Comparative Study of Learning Machine Predictors for Half-Hour and Day-Ahead Electricity Price Forecast in Deregulated Markets

Ali Lahouar and Jaleleddine Ben Hadj Slama
Advanced Systems in Electrical Engineering
National Engineering School of Sousse, BP 264, Sousse Erriadh 4023, University of Sousse, Tunisia
National Engineering School of Tunis, BP 37, Le Belvedere 1002 Tunis, Tunisia
Email: ali.lahouar@gmail.com, bhslama@yahoo.fr

***Abstract*—The electricity markets all over the world have received several changes since the arrival of smart meters, real time pricing and deregulation waves. In such conditions, electricity price is a key index that reflects competition between generators, percentage of renewable generation and load behavior. An accurate forecast of the future price allows both producers and customers to plan their strategies, increase their profits and avoid contingencies. For this purpose, this paper proposes two predictors that perform half-hour-ahead and day-ahead price forecast, respectively. Both predictors are based on three learning machines, namely artificial neural network, support vector regression and random forest, in addition to persistence model which is used as reference. The built models are tested through real historical time series from New South Wales in Australia, which include half-hourly price and load. The profiles are analyzed with the aim of defining correlations and choosing adequate inputs. Then, the performance of each machine is assessed using the mean absolute percentage error, in order to determine the most suitable prediction machine according to the case. Furthermore, the best forecast situations are plotted to evaluate the ability of predicting sudden and sharp price peaks.**

***Keywords*—Short term price forecasting, artificial intelligence, day-ahead market, deregulation.**

I. Introduction

With deregulation of energy markets, the electricity price becomes a very important issue. The main goal of deregulation is to maximize effective generation and reduce costs at the same time. With the introduction of smart meters and real time pricing, the notion of demand response (DR) appeared. DR means changing the consuming patterns of end-user customers willingly to match peak-off and low-cost periods. The effective demand response depends on the ability of forecasting load and renewable generation from supplier side, and forecasting prices from market and clients side. In general, this forecast must be done for short times intervals in advance, up to an hour or a half-hour or even 5 minutes in some markets. These intervals are shortened because the price profile is highly volatile and non-smooth, as a consequence of interactions between price, load and penetration degree of renewable generation. The price profile has a strongly seasonal nature, and is correlated to the load to some extent [1]. Generally speaking, forecasting load and price in wholesale markets are intertwined tasks. Both producers and customers rely on price predictions to prepare their strategies. Load depends essentially on meteorological factors, while price is driven specially by load, even if some other market factors have also their impact on electricity price behavior; such as economic growth and fuel prices. However, load curves have similar patterns, whereas price curves are very volatile and spike-prone. The price in particular, may jump to tens or hundreds of times its normal value for very short periods, as it may also drops to zero or even negatives values. In such conditions, forecasting prices along with load is crucial, since it allows producers as well as customers to optimize their affairs [2].

The role of an organized market is to match electricity offer and demand in a way to determine the market clearing price (MCP). The MCP is generally determined once a day by an auction sale, as the intersection between the curve of offer and demand. The negative prices are authorized in many markets, generally when the demand is very low (losses due to shutting down a production unit may be greater than losses caused by accepting negative prices) or when renewable generation is very high. In day-ahead markets, offers are submitted for each hour (or half-hour) of the next day. Some markets follow a real time settlement structure. In such system, prices are fixed by the system operator every 5 minutes, and spot prices are determined every half-hour. This structure makes the market very volatile, and this the case of the Australian market, subject of this study. The electricity price depends on several factors such as system load, weather conditions, fuel prices, spinning reserve (available production minus expected load), planned maintenance and forced cut-off of important grid components. The historical time series of all these factors as well as their expected future values for the considered forecast horizon are very important for the construction and calibration of price predictors [3].

The different approaches proposed by the literature are:

- Multi-agents models: they are models that simulate the operation of heterogeneous agent system (generating units, companies) interacting with each other, in order to build the price process by matching offer and demand in the market. They are production-cost based and well adapted to regulated markets with little incertitude and stable structure.

