

# Classification of Future Electricity Market Prices

Hamidreza Zareipour, *Senior Member, IEEE*, Arya Janjani, *Student Member, IEEE*, Henry Leung, *Member, IEEE*, Amir Motamedi, *Student Member, IEEE*, and Antony Schellenberg, *Member, IEEE*

**Abstract**—Forecasting short-term electricity market prices has been the focus of several studies in recent years. Although various approaches have been examined, achieving sufficiently low forecasting errors has not been always possible. However, certain applications, such as demand-side management, do not require exact values for future prices but utilize specific price thresholds as the basis for making short-term scheduling decisions. In this paper, classification of future electricity market prices with respect to pre-specified price thresholds is introduced. Two alternative models based on support vector machines are proposed in a multi-class, multi-step-ahead price classification context. Numerical results are provided for classifying prices in Ontario’s and Alberta’s markets.

**Index Terms**—Classification, demand-side management, forecasting, scheduling, smart grid, support vector machines.

## I. INTRODUCTION

ELECTRICITY price is a key factor in determining short-term operating schedules and bidding strategies in competitive electricity markets [1]. Consequently, numerous data-driven approaches have been proposed for modeling and forecasting short-term electricity market prices [2]–[16]. The reported price forecasting errors generally range from approximately 5% to 36% and vary based on the technique used and the market analyzed. This range of error, however, is relatively high when compared to that of short-term electric load forecasting where errors usually range from 1% to 3% [17].

Various factors contribute to reduced accuracy of electricity price forecasting models; unpredictable forced outages [16], complex and changing price regimes [18], integration of intermittent energy sources [19], and implementation of reliability-based demand response programs [20] all introduce fluctuations and changes in electricity prices that may be extremely difficult to model accurately and consistently.

It is observed from the existing literature that traditional price forecasting models are generally developed for numerical prediction or point-forecasting. That is, existing models try to predict the exact value of prices at future hours by approximating the true underlying price formation process. However, not all market participants need to know the exact value of future prices in their decision-making process. For example, through the in-

roduction of “smart grid” technologies and new marketplace initiatives, it is expected that the demand-side interactions will be enabled to widely participate in electricity markets at the residential, commercial, and industrial levels [21]. Considering the on/off nature of most of electric loads, especially at the residential level, demand-side market participants may primarily react when prices exceed specific thresholds. Beyond these thresholds, the exact price of electricity would be considered unimportant since it is simply “too expensive”. Moreover, many demand-response products are designed having certain thresholds for electricity prices in mind [20], such as the hour-ahead dispatchable load program in the Ontario market [22]. Another example of threshold-based decision making can be found in electricity consumers with on-site generation facilities. These facilities only purchase electricity from the grid if the electricity market price are below the marginal cost of operating the on-site electricity generation equipment [12]. In these types of applications where the exact value of prices is not primarily required, the point-price forecasting problem can be reduced to price classification subproblems in which the class of future prices is of interest.

This paper proposes a short-term price classification method as an alternative to numerical price forecasting. In price classification, predictions are made with respect to whether the price is above or below pre-specified price thresholds defined by users based on their operation and planning objectives. Price classification is specifically useful when the exact value of future prices is not critically important. The main contribution of this paper is to propose a customized approach to predict the behavior of future prices where the specific forecasting needs of the users are taken into consideration.

The remainder of this paper is organized as follows: In Section II, a review of the background pertaining to this work is presented. The proposed models are discussed in Section III followed by the numerical results in Section IV. Finally, the main findings of this paper are summarized in Section V.

## II. BACKGROUND REVIEW

In general, data-driven predictive models are built for either numerical prediction or classification. A numerical prediction model approximates the underlying process under consideration and is used to forecast future values for the variable of interest. Classification refers to the assignment of class labels to unlabeled data. This section presents a background review of literature pertaining to short-term electricity price forecasting, and discusses the steps of predictive model building.

### A. Review of Short-Term Price Forecasting Literature

There are a wide variety of publications regarding numerical or point forecasting of future electricity prices; these works em-

Manuscript received November 11, 2009; revised November 15, 2009, March 19, 2010, and April 23, 2010; accepted May 12, 2010. Date of publication July 08, 2010; date of current version January 21, 2011. This work was supported by NSERC Canada and the University of Calgary’s Research Grant Committee (URGC). Paper no. TPWRS-00876-2009.

