest sectors were renewable energy and electrified transport due to increases in wind and solar installations and large increases in electric vehicle sales. At current trends, the electrified transport sector will overtake renewable energy as the top sector by the year 2024. Investment levels increased to a new record in every major region of the world. On a country basis, China’s energy transition investment in 2021 was the highest ($268 billion), the US was second ($114 billion), and Germany was third ($47 billion).

Increases in clean energy usage are creating new opportunities for clean energy equity investing. The existing literature mostly focuses on the dynamic relationship between clean energy equities, oil prices, technology stock prices, and other important macroeconomic variables like financial market volatility and economic policy uncertainty (Bondia et al., 2016; Dutta, 2017; Dutta et al., 2018; Elie et al., 2019a, 2019b; Ferrer et al., 2018; Geng et al., 2021; Gupta, 2017; Henriques and Sadorsky, 2008; Kumar et al., 2012; Le et al., 2021; Maghyereh et al., 2019; Managi and Okimoto, 2013; Nasreen et al., 2020; Pham, 2021, 2019; Reboredo, 2015; Reboredo et al., 2017; Reboredo and Ugolini, 2018; Saeed et al., 2021; Uddin et al., 2019; Wen et al., 2014). This research shows that oil prices and the prices of technology stocks affect clean energy stock prices while the impact of economic policy uncertainty on clean energy stock prices is mixed. However, forecasting the prices of clean energy stocks is a topic that is understudied. Forecasting clean energy equity prices is imperative for making well informed investment decisions about this asset class. This raises several important questions. Which forecasting methods are the best to use? How important are market volatility and economic policy uncertainty (EPU) for predicting clean energy stock prices? Does variable importance change across the forecast horizon? These are the questions that this paper addresses.

In answering these questions, the following approach is followed. First, this paper focusses on predicting the direction of clean energy stock prices. In practice, investors are more interested in the direction of asset prices rather than the actual values when determining asset allocation (Pesaran and Timmermann, 2002). Numerous studies have reported high accuracy from predicting asset price direction (Ballings et al., 2015; Basak et al., 2019; Leung et al., 2000; Lohrmann and Luukka, 2019; Nyberg, 2011; Nyberg and Pönkä, 2016; Pönkä, 2016; Sadorsky, 2022). Second, machine learning methods are used to predict the price direction of clean energy stocks. Compared to standard regression methods, machine learning methods can achieve higher prediction accuracy when the relationship between the predictor and features is complex (Hastie et al., 2009; James et al., 2013; Mullainathan and Spiess, 2017). Machine learning methods like random forests and support vector machines have demonstrated high accuracy when predicting stock prices (Ampomah et al., 2020; Ballings et al., 2015; Basak et al., 2019; Ghosh et al., 2021; Khan et al., 2020; Lohrmann and Luukka, 2019; Weng et al., 2018). Building on this existing literature, the machine learning methods used in this present paper include random forests, extremely randomized trees, stochastic gradient boosting, support vector machine, Lasso, and Naïve Bayes. Third, the features (also referred to as explanatory variables or predictors) includes technical indicators, market volatility and policy uncertainty. Technical indicators are important for predicting asset prices (Bustos and Pomares-Quimbaya, 2020; Neely et al., 2014; Sadorsky, 2022, 2021a, 2021b; Wang et al., 2020; Yin et al., 2017; Yin and Yang, 2016). Including market volatility and uncertainty is a new contribution to the literature on predicting clean energy stock prices. The research that comes closest to this present study are the papers by Sadorsky (2021a) and Sadorsky (2022). Sadorsky (2021a) uses random forests to predict clean energy stock price direction. He finds that random forests and bagging have higher accuracy than logit models. The feature space includes technical indicators but does not include business cycle variables. Sadorsky (2022) uses random forests and support vector machine to predict solar stock prices. Random forests, tree bagging, extremely randomized trees, and support vector machine produce more accurate predictions than logit or boosted logit. The feature space includes technical indicators and the volatility of oil prices and silver prices. The most important predictors are technical indicators. Silver price volatility and oil price volatility rank in the top third. Neither of these two studies include a more comprehensive set of market volatility or policy uncertainty variables.

Several important findings emerge from the analysis reported in this paper. First, random forests, extremely randomized trees, stochastic gradient boosting, and support vector machines produce much higher accuracy than that of Lasso or Naïve Bayes when predicting the direction of clean energy equity price movements. These results are consistent with the literature showing the high prediction accuracy of machine learning models for predicting stock price direction. Second, the analysis of this paper finds that technical indicators like MA200, WAD, and MA20 are important features for predicting the prices of clean energy stocks. This result is consistent with the literature showing the importance of technical indicators for forecasting stock prices. The stock market volatility (VIX) and oil price volatility (OVX) are important non-technical predictors. Third, at the 15 day forecast horizon, the most important features are, in most cases, also the most important features at the 20 day forecast horizon. This is important since the highest prediction accuracy is observed between 15 and 20 days.

This paper is organized as follows. The next section of the paper presents the background literature. Section 3 presents the machine learning methods while Section 4 describes the data. The results are reported in Section 5. Section 6 concludes the paper and offers some practical implications stemming from the results.

**2. Background literature**

Stock market volatility can have a major impact on stock prices because higher volatility implies more risk (Schwert, 1990). Investments that looked attractive in a low risk environment may not be as attractive in a high risk environment. Oil price volatility and technology stock market volatility may also have an important impact on clean energy stock prices. Since clean energy and fossil fuels are substitutes, higher oil prices should encourage a shift away from fossil fuels to clean energy. Higher oil price volatility can, however, create delays in decision making because of uncertainty over the future profitability of renewable energy investments. Clean energy relies on technological innovation to bring down the cost of generating energy from renewable sources and uncertainty about technology can delay the decision to invest in clean energy. This uncertainty can be captured in technology stock market volatility (Henriques and Sadorsky, 2008).

Dutta (2017) finds a positive relationship between the realized volatility of clean energy equities and oil volatility (OVX). Ahmad et al. (2018) find a negative correlation between renewable energy stock volatility and OVX. In studying the stock price dynamics of solar energy firms, Dutta (2019) finds that OVX has a larger impact on solar stock prices than either the gold VIX or the silver VIX. Dutta (2018) finds that OVX is useful for hedging the downside risk of clean energy stock prices. Sadorsky (2022) finds that the volatility of oil prices along with the volatility of silver prices are important predictors of solar energy stock prices. Dutta et al. (2020) study how useful gold, silver, and oil volatility indices are for hedging downside risk of clean energy equities. They find that the crude oil volatility index is the most effective hedging asset followed by gold and silver.

Economic policy uncertainty in the form of uncertainty about fiscal policy, monetary policy and regulatory uncertainty can affect business decisions and financial markets (Baker et al., 2016). In the face of uncertainty, businesses are reluctant to commit to capital spending which affects current and future productivity and profitability (Pástor and Veronesi, 2012). An increase in economic policy uncertainty can indicate worsening economic performance which will affect the cost of capital and other real option decisions (Arouri et al., 2016).

Lundgren et al. (2018) find that uncertainty variables (VIX and financial stress) have a significant impact on clean energy stock returns and volatility. Uncertainty is mostly a net transmitter of connectedness during the global financial crisis and the European debt crisis. Ferrer et al. (2018) using time and frequency connectedness and controlling for the impact of financial factors find that clean energy stocks are similar to technology stocks. In comparing the impact of financial market uncertainty, oil market uncertainty, and economic policy uncertainty on clean energy stocks, Ji et al. (2018) find that economic policy uncertainty has a weaker effect than the other two uncertainty variables. Liu and Hamori (2020) Find that the VIX has a significant impact on renewable energy stocks and the impact of stock market uncertainty on clean energy stocks is greater in the US as compared to Europe. Zhao (2020) finds that economic policy uncertainty shocks have a negative impact on clean energy stock prices. Chakrabarti and Sen (2021) study the market risk of green energy equities. They find that green stocks are defensive and may be of interest to investors who want protection against downside risk. Volatility spills over from domestic markets to green equities. Market risks and spillovers are more variable in the US and Europe than they are in Asia Pacific. Green stocks have shown resilience to the COVID-19 period. Saeed et al. (2021) find that economic policy uncertainty has no statistically significant impact on the total spillover index for clean energy stocks, green bonds, crude oil., and fossil fuel energy stocks. Liu et al. (2022) find that the impact of economic policy uncertainty on clean energy stocks in the US, Europe, and the world tends to be concentrated at a high frequency. Uncertainty caused by COVID-19 is more significant than that caused by the global financial crisis.

In summary, oil price volatility has a larger impact on clean energy stock prices than does stock market volatility. The impact of EPU on clean energy stock prices is mixed. There is, however, little known about how useful market uncertainty and economic policy uncertainty are for forecasting clean energy stock prices. This is the gap in the literature that this paper addresses.

**3. Methods**

*3.1. Machine learning methods*

This section provides a brief non-technical description of the machine learning methods (random forests, boosting, extremely randomized trees, SVM, Lasso, and Naïve Bayes) used to predict clean energy stock price direction. James et al. (2013) provide a very informative and easy to read overview of these methods while a more technical description is provided by Hastie et al. (2009).

Random forests and extremely randomized trees are based on ensembles of decision trees. Decision trees can be trained to have high prediction accuracy on a training data set, but decision trees are susceptible to high variance. High variance means that small changes in the data set can produce different predictions. One way to address the high variance problem is to use ensemble methods to create many de-correlated trees and then average the predictions across the trees. Random forests are collections of decision trees created using bootstrapping (Breiman, 2001). Bootstrapping is used to create many training samples and a decision tree is fit to each sample. Each time a split in a tree occurs a random sample of predictors is chosen from the full set of predictors. The random choice of features helps to reduce the correlation between the trees. The number of predictors chosen at random is calculated as the floor of the square root of the total number of predictors (James et al., 2013). The predictions from each tree are averaged to provide an overall prediction. Extremely randomized trees (Extra Trees) is a tree based method that is like random forests (Geurts et al., 2006). At each split, Extra Trees, similar to random forests, also selects a subset of predictors chosen at random. For each of the selected predictors a small number of randomly chosen split points (often one) is made. The best split is chosen from this small number of choices. Extra Trees does not use bootstrapping but samples from the original data set. Extra Trees creates an ensemble of trees containing trees that have more variability but are less correlated than the trees from a random forests (Geurts et al., 2006). Extra Trees takes much less computational time than random forests because splits are chosen at random and bootstrapping is not used.

Instead of creating an ensemble of trees, tree boosting works by making a sequence of adjustments to one decision tree (James et al., 2013). A boosted decision tree is constructed as follows. First, a decision tree is fit to the data. Then a decision tree is fit to the residuals. This new decision tree is added to the previous tree and a new model is fit. The updated residuals are again fit with a decision tree. This sequential approach continues until a stopping criterion is met. The estimation process fits small trees to the residuals at each step and in this way slowly improves the fit of the original decision tree. There are three main tuning parameters for boosting; the number of trees, the shrinkage parameter which controls the rate at which learning takes place, and the number of splits in each tree (interaction depth) which controls the complexity of the tree. There is an inverse relationship between the learning rate and the number of trees and in many practical applications boosting can take much more computational time than random forests.

Support Vector Machine (SVM) is a machine learning method that partitions a data set into groups. The separation between groups is achieved using a boundary (hyperplane) (James et al., 2013). There are many different hyperplanes which can be used to separate the data and the goal of SVM is to find the maximum margin hyperplane (MMH). This is the hyperplane that has the greatest separation between the groups of data. The points that are closest to the MMH are the support vectors. For linearly separable data the support vectors can be found easily. For more complex problems the support vectors can be found by using advanced vector geometry. Kernels can be used to map the data into higher dimensionally space. Mapping the data into higher dimensional space creates the possibility that a nonlinear relationship in a specific dimension may become linear in higher dimensional space.

Two classification methods, Naïve Bayes and Lasso are included as benchmarks (James et al., 2013). Modelling stock price direction is a classification problem because the direction of stock price change can be classified as a binary variable. Stock price change from one period to the next can be classified as either positive or non-positive. The Naïve Bayes classifier uses Bayes theorem to make predictions.

The least absolute shrinkage and selection operator (Lasso) is like ridge regression (James et al., 2013). Whereas ridge regression shrinks parameter estimates down to small but non-zero values, Lasso may set parameter values equal to zero which results in a sparse model.

*3.2. Setup of the models*

For each model, multi-step forecasts from one day to twenty days are produced. Twenty days is approximately the average number of trading days in a month. The data were randomly split into a training set consisting of 70% of the data and a testing set consisting of 30% of the data. Five hundred trees were used to estimate the random forests and extremely randomized trees. At each split the number of predictors was randomly chosen from the square root of the total number of features. The number of trees does not affect the prediction accuracy of random forests so long as a large enough number of trees are chosen. A small number of trees, however, results in high test error. In the case of a 20-day forecast for the clean energy ETF PBW, the test error drops sharply as the number of trees approaches 200 trees. The test error does not change much as the number of trees passes 200. The results for other forecast horizons yield similar results. Results for the ETFs ICLN and QCLN are similar to that of PBW. The stochastic gradient boosted model was estimated with 1,000 trees, a bag fraction of 0.5, an interaction depth of 4, a minimum number of observations in each node of 10, and a shrinkage of 0.1. These values were chosen based on bootstrap re-sampling of a grid search with the number of trees (500, 1,000, 2,000, 3,000), the interaction depth (2, 4, 5) and shrinkage (0.01, 0.1, 0.2). The SVM was estimated with a radial basis and two tuning parameters. Ten-fold cross validation with 10 repeats was used to determine the optimal values. The tuning grid for the cost parameter contained the values (0.1, 1, 10, 100, 1000). The tuning grid for the gamma parameter contained the values (0.2, 0.5, 1, 2, 3, 4). Distance-based algorithms like the ones used by SVM are susceptible to the range of the data. Consequently, the features were scaled to have a mean of zero and standard deviation of unity. Data scaling was not applied to the tree-based methods because splitting takes place on a single feature which is not affected by scaling.

For classification problems the confusion matrix provides the main output from which to create measures of prediction accuracy. Prediction accuracy is the number of true positives and true negatives divided by the total number of predictions. This value ranges from zero to one. The kappa statistic is a better measure of prediction accuracy when there is a considerable unbalance between the classification categories because kappa adjusts prediction accuracy by accounting for the possibility of a chance occurrence of a correct prediction. The F1 score, sometimes called the F measure, is a prediction accuracy measure that combines precision and recall into a single number which also takes into account class imbalance.

Shapley values are used to explain feature importance. Shapley values are derived from conditional game theory. Essentially, the Shapley value signifies the contribution of each feature towards the predicted value compared to the average prediction for the data set (Lundberg and Lee, 2016, 2017). As an example, consider the Shapley calculation for one feature say X1. The accuracy of every combination of features not including X1 is recorded and then a test is conducted to see how adding X1 to each combination improves the accuracy. Shapley values are theoretically consistent and more robust than permutation based variable importance. The tradeoff is that Shapley values are much more computationally intensive and have a much longer computation time.

All calculations were done in R (R Core Team, 2022) using the ranger machine learning package (Wright et al., 2022), the e1071 package (Meyer et al., 2021), the Extra Trees package (Simm and Abril, 2014), the GBM package (Greenwell et al., 2020), the caret package (Kuhn et al., 2020) and the fastshap package (Greenwell, 2021).

**4. Data**

The study used data on clean energy stock prices, stock market volatility, oil market volatility, technology stock market volatility, economic market uncertainty, economic policy uncertainty, and the equity market infectious disease tracker. Clean energy stock prices are measured using three of the most widely traded clean energy exchange traded funds (ETFs). The Invesco WilderHill Clean Energy ETF (PBW) is based on the WilderHill Clean Energy Index. The WilderHill Clean Energy Index is a collection of stocks of publicly traded companies listed on a major exchange in the U.S. that are engaged in the business of the advancement of cleaner energy and conservation or are important to the development of clean energy. PBW holds about 80 stocks. As of April 30, 2022, the top 10 holdings comprised 22% of total assets. The second clean energy ETF is the ICLN. The ICLN is based on the iShares Global Clean Energy index which is designed to track the performance of approximately 100 global clean energy-related companies. As of April 30, 2022, the top 10 holdings accounted for 49% of total assets. The third clean energy ETF is QCLN which is based on the NASDAQ Clean Edge Green Energy Index. The fund tracks the NASDAQ Clean Edge Green Energy Index. The index is designed to track the performance of small, mid and large capitalization clean energy companies that are publicly traded in the United States. This ETF holds about 65 stocks. As of April 30, 2022, the top 10 holdings consisted of 56% of total assets.

Stock market volatility is measured using the well-known CBOE VIX index that tracks the S&P 500 volatility. The VIX is calculated using a complicated formula that averages the weighted prices of out of the money puts and calls that expire in 16 and 44 days on the S&P 500. Technology stock market volatility is measured using VXN which is computed in a similar way as VIX but for the NASDAQ 100 stocks. The CBOE oil price volatility index (OVX) is used to measure oil price volatility. The OVX is constructed using options prices on crude oil futures.

US equity market uncertainty (EMU), US economic policy uncertainty (EPU) and the infectious disease equity market volatility (EMV\_IDT) are obtained from the economic policy uncertainty website (<https://www.policyuncertainty.com/>) (Baker et al., 2016). These indices are based on key word searches in newspapers and media.

The data for this study is daily and starts on July 1, 2008 (ICLN began trading on June 24, 2008) and ends on June 30, 2022. Stock price and market volatility data were collected from Yahoo Finance. Data on the uncertainty variables was obtained from FRED. The technical indicators used in this study include the 50-day and 200-day moving averages, the relative strength indicator (RSI), Williams accumulation and distribution (WAD), stochastic oscillator (slow, fast), advance – decline line (ADX), moving average cross-over divergence (MACD), price rate of change (ROC), on balance volume, and the money flow index (MFI).. These technical indicators have shown predictive power for predicting stock prices (Bustos and Pomares-Quimbaya, 2020; Neely et al., 2014; Wang et al., 2020; Yin et al., 2017; Yin and Yang, 2016). Technical indicators were calculated using the default settings in the R package TTR (Ulrich, 2020).

Time series patterns of the clean energy ETFs are shown in Figure 1a. Clean energy stocks, like most stocks, dropped considerably during the 2008-2009 global financial crisis. Except for 2014, clean energy stock prices did not fluctuate very much between January 2, 2013 and January 2, 2019. The year 2014 was important because global investment in clean energy increased 16% that year mostly due to increased solar power in the US and China. Wind power investment in Europe was also strong[[1]](#footnote-1). After January 2019, clean energy stock prices increased and peaked in early January of 2021. This time period reflects the onset of the COVID19 pandemic. After January of 2021 prices declined somewhat but as of March 2022 are still above the values recorded in 2019.

VIX and VXN display similar time paths (Figure 1b). These market volatility indices recorded high values during the 2008 to 2009 global financial crisis and the beginning of the COVID19 pandemic (March 2020). Notice that in March of 2020, there was a large increase in oil price volatility. This was when the price of the front month oil futures contract turned negative.

Equity market uncertainty and economic policy uncertainty recorded high values during the global financial crisis and March 2020 (beginning of the COVID19 pandemic) (Figure 1c). The COVID19 pandemic had been detected in late 2019 and was a topic of concern in early 2020 but it was declared a global pandemic by the World Health Organization in March 2020. This is reflected in the time pattern of the equity market volatility infectious disease tracker (EMV\_IDT).

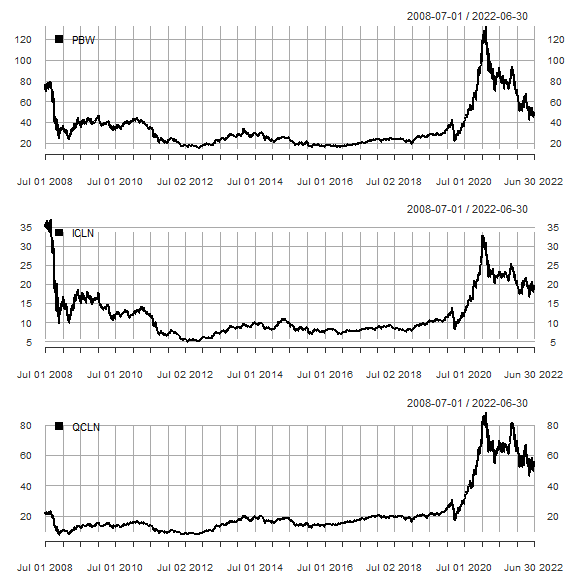


Figure 1a. Time series plots of the clean energy ETFs.

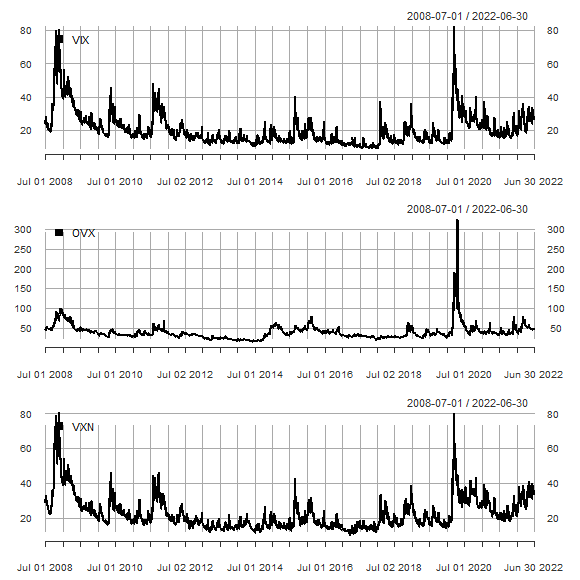


Figure 1b. Time series plots of financial market volatility.

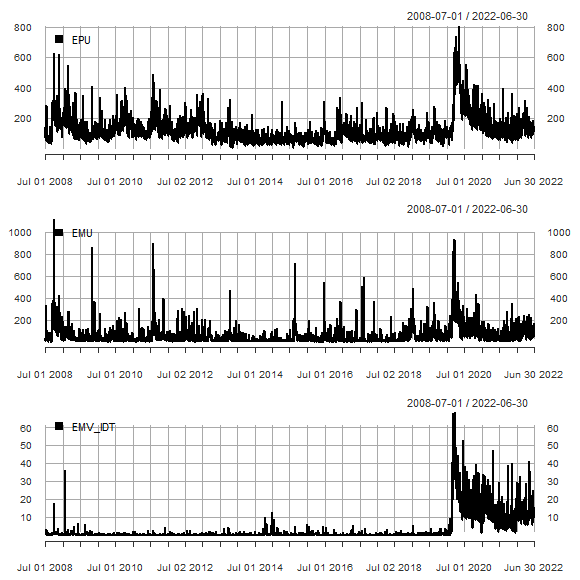


Figure 1c. Time series plots of uncertainty indicators.

Of the three ETFs, only QCLN had a positive average return over the sample period (Table 1). Returns are calculated using the formula 100\*ln(pt/pt-1) where pt is the adjusted closing price of the asset, or index value, at time period t. The coefficient of variation values indicates that OVX is more variable than VIX and VIX is more variable than VXN. Among the uncertainty indices, EMV\_IDT is the most variable and EPU is the least variable. The data display non-normality as for each variable the median is different from the mean, kurtosis and skewness are present, and the null hypothesis of normality is rejected.

Table 1. Summary statistics.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | median | mean | std.dev | coef.var | skewness | kurtosis | normtest.W | normtest.p |
| PBW | 0.085 | -0.015 | 2.351 | -161.482 | -0.411 | 5.278 | 0.936 | 0.000 |
| ICLN | 0.000 | -0.018 | 2.159 | -120.305 | -0.619 | 10.580 | 0.878 | 0.000 |
| QCLN | 0.135 | 0.024 | 2.218 | 92.522 | -0.432 | 5.431 | 0.937 | 0.000 |
| VIX | 17.355 | 20.234 | 9.747 | 0.482 | 2.333 | 7.431 | 0.783 | 0.000 |
| OVX | 35.045 | 39.014 | 19.090 | 0.489 | 4.253 | 34.561 | 0.705 | 0.000 |
| VXN | 19.700 | 22.603 | 9.433 | 0.417 | 2.060 | 6.073 | 0.818 | 0.000 |
| EPU | 101.570 | 124.721 | 86.497 | 0.694 | 2.333 | 8.210 | 0.802 | 0.000 |
| EMU | 32.750 | 59.908 | 82.233 | 1.373 | 4.150 | 28.319 | 0.615 | 0.000 |
| EMV\_IDT | 0.350 | 3.108 | 7.363 | 2.369 | 3.536 | 15.658 | 0.483 | 0.000 |

PBW, ICLN, and QCLN are expressed in continuous returns. Other variables are in original units. Data are for the period July 2, 2008 to June 30, 2022.

The percentage of up (positive) days for each forecast day and each clean energy ETF are displayed in Figure 2. The values range between 0.50 and 0.55. This indicates a relatively even class distribution.

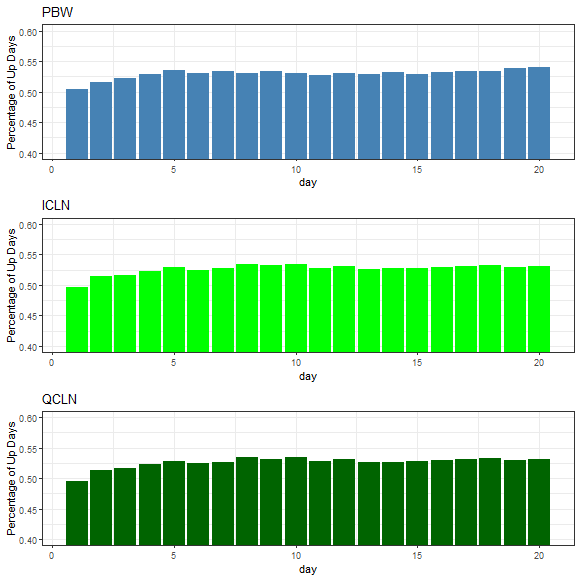


Figure 2. The percentage of positive days.

**5. Results**

Looking first at the accuracy for predicting PBW, random forests, stochastic gradient boosting, extremely randomized trees and SVM produce higher prediction accuracy than Lasso or the Naïve Bayes models (Figure 3). The prediction accuracy of random forests (RF), stochastic gradient boosting model (GBM), SVM, and extremely randomized trees (Extra Trees) increases quickly up to eight days. After 10 days the accuracy of random forests, GBM, SVM, and Extra Trees is over 80%. Between 15 and 20 days, random forests, GBM, and extremely randomized trees obtain accuracies of 90% or in some cases slightly greater. The pattern of accuracy observed for PBW across the number of days forecast is similar to that of ICLN and QCLN. Overall, random forests, GBM, and Extra Trees have the highest accuracy followed by SVM, Lasso, and Naïve Bayes. These results are supportive of the findings by Sadorsky (2021a) who finds that random forests have high accuracy for predicting the price direction of clean energy ETFs.

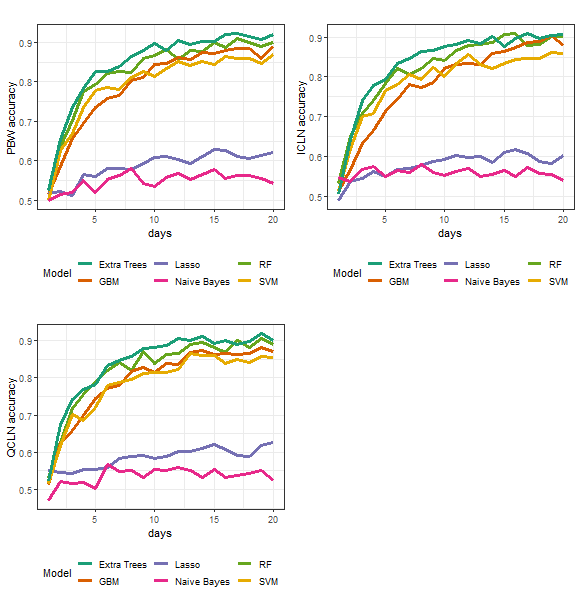


Figure 3. Prediction accuracy.

The pattern of kappa (Figure 4) is like that of the pattern of prediction accuracy shown in Figure 3. Random forests, GBM, SVM, and Extra Trees have the highest kappa values while Lasso and Naïve Bayes have lower kappa values.

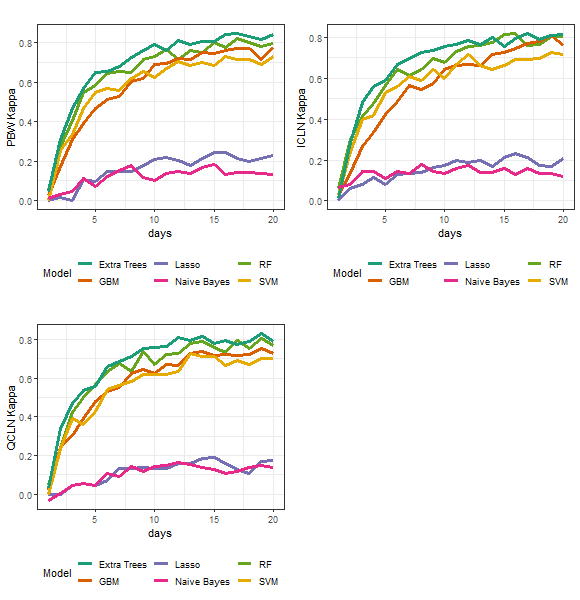


Figure 4. Kappa values.

After 5-days, the F1 values for random forests, GBM, Extra Trees and SVM are mostly greater than 0.80 (Figure 5). By comparison, Naïve Bayes has the lowest F1 values and Lasso has the second lowest F1 values. In summary, overall accuracy, kappa, and F1 are in agreement showing random forests, GBM, Extra Trees and SVM have higher prediction accuracy than Lasso or Naïve Bayes.

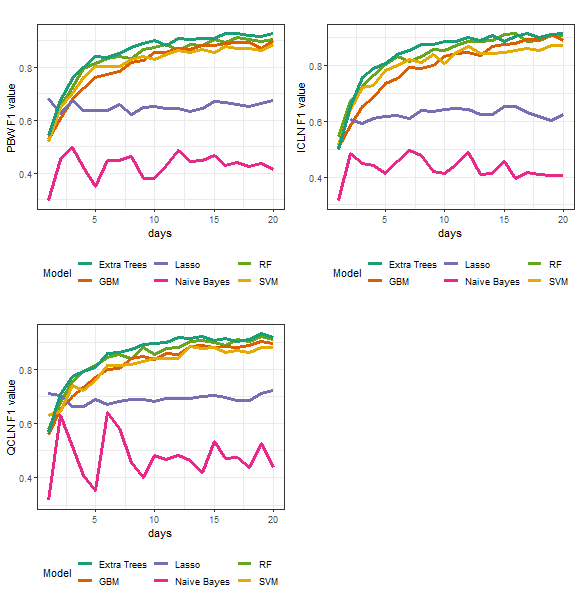


Figure 5. F1 value.

Mean absolute Shapley values for PBW reveal the importance of technical indicators (Figure 6a). For each plot in Figure 6a, features are ordered in descending order of importance. The horizontal axis denotes probability associated with the up classification. For a 10-day forecast for example (Figure 6a, top right), WAD increases the probability of being in the up classification by 3.25% on average. For forecasts of 10, 15, and 20 days, WAD, MA50, and MA200 are top five most important features while VIX is the most important non-technical feature. Focusing on the 20-day forecasts (because from Figure 3 forecast accuracy is highest for the 15 to 20-day forecasts), shows that the top four important features are technical indicators (WAD, MA200, MA50, OnBalanceVolume). The most important non-technical feature is VIX followed by OVX.

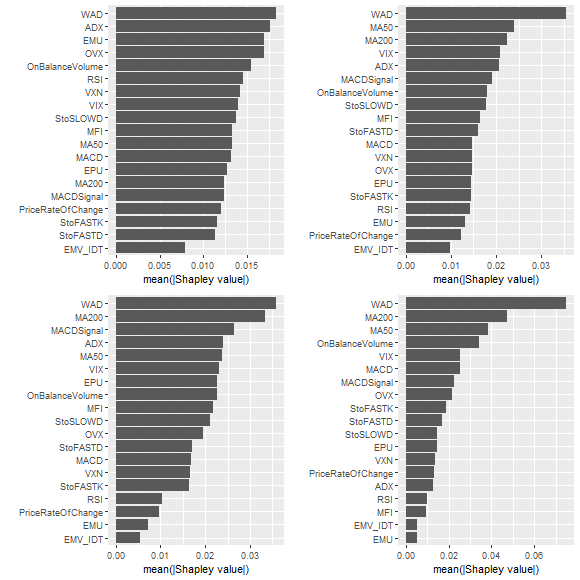


Figure 6a. Shapley values for PBW. Top left (5 day forecast horizon), top right (10 day forecast horizon), bottom left (15 day forecast horizon), and bottom right (20 day forecast horizon).

For ICLN, WAD, MA200, and MA50 are among the top five features for forecasting 10, 15, and 20 days (Figure 6b). For 20-day forecasts, the top five features in terms of importance are the same as the those for PBW (WAD, MA200, MA50, VIX, OnBalanceVolume).

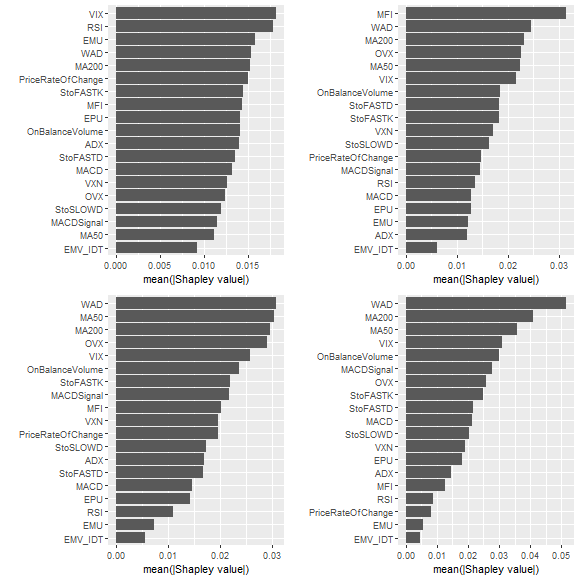


Figure 6b. Shapley values for ICLN. Top left (5 day forecast horizon), top right (10 day forecast horizon), bottom left (15 day forecast horizon), and bottom right (20 day forecast horizon).

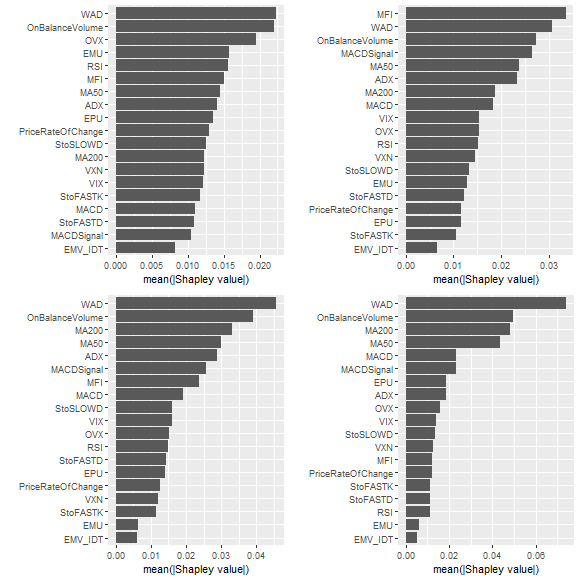


Figure 6c. Shapley values for QCLN. Top left (5 day forecast horizon), top right (10 day forecast horizon), bottom left (15 day forecast horizon), and bottom right (20 day forecast horizon).

For QCLN, the most important features for predicting 15 and 20 days are WAD, OnBalanceVolume, MA200, and MA50 (Figure 6c). For 20-day forecasts, the most important non-technical feature is EPU which is different from PBW and ICLN where EPU ranks below VIX and OVX in terms of importance.

The Figures 6a, 6b, and 6c show variable importance calculated using the mean of the absolute Shapley value but these figures do not show the relationship between the sign of the feature values and importance. A bee swarm plot for PBW 20-day forecast SHAP (Shapley additive explanations) variable importance is also revealing (Figure 7). The plot is organized so that the most important features are listed in order from the top to the bottom. The ranking of each feature in terms of importance in Figure 7 is the same as that in Figure 6a (bottom right) but Figure 7 includes additional information on the relationship between the importance and feature value. Figure 7 is color coded so that high feature values are warmer colors and low feature values are cooler colors. The most important feature is WAD and lower values of WAD have a greater importance. It is also the case that lower values for MA200 and MA50 have higher importance.

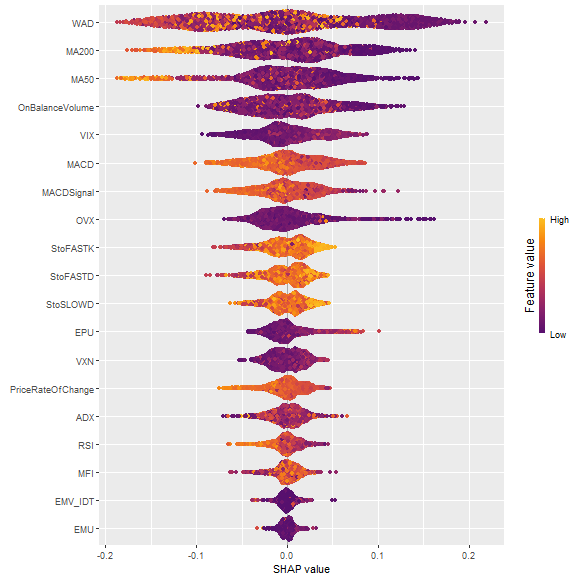


Figure 7. PBW SHAP variable importance (random forest 20-day forecast).

In the case of ICLN and QCLN, lower values of WAD, MA200, and MA50 have greater importance (Figures 8 and 9). For each of PBW, ICLN, and QCLN, higher values of VIX tend to have more importance than lower values (Figures 7 – 9). These results on VIX are supportive of the findings by Chakrabarti and Sen (2021) who find that green stocks offer protection against downside risks.

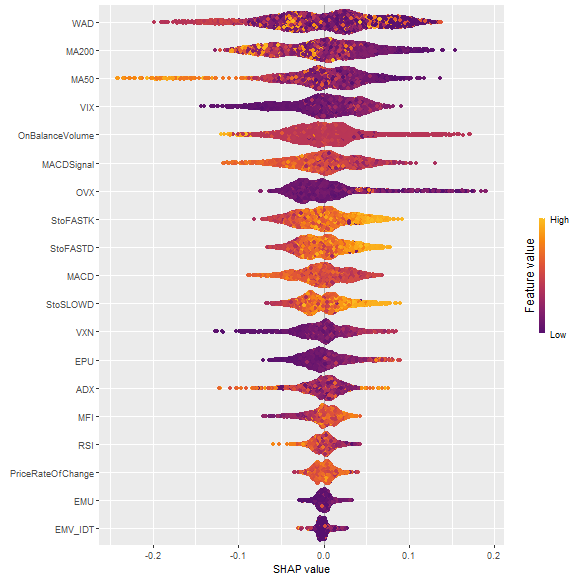


Figure 8. ICLN SHAP variable importance (random forest 20 -day forecast).

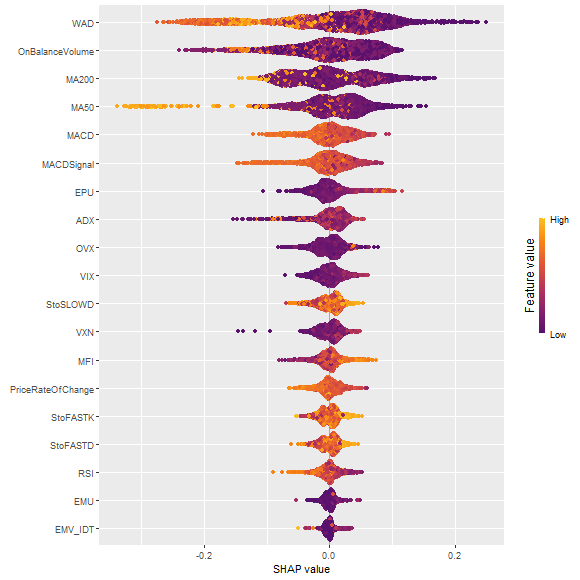


Figure 9. QCLN SHAP variable importance (random forest 20-day forecast).

The analysis reported so far has relied on cross-validation (CV) where the data is randomly split into a training data set (70% of the data) and a test data set (30% of the data). This approach may be problematic for time series data unless the time series properties of the data are preserved. This is not likely a problem for the applicating in this paper because the models used in this paper use a classification dependent variable with technical indicators as features which contain important time series information.

To gain further insight, time-series cross-validation (tsCV) was applied to the random forests model. With tsCV, the training set occurs immediately prior to the test set. A fixed rolling window approach was used where the length of the rolling window was fixed to use 70% of the data. The random forests model is selected because it has high accuracy over all forecast horizons and is easy to use to calculate variable importance.

The prediction accuracy from tsCV is very similar to that from the CV (Table 2). For PBW, CV accuracy is 0.8992 while for tsCV it is 0.9020. For ICLN, CV accuracy is 0.9012 while for tsCV it is 0.9040. For QCLN, CV accuracy is 0.8901 while for tsCV it is 0.8737. For each ETF, the CV and tsCV approaches yield similar values for predictive accuracy. Although not reported, the CV and tsCV approaches yield similar values for predictive accuracy for the Extra Trees. This shows that the accuracy measures are robust to the choice of data splitting (CV or tsCV).

Table 2. Comparing random forest prediction accuracy from CV and tsCV.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | 20-day | 20-day |
|  |  | CV | tsCV |
| PBW |
| Accuracy |  | 0.8992 | 0.9020 |
| Kappa |  | 0.7972 | 0.7960 |
| F1 |  | 0.9064 | 0.8777 |
| ICLN |
| Accuracy |  | 0.9012 | 0.9040 |
| Kappa |  | 0.8009 | 0.7949 |
| F1 |  | 0.9093 | 0.8714 |
| QCLN |
| Accuracy |  | 0.8901 | 0.8737 |
| Kappa |  | 0.7714 | 0.7279 |
| F1 |  | 0.9083 | 0.8271 |

In summary, for each clean energy ETF, Shapley feature importance indicates that technical indicators like WAD, MA50, and MA200 are consistently ranked among the most important features for the 15-day and 20-day forecast horizons. These results that technical indicators are important predictors of clean energy ETFs are consistent with some recent research that uses machine learning methods to predict clean energy stock prices. For example, Sadorsky (2021a) finds that technical indicators are important features for predicting clean energy ETF price direction but he did not include non-technical features. Sadorsky (2022) finds that technical indicators are the most important features for predicting solar stock price direction.

Among the non-technical indicators, VIX and OVX are important features for the 15 and 20 day forecast horizons. The VIX measures the volatility of the S&P 500 and it is not surprising that VIX is an important feature for predicting clean energy ETFs since market volatility can have a pronounced impact on asset prices. The importance of OVX for predicting clean energy stock prices is consistent with recent research demonstrating that oil price volatility has been identified as an important leading economic indicator (Chatziantoniou et al., 2021). The EPU is the most important non-technical feature for predicting 20-day QCLN. This may be due to QCLN having a high concentration (56% of the total assets are accounted for by the top 10 stocks) that makes it more susceptive to changes in economic policy uncertainty. Perhaps somewhat surprising is that technology stock market volatility (VXN) is not a very important feature for predicting clean energy stock price direction. Historically, one of the most cited criticisms of the adoption of clean energy is that it is too expensive and very much dependent upon technological innovation. But things have changed. The levelized cost of clean energy like solar and on shore wind has fallen dramatically since 2010 and these clean energy sources are now less expensive than that of coal (The Economist, 2020).

**6. Conclusions and practical implications**

Clean energy is, currently, the fastest growing energy source. This growth in clean energy is expected to continue as countries transition to lower carbon energy sources. Clean energy equities are growing in popularity and investors need accurate clean energy stock price forecasts in order to make sound investment decisions. The objective of this paper is to determine how important volatility and uncertainty are for forecasting clean energy stock price direction. Clean energy stock prices are measured using three widely traded ETFs.

The introduction to this paper posed three important questions regarding forecasting clean energy equity prices. In terms of the first question, “which forecasting method to use?”, this research finds random forests, stochastic gradient boosting, and extremely randomized trees produce higher accuracy forecasts than those from Lasso or Naïve Bayes models. Random forests, GBM, and extremely randomized trees have prediction accuracy greater than 80% after eight days. After fifteen days, the forecast accuracy from these models is greater than 85%. In some cases, prediction accuracy is 90% or even slightly higher. SVMs have high prediction accuracy but not as high as the predictions achieved using random forests, GBM, and extremely randomized trees. Extra Trees have the highest prediction accuracy for most forecast horizons. The Lasso prediction accuracy is never higher than 65%. The prediction accuracy of Naïve Bayes is mostly less than that of Lasso. The results from this research add support to the literature demonstrating the accuracy of predicting the direction of stock prices with machine learning methods.

In answering the question of how important volatility and uncertainty are for predicting clean energy stock prices, the analysis reveals that the most important features for predicting the direction of clean energy stock prices are the technical indicators WAD, MA50, and MA200. These results are consistent with the literature that establishes technical indicators as important features for predicting stock prices. In most cases, the most important non-technical features for the 15 to 20 day forecast period are VIX and OVX. Unlike in the case of PBW or ICLN, economic policy uncertainty is important for predicting 20-day QCLN. In other cases, economic policy uncertainty and market uncertainty variables are not that important. Economic market uncertainty is one of the least important features. Somewhat surprising is that technology stock market volatility is not a very important predictor of clean energy stock prices. This may be due to the dramatic drop in costs of supplying clean energy which lessens the reliance on technological innovation.

With regards to the question as to whether variable importance changes across the forecast horizon, the answer is mixed. Variable importance tends to vary little between the 15 day and 20 day forecast horizons. This is important since the highest prediction accuracy is observed between 15 and 20 days. Variable importance does change between five days and ten days and is most noticeable when comparing the 5 day forecasts with the 20 day forecasts but forecast accuracy is poorest at the one to five day forecast horizon.

This research offers several practical implications. First, random forests, stochastic gradient boosting, and extremely randomized trees should be used to forecast clean energy stock price direction because these methods provide more accurate predictions than those computed from Lasso and Naïve Bayes models. Random forests and extremely randomized trees are particularly worth using because they are easy to estimate and take relatively little computation time. One recommendation is to use extremely randomized trees to forecast clean energy stock price directions because this method is computationally fast and produces high accuracy. If one is interested in high forecast accuracy and a closer examination of variable importance than random forests are a good choice. Second, the features used for clean energy stock price prediction should include technical indicators (like WAD and moving averages), VIX and oil price volatility. The literature on clean energy stock prices has verified that oil is an important variable impacting clean energy stock prices and the results from this paper support the importance of the volatility of oil prices as a feature for forecasting clean energy stock price direction.

There are several avenues for future research. One avenue for future research would be to expand the predictor space to include interest rate and exchange rate variables. Including these variables may offer a more complete picture of how business cycle conditions affect clean energy stock price predictability. Another possible avenue for future research would be to forecast the stock prices of individual clean energy companies.

**References**

Ahmad, W., Sadorsky, P., Sharma, A., 2018. Optimal hedge ratios for clean energy equities. Economic Modelling 72, 278–295. https://doi.org/10.1016/j.econmod.2018.02.008

Ampomah, E.K., Qin, Z., Nyame, G., 2020. Evaluation of Tree-Based Ensemble Machine Learning Models in Predicting Stock Price Direction of Movement. Information 11, 332. https://doi.org/10.3390/info11060332

Arouri, M., Estay, C., Rault, C., Roubaud, D., 2016. Economic policy uncertainty and stock markets: Long-run evidence from the US. Finance Research Letters 18, 136–141. https://doi.org/10.1016/j.frl.2016.04.011

Baker, S.R., Bloom, N., Davis, S.J., 2016. Measuring Economic Policy Uncertainty. Q J Econ 131, 1593–1636. https://doi.org/10.1093/qje/qjw024

Ballings, M., Van den Poel, D., Hespeels, N., Gryp, R., 2015. Evaluating multiple classifiers for stock price direction prediction. Expert Systems with Applications 42, 7046–7056. https://doi.org/10.1016/j.eswa.2015.05.013

Basak, S., Kar, S., Saha, S., Khaidem, L., Dey, S.R., 2019. Predicting the direction of stock market prices using tree-based classifiers. The North American Journal of Economics and Finance 47, 552–567. https://doi.org/10.1016/j.najef.2018.06.013

BloombergNEF, 2022. Energy Transition Investment Trends 2022.

Bondia, R., Ghosh, S., Kanjilal, K., 2016. International crude oil prices and the stock prices of clean energy and technology companies: Evidence from non-linear cointegration tests with unknown structural breaks. Energy 101, 558–565. https://doi.org/10.1016/j.energy.2016.02.031

Breiman, L., 2001. Random Forests. Machine Learning 45, 5–32. https://doi.org/10.1023/A:1010933404324

Bustos, O., Pomares-Quimbaya, A., 2020. Stock market movement forecast: A Systematic review. Expert Systems with Applications 156, 113464. https://doi.org/10.1016/j.eswa.2020.113464

Chakrabarti, G., Sen, C., 2021. Dynamic market risk of green stocks across regions: Where does the devil lie? Journal of Cleaner Production 303, 127028. https://doi.org/10.1016/j.jclepro.2021.127028

Chatziantoniou, I., Filippidis, M., Filis, G., Gabauer, D., 2021. A closer look into the global determinants of oil price volatility. Energy Economics 95, 105092. https://doi.org/10.1016/j.eneco.2020.105092

Dutta, A., 2019. Impact of silver price uncertainty on solar energy firms. Journal of Cleaner Production 225, 1044–1051. https://doi.org/10.1016/j.jclepro.2019.04.040

Dutta, A., 2018. Oil and energy sector stock markets: An analysis of implied volatility indexes. Journal of Multinational Financial Management 44, 61–68. https://doi.org/10.1016/j.mulfin.2017.12.002

Dutta, A., 2017. Oil price uncertainty and clean energy stock returns: New evidence from crude oil volatility index. Journal of Cleaner Production 164, 1157–1166. https://doi.org/10.1016/j.jclepro.2017.07.050

Dutta, A., Bouri, E., Das, D., Roubaud, D., 2020. Assessment and optimization of clean energy equity risks and commodity price volatility indexes: Implications for sustainability. Journal of Cleaner Production 243, 118669. https://doi.org/10.1016/j.jclepro.2019.118669

Dutta, A., Bouri, E., Noor, M.H., 2018. Return and volatility linkages between CO2 emission and clean energy stock prices. Energy 164, 803–810.

Elie, B., Naji, J., Dutta, A., Uddin, G.S., 2019a. Gold and crude oil as safe-haven assets for clean energy stock indices: Blended copulas approach. Energy 178, 544–553.

Elie, B., Naji, J., Dutta, A., Uddin, G.S., 2019b. Gold and crude oil as safe-haven assets for clean energy stock indices: Blended copulas approach. Energy 178, 544–553. https://doi.org/10.1016/j.energy.2019.04.155

Ferrer, R., Shahzad, S.J.H., López, R., Jareño, F., 2018. Time and frequency dynamics of connectedness between renewable energy stocks and crude oil prices. Energy Economics 76, 1–20. https://doi.org/10.1016/j.eneco.2018.09.022

Geng, J.-B., Liu, C., Ji, Q., Zhang, D., 2021. Do oil price changes really matter for clean energy returns? Renewable and Sustainable Energy Reviews 150, 111429. https://doi.org/10.1016/j.rser.2021.111429

Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Mach Learn 63, 3–42. https://doi.org/10.1007/s10994-006-6226-1

Ghosh, P., Neufeld, A., Sahoo, J.K., 2021. Forecasting directional movements of stock prices for intraday trading using LSTM and random forests. Finance Research Letters 102280. https://doi.org/10.1016/j.frl.2021.102280

Greenwell, B., 2021. fastshap: Fast Approximate Shapley Values.

Greenwell, B., Boehmke, B., Cunningham, J., Developers (https://github.com/gbm-developers), G.B.M., 2020. gbm: Generalized Boosted Regression Models.

Gupta, K., 2017. Do economic and societal factors influence the financial performance of alternative energy firms? Energy Economics 65, 172–182. https://doi.org/10.1016/j.eneco.2017.05.004

Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition, 2nd ed, Springer Series in Statistics. Springer-Verlag, New York. https://doi.org/10.1007/978-0-387-84858-7

Henriques, I., Sadorsky, P., 2008. Oil prices and the stock prices of alternative energy companies. Energy Economics 30, 998–1010. https://doi.org/10.1016/j.eneco.2007.11.001

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning: with Applications in R, Springer Texts in Statistics. Springer-Verlag, New York. https://doi.org/10.1007/978-1-4614-7138-7

Ji, Q., Liu, B.-Y., Nehler, H., Uddin, G.S., 2018. Uncertainties and extreme risk spillover in the energy markets: A time-varying copula-based CoVaR approach. Energy Economics 76, 115–126. https://doi.org/10.1016/j.eneco.2018.10.010

Khan, W., Ghazanfar, M.A., Azam, M.A., Karami, A., Alyoubi, K.H., Alfakeeh, A.S., 2020. Stock market prediction using machine learning classifiers and social media, news. J Ambient Intell Human Comput. https://doi.org/10.1007/s12652-020-01839-w

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., R Core Team, Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., Candan, C., Hunt, T., 2020. caret: Classification and Regression Training.