- Fundamental models (structural): they describe the price

978-1-4673-9768-1/16/$31.00 © 2016 IEEE

dynamic by modeling the impact of important physical and economic factors on it.

- Reduced-form models (quantitative, stochastic): they characterize the statistic properties of price over time, with the ultimate goal of risk management.

- Statistic models (econometric): they are direct applications of statistic techniques of load or energy forecast, thus implementations of econometric models.

- Computing intelligence (artificial learning, nonlinear statistics): they are techniques combining elements of learning, evolution and fuzzy logic to create approaches adapted to complex dynamic systems [3].

The two last families of models (statistics and artificial intelligence) are generally established by electrical engineering researches; since they are considered signal processing techniques. Contrariwise, the three first families are especially utilized in economics reviews. To establish a fair comparative study between different methods, it is essential to use the same data sets and the same evaluation criteria.

High frequency, non-stationarity, multiple seasonality, calendar effect, high volatility, spikiness and nonlinearity make the accurate forecast a real challenge. Several techniques are developed to forecast electricity prices, like artificial neural networks (ANN) and their derivatives, support vector regression (SVR), hidden Markov models, autoregressive integrated moving average (ARIMA), generalized autoregressive conditional heteroskedasticity (GARCH), linear regression, Kalman filter and regression tree [4], [5]. Recently, the empirical mode decomposition has been used as a promising alternative [6], [7]. In this paper, three of most effective learning machines are chosen to construct the predictors, namely ANN, SVR and the random forest (RF), in order to compare their abilities under the same conditions and facing the same inputs, with no optimizations. The signal to be predicted is the electricity price from New South Wales in Australia, for a forecast horizon of half-hour and 24 hours, respectively. The remainder of the paper is organized as follows: section II defines some mathematical notations, section III presents the Australian energy market, section IV analyzes the price and load profiles, section V details the forecast strategy and interprets the results, and section VI concludes the study.

II. Mathematical development

The forecast problem is presented as follows. Let F be the prediction model where $X = (X^1, ..., X^m)$ are its m inputs and $\hat{Y}$ is its output.

$$\hat{Y} = F(X),\ X \in \mathbb{R}^m,\ \hat{Y} \in \mathbb{R} \tag{1}$$

We denote Y the real (measured) value of the predicted signal, $Y \in \mathbb{R}$. X, Y and $\hat{Y}$ are in terms of time t, which is discrete and hourly stepped. Predicting half-hour in advance for example means determining $\hat{Y}(t+1/2)$. Let $S_n = \{(X_1, X_1), ..., (X_n, X_n)\}$ be the training set used to create the predictor F from n observations (X_i, Y_i). Of course, X_i and Y_i are past samples of X and Y. Let s be the length of forecast horizon in hours. For day-ahead prediction by half-hour steps, the problem is to determine $\hat{Y}(t+1/2), \hat{Y}(t+1), ..., \hat{Y}(t+s)$ where $s = 24$.

Each predicting machine including persistence should give $\hat{Y}$, most cases in terms of X. For half-hour persistence for example, $\hat{Y}(t+1/2) = Y(t)$. The development of the other machines is not presented for the purpose of brevity. See reference [8] for ANN development, [9], [10] for SVR and [11], [12] for RF. The proposed error evaluation criterion is mean absolute percentage error (MAPE). Let $E = \hat{Y} - Y$ be the error signal, where $E(t)$ is the error at time t. We assume getting k errors (samples of $E(t)$) after performing β predictions (the relationship between k and β will be presented later in section V). The MAPE is then defined by:

$$MAPE = \frac{1}{k}\sum_{i=1}^{k}\frac{|E(t)|}{Y(t)} \times 100 \tag{2}$$

The MAPE criterion is one of the most used in literature. It gives good quantitative interpretation of the prediction as long as $Y(t)$ is far away from zero.

