39th International Symposium on Forecasting (ISF), 2019

Application of a SVM-based model for day-ahead electricity price prediction for the single electricity market in Ireland

C. Lynch[1,*], J. Kehoe, R. Bain, F. Zhang, J. Flynn, C. O'Leary, G. Smith, R. Linger, K. Fitzgibbon, F. Feijoo[**]

[*] *Nimbus Research Centre, Cork Institute of Technology, Bishopstown, Co. Cork, Ireland*
[**] *School of Industrial Engineering, Pontificia Universidad Católica de Valparaiso, Chile*

***Abstract* — This paper espouses an innovative approach to predict the day-ahead electricity prices for the single electricity market (SEM) in Ireland. An upsurge in demand response and the proliferation of distributed energy resources continue to drive the requirement for more accurate and computationally efficient models for forecasting the system marginal price (SMP).**

This paper presents such a model, a k-means, Support Vector Machine and Support Vector Regression (k-SVM-SVR) model. While obtaining prediction accuracy comparable with the best-known models in the literature, the k-SVM-SVR model requires limited computational effort. The computational efficiency is achieved by eliminating the use of a price feature selection process, which is commonly used by existing models in the literature. The developed model achieved approximately 20% improvement and reduced error variances over the existing predictions available to market participants in Ireland.

The k-SVM-SVR model is tested using SMP electricity market data from the periods 2010–11, 2015-16 and 2016–17 respectively.

Keywords: Energy Price Forecasting, Single Electricity Market (SEM) and Single Marginal Price (SMP)

I. Introduction

Deregulation of the electricity sector has led to the development of a sophisticated and competitive market structure [1] . In the market pool [known as a Power eXchange (PX)], power producers [generating companies (GENCOs)] submit generation bids and their corresponding bidding prices, and consumers [consumption companies (CONCOs)] do the same with consumption bids [2]. The market operators use a market-clearing tool to clear the market. This tool is normally based on single-round auctions, and considers each hour / half-hour of the market horizon one at a time to give an hourly / half-hourly Market Clearing Price (MCP) or "spot price" [1]. Alternatively, the companies that wish to hedge against the risk of daily price volatility can do so through physical bilateral over-the-counter (OTC) contracts [3].

The day-ahead forecast of spot prices is the main requirement for market participants [4] and is therefore the focus of this research paper. Energy service companies (ESCOs) buy energy from the PX and from bilateral contracts to sell it to their clients. Having reliable daily price forecast information enables producers or ESCOs to delineate proficient contacts and make better financial decisions to maximise profits [5]. Consumers have to make similar decisions on buying energy through bilateral contracts, or from the PX. Additionally, if a consumer has self-production capability, it can use it to protect itself against high prices in the pool [2][6]–[8]. For these types of portfolio decisions, it is desirable to have forecasts of the next-days' hourly or half-hourly average price values available. Furthermore, the accuracy of day-ahead electrical unit price estimates is of strategic importance for the cost-effective return on power generator capital investment with a consequential beneficial economic return on electrical network operation. Driven by the importance of future prices and the complexities involved in determining them, detailed modelling and forecasting of electrical unit pricing has become a major research field in electrical engineering [1].

However, unlike load forecasting, electricity spot prices tend to exhibit extreme volatility, popularly identified as spikes. Eydeland et al. [9] have suggested that the main reason behind the spikes in the electricity prices are high network transmission grid, unanticipated high load and bidding behaviour of market participants. Additionally, non-storability, seasonal behaviour and transportability are the major issues which make electricity prices so specific [10]. These issues make it impossible to treat electricity on a par with any other commodity and forbids the application of forecasting methods common in other commodity markets [5]. Consequently, electricity price forecasting is much more complex because of the unique characteristics and uncertainties in operation as well as bidding strategies [11]. Although electricity prices can appear unpredictable in their behaviour, Singh et al. [12] state that they are non-random in nature making it possible to identify underlying patterns and periodicities based on the historical data and forecasts.

In this work, the focus is on the Irish Single Electricity Market (SEM) in Ireland. The paper presents a hybrid k-SVM-SVR model, comprising the k-means, the Support Vector Machine (SVM) and the Support Vector Regression (SVR) algorithms. Essentially, implementing the SVM algorithm as both a classifier and predictor tool to accurately predict day-

[1] Corresponding author. Tel.: +353-21-4335166
E-mail address: conor.lynch@cit.ie

ahead prices for the SEM. With the help of data from the Single Electricity Market Operator (SEMO), Eirgrid Group, an analysis of the daily €/MWh prices for day-ahead price forecasting was performed.

