Price Forecasting and Optimal Operation of Wholesale Customers in a Competitive Electricity Market

by

Hamidreza Zareipour

A thesis

presented to the University of Waterloo

in fulfilment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2006

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis addresses two main issues: first, forecasting short-term electricity market prices; and second, the application of short-term electricity market price forecasts to operation planning of demand-side Bulk Electricity Market Customers (BEMCs). The Ontario electricity market is selected as the primary case market and its structure is studied in detail. A set of explanatory variable candidates is then selected accordingly, which may explain price behavior in this market. In the process of selecting the explanatory variable candidates, some important issues, such as direct or indirect effects of the variables on price behavior, availability of the variables before real-time, choice of appropriate forecasting horizon and market time-line, are taken into account. Price and demand in three neighboring electricity markets, namely, the New York, New England, and PJM electricity markets, are also considered among the explanatory variable candidates.

Electricity market clearing prices in Ontario are calculated every five minutes. However, the hourly average of these 5-minute prices, referred to as the Hourly Ontario Energy Price (HOEP), applies to most Ontario market participants for financial settlements. Therefore, this thesis concentrates on forecasting the HOEP by employing various linear and non-linear modeling approaches.

The multivariate Transfer Function (TF), the multivariate Dynamic Regression (DR), and the univariate Auto Regressive Integrated Moving Average (ARIMA) are the linear time series models examined. The non-linear approaches comprise the Multivariate Adaptive Regression Splines (MARS), and the Multi-Layer Perceptron (MLP) neural networks. Multivariate HOEP models are developed considering two forecasting horizons, i.e., 3 hours and 24 hours, taking into account the case market time-line and the ability of market participants to react to the generated forecasts. Univariate ARIMA models are also developed for day-ahead market prices in the three neighboring electricity markets. The developed models are used to generate price forecasts for low-demand, summer peak-demand, and winter peak-demand periods.

The HOEP forecasts generated in this work are significantly more accurate than any other available forecast. However, the accuracy of the generated HOEP forecasts is relatively lower than those of the price forecasts for Ontario's neighboring electricity markets. The low accuracy of the HOEP forecasts is explained by conducting a price volatility analysis across the studied electricity markets. This volatility analysis reveals that the Ontario electricity market has the most volatile prices compared to the neighboring electricity markets. The high price volatility of the Ontario electricity market is argued to be the direct result of the real-time nature of this market. It is further observed that the inclusion of the just-in-time publicly available data in multivariate HOEP models does not improve the HOEP forecast accuracy significantly. This lack of significant improvement is attributed to the information content of the market data which are available just-in-time.

The generated HOEP forecasts are used to plan the short-term operation of two typical demand-side case-study BEMCs. The first case-study BEMC is a process industry load with access to on-site generation facilities, and the second one is a municipal water plant with controllable electric demand. Optimization models are developed for the next-day operation of these BEMCs in order to minimize their total energy costs. The optimization problems are solved when considering market price forecasts as the expected future prices for electricity. The economic impact of price forecast inaccuracy on both the case study is analyzed by introducing the novel Forecast Inaccuracy Economic Impact (FIEI) index. The findings of this analysis show that electricity market price forecasts can effectively be used for short-term scheduling of demand-side BEMCs. However, sensitivity to price forecast inaccuracy significantly varies across market participants. In other words, a set of price forecasts may be considered "accurate enough" for a customer, while leading to significant economic losses for another.

Acknowledgments

First, I would like to sincerely express my gratitude to Professor Claudio A. Cañizares and Professor Kankar Bhattacharya for their continuous support and encouragement, and invaluable guidance and supervision throughout the period of my PhD studies. I learned priceless lessons from their vision, personality, and professionalism.

I would also like to offer my appreciation to my Doctoral Committee Members, Professor Andrew Heunis and Professor Ehab El-Saadani, from the Electrical and Computer Engineering Department, and Professor David Fuller from the Management Science Department.

The unique research environment at the University of Waterloo, and the great services and support provided by the Department of Electrical and Computer Engineering, more specifically, the personal of the Departmental Graduate Office, is acknowledged.

Also acknowledged is the partial financial support of the Natural Sciences and Engineering Research Council (NSERC) of Canada, NRGen Inc. of Canada, and the Iranian Ministry of Science, Research, and Technology.

The valuable comments, information, and data provided by the personal of the Ontario Independent Electricity System Operator (IESO), specifically Mr. Peter Lafoyiannis, and Mr. Cristian Dragnea, are greatly appreciated.

Also appreciated is the information and technical collaborations provided by the Works Department, Region of Durham, Ontario, Canada, specifically the personal of the Municipal Water Plant in Ajax, Ontario, Canada.

Special thanks to my colleagues at the Electricity Market Simulation & Optimization Lab, University of Waterloo, for everything they did to make our lab a friendly place to work in: Hong, Valery, Sameh, Warren, Adrian, David, Gregor, Hassan, Rafael, Ismael, Hemant, Mingbo, and Amir.

My warmest thanks and appreciation to my parents who have undergone all the hardships in bringing me up and have offered me their support, encouragement, and love during the long years of my studies.

My endless thanks also to my parents-in-law, to my brothers and sister, and to my brothers- and sisters-in-law for their support and kindness.

Last but not the least, I would like to express my deep appreciation to my beloved wife, Maryam, for her understanding, patience, unconditional support, and true love during all sunny and cloudy days ever since we met.

Dedication

This thesis is dedicated to my wife, Maryam, and to our lovely angels Helya and Ava.

Contents

List of Tables

List of Figures

Acronyms

ADE Availability Deceleration Envelope

ANN Artificial Neural Network

ARIMA Auto Regression Integrated Moving Average

BEMC Bulk Electricity Market Customer

CAOR Control Action Operating Reserve

CMSC Congestion Management Settlement Credit

DACP Day-Ahead Commitment Process

DAIOG Day-Ahead Intertie Offer Guarantee

DAGCG Day-Ahead Generation Cost Guarantee

DG Distributed Generation

DSPS Dispatch Scheduling and Pricing Software

DR Dynamic Regression

EDRP Emergency Demand Response Program

ELRP Emergency Load Reduction Program

FIEI Forecast Inaccuracy Economic Loss

HADL Hour-Ahead Dispatchable Load

HADLOG Hour-Ahead Dispatchable Load Offer Guarantee

HOEP Hourly Ontario Energy Price

ICP Intertie Congestion Price

IESO Independent Electricity System Operator

i.i.d. identically and independently distributed

IOG Intertie Offer Guarantee

ISO Independent System Operator

LMIP Linear Mixed-Integer Programming

MAPE Mean Absolute Percentage Error

MARS Multivariate Adaptive Regression Splines

MCP Market Clearing Price

MIO Multi-Interval Optimization

MLP Multi-Layer Perceptron

MMCP Maximum Market Clearing Price

MPMA Market Power Mitigation Agreement

NIS Net Interchange Schedule

NISL Net Interchange Schedule Limit

OPG Ontario Power Generation Inc.

PDP Pre-dispatch Price

PDR Pre-dispatch Report

PSC Predicted Supply Cushion

SGOL Spare Generation On-Line

SSR System Status Report

TF Transfer Function

10N 10 Minute Non-Synchronized Operating Reserve

30R 30 Minute Non-Synchronized Operating Reserve

10S 10 Minute Synchronized Operating Reserve

Chapter 1

Introduction

1.1 Research Motivation

During the past decade, many countries have moved from a vertically integrated operating environment and have introduced privatization, competition and hence deregulation of their power industry. In vertically integrated systems, electricity price is usually regulated and customers are offered pre-determined tariffs. In competitive electricity markets, however, market operators determine the electricity price for specific intervals during a day (e.g., 5-minute or hourly prices), taking into account various economical and operational factors. Thus, the supply and demand side market participants are faced with the new challenge of electricity market price uncertainty in their daily operations.

In a competitive electricity market, participants have the option of trading electricity through spot markets, forward markets, or physical bilateral contracts. Furthermore, Bulk Electricity Market Customers (BEMCs) may choose to supply their energy needs using on-site Distributed Generation (DG) facilities. BEMCs may also adopt other feasible load management strategies (e.g., load shifting) in order to minimize their electricity costs. Given such a wide variety of options, conjecture of the future electricity market

prices is essential for market participants in order to optimize their operation.

A prior knowledge of electricity market price fluctuations helps power suppliers in setting up rational offers in the short-term, as well as designing physical bilateral contracts in the medium-term. In addition, generation expansion plans are directly influenced by the trend of electricity market prices in the long-term. For the demand-side, an insight into the market price trends and fluctuations is crucial in order to design optimal operational strategies in the short-term. Furthermore, this insight can help the demand-side to hedge against the risk of price volatility through physical bilateral contracts in the medium-term, and aid in the planning of their investments in DG options over the long-term. In view of the these facts, electricity market price forecasting has gained a critical significance in electricity market research during the past recent years. However, there are many aspects of the price-forecasting problem that remain to be addressed, as discussed in the following sections.

Load management programs are receiving a great deal of attention from both the Independent System Operators (ISOs) and the BEMCs. The ISOs look for reliable tools to reduce the system demand during stressed operating hours. The BEMCs, on the other hand, look for feasible options to avoid the high electricity market prices during peak hours. Employing on-site electricity generation facilities and managing the controllable part of their load are two viable options being promoted intensively by the ISOs for the BEMCs. These options call for new research to derive optimal short-term operational strategies for the BEMCs while taking into account the uncertain behavior of future electricity market prices.

This thesis addresses the important issues of forecasting future electricity market prices and scheduling the short-term operation of BEMCs. This research is novel because no previous work has been reported that deals with *integrating* the operation of the demand-side with the competitive electricity markets to achieve increased benefits for the system as a whole. The Ontario electricity market is selected as the primary case

market and its structure is studied in detail. Thereafter, a set of explanatory variable candidates are selected from Ontario and its neighboring markets, and various methods are employed to forecast future Ontario electricity market prices. The generated Ontario market price forecasts are then used for short-term operation planning of two case study BEMCs, one with on-site generation facilities and one without. Next-day operation of the case study BEMCs is formulated and the corresponding optimization problems are solved to minimize their total energy costs. Economic impact of price forecasting inaccuracy in both the case studies is also analyzed.

The "big picture" of this research is presented in Figure 1.1. Historical data from the electricity market is used by the forecasting system to generate electricity market price forecasts. An optimization model is developed considering the BEMC's technical characteristics, generated electricity market price forecasts, and any other relevant information from the electricity market. The optimization problem is solved and optimal operation schedules are derived, so that the net cost of electricity transactions with the electricity market is minimized. Effectiveness of using price forecasts for operation planning is then assessed by an economic impact analysis procedure.

The Ontario electricity market is selected as the case market because of its various unique features. For example, several kinds of price and revenue caps exist for wholesale market participants and retail customers. Also, even after its deregulation, about 75% of the generation capacity is held by one generating company. In addition, Ontario is a single-settlement uniform-price real-time market, while the other four adjacent North American markets are two-settlement markets with nodal prices. Furthermore, the Ontario power grid is directly connected to the New York and Midwest electricity markets and indirectly connected to the New England and PJM markets. It is also connected to the regulated utilities in Quebec and Manitoba, both having significant energy transactions with utilities in the United States. As a result of these significant interconnections, the operation of the Ontario electricity market can greattly affect the Northeast and the Mid-West power interconnections in North America.

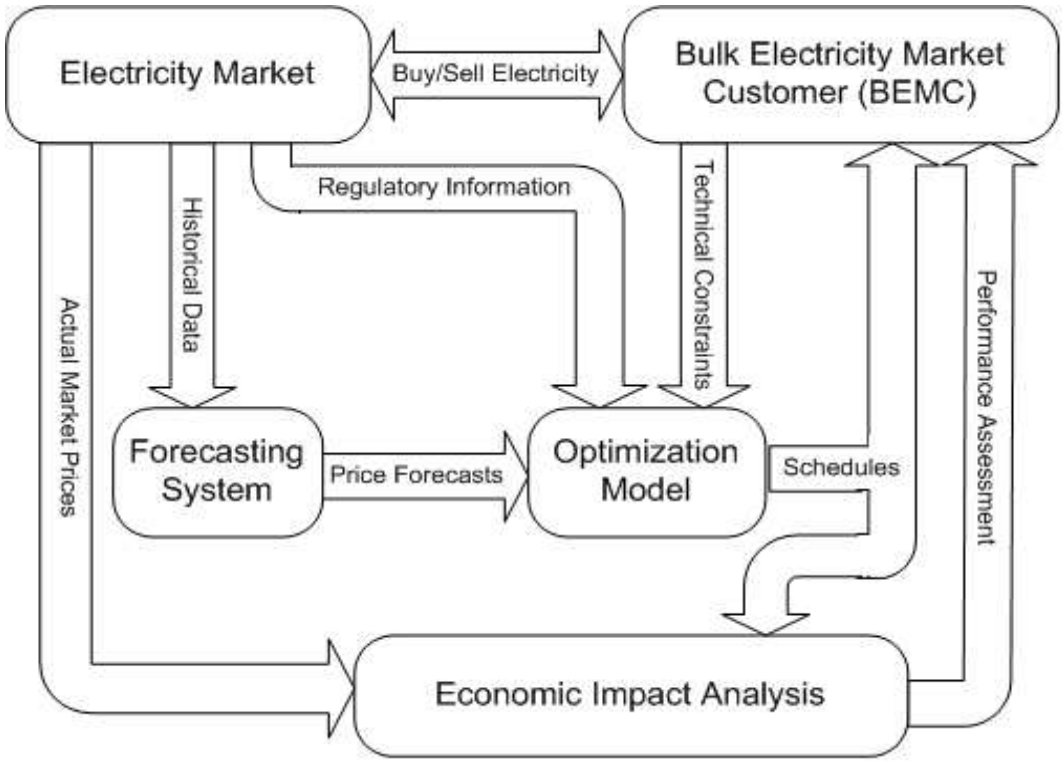

Figure 1.1: The “big picture”.

1.2 Literature Review

1.2.1 Electricity Market Price Forecasting

Electricity market price forecasting is a relatively new area of research, unlike the electric load forecasting problem [1]. Chronological load variations show a high degree of seasonality and dependence on exogenous factors, especially weather-related variables. These dependencies are well studied and addressed and load patterns for various situations are well known. However, the relationship between electricity market price and other factors (e.g., demand) have not been clearly addresses yet, if at all. For example, an analysis by Vucetic *et.al* in [2] to understand the relationship between the demand and price in the electricity market of California shows that several price behavior regimes

may exist, and numerous characterizing models are needed to illustrate the approximate price-demand relationship in different situations.

Electricity market prices are highly volatile, suffering from unusually high or low price spikes [2,3,4]. Moreover, prices are shown to be more volatile in electricity markets than other financial markets [5]. These are mainly because of the fact that electrical supply and demand need to be on a real-time balance and, unlike other commodities, it is practically impossible to store electricity economically. In general, various factors may affect electricity market price volatility such as, unexpected physical problems in generation and transmission systems, sudden changes in weather conditions, availability of relatively inexpensive generation facilities (e.g., nuclear and hydro), volatility in fuel markets, and possible collusion among market players.

The factors which may specifically affect price volatility, however, vary across electricity markets. A study by Benini *et al.* [3] shows that in California, PJM and Spain electricity markets, price volatility is strongly connected to the installed generation capacity. However, in the erstwhile UK power pool, market regulations, such as the inclusion of a "capacity payment" factor in the spot price, are found in [3] to be the dominant factors leading to highly volatile prices. It is also demonstrated in [3] that proper market regulations can restrict any possible collusive behavior among generation companies and hence reduce electricity price volatility, as in the case of the Spain's electricity market. These facts imply that it would be very difficult to generalize a pattern in price behavior across electricity markets, and hence, price forecasting is a complex *market-specific* phenomenon.

Electricity market prices may be forecasted by using either simulation-based or analysis-based methods. In simulation-based methods, which are mainly used by power utilities and market operators, the actual market dispatch is mimicked by considering initial supply offers, demand bids, and system operating constraints [6, 7]. Although simulation-based price forecasting can provide a more detailed view of the price fluctuations, they

require full insight into the system operation and hence are not practical for market participants. Analysis-based methods, however, use historical operation data to forecast future prices. The present thesis employs analysis-based methods for short-term electricity market price forecasting with a maximum 24-hour forecasting horizon.

The first attempt to forecast electricity market prices was reported in 1997 [8], where Artificial Neural Networks (ANNs) have been used to predict the future System Marginal Prices (SMPs) in the erstwhile UK power pool. Subsequently, several methods have been reported in the literature for short-term electricity market price forecasting. Among these, artificial intelligence-based methods [9, 10, 11, 12, 13, 14, 15], univariate Auto Regression Integrated Moving Average (ARIMA) models [14,16,17,18], multivariate Dynamic Regression (DR) models [14,19], multivariate Transfer Function (TF) models [14,19,20], input/output hidden Markov models [21], wavelet models [14, 17], and General Auto Regressive Conditional Heteroscedastic (GARCH) models [22] are some of the methods that have been reported.

Szkuta *et al.* [9] have developed ANN models using historical price, demand, and system reserve data, and 1-hour-ahead price forecasts are generated for the Victorian (Australia) power market. Multi-Layer Perceptron (MLP) and Radial Basis Function (RBF) neural networks have been employed in [10, 11, 12, 13] to forecast the average on-peak (from 7 AM to 11 PM) and average off-peak (from 1 AM to 7 AM and from 11 PM to 12 PM) New England electricity market prices. Historical price, market demand, fuel prices, and system reserve data are considered as input factors in these studies, and 1-step-ahead forecasts are generated. Employing different network structures and implementing improved training algorithms have led to slightly better forecasting accuracy in these studies.

Contreras *et al.* [16] have developed univariate ARIMA models to forecast electricity market prices in California and Spain. Multivariate TF and DR models have been applied by Nogales *et al.* [19] to the prices in Californian and Spanish electricity markets, where

demand is the only explanatory variable used. Comparing the forecast results obtained in [16] with those obtained in [19] shows that the forecasts generated by using multivariate TF and DR models have gained higher accuracy than those with ARIMA models. Nogales and Conejo [20] have used TF models to forecast PJM electricity market prices, where demand is again the only explanatory variable studied. Univariate ARIMA models are also used in [18] for predicting Leipzig's (Germany) electricity market prices. Numerous forecast models are developed in [18], based on which it is concluded that in the case of the Leipzig electricity market, forecasting hourly prices using separate models for each hour yields better results than modeling all 24 hourly prices as a whole.

Wavelet transformation used by Conejo *et al.* in [17] is observed to slightly improve the forecasting accuracy of univariate ARIMA models. GARCH models are used by Garcia *et al.* in [22] to account for fluctuations in the variance of electricity prices in the Spanish electricity market. Input/Output hidden Markov models are employed by Gonzalez *et al.* in [21] to generate 1-hour-ahead forecasts for Spanish electricity market prices; however, their forecast results do not depict any significant improvement compared to previous works. In [14], TF, DR, ARIMA, wavelet, and ANN models are used to generate 24-hour-ahead price forecasts for the PJM electricity market, where the historical demand is the only explanatory variable considered in this study.

It can be concluded from the above-cited studies that multivariate TF and DR models have yielded more accurate results than other methods, where comparison has been possible. For example, while the price forecasts generated by using TF and DR models for the Spanish electricity market report a weekly Mean Absolute Percentage Error (MAPE) on the order of 5% [19], the weekly MAPE of the forecasts generated by the input/output hidden Markov models for the same market is 15.8% [21]. It is also shown in [14] that the TF and DR models outperform the ARIMA, ANN, and wavelet models for forecasting PJM electricity market prices.

The Hourly Ontario Energy Price (HOEP) is a province-wide uniform market price,

applicable to wholesale electricity customers in Ontario. Forecasting the HOEP has been a challenging issue for both market participants and the Ontario Independent Electricity System Operator (IESO) [23]. Simulation-based HOEP forecasts are published by the IESO, referred to as Pre-Dispatch Prices (PDPs), and these are updated every hour until real-time [24]. The last published PDP for a given hour, called 1-hour-ahead PDP, is considered the final price signal to be sent from the IESO to Ontario market participants before real-time. Analysis of the historical market data reveals that there exists a significant deviation of the HOEP from 1-hour-ahead PDPs, with an MAPE of 35.2% over the first four years of market operation (May 1, 2001-April 30, 2005). In the only other reported research on HOEP forecasting [15], 1-hour-ahead HOEP forecasts are generated by using a non-linear neuro-fuzzy model. Ontario demand, forced outages, and capacity excess/shortfall are considered as input factors. The forecasts have resulted in daily MAPEs to be varying between 19.83% to 24% across different scenarios. To the best of the author's knowledge, no work has been reported so far that applies time series models to HOEP forecasting.

From the above cited studies, it can also be concluded that price forecasting accuracy varies significantly across the electricity markets. For example, the MAPE of the forecasts generated for Ontario market prices is in the order of 20-24% [15], and in the range of 10-15% for PJM prices [14], while varying between 3 to 8% for Spanish market prices [16, 19]. No studies have been reported that investigate and explaine these differences.

Furthermore, an important factor to observe is that the above-cited references have employed *after-the-fact* data for model building and out-of-sample forecasting. Although building price models using *ex-post* data is important in order to discover the factors influencing price behavior, these data are not available before real-time for a practical price-forecasting tool. This calls for research on developing forecasting models based on publicly available *before-the-fact* electricity market data.

The forecasting horizon is considered to be one hour in some of the above studies, since one-hour-ahead forecasting is very useful in examining the efficiency of the developed forecasting models. In addition, there may be some market participants who are able to refine their operation based on 1-hour-ahead market price forecasts. However, in practice, the majority of market participants are not able or allowed to change their operation schedules one hour before real-time, depending on the market rules and structure. Selecting appropriate forecasting horizons, when considering market time-line and the participants' ability to react to price forecasts, has not been systematically addressed in the literature yet.