III. Presentation of Australian energy market

The Australian Energy Market Operator (AEMO) has been created to manage electricity and gas markets since 1st July 2009. Since the beginning and the middle of the Nineties, the deregulation and privatization of state assets have conducted to markets more open and transparent, thereby facilitating trade and concurrence.

Since 1998, the production, distribution and supply of electricity in eastern and southern Australia have been led by the National Electricity Market (NEM). The NEM connects five regional markets (Queensland, New South Wales, Victoria, Southern Australia and Tasmania). It ensures also wholesale electricity production which is carried through high voltage transmission lines to distributors, and then to houses and companies.

The transport of electricity from all producers to all consumers is facilitated through a pool, in which the production of all generators is aggregated and programmed at intervals of five minutes in order to respond to demand. The pool is not a physical thing, but a set of procedures that the AEMO manages in conformity with the law and national electricity regulation. The market makes use of sophisticated systems to send signals to producers informing them how match energy they must produce every five minutes. This information is mandatory to adapt production to consumers requirements, maintain the reserves ready to contingencies and calculate the actual price. The NEM contains both state and private assets managed by several participants.

IV. Profile analysis

The restructuring and deregulation of Australian electricity markets entailed major changes in wholesale price behavior. These prices are always characterized by high volatility (high variance), a strong return to the mean (prices have tendency to fluctuate around a balance point in the long term), and sudden and expected spikes or peaks which disintegrate rapidly. In fact, they reflect the inherent features of competitive electricity markets; such as seasonality, low marginal cost of production, impact of failure or system breakdown, interconnection constraints between markets, limited storage and so on. As a consequence, the major participants at these markets, including generators, retailers and big industrial consumers, are exposed to significant market risks and are obliged to take costly measures of risk management [13].

The New South Wales is one of the five markets managed by the AEMO, and is the subject of this study. The load and price profile of this market during three days of March 2013 are given in Fig. 1.

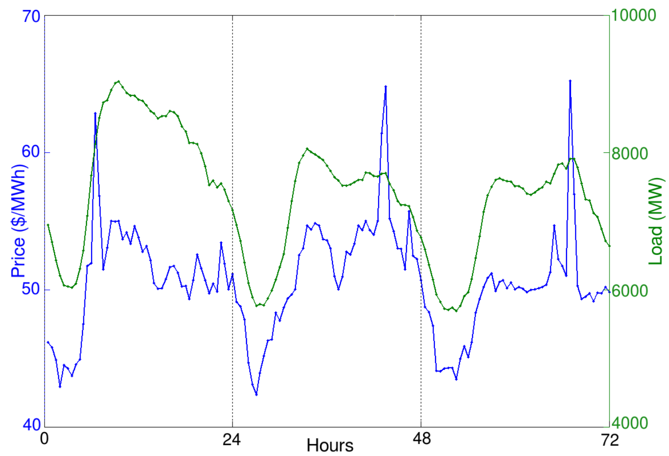

Fig. 1. Load and price profile, 1-3 March 2013, New South Wales

The load profile is more or less regular, it is possible to distinguish vaguely through the curve the two consumption peaks of morning and evening, and the drop after midnight. Variations are smooth and progressive. However, the price profile is much more chaotic; variations are sharp and almost random, and yet correlation with load remains visible with the noticeable price drop at the first hours of the morning. Generally speaking, it is not possible to explain all the price variations without taking into account pricing procedures and regulation, hence the price predictor may be much more accurate when it is refined according the market characteristics.

The Fig. 2 shows the same curves in three days of July 2013. Likewise, the load curve contains two peaks at morning and evening, clearly more visible. However, price is further more tough to predict; in a way that it keeps nearly constant value for long periods, but peaks are spectacular and very sharp when they occur. These peaks are generally accompanied by steep slope or tendency change in the load curve. A possible reason is that prices are intentionally increased during transitions in order to discourage consumers and give the generators enough time to start or stop. The season effect should not be neglected too, since July is winter in Australia, and the consumers need for warm may be the main reason of this complex profile.

V. PRICE FORECAST

Since much specific information are not included in this study, such as Australian production plants, pricing procedure, dominant weather, population distribution and customers

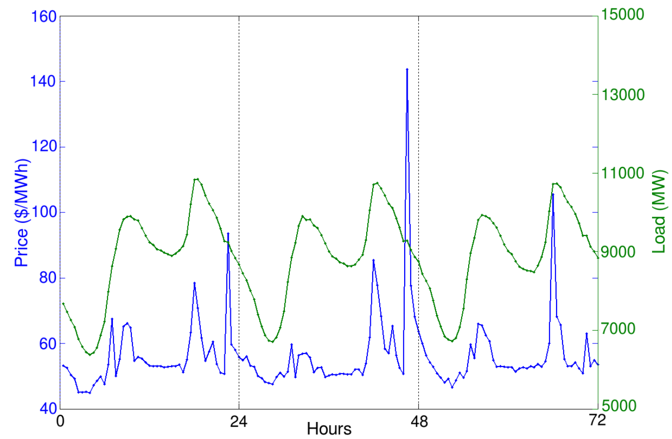

Fig. 2. Load and price profile, 1-3 July 2013, New South Wales

behavior, prediction will be blindly performed; aiming only to assess the performance of learning machines, regardless of all these exogenous factors that may affect the forecast accuracy. The main goal is hence to compare effectiveness of the machines operating under the same conditions and handling the same data sets.

A. *Half-hour-ahead price forecast*

In this section, prediction will be performed for a horizon of half-hour in advance, thus very short term ($s = 1/2$). The time step is also a half-hour. In order to construct the predictor, inputs are classified; endogenous and exogenous. Endogenous inputs are internal, which means historical price time series. In other words, past price values that may affect future ones. Contrariwise, exogenous inputs are external, meaning any signal that may affect the price evolution, except the price signal itself. The chosen inputs are as follows: three past values of price taken every half-hour (endogenous), and four inputs of load; three past values and one future value also taken every half-hour (exogenous). No other endogenous factors are considered like temperature or fuel price despite their potential influence, in order to simplify the model design. In addition, the future value of load (of half-hour-ahead, which is used as input) is assumed to be exact and performed through another predictor, to avoid adding uncertainties. The price of half-hour-ahead will then be in terms of this future load value, 3 past load values, and 3 past price values, resulting in a total number of inputs equal to 7 ($m = 7$). The Fig. 3 shows the designed predictor with its different inputs. The prediction model is written through these equations:

$$X_L = \left[L\left(t+\frac{1}{2}\right)\ L(t)\ L\left(t-\frac{1}{2}\right)\ L(t-1)\right] \quad (3)$$

$$X_P = \left[P(t)\ P\left(t-\frac{1}{2}\right)\ P(t-1)\right] \quad (4)$$

$$X\left(t+\frac{1}{2}\right) = [X_L\ X_P]^T \quad (5)$$

$$Y(t) = P(t) = X^5(t) \quad (6)$$

$$\hat{Y}\left(t+\frac{1}{2}\right) = F\left(X\left(t+\frac{1}{2}\right)\right) \quad (7)$$

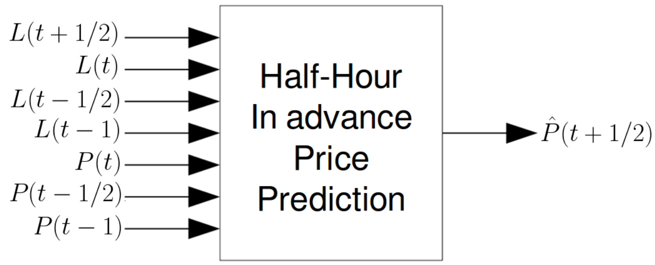

Fig. 3. The half-hour-ahead price predictor model

$L(t)$ stands for load and $P(t)$ stands for price at hour t. Then $P(t+1/2)$ is the future price (of half-hour-ahead), $P(t)$ is the actual price, $P(t-1/2)$ is a past price (of half-hour ago) and $P(t-1)$ is a past price (of hour ago). Same thing for the load L.