The paper is organised as follows. In section 2, an overview of the Irish SEM is presented. Section 3 articulates the stages within the proposed k-SVM-SVR model. The forecast results from the k-SVM-SVR model and a comparative analysis are presented in Section 4. Concluding remarks are placed in Section 5.

II. Overview of the Irish Single Electricity Market

The SEM in Ireland is a gross mandatory auction-based market or pool. Under the pool arrangements, the sale and purchase of electricity occurs on a gross basis with all generators receiving and all suppliers paying the same price for electricity sold into and bought via the pool in a given Trading Period (TD). Figure 1, along with the subsequent paragraphs, provide a brief overview of how the existing market runs, from initial Commercial Offer Data (COD) and Technical Offer Data (TOD) submissions by participants commencing up to 29 days before Trading Day (TD-29) until publishing of the Ex-Post Initial Schedule four days after Trading Day (TD+4).

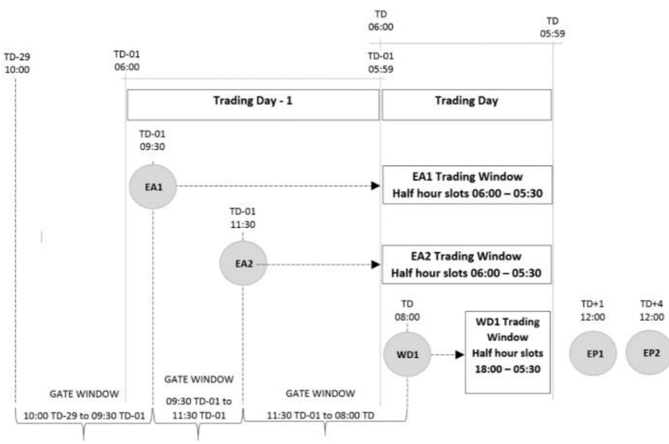

Figure 1: Existing SEM structure in Ireland

Similar to Grimes et al. [13], the two primary runs by the SEMO's market scheduling and pricing software, relevant to the developed model are as follows:

1. The Ex-Ante (EA1) run, which is carried out one day prior to the TD being scheduled – this releases a schedule of half-hourly forecasted System Marginal Price (SMP) produced by the SEMO for the coming TD. The SMP composes the Shadow price + the Uplift price. The Shadow Price comprises the cost of the marginal MW required to meet demand in a TD within the context of an unconstrained schedule. A Generator Unit that can increase its generation in order to meet demand is considered to be marginal. The Uplift cost relates to the operating costs associated with Start-Up-Costs and No-Load-Costs that a generator will need to recover.

2. The Ex-Post Initial (EP2) run, carried out four days after the TD which is being scheduled for. The system marginal prices produced in the EP2 run are used for weekly invoicing and the SMP determined in the EP2 run for a given half hour trading period is the price applicable to both generators and suppliers active in such a trading period.

Each of the five auction categories is held once daily. The **Ex-Ante One** (**EA1**) is held at 09:30 on the day prior to the TD of interest, i.e. TD-01. The EA1 auction covers the delivery hours of 06:00 to 06:00 i.e. the full Trading Day. The first '***intraday***' auction is called the **Ex-Ante 2** (**EA2**) auction and is held at 11:30 on TD-01 which cover the delivery hours of 06:00 to 06:00, which also covers the full Trading Day. The second intraday auction is called the ***Within-Day*** (**WD1**) auction and is held at 08:00 within the TD and covers the delivery hours of 18:00 to 06:00, i.e. the second half of the Trading Day. **Ex-Post Indicative** (**EP1**) provides the market with indicative settlement values regarding the final price and schedule. As previously discussed, the **Ex-Post Initial** (**EP2**) provides market participants with final price and schedule [14].

Imprecise estimation of SMPs can lead to inappropriate quantity bidding strategies by the generators, over/under supply of planned generation, non-optimal demand response, rise in cost of meeting demand, and an increase in real time operation challenges. Hence, it is imperative to have a methodology to precisely predict real time (and day-ahead) SMPs [15].

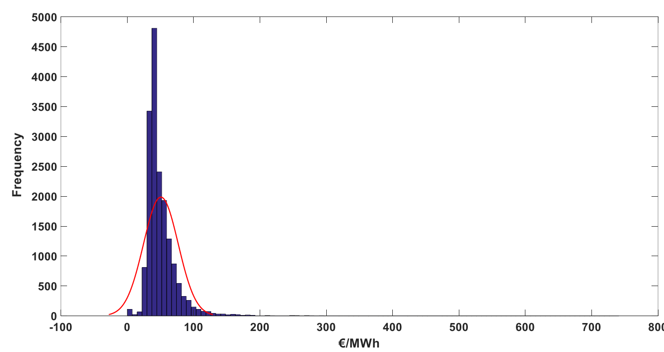

Figure 2(a): Distribution of the 2015 half-hourly SEM EP2 €/MWh data

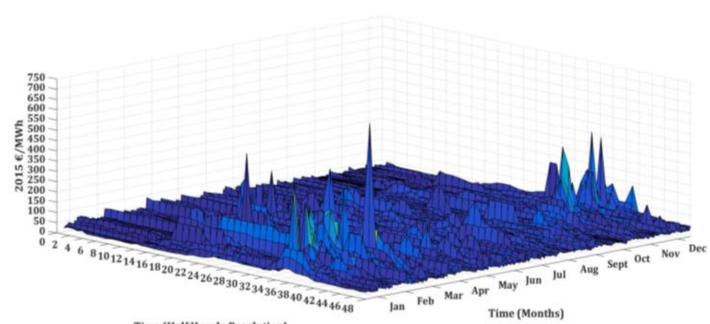

Figure 2(b): 3D plot of the 2015 half-hourly SEM EP2 €/MWh data

This section will take an in-depth look at the price volatility within the SEM and the challenge to accurately forecast for the energy market in Ireland. The study of the SEM EP2 €/MWh price data will focus on the years 2015 and 2016 respectively. Essentially, 2015 data will be used for model training and 2016 data will be used for model prediction validation and analysis.

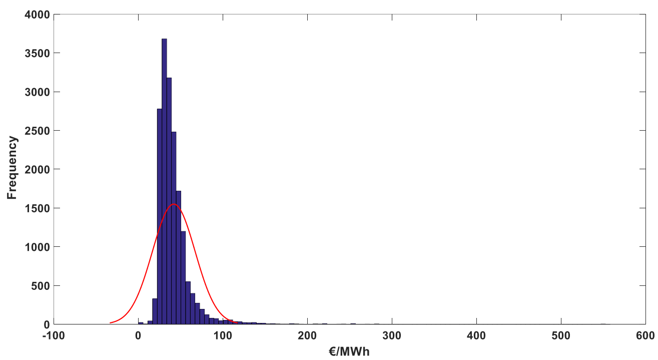

Figure 2(c): Distribution of the 2016 half-hourly SEM EP2 €/MWh data

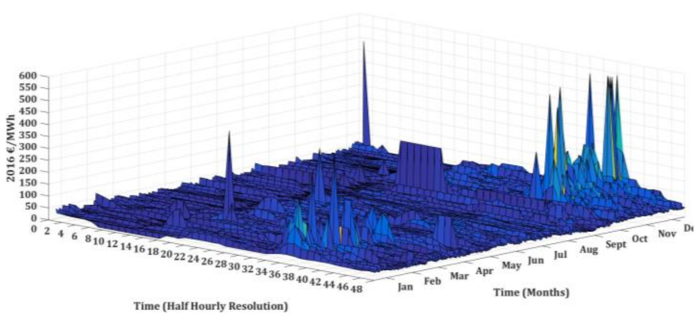

Figure 2(d): 3D plot of the 2016 half-hourly SEM EP2 €/MWh data

Figures 2(a) and 2(c) detail a statistical distribution of the half-hourly SEMO SEM EP2 €/MWh for 2015 and 2016 respectively. While Figures 2(b) and 2(d) present a 3D perspective of same. The aforementioned plots, along with the statistical metrics in Table 1 below, clearly portray the daily periodicities and the volatility that exists within the energy market, thereby indicating the difficulty in deriving accurate day-ahead predictions for the Irish SEM.

