Load - Price Forecasting Model Employing Machine Learning Techniques

R. A. Swief, Y. G. Hegazy, T. S. Abdel-Salam, and M.A Bader
Electrical Power and Machine Department, Faculty of Engineering, Ain-Shams University, Cairo, Egypt

Abstract-- Short term load forecasting is always an important study from operational and planning point of view. But short term price forecasting is a new topic. In this study, with the implementation of machine learning techniques, a new algorithm is proposed to predict both load and price values. A machine learning techniques such as Principal Component Analysis, and K nearest neighbor points, are applied as preprocessing techniques to assist the Support Vector Machine. The proposed model is implemented to build a closed loop dynamic model to forecast both 24 hours a head short term load and price values, which is a common study in markets. Because of leak of data, data proposed in this study are load, price time series only which are available in markets sites. Load and price volatilities time series are also needed to be obtained. The model has been trained and validated using data from, Australia electricity market. Moreover, the model has been tested using data from North England electricity market.

Keywords-- Deregulated electricity market, load forecasting, load-price elasticity, price forecasting, Machine learning techniques, SVM.

I. INTRODUCTION

Deregulated market is attracting more topics and studies under the new construction of the electrical power market. The targets of such changes are to increase the efficiency and the benefit to both producers and consumers. To achieve these targets, the prices and loads should be predicted to take the right strategic decisions in the market. Long term, middle, or short term forecasts are important for operational and planning studies. From operational point of view, daily target of power marketers is recommended be short term forecasting. One of the merits of the deregulated power market is the load/price mutual sensitivity. For price forecasting, it is a new study related to the new market. Points of view, economical and electrical, have been dealt with price forecasting differently. Dealing with load and price as two time series had encouraged many researchers to find the link between both [1-10]. The price volatility, the difference between two successive points in price time series is relatively high so the price time series looks like saw tooth wave form. On the other hand the changes in load are comparably small with respect to price changes, i.e. small sensitivity. Thus, no well defined relation could be easily detected between load and price time series. As an accepted approximation and with the help of statistical concepts, forecasting models had been conducted to link between both price and load together [1]. The model is relating the load and price at every hour, $t$, with the load and price at $t$ in a comparable day. Comparable day is the closest day which has a similar load pattern [1]. The aim of that study was to calculate the price and price volatilities based on a random walk pattern which was describing the price time series response. With the implementation of neural network algorithm, two neural networks have been implemented to forecast both load and price values each hour separately without linking both, only based on the data of the days with similar pattern [2]. Days of a similar pattern have been determined based on Euclidean norm [2]. With the help of load data, wavelet has been applied to decompose the price time series into three levels; high, medium, and low levels according to their loading conditions to facilitate the prediction for each level [3]. Fuzzy logic pattern has been implemented to separate the data into 10 levels according to their loading conditions [4]. Then, the forecasted price value was related to one of these levels and then; forecasted using the points of this level [4]. Both load and price data have been combined together in spike classification [5-7] and determination with the help of classification technique such as: data miming [5], neural network with the implementation of extreme learning machine algorithms, ELM [6], or Naïve Bayesian classifier [7]. A generalized autoregressive conditional heteroskedasticity, GARCH, which is a very well known regressive model in the stock market studies, has been implemented to forecast the prices independent of load conditions [8]. Load forecasting studies have involved prices time series not only in short term load forecasting but also in middle and long terms [9]. Recently, a new closed loop proposed model has been utilized to forecast both load and price values applying "co-co integration" technique [10]. Most probably, price in short term load forecasting study is not involved due to classification process which is implemented before conducting any study. With the help of NEMMCO data, an analysis of one day ahead price and load forecasting has been performed. In this study, with the aid of both load and price volatilities, a dynamic closed loop load-price model has been proposed to build a load price forecasting model with the help of machine learning techniques: PCA, Knn, and SVM.

II. VOLATILITY OF BOTH LOAD AND PRICE

Price changes have a descriptive meaning of market status. Equilibrium is achieved when the market satisfies the demand, this state has been translated into price which is called clearing price. Change in two successive clearing market prices is defined as price volatility. Using the same
definition, load volatility could be defined. From engineering point of view, volatility terms could be expressing the first derivative for both load and price time series with respect to time.

III. THE PROPOSED ALGORITHM

The proposed algorithm in this paper has been implemented to link both load and price forecasting models. For time series studies, the point at time t is related to a set of previous points from the same time series. In this study the number of previous points is equal to 7. This number has been selected using trial and error technique. Due to less volatility in load

Figure(1): The proposed algorithm

In the first stage, inputs are the previous 7 load and price points and their volatilities. In each step of the first stage, the output will be feeding the next step as illustrated in Figure(1). Sensitivities appear in the first stage between the forecasted price and the product of load and price, load * price index. In the second stage, another index arises, the model shows also sensitivity to the square of the forecasted values which are given from first stage. These indices have been investigated using different combinations between load and price data as helping indices. However, the model is described dynamic because for each tested point the model parameters will be calculated to generate the best performance. Stage two will continue until two successive unchanged values are reached. Forecasting studies is based on selecting the best values of the model parameters, and the features which affect the forecasted point. To fully achieve the expected output, preprocessing steps must be preceded on the data first.

A. Step1: Normalization

All observations, time series, must be normalized from 0 to 1. Therefore, observation will be with 0 mean and standard deviation equals to 1. In engineering term, it is like transferring the problem into per unit values, to neglect the effect of different units and types of observations.