It is interesting to view the response of the persistence model before proceeding to other predictors. This model, in which the predicted value is simply equal to the actual one, is a reference for short-term prediction methods. The designed predictor must perform a better forecast than the persistence, otherwise is is useless. The calculated MAPE of persistent forecast is given in Table I, and obviously, it is rather low, about 2.6%. The challenge for other models is then to perform further more accurate forecast.

	Persistence	ANN	SVR	RF
Half-hour-ahead March	2.63%	2.25%	2.21%	2.38%
Half-hour-ahead July	7.61%	5.39%	-	5.45%
Day-ahead March	5.32%	11.23%	3.80%	4.11%
Day-ahead July	7.51%	10.24%	5.68%	5.72%

TABLE I. THE MAPE RESULTS FOR ALL PREDICTION CASES

January and February 2013 are chosen for training for all predicting machines. Offline training is enough, since the test period is short; only the first three days of March 2013. The number of observations in the training test is $n = 2829$, the number of predictions $\beta = 144$ and the overall number of errors $k = \beta$. No optimization process is proposed to ensure fair comparison; hence the machines settings are empirical and manual. The proposed configurations are as follows:

- The random forest has $ntree = 500$ and $mtry = 2$.

- The support vector regression owns a polynomial kernel of degree 1, $C = 10$ and $\gamma = 3.10^{-4}$.

- The artificial neural network has feedforward architecture, with one hidden layer of 20 neurons, unipolar sigmoid activation function and Levenberg-Marquardt training algorithm.

After testing and checking the results of Table I, we notice that the error reduction is minimal; it does not even fall under the barrier of 2%.

In order to get more accurate results, the training test is widened; it contains henceforth 6 months, from January to July 2013. The testing period covers the three first days of July. As a consequence, $n = 8685$, $\beta = 144$ and $k = \beta$. The profile is a bit different, as explained in the previous section. The persistence model entailed rather high error, due the spectacular price spikes. The other machines keep their configurations. The RF reduces the MAPE of more than 2%, as well as the ANN. However, the SVR does not converge anymore with its actual configuration, since the number of observations n became very high. The predicted curve has a more intelligent behavior with RF and ANN, and does not follow by rote the measured curve anymore, like it was the case with persistent forecast. The ANN succeeded to predict effectively the peaks many times, and gave the best results in predicting half-hour-ahead prices, as shown in Fig. 4.

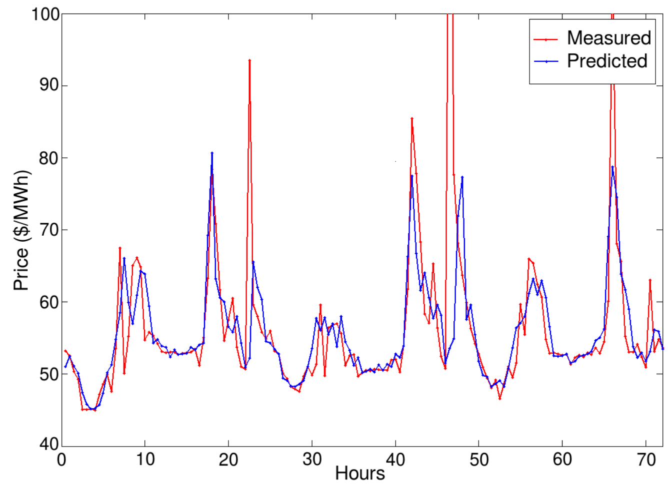

Fig. 4. Half-hour-ahead price prediction using ANN, 1 to 3 July 2013, New South Wales