Table 1: Statistics of the Irish SMP (€/MWh) data for 2010-11 and 2015-17

Year	Min.	Max.	Mean	Median	St. Dev.
2010	-88.12	766.35	53.85	46.40	35.49
2011	0.00	649.48	63.18	54.45	35.79
2015	0.00	740.26	50.74	43.69	26.04
2016	0.00	557.94	41.82	36.73	25.23
2017	-85.35	715.17	47.41	42.08	25.83

In order to perform rigorous model testing and validation, and to facilitate the comparison of the models performance with results elsewhere in the published literature [13], data for the years 2010-11 and 2015-17 respectively have been used.

III. K-SVM-SVR PRICE FORECASTING MODEL

SMPs can be forecasted using time series information of SMP and other variables such us electricity demand, temperature, wind energy, and fuel prices, etc. Many models in the open literature utilise feature selection and data selection techniques [16]. For large data sets, feature selection techniques (e.g., backward and forward selection, stepwise selection) may become computationally challenging [15].

In this paper, as described in [15], the authors applied a computationally simpler feature selection method that uses signals from the autocorrelation and partial autocorrelation functions. Thus, for the SMP forecast at time t+1, features including previous SMP values, for all significant lags in the ACF (see Figure 3) plot along with the attributes with seasonal information from the PACF (see Figure 4) plot are utilised.

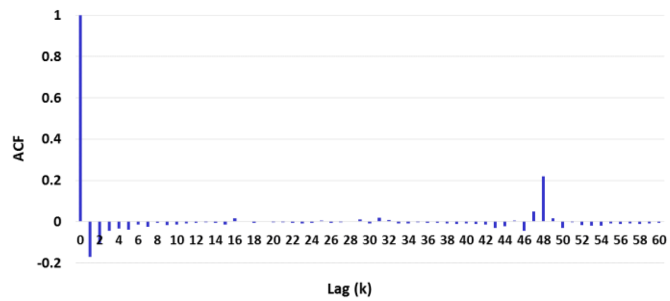

Figure 3: Autocorrelation function (ACF) for 2015 SEMO SMP

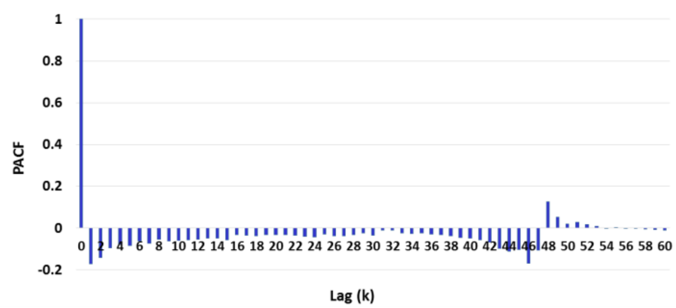

Figure 4: Partial autocorrelation function (PACF) for 2015 SEMO SMP

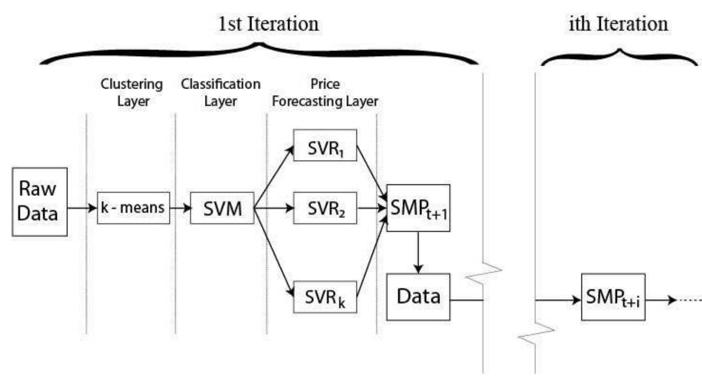

Figure 5: Schematic of k-SVM-SVR model

Before deriving the SVR models, the SMP training data is classified into k clusters of different prices. For each price cluster, a separate SVR model is developed. The forecast for time t + 1 is obtained from one of these k-SVR models. If the current time is t, which the actual SMP is known, the forecast for the next t + *n* (*n* represents the number of future periods to forecast) periods is obtained as follows. The expected SMP for the period t + 1 is assigned to one of the k clusters using a SVM model. Figures 6 and 7 respectively, detail a schematics for the k-means algorithm and the SVM model architecture.

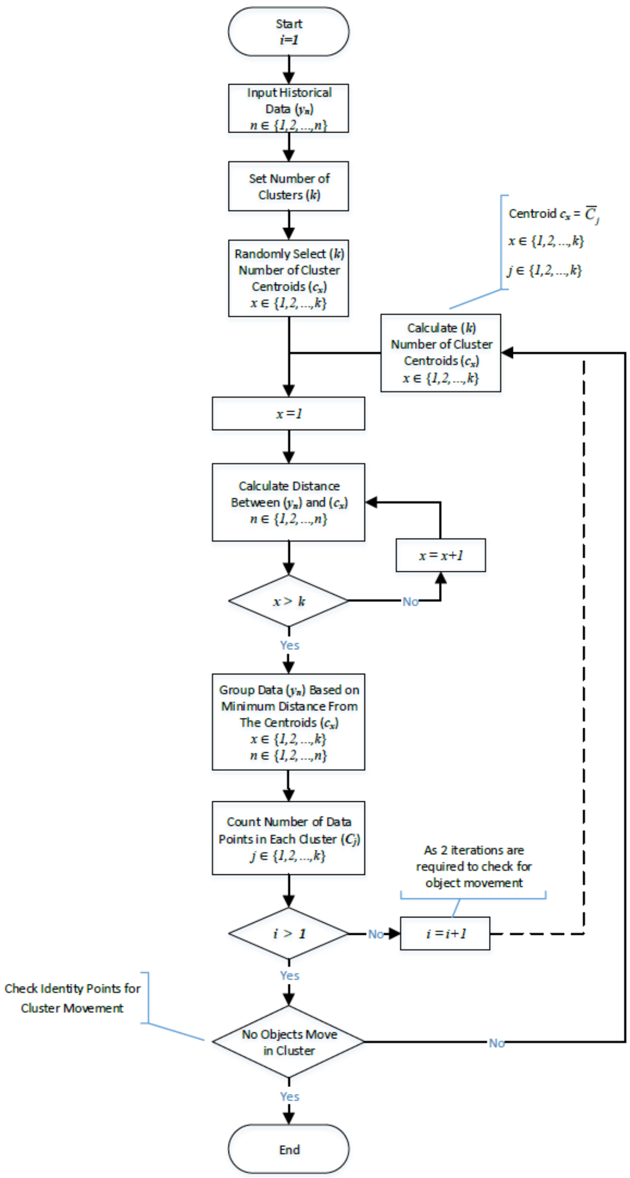

Figure 6: Schematic of k-means algorithm architecture

Then the SVR model for the corresponding cluster is used to obtain the SMP forecast for period t + 1. The forecast for period t + 1 is considered as the actual SMP for forecasting for period t + 2. This process continues until period t + *n* is attained. A depiction of the proposed model is shown in Figure 5.

A step-by-step summary of the training and testing of the model to forecast the price at time t + n is given as follows.

Training:

(1) Perform k-means using the historical data in order to identify the price clusters.
(2) Select the appropriate number of clusters beyond which the decrease in the sum of squares within the clusters is below a chosen tolerance level.
(3) For each price cluster, train a SVR model for prediction.
(4) Train a SVM model for classification using information of all k clusters.

Testing:

Given the trained SVM and SVR models, consider a forecast horizon period of t + hz. Set hz = 1.

(1) Classify the SMP for t + hz as belonging to one of the k Cluster groups.
(2) Obtain the prediction for period t + hz using the SVR model for the cluster identified or selected in step 1 above.
(3) Use the forecast value for t + hz to predict the SMP for t + hz + 1.
(4) Set hz < -hz + 1. Repeat steps 1 through 3 till hz = *n*.

For training purposes, the information of the past d days of the forecast day, and a window of D days of the same forecast day in the previous year (±d days) are used as training data for the SVM and SVR models.

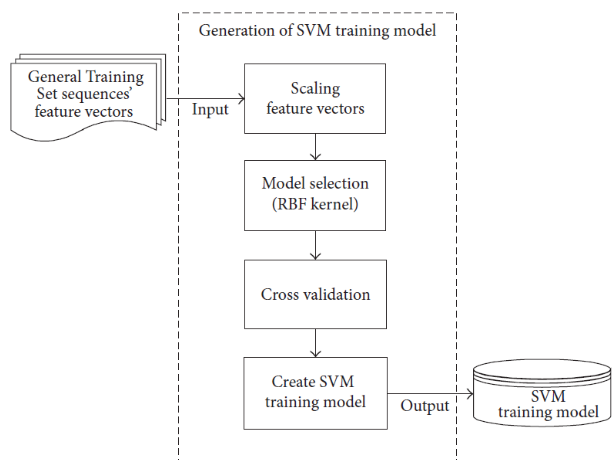

Figure 7: Generic schematic of a SVM model architecture [17]