The Multivariate Adaptive Regression Splines (MARS) approach was first introduced by Friedman [25] to approximate the relationship between a dependent variable and a set of explanatory variables in a piece-wise regression. Capability of MARS for modeling time series data was subsequently demonstrated in [26], where lagged values of the time series were treated as explanatory variables. Although application of MARS has been reported for modeling a variety of data with promising performance in recent years, such as for speech modeling [27] and mobile radio channels prediction [28], no work has been reported on applying MARS for electricity market price forecasting.

1.2.2 Demand-side Scheduling in Competitive Electricity Markets

Optimizing electricity consumption and load management is a mature area of research which was originally developed in the 1980s [29]. The core idea in load management is to change the shape of the load curve of electricity consumers, and hence of the system as a whole, so that less electricity is consumed during the system peak demand hours. In general, load management programs are promoted by energy utilities to provide their consumers with economic incentives to reduce their load during the stressed periods of system demand. Electricity consumers are induced to take advantage of the available economic incentives by changing their load patterns and responding to high electricity

prices. The *price-responsive* group of electricity consumers either have on-site generation facilities as an alternative source of electricity, or the nature of their controllable load enables them to reduce their demand during high price hours [30, 31]. Scheduling the short-term operation of this group of electricity consumers in a competitive electricity market environment is the focus of this research.

On-site generation of electricity using DG facilities has already been widely adopted in several countries, and estimations show that it will contribute more in the future to the power generation business [32]. Major manufacturers, market research organizations, and consulting companies believe that by the year 2010, the DG market size would be in the range of $10 to $30 billion in the US and $75 billion worldwide. Furthermore, developments in small size generation unit markets (under 5 MW) also represent the growing trends toward DG in power industry. Potential economic benefits from on-site electricity generation are: impact on electricity prices; deferral of upgrades to the transmission and distribution systems; utilization of waste energy resources and fuel flexibility; improvement in power quality; provision of ancillary services, cogeneration or Combined Heat and Power (CHP) production; providing reliable power; off-grid applications, and microgrid benefits [33, 34, 35, 36, 37].

Various aspects of on-site generation of energy such as sizing, reliability, and investment have been addressed in the literature. The electrical and thermal energy needs of industrial loads can be separately met by the grid and a conventional thermal energy generator, respectively. However, several industries choose to have their own on-site generation facilities with CHP capability, or cogeneration, to improve their energy efficiency. The overall efficiency of available cogeneration systems is typically around 75% to 90%, which makes them a viable option for energy production, specifically when electricity prices are high [31, 38, 39, 40, 41, 42, 43].

The common objective for optimal scheduling of cogeneration systems is to minimize the total cost of meeting thermal and electrical energy needs, subject to equipment

specifications and operational constraints. Pre-defined electricity tariffs and bilateral energy contracts are considered among the optimization constraints in [30, 31, 39, 40], in a regulated electricity sector environment. Optimal operation of cogeneration systems is studied in [41, 42, 43] within the context of liberalized energy markets. In these studies, optimization models are developed to decide the mid-term or long-term energy contract options available to cogeneration system owners. Scheduling cogeneration systems when considering short-term electricity market prices has not been reported in the literature.

Electricity consumers with controllable load are capable of adopting suitable load management strategies to optimize their usage of grid electricity [30, 44, 45]. In [30], an optimization model is developed for a price-responsive process industry load with no access to on-site generation facilities. The studied load is an industrial flourmill with storage facilities and night-work shift options. Production commitments are optimally scheduled to minimize electricity costs under time-of-use electricity rates. In [44], a fuzzy-based decision making algorithm is developed for industrial load management; the algorithm is applied to a cool energy storage air conditioning system for minimizing electricity costs under the Taipower's time-of-use rates. In [45], a multi-objective optimization model is developed for municipal water plant management. Minimizing energy costs is part of the objective function, given the fact that pumping can be scheduled at off-peak hours. To the best of the author's knowledge, no studies have been reported which address the short-term operational scheduling of electricity consumers with controllable load in a competitive electricity market environment.

Participants in a competitive electricity market need to consider the challenge of electricity market price uncertainty in their operation planning. One feasible approach to deal with this challenge is to forecast short-term electricity market prices and to schedule their operations accordingly. This approach has been exercised in [46] to derive optimal bidding strategies for a thermal-based power supplier. To the best of the author's knowledge, no studies regarding the applicability of electricity market price forecasts for short-term operation planning of BEMCs have been reported in the literature.

Electricity demand forecasting is the core component of supply scheduling programs in power systems, and economic impact of demand forecasting inaccuracy has been addressed in the literature [47, 48, 49]. Studies have shown that improving demand forecasting accuracy, even by 1%, can help power utilities save millions of dollars [48]. In the same way, an accurate forecast of future electricity market prices may possibly save the electricity market participants significant costs. However, no studies have been reported in the literature that examine the economic impact of employing price forecasts for short-term scheduling of demand-side market participants.

1.3 Objectives

Considering the state-of-the-art of research discussed above, the main objectives in this thesis can be stated as follows:

1. Review rules, regulations, and structure of a selected case market, which is the Ontario electricity market, in order to gain a clear understanding of its operation.
2. Detect explanatory variables candidates which are available *before real-time* from publicly available sources, and examine if these variables could potentially explain market price behavior.
3. Employ well-established linear and non-linear forecasting methods, including MARS, to relate price behavior in the case market to the selected explanatory variables.
4. Generate price forecasts for the case market considering forecasting horizons that are practical from the viewpoint of market participants.
5. Develop forecasting models for the case market's neighboring electricity markets and compare their accuracy with that of the models developed for the case market.

6. Explain the differences observed in price predictability across the case market and its neighboring markets.

7. Apply the generated price forecasts to short-term scheduling of two case study BEMCs, one with on-site generation facilities and one without.

8. Analyze the economic impact of price forecasts inaccuracy on the BEMC case studies.

1.4 Thesis Outline

Chapter 2 presents a detailed overview of the operation of the Ontario electricity market. The Ontario market outcomes, such as market prices for energy and operating reserves, and demand, are also analyzed in this chapter for the first four years of market operation.

Chapter 3 describes the process of selecting explanatory variable candidates, and applying linear time series models to forecast the HOEPs. Price forecasting models for the neighboring electricity markets, i.e., the New England, New York, and PJM markets are also developed in this chapter.

Chapter 4 discusses the application of two non-linear approaches, i.e., MARS and MLP networks, to HOEP forecasting.

Chapter 5 presents a comprehensive price volatility analysis for Ontario and its neighboring markets. This volatility analysis explains the differences observed between the accuracy of the price forecasts generated for the Ontario electricity market and those generated for the three neighboring electricity markets.

Chapter 6 presents the application of electricity market price forecasts to short-term scheduling of two BEMC case studies. Economic impact of price forecasting inaccuracy on the studied BEMCs is also analyzed in this chapter.

Finally, Chapter 7 summarizes the main content and contributions of this thesis, and suggests directions for possible future research work.

Chapter 2

An Overview of the Operation of the Ontario Electricity Market[1]

2.1 Introduction

In order to gain a clear understanding of the operation of the Ontario electricity market, which is selected as the case market in this thesis, a detailed overview of this market is presented in this chapter. The procedures for clearing the energy and operating reserve markets, pre-dispatch and real-time dispatch of the supply and demand sides, inter-jurisdictional energy trading, and procurement of ancillary services are discussed. Furthermore, the main market outcomes, namely pre-dispatch and real-time energy prices, operating reserve prices, and market demand, are studied for the period May 1, 2002 to April 30, 2006. The programs introduced by the IESO to enhance the operational aspects of the Ontario electricity market are also analyzed and their effectiveness is discussed.

[1]Findings of this chapter have been partly presented in the 2005 IEEE PES General Meeting, San Francisco, USA [24], and submitted to in the *IEEE Transactions on Power Systems* [50].

2.2 Ontario Power Industry at a Glance

In the Ontario electricity sector prior to deregulation, Ontario Hydro along with some small municipal utilities generated, transmitted, and distributed electricity to their customers across the province. In that era, electricity prices were regulated by the provincial government. The Ontario Electricity Act of 1998 reorganized Ontario Hydro into five companies, and on April 1, 1999, these new companies were created, namely, the Independent Market Operator (IMO), Hydro One Inc., Ontario Power Generation Inc. (OPG), the Electrical Safety Authority (ESA), and the Ontario Electricity Financial Corporation (OEFC). The ESA is responsible for the electric industry standards, and the OEFC manages financial services of the erstwhile Ontario Hydro and its successors.

The Ontario wholesale electricity market opened on May 1, 2002, two years after the originally scheduled date. This market consists of a real-time physical market for energy and operating reserves, and a financial transmission rights market. The Electricity Act of 2004, renamed the IMO as the Independent Electricity System Operator (IESO). The IESO is a non-profit company which is regulated by the Ontario Energy Board and its core responsibility is to operate the Ontario wholesale electricity market.

Hydro One Inc., wholly owned by the Government of Ontario, is the major transmission company that owns and operates Ontario's transmission network. The transmission system has remained regulated and the Ontario Energy Board determines the transmission and distribution tariffs. The distribution system is also regulated by the Ontario Energy Board with 91 local distribution companies delivering electricity to the retail customers. The Ontario's high voltage transmission system has interconnections with Manitoba, Quebec, New York, Michigan and Minnesota control areas through 12 lines. These high voltage interties allow 4,000 MW of electric power transactions.

The OPG owns about 75% of the 30,662 MW installed capacity in Ontario. The total installed generation capacity consists of: 11,397 MW of nuclear power plants (37.2%); 7,855 MW of hydro and other renewable resources (25.6%); 6,434 MW of coal-fired

generation facilities (21%); and 4,976 MW of oil/gas-fired power stations(16.2%). Energy imports from the neighboring areas are also an important part of Ontario's supply portfolio. The highest Ontario summer peak demand was recorded in August 2006 at 27,005 MW, an about 3% increase with respect to the previous peak demand recorded in July 2005 at 26,160 MW.

A target of refurbishing, rebuilding, or replacement of 25,000 MW of generating capacity by the year 2020 has been set by the government to meet the Ontario demand while replacing polluting coal-fired generation. The government has also set a target of reducing Ontario's energy consumption by 5% by 2007. Furthermore, the Ontario Energy Board is developing a plan for installation of smart meters for all consumers by 2010. The Conservation Bureau of the Ontario Power Authority is responsible to pursue the Government's energy conservation and demand management programs. The Ontario Power Authority was established in 2005, under directions of the Electricity Act of 2004, to ensure the long-term adequacy, reliability, security, and efficiency of the Ontario electricity sector.

The supply and demand side entities within the province having direct connection to the transmission network must participate in the Ontario electricity market [51]. This group of entities consists of generation companies, large industrial loads, and local distribution companies. Other parties with physical assets which are connected to the distribution network are referred to as "embedded" facilities and can choose to either participate in the market or buy/sell power through contracts with power retailers. There are, however, other market participants without a physical connection, such as power traders, or boundary entities who import/export power to/from Ontario and may participate in the physical or financial markets.

Energy market participants in Ontario can choose to buy or sell energy through bilateral contracts. However, bilateral contracts may not be necessarily reported to the IESO. Bilateral contracts are not considered in the process of scheduling and dispatch of energy,

and have a small share in the whole electricity trading in Ontario.

Market participants are grouped into dispatchable and non-dispatchable. Dispatchable market participants actively bid into the market and receive dispatch commands every five minutes to reach a specified level of generation or consumption. In contrast, non-dispatchable market participants are "price-takers" and accept to produce or consume power at real-time and be paid or charged at the hourly price prevailing at that time. Most of the loads in Ontario are non-dispatchable, and most of the generation facilities are dispatchable. Non-dispatchable generators are those generation facilities which cannot follow dispatch instruction as required by the IESO; this group of generators are either small self-scheduling units, such as hydro plants running on a small river, or intermittent generators such as wind farms.

A uniform, province-wide, Market Clearing Price (MCP) is determined for Ontario every five minutes. The hourly average of these five-minute MCPs is defined as the Hourly Ontario Energy Price (HOEP). For financial settlements, the MCP applies to dispatchable market participants, whereas the HOEP is applicable to non-dispatchable participants. Zonal MCPs are also calculated for each of the 12 intertie zones. The pre-dispatch and real-time Ontario MCPs and zonal MCPs are the basis of settling the imports and exports.

The Ontario electricity market has 289 market participants (May 2006). Wholesale prices apply to most of the electricity consumers having more than 250 MWh/year of electricity consumption, whereas, prices are capped at the retail level. The capped prices are determined based on the Regulated Price Plan (RPP) which was initiated by the Electricity Act of 2004. Residential customers pay 5.8 cents for the first 600 kWh per month and 6.7 cents for the consumption over this threshold, as of May 2006. Designated large-volume consumers such as schools, universities, hospitals, farms and specified charities also pay the RPP rates.

2.3 The Physical Market for Energy and Operating Reserves

The physical market is jointly optimized for energy and operating reserves. Three separate operating reserve classes are used in the Ontario market, namely, 10 minute synchronized operating reserve (also called 10 minute spinning or 10S), 10 minute non-synchronized operating reserve (also called 10 minute non-spinning or 10N), and 30 minute non-synchronized operating reserve (30R). Only dispatchable generators are authorized to offer the 10S, while dispatchable generators and loads, and boundary entities can participate in the market for 10N and 30R operating reserves. These three operating reserve classes are requirements determined by the North American Electric Reliability Council (NERC), and the Northeast Power Coordinating Council (NPCC).

The physical market is optimized to maximize the market's "Economic Gain", which is conceptually same as the social welfare. The market optimization program, referred to as Dispatch Scheduling and Pricing Software (DSPS), consists of several system and data analysis blocks, with a dc-based security-constrained optimal power flow block together with an ac-power-flow-based contingency analysis tool as its heart [52, 53]. Several penalty functions and violation variables are also defined to allow the DSPS to automatically violate system constraints when a solution is not found otherwise. A separate ac power flow is run to calculate the transmission losses, which are incorporated in the power balance requirements constraint using appropriate penalty factors, and reactive power dispatch and voltage profiles.

The market Economic Gain is defined as the difference between the value of the

electricity produced and the cost of producing that electricity, as follows:

$$\begin{aligned}\text{Economic Gain} = &\sum_{j} \rho_{D,j} \times P_{D,j} \times PF_{D,j} \\ &- \sum_{i} \rho_{S,i} \times P_{S,i} \times PF_{S,i} - \sum_{k,c} \rho^{OR}_{k,c} \times P^{OR}_{k,c} \\ &- \text{Violation Variables} - \text{Tie Breaking} \end{aligned} \tag{2.1}$$

where $P_{D,j}$ and $P_{S,j}$ are demand bid and supply bid blocks respectively; $\rho_{D,j}$ and $\rho_{S,j}$ are the prices associated with the $P_{D,j}$ and $P_{S,j}$; PFD and PSF are the defined loss penalty factors associated with each demand or supply bid; $P^{OR}_{k,c}$ is a bid block for class c of operating reserves with a price $\rho^{OR}_{k,c}$; the Violation Variables are defined to represent the cost of violating respective constraints; and the Tie Breaking function deals with the bids that have the same price. The algorithm determines the best trade-off between energy and operating reserves using appropriate constraints and operational functions.

The DSPS is run in two time-frames, i.e., the pre-dispatch and real-time (dispatch), and in two modes, i.e., unconstrained and constrained. The pre-dispatch run is used to provide the market participants with the "projected" schedules and prices for advisory purposes in advance, while the final schedules and prices for financial settlement are determined in the real-time run. In the "unconstrained" algorithm, the Economic Gain is optimized based on supply and demand bids, but most of the physical power system constraints are neglected except for some of the operational constraints, such as intertie energy trading limits and ramping constraints. In the "constrained" algorithm however, system security limits together with a representation of the Ontario transmission network model are considered.

2.3.1 Market Time-line

Hourly supply and demand bids as well as operating reserves bids for a dispatch day must be submitted to the IESO between 6:00 and 11:00 on the pre-dispatch day. The

bids may be revised up until two hours prior to the dispatch hour without any restriction. Furthermore, the quantity of bids can be revised up until 10 minutes before dispatch hour (for imports and exports 60 minutes prior the dispatch hour) with the permission of the IESO (see Fig. 2.1).

Pre-dispatch

From 11:00 of the pre-dispatch day, the pre-dispatch version of DSPS is run hourly for the remaining hours of the pre-dispatch day and for 24 hours of the dispatch day. This procedure uses the unconstrained algorithm to determine the projected market clearing prices for energy and operating reserves, referred to as the Pre-Dispatch Prices (PDPs), and unconstrained schedules. The resulting schedules are then analyzed for any network constraint violations iteratively until all violations are resolved. If violations exist, the associated constraint equations are incorporated in the constrained algorithm. The economic gain is optimized again using the constrained algorithm and final schedules are generated and sent to each market participant.

The pre-dispatch run covers a range of 37 hours (at 11:00 on the pre-dispatch day) to 14 hours (at 10:00 on the dispatch day), and provides a first glance on future schedules and prices. Every hour after 11:00 on the pre-dispatch day, revised pre-dispatch schedules and prices are derived for the rest of the pre-dispatch day and/or dispatch day, until 11:00 on the dispatch day, which then becomes the pre-dispatch day for tomorrow (see Fig. 2.1). The results for energy prices and total market demand at each pre-dispatch run are publicly available by the end of the hour or during the next hour.

Real-time

In real-time, the dispatch version of DSPS is run every five minutes to derive prices, schedules and dispatch instructions for each interval. Both the unconstrained and constrained algorithms start at the beginning of each interval. The unconstrained algorithm

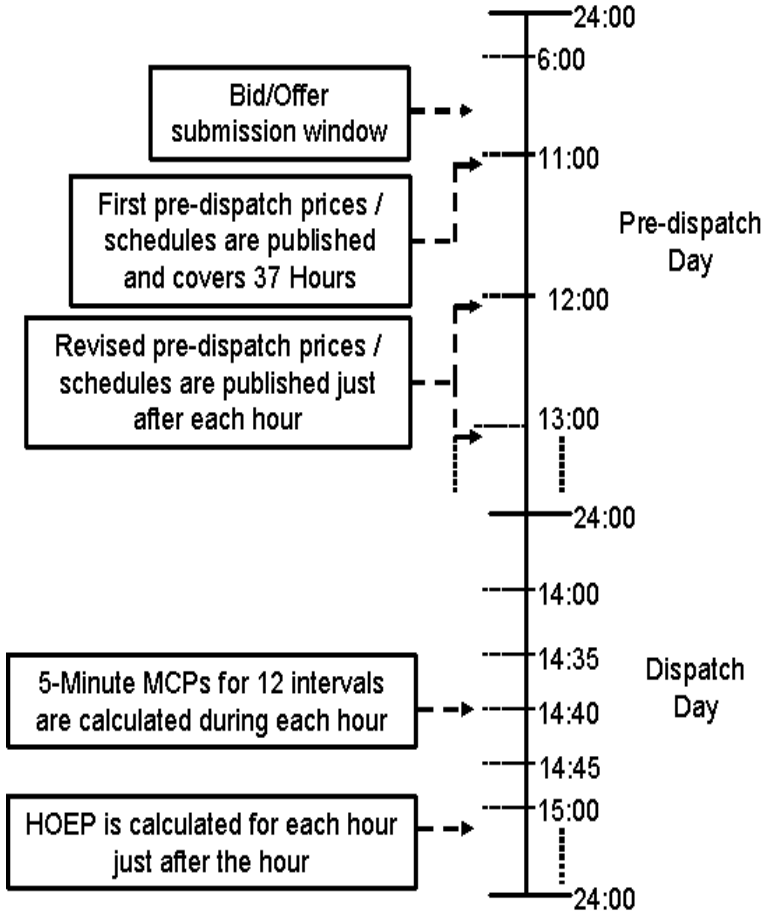

Figure 2.1: Market time-line: pre-dispatch and dispatch days.

determines the market schedule and prices for the interval that just passed, based on real-time supply and consumption (see Fig. 2.2). The constrained algorithm provides real-time schedules and dispatch instructions for the next interval. Schedules and prices obtained in real-time are the basis of all financial settlements. It is to be noted that after June 2004, a Multi-Interval Optimization (MIO) algorithm was implemented by the IESO, thereby the constraint algorithm derives real-time schedules for an interval while also considering four other advisory intervals. The MIO project is described in more details in Section 2.4.4. The differences of the pre-dispatch and dispatch versions of DSPS

are mostly on the time frame and the type of inputs used, but the algorithms remain the same.

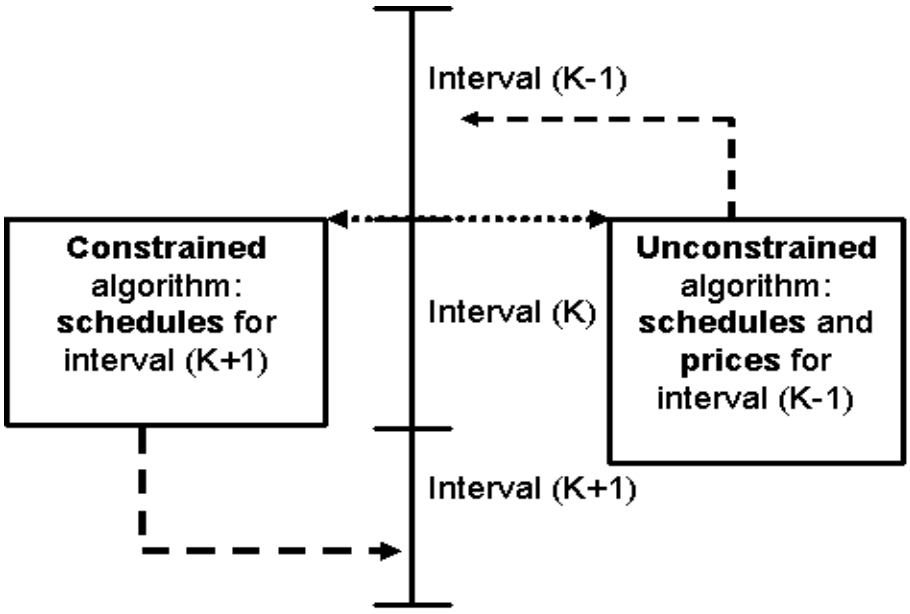

Figure 2.2: Unconstrained and constrained algorithm in real time.