Kumar, S., Managi, S., Matsuda, A., 2012. Stock prices of clean energy firms, oil and carbon markets: A vector autoregressive analysis. Energy Economics 34, 215–226. https://doi.org/10.1016/j.eneco.2011.03.002

Le, T.-L., Abakah, E.J.A., Tiwari, A.K., 2021. Time and frequency domain connectedness and spill-over among fintech, green bonds and cryptocurrencies in the age of the fourth industrial revolution. Technological Forecasting and Social Change 162, 120382. https://doi.org/10.1016/j.techfore.2020.120382

Leung, M.T., Daouk, H., Chen, A.-S., 2000. Forecasting stock indices: a comparison of classification and level estimation models. International Journal of Forecasting 16, 173–190. https://doi.org/10.1016/S0169-2070(99)00048-5

Liu, T., Hamori, S., 2020. Spillovers to Renewable Energy Stocks in the US and Europe: Are They Different? Energies 13, 3162. https://doi.org/10.3390/en13123162

Liu, T., Nakajima, T., Hamori, S., 2022. The impact of economic uncertainty caused by COVID-19 on renewable energy stocks. Empir Econ 62, 1495–1515. https://doi.org/10.1007/s00181-021-02087-3

Lohrmann, C., Luukka, P., 2019. Classification of intraday S&P500 returns with a Random Forest. International Journal of Forecasting, Special Section: Supply Chain Forecasting 35, 390–407. https://doi.org/10.1016/j.ijforecast.2018.08.004

Lundberg, S., Lee, S.-I., 2016. An unexpected unity among methods for interpreting model predictions. Presented at the Advances in Neural Information Processing Systems, arXiv. https://doi.org/10.48550/arXiv.1611.07478

Lundberg, S.M., Lee, S.-I., 2017. A Unified Approach to Interpreting Model Predictions, in: Advances in Neural Information Processing Systems. Curran Associates, Inc.

Lundgren, A.I., Milicevic, A., Uddin, G.S., Kang, S.H., 2018. Connectedness network and dependence structure mechanism in green investments. Energy Economics 72, 145–153. https://doi.org/10.1016/j.eneco.2018.04.015

Maghyereh, A.I., Awartani, B., Abdoh, H., 2019. The co-movement between oil and clean energy stocks: A wavelet-based analysis of horizon associations. Energy 169, 895–913.

Managi, S., Okimoto, T., 2013. Does the price of oil interact with clean energy prices in the stock market? Japan and the World Economy 27, 1–9. https://doi.org/10.1016/j.japwor.2013.03.003

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., C++-code), C.-C.C. (libsvm, C++-code), C.-C.L. (libsvm, 2021. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien.

Mullainathan, S., Spiess, J., 2017. Machine Learning: An Applied Econometric Approach. Journal of Economic Perspectives 31, 87–106. https://doi.org/10.1257/jep.31.2.87

Nasreen, S., Tiwari, A.K., Eizaguirre, J.C., Wohar, M.E., 2020. Dynamic connectedness between oil prices and stock returns of clean energy and technology companies. Journal of Cleaner Production 260, 121015. https://doi.org/10.1016/j.jclepro.2020.121015

Neely, C.J., Rapach, D.E., Tu, J., Zhou, G., 2014. Forecasting the Equity Risk Premium: The Role of Technical Indicators. Management Science 60, 1772–1791. https://doi.org/10.1287/mnsc.2013.1838

Nyberg, H., 2011. Forecasting the direction of the US stock market with dynamic binary probit models. International Journal of Forecasting 27, 561–578. https://doi.org/10.1016/j.ijforecast.2010.02.008

Nyberg, H., Pönkä, H., 2016. International sign predictability of stock returns: The role of the United States. Economic Modelling 58, 323–338. https://doi.org/10.1016/j.econmod.2016.06.013

Pástor, L., Veronesi, P., 2012. Uncertainty about Government Policy and Stock Prices. The Journal of Finance 67, 1219–1264. https://doi.org/10.1111/j.1540-6261.2012.01746.x

Pesaran, M.H., Timmermann, A., 2002. Market timing and return prediction under model instability. Journal of Empirical Finance 9, 495–510. https://doi.org/10.1016/S0927-5398(02)00007-5

Pham, L., 2021. Frequency connectedness and cross-quantile dependence between green bond and green equity markets. Energy Economics 98, 105257. https://doi.org/10.1016/j.eneco.2021.105257

Pham, L., 2019. Do all clean energy stocks respond homogeneously to oil price? Energy Economics 81, 355–379. https://doi.org/10.1016/j.eneco.2019.04.010

Pönkä, H., 2016. Real oil prices and the international sign predictability of stock returns. Finance Research Letters 17, 79–87. https://doi.org/10.1016/j.frl.2016.01.011

R Core Team, 2022. R: A Language and Environment for Statistical Computing. The R Project for Statistical Computing, Vienna, Austria.

Reboredo, J.C., 2015. Is there dependence and systemic risk between oil and renewable energy stock prices? Energy Economics 48, 32–45. https://doi.org/10.1016/j.eneco.2014.12.009

Reboredo, J.C., Quintela, M., Otero, L.A., 2017. Do investors pay a premium for going green? Evidence from alternative energy mutual funds. Renewable and Sustainable Energy Reviews 73, 512–520. https://doi.org/10.1016/j.rser.2017.01.158

Reboredo, J.C., Ugolini, A., 2018. The impact of energy prices on clean energy stock prices. A multivariate quantile dependence approach. Energy Economics 76, 136–152. https://doi.org/10.1016/j.eneco.2018.10.012

Sadorsky, P., 2022. Forecasting solar stock prices using tree-based machine learning classification: How important are silver prices? The North American Journal of Economics and Finance 61, 101705. https://doi.org/10.1016/j.najef.2022.101705

Sadorsky, P., 2021a. A Random Forests Approach to Predicting Clean Energy Stock Prices. Journal of Risk and Financial Management 14, 48. https://doi.org/10.3390/jrfm14020048

Sadorsky, P., 2021b. Predicting Gold and Silver Price Direction Using Tree-Based Classifiers. Journal of Risk and Financial Management 14, 198. https://doi.org/10.3390/jrfm14050198

Saeed, T., Bouri, E., Alsulami, H., 2021. Extreme return connectedness and its determinants between clean/green and dirty energy investments. Energy Economics 96, 105017. https://doi.org/10.1016/j.eneco.2020.105017

Schwert, G.W., 1990. Stock Market Volatility. Financial Analysts Journal 46, 23–34. https://doi.org/10.2469/faj.v46.n3.23

Simm, J., Abril, I.M. de, 2014. extraTrees: Extremely Randomized Trees (ExtraTrees) Method for Classification and Regression.

The Economist, 2020. The world’s energy system must be transformed completely. The Economist May 23 edition.

Uddin, G.S., Rahman, M.L., Hedström, A., Ahmed, A., 2019. Cross-quantilogram-based correlation and dependence between renewable energy stock and other asset classes. Energy Economics 80, 743–759. https://doi.org/10.1016/j.eneco.2019.02.014

Ulrich, J., 2020. TTR: Technical Trading Rules.

Wang, Y., Liu, L., Wu, C., 2020. Forecasting commodity prices out-of-sample: Can technical indicators help? International Journal of Forecasting 36, 666–683. https://doi.org/10.1016/j.ijforecast.2019.08.004

Wen, X., Guo, Y., Wei, Y., Huang, D., 2014. How do the stock prices of new energy and fossil fuel companies correlate? Evidence from China. Energy Economics 41, 63–75. https://doi.org/10.1016/j.eneco.2013.10.018

Weng, B., Lu, L., Wang, X., Megahed, F.M., Martinez, W., 2018. Predicting short-term stock prices using ensemble methods and online data sources. Expert Systems with Applications 112, 258–273. https://doi.org/10.1016/j.eswa.2018.06.016

Wright, M.N., Wager, S., Probst, P., 2022. ranger: A Fast Implementation of Random Forests.

Yin, L., Yang, Q., 2016. Predicting the oil prices: Do technical indicators help? Energy Economics 56, 338–350. https://doi.org/10.1016/j.eneco.2016.03.017

Yin, L., Yang, Q., Su, Z., 2017. Predictability of structural co-movement in commodity prices: the role of technical indicators. Quantitative Finance 17, 795–812. https://doi.org/10.1080/14697688.2016.1225977

Zhao, X., 2020. Do the stock returns of clean energy corporations respond to oil price shocks and policy uncertainty? Economic Structures 9, 53. https://doi.org/10.1186/s40008-020-00229-x

1. https://www.theguardian.com/environment/2015/jan/09/solar-power-drives-renewable-energy-investment-boom-2014 [↑](#footnote-ref-1)