IV. COMPARATIVE MODEL ANALYSIS

In order to benchmark the developed forecast model with Irish SEM data, the 2011 SEMO SEM data set adopted by Grimes et al [13] was utilised. In [13] the authors developed two SVR-based algorithms to forecast with a day-ahead time horizon. They used the SEMO SEM EA1 price as an indicative estimate of the final value of the SEM EP2 price. Their two constructed models were then compared to the raw SEMO SEM EA price estimates for February, April and June 2011. These results have been summarised in the table below:

Table 2: Mean Square Error (MSE) for predictions from the developed k-SVM-SVR model versus SEMO, FM1 and FM2 model forecasts for 2011

Method	MSE	% Improvement
SEMO EA1	1086.25*	
Forecast Model (FM) #1 [13]	821.01	24.41%
Forecast Model (FM) #2 [13]	781.72	28.03%
k-SVM-SVR Model	712.89	34.37%

*Based on SEMO EA1 v EP2 prices for February, April and June 2011 [18]

$$\text{MSE} = \textit{Mean Square Error} = \frac{1}{n}\sum_{i=1}^{N}(\hat{Y}_i - Y_i)^2$$

$\hat{Y}_i$ and Y_i denote the real and forecast SMP values, N the number of prediction periods and n the number of days.

Considering the mean square error (MSE), Table 2 demonstrates that the k-SVM-SVR model outperforms the two proposed SVR-based models, FM1 and FM2 respectively, as described by Grimmes et al. [13] developed by University College Dublin (UCD). The two UCD models achieved a 24.41% and a 28.03% improvement over the simple utilisation of the EA1 SEM price as a prediction estimate for the final EP2 price. The developed model, described in Section III, based on the same period of February, April and June 2011 accomplished a 34.37% improvement over the existing EA1 price estimates. Figures 8(a) and 8(b) show the predictive performance of the SEMO EA1 estimates and the k-SVM-SVR model against the SEMO EP2 €/MWh data for a random day in February and April 2011.

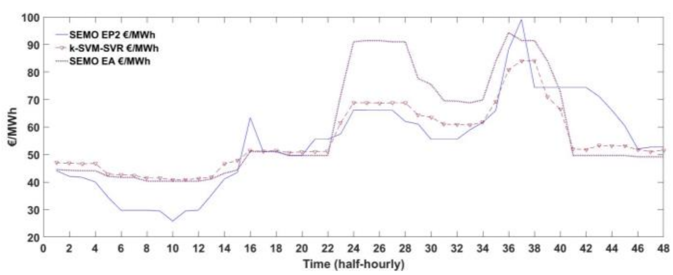

Figure 8(a): EP2 v EA1 v k-SVM-SVR €/MWh data for 02.02.2011

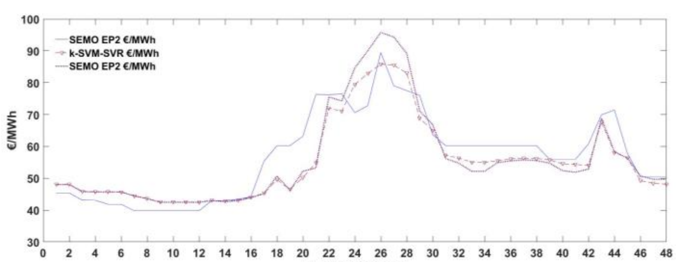

Figure 8(b): EP2 v EA1 v k-SVM-SVR €/MWh data for 24.04.2011

A further benchmark of the developed model, using a more recent SEMO SEM data set was also used. The k-SVM-SVR model was configured in order to predict the day-ahead EP2 price data for 2016 and 2017. Again, as per [13], the raw SEMO SEM EA data served as a benchmark to forecast the EP2 prices. The results are outlined in the Table 3 below.

Table 3: Mean Square Error (MSE) for predictions from the developed k-SVM-SVR model versus SEMO model forecasts for 2016 and 2017

Year	Method	MSE	% Improvement
2016	**SEMO EA1**	457.61*	
	k-SVM-SVR Model	372.57	18.60%
2017	**SEMO EA1**	565.99*	
	k-SVM-SVR Model	488.03	13.77%

*Based on SEMO EA1 v EP2 prices for the full year [18]

$$\text{MSE} = \textit{Mean Square Error} = \frac{1}{n}\sum_{i=1}^{N}(\hat{Y}_i - Y_i)^2$$

Here, the k-SVM-SVR used 2015/16 data for model training. Subsequently day-ahead predictions from the model for 2016/17 showed ≈14-19% improvement over utilising the raw SEMO EA1 €/MWh SEM data as an EP2 price estimate.

The mean absolute percentage error (MAPE) is among the most common measures used to evaluate forecast performance. MAPE is defined as the average of the absolute value of the error forecast, real minus predicted, over the real value. Authors refer to this as MAPE1. Some authors have redefined the concept of the MAPE due to the highly volatile nature of electricity prices [19]. Authors will refer to this as MAPE2. In expressions for MAPE1 and MAPE2, $\hat{Y}_i$ and Y_i denote the real and forecast SMP values respectively and N denotes the number of periods predicted. Other performance measures that are also presented in the literature are the forecast mean square error (FMSE) or root mean square error (RMSE). In order to allow ease of comparison with any existing and future work, these statistical metrics for 2016 and 2017 data in relation to the developed k-SVM-SVR model are shown in Table 4 [15][20].

Table 4: MAE, RMSE, MAPE1 and MAPE2 for predictions from the developed k-SVM-SVR model versus SEMO model forecasts for 2016 and 2017

Year	Model	MAE	RMSE	MAPE1	MAPE2
2016	**SEMO EA1***	8.67	15.12	18.01	20.14
	k-SVM-SVR	**8.28**	**14.55**	**17.21**	**19.24**
2017	**SEMO EA1***	9.92	17.90	19.50	20.28
	k-SVM-SVR	**9.76**	**17.36**	**19.34**	**19.89**

* Based on EA1 v EP2 prices for the full year [18]

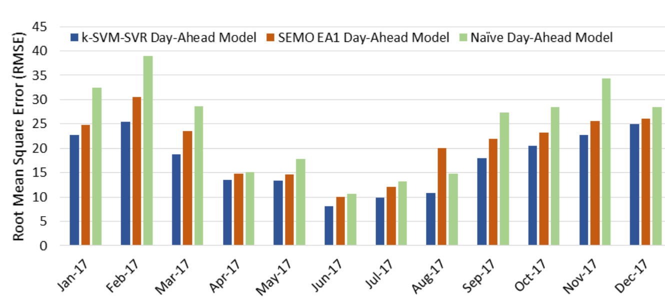

Figure 9: RMSE of k-SVM-SVR, SEMO EA1 and Naïve model for 2017

$$\text{MAE} = \text{Mean Absolute Error} = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{Y}_i - Y_i\right| \quad (1)$$

$$RMSE = \text{Root Mean Square Error} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{Y}_i - Y_i\right)^2} \quad (2)$$

$$\text{MAPE1} = \text{Mean Absolute Percentage Error} = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|\hat{Y}_i - Y_i\right|}{Y_i} \quad (3)$$

$$\text{MAPE2} = \text{Mean Absolute Percentage Error} = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|\hat{Y}_i - Y_i\right|}{Y_i^a} \quad (4)$$

$$\text{where} \quad Y_i^a = \frac{1}{N}\sum_{i=1}^{N} Y_i \quad (5)$$

Using available 2017 yearly data, Figure 9 presents the monthly RMSE values for the superior k-SVM-SVR model developed and contrasts it, for benchmarking purposes, with the existing SEMO EA1 day-ahead price estimates along with the energy price estimates from a conventional Naïve day-ahead model.

V. CONCLUSIONS

This paper discusses a k-SVM-SVR model applied to the Single Electricity Market in Ireland to forecast accurate day-ahead €/MWh energy prices. Data from 2010-11, 2015-16 and 2016-17 respectively was used for training, testing and comparison purposes. For the majority of cases, results showed that the k-SVM-SVR model provided comparable forecast errors with reduced error variances over the existing predictions available to market participants. These decreased variances can potentially translate into large financial saving to electricity utility traders. Future work will investigate the application of such a model to other energy markets including the now Integrated-Single Electricity Market (I-SEM) in Ireland.

VII. ACKNOWLEDGEMENTS

The authors would like to thank Enterprise Ireland (EI) for project funding. Eirgrid Group, the SEMO in Ireland for usage of data and the Cork Institute of Technology for its continued support.

VIII. REFERENCES

[1] N. A. Shrivastava and B. K. Panigrahi, "A hybrid wavelet-ELM based short term price forecasting for electricity markets," *Int. J. Electr. Power Energy Syst.*, vol. 55, pp. 41–50, 2014.

[2] F. J. Nogales, J. Contreras, A. J. Conejo, and R. Espínola, "Forecasting next-day electricity prices by time series models," *IEEE Trans. Power Syst.*, vol. 17, no. 2, pp. 342–348, 2002.

[3] J. Contreras, R. Espínola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1014–1020, 2003.

[4] M. Negnevitsky, S. Member, P. Mandal, and A. K. Srivastava, "Machine Learning Applications for Load , Price and Wind Power Prediction in Power Systems," 2009.

[5] P. Mandal, T. Senjyu, S. Member, and N. Urasaki, "A Novel Approach to Forecast Electricity Price for PJM Using Neural Network and Similar Days Method," vol. 22, no. 4, pp. 2058–2065, 2007.

[6] Q. A. Phan, T. Scully, M. Breen, and M. D. Murphy, "Determination of optimal battery utilization to minimize operating costs for a grid-connected building with renewable energy sources," *Energy Convers. Manag.*, vol. 174, no. April, pp. 157–174, 2018.

[7] Q. A. Phan, M. D. Murphy, M. C. Breen, and T. Scully, "Economic optimisation for a building with an integrated micro-grid connected to the national grid," *2015 World Congr. Sustain. Technol. WCST 2015*, pp. 140–144, 2016.

[8] Q. A. Phan, M. D. Murphy, M. C. Breen, and T. Scully, "One-day-ahead cost optimisation for a multi-energy source building using a genetic algorithm," *2016 UKACC Int. Conf. Control. UKACC Control 2016*, 2016.

[9] W. K. Eydeland A, *Energy and Power Risk Management: New Developments in Modeling, Pricing, and Hedging*. Wiley Publivations, 2012.

[10] L. Z. Shahideehpour M., Yamin H., *Market operations in electric power systems: forecasting, scheduling, and risk assessment*. John Wiley & Sons, Inc., New York, 2002.

[11] D. W. Bunn, "Forecasting loads and prices in competitive power markets," *Proc. IEEE*, vol. 88, no. 2, pp. 163–169, 2000.

[12] N. Singh and S. R. Mohanty, "A Review of Price Forecasting Problem and Techniques in Deregulated Electricity Markets," *J. Power Energy Eng.*, vol. 3, no. September, pp. 1–19, 2015.

[13] D. Grimes, G. Ifrim, B. O'Sullivan, and H. Simonis, "Analyzing the impact of electricity price forecasting on energy cost-aware scheduling," *Sustain. Comput. Informatics Syst.*, vol. 4, no. 4, pp. 276–291, 2014.

[14] SEMO, "Intra-Day Trading Project Market Participant Technical Overview 3," 2011.

[15] F. Feijoo, W. Silva, and T. K. Das, "A computationally efficient electricity price forecasting model for real time energy markets," *Energy Convers. Manag.*, vol. 113, no. April, pp. 27–35, 2016.

[16] G. J. Osório, J. C. O. Matias, and J. P. S. Catalão, "Electricity prices forecasting by a hybrid evolutionary-adaptive methodology," *Energy Convers. Manag.*, vol. 80, pp. 363–373, 2014.

[17] X. Y. Ng, B. A. Rosdi, and S. Shahrudin, "Prediction of antimicrobial peptides based on sequence alignment and support vector machine-pairwise algorithm utilizing LZ-complexity," *Biomed Res. Int.*, vol. 2015, no. April, 2015.

[18] Eirgrid, "SEMO Single Electricity Market Operator." [Online]. Available: http://www.sem-o.com/Pages/default.aspx.

[19] P. Mandal, T. Senjyu, N. Urasaki, T. Funabashi, and A. K. Srivastava, "A novel approach to forecast electricity price for PJM using neural network and similar days method," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 2058–2065, 2007.

[20] S. K. Aggarwal, L. M. Saini, and A. Kumar, "Electricity price forecasting in deregulated markets: A review and evaluation," *Int. J. Electr. Power Energy Syst.*, vol. 31, no. 1, pp. 13–22, 2009.