2.3.2 Clearing Energy and Operating Reserves Markets

Pre-dispatch

The IESO forecasts the aggregate non-dispatchable Ontario demand and estimates the amount of generation capacity available from non-dispatchable generators for the dispatch hour. Recall that non-dispatchable loads and generators consume/generate the amount of energy they need/can regardless the market price. Therefore, the predicted amount of price-taker demand is considered as an energy buying bid at the Maximum Market Clearing Price (MMCP), and the aggregated predicted of non-dispatchable generation capacity available is considered as an energy sell at the -MMCP. The MMCP is currently $2000/MWh.

All price-sorted energy buying bids from entities within Ontario plus all exports are stacked in decreasing order, and all price-sorted supply bids from generators inside Ontario plus all import bids are stacked in increasing order. Operating reserve offers from inside Ontario and boundary entities for the three reserve classes are also stacked in increasing order, and the operating reserves requirements are specified by the IESO for each hour. The point of intersection between energy supply and demand bid stacks, while honoring all applicable constraints, determines the uniform Ontario energy MCP (see Fig. 2.3 for a simple illustration). The projected price of each class of operating reserves are also calculated in a similar manner, except that a 'single auction' market structure [54] is used (see Fig. 2.4). Energy and operating reserves MCPs are calculated jointly and the algorithm determines the best trade-off between energy and operating reserves. The calculated MCPs apply to all the twelve 5-minute intervals of the dispatch hour. It should be noted that for import and export bids, the intertie physical capacity limits as well as the Net Interchange Schedule Limit (NISL) (described in Section 2.3.3) are considered in the algorithm.

Real-time

The dispatch version of the unconstrained algorithm is run to determine the Ontario energy and operating reserves MCPs in real-time. This version is basically the same as the pre-dispatch unconstrained algorithm, except for the differences in inputs and time horizon. For example, the import/export quantities for energy and operating reserves cleared in the one-hour ahead pre-dispatch run are assumed constant and treated as supply/demand bids with the prices of -MMCP/MMCP respectively. Furthermore, actual metered non-dispatchable "primary" demand, as well as the system losses for the previous interval is used as an energy bid with the price of MMCP (see Fig. 2.5). Also, the non-dispatchable generators' capacity forecast is assumed as a supply bid with the price of -MMCP, similar to the pre-dispatch run.

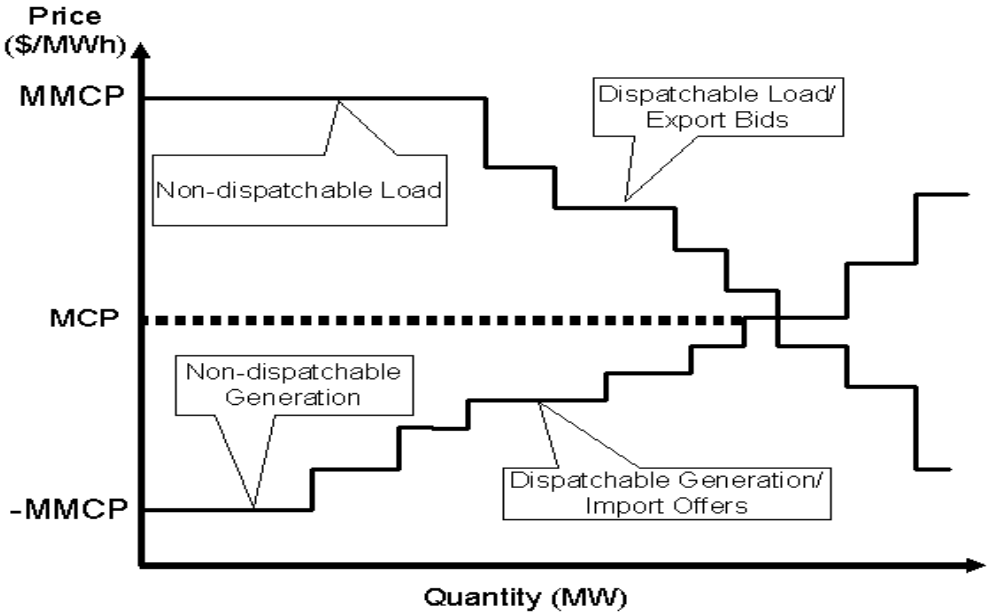

Figure 2.3: Energy MCP in pre-dispatch.

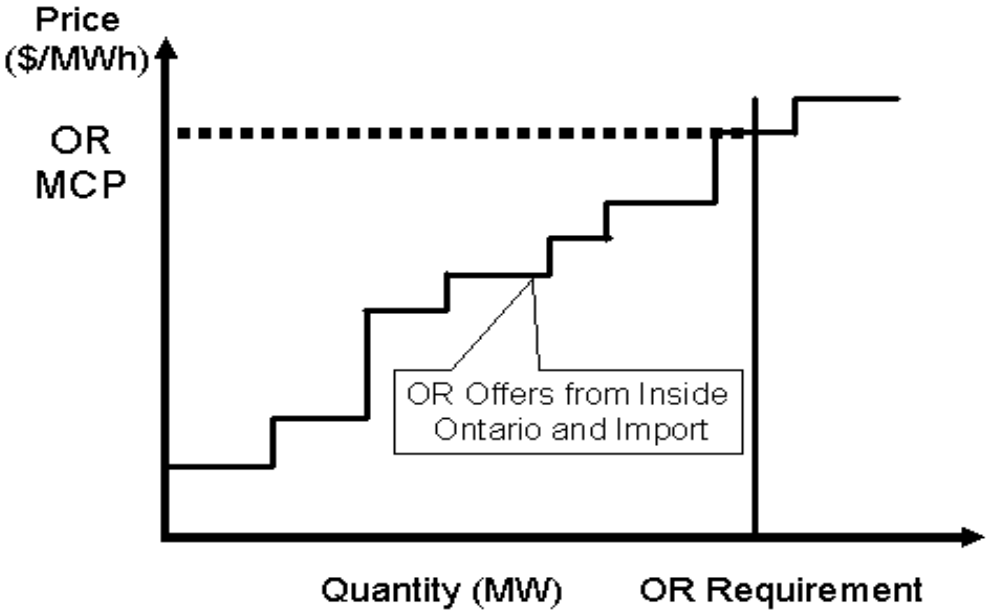

Figure 2.4: Determining operating reserves (OR) MCP for a specific class of operating reserves in pre-dispatch.

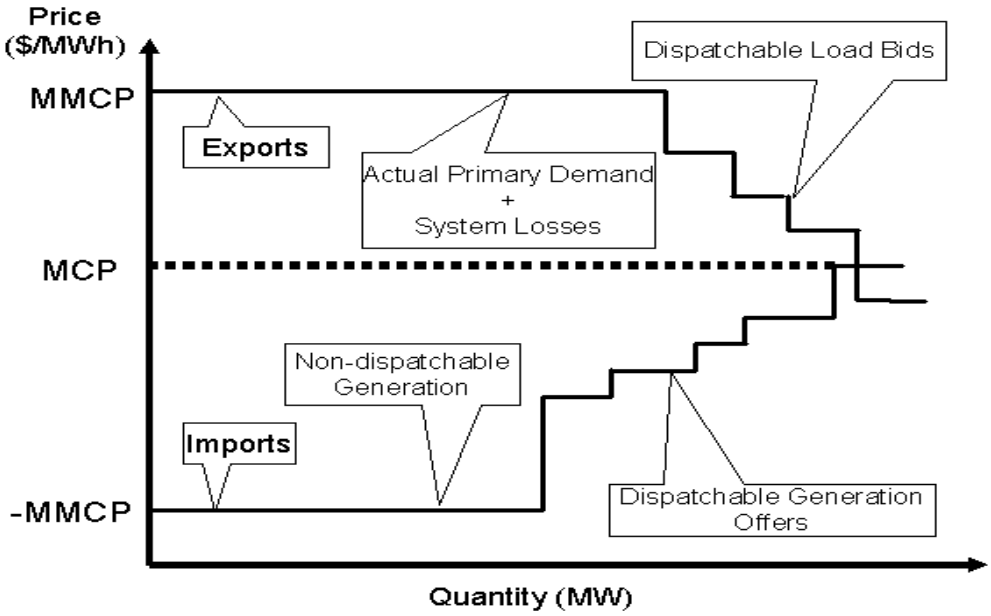

Figure 2.5: Energy MCP in real-time.

2.3.3 Inter-jurisdictional Energy Trading

The Ontario electricity market is interconnected with the New York electricity market, and Quebec, Michigan, Manitoba, and Minnesota control areas. The last three control areas are now part of the Midwest market. The New York electricity market is also interconnected with the PJM and New England electricity markets, and New England and PJM trade energy with Quebec and Michigan (see Fig. 2.6). Energy transactions take place among all these interconnected control areas.

Imports and exports to and from Ontario are treated in the same manner as the local supply and demand in many aspects. However, there are two major exceptions: first, as previously described in Section 2.3.2, imports and exports are scheduled in the 1-hour-ahead pre-dispatch run and they are considered as constant supply offers and demand bids in real-time; second, physical intertie limitations, as well as the NISL are honored by the DSPS when scheduling the imports and exports.

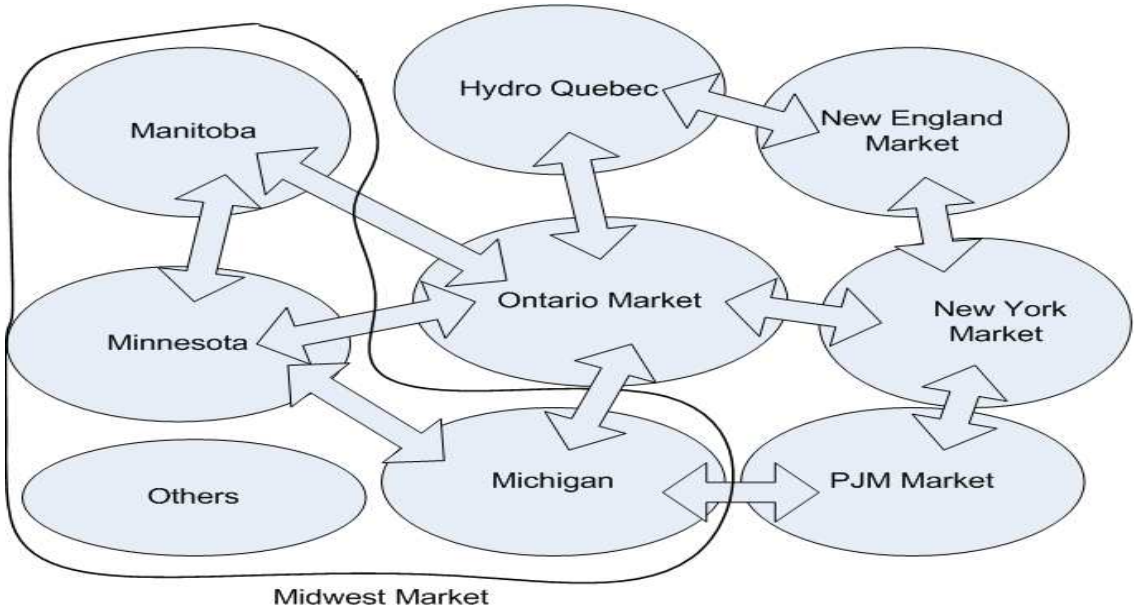

Figure 2.6: Ontario's interconnections with other areas.

Net Interchange Schedule Limit

Sharp changes in import/export schedules during consecutive hours can expose the IESO-controlled grid to reliability risks. To prevent this possibility, the Net Interchange Schedule (NIS) is defined as the total imports minus total exports, and the change in NIS across two consecutive hours is limited to 700 MW. This limitation is referred to as the NISL and is automatically respected by the dispatch algorithm. Because of the NISL, there might be some uneconomical supply/demand bids scheduled (or economical supply/demand bids not being scheduled) which should not have been scheduled (should have been scheduled) in the absence of NISL. If there is insufficient import bids and export bids for the algorithm to come up with a feasible solution, the IESO asks importers and exporters to change their import/export bids.

Zonal MCPs for the Interties

In order to find the zonal MCPs, the DSPS passes the import/export bids to the Ontario bid stacks while honouring both physical capacity limits and the NISL. If all economic bids from an intertie can be used in the Ontario market without violating both limits, or if the economical bids cannot be used due to the NISL, there is no congestion in the intertie, and the zonal MCP is equal to the Ontario MCP; otherwise, the intertie is assumed congested and the marginal price of energy in the intertie zone is considered as the zonal MCP. For example, assume that the New York intertie physical limit for import is 1000 MW and there are 1500 MW of import bids, all with prices under $300/MWh, and the NISL is met; up to 1000 MW of bid blocks are being passed to the Ontario supply bid stack, and the Ontario MCP clears at $300/MWh (see Fig. 2.7). In this case, since the next MWh not scheduled due to intertie limit is valued at $100, the zonal MCP is $100.

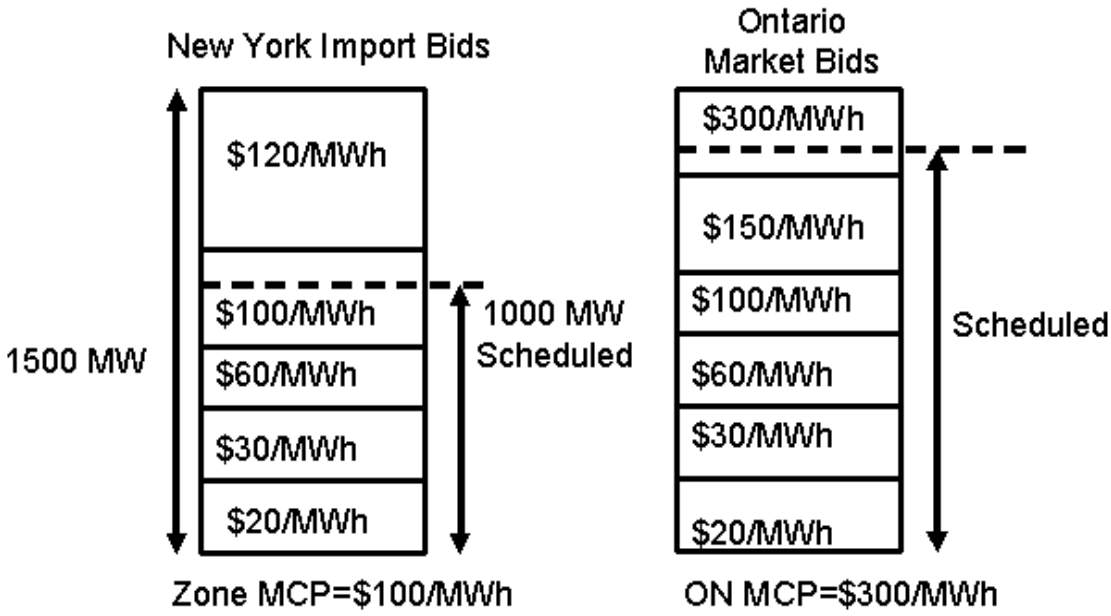

Figure 2.7: Determining zonal MCP

Intertie Congestion Price

The Intertie Congestion Price (ICP) is defined as the net costs incurred by the Ontario market because of congestion in an intertie. The ICP is calculated based on pre-dispatch zonal and Ontario MCPs and then used in real-time to determine the final zonal MCPs for financial settlements.

In pre-dispatch, the ICP is determined considering two scenarios: In the first scenario, where only the physical intertie capacity limits the scheduling of intertie offers/bids, the ICP is calculated as the difference between the 1-hour-ahead pre-dispatch zonal MCP and the pre-dispatch Ontario MCP, as follows:

$$\text{ICP} = \text{MCP}_{PD}^{Zone} - \text{MCP}_{PD}^{ON} \tag{2.2}$$

where PD indicates pre-dispatch, MCP^{Zone} is the zonal MCP and MCP^{ON} is the Ontario MCP. For the example in Section 2.3.3, if the intertie limit were 1001 MW, one more MWh would be supplied from the $100 import bid, instead of using the $300 Ontario bid; therefore, the congestion has cost the Ontario market $200/ MWh, i.e., ICP=-$200/MWh.

In the second scenario, both the physical intertie capacity and the NISL limit the scheduling of intertie offers/bids. It is to be noted that when the NISL is violated and the intertie is congested, relaxing the physical limit for an intertie leads to decreasing the physical limit for another intertie. In this scenario, the global cost of congestion from both the increase and decrease of the intertie capacities is calculated for determination of the intertie ICP. For example, let assume that an export congested intertie is relaxed by one MW and that this will save the market $300; at the same time, assume that decreasing another intertie limit by one MW, to meet the NISL, will cost the market $200. This is the total cost of the intertie congestion or, ICP=$100/MWh.

In real-time, the ICP is used to determine the final price for intertie transactions, as follows:

$$\text{MCP}_{RT}^{Zone} = \text{ICP} + \text{MCP}_{RT}^{ON} \tag{2.3}$$

where RT indicates real-time; a similar process is used to determine zonal MCP for 10N and 30N operating reserve classes. It should be noted that when an intertie is export congested, the exporters should pay a price higher than the Ontario MCP for the energy purchased from the Ontario market and hence ICP$>$0. On the other hand, when the intertie is import congested, the importers should receive a price lower than the Ontario MCP for the energy sold to the Ontario market and thus ICP$<$0.

The ICPs for energy and operating reserves for the 12 interties are zero most of the time. Table 2.1 shows the maximum and minimum energy ICPs for the period of January 12 to October 31, 2004 for the Manitoba, Michigan, Minnesota and New York interties.

Table 2.1: Energy ICPs ($/MWh) for January 12 to October 31, 2004.

	Manitoba	Michigan	Minnesota	New York
Max	4.54	177.62	55.8	167.33
Min	-56.83	-825	-91.99	-5.497

Intertie Offer Guarantee

To ensure adequate supply and encourage power imports to Ontario, the Intertie Offer Guarantee (IOG) mechanism is designed to pay the power importers at least the average price of their bid and prevent importers from incurring negative operating profit. One of the main assumptions in the Ontario market design is that supply and demand bids are based on marginal costs and marginal benefits. It means that if the MCP for a given interval is equal to a bid price, the operating profit of the respective market participant is zero and it would not be better off either scheduled or not[2]. Therefore, if under any

[2]This assumption is not necessarily true in an entirely competitive market in which the market participants design their bidding strategies to maximize profit.

circumstances the actual operating profit for a power importer is negative, the IOG payments return it to zero. Of course, this payment does not hedge the risk of having a lower operating profit in real-time than what was expected in pre-dispatch.

For example, assume the pre-dispatch Ontario MCP is equal to $25/MWh and the ICP is zero. The expected operating profit for a 100 MW power import at the bid price of $20/MWh for a given hour would be:

$$\text{OP=100MWh} \times (\$25/\text{MWh-}\$20/\text{MWh}) = \$500 \tag{2.4}$$

where OP is the operating profit. If in real-time the Ontario MCP turns out to be equal to $15/MWh, the actual operating profit would be:

$$\text{OP=100MWh} \times (\$15/\text{MWh-}\$20/\text{MWh}) = -\$500 \tag{2.5}$$

In this case, an IOG payment equal to $\$500$ will be made by the IESO to the power importer to return it to zero operating profit.

2.3.4 Congestion Management Settlement Credit

In the Ontario electricity market, real-time unconstrained prices and schedules are the basis of the financial settlements. If power system constraints force a market participant to generate/consume more/less than what it was supposed to in the unconstrained schedule, the market participant is treated as 'constrained on/off', and the Congestion Management Settlement Credit (CMSC) is used to provide the market participant with the same operating profit as it would gain in the absence of power system constraints.

For example, assume that generator A bids to generate 100 MW of energy at a price of $20/MWh for a given hour. Assume also that the Ontario MCP is equal to $30/MWh, and generator A is scheduled by the unconstrained algorithm for its entire bid for all 5-minute intervals of the hour. In this case the "operating profit" of generator A would

be:

$$\text{OP} = 100\text{MWh} \times (\$30/\text{MWh} - \$20/\text{MWh}) = \$1000 \tag{2.6}$$

However, if in the constrained algorithm run the generator is scheduled to inject only 50 MW at all 5-minute intervals, the actual operating profit would be:

$$\text{OP} = 50\text{MWh} \times (\$30/\text{MWh} - \$20/\text{MWh}) = \$500 \tag{2.7}$$

and hence the lost profit is \$500. In such a case, a \$500 CMSC payment will be given to generator A to bring it to the same level of operating profit as obtained from the unconstrained schedule.

When a market participant has gained some profit or prevented loss in real time as the result of being constraint on/off, it has to pay the extra operating profit to the IESO as CMSC. For example, an exporter is scheduled to export 100 MW during the next hour in the pre-dispatch unconstrained run, but network constraints force the exporter not to export at all. Let then assume the exporter's bid price is \$40/MWh, the ICP is equal to zero, and the real-time MCP is \$60/MWh for the all 5-minute intervals of the hour. Under these circumstances, the exporter would lose \$2000 if it were not constrained off. Therefore, the exporter has to pay \$2000 to the IESO as CMSC.

It was observed by the IESO that in the cases when the offer prices were negative, constrained-off payments could be very high and unjustifiable. Therefore, changes were made to the financial settlement algorithm in June 2003 to use \$0/MWh offer prices in such cases.

2.3.5 Contracted Ancillary Services

Ancillary services are required to ensure the reliability of the IESO-controlled grid. Ancillary services may be procured either through physical markets, such as operating reserves or through contracts with eligible service providers. The IESO procures five dif-

ferent ancillary services through contracts with various service providers in addition to the three classes of operating reserves discussed earlier; these are:

- Regulation/Automatic Generation Control Service: The IESO contracts with eligible generators to provide regulation service for the period beginning May 1 of each year to April 30 of the following year. Minimum requirements are calculated by the IESO and control signals are sent to the generators under contract to raise or lower their output as required.

- Reactive Support and Voltage Control: Reactive support and voltage control is contracted to ensure that the IESO is able to maintain the voltage level of its grid within acceptable limits. Generation facilities are the major provider of this service in Ontario.

- Black Start Service: Black start service is contracted to meet the requirements of restoring Ontario's power system after a major contingency. Generators that wish to provide this service must meet specific requirements determined by the IESO.

- Emergency Demand Response Load: Emergency demand response loads are the loads that can be called upon by the IESO to cut their demand on short notice in order to maintain the reliability of the IESO-controlled grid; this service is envisaged for emergency operating conditions.

- Reliability Must-Run Resources: Whenever sufficient resources to provide physical services in a reliable way are not available, the IESO may need to call registered facilities, excluding non-dispatchable loads, to maintain the reliability of the grid through reliability must-run resources contracts.

2.3.6 Market Uplift

The Ontario electricity market has been designed so that the consumers of electricity pay for all costs associated with operating the market in a reliable way. The operating costs are categorized under hourly and monthly components and recovered through market uplift. The market uplift is collected from the loads based on their share of the total demand. Congestion management costs, operating reserve costs and the costs associated with system losses are the hourly components of the market uplift. However, other components of the market uplift, including contracted ancillary services, IESO administration fees and miscellaneous charges, are calculated monthly. Some costs are regulated by the Ontario energy authorities and have a fixed price per MWh; for example, the IESO administration fees are $0.909/MWh. The market uplift appears in the customers' monthly invoice under separate charges.

2.3.7 Market Data

The Ontario IESO publishes two sets of system operation data prior to real-time dispatch of energy. The first set consists of conventional forecasts for some of the market variables, and is published as the "System Status Reports" (SSR). The SSR provides forecasts for Ontario demand and supply, energy imports, and capacity excess or shortfall. The SSR also contains total planned transmission and generation outages and other market advisory notices. The SSR is released for each day at least 24 hours in advance, and it is updated in case of any change in the system status or forecasts. The second set is referred to as the "Pre-Dispatch Reports" (PDR) and it provides the market participants with simulation-based forecasts of market outcomes, generated by the pre-dispatch run of the market, as explained in Section 2.3.1.

2.4 Programs to Improve Market Operation

Subsequent to the opening of the Ontario electricity market, several programs have been introduced by the IESO in order to improve reliability, efficiency, and transparency of the Ontario electricity market. These are briefly discussed in this section.

2.4.1 Hour-Ahead Dispatchable Load Program

Most of the loads in Ontario are non-dispatchable and therefore they do not respond to high market prices in real-time. The Hour-Ahead Dispatchable Load (HADL) program was launched in June 2003 for three main reasons: First, to make non-dispatchable loads more price-responsive; second, to allow the IESO to include future load reductions in the scheduling process; and third, to encourage load curtailment during peak operating hours.

The non-dispatchable loads would have an upper limit on the energy costs associated with their production process in most cases. If electricity price exceeds a specific upper cap, the load would choose to shut down its production. Non-dispatchable loads who wish to participate in the HADL program offer their price cap to the IESO and the quantity of demand that would be curtailed. If the 3-hour-ahead PDP is higher than the price cap offered by the load, the IESO will send dispatching instructions to the load to reduce its demand by the amount of its HADL offer. If the real-time HOEP turns out to be equal or more than the loads price cap, there will not be any payment. However, if the real-time HOEP is lower than the load's price cap, the load would have been better off to operate than shutting down its processes. In this situation, there would be a lost operating profit and the Hour-Ahead Dispatchable Load Offer Guarantee (HADLOG) is payable to the load to bring it to the same operating profit as it would have been gained when operating. The HADLOG is calculated as follows:

$$\text{HADLOG} = \max\{0, (\text{PC-HOEP}) \times \text{Q}\} \tag{2.8}$$

where Q is the quantity of demand that is offered to be cut, and PC is the load price cap. For example, Load A bids to cut 100 MW of its demand if the 3-hour-ahead PDP for energy is more than \$45/MWh. If in the 3-hour-ahead pre-dispatch run the energy price for the dispatch hour clears at \$50/MWh, dispatch commands are sent to Load A by the IESO to cut its load by 100 MW. If in real-time, the HOEP clears at \$40/MWh, an HADLOG payment of \$500 will be credited to Load A by the IESO in this case, as per equation (2.8).

2.4.2 The Spare Generation On-Line

Fossil-based generation units usually require a long and expensive start-up process, and thus they require a reasonably long operation period in order to recover the start-up costs. During the off-peak periods, these units are exposed to the risk of not being scheduled for a long enough period, and hence, they may decide not to put bids for the risky off-peak periods. On the other hand, if during the off-peak period, a large decrease in supply or increase in demand occurs, the IESO has to buy power from more expensive units or import expensive power; these lead to unusually high price spikes.

To increase the reliability of the IESO-controlled grid and to reduce price volatility, the IESO launched the Spare Generation On-Line (SGOL) program in August 2003, which offers eligible generators a guarantee of their start-up costs. Eligible generators submit their minimum loading point, minimum up-time and combined guaranteed costs to the IESO. If an eligible generator registered in the SGOL program submits a supply bid and is scheduled to run but the revenue earning over its minimum up-time is lower than its combined guaranteed costs, it will receive compensation from the IESO to cover its minimum combined guaranteed costs. The IESO recovers the SGOL payments through monthly uplift charges to loads.

2.4.3 Control Actions Operating Reserves

Under the Ontario market rules, the IESO is allowed to use out-of-market control actions when there no sufficient operating reserve offered in the market. These control actions include a 3% voltage reduction, a 5% voltage reduction, and a reduction of 30R requirements. In the initial Ontario market design, the market operator manually put in place these actions to maintain system reliability in stressed situations. This was obviously an out-of-market action and there was no cost associated with it to the market participants. This 'free' service could affect integrity of the price signals sent to the supply side, putting unrealistic downward pressure on the HOEP. Furthermore, out-of-market control actions have been recognized as one of the main sources of discrepancy between pre-dispatch and real-time prices.

To mitigate potential implications of the 'free' out-of-market control actions, the Control Action Operating Reserve (CAOR) was introduced in the market in August 2003. The first 200 MW CAOR was priced at $30/MWh as 10N operating reserve, and at $30.1/MWh as a reduction in 30R operating reserve requirements. In October 2003, an additional 200 MW CAOR was implemented in the market with the same pricing scheme. This 400 MW CAOR resulted in a significant reduction in the rate of out-of-market control actions. An additional 400 MW CAOR was later brought into the market in November 2005 at the price of $75/MWh for the first 200 MW and $100/MWh for the next 200 MW. Although it is expected that the CAOR result in slightly higher HOEPs, it would provide more realistic price signals during stressed conditions.

2.4.4 Multi-Interval Optimization Project

The MIO project was implemented in two stages. The initial stage was implemented in March 2004, by which a change was made to the DSPS to recognize 'effective unit ramp rates'. Before this stage, the DSPS assumed that generators can only operate under their

offered ramp rates. If a unit could not reach the dispatched level for a specific interval for any reason, dispatch instructions for next intervals could be undesirably affected; this problem is referred to as the 'stutter step' by the Ontario market participants. On the other hand, it was observed that some non-quick start generators can ramp up higher than their offered ramp rates for a short period of time. Therefore, in order to prevent some undesirable dispatch instructions, the DSPS was modified to use effective unit ramp rate, which is the lesser of the offered ramp rate for the interval multiplied by 1.2, and the maximum registered ramp rate. For example, if the maximum registered ramp rate of a facility is 4 MW/min, and the offered ramp rate for a given interval is 2.0 MW/min, the effective ramp rate that is used by the DSPS is 2.4 MW/min.

In the initial DSPS, dispatch instructions were derived for each interval independently. This caused some dispatching difficulties because the IESO had to dispatch generators on and off to maintain system reliability. However, frequent ramp up and down instructions are costly. In order to address this issue of dispatch volatility, the second stage of MIO was implemented in June 2004 [55], through which dispatch instructions for a given interval is calculated considering four other advisory intervals. These four intervals are selected out of a rolling 11-interval 'study period', based on some pre-defined selection criteria. These criteria are designed with the intention of providing the most efficient optimal solution, as well as providing the unit operators with an insight into the upcoming operating instructions. The four advisory intervals are not necessarily the same for each study period, and unit operators are provided with advisory dispatch targets for these intervals. The new MIO algorithm is further expected to improve system reliability, market efficiency, and market transparency.

2.4.5 Demand Response Programs

Emergency Demand Response Program

The Emergency Demand Response Program (EDRP), announced in June 2002, is intended to enhance system reliability by providing the IESO with a control action option prior to non-dispatchable load shedding. Terms and conditions of the EDRP are agreed upon by the IESO and the interested market participant through a 18-month contract. In case the IESO anticipates an emergency situation, it will give the EDRP participants a notice indicating the possibility of EDRP activation. EDRP participants are required to inform the IESO whether they intend to curtail their load. If EDRP participants reduce their demand in practice, they will receive financial compensation for the costs they incurred in responding to the IESO's request, based on the contract rates. Of the several occasions that the EDRP participants were given notice for EDRP activation (e.g., during summer 2003, winter 2004, and summer 2005), there was only one actual load reduction.

Emergency Load Reduction Program

In view of the EDRP experience and feedback from stakeholders, and to address the reliability concerns arisen from the shortage of supply during summer 2005, the Emergency Load Reduction Program (ELRP) was approved for launching on June 15, 2006. The ELRP is intended to provide Ontario market participants with an opportunity to improve the reliability of the electricity grid during stressed system conditions, particularly in the summer. If the program can attract a dependable amount of the demand side, it will enable the IESO to reduce usage of other more costly control actions, such as emergency energy purchases and voltage reductions. In fact, the ELRP is part of the IESO Emergency Operating State Control Action (EOSCA) list prior to public appeals for energy conservation, 3% and 5% voltage reductions, emergency energy purchases, and EDRP.

The ELRP is an ongoing program which will be activated only on Mondays to Fridays

from 8 AM to 8 PM, and market participants with capability of reducing their consumption for at least 1 MW during a minimum 2-hour time window can participate. The ELRP will be implemented in three steps. In the notification step, market participants will be informed that the ELRP will be implemented for a given day. The notice may be issued a day ahead or on the dispatch day, and can be for any number of hours within the program's window. In the offering step, the interested market participants notify the IESO by submitting their MW offer of load reduction. Although participating in the ELRP is not mandatory, a market participant is committed to reduce its specified load upon submitting an ELRP offer to the IESO. In the activation step, the IESO contacts ELRP participants to reduce their consumption, and non-compliance penalties may apply in case of under performance of greater than 20%.

ELRP participants will receive two types of financial compensations. A standby fee of $15 per MW per hour will be paid for participating in the program up until the activation hour. Upon activation, the participants will receive a load reduction payment based on the greater of the HOEP and the applicable following options:

- $400/MWh, for 2 hours of consecutive load reduction,
- $500/MWh, for 3 hours of consecutive load reduction,
- $600/MWh, for 4 hours of consecutive load reduction.

Load reduction for a period longer than 4 hours is not considered in the program. Actual load reductions, measured and verified by the IESO, are the basis for payments.

2.4.6 Mitigating OPG's Market Power

In order to limit OPG's market power, the Market Power Mitigation Agreement (MPMA) was put in place by the Ontario government before opening the market in May 2002. According to the MPMA, OPG had to pay the IESO a rebate if the HOEP exceeded

$38/MWh. However, it was observed that the MPMA seriously affected OPG's efficiency and led to financial problems; it cost OPG about $100 million per month, as OPG was not able to recover the overall costs of producing electricity.

The Electricity Act of 2004 replaced the MPMA with a new plan that sets capped prices and revenue for most of the OPG's generation facilities. Effective April 1, 2005, generation from OPG's base load hydroelectric and nuclear facilities, referred to as regulated assets, was capped at $33/MWh, and $49.5/MWh, respectively; these regulated assets represent about 40% of the Ontario's total generation capacity. Furthermore, OPG's revenues from about 85% of its unregulated assets, i.e., non-base load hydroelectric, coal and gas-fired stations, were set at an upper limit of $47/MWh; OPG's unregulated assets represent about 33% of the total generation capacity in Ontario. Under this new pricing regime, most of the demand side participants are eligible for two types of rebates. These rebates are referred to as the OPG Rebate and the Global Adjustment Rebate, and are calculated based on the HOEP and the mentioned limits on OPG assets. The limits on prices and revenues are subjected to change each year and are temporarily in effect until Ontario government develops a mechanism for pricing OPG's output, no later than March 2008.

2.4.7 Day-Ahead Commitment Process

In late 2003, the Day-ahead Market Working Group was established by the IESO in order to assess feasibility and features of a day-ahead market in Ontario. The group proposed a comprehensive day-ahead market with nodal pricing mechanism; however, it was finally not implemented by the IESO because of various political, economical, regulatory, and design issues raised by stakeholders. Instead of the comprehensive day-ahead market, a Day-Ahead Commitment Process (DACP) with reliability guarantees was approved by the IESO Board of Directors in September 2005.

The DACP is intended to improve system reliability by providing the supply side

with financial incentives, 24 hours before real-time dispatch of energy. The fundamental targets of designing the DACP are to address frequent real-time failure of import transactions, and to optimally manage next-day available energy resources. The DACP was launched on May 31, 2006. The IESO will evaluate effectiveness of the DACP based on its impact on system reliability, import failures reduction, and market uplifts, and will decide whether to continue with the DACP after November 2006.

The DACP is aimed to provide the IESO with a reliable anticipation of next day's available supply. For this purpose, dispatchable generators/loads who intend to participate in next day's real-time market must submit their operational data to the IESO by 11 AM on pre-dispatch day. Dispatchable facilities are also required to submit an Availability Deceleration Envelope (ADE). The ADE specifies the hours, energy, and capacity limits within which a dispatchable facility intends to operate during real-time. Although the dispatchable facilities are allowed to change offered prices, quantity of bids have to remain within the limits specified in the ADE. Importers are not obliged to submit import data into the DACP; however, they must do so in order to be qualified for the DACP financial incentives. The importer participating in the DACP will have to pay a day-ahead import failure charge if they do not follow their DACP obligations. It is to be noted that under the pre-DACP data submission rules, market participants were allowed to make any change to their submitted data, up to two hours before real-time.

The DACP is also designed to help dispatchable generators and importers to manage the financial risks associated with supplying energy into the Ontario grid. A Day-Ahead Generation Cost Guarantee (DAGCG) is offered to dispatchable generators to ensure they recover certain combined costs if they have not recovered their costs through market revenues. The DAGCG is the day-ahead version of the SGOL, except it also covers eligible maintenance and operation costs. A Day-Ahead Intertie Offer Guarantee (DAIOG) is also offered to imports to guarantee their 'as-offered' costs, and is basically the day-ahead version of the IOG. Aside from some minor differences in the way the DAGCG and the DAIOG are calculated and financially settled, they are basically the same as the

real-time SGOL and the IOG payments, respectively. If a market participant is entitled to both day ahead (DAGCG and DAIOG) and real-time (SGOL and IOG) credits, only the higher one will be credited. The costs of paying the DAGG and DAIOG will be recovered through market uplifts.

The existing pre-dispatch algorithm, discussed earlier, is the calculation engine for the DACP. The first four runs of the pre-dispatch algorithm after 11 AM on pre-dispatch day are used to generate DACP schedules. The first three runs are to generate initial schedules and necessary reliability refinements are carried out. Also, energy limited generators can change their submitted data during the first three runs. The 4^{th} run starts at 14:00 and produces final schedules, referred to as the Pre-dispatch of Record. The Pre-dispatch of Record is the basis of financial guarantees, and may be rejected by the committed participants by hour 15:15.

2.5 Analysis of Market Outcomes and Discussion

The main market outcomes, namely pre-dispatch and real-time energy prices, operating reserve prices, and market demand are studied in this chapter for the period May 1, 2002 to April 30, 2006.

2.5.1 Energy Price

The monthly demand-weighted averages of the HOEP for the period of May 1, 2002 to April 30, 2006 are shown in Figure 2.8. The Ontario market experienced a high record demand during summer 2002, which coincided with some supply limitations. Early in winter 2003, extremely cold weather resulted in an increase in demand as well. Furthermore, during winter 2003, marginal cost of supplying electricity into the grid soared as a result of unusually high natural gas prices. Also, some of the gas/oil stations were not

available to the market operator as they experienced difficulties in their fuel procurement systems in this period. Thus, the high demand and limited supply availability resulted in high energy prices during summer 2002 and winter 2003.

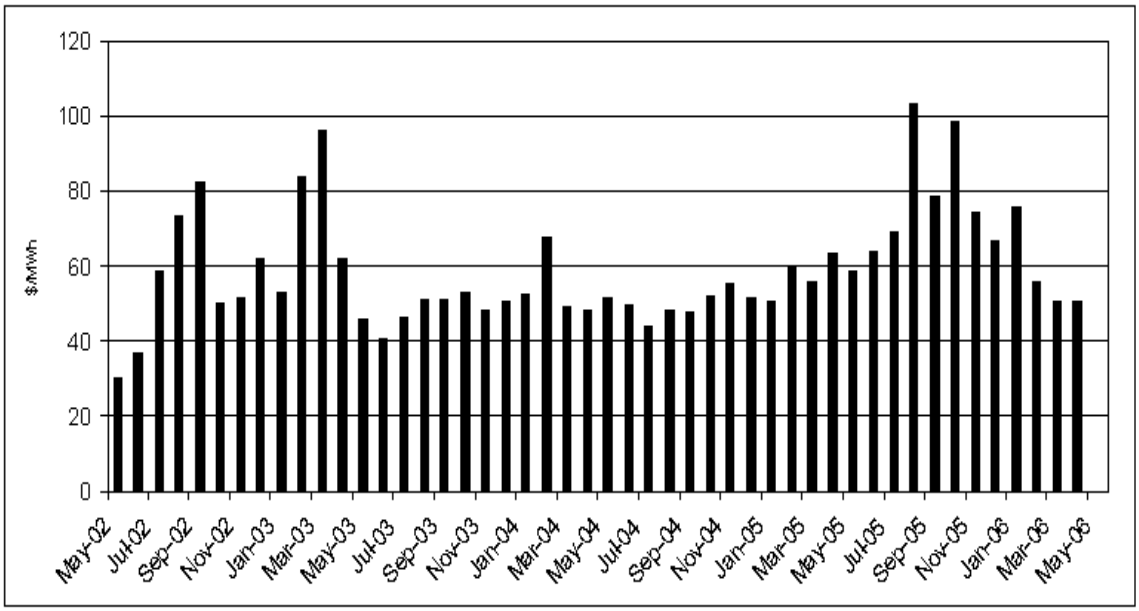

Figure 2.8: Monthly weighted HOEP averages, 2002-2006

The average HOEPs remained steady until summer 2005, when high temperature and humidity levels led to a record peak demand. On the other hand, reliance on gas-fired stations during this period was increased as a result of reduced hydroelectric outputs and the shutdown of a large coal-fired station. Furthermore, since natural gas prices were high during this period, some generators preferred to sell their gas contracts in the natural gas spot markets rather than producing electricity. Hence, the costs of producing electricity in Ontario as well as in the neighboring areas increased, resulting high and volatile HOEPs.

Energy prices in Ontario on average have been in the same order as the wholesale energy prices in New York and PJM. However, New England prices have been persistently higher than Ontario prices. Michigan, Manitoba, and Minnesota control areas have

joined the Mid-West electricity market, opened in April 2005, with the energy prices being always lower than the HOEP. It is usually expected that in a fully competitive environment, arbitrage results in elimination of the price differences in the neighboring areas; however, transmission line constraints, different scheduling protocols, and physical power flow rules have limited the ability of power traders to arbitrage away the price differences.

The HOEP has been in general highly volatile, varying from as low as $4/MWh to as high as $1,028.4/MWh during the 4-year period. About 82% of the hours, the HOEP has remained in the range of $20/MWh to $80/MWh, and for about 15% of the hours, the HOEP has varied in the $80/MWh-$200/MWh range. Furthermore, during the first four years, there have been 196 hours at which the HOEP has exceeded $200/MWh. Finally, for about 2% of the hours, prices have been low, in the range of $4/MWh to $20/MWh.

2.5.2 Demand

During the period May 1, 2002 to April 30, 2006, the highest recorded demand in Ontario was 26,160 MW, attained in July 2005. The monthly maximums and averages of the Ontario demand over the 4-year period are displayed in Figure 2.9. It can be observed from Figures 2.8 and 2.9 that demand is the main driver of energy price; however, unusually high energy prices occurred when peak demand coincided with supply limitations. For instance, while maximum demand in winter 2004 was slightly higher than that of winter 2003, energy prices were higher in winter 2003 as a result of supply limitations.

2.5.3 Operating Reserve Prices

The yearly average prices of the three classes of operating reserves are displayed in Figure 2.10. Unlike the energy prices, operating reserves prices have declined over the 4-year period. It was observed that the operating reserve prices were as high as energy

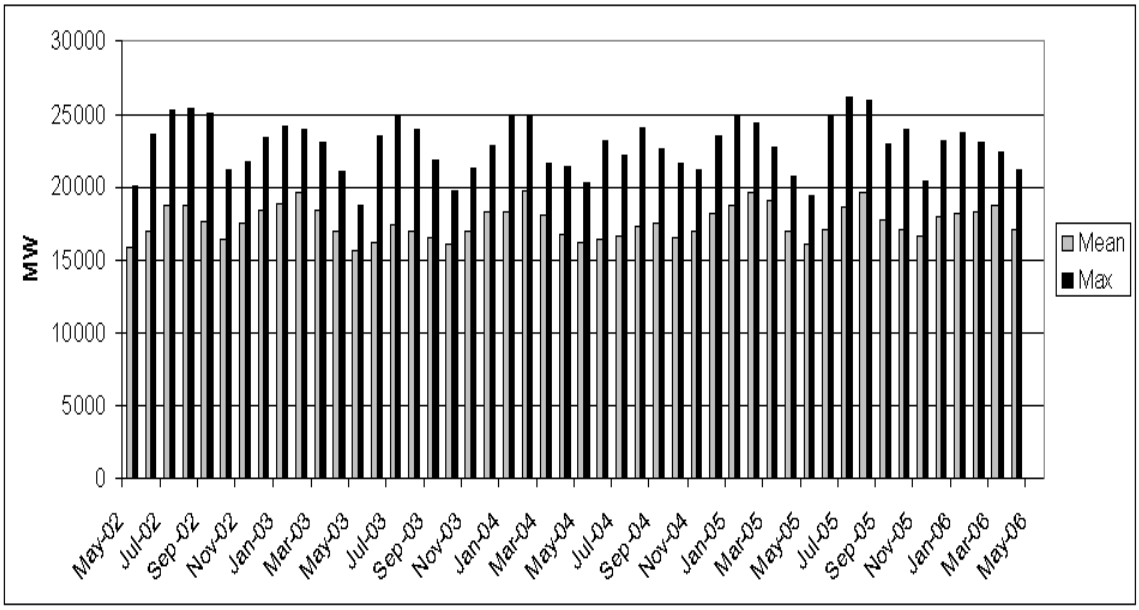

Figure 2.9: Mean and maximum Ontario demand.

prices on a few days during September 2003 because of a series of unusual events; these unusually high prices interrupted the reducing pattern of the 10N and 30R prices during the second year. Reduction in operating reserve prices can be attributed to the fact that about 600 MW of dispatchable load has emerged in the market. This group of loads are allowed to offer 10N and 30R into the market, resulting in a more competitive and lower 10N and 30R prices. It should be noted that despite the high energy prices during summer 2005, the 10S prices have continued to decline; these low 10S prices can be explained by the fact that the limited water supply due to drought during this period shifted some of the hydroelectric units from energy production to providing 10S reserve.

2.5.4 Discrepancy between the HOEP and the PDPs

The PDPs are generated based on the most recent available market information in order to provide the market participants with an estimate of the real-time HOEPs. However, it

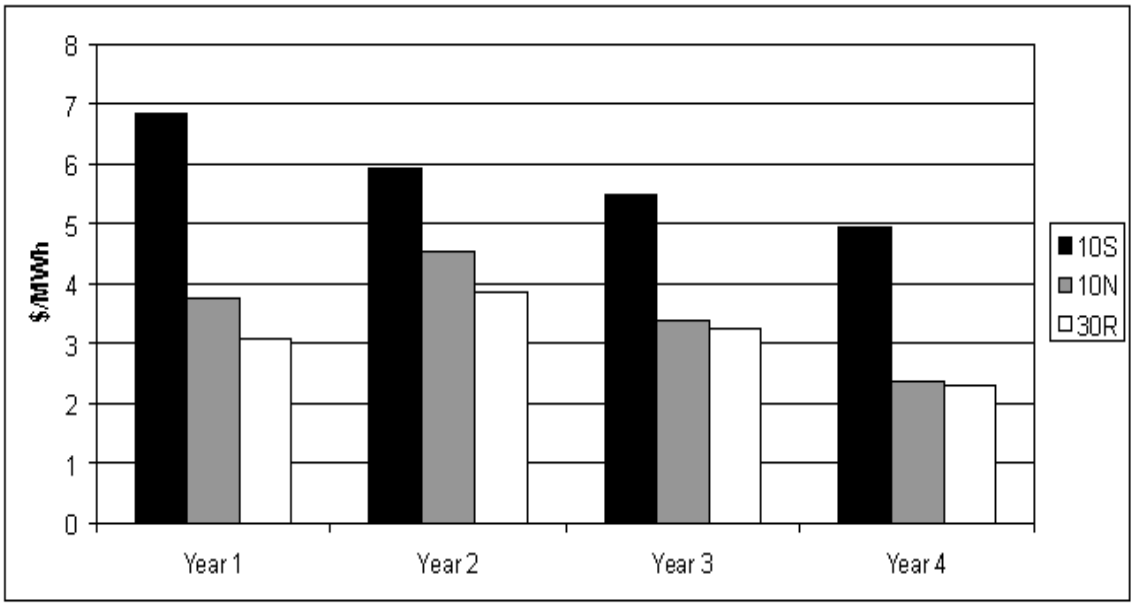

Figure 2.10: Yearly operating reserve (OR) price averages

has been consistently observed that there is a large discrepancy between the PDPs and the HOEP [23]. Let define the yearly Mean Absolute Percentage Error (MAPE) of the PDPs as:

$$\text{MAPE} = \frac{100}{N} \times \sum_{t=1}^{N} \frac{|\text{HOEP}_t - \text{PDP}_t|}{\text{HOEP}_t} \tag{2.9}$$

where HOEP_t and $\text{PDP}_{a,t}$ are the values of the HOEP and PDP for hour t, respectively, and N is the number of hours in a year. The yearly MAPEs of the 1-hour-ahead and 3-hour-ahead PDPs for the first four years of market operation are depicted in Figure 2.11. One can observe from this figure that the discrepancy between the HOEP and PDPs has declined, to some extent, through the first three years. Moreover, the highest discrepancies happened during the first year of market operation, mainly because of the volatile prices during summer 2002, and probably due to market immaturity. Note that the deviation of HOEP from the PDPs has increased in the 4^{th} year, which is because of the unstable and unusually high prices during summer 2005.

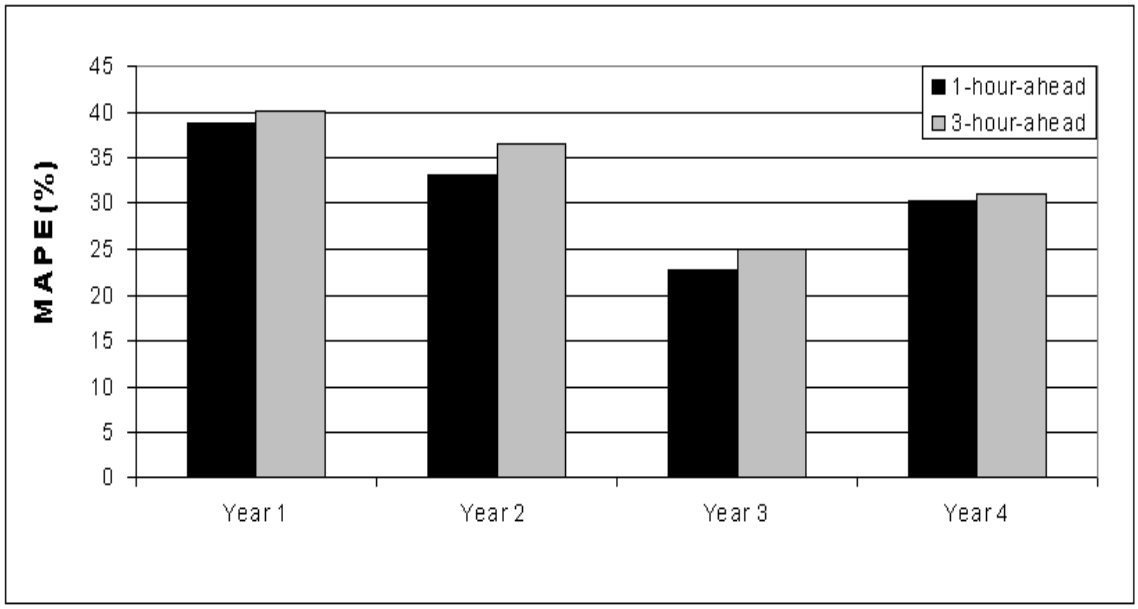

Figure 2.11: Yearly MAPE of the discrepancy between the HOEP and PDPs.

The high discrepancy between the HOEP and the PDPs can be explained by operational aspects of the Ontario market. The real-time nature of this market makes it vulnerable against unpredictable events. The generation offer curve in Ontario is "hockey-stick" shaped [23]. Consequently, demand over-forecasting, demand under-forecasting, errors in forecasting the output of self-scheduling generators, and import/export failures oblige the market operator to commit expensive units on the "blade" portion of the offer curve, or to de-commit some of the already committed units and move back on the "shaft" portion of the offer curve. This requirement puts upward or downward pressure on the HOEP, leading to price spikes and deviation of HOEP from the PDPs. Furthermore, out-of-market control actions affect the consistency between the real-time and pre-dispatch market clearing procedures, leading to deviation of the HOEP from the PDPs.

The improvements in consistency between the HOEP and PDPs can be attributed to changes and enhancements gradually being implemented in the market. Specifically, the processes of forecasting the load, as well as forecasting the output of self-scheduling

generators have been modified. Furthermore, 800 MW of out-of-market control actions are introduced in the market by the CAOR program.

Deviation of the HOEP from the PDPs has many implications, and seriously affects market efficiency. For example, analysis of PDP data for the first four years of market operation shows that for about 81% of the hours, real-time HOEPs have been less than the corresponding 1-hour-ahead PDPs. For such hours, eligible importers are entitled to IOG payments, recalling the fact that imports are scheduled based on 1-hour-ahead PDPs. Furthermore, too many imports are scheduled while the cheap local supply is dispatched off, and too few exports are scheduled while the demand side in the neighboring areas are willing to pay more for Ontario energy. Another example is the HADL program which is also designed based on the 3-hour-ahead PDPs. If the real-time HOEPs turn out to be lower than the 3-hour-ahead PDPs, which has been the case for about 79% of the hours for the first four years of the market operation, the HADL program participants may be eligible for HADLOG payments.

2.5.5 Effectiveness of the Market Improvement Programs

Although the programs implemented by the IESO to enhance market operation have positively affected market outcomes, the goals of the programs have not been entirely reached in some cases [23]. The following conclusions can be reached considering the publicly available data and reports:

- By using the MIO and looking-ahead scheduling, the use of out-of-market control actions is reduced. However, the problem of dispatch volatility, which was one of the main objectives to be addressed by the MIO program, still exists.
- The HADL program was implemented to boost load responsiveness to price signals. Only a total of 240 MW load participated in the program, and the participants have been scheduled for load reduction for 110 hours. Analyses of market data

have revealed that the overall benefit to the market form the HADL program has been minimal, given the high discrepancy between the HOEP and the 3-hour-ahead PDPs.

- The SGOL program was mainly designed to improve market reliability. While this objective has been met, market efficiency has been reduced by the payment of more than $33 million to eligible generators.

- There are only 600 MW of dispatchable loads biding into the market. These loads usually bid high prices compared to the normal range of the HOEP, and demand-side involvement in market enhancement programs has not been very significant. This high level of load inelasticity in Ontario affects market efficiency in general.

- The MPMA program was designed to improve market efficiency by mitigating OPG's obvious market power. However, it resulted in inefficient operation of OPG and subsidized electricity prices for the consumers. Considering the dominant share of OPG in Ontario supply, these side effects are against market efficiency and transparency. On the other hand, non-utility generators (NUGs) have been holding long-term power purchase agreements with the Ontario government that have excluded them from openly competing in the market. Also, the Ontario Power Authority has been assigned to manage generation and load management contracts with supply and demand side entities in order to ensure availability of reliable power for Ontario. These contracts are refereed to as the "Request for Proposal" (RFP) contracts. Under the RFP contracts, while the generators sell their output into the market, they will be provided with guaranteed revenue to ensure they can recover their costs. The generators will be financially settled based on the net revenue they received from the market, and the revenues agreed upon in the RFP contract. Given the NUG and RFP contracts, and the new capped prices and revenues over most of the OPG's output, only about 25% of the total Ontario generation capacity is fully open to compete in the market.

- Despite the reduction in the deviation of the HOEP from the PDPs, the discrepancy between the two and the overall price volatility is still high. The highly volatile pre-dispatch and real-time prices during the summer 2005 highlighted the limited effectiveness of the implemented market programs in maintaining consistency between the pre-dispatch and real-time prices.

2.6 Summary

This chapter presents a unique overview of the operation of the Ontario electricity market, along with an analytical discussion of the market's outcomes. The Ontario electricity market is the only real-time market in North America, and it is interconnected with the New England, New York, PJM, and Midwest competitive electricity markets, as well as Quebec and Manitoba regulated power markets. The physical system is not fully considered in the process of clearing the market prices, and a province-wide uniform price applies to all market participants. Most of the load in Ontario is not price-responsive, which has adversely affected the load management programs initiated by the Ontario IESO. Many programs are being implemented by the IESO to improve the market operation; however, some of the challenges behind the implementation of these programs have not been fully addressed yet.

Chapter 3

Forecasting the HOEP Using Linear Time Series Models[1]

3.1 Introduction

In Chapter 2, the operation of the Ontario electricity market is reviewed and the process of clearing energy prices in this market is discussed. Recall that the HOEP is the province-wide uniform price that applies to non-dispatchable market participants, mainly the demand-side BEMCs. Keeping in mind the findings and discussions of Chapter 2, an attempt is made in the present chapter to forecast the short-term behavior of the HOEP by employing the well-established linear time series models. Relevant data from Ontario and its neighboring electricity markets, namely the New York, New England, and PJM electricity markets, are investigated and a final set of explanatory variable candidates are selected. The focus of the variable selection procedure is on those variables that are available before real-time. This makes the selected variables capable of being used in practical price forecasting tools.

[1]Findings of this chapter have been accepted for publication in the *IEEE Transactions on Power Systems* [56].

The multivariate TF and DR models are employed to relate HOEP behavior to the selected explanatory variable candidates. Univariate ARIMA models are also developed for HOEP forecasting. The HOEP models are developed on the basis of two forecasting horizons, i.e., 3 hours and 24 hours, and forecasting performance of the multivariate models is compared with that of the univariate ARIMA models. Univariate ARIMA models are also developed for three day-ahead Locational Marginal Prices (LMPs) of Ontario's neighboring markets, and the accuracy of these models is compared with those for the HOEP.

3.2 Selecting Explanatory Variable Candidates

While demand has been the most commonly examined explanatory variable in the reported price forecasting studies (e.g., in [19, 14]), the present research evaluates a wide range of system information to develop price forecasting models. The explanatory variable candidates are selected from information publicly available before real-time, based on two main criteria. The first criterion is the consideration of implicit and/or explicit effects of the variables on the Ontario market clearing process. The second criterion is the consideration of linear correlations between current HOEP values and current and past values of the variables, given the linear nature of TF and DR models. These correlations are measured by the Cross Correlation Functions (CCF) [57], however, the linear correlation coefficients between current HOEP values and current values of each explanatory variable candidate, referred to as ρ here, are the only correlation coefficients discussed here.

The Ontario Market Surveillance Panel (MSP) reports [23] reveal that coal and gas fired generators are the main price setters in the Ontario electricity market. However, although fuel prices have shown to affect the long-term HOEP trends, no short-term relationship between the HOEP and fuel prices was found in [23]. Therefore, the present

study does not consider fuel prices among the explanatory variable candidates.

The ability of market participants to react to price forecasts depends on the forecasting horizon. For example, in the Ontario electricity market, dispatchable generators are restricted from changing their bids two hours before real-time [51], as discussed in Section 2.3.1. This requires that HOEP forecasts should be generated at least three hours before real-time so as to make them useful to this group of generators. On the other hand, when the forecasting horizon is long, say more than 24 hours, critical market information is either not available, or available but likely subject to significant changes. Thus, 3 hours and 24 hours are the reasonable HOEP forecasting horizons used here to which market participants can properly react.

As described in Section 2.3.7, The Ontario IESO publishes two sets of system operation data prior to real-time dispatch of energy, namely, the System Status Reports (SSR), and the Pre-Dispatch Reports (PDR). These data sets are publicly available on the IESO's web site at www.ieso.ca, and are mined here to explain HOEP behavior.

3.2.1 Explanatory Variables from the SSR

Demand Forecasts

Ontario demand is one of the main factors involved in the process of clearing the HOEP. The relationship between Ontario demand and the HOEP is presented in Figure 3.1-a. The corresponding linear correlation coefficient is 0.73. The most accurate forecast of Ontario demand available 24 hours before real-time is the SSR demand forecast.

Let define the annual MAPE of demand forecasts as:

$$\text{MAPE} = \frac{100}{N \times 24} \sum_{t=1}^{N \times 24} \frac{|\text{Demand}_{f,t} - \text{Demand}_{a,t}|}{\text{Demand}_{a,t}} \tag{3.1}$$

where $\text{Demand}_{f,t}$ and $\text{Demand}_{a,t}$ are the forecast and the actual values of demand at hour t, respectively, and N is the number of days in the studied year ($N = 366$ for 2004).

The 24-hour-ahead SSR demand forecasts have annual MAPEs of 2.1% and 4.8% for 2004, when compared with actual Ontario demand and actual market demand (demand plus exports and losses), respectively. Furthermore, these forecasts show a significant linear correlation with the HOEP, as shown in the scatter plot presented in Figure 3.1-b. The linear correlation coefficient between SSR demand forecasts and the HOEP is 0.68. Thus, the SSR demand forecast variable is included in the set of explanatory variable candidates.

Predicted Supply Cushion

The SSR energy supply forecast variable shows no meaningful linear correlation with the HOEP, as shown in Figure 3.2-a. Nevertheless, the concept of supply cushion (SC) [23] is used here, and is defined as follows:

$$\text{SC} = \frac{\text{EO-(TD+OR)}}{\text{TD+OR}} \times 100 \tag{3.2}$$

where EO is the actual energy offered, TD is the actual market demand, and OR is the operating reserve requirement. It was observed in [23] that price spikes are more likely when the SC is below 10%. In the present work, (3.2) is modified and actual quantities are substituted with respective forecasts from the SSR; the resulting SC is referred here to as the Predicted Supply Cushion (PSC). The PSC is found to be linearly correlated with the HOEP, as shown in Figure 3.2-b, with $\rho = -0.60$; hence, it is added to the set of explanatory variable candidates.

Planned Outages

Although the physical power system is not directly considered in the process of determining the HOEP, as discussed in Chapter 2, the physical system can influence the HOEP behavior indirectly. For example, outage of cheap generation facilities can result in higher energy prices, especially during low-demand hours [23]. However, the total

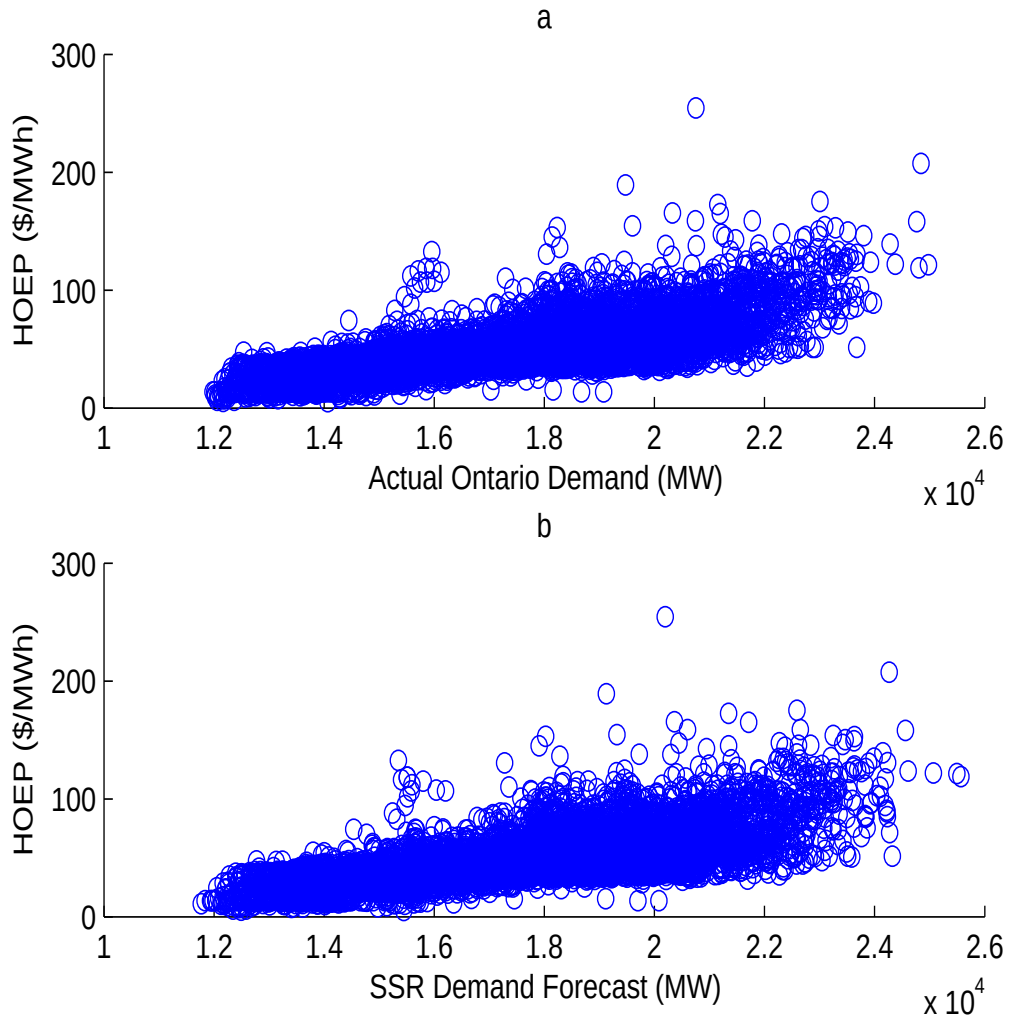

Figure 3.1: Relationship between: a) Ontario demand and the HOEP; b) SSR demand forecast and the HOEP.

outages reported in the SSR is the aggregation of various planned generation and transmission system outages, and this total is found to be not meaningfully correlated with the HOEP ($\rho = 0.18$). Hence, the SSR planed outages variable is not considered in the

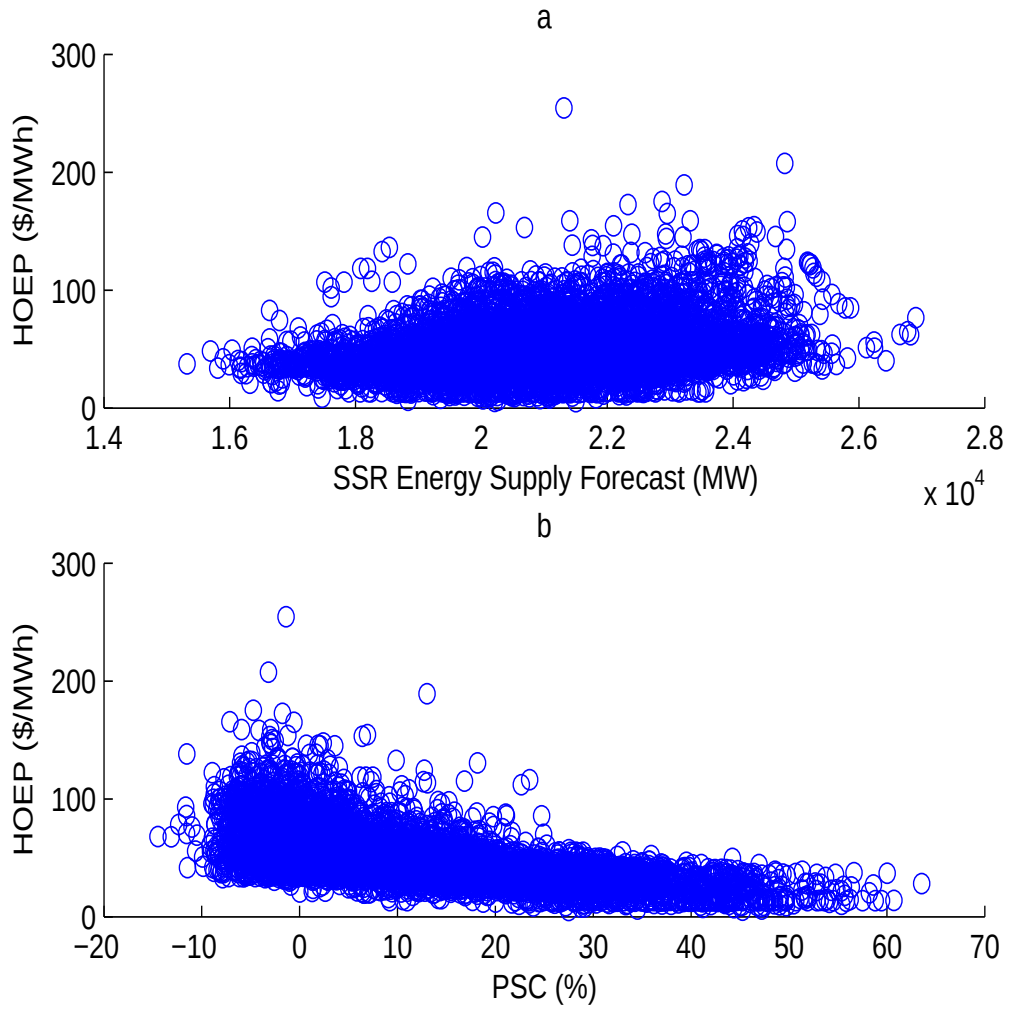

Figure 3.2: Relationship between: a) SSR supply forecast and the HOEP; b) PSC and the HOEP.

model building process.

Capacity Excess or Shortfall

The SSR capacity excess or shortfall variable is found to be linearly correlated with the HOEP, with a linear correlation coefficient of $\rho = -0.65$; hence, it is included in the set of explanatory variable candidates. However, it should be noted that when demand is low, capacity excess is high and vice versa, a fact confirmed by the high negative correlation between the SSR capacity excess or shortfall and demand ($\rho = -0.75$). Therefore, the SSR capacity excess or shortfall variable is highly collinear with the SSR demand forecast variable and should be included in the model only after the possible effects of demand have been modeled.

Imports

The SSR import forecast variable deviates significantly from actual values. Therefore, the import forecasts are not considered in the set of explanatory variable candidates. No export forecasts are published in the SSR.

3.2.2 Explanatory Variables from the PDR

Recall from Section 2.3 that the Ontario market clearing algorithm is run in pre-dispatch and real-time (dispatch). The pre-dispatch run provides the market participants with the "projected" schedules and prices, based on the most recent available market information. Outcomes of the pre-dispatch run are published by the IESO as the PDR for a variety of variables, including energy and operating reserves prices, total load, dispatchable load not served, system losses, and some of the system security constraints. From the PDR variables, the PDP and Pre-Dispatch Demand (PDD) variables carry the latest information about demand and price in the coming hours; hence, they are examined here for their role in improving accuracy of HOEP forecasting.

The PDP/PDD values that correspond to hour t and that are published k hours before real-time are called k-hour-ahead PDPs/PDDs [24]. Linear correlation coefficients between the HOEP and k-hour-ahead PDPs ($\rho_{\text{HOEP,PDP}}$), and between the HOEP and k-hour-ahead PDDs ($\rho_{\text{HOEP,PDD}}$), for $k = \{1, 2, 3, 24\}$, are presented in Table 3.1. Linear correlation coefficients between actual Ontario market demand and k-hour-ahead PDDs ($\rho_{\text{PDD,Demand}}$) are also presented in Table 3.1.

k-hour-ahead PDPs

It can be inferred from the correlation coefficients presented in Table 3.1 that when k is small, the k-hour-ahead PDPs are closer to the HOEP. Hence, considering k-hour-ahead PDPs as explanatory variables depends on the forecasting horizon. For 24-hour-ahead forecasting, the 24-hour-ahead PDP variable is clearly not useful and hence is not considered as an explanatory variable candidate. However, for shorter forecasting horizons, k-hour-ahead PDPs become more relevant; thus, in this work the 3-hour-ahead PDP variable is included in the set of explanatory variable candidates for 3-hour-ahead forecasting.

k-hour-ahead PDDs

It can be observed from Table 3.1 that the k-hour-ahead PDDs do not deviate significantly from actual market demand, thus these should be included in the set of explanatory variable candidates. Therefore, the 3-hour-ahead PDD variable is considered an explanatory variable candidate for 3-hour-ahead forecasting. However, since the accuracy levels of the 24-hour-ahead PDDs and the SSR demand forecasts are very close, only one of them, namely the SSR demand forecast, is included in the set of explanatory variable candidates. It is worth mentioning that the 24-hour-ahead PDD variable was also considered for model building, in lieu of the SSR demand forecast variable, but no significant difference in the overall performance of the developed models was observed.

Table 3.1: Correlation between HOEP, k-hour-ahead PDPs, and PDDs

k	24	3	2	1
$\rho_{\text{HOEP,PDP}}$	0.16	0.74	0.77	0.78
$\rho_{\text{HOEP,PDD}}$	0.62	0.63	0.63	0.64
$\rho_{\text{PDD,Demand}}$	0.97	0.98	0.98	0.98

3.2.3 Demand and Energy Price in the Neighboring Areas

As previously mentioned in Section 2.3.3, the Ontario electricity market is interconnected with the New York electricity market, and Quebec, Michigan, Manitoba, and Minnesota control areas. The last three control areas are now part of the Midwest market. The New York electricity market is also interconnected with the PJM and New England electricity markets, and New England and PJM trade energy with Quebec and Michigan (see Figure 2.6). With such a complex interconnection between neighboring areas, it is difficult to assess the effects of energy price and demand of the neighboring areas on the HOEP. Furthermore, lack of publicly available information on quantity and price of energy transactions between Ontario and Quebec, Michigan, Manitoba, and Minnesota constrained the author to consider only the effects of demand and price of the New York, New England and PJM electricity markets on the HOEP. The data employed are available online at www.nyiso.com, www.iso-ne.com, and www.pjm.com.

Demand

To evaluate the possible effects of New England and PJM market demands on HOEP, actual demand data from these markets are considered. But, since these data are not available before real-time, they cannot be considered in the final models even if they turn out to be significant. For the New York market, historical demand forecasts are available

and hence are used in this study.

Linear correlation coefficients between demand in the neighboring markets and demand and price in Ontario market are presented in Table 3.2. Climatic conditions that are similar across New York, Ontario, and New England could be a reason for the collinearity in demand between these markets. Consequently, demand collinearity could be the reason for high correlation between the HOEP and the New York and New England demands. On the other hand, the low correlation between the Ontario and PJM demands can be attributed to variations in the residential and industrial load distribution pattern across these markets, plus climatic differences between the two.

In this work, the New York and New England markets demands are considered as explanatory variable candidates. However, due to the collinearity between the Ontario demand and the other demands, they should be included in the model building process only after the effects of the Ontario demand on the HOEP are modeled. Given its small correlation with the HOEP, the PJM market demand is not considered an explanatory variable candidate.

It was also observed that actual quantities of power transactions through the Ontario-New York intertie had no meaningful correlation with demand or price in the neighboring markets. This lack of correlation is due to the fact that much of the overall transactions constitute power wheeling transactions from different parties taking place through this intertie.

Price

Only day-ahead prices in the neighboring markets are examined for their possible effects on the HOEP, because they are known before real-time. Note that the main components of the costs of any energy transactions between Ontario and the neighboring markets are the HOEP and the LMPs at the pricing points in those markets involved in the trade; thus, only these three LMPs are studied here. These LMPs are denoted as LMP_{NYON} for the

Table 3.2: Correlation between demand in the neighboring markets, Ontario price and Ontario demand, year 2004

	New York demand	New England demand	PJM demand
HOEP	0.54	0.63	0.37
Ontario Demand	0.83	0.89	0.52

New York to Ontario interface in the New York market, LMP_{NENY} for the New England to New York interface in the New England market, and $\text{LMP}_{\text{PJMON}}$ for the PJM to Ontario interface in the PJM market.

The HOEP is correlated to LMP_{NYON}, LMP_{NENY} and $\text{LMP}_{\text{PJMON}}$, with ρ values of 0.69, 0.67, and 0.67, respectively; hence, they are considered as explanatory variable candidates. Since market prices are influenced mainly by demand, the high correlations may be due to the similar demand patterns in the neighboring areas. Therefore, the mentioned LMPs need to be included in the HOEP models only after the possible effects of demand in Ontario and other markets on the HOEP are properly modeled.

For simplicity, the final and total set of explanatory variable candidates are denoted by x_1 to x_{10}, and are summarized in Table 3.3.

3.3 Review of Time Series Models

3.3.1 ARIMA Model

Let denote the equally sequenced values of a stationary stochastic process z by $z_t, z_{t-1}, ...$ An Auto Regressive Moving Average model ARMA(p, q) for this process can be ex-

Table 3.3: The final explanatory variable candidates

Variable	ρ
x_1: 3-hour-ahead PDP	0.74
x_2: 3-hour-ahead PDD	0.63
x_3: Predicted supply cushion (PSC)	-0.60
x_4: The SSR Ontario demand forecast	0.68
x_5: New England market demand	0.63
x_6: New York market demand	0.56
x_7: LMP_{NENY}	0.67
x_8: LMP_{NYON}	0.69
x_9: $\text{LMP}_{\text{PJMON}}$	0.67
x_{10}: The SSR capacity excess	-0.65

pressed as [58]:

$$z_t = c + \sum_{i=1}^{p} \phi_i z_{t-i} + \epsilon_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} \tag{3.3}$$

where c, ϕ_i and θ_j are the model parameters to be estimated, and ϵ_t is assumed to be an independently and identically distributed (i.i.d.) normal random variable (shock) with mean zero and variance σ_ϵ^2. Using the backward shift operator B, i.e., $Bz_t = z_{t-1}$, model (3.3) may be represented as:

$$\phi(B)z_t = c + \theta(B)\epsilon_t \tag{3.4}$$

where $\phi(B) = 1 - \phi_1 B - ... - \phi_p B^p$ is the auto regressive operator AR(p), $\theta(B) = 1 - \theta_1 B - ... - \theta_q B^q$ is the moving average operator MA(q).

Stationarity condition requires stability in the mean and variance of the process; most

real-life processes do not meet these requirements. Non-stationarity in variance is dealt with by the Box-Cox power transformations which is defined as $u_t = (z_t^\lambda - 1)/\lambda$ for $\lambda \neq 0 \in \mathbb{IR}$. For a given model, the optimal value of λ is found by minimizing the sum of squares of the residuals of the model. In case λ turns out to be close to or equal to zero, a natural logarithmic transformation $u_t = \ln(z_t)$ is used [58]. If non-stationarity is the result of a variable mean, the d^{th} order differenced process $v_t = (1-B)^d z_t$ is modeled. Setting $d = 1$ or $d = 2$ usually induces constant mean; the ARMA(p, q) model for the differenced process v is referred to as the Auto Regressive Integrated Moving Average model ARIMA(p, d, q) for the process z.

A time series with potential seasonality, indexed by s, is represented by a general ARIMA$(p, d, q)(P, D, Q)_s$ model:

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D z_t = c + \theta_q(B)\Theta_Q(B^s)\epsilon_t \tag{3.5}$$

where $\phi_p(B)$ and $\theta_q(B)$ are nonseasonal AR(p) and MA(q) operators; $\Phi_P(B^s)$ and $\Theta_Q(B^s)$ are seasonal AR(P) and MA(Q) operators; and B^s is the seasonal backward shift operator which is defined as $B^s z_t = z_{t-s}$. For hourly data, $s = 24$ and $s = 168$ indicate daily and weekly seasonality, respectively.

3.3.2 Dynamic Regression Model

The relationship between a dependent variable y and a set of explanatory variables $x_i, i = 1, 2, ..., n$, at time t can be expressed by a constant c, a transfer function term f, and a disturbance term N_t, as follows:

$$y_t = c + f(x_{1,t}, ..., x_{n,t}) + N_t \tag{3.6}$$

In ordinary linear regression (OLR) models, f in (3.6) is a linear function of the explanatory variables $x_i, i = 1, 2, ..., n$, and the disturbance term is assumed to be an i.i.d.

normal random shock, i.e.,

$$y_t = c + \omega_1 x_{1,t} + \omega_2 x_{2,t} + ... + \omega_n x_{n,t} + \epsilon_t = c + \sum_{i=1}^{n} \omega_i x_{i,t} + \epsilon_t \tag{3.7}$$

where ω_i are the coefficients to be estimated. If the current value of the dependent variable is affected by up to r_i past values of the i^{th} explanatory variable, in addition to the current values of the explanatory variables, the OLR model can be represented as:

$$\begin{aligned} y_t = c &+ (\omega_{1,0} x_{1,t} + \omega_{1,1} x_{1,t-1} + ... + \omega_{1,r_1} x_{1,t-r_1}) + \\ &(\omega_{2,0} x_{2,t} + \omega_{2,1} x_{2,t-1} + ... + \omega_{2,r_2} x_{2,t-r_2}) + ...+ \\ &(\omega_{n,0} x_{n,t} + \omega_{m,1} x_{n,t-1} + ... + \omega_{n,r_n} x_{n,t-r_n}) + \epsilon_t \end{aligned} \tag{3.8}$$

where $\omega_{i,j}$ corresponds to the coefficients for x_i at lag j to be estimated. Using the backward shift operator, model (3.8) can be represented as follows:

$$y_t = c + \sum_{i=1}^{n} \sum_{j=0}^{r_i} \omega_{i,j} B^j(B) x_{i,t} + \epsilon_t \tag{3.9}$$

Models (3.8) or (3.9) are referred to as distributed lag models [59].

It is also possible to include up to p past values of the dependent variable in this model as explanatory variables as follows:

$$y_t = c + (\phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p}) + \sum_{i=1}^{n} \sum_{j=0}^{r_i} \omega_{i,j} B^j x_{i,t} + \epsilon_t \tag{3.10}$$

where ϕ_is are model parameters to be estimated. By applying the backward shift operator on the dependent variable y, model (3.10) can be represented as follows:

$$\phi(B) y_t = c + \sum_{i=1}^{n} \sum_{j=0}^{r_i} \omega_{i,j} B^j x_{i,t} + \epsilon_t \tag{3.11}$$

where $\phi(B)$ is defined in (3.4). The models (3.10) and (3.11) are sometimes referred to as a dynamic regression (DR) [14, 19] or autoregressive dynamic model [59].

3.3.3 Transfer Function Model

In a more general form than the DR model, the relationship between the dependent variable y and the independent variables x_is, can be defined as a rational transfer function term and a disturbance term N_t as follows [57, 60]:

$$y_t = c + \sum_{i=1}^{n} \frac{\omega_i(B)B^{b_i}}{\delta_i(B)} x_{i,t} + N_t \tag{3.12}$$

where $\omega_i(B) = \sum_{j=0}^{r_i} \omega_{i,j} B^j$; $\delta_i(B) = 1 - \sum_{k=1}^{k_i} \delta_{i,k} B^k$; k_i is the order of the polynomial $\delta_i(B)$; b_i is referred to as the delay time for variable x_i; the disturbance term N_t is expressed by an ARMA model, i.e., $N_t = \theta(B)\epsilon_t/\phi(B)$; and the polynomial operators $\phi(B)$ and $\theta(B)$ are defined in (3.4). The model in (3.12) is referred to as a TF model.

3.3.4 Building Time Series Models

The Box-Jenkins three-stage procedure, comprising identification, estimation and diagnostic checking, for the ARIMA model building has been used here [58]. In the identification stage, the time series data is analyzed using the estimated autocorrelation function (ACF) and partial autocorrelation function (PACF), and a tentative model is selected. In the estimation stage, the parameters of the tentative model are estimated using the maximum likelihood method. In the diagnostic checking, the residuals of the model are examined for the i.i.d. assumption and, in case of failure, the tentative model is improved accordingly until it is acceptable. Normalized residuals time domain plots, residuals ACF, the Ljung-Box statistics [60], residuals probability plots, and plotting residuals against the fitted values are popular tests in the diagnostic checking stage; all of these are employed in the current work.

For building the TF and DR models, the three-stage linear transfer function (LTF) method [57] is used here. In the identification stage, a tentative LTF is selected for the

transfer function term. For TF models, a simple ARMA model is selected for the disturbance term, while for DR models, the disturbance term is assumed to be an i.i.d. normal random shock and a tentative AR term is specified for the autoregressive term $\phi(B)$. The tentative model is estimated and the disturbance term is checked for stationarity. In case of non-stationarity of the disturbance, the response and explanatory variables are properly transformed. The tentative terms are modified according to the behavior of the residuals of the model. A tentative rational form for the transfer function term may be identified at this stage for the TF models. The estimation and the diagnostic checking stages are similar to the aforementioned Box-Jenkins procedure for the ARIMA model.

3.4 Modeling Market Prices by Time Series Models

3.4.1 General Considerations

Three time periods, each of two weeks duration, are selected for building the time series models and generating HOEP forecasts. The first period comprises two consecutive weeks from April 26 to May 9, 2004, referred to as Week_1 and Week_2; during this period, the Ontario market demand reached its spring low point. The second period comprises two consecutive summer peak-demand weeks from July 26 to August 8, 2004, and are referred to as Week_3 and Week_4. The last period includes two high-demand winter weeks in 2004, spanning December 13-26, and are referred to as Week_5 and Week_6.

Models for each of the six weeks have been individually identified, estimated and checked. The ARIMA models are built using four weeks of historical data, while the TF and DR models are developed based on ten weeks of historical data. The main criteria for identifying the final models are as follows: diagnostic checking tests (Section 3.3.4); the principle of parsimony [60]; t-value of the estimated model parameters; Akaike Information Criterion (AIC) [57]; out-of-sample forecasts accuracy; and reality of the identified models.

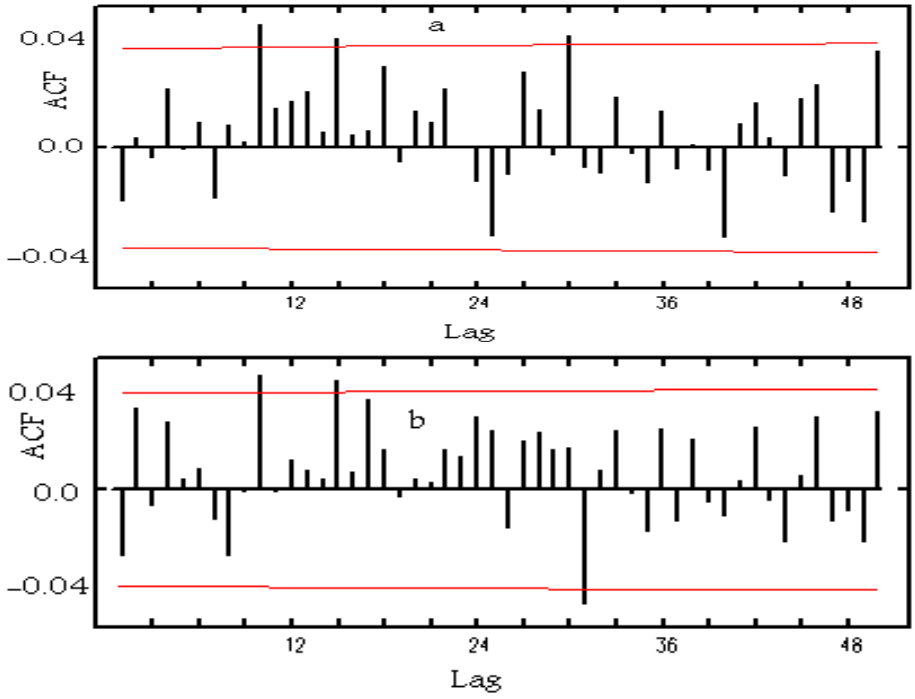

Figure 3.3: Week_3: a) Residuals ACF of the DR model. b) Residuals ACF of the TF model.

To illustrate some results of the diagnostic checking stage, the residuals ACFs of the 24-hour-ahead forecasts by the TF model and the 3-hour-ahead forecasts by the DR model developed for Week_3 (in Section 3.4.3) are presented in Figure 3.3; the horizontal bands in this figure represent the significance limits of the ACFs. Observe that no significant correlations for the first few lags and the relevant seasonal lags (e.g., 24, 48) exist.