B. Day-ahead price forecast

In order to predict 24 hours in advance, it is enough to build $2s$ instances of the previous model, dedicated to predict half-hour-ahead, where $s = 24$. Then 48 models will be constructed and called for this purpose. Fig. 5 describes the architecture of each model, and the inputs are the same ($m = 7$). Obviously, such a procedure needs a prior knowledge of 48 load values in advance (day-ahead), which is assumed to be exact and gathered through another predictor. The equations are the following:

$$X_{L1} = \left[L\left(t+\frac{i}{2}\right) \; L\left(t+\frac{i-48}{2}\right)\right] \quad (8)$$

$$X_{L2} = \left[L\left(t+\frac{i-96}{2}\right) \; L\left(t+\frac{i-144}{2}\right)\right] \quad (9)$$

$$X_{P1} = \left[P\left(t+\frac{i-48}{2}\right) \; P\left(t+\frac{i-96}{2}\right)\right] \quad (10)$$

$$X_{P2} = \left[P\left(t + \frac{i-144}{2} \right) \right] \quad (11)$$

$$X\left(t+\frac{i}{2}\right) = [X_{L1}\ X_{L2}\ X_{P1}\ X_{P2}]^T \quad (12)$$

$$Y(t) = P(t) = X^5(t) \quad for\ i = 2s \quad (13)$$

$$\hat{Y}\left(t+\frac{i}{2}\right) = F\left(X\left(t+\frac{i}{2}\right)\right) \quad (14)$$

Where $i = 1, 2, 3, ..., 2s$.

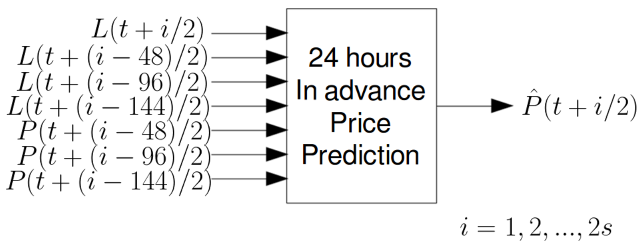

Fig. 5. The day-ahead price predictor model

This strategy was applied using the same methods of previous section, namely persistence, RF, ANN and SVR. Learning and test sets are also the same; January and February for training, and three days of March for testing. Obviously, each model among the 48 has its own training set S_n, with $n = 56$ and $\beta = 3$. The overall number of errors is $k = 2s\beta = 144$. The persistence in this case is a shift of all the previous day, and not just a shift of the previous half-hour.

The first conclusion that can be made after seeing the results of Table I is that ANN is not suitable at all for this kind of forecast; it is even worse than persistence. This may be due to stochastic training procedure of ANN, which makes the 48 built models very different. However, RF and ANN improve noticeably the persistence model, even if the MAPE remains high. As regards the July results with widened training set, there is no big difference. For each model among the 48, $n = 178$. The 3 days testing period give $\beta = 3$ and $k = 2.s.\beta = 144$. Obviously, the enlargement of the training period gave more accurate results (compared to persistence). The SVR has slightly lower MAPE than RF, but RF is more able to predict sudden peaks, as shown in Fig. 6. The forecast error in literature varies enormously according to market specifications. But in general, it is between 3% and 14% in terms of MAPE [5], [14]. The results found in this paper are sufficiently acceptable, taking into account the high volatility of the Australian market.

VI. Conclusion

In this paper, electricity price forecast was proposed in a deregulated market, where a real time pricing procedure is applied, for two different time scales. Since the research tendency moves toward artificial intelligence methods, three of the most effective learning machine are proposed, namely artificial neural networks, support vector regression and random forest. After presenting the market specification, the price profile is analyzed along with load, in order to detect correlations and

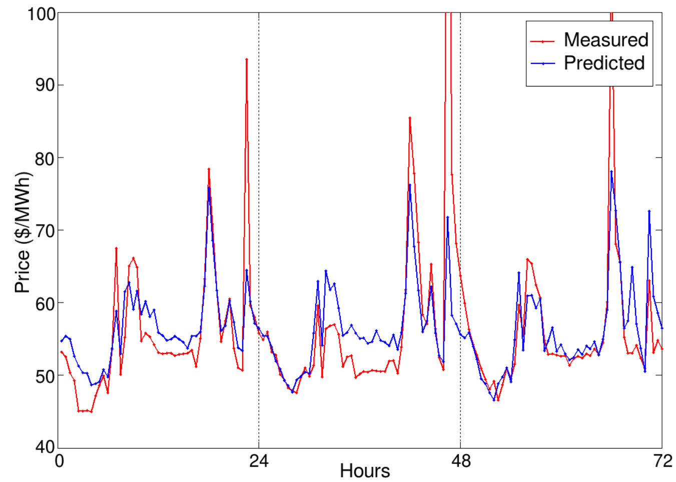

Fig. 6. Day-ahead price prediction using RF, 1 to 3 July 2013, New South Wales

define suitable model inputs. Then, two predictors are designed for half-hour and day-ahead forecast respectively, based on the three proposed machines. All the responses are assessed and compared with persistence forecast, in order to determine the most effective machine for each context. Results show that ANN performs the most accurate results for half-hour-ahead, whereas RF is the most effective for day-ahead forecast, especially for detecting spikes. The SVR may give lower errors than RF in that case, but may also have technical problems of convergence. Obviously, this comparison does not include any optimization process, hence the prospect of this work is to compare optimized methods using the same model and inputs, and conclude about the effectiveness of each optimization process.

References

[1] S. Chan, K. Tsui, H. Wu, Y. Hou, Y.-C. Wu, and F. Wu, “Load/price forecasting and managing demand response for smart grids: Methodologies and challenges,” *Signal Processing Magazine, IEEE*, vol. 29, no. 5, pp. 68–85, Sept 2012.

[2] P. Mandal, T. Senjyu, and T. Funabashi, “Neural networks approach to forecast several hour ahead electricity prices and loads in deregulated market,” *Energy Conversion and Management*, vol. 47, no. 1516, pp. 2128 – 2142, 2006.

[3] R. Weron, “Electricity price forecasting: A review of the state-of-the-art with a look into the future,” *International Journal of Forecasting*, vol. 30, no. 4, pp. 1030 – 1081, 2014.

[4] N. Shrivastava, A. Khosravi, and B. Panigrahi, “Prediction interval estimation of electricity prices using pso-tuned support vector machines,” *Industrial Informatics, IEEE Transactions on*, vol. 11, no. 2, pp. 322–331, April 2015.

[5] P. Pany and S. Ghoshal, “Dynamic electricity price forecasting using local linear wavelet neural network,” *Neural Computing and Applications*, pp. 1–9, 2015.

[6] Z. Shao, F. Gao, S.-L. Yang, and B. gong Yu, “A new semiparametric and {EEMD} based framework for mid-term electricity demand forecasting in china: Hidden characteristic extraction and probability density prediction,” *Renewable and Sustainable Energy Reviews*, vol. 52, pp. 876 – 889, 2015.

[7] K. He, L. Yu, and L. Tang, “Electricity price forecasting with a {BED} (bivariate {EMD} denoising) methodology,” *Energy*, vol. 91, pp. 601 – 609, 2015.

[8] M. Benrejeb, P. Borne, and J. Haggege, *Les reseaux de neurones : Presentation et applications*. Edition TECHNIP, 2007.

[9] H. Frezza-Buet, “Machines a vecteurs supports,” Supelec, Tech. Rep., November 2012.

[10] H. Drucker, C. J. C. Burges, L. Kaufman, A. J. Smola, and V. Vapnik, “Support vector regression machines,” in *NIPS*, 1996, pp. 155–161.

[11] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[12] R. Genuer, “Forets aleatoires : aspects theoriques, selection de variables et applications,” Ph.D. dissertation, Universite Paris Sud - Paris XI, 2010.

[13] H. Higgs and A. Worthington, “Stochastic price modeling of high volatility, mean-reverting, spike-prone commodities: The australian wholesale spot electricity market,” *Energy Economics*, vol. 30, no. 6, pp. 3172 – 3185, 2008, technological Change and the Environment.

[14] H. Shayeghi, A. Ghasemi, M. Moradzadeh, and M. Nooshyar, “Simultaneous day-ahead forecasting of electricity price and load in smart grids,” *Energy Conversion and Management*, vol. 95, no. 0, pp. 371 – 384, 2015.