The authors are with the Department of Electrical and Computer Engineering, University of Calgary, Calgary, AB T2N 1N4, Canada (e-mail: h.zareipour@ucalgary.ca; janjania@ucalgary.ca; leungh@ucalgary.ca; amotamed@ucalgary.ca; awshell@ucalgary.ca).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TPWRS.2010.2052116

ploy a variety of different models and have had varying accuracy in their predictions. For example, the work presented in [4] describes neural networks-based models for forecasting prices in the Spanish and Pennsylvania, New Jersey, and Maryland (PJM) markets with an overall forecasting error of about 5%; however, the model accuracy was reported to collapse when only high price hours were concerned. Weighted nearest neighbors techniques are proposed in [5] and forecasting errors ranging between 5% and 16% were reported for the Spanish market; the variations in forecast accuracy over the studied period were attributed to various unpredictable factors including extreme weather conditions. A dynamic model based on system identification techniques is detailed in [6] with forecasting errors varying between 5% and 36% for the Italian, New England, and New York markets. Several statistical parametric and semi-parametric models are applied to forecasting in California and Nordic electricity market prices in [7] and errors ranging from 3% to 15% were reported. It was concluded in [7] that no single model could presently be chosen as the single best approach. A hybrid method, composed of support vector machines (SVMs) and a self-organized map, is applied to New England market prices in [8] with resulting forecasting errors of about 10% and 7% reported for the prices before and after the implementation of the standard market design, respectively. In [9] and [10], time series and neuro-fuzzy models are applied to forecasting Ontario's electricity prices; the forecast errors were reported to vary between 16% and 22% and the high forecast errors were attributed to the high volatility of prices in Ontario [9]. In [11], several linear and nonlinear models are employed to forecast Ontario prices for a three-year study period. The average forecasting errors in [11] vary between 22.98% and 31.86%, with a SVM-based model yielding the lowest errors. In another study [12] on forecasting prices in National Electricity Market of Australia, an SVM-based model is optimized using genetic algorithms and the resulting forecast errors are reported to vary between 16.39% and 23.26% for different time periods.

In addition to publications regarding numerical electricity price forecasting, several papers have included the estimation of prediction or confidence intervals. Among those, a method based on neural networks and Kalman filters is described in [13] and a different approach based on SVMs is detailed in [14].

Finally, [15] and [16] focus on the treatment and handling of price spikes and propose hybrid models to predict their occurrence.

## B. Data-Driven Model Building

Building a data-driven predictive model has three main steps: data preprocessing, feature selection, and model selection. This section provides a literature review and discussion regarding these three steps. When reviewing each of the steps, both the pertinent background information as well as the methods employed in the present work are discussed.

1) *Data Preprocessing*: Data preprocessing focuses on the initial treatment of data and includes gathering of information on data statistics, anomalies, missing values, and necessary data transformations. In the context of modeling electricity market price data, the reported studies highlight two aspects applicable

to price data in this step. First is the problem of outliers, where prices do not follow the observed historical patterns [2]. Outliers or abnormal prices generally result from supply scarcity or unexpected operational events such as the forced outage of a generation unit. Manipulating the outliers and data smoothing has been reported [7]; however, it has also been argued that unusual prices in electricity markets reflect the reality of price volatility in these markets and thus should neither be removed nor manipulated [24].

Second, electricity prices are not stationary and show strong daily and weekly seasonalities [2], [5]. In order to achieve better stationarity in the data, several data transformation approaches such as differencing, Box-Cox, and wavelet transformations have been utilized [2], [7], [24]. However, stationarity is not always a necessary condition, depending on the underlying assumptions of the employed models; for example, time series models are limited to stationarity data, but not neural networks. In the present work, only data normalization is applied since it has been found to improve classification accuracy.

2) *Feature Selection*: In this step, a subset of features (i.e., inputs or explanatory variables) is chosen from an initial feature set that efficiently captures patterns in the data. Two major groups of feature selection techniques are filter and wrapper methods. In filter methods [25], features are assessed for their relevance in explaining the target variable and those with the highest relevance are selected. Filter methods are fast and simple but the potential disadvantage of them is that feature selection is isolated from the prediction model. In wrapper methods, the complete feature set is explored for a near-optimal subset and the relevance of features is evaluated by the accuracy of the final predictions. While wrapper methods have been shown to provide high prediction accuracy [25], they are computationally expensive when compared with filter methods.

In the context of forecasting electricity prices, the most popular features are historical price and load data. Other features such as day and hour indexes, transmission constraints, load levels of neighboring systems [9], [10], [26], variants of reserve margin [9], [16], generator outages and temperature [10], and availability of different types of generation resources [3], [4], [27] have also been reported with varying degrees of effectiveness [10], [26]. Filter methods are the most frequently reported feature selection technique where linear measures including cross-correlation and auto-correlation [6], [7], [9], [10], [26], [28] as well as nonlinear measures such as mutual information [4] are used to evaluate the relevance of candidate features to price.

In the present work, historical price, load, and reserve information are considered in the initial set of features. The support vector machine recursive feature elimination (SVMRFE) [29] and kernel-based feature vector selection (KFVS) [30] techniques are employed for feature selection. In SVMRFE [29], a ranked list of features is created, and through a forward search process, the subset of features with the highest impact on classification accuracy is selected. KFVS [30] maps feature vectors into a lower dimensional space and their effectiveness in improving classification accuracy is evaluated through a forward search method. Note that KFVS is presented in [29] as a filter method, but it is customized and used as a wrapper method in

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات