Scientific Annals of Economics and Business
67 (SI), 2020, 1-17
DOI: 10.47743/saeb-2020-0024

Using Machine Learning to Profit on the Risk Premium of the Nordic Electricity Futures

Helder Sebastião*, Pedro Godinho**, Sjur Westgaard***

Abstract

This study investigates the use of several trading strategies, based on Machine Learning methods, to profit on the risk premium of the Nordic electricity base-load week futures. The information set is only composed by financial data from January 02, 2006 to November 15, 2017. The results point out that the Support Vector Machine is the best method, but, most importantly, they highlight that all individual models are valuable, in the sense that their combination provides a robust trading procedure, generating an average profit of at least 26% per year, after considering trading costs and liquidity constraints. The results are robust to the different data partitions, and there is no evidence that the profitability of the trading strategies has decreased in recent years. We claim that this market allows for profitable speculation, namely by using combinations of non-linear signal extraction techniques.

Keywords: Nord Pool; electricity futures; risk premium; machine learning; trading.

JEL classification: G13; G14; Q40.

1. INTRODUCTION

There are mainly two types of participants in commodity futures markets. Hedgers, who are involved in the production, commercialization, or consumption of the underlying commodity, use futures to manage the price risk of their existing or anticipated spot positions. While speculators gather information on the future spot price and trade on their expectations.

Speculation may provide two economic benefits. First, the information gathered by speculators is impounded through trading into the futures prices, increasing its informational efficiency and accelerating the spot price discovery process. Second, competition between speculators tends to lower the risk premium of the futures contract, hence decreasing the

* Univ Coimbra, CeBER, Faculty of Economics, Portugal; e-mail: *helderse@fe.uc.pt* (corresponding author).

** Univ Coimbra, CeBER, Faculty of Economics, Portugal; e-mail: *pgodinho@fe.uc.pt*.

*** Norwegian University of Science and Technology, Department of Industrial Economics and Technology Management, Norway; e-mail: *sjur.westgaard@ntnu.no*.

implicit cost of hedging. Therefore, speculation is the main force driving liquidity and informational efficiency in commodity futures.

This paper analyses the profitability of trading strategies on the risk premium of the Nordic Electricity Base Week Futures. The information set used to create those strategies is composed of variables retrieved from the daily spot and futures prices. The paper uses several Machine learning (ML) methods to create forecasts of the signal of the risk premium, according to which speculative trading positions are initiated. These methods are characterized by a high level of flexibility, allowing the consideration of non-linear dynamics. Several papers have already address the issue of forecasting future electricity prices using non-linear models and some of them use ML, however the literature on non-linear models applied to the risk premium of electricity futures is still scarce, and, to the best of our knowledge, this is the first paper to use ML to devise trading strategies on the risk premium of these contracts.

The existence of a risk premium, and even its predictability, does not necessarily imply that there are profitable trading opportunities. An additional step must be taken in devising such trading strategies and proving that they offer economically significant risk-adjusted abnormal returns after controlling for trading costs and liquidity constraints. This paper provides compelling evidence that those trading opportunities do existed for the risk premium of Nordic electricity base-load week futures. More precisely, combinations of ML methods, such as Regression Trees, Random Forests and Support Vector Machines, using a financial information set, from which stands out the futures returns, would have generated a profit of at least 26% per year, in the period from October 2013 to November 2017.

This paper is organized as follows. Section 2 presents a brief literature review. Section 3 presents the Nordic power market and the specification of the Nord Pool electricity base-load week futures contract, describes the data, and provides a preliminary analysis. Section 4 describes the methods used to construct the trading strategies and explains the sample partition used to train, validate, and test them. Section 5 describes the parametrization of the models and present the main results on the performance of the trading strategies. Section 6 conducts several robustness checks, aiming to assess the sensitivity of the performance results to the sample partitioning and possible liquidity constraints. Section 7 concludes the paper.

2. LITERATURE REVIEW

This literature review is three-pronged. First, it looks at the main features and predictability of the risk premium in the Nordic electricity futures. Second, it presents some findings on the importance of financial data, especially futures prices, in forecasting the risk premiums of electricity futures. And third, it gathers supportive information on the ML methods used in our empirical application.

Evidence on the risk premium of Nordic electricity futures and forward contracts can be found in several papers, such as Botterud *et al.* (2002), Mork (2006), Lucia and Torró (2011), Haugom *et al.* (2014), Fleten *et al.* (2015) and Smith-Meyer and Gjolberg (2016). Although these papers show the existence of a negative risk premium, they also highlight its dynamic nature. The risk premium seems to be conditional on the holding period, changes in the cost structure of power producers and market characteristics, such as maturity and liquidity.

Several studies, including Lucia and Torró (2011), conclude that the risk premium in the Nordic weekly futures is persistent and that, to some extent, can be predicted. Fleten *et al.*

(2015) show that the risk premium has persisted for many years for the Nordic and German monthly forward contracts, but its size has reduced recently. High entry costs and low speculation capital (as the market is dominated by a limited number of producers, consumers, and retailers) can be the reasons for a less efficient market. Conversely, Smith-Meyer and Gjolberg (2016) find that after 2008 the Nordic monthly power futures became unbiased and claim that the market has matured and now appears to be at least weak form efficient.

There is a stream of literature that considers the fundamentals to explain the dynamics of the risk premium in the Nordic electricity futures. Amongst these variables, stand out the reservoir levels, deviations in inflow and consumption from a long-term average, and volatility of electricity consumption (Botterud *et al.*, 2002; Cartea and Villaplana, 2008; Lucia and Torró, 2011; Weron and Zator, 2014; Haugom *et al.*, 2018). The framework of this paper is different, in the sense that it only uses financial data, i.e. the emphasis is on the ability of financial data, especially futures prices, to predict the risk premium. This is indirectly documented in these papers, as lagged financial variables are also included in the regression models, with positive results. In fact, some papers highlight the predictive power of futures prices (e.g., Huisman and Kilic, 2012; Paraschiv *et al.*, 2015; Aoude *et al.*, 2016; Ferreira and Sebastião, 2018; Steinert and Ziel, 2019).

Several statistical methods have been proposed for forecasting electricity spot prices. Weron (2014) provides a comprehensive review of these methods, covering standard time series models and some ML methods, such as Neural Networks and SVMs. This last method has been successfully used, for instance, by Gao *et al.* (2007); Zhao *et al.* (2008); Saini *et al.* (2010); Shrivastava *et al.* (2015). Other ML techniques, such as Random Forests (RFs), have also been successfully used to forecast electricity prices (see, for instance, Mei *et al.*, 2014; González *et al.*, 2015; Ludwig *et al.*, 2015; Sadeghi-Mobarakeh *et al.*, 2017). Besides these two methods, we also consider Regression Trees (RT). This method is simpler than the previous ones and is the basis of RFs. The main idea is to assess if simpler methods might also have a good perform in this framework.