B. Step2: Principle component analysis, PCA

Principle component analysis is analytical process which is determining the eigen-values of the inputs and ranking them according to their impact on the data. In other words, PCA is analyzing the signal into its principle components. Then, certain threshold will be determined to select the chosen components to reduce the selected data. The problem is that PCA works like low pass filter but the signal main characteristics could be existed in the filtered components. In this work, PCA has been employed with number of features equal to 1 which means that it works as a classifier. PCA only splits the data set into related or unrelated subsets to the forecasted point.

C. Step3: K nearest neighbor points, Knn

Among the selected points using PCA, Knn will relate the output of the model to the average of the K points of the previously selected points using PCA. The selection of K points is based on Euclidean norm. Afterwards, the K nearest points will be fed to the SVM model.

D. SVM model

SVM is one of the best classifying and forecasting techniques which is classified as a machine linear technique. The idea is to remap the data into linear space using kernel. Kernel is a function applied to transform the data into linear feature space such as: Gaussian, radial basis function, linear, quadratic, etc…. Thus, if these features could be separated regardless the kernel type, a linear relation could be represented for the feature space f as shown in Eq. (1).

\[ f = \langle \omega, x \rangle + b \]  \hspace{1cm} (1)

Where

\( x \) is the input data set to the model.
\( \omega \) is vector perpendicular to the plane.
\( b \) is a variable scanning the space.

Around the linear hyper-plane f, \( \varepsilon \) will be added, where \( \varepsilon \) is the radius of the tube which its axis is the linear hyperplane f. The mathematical formulation of this problem could be described as the primal optimization problem as shown in Eqs. (2-5). The optimization problem, which is a curve fitting problem, is to minimize the sum of square error. In the dual form, the target is to determine the Lagrange multipliers which are indicating the location of each feature space point, inside, on the upper hand, and on the lower hand of the \( \varepsilon \) intensive tube. The points which are outside the intensive tube are called support vector machine [11]. Figure 2 illustrates the positions of different feature space points.

Reference Number: W09-0009
Reference Number: W09-0009

The Online Journal on Power and Energy Engineering (OJPEE) Vol. (1) – No. (2)

SVM model has three free parameters which are controlling its performance, \( \varepsilon \), \( \sigma \), and \( C \) where \( \varepsilon \) is the radial basis function parameter, kernel parameter, \( C \) regularization parameter. The values of these parameters differ from case study to another one. A new SVM, \( \nu \)-SVM, model has proposed in [12] conducted another variable \( \nu \) which is optimizing the number of support vectors points. \( \nu \) is controlling over \( \varepsilon \). Cross validation is utilized to determine the best parameters used to minimize the error between the actual and the forecasted values. To facilitate cross validation job, \( C \), \( \nu \), and \( \sigma \) will be previously defined and leaving cross validation to determine the best \( \varepsilon \). All machines learning algorithms have been applied using Spider software which is operated using MATLAB.

IV. RESULTS AND SIMULATION

Electricity markets have been initiated since more than 10 years. NEMMCO [13] and NePool [14] are two electricity markets in Australia and North of England respectively. The implemented samples are half hourly data of 2008, NEMMCO, and an hourly data of 2001, NePool. In the proposed algorithm, some parameters must be determined previously such as SVM free parameters, K for Knn, and the number of previous points in the load and price time series. SVM free parameters are set to: \( C = 10000 \), \( \sigma = 0.0001 \), \( \nu = 1 \) and \( \varepsilon \) will be determined using cross validation. Trial and error will be utilized to determine K and the appropriate number of the previous points. By applying number of trials, the best previous number of the historical points equals to 7 and the k number is equal to 30 points. Based on Chaos theory, in any time series there are repetitive phenomena on small or large scale. Thus, to a certain extend we may assume the generalization of these values, because 7 is standing for the number of days in the week and 30 stands for the number of days in the month. Also based on Chaos the time series may be controlled from minor or major attractor, phenomena.

In the proposed model we will consider the time series to be the load or price values in hour, \( t \), throughout the whole year. The evaluation of the proposed model is based on \( \text{MAPE} \) which is described in Eq. (6).

Vertical time series contains the load or price values at hour, \( t \), in the all days of the year. The inputs for each point contain the 7 previous data from the vertical time series. The sample window is fixed number but the window is moving. This window contains 150 training points for testing the proceeded point to the window limit. To forecast the point numbered 151, the window covers points 1 to 150 and to forecast point 152, the window starts from point 2 and ends at point 151 and so on. The model forecasts two types of days, holiday and working day.

Tables 1 and 2 show the load and actual price, forecasted, and their MAPE\%, for NEMMCO 2008. The forecasted 12 hours, from mid day to mid night, is shown in Table 1 for Friday 26 of September, working day. The forecasted 12 hours, from mid night to mid day, is shown in Table 2 for Sunday 28 of September, holiday after updating the time series with the predicted values from Table 1.
MAEP \text{P over 12hours} = 1.6171\% \\
MAEP \text{L over 12hours} = 0.5424\%

V. CONCLUSION

In this paper, an algorithm consisted of machine learning techniques, PCA, Knn, and SVM is presented to predict the load and price values based on their time series only. The algorithm is dynamic closed loop which changes its values based on the new data entry. A very important entry for both load and price in second stage is the square of the value forecasted in stage one. With the same parameters, many samples have been validated and tested under different conditions and from different sources. The results show promising results. From this work, we may globalize this algorithm for load and price forecasting. The algorithm shows promising results facing different ways of handling time series.

REFERENCES