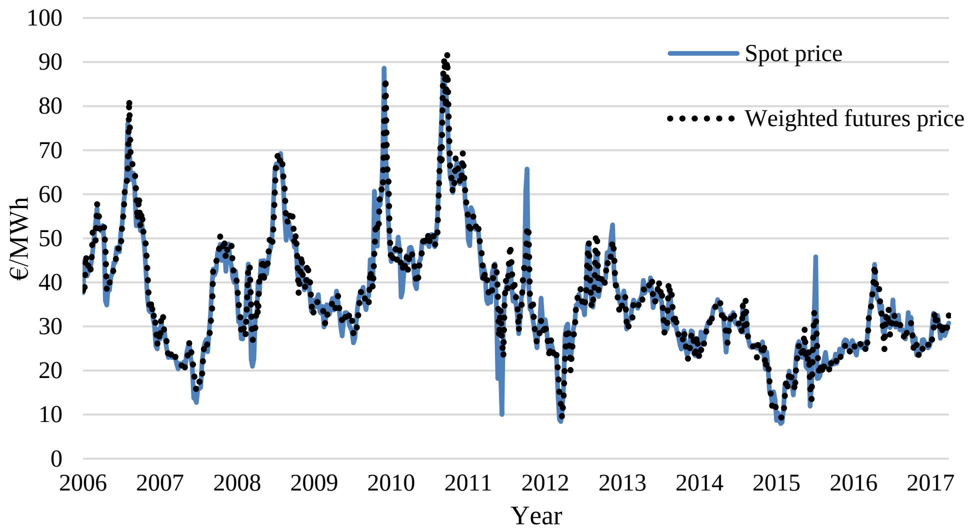
3. DATA AND PRELIMINARY ANALYSIS

This paper investigates the Nordic Electricity Base Week Futures (market designation ENOW[WW]-[YY]). Weekly futures were chosen amongst other maturities due to its sample size, which is large enough to carry out meaningful analyses. These contracts were traded on the Nord Pool ASA exchange until November 01, 2010. At that time, the exchange was bought by the Nasdaq OMX, and changed its name to NASDAQ OMX Oslo ASA, belonging from that date forward to the NASDAQ OMX Commodities Europe.

The underlying asset of these contracts is a supply of electric energy at a constant power of 1 MWh during all hours in a week time for delivery in Norway, Denmark, Sweden and Finland (normally 168 hours, except when changes to or from Daylight Savings Time occur, which imply a delivery period one hour shorter or longer, respectively). These contracts are subjected to cash settlement according to the "Elspot System Price" for the Nordic region, published by Nord Pool Spot AS. The contracts are quoted in EUR, have a tick size of 0.01€, and are traded during Norway business days. Although during the period under scrutiny there were at least six series of weekly futures available for trading and clearing, this paper only considers the nearby contract (assuming a roll over procedure in a weekly basis, at the last trading day, to the next contract). The decision to only use the nearby contract is driven by liquidity concerns, as most of the trading volume is concentrated in the last week before delivery. Hence, there are at most five daily prices for each delivery calendar week. The delivery week is from Monday to Sunday, while the last trading day is the last business day before the delivery week, hence between the last trading day and the beginning of the delivery period there are at least two calendar days (a weekend).

The raw daily spot and futures prices were collected from the Montel AS Database from January 02, 2006 to November 15, 2017, covering a total of 619 weekly calendar deliveries. For each week, the delivery spot price is computed as a simple arithmetic average of the daily spot prices, which in turn, are the simple arithmetic average of all hourly prices (this is the procedure used by the market for the delivery setting). In order to fill in the gaps, the following rule was applied: If the market is closed (holiday), or if it is open but there are no trades, the previous closing price is used.

Figure no. 1 plots the time series of the weekly averages of the spot price and daily closing price of futures contracts in the last trading week. The differences between the spot and futures prices are almost indiscernible at the presenting scale. However, even at this scale, one may see that spot prices are more jagged than futures prices, there is a slightly decreasing trend in the long run, and, from the middle of the sample, the paths of prices are smoother.



Source: Montel AS Database.
Notes: Spot prices are the weekly spot electricity prices (simple average of hourly spot prices in week *T*). Weighted futures prices are the average of electricity base week futures prices traded on Nord Pool. These last series are computed as weighted averages of the daily closing futures prices in the last trading week, i.e. in week *T-1*, using the daily trading volume as weighting scheme. The sample covers the period between January 02, 2006 and November 15, 2017.

Figure no. 1 – Weekly spot and futures prices (Jan. 02, 2006 - Nov. 15, 2017)

Table no. 1 shows some summary statistics of the spot and futures prices, and of the risk premium computed as the difference between the spot price and the last closing futures prices before delivery. As expected, the futures prices and the spot price statistics are quite similar. The mean, median, minimum, and maximum prices only differ a few cents. Prices show mild positive skewness and excess kurtosis and are highly autocorrelated, but there is no evidence on the existence of a unit root in those series. There is no discernible pattern in the statistics of the futures prices considering the time to delivery. For instance, the statistics of the futures price in the last trading day are not always closer to the statistics of the spot price than those of the other daily futures prices, except for the mean. The average futures price decreases as delivery approaches, rendering a decrease in the absolute value of the risk premium. The correlations between the average weekly spot price and the futures prices in the last week before delivery are quite high, with this statistic assuming the value 0.97 for the average, weighted by trading volume, of the futures closing prices.

Table no. 1 – Summary statistics of spot prices and futures prices

	S_T	$F_{0,T}$	$F_{1,T}$	$F_{2,T}$	$F_{3,T}$	$F_{4,T}$	$F_{[0,4],T}$	RP_T
Mean	36.08	36.31	36.39	36.50	36.51	36.58	36.49	-0.24**
Median	34.12	34.30	34.35	34.25	34.40	34.35	34.20	-0.05
Min.	7.95	8.10	8.00	9.25	9.25	9.25	9.09	-16.68
Max.	88.64	90.50	92.30	91.75	92.00	94.50	91.83	13.40
Std. dev.	13.31	13.51	13.51	13.42	13.37	13.39	13.45	2.80
Skewness	0.85	0.91	0.94	0.91	0.90	0.91	0.91	-0.29
Ex. kurt.	1.07	1.23	1.35	1.28	1.24	1.25	1.22	6.78
$\rho(1)$	0.95***	0.96***	0.95***	0.95***	0.95***	0.96***	0.96***	-0.002
$Corr(S_T, F_{t,T})$	--	0.978	0.971	0.965	0.958	0.978	0.971	--
ADF	-3.57**	-3.97***	-4.38***	-3.95***	-4.41***	-4.21***	-3.97***	-12.48***

Source: Montel AS Database.

Notes: This table shows the descriptive statistics on the average weekly spot price, S_T, the futures daily closing price in the last week before delivery, $F_{t,T}$, (t is the number of days before delivery, which occurs at time T (the last trading day is assumed to be $t = 0$), the weighted average by trading volume of the futures closing prices, $F_{[0,4],T}$, and the *ex post* risk premium using the last closing price, that is $RP_T = S_T - F_{0,T}$. The sample period is from January 02, 2006 to November 15, 2017The significance of the mean risk premium is assessed using the t-statistic with Newey-West HAC standard error, with a Bartlett kernel bandwidth of 6. $\rho(1)$ is the first order autocorrelation, $Corr(S_T, F_{t,T})$ is the correlation between the average weekly spot price and the daily futures price in the last week before delivery. ADF is Augmented Dickey-Fuller test on the null hypothesis of a unit root, considering a constant, a time trend and a number of lags chosen by the AIC criterion. The significance of these two last statistics is denoted by *, ** and ***, for the levels of 10%, 5% and 1%, respectively.

The risk premium is negative, with a median value of -0.05€ and an average value of -0.24€ (significant at the 5% level). The risk premium is highly volatile, achieving a minimum of -16.68€ and a maximum of 13.4€, has negative skewness and excess kurtosis (the Jarque-Bera statistic is 1196.41, clearly rejecting the null hypothesis of normality). The risk premium has no significant first order autocorrelation and no unit root.

4. METHODOLOGY

This paper addresses the issue of predicting and trading on the futures risk premium. The analysis starts with linear models, but arguably the relationship between the risk premium and the information set may be non-linear. Since there is no prior information on the structure of these interactions, we resort to three ML methods: Regression Trees (RT), Random Forests (RF) and Support Vector Machines (SVM).

RTs are relatively simple models based on the recursive partition of the space defined by the independent variables into smaller regions. Each partition constitutes a node of the tree and has two branches. When using the tree to make a prediction (in this case, to predict the risk premium), the branch to be chosen depends on the outcome of a test, defined for the values of a given independent variable. For instance, the test based on a variable *v*, can be something like "if v>5 then go to the right branch, else go to left branch of the tree". In making a prediction, the tree is thus read from the first node – the root node. Successive tests are made, and successive branches are chosen, until a terminal node – a leaf node – is reached. This leaf node defines the predicted value of the dependent variable. RTs are usually constructed using a two-step process. The first step consists on growing a large tree, choosing, at each node, the best independent variable and the best condition based on that variable. Such a large tree may lead to overfitting, capturing spurious relationships in the data set used to grow it (the training data), and thus performing badly in new data. To avoid that, a second step is usually applied, which consists of deleting unimportant leaf nodes by a process of statistical estimation ("pruning the tree").

RFs are combinations of RTs, such that each tree is built using just a subset of the training data, which is sampled independently for each tree (Breiman, 2001). This means that, in each tree node, a random subset of the independent variables and a random subset of the observations in the training dataset are used to define the test that will lead to the choice of the branch. RFs forecasts are then an average of the forecasts made by the different trees that compose the "forest". By using subsets of the data and independent variables in each tree node, RFs can reduce the potential overfitting problem.

SVMs can be used for regression or classification tasks. In the first case, the objective is to avoid estimation errors that are larger than a pre-defined value ε. This is achieved by minimizing a function that penalizes the deviations between the predicted and the original values of the output larger than ε. SVMs handle non-linear models by using the "kernel trick". First, the original data is mapped into a new high-dimensional space, where it is possible to apply linear models. Such mapping is based on kernel functions, and SVMs operate on the dual representation induced by those functions. The model is linear in the new space but non-linear in the original data space. Although it is also possible to use the original linear models ("linear kernels"), Gaussian and polynomial kernel functions are commonly used in SVMs. According to Tay and Cao (2001) Gaussian kernels tend to have good performance under general smoothness assumptions.

In our applied work, RTs, RFs and SVMs are implemented in R, using packages rpart (Therneau *et al.*, 2018), which implements the CART algorithm (Breiman *et al.*, 1984), randomForest (Liaw and Wiener, 2002) and e1071 (Meyer *et al.*, 2017), respectively. For a reference on the practical application of these methods in R see Torgo (2016).

In ML applications, data is often split into a training set, used to estimate the different models, a validation set, in which the best model/parameterization is chosen, and a test set, where the results of the best model are assessed. In this work, the main concerns when defining the different data subsets were to avoid all risks of data snooping and to make sure that the results obtained in the test set can be considered representative. Figure no. 2 illustrates the sample partition used in the present work.

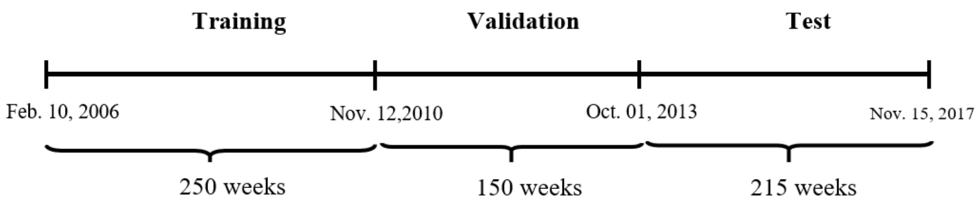


Figure no. 2 – Partition of the overall sample into training, validation, and test periods

The results may depend on the data partition used, therefore we opted for two other partitions in order to check the robustness of the results (see Section 6). These partitions consider different sizes for the test set, roughly keeping the ratio between the sizes of the training and validation sets. The first alternative partition uses 187/113/ 315 observations and the second one uses 312/188/115 observations, respectively.

Choosing a model means defining a method, a set of parameters and a set of explanatory variables. Parameter setting in ML applications is usually performed by trying many different configurations and choosing the one that leads to the best results. However, the main goal of this study is not to extensively test the alternative parameterization of different methods, but instead to simply find out if ML can, in general, lead to profitable strategies. So, only a small number of alternative parameterizations are initially chosen to be tested. In the case of linear regressions and RTs, there are no alternative parameterizations, simply different sets of explanatory variables. For the RFs, it is used 500, 1000 and 1500 trees and the number of variables randomly sampled at each split is initially set to one half and one third of the total number of explanatory variables. In the case of SVMs, the only alternative parameterizations consist of using different kernels: Linear Gaussian, and polynomial.

5. RESULTS

To get a first look at the potential importance of the variables, a linear regression is estimated using the first 400 observations (corresponding to the training and validation samples). In this regression the dependent variable is the logarithmic risk premium and the information set is composed by the logarithmic returns of the futures contracts in the last four trading days, the logarithmic spot price in the previous week, the logarithmic premium lagged one to four weeks, and dummies for the quarter of the year to which the first day of the delivery week belongs. The results are shown in Table no. 2.

It is noteworthy that this simple linear regression achieves a coefficient of determination of around 9% using just financial data for the period between February 10, 2006, and September 27, 2013. This is an interesting result. Similar regressions, using not only financial data but also fundaments, achieved coefficients of determination between 3% and 22% – Botterud *et al.* (2010) obtained a $R^2 = 0.22$ for the period 1996-2006; Weron and Zator (2014), obtained a $R^2 = 0.19$ for 1998-2010, and Haugom *et al.* (2018) got a R^2 around 0.03 for 2004-2013.

Table no. 2 – Linear regression of the logarithmic risk premium on financial and deterministic variables

	Coefficient	Std. Error
Constant	0.015	0.057
$r_{0,T}$	0.988***	0.165
$r_{1,T}$	-0.138	0.153
$r_{2,T}$	0.045	0.161
$r_{3,T}$	0.004	0.146
s_{T-1}	-0.007	0.015
rp_{T-1}	0.032	0.066
rp_{T-2}	0.019	0.051
rp_{T-3}	0.077	0.049
rp_{T-4}	-0.100*	0.049
2nd Quarter	0.015	0.013
3rd Quarter	-0.012	0.013
4th Quarter	-0.003	0.013
$\mathbf{R^2}$	0.092	
F-statistic	5.000***	

Notes: This table shows the results of the linear regression for the logarithmic risk premium $rp_T = \ln(S_T) - \ln(F_{0,T})$ on the daily logarithmic returns of the futures contract, with delivery at week T, in the last four trading days, i.e. $r_{t,T} = \ln(F_{t,T}) - \ln(F_{t+1,T})$, with t = 0, 1, 2, 3, 4, (the last trading day is denoted by t = 0), the logarithmic spot price in the previous week, $s_{T-1} = \ln(S_{T-1})$, the logarithmic premium up to one to four weeks before, rp_{T-j}, with j = 1, 2, 3, 4, and quarterly dummies, that assume the unit value if the first day used for calculating the spot belongs to that period of the year. The data used in this regression is from February10, 2006 to September 27, 2013, corresponding to the training and validation periods. Significance at the 10%, 5% and 1% levels is denoted by *, ** and ***, respectively.

These results highlight that the return of the futures contract in the last trading day and the logarithmic premium one month before are the most significant variables in explaining the risk premium. Keeping in mind that it is important to test a variety of explanatory variables, the input information was divided into 11 subsets, considering different combinations of the logarithmic futures returns t days before delivery occurring at time T, i.e. $r_{t,T}$, with t = 0, 1, 2, 3, 4, (the last trading day is denoted by t = 0), the logarithmic spot price one week prior to delivery, s_{T-1}, the logarithmic risk premium j weeks prior to the delivery week T, rp_{T-j}, with j = 1, 2, 3, 4, and quarter dummies which assume the unity if the first day of T belongs to that quarter of the year, Q_s. The most restricted sets are $\{r_{0,T}, rp_{T-4}, s_{T-1}\}$ and $\{r_{0,T}, rp_{T-4}, Q_s\}$, while the most enlarged set has all the variables described above.

Each method is run for each set of parameters and for each of those 11 sets of variables. For each observation in the validation sample, a model is estimated using the previous 250 observations – that is, using a rolling window of 250 observations. The next step is to compute the returns of a trading strategy that uses the sign of the risk premium forecast to devise a position in the futures market in the last trading day. If the risk premium forecast is positive (negative) the strategy prescribes a long (short) position in the futures contract, which is held during the delivery period, hence capturing the risk premium (symmetrical of the risk premium). At this point of the analysis, it is assumed that all prescribed trades are feasible, and hence that there are no liquidity constraints. The models are assessed using the time series of 150 outcomes (the number of observations in the validation sample).

For each method, the combination of the set of parameters and the set of variables that lead to the best performance are chosen according to the average return per trade during the validation sample. Because the models always prescribe a position in the futures markets, these values can also be interpreted as weekly averages. Table no. 3 presents, for each method, the parameters and set of variables to be used afterwards in the test sample.

Table no. 3 – Models, parameters and sets of variables to be used in the test sample

Method	Parameters	Set of variables	Weekly Average Return (Validation sample)
Linear	-	$r_{0,T}$, rp_{T-4}, s_{T-1}	2.83%
RT	-	$r_{0,T}$, rp_{T-4}, Q_s	2.16%
RF	1000 trees, 1/3 of the variables at each split	$r_{0,T}$, rp_{T-4}, s_{T-1}, Q_s	2.26%
SVM	Linear kernel	$r_{0,T}$, rp_{T-4}, s_{T-1}, Q_s	2.75%

Notes: The last column shows the weekly average return obtained during the validation sample (November 15, 2010 to September 27, 2013). The parameters and variables sets are chosen according to the weekly average return. $r_{0,T}$ is the logarithmic futures return in the last trading day before delivery week *T*, rp_{T-4} is the logarithmic risk premium 4 weeks prior to *T*, s_{T-1} is the logarithmic spot price one week prior to delivery and Q_s are quarter dummies that assume the unity if the first day of *T* belongs to that quarter of the year.

The analysis is also conducted on two additional "ensemble models", based on the four initial models – Linear, RT, RF and SVM. Ensemble 3 and Ensemble 4 consider that a trade is made only if at least three and if all four models agree on the sign of the risk premium, respectively. Hence, while trades prescribed by individual models occur in all weeks, in these ensemble models there are several weeks without trades.

In the test sample, only the sets of parameters and variables with the best performance in the validation sample are considered. Again, the models used to obtain the 1-step forecast at each observation point are estimated using a rolling window consisting of the previous 250 observations. The performances of the models, without trading costs and with a proportional trading cost of 0.5%, are presented in Table no. 4 and Table no. 5, respectively. Notice that the trading strategies only involve taking one position in the futures market that is cleared at delivery. According to the information obtained from a broker, this would imply 0.0075€/MWh in explicit transaction cost and 0.01€/MWh in clearing costs. This is just about 0.0175/36.31 ≅ 0.05% of the average futures price at the last trading day. We decide to be conservative and assume a proportional trading cost of 0.5%. This is the most used figure for round-trip trading costs in the stock market and it is a higher value than the one used by most researchers when dealing with round-trip costs in futures markets (see, for instance, Lubnau and Todorova, 2015). Notice that in our case there is no bid-ask spread or price slippage (price impact) when the trading position is closed.

The trading strategies are compared with a Naïve strategy that serves as a benchmark. This strategy assumes that the risk premium will always have the same sign as its unconditional average during the test sample. In the test sample, the average risk premium is 0.32%, therefore, the naïve trader is always long in the futures contract and the strategy has an average profit equal to that value. The strategy is not so naïve as one would think at a first glance, because it is built on the assumption that the investor has a perfect expectation about the sign of the future average of the risk premium[1].

Table no. 4 – Performance in the test sample (without trading costs)

Model	% Profitable trades (# trades)	Average profit per trade (%)	Standard deviation of profit (%)	Skewness of profit	Sharpe ratio (%)	Average profit per year (%)
Naïve (long position)	53.02 (215)	0.32	7.03	0.92	4.51	17.92
Linear	55.35 (215)	0.84	6.98	-0.73	12.01	54.66
RT	52.56 (215)	0.18	7.03	-0.79	2.60	9.98
RF	59.53 (215)	1.07	6.95	0.21	15.35	74.15
SVM	60.00 (215)	1.18*	6.93	0.51	17.01	84.63
Ensemble3	58.66 (179)	1.19*	6.73	-0.17	17.61	67.04
Ensemble4	65.91 (88)	1.58**	6.14	-0.74	25.77	40.00

Notes: This table shows the performance of the trading strategies in the test sample (October 01, 2013 – November 15, 2017). The average, standard deviation, skewness of profit and Sharpe ratio are computed per trade. The average profit per year is computed using the accumulated return in the test period (52 weeks) and indicates the expected annual growth rate that compounds to the total accumulated return. The two last rows refer to the ensemble classification models based on the first four models. In these ensemble models a trade is only made if at least 3 models (Ensemble3) or if the four models (Ensemble4) give the same indication. The models are compared with the Naïve trading strategy that always considers a long position in the futures contract (the significance at the 10%, 5%, and 1% levels are denoted by *, **, ***, respectively). The significance levels are obtained using 20000 bootstrap samples, created with the circular block procedure of Politis and Romano (1994), with an optimal block size chosen according to Politis and White (2004, 2009).

The ranking of the strategies, according to the Sharpe ratio is the following: 1st Ensemble4, 2nd Ensemble3, 3rd SVM, 4th RF, 5th Linear and 6th RT. The RT is the only model that has a worse performance than the Naïve strategy. For the other models, the percentage of successful trades ranges from 55.35% for the Linear model to 65.91% for the Ensemble4, the average profit per trade ranges from 0.84% for the Linear model to 1.58% for the Ensemble4, and the standard deviation per trade ranges from 6.14% for the Ensemble4 to 6.98% for the Linear model. The comparison of the trading strategies reveals an interesting pattern: A higher average return is associated with a lower risk (measured by the standard deviation per trade). The profit asymmetry is low, and it is only positive for the Naïve, RF and SVM models. In terms of average profit per trade, only the last three models are statistically better than the Naïve strategy: the SVM and the Ensemble3 are statistically better at the 10% significance level and the Ensemble4 is significantly better at the 5% significance level. Excluding the RT strategy, the average profit per year presents high values, ranging from 40.00% for the Ensemble4 strategy to 84.63% for the SVM strategy. The most robust strategy, Ensemble4, has a lower performance in terms of yearly profit than all the other successful strategies due to the reduced number of trades (only 88 trades). But one should notice that a yearly profit of 40% amounts to an accumulated return of 402% during the 4 years of the test sample. This is clearly an indication that the strategies are successful in predicting the sign of the risk premium in the Nordic electricity base week futures. But this only means profitability once trading costs are considered (see Table no. 5).

Table no. 5 – Performance in the test sample (proportional trading costs of 0.5%)

Model	% Profitable trades (# trades)	Average profit per trade (%)	Sharpe ratio per trade (%)	Average profit per year (%)
Naïve (long position)	49.77 (215)	-0.18	-2.61	-9.08
Linear	50.70 (215)	0.34	4.85	19.25
RT	49.77 (215)	-0.32	-4.51	-15.20
RF	55.81 (215)	0.57	8.15	34.28
SVM	55.81 (215)	0.68*	9.80	42.36
Ensemble3	55.31 (179)	0.69*	10.18	34.53
Ensemble4	62.50 (88)	1.08**	17.62	25.87

Notes: This table shows the performance of the trading strategies in the test sample (October 01, 2013 – November 15, 2017) after trading costs, i.e. after deducting 0.5% of the contract price from the profit series. The average profit and Sharpe ratio are computed per trade. The average profit per year is computed using the accumulated return in the test period (52 weeks), and indicates the expected annual growth rate that compounds to the total accumulated return. The two last rows refer to the ensemble classification models based on the first four models. In these ensemble models a trade is only made when at least 3 models (Ensemble3) or the four models (Ensemble4) give the same indication. Average profits per trade better than those of the Naïve strategy at the significance levels of 10%, 5% and 1% are denoted by *, **, ***, respectively. These significance levels are obtained using 20000 bootstrap samples, created with the circular block procedure of Politis and Romano (1994), with an optimal block size chosen according to Politis and White (2004, 2009).

With proportional trading costs of 0.5%, the average profit per trade of the Naïve strategy becomes negative, around -0.18%. This implies that it is better for the naïve investor to stay-out of the market. The ranking of the strategies is the same as before without trading costs, but obviously these costs lower the percentage of profitable trades, the average profit per trade, the Sharpe ratio and the yearly average profit of all strategies (the standard deviation and the skewness of profits remain the same). After trading costs, only the SVM, Ensemble3 and Ensemble4 models perform significantly better than the Naïve strategy. Ensemble4 continues to be the most robust model. The difference between the yearly profit of Ensemble4 and the other successful models decreases substantially, because the overall trading costs SVM are lower due to the lower number of trades. The average profit per year after trading costs of Ensemble4 is 25.87%, which accumulates to a total return of 259% during the four years of the test period.

6. ROBUSTNESS CHECK

The robustness of the previous results is tested on two aspects: the stability of results across the overall sample, which depends on the partition of the data, and the impact of eventual liquidity constraints. The first issue is addressed by considering two alternative data partitions besides the baseline partition (250/150/215 weeks). The first alternative split into training/validation/test samples is 187/113/315 weeks and the second one is 312/188/115 weeks.

One may argue that the Nordic electricity futures is a thin traded market, and therefore execution risk is high. In this sense, our results are not valid for any trade size, as the profitability of the trade depends on the price slippage produced by the trade itself. A rigorous robustness test to this type of implicit trading cost is only possible with order book information (quotes and quantities) during the last trading day. We do not have that data, but we can still conduct some tests using proxies for the execution prices. In order to guarantee that the trading strategies are feasible throughout the last trading session, in each week the forecast of the risk premium is obtained at the beginning of the last trading day, using the futures return open-to-open instead of the return close-to-close, i.e., using the open price in day $t = 1$ and the open price in day $t = 0$. This implies that the return in the last trading day, $r_{0,T}$, is measured with an error.

A first test, called Scenario 1, considers the following execution rules: (i) if trading volume in the last trading day before delivery is less than 10 contracts, there is no trade; (ii) if the model prescribes a position in the futures market, the opening trade occurs at the opening price, and, (iii) the profit of each trade is obtained after deducting proportional trading costs of 0.5%. The second test, called Scenario 2 is a worst-case scenario, with the following execution rules: (i) if trading volume in the last trading day is less than 10 contracts, there is no trade; (ii) if the model prescribes a position in the futures market, the opening trade occurs at the worst daily price, i.e., if the model signals a long (short) position, the highest (lowest) price of the day is used, and, (iii) the profit of each trade is obtained after deducting explicit trading costs of 0.0175€/MWh. Table no. 6 presents the profit results of the trading strategies in the two scenarios and considering the three data splits.

Panel A of Table no. 6 considers the initial data split. In the initial test sample there are 10 last trading days with less than 10 traded contracts, hence the number of trades for the first 4 models in the worst-case scenario is only 205, but some of these days are non-trade days for Ensemble3 and Ensemble4, and the number of trades are 170, i.e. minus 9 trades, and 86, i.e. minus 2 trades, than in the baseline case (Table no. 5). Given these small differences, the percentage of profitable trades increase and decrease slightly in Scenario 1 and Scenario 2, respectively, with the exception being Ensemble4, for which the percentage of profitable trades decrease -0.87% in Scenario 1 and -5.52% in Scenario 2. The average profit per trade and per year in Scenario 1 increase noticeably, except for Naïve, in which the profits decrease marginally, and for Ensemble4, in which they only increase slightly (for this strategy the average profit per year increases from 25.87% in the baseline case to 27.49% in Scenario 1). The results in Scenario 2 are worse for most strategies. However, it is surprising to see that, even in this worst-case scenario, the RF strategy has an increase in the yearly profit from 34.28% to 45.17%, and the Ensemble 3 strategy has an increase from 34.53% to 48.42%. The yearly profit of Ensemble4 strategy decreases to almost half in Scenario 2, achieving the value of 13.99%, but the profit per trade is still significantly better than that of the Naïve strategy, at the level of 10%. The removal of the best and worst 10% trades decrease the average profit per trade, except for the SVM model in Scenario 2, without questioning the profitability of the strategies that are profitable in the baseline case. This result also highlights that the profitability of the strategies is not due to just a few extreme values in the profit per trade series.

Table no. 6 – Performance in different test samples with trading costs and liquidity constrains

	Scenario 1			Scenario 2		
Model	**% Profitable trades (# trades)**	**Avg. profit per trade (%)**	**Avg. profit per year (%)**	**% Profitable trades (# trades)**	**Avg. profit per trade (%)**	**Avg. profit per year (%)**
Panel A - Training: 250 weeks; Validation: 150 weeks; Test: last 215 weeks						
Naïve (long position)	50.24 (205)	-0.16 (-0.24)	-8.10	48.29 (205)	-0.50 (-0.54)	-22.77
Linear	52.68 (205)	0.53 (0.40)	31.58	49.76 (205)	0.08 (0.02)	4.06
RT	52.68 (205)	0.02 (-0.03)	0.97	49.76 (205)	-0.40 (-0.44)	-18.98
RF	56.59 (205)	1.04** (0.74)	72.06	53.66 (205)	0.72** (0.48)	45.17
SVM	58.05 (205)	0.79* (0.77)	50.73	55.12 (205)	0.44* (0.52)	25.98
Ensemble3	60.00 (170)	1.32** (1.12)	77.00	56.47 (170)	0.92*** (0.82)	48.42
Ensemble4	61.63 (86)	1.11* (0.96)	27.49	56.98 (86)	0.60* (0.52)	13.99
Panel B - Training: 187 weeks; Validation: 113 weeks; Test: last 315 weeks						
Naïve (long position)	51.15 (305)	-0.15 (-0.07)	-7.41	49.84 (305)	-0.59 (-0.41)	-26.40
Linear	53.11 (305)	0.53 (0.30)	31.85	50.16 (305)	0.06 (-0.01)	3.31
RT	53.44 (305)	0.32 (0.16)	18.10	51.48 (305)	-0.16 (-0.16)	-7.80
RF	53.11 (305)	0.41 (0.23)	23.45	50.49 (305)	-0.08 (-0.13)	-3.94
SVM	54.10 (305)	0.39 (0.21)	22.56	51.15 (305)	-0.07 (-0.12)	-3.79
Ensemble3	57.20 (243)	0.73* (0.59)	35.37	54.32 (243)	0.19 (0.22)	8.06
Ensemble4	59.84 (127)	1.53** (1.00)	39.13	56.69 (127)	0.94** (0.63)	22.62
Panel C - Training: 312 weeks; Validation: 188 weeks; Test: last 115 weeks						
Naïve (long position)	52.78 (108)	0.21 (0.11)	11.73	50.00 (108)	-0.16 (-0.19)	-8.04
Linear	62.96 (108)	1.29* (1.13)	95.38	57.41 (108)	0.99* (0.90)	67.59
RT	53.70 (108)	0.15 (0.03)	8.00	49.07 (108)	-0.07 (-0.15)	-3.81
RF	58.33 (108)	0.50 (0.59)	29.77	52.78 (108)	0.22 (0.29)	11.95
SVM	63.89 (108)	1.11 (1.13)	78.00	59.26 (108)	0.81 (0.92)	52.50
Ensemble3	63.37 (101)	1.23* (1.07)	82.13	57.43 (101)	0.92* (0.79)	56.60
Ensemble4	66.67 (63)	1.05 (1.05)	37.38	60.32 (63)	0.84 (0.87)	28.83

Notes: This table shows the trading performance considering trading costs and liquidity constrains, and different splits of the series into training, validation, and test periods. The trading strategies are devised upon the forecasts of the sign of the risk premium using the open-to-open return of the futures contract in the last trading day. If the trading volume in the last trading day before delivery is less than 10 contracts, there is no trade. Scenario 1 is defined by the following additional rules: if the model prescribes a position in the futures market, the opening trade occurs at the opening price, and, the profit of each trade is obtained after deducting implicit trading costs of 0.5%. Scenario 2 is defined by the following additional rules: if the model prescribes a position in the futures market, the opening trade occurs at the worst daily price, i.e., if the model signals a long (short) position, the highest (lowest) price of the day is used, and the profit of each trade is obtained after deducting explicit trading costs of 0.0175€/MWh. Below the average profit per trade, in parenthesis, is the average profit per trade after removing the highest and lowest 10% profits (no significance tests are performed for these values). The average profit per year is computed using the accumulated return in the test period (52 weeks) and indicates the expected annual growth rate that compounds to the total accumulated return. The two last rows of each panel refer to the ensemble classification models based on the four individual models. In these ensemble models a trade is only made if at least 3 models (Ensemble3) or if all models (Ensemble4) give the same indication. Average profits per trade better than those of the Naïve strategy (always being long in the futures contract) at the significance level of 10%, 5%, and 1% are denoted by *, **, ***, respectively. These significance levels are obtained using 20000 bootstrap samples, created with the circular block procedure of Politis and Romano (1994), with an optimal block size chosen according to Politis and White (2004, 2009).

Other data splits (Panel B and Panel C of Table no. 6) reinforce two earlier important claims. First, Ensemble4 is quite a robust procedure. The average profit per year is higher when the size of the test sample increases to 315 weeks or decreases to 115 weeks, in comparison with the initial test sample size. In these two other data splits, the profits per year of Ensemble4 are 39.13% and 37.38% in Scenario 1 and 22.62% and 28.83% in Scenario 2, which are higher than the values of 27.49% and 13.99% obtained in the initial test sample. Second, profits are generally higher in the last split, which considers only the last 115 weeks, so there is no reason to believe that the profitability of these strategies, in particular Ensemble4, has decreased in recent years.

7. CONCLUSIONS

This paper gives some insights on predicting the sign of the risk premium of the Nordic electricity base-load week futures and checks if this predictability can be appropriated by speculative trading on the risk premium. The study uses daily financial data in the last trading week from January 02, 2006, until November 15, 2017, covering a total of 619 weekly calendar deliveries.

The risk premium, at the last trading day of the futures contracts, is negative with a median value of just -0.05€/MWh and an average value of -0.24€/MWh. The properties of the risk premium have changed considerably in the sample. For instance, the mean risk premium is -1.32% (significant at the 1% level) from January 02, 2006 to November 12, 2010, it increases to -1.05% from November 15, 2010 to September 27, 2013, and turn positive in the last two years, when it is 0.32%. So, there is a substantial change in the risk premium and the market, which was in contango in the early years, went to backwardation in the last years.

The linear regression of the weekly logarithmic risk premium on a small set of financial variables obtains a coefficient of determination of around 9%, with the futures returns in the last day achieving the highest contribution to explaining the variability of the risk premium. This is better than initially expected, as some analogous models in the literature, which consider not only financial but also fundamental information, achieved on average similar results.

The analysis on predicting the sign of the risk premium begin with linear models. However, we conjecture that the relationship between the explanatory variables and the risk premium may be more complex, and resort to three ML techniques: Regression Trees (RTs), Random Forests (RFs) and Support Vector Machines (SVMs).

The results show clearly that SVM provides the best of those models, but, most importantly, it is shown that all models produce valuable information, in the sense that its combination produces a robust procedure to devise a trading strategy designed to capture the risk premium. From 2013 until 2017, this strategy would have a yearly profit of roughly 26%, after considering a proportional trading cost of 0.5%. Arguably this figure overstates the trading costs, as the strategies aiming to capture de risk premium of the futures contracts only imply one weekly trade in the futures market.

These results were subjected to several robustness checks. Basically, we have considered two additional data splits with 315 and 115 weeks in the test sample (besides the baseline split of 215 weeks) and devised a worst-case scenario. In this scenario investors gather the information at the opening of the last trading session before the delivery week, (this implies that the last daily return of the futures contract – the most important variable in all models – is measured with an error using open-to-open prices) and the position in the futures market is open at the worst price of the day (the daily high if it is long, the daily low if it short). For the individual models, the results in this worst-case scenario, in the different data splits, are mixed. However, the Ensemble4 strategy shows a robust performance, achieving an average return per year even higher in the additional data splits.

The smoothness of the risk premium in the last half of the sample suggests that the Nordic power market has matured since 2013. But this also implies that the performance of our trading strategies was subjected to a harsh test. Hence, we may conclude that there is still some space for profitable speculative trading on the risk premium, and most particularly for the use of trading schemes built on the combination of non-linear signal extraction techniques.

Acknowledgements

This work has been funded by national funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., Project UIDB/05037/2020.

ORCID

Helder Sebastião http://orcid.org/0000-0002-1743-6869
Pedro Godinho http://orcid.org/0000-0003-2247-7101

References

Aoude, S., De Mello, L., and Truck, S., 2016. *Electricity futures markets in Australia - An analysis of risk premiums during the delivery period*. Paper presented at the Meeting Asia's Energy Challenges, 5th IAEE Asian Conference. International Association for Energy Economics. https://www.mq.edu.au/__data/assets/pdf_file/0007/583288/working_paper_16-03.pdf.

Botterud, A., Bhattacharyya, A. D., and Ilic, M. D., 2002. *Futures and spot prices – An analysis of the Scandinavian electricity market.* Paper presented at the 34th Annual North American Power Symposium (NAPS 2002).

Botterud, A., Kristiansen, T., and Ilic, M. D., 2010. The relationship between spot and futures prices in the Nord Pool electricity market. *Energy Economics, 32*(5), 967-978. http://dx.doi.org/10.1016/j.eneco.2009.11.009

Breiman, L., 2001. Random Forests. *Machine Learning, 45*(1), 5-32. http://dx.doi.org/10.1023/A:1010933404324

Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A., 1984. *Classification and Regression Trees*: CRC press.

Cartea, A., and Villaplana, P., 2008. Spot price modelling and the valuation of electricity forward contracts: The role of demand and capacity. *Journal of Banking & Finance, 32*(12), 2502-2519. http://dx.doi.org/10.1016/j.jbankfin.2008.04.006

Ferreira, M., and Sebastião, H., 2018. The Iberian electricity market: Analysis of the risk premium in an illiquid market. *Journal of Energy Markets, 11*(2), 1-22. http://dx.doi.org/10.21314/JEM.2018.176

Fleten, S. E., Hagen, L. A., Nygard, M. T., Smith-Sivertsen, R., and Sollie, J. M., 2015. The overnight risk premium in electricity forward contracts. *Energy Economics, 49*, 293-300. http://dx.doi.org/10.1016/j.eneco.2014.12.022

Gao, C., Bompard, E., Napoli, R., and Cheng, H., 2007. Price forecast in the competitive electricity market by Support Vector Machine. *Physica A, 382*(1), 98-113. http://dx.doi.org/10.1016/j.physa.2007.03.050

González, C., Mira-McWilliams, J., and Juarez, I., 2015. Important variable assessment and electricity price forecasting based on regression tree models: Classification and regression trees, bagging and random forests. *IET Generation, Transmission & Distribution, 9*(11), 1120-1128. http://dx.doi.org/10.1049/iet-gtd.2014.0655

Haugom, E., Hoff, G. A., Mortensen, M., Molnar, P., and Westgaard, S., 2014. The forecasting power of medium-term futures contracts. *Journal of Energy Markets, 7*(4), 47-69. http://dx.doi.org/10.21314/jem.2014.108

Haugom, E., Hoff, G. A., Mortensen, M., Molnar, P., and Westgaard, S., 2018. The forward premium in Nord Pool power market. *Emerging Markets Finance & Trade, 54*(8), 1793-1807. http://dx.doi.org/10.1080/1540496X.2018.1441021

Huisman, R., and Kilic, M., 2012. Electricity futures prices: Indirect storability, expectations, and risk premiums. *Energy Economics, 34*(4), 892-898. http://dx.doi.org/10.1016/j.eneco.2012.04.008

Liaw, A., and Wiener, M., 2002. Classification and regression by randomForest. *R News, 2*(3), 18-22.

Lubnau, T., and Todorova, N., 2015. Trading on mean-reversion in energy futures markets. *Energy Economics, 51*, 312-319. http://dx.doi.org/10.1016/j.eneco.2015.06.018

Lucia, J. J., and Torró, H., 2011. On the risk premium in Nordic electricity futures prices. *International Review of Economics & Finance, 20*(4), 750-763. http://dx.doi.org/10.1016/j.iref.2011.02.005

Ludwig, N., Feuerriegel, S., and Neumann, D., 2015. Putting big data analytics to work: Feature selection for forecasting electricity prices using the LASSO and Random Forests. *Journal of Decision Systems, 24*(1), 19-36. http://dx.doi.org/10.1080/12460125.2015.994290

Mei, J., He, D., Harley, R., Habetler, T., and Qu, G., 2014. *A random forest method for real-time price forecasting in new york electricity market*. Paper presented at the PES General Meeting Conference & Exposition, 2014 IEEE. http://dx.doi.org/10.1109/PESGM.2014.6939932

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., and Leisch, F., 2017. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. from https://CRAN.R-project.org/package=e1071

Mork, E., 2006. The dynamics of risk premiums in Nord Pool's futures market. *Energy Studies Review, 14*(1), 170-185. http://dx.doi.org/10.15173/esr.v14i1.485

Paraschiv, F., Fleten, S. E., and Schurle, M., 2015. A spot-forward model for electricity prices with regime shifts. *Energy Economics, 47*, 142-153. http://dx.doi.org/10.1016/j.eneco.2014.11.003

Politis, D. N., and Romano, J. P., 1994. The stationary bootstrap. *Journal of the American Statistical Association, 89*(428), 1303-1313. http://dx.doi.org/10.1080/01621459.1994.10476870
Politis, D. N., and White, H., 2004. Automatic block-length selection for the dependent bootstrap. *Econometric Reviews, 23*(1), 53-70. http://dx.doi.org/10.1081/etc-120028836
Politis, D. N., and White, H., 2009. Correction to "Automatic block-length selection for the dependent bootstrap. *Econometric Reviews, 28*(4), 372-375. http://dx.doi.org/10.1080/07474930802459016
Sadeghi-Mobarakeh, A., Kohansal, M., Papalexakis, E. E., and Mohsenian-Rad, H., 2017. *Data mining based on random forest model to predict the California ISO day-ahead market prices*. Paper presented at the Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), 2017 IEEE. http://dx.doi.org/10.1109/ISGT.2017.8086061
Saini, L. M., Aggarwal, S. K., and Kumar, A., 2010. parameter optimisation using genetic algorithm for support vector machine-based price-forecasting model in national electricity market. *IET Generation, Transmission & Distribution, 4*(1), 36-49. http://dx.doi.org/10.1049/iet-gtd.2008.0584
Shrivastava, N. A., Khosravi, A., and Panigrahi, B. K., 2015. Prediction interval estimation of electricity prices using PSO-tuned support vector machines. *IEEE Transactions on Industrial Informatics, 11*(2), 322-331. http://dx.doi.org/10.1109/TII.2015.2389625
Smith-Meyer, E., and Gjolberg, O., 2016. The Nordic futures market for power: Finally mature and efficient? *Journal of Energy Markets, 9*(4), 1-15. http://dx.doi.org/10.21314/jem.2016.151
Steinert, R., and Ziel, F., 2019. Short-to mid-term day-ahead electricity price forecasting using futures. *Energy Journal, 40*(1), 105-127. http://dx.doi.org/10.5547/01956574.40.1.rste
Tay, F. E., and Cao, L., 2001. Application of support vector machines in financial time series forecasting. *Omega, 29*(4), 309-317. http://dx.doi.org/10.1016/s0305-0483(01)00026-3
Therneau, T., Atkinson, B., and Ripley, B., 2018. rpart: Recursive Partitioning and Regression Trees. from https://CRAN.R-project.org/package=rpart
Torgo, L., 2016. *Data Mining with R: Learning with Case Studies*: CRC press. http://dx.doi.org/10.1201/9781315399102
Weron, R., 2014. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting, 30*(4), 1030-1081. http://dx.doi.org/10.1016/j.ijforecast.2014.08.008
Weron, R., and Zator, M., 2014. Revisiting the relationship between spot and futures prices in the Nord Pool electricity market. *Energy Economics, 44*, 178-190. http://dx.doi.org/10.1016/j.eneco.2014.03.007
Zhao, J. H., Dong, Z. Y., Xu, Z., and Wong, K. P., 2008. A Statistical approach for interval forecasting of the electricity price. *IEEE Transactions on Power Systems, 23*(2), 267-276. http://dx.doi.org/10.1109/TPWRS.2008.919309

Notes

[1] Initially, we also considered two additional naïve strategies (a first one that assumed thatthe risk premium in week *T* would have the same sign as the average risk premium over the previous 250 weeks, i.e. the strategy uses the sign of a long run moving average to signal the trade, and a second one that assumed that the risk premium would have the same sign as the realized risk premium during the previous week). These two naïve strategies were significantly worse than the former naïve strategy and hence their results are not reported here.