The Scientific Computing Associates (SCA) statistical system is used here to build the proposed models [61]. To deal with outliers, the Chen-Liu algorithm for joint estimation of model parameters and outliers [62], implemented in the SCA system, was employed; however, no significant improvement in the overall forecasting accuracy was observed. Natural logarithmic transformation is found to be the optimal Box-Cox transformation for variance stability in this study, given the historical data and the identi-

Table 3.4: The ARIMA models for the HOEP

Week_1	$(1)(24,25,72,119)(168,169)Z_t = (1)(24)\epsilon_t$
Week_2	$(1)(24,25,72)(168,169,336)Z_t = (1)(24)\epsilon_t$
Week_3	$(1,2)(24,25)(168)Z_t = (1,2,3)(24)\epsilon_t$
Week_4	$(1,2)(24)Z_t = (1,2)(24)\epsilon_t$
$\text{Week}_{5,6}$	$(1)(23)Z_t = (1,2,3,4)(24)\epsilon_t$

fied models; hence, the forecasts are untransformed using the unbiased untransformation method in [63]. Furthermore, a seasonal differencing with $s = 24$ is applied to induce mean stationarity in all models. The transformed differenced HOEP time series, i.e., $(1 - B^{24}) \ln(\text{HOEP}_t)$, is referred to as Z_t here onward.

3.4.2 ARIMA Models for the HOEP

A shorthand convention, also used in [61], is employed here for simplicity to show the developed ARIMA models. According to this convention, an AR or an MA operator is represented by the orders of the respective backward shift operator. For example, the ARIMA model $(1-\phi_1 B)(1-\Phi_{24}B^{24}-\Phi_{47}B^{47})z_t$= $(1-\Theta_2 B^2)(1-\Theta_{24}B^{24})\epsilon_t$ is shown as $(1)(24,47)z_t = (2)(24)\epsilon_t$. The ARIMA models developed for each of the six studied weeks are listed in Table 3.4.

3.4.3 TF and DR Models for HOEP

Multicollinearity arises in a regression problem if there is a linear dependency among the explanatory variables. Popular methods to deal with the problem of multicollinear-

ity, such as ridge regression [64] and principal component regression [65], are developed in the ordinary regression framework. Hence, they are not applicable to multivariate time series models, given the inherent differences between model definitions and estimation for time series models and ordinary regression models. In this study, the following two-step procedure is designed for building the TF and DR models in the presence of multicollinearity among the explanatory variables:

1. In the first step, market knowledge, theoretical justifications, and linear correlation between the HOEP and the explanatory variable candidates are used to choose the most influential explanatory variable, referred to as the "first-step variable". TF and DR models for the HOEP are fully built assuming that the first-step variable is the only explanatory variable. In this step, the power transformation and the differencing order, which are needed for stabilizing the variance and the mean of the time series are identified.

2. In the second step, the general form of the transfer function term associated with the first-step variable is kept constant and other variables are added to the model in a step-wise manner; variables with collinearity with the first-step variable are considered first. The performance of the new models is monitored using the identification criteria mentioned in Section 3.4.1. The transfer function terms associated with each significant variable, as well as the disturbance terms, are modified appropriately in this step, and the final model is identified by adding other explanatory variable candidates and repeating this step.

Since the PDP variables are produced by mimicking the market clearing process using the most recent market data, they implicitly carry the information of inherent interactions among influential market variables. Hence, the 3-hour-ahead PDP variable, namely x_1, is considered as the first-step variable in the TF and DR models when the forecasting horizon is three hours. In addition, given the critical effect of Ontario demand on the

HOEP, the SSR demand forecast variable, namely x_4, is selected as the first-step variable when the forecasting horizon is 24 hours.

Lags 1, 2, 3, 4, 23, 24, 25, 47, 48, 49, 71, 72, 73, 95, 96, 97, 119, 120, 121, 143, 144, 145, 167, 168, 169, 335, and 336, of each explanatory variable candidate are considered for TF and DR model building. The inclusion of trading-day effects in the TF and DR models was not found to improve overall forecast accuracy; this can be attributed to the fact that demand forecasts are already used as model inputs, carrying the corresponding trading-day information.

TF Models

The final TF models in the first step for 3-hour-ahead and 24-hour-ahead forecasting hold the following form:

$$Z_t^{3h} = (\omega_{1,0} + \omega_{1,1}B + \omega_{1,2}B^2)(1 - B^{24}) \ln(x_{1,t}) + N_t^3 \tag{3.13}$$

$$Z_t^{24h} = (\omega_{1,0} + \omega_{1,1}B + \omega_{1,2}B^2)(1 - B^{24}) \ln(x_{4,t}) + N_t^{24} \tag{3.14}$$

Observe in models (3.13) and (3.14) that only the two latest values of the first-step variables affect HOEP behavior. The LTF method did not yield a rational form for the TF models in all cases in this step.

In the second step, the following TF models were finally identified for 3-hour-ahead and 24-hour-ahead forecasting:

$$Z_t^{3h} = \sum_{i=1}^{2} \sum_{j=0}^{2} (\omega_{i,j}B^j)(1 - B^{24}) \ln(x_{i,t}) + N_t^3 \tag{3.15}$$

$$\begin{aligned} Z_t^{24h} = {} & \sum_{j=0}^{2} ((\omega_{3,j}B^j)(1 - B^{24})x_{3,t} + \\ & (\omega_{4,j}B^j)(1 - B^{24}) \ln(x_{4,t})) + N_t^{24} \end{aligned} \tag{3.16}$$

No rational form was identified for the models in this step. The disturbance terms corresponding to models (3.15) and (3.16) are presented in Table 3.5 and 3.6, respectively.

Table 3.5: Disturbance terms for 3-hour-ahead forecasting TF models

$\text{Week}_{1,2}$	$(1)(24)N_t^3 = (24, 48)\epsilon_t$
$\text{Week}_{3,4}$	$(1, 2, 3)(23, 24)N_t^3 = (1)(24, 48)\epsilon_t$
$\text{Week}_{5,6}$	$(1)(23, 24, 25)N_t^3 = (1, 2)(24)\epsilon_t$

Table 3.6: Disturbance terms for 24-hour-ahead forecasting Transfer Function (TF) models

Week_1	$(1)(24, 25, 72)N_t^{24} = (1, 2, 3, 4)(24)\epsilon_t$
Week_2	$(1)(23, 24, 25, 72)N_t^{24} = (24)\epsilon_t$
$\text{Week}_{3,4}$	$(1)(24, 25)(168, 169)N_t^{24} = (1, 2, 3)(24)\epsilon_t$
$\text{Week}_{5,6}$	$(1)(24)N_t^{24} = (2)(24)\epsilon_t$

DR Models

The general forms of the final identified DR models for the six weeks under study, and for both forecasting horizons, are:

$$\phi(B)Z_t^{3h} = \sum_{i=1}^{2}\sum_{j=0}^{2}(\omega_{i,j}B^j)(1 - B^{24})\ln(x_{i,t}) + \epsilon_t \tag{3.17}$$

$$\phi(B)Z_t^{24h} = \sum_{j=0}^{2}((\omega_{3,j}B^j)(1 - B^{24})x_{3,t} + (\omega_{4,j}B^j)(1 - B^{24})\ln(x_{4,t})) + \epsilon_t \tag{3.18}$$

The final identified $\phi(B)$s turn out to be similar for both forecasting horizons, and can be presented as $\phi(B) =$ (1, 2, 3, 4, 23, 24, 25, 47, 48, 49, 71, 72, 73, 95, 96, 97, 119, 120, 121, 143, 144, 145, 167, 168, 169, 335, 336). However, the following lags were not found to be significant: lags 2, 3, 4, 23, 335 for $\text{Week}_{1,2}$; lags 3, 4, 25, 336 for $\text{Week}_{3,4}$; and lags 2, 335, 336 for $\text{Week}_{5,6}$.

3.4.4 ARIMA Models for the Neighboring Markets' LMPs

In order to compare the price behavior in Ontario with that in the neighboring markets, ARIMA models are also developed for LMP_{NENY}, LMP_{NYON}, and $\text{LMP}_{\text{PJMON}}$ (x_7, x_8, and x_9). Ten weeks of historical data are used to identify and estimate the following models for Week_1:

1. New England:

$$\begin{aligned}&(1)(23,24,25,48,72,96,120,144)(167,168,169,170)\times\\&(1-B^{24})\ln(x_{7,t}) = (1,2,3,4,5)(24,25)\epsilon_t\end{aligned} \tag{3.19}$$

2. New York:

$$\begin{aligned}&(1,2)(24,48,49,72,96)(168,169,336,337,504)\times\\&(1-B^{24})\ln(x_{8,t}) = (1)(24,48,72,96)\epsilon_t\end{aligned} \tag{3.20}$$

3. PJM:

$$\begin{aligned}&(1,2,3)(24,25,26,47,72)(167,168,169)\times\\&(1-B^{24})\ln(x_{9,t}) = (24)(167,168,169)\epsilon_t\end{aligned} \tag{3.21}$$

It is observed that the studied LMPs exhibit a stable behavior; in other words, models (3.19), (3.20), and (3.21) fit the data well for Week_2, and even for Week_3, and Week_4. Similar behavior is reported in [19, 16, 14, 20], where the developed models are shown to perform stably for long periods of time, in some cases for a full year. However, the final identified ARIMA, TF, and DR models for the HOEP are different for the studied weeks. The unstable behavior of the HOEP models highlights the fact that these models have to be re-identified and re-estimated after new observations are available. The need for model re-identification implies that market participants cannot count on a single model in order to produce HOEP forecasts in a non-supervised automatic manner, a fact that must be taken into account for commercialization of the forecasting models.

3.5 Numerical Results and Discussion

3.5.1 Final Identified Explanatory Variables

The final identified sets of explanatory variables differ for the two forecasting horizons. When the forecasting horizon is three hours, 3-hour-ahead PDD and New England demand are the significant explanatory variables for both of the TF and DR models identified in the second step. For the 24-hour forecasting horizon, PSC and New England demand are identified as significant variables in the second step. It is observed in [23] that the New England electricity market prices are generally higher than the HOEP, a factor affecting exports from Ontario, which in turn affects HOEP behavior. This would explain why the New England demand appears in the developed TF and DR models, hence improving the forecast MAPEs by about 1%. However, since after-the-fact New England demand data are used in the model building process, it is expected that if demand forecast is used in a practical implementation, it would offset this improvement; hence, the New England demand is excluded from the final presented models.

Because of the presence of the 3-hour-ahead PDP variable in the 3-hour-ahead forecasting models, the PSC variable does not appear in these models. This complies with the nature of PDPs, which are the final projected values of the HOEP taking into account all available market information and market variable interactions. Demand and price from other markets, and the SSR capacity excess or shortfall variable are also insignificant variables in the developed models, thereby implying that they do not carry additional information once the effects of the Ontario and New England demands are modeled.

3.5.2 Accuracy Measures

Two measures are used in this work to quantify out-of-sample forecast accuracy of the developed models. Thus, the absolute error (AE) is defined as:

$$\mathrm{AE}_t = \left| \mathrm{P}_t^f - \mathrm{P}_t^a \right| \tag{3.22}$$

where P_t^f and P_t^a are the forecast and the actual values of market price P_t, respectively. The absolute percentage error (APE) is defined as:

$$\mathrm{APE}_t = \frac{\left| \mathrm{P}_t^f - \mathrm{P}_t^a \right|}{\mathrm{P}_t^a} \tag{3.23}$$

The weekly mean absolute error (MAE) and mean absolute percentage error (MAPE) can be defined as follow:

$$\mathrm{MAE} = \frac{1}{168} \sum_{t=1}^{168} \mathrm{AE}_t \tag{3.24}$$

$$\mathrm{MAPE} = \frac{100}{168} \sum_{t=1}^{168} \mathrm{APE}_t \tag{3.25}$$

Neither the HOEP nor the other studied LMPs have ever shown a zero value during the studied period; therefore, the values of APE_t are finite values for all hours. In cases where market prices may take a zero value, modified versions of the above error measures can be used [7].

3.5.3 Forecasting Results for Ontario

The weekly MAPEs and MAEs of the generated HOEP forecasts, and those of the IESO-generated PDPs for the six studied weeks are presented in Tables 3.7 and 3.8. The forecasting origin for a set of 24-hour-ahead forecasts is the last hour of the previous day; for example, for the 24-hour-ahead forecast horizon spanning from today midnight to

tomorrow midnight, the forecasting origin is 11 PM today. On the other hand, forecasting origins move through the day in case of 3-hour-ahead forecasting. It is to be noted that the PDP values used in the present work have the same forecasting origins as the generated HOEP forecasts. The mean, standard deviation (STD), minimum (Min), and maximum (Max) of the HOEP for the six studied weeks are presented in Table 3.9.

Table 3.7: Weekly MAPEs (%), and weekly MAEs ($/MWh) fo the HOEP models

	3-hour-ahead							
	ARIMA		TF		DR		PDP	
	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE
$Week_1$	13.6	6.5	12.4	6.0	12.3	6.0	26	11.2
$Week_2$	15.5	7.0	14.7	6.7	12.9	6.2	26.4	10.6
Average[a]	**14.5**	**6.8**	**13.5**	**6.4**	**12.6**	**6.1**	**26.2**	**10.9**
$Week_3$	11	5.6	10.5	5.4	9.6	5.0	15.2	7.4
$Week_4$	14.3	5.6	11.9	4.8	12.2	5.3	15.1	6.0
$Week_5$	12.5	7.9	10	6.4	10.8	6.8	28.8	16.7
$Week_6$	17.6	9.75	13.2	7.6	12.5	7.1	28.9	16.5
Average[b]	**13.9**	**7.2**	**11.4**	**6.0**	**11.3**	**6.1**	**22.0**	**11.7**
Grand Average	**14.1**	**7.1**	**12.2**	**6.2**	**11.7**	**6.1**	**23.4**	**11.4**

[a] Average for the low-demand period ($Week_1$ and $Week_2$).

[b] Average for the high-demand period ($Week_3$ to $Week_6$).

The results presented in Tables 3.7 and 3.8 show that the accuracy of the gener-

Table 3.8: Weekly MAPEs (%), and weekly MAEs ($/MWh) fo the HOEP models

	24-hour-ahead							
	ARIMA		TF		DR		PDP	
	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE
$Week_1$	15.9	7.2	15.6	7.1	15.9	7.3	39.7	17.5
$Week_2$	18.6	8.2	18	8.2	18.1	8.2	30.3	12
Average[a]	**17.2**	**7.7**	**16.8**	**7.7**	**17**	**7.8**	**35**	**14.7**
$Week_3$	13.6	6.9	12.3	6.4	13	7.2	36.9	20.6
$Week_4$	21.5	8.7	18.3	7.3	19	7.6	31.6	12.3
$Week_5$	15.4	9.6	14.8	9.2	14.7	9.3	60.2	34.3
$Week_6$	20.8	12.0	17.5	10.1	18.5	10.7	37.3	22.8
Average[b]	**17.8**	**9.3**	**15.7**	**8.2**	**16.3**	**8.7**	**41.5**	**22.5**
Grand Average	**17.6**	**8.8**	**16.1**	**8.1**	**16.5**	**8.4**	**40**	**19.9**

[a] Average for the low-demand period ($Week_1$ and $Week_2$).

[b] Average for the high-demand period ($Week_3$ to $Week_6$).

ated HOEP forecasts is significantly higher than that of the IESO-generated PDPs. The overall MAPE of the generated forecasts are 23.9% and 11.7% lower than those of the IESO-generated PDPs for 24-hour-ahead and 3-hour-ahead forecasts, respectively. Also, the overall weekly MAE of the 24-hour-ahead and 3-hour-ahead forecasts generated by the TF and DR models are $11.8/MWh and $5.3/MWh lower, respectively, than those of the 3-hour and 24-hour-ahead IESO-generated PDPs. Furthermore, observe that, as

Table 3.9: HOEP statistics for each week

	Week$_1$	**Week**$_2$	**Week**$_3$	**Week**$_4$	**Week**$_5$	**Week**$_6$
Mean	45.4	46.4	51.2	46.7	54.9	51.6
STD	16.4	16.9	13.8	18.9	23.1	23.7
Min	18.9	13.9	20.6	10	35.4	30.3
Max	153.1	106.1	101.8	94.8	175.2	135

expected, the accuracy of the forecasts is generally higher for the shorter forecasting horizons.

The results in Tables 3.7 and 3.8 clearly show that for the high-demand period (Week$_3$ to Week$_6$), multivariate models outperform the univariate models; one can observe that the MAPEs of the 24-hour-ahead forecasts generated by the TF models and the 3-hour-ahead forecasts made by the DR models improve by 2.1%, and 2.6%, respectively. In terms of MAE, the improvements achieved over the ARIMA models by the multivariate models are \$1/MWh for the 24-hour-ahead forecasts and \$1.2/MWh for the 3-hour-ahead forecasts. However, for the low-demand period (Week$_1$ and Week$_2$), inclusion of the market data in the multivariate models does not improve forecast accuracy. This result implies that during low-demand periods, the market data carry less useful information than during high-demand periods. The improvements made by the multivariate models are lower when all the six studied weeks are considered.

Although inclusion of the *before-the-fact* market data into the forecasting models improves forecast accuracy to some extent, this improvement is not significant. This not-so-significant improvement in accuracy of the multivariate HOEP models can be attributed to the real-time nature of the Ontario market. Recall from Section 2.5.4 that the entire demand obligation has to be cleared in real-time in Ontario, which puts upward

or downward pressures on the HOEP, leading to price spikes. Furthermore, out-of-market actions by which the market operator manipulates the market clearing procedure affect the patterns behind the price behavior. Hence, the HOEP is highly volatile, as analyzed in more detail in Chapter 5, and the information contents of the before-the-fact market data have a high level of uncertainty.

In some studies, such as [14] and [19], multivariate models are reported to significantly outperform univariate models. These two studies use *after-the-fact* demand data for developing multivariate models; although building price models using actual demand data is important in order to discover the true demand-price relationships, these data are not available before real-time for a practical price forecasting tool. In contrast, the multivariate models developed in the present work do not significantly improve forecast accuracy compared with univariate models. This deficiency can be attributed to the lack of accuracy in information content of the *before-the-fact* data used, which is not the case when using *after-the-fact* data.

The forecasts obtained for Week_3, one of the highest demand weeks of 2004, have the lowest error among the studied weeks. Actual values of HOEP, and the most accurate HOEP obtained using the corresponding TF model are depicted in Figure 3.4. During this week, prices on all seven days were within the normal range, except for three unusual price spikes. Although the general price trend during this week could be forecasted with considerable accuracy, none of the three price spikes were captured by the models.

The 24-hour-ahead HOEP forecasts generated by the ARIMA and the TF models are plotted against the corresponding actual HOEP values for Week_4 in Figure 3.5, which presents the poorest forecasting results; the results for the DR model are similar to those of the TF and hence are not shown here. During this week, the prices are unusually high for the first two days, and relatively low for the rest of the week. Note that none of the models can reasonably forecast the unusually high or low prices, although the TF and DR models predict high/low prices relatively better than ARIMA.

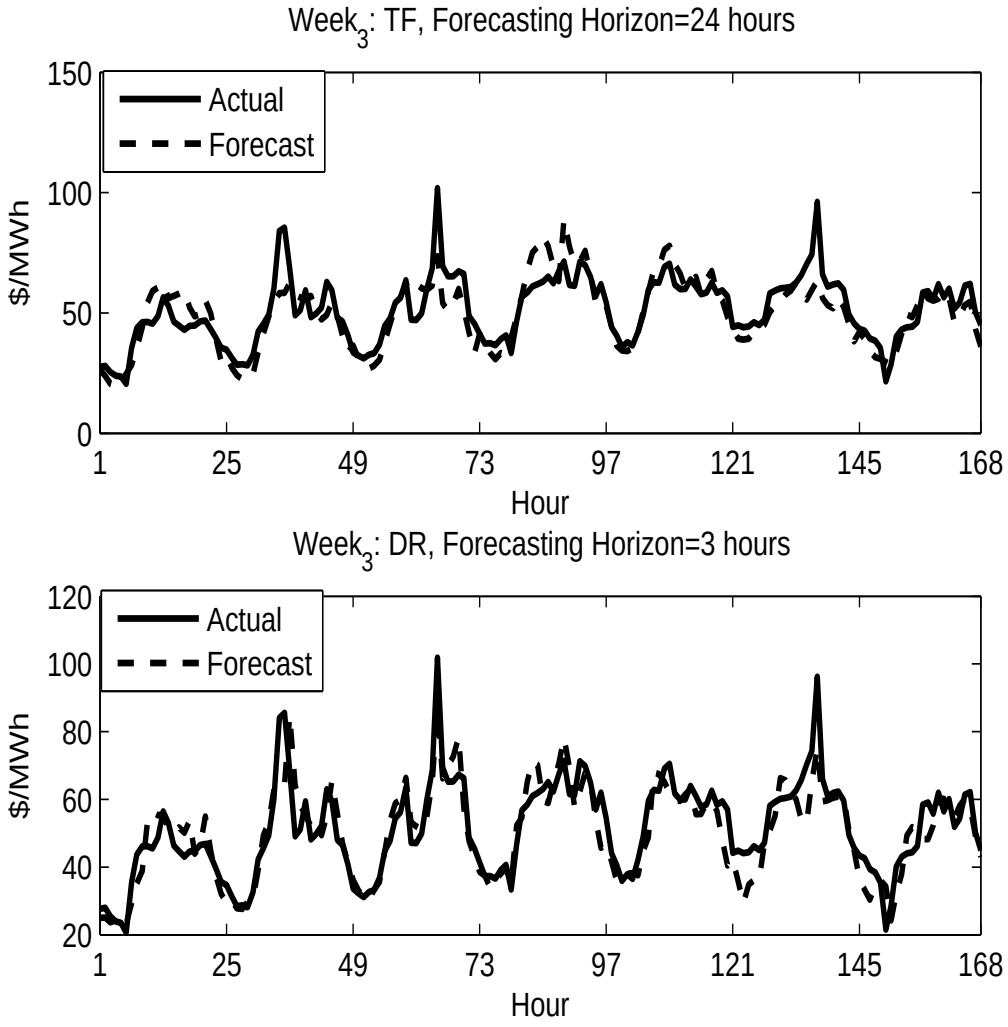

Figure 3.4: 24-hour-ahead and 3-hour-ahead HOEP forecasts for $Week_3$ by the TF and DR models.

During $Week_6$ the prices also behave irregularly, as depicted in Figure 3.6. For this week, severe mid-day price fluctuations happened in the first two days, and the Christmas holidays at the end caused unusual flat prices during the last three days; observe also the price spike on Christmas Eve day. When forecasting horizon is 24 hours, none of the models could predict the unusual prices; however, the 3-hour ahead forecasting models

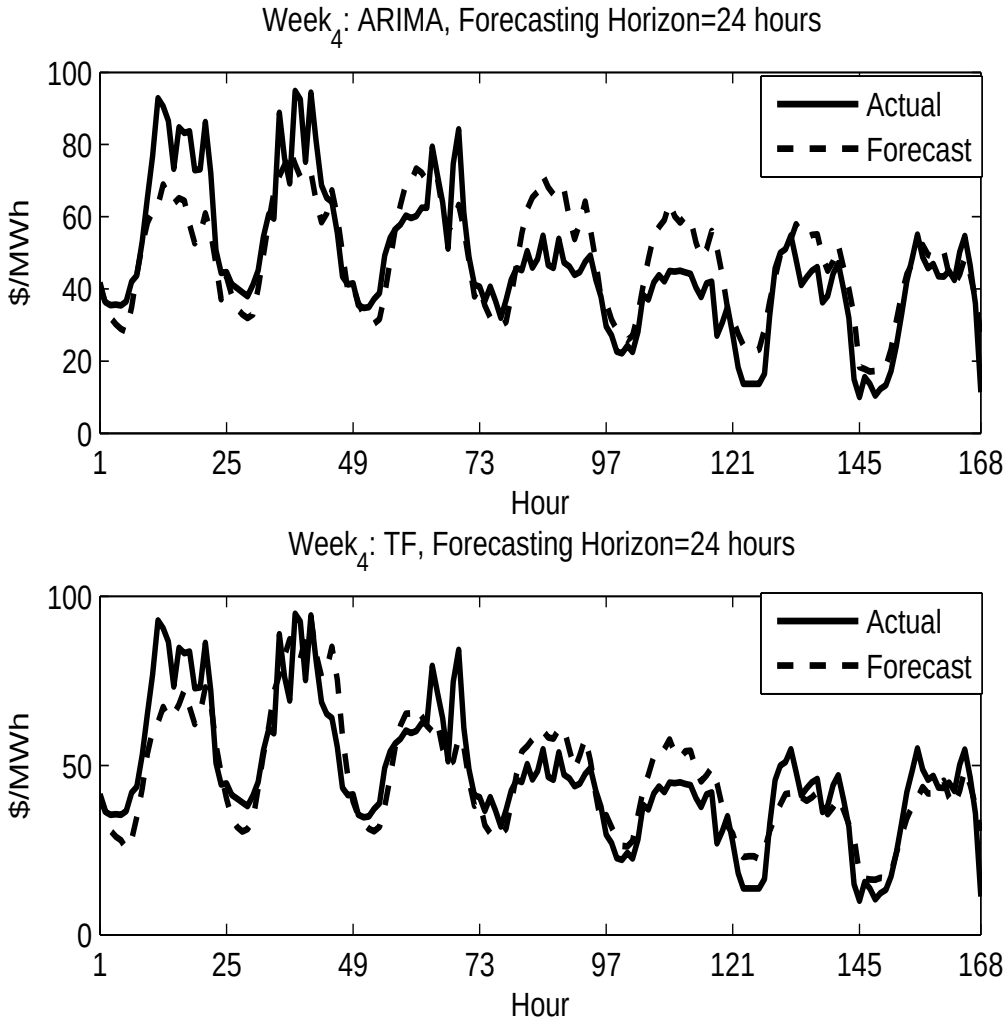

Figure 3.5: 24-hour-ahead HOEP forecasts for Week_4 by the ARIMA and TF models.

predict better the unusual prices in the market.

Histograms of APE_t values for the 24-hour-ahead forecasts by the TF models, and for the 3-hour-ahead forecasts by the DR models for the entire six-week period are shown in Figure 3.7; histograms of the corresponding PDPs for each forecasting horizon are also presented for comparison purposes. Although the accuracy of the generated HOEP forecasts is significantly higher than the PDPs, there still exist some hours during which

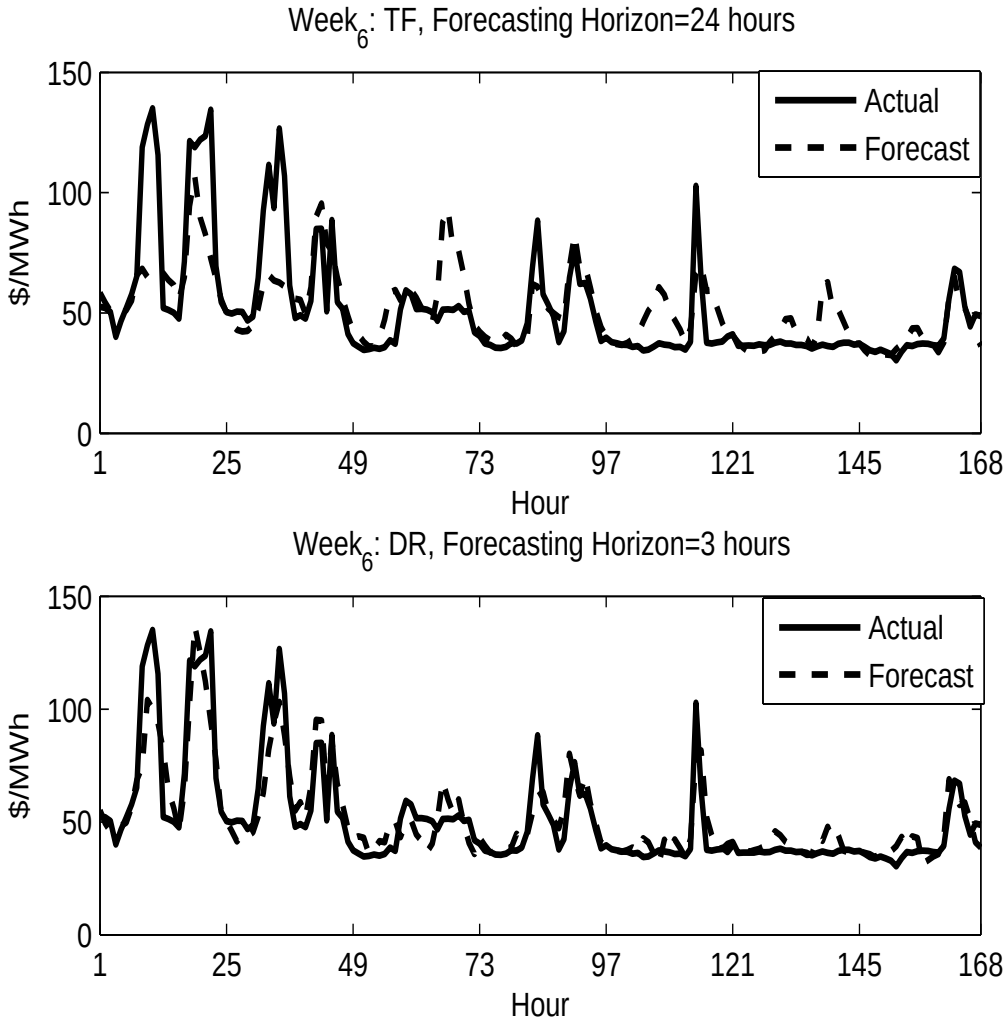

Figure 3.6: Forecasts by TF models for Week_4 for forecasting horizons of 3 and 24 hours.

the forecasting errors are relatively high. For example, for 28% of the hours in the case of 24-hour-ahead forecasts, and for 17% of the hours in the case of 3-hour-ahead forecasts, the APE_ts are more than 20%. If only the high-demand period, i.e., Week_3-Week_6, is considered, these numbers improve to 19% for 24-hour-ahead forecasts, and to 11% for 3-hour-ahead forecasts.

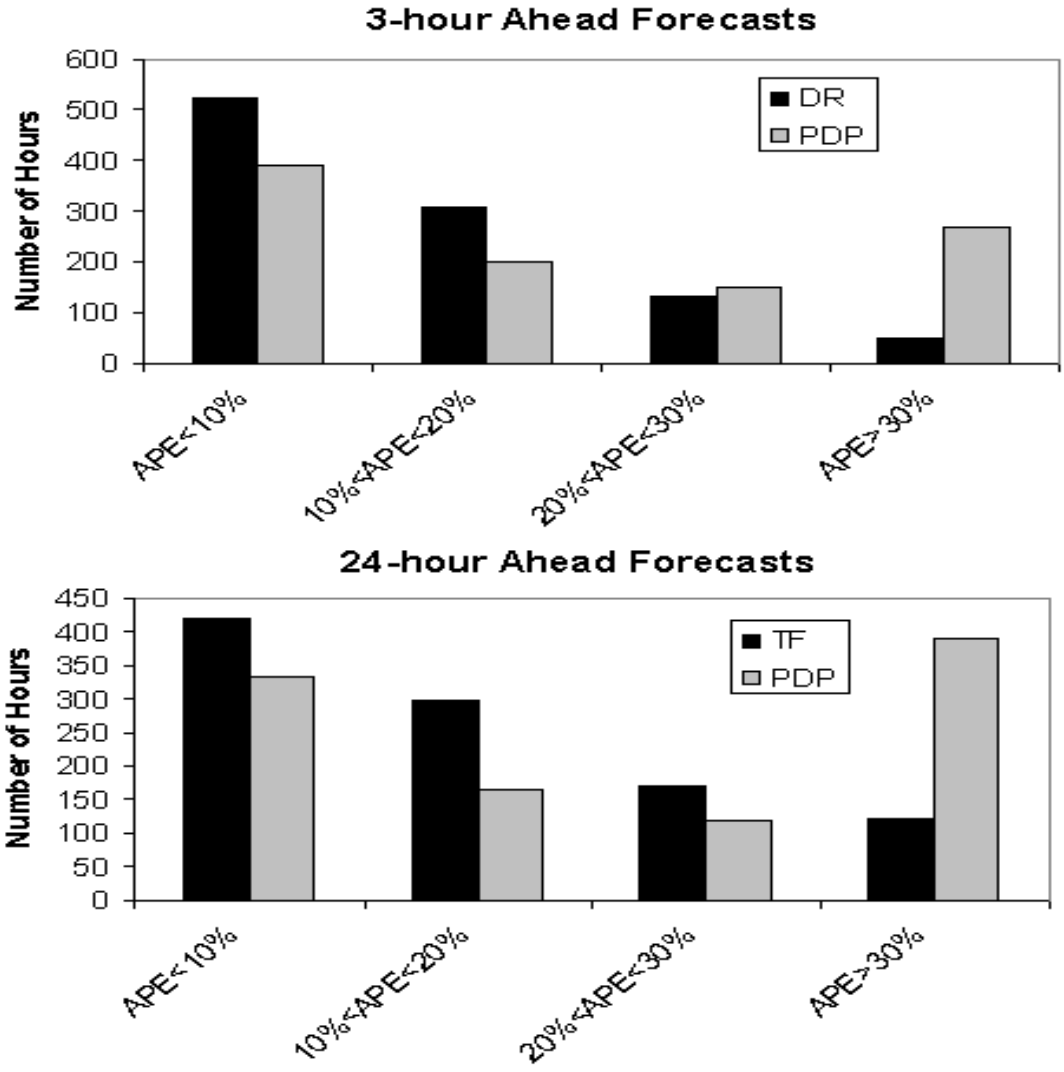

Figure 3.7: APE_ts histograms for the TF and DR models.

3.5.4 Forecasting Results for the Neighboring Markets

The models developed in Section 3.4.4 are used to generate 24-hour-ahead forecasts for the studied LMPs for $Week_1$ to $Week_4$. The time duration of these weeks is the same as that defined in Section 3.4.1, and correspond to two typical low-demand and two typical high-demand weeks in these markets. The calculated weekly MAPEs of the forecasts

Table 3.10: Weekly MAPEs (%) of the forecasts for the neighboring LMPs

	New England	**New York**	**PJM**	**Ontario**
Week_1	6.2	6.9	10.1	15.9
Week_2	5.4	7.1	11.4	18.6
Week_3	3.7	6.1	8.7	13.6
Week_4	7.1	8.1	17.3	21.5
Average	**5.6**	**7.1**	**11.9**	**17.4**

are presented in Table 3.10, along with the respective results for the HOEP forecasts for comparison. Observe that the accuracy of the forecasts generated for the New England, New York, and PJM day-ahead market LMPs, i.e., x_7, x_8, and x_9, is higher than the accuracy of the HOEP forecasts. As well, the HOEP forecasts reported in the present paper, and elsewhere [15, 23], have a much lower accuracy level than price forecasts for other markets (e.g., Spanish [19] and PJM [14] markets). A volatility analysis is presented in Chapter 5 in order to partly explain the relatively low accuracy level of the HOEP forecasts.

3.6 Summary

In this chapter, publicly available electricity market data from Ontario and its neighboring markets are evaluated to determine how efficiently they can be employed for improving HOEP forecast accuracy. A wide range of market information is studied in detail, and a final set of explanatory variable candidates are selected. Important practical issues, such as the availability of explanatory variables before real-time, and the market structure and operational time-line are considered in the process of selecting the explanatory variables,

which have not been addressed in previous research work. The novel concept of Predicted Supply Cushion (PSC) is introduced as an explanatory variable candidate in this study.

Multivariate TF and DR models are employed to relate HOEP behavior to the selected market variables. Univariate ARIMA models are also developed for the HOEP and their accuracy is compared with that of the multivariate models. The problem of multicollinearity among the explanatory variable candidates is addressed by a two-step model building procedure. The developed models are used to generate HOEP forecasts for low-demand, summer peak-demand, and winter peak-demand periods.

The results of this chapter demonstrate that the accuracy of the HOEP forecasts generated in this research is significantly higher than any other reported HOEP forecasts. However, that is still relatively low compared to the accuracy of forecasts generated for Ontario's neighboring electricity markets. Furthermore, it is observed that inclusion of the *just-in-time* publicly available data in multivariate HOEP models does not significantly improve the forecast accuracy, compared to the HOEP forecasts generated using univariate ARIMA models. It is also observed that the available market data is not very useful in predicting unusual upward or downward price spikes. The small improvements gained by using the multivariate HOEP models can be attributed to the poor information content of the market data available in practice. Finally, it is found that the final set of informative explanatory variables is different for each forecasting horizon, leading to different forecast accuracies. Furthermore, the developed HOEP models are less stable than the similar models developed for other electricity market prices. This model-instability highlights the difficulties of developing practical forecasting tools to predict short-term HOEP behavior.

Chapter 4

Forecasting the HOEP Using Non-linear Models[1]

4.1 Introduction

In Chapter 3, the application of linear time series models to forecasting the HOEPs is presented. Although the accuracy measures of the forecasts obtained by the linear models are significantly improved compared to previously reported HOEP forecasts, it is important to examine if any non-linear relationship exists between the explanatory variable candidates and the HOEP. In order to investigate this issue, two well-established non-linear techniques, namely, MARS and MLP neural networks, are examined in the present chapter.

MARS is an adaptive piece-wise regression approach and has been successfully employed for various prediction and data mining applications in recent years. This method has been shown to be particularly useful when a large number of explanatory variables

[1]Some of the findings of this chapter have been presented in the 2006 IEEE PES General Meeting, Montreal, Canada [66].

are involved. MLP networks have also been successfully applied to various function approximation problems in several areas. These networks are reported to be able to model virtually any function, provided an appropriate architecture is used. Neural network-based models have been applied to the HOEP forecasting problem in [15]; however, they are considered here again since a new set of explanatory variable candidates is used in the present work.

4.2 Multivariate Adaptive Regression Splines for Forecasting

4.2.1 MARS Review

Shortly after Friedman introduced MARS as a piece-wise non-linear regression method [25], it was demonstrated in [26] that MARS can be efficiently used to model time series. Subsequently, MARS has been applied to a variety of modeling and data mining problems; among these, MARS has been used for speech modeling [27], mobile radio channels prediction [28], and intrusion detection in information systems security [67]. In addition, MARS has been employed to model the relationship between retention indices and molecular descriptors of alkanes [68], to describe pesticide transport in soils [69], to predict the average monthly foreign exchange rates [70], to model credit scoring [71], and for data mining on breast cancer pattern [72]. In all of the cited studies, promising results are reported for both modeling and data mining applications.

MARS is a non-parametric modeling approach, different from the well-known global parametric modeling methods such as linear regression [25]. In global parametric approaches the underlying relationship between a target variable and a set of explanatory variables is approximated using a (usually simple) global parametric function which is fitted to the available data. While global parametric modeling methods are relatively easy

to develop and interpret, they have a limited flexibility and work well only when the true underlying relationship is close to the pre-specified approximated function in the model. To overcome the weaknesses of global parametric approaches, non-parametric models are developed locally over specific subregions of the data; the data is searched for optimum number of subregions and a simple function is optimally fit to the realizations in each subregion. Recursive Partitioning Regression (RPR) is one of the most studied paradigms of the non-parametric modeling approaches.

RPR is an adaptive algorithm for function approximation with the capability of handling a large number of explanatory variables. Let consider a set of explanatory variables $X = \{x_1, x_2, ..., x_n\}$ over a domain $D \subset \mathbb{IR}^n$, and a target variable y. The true relationship between y and X can be described as:

$$y = f(x_1, x_2, ..., x_n) + \epsilon \tag{4.1}$$

where f is an unknown function, and the error term ϵ is a white noise. Briefly, $f(X)$ is approximated in RPR as:

$$\hat{f}(X) = \sum_{m=1}^{M} a_m B_m(X) \tag{4.2}$$

where $\{a_m\}_1^M$ are the coefficients of the model which are estimated to yield the best fit to the data; M is the number of subregions $R_m \subset D$, or equivalently the number of basis functions in the model; and the basis function B_m is defined as:

$$B_m(X) = \begin{cases} 1 & X \in R_m \\ 0 & otherwise \end{cases} \tag{4.3}$$

Although RPR is a powerful method, it suffers from various shortcomings such as discontinuity at the subregion boundaries. MARS is a generalized version of RPR that overcomes some of the limitations of the original version.

The main core of the MARS modeling approach is the hockey-stick spline basis

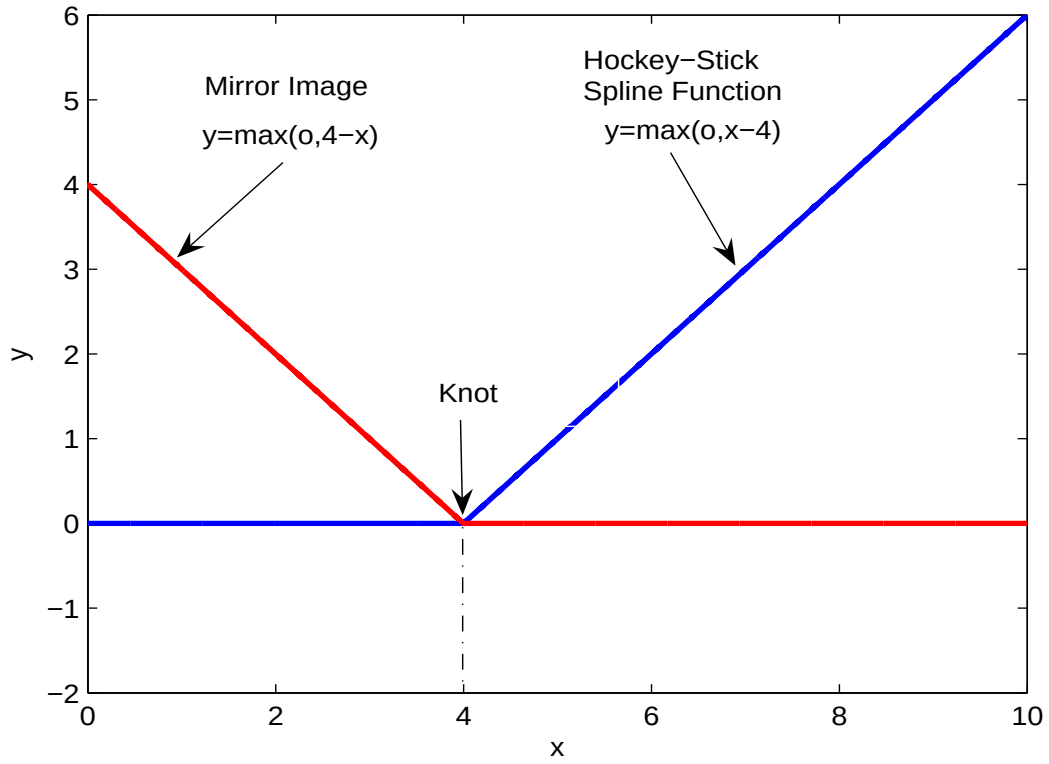

Figure 4.1: Hockey-stick spline basis function and its image for $c = 4$.

function, which maps a variable x to x^* as:

$$x^* = \max(x - c, 0) \tag{4.4}$$

where c is referred to as the basis function knot. The mirror image of the hockey-stick spline basis function is also exploited in MARS to handle non-zero slope for values below the knot, and it can be expressed as:

$$x^* = \max(-(x - c), 0) \tag{4.5}$$

A hockey-stick spline basis function and its mirror image are illustrated in Figure 4.1.

In MARS, the RPR model represented in (4.2) is modified as follows:

$$\hat{f}(X) = a_0 + \sum_{m=1}^{M} a_m H_m(X) \tag{4.6}$$

where a_0 is a constant, a_m and M are defined in (4.2), and the spline basis functions $H_m(X)$ are defined as:

$$H_m(X) = \prod_{k=1}^{K_m} [\max(s_{k,m}(x_{v(k,m)} - t_{k,m}), 0)] \tag{4.7}$$

where the explanatory variables associated with the basis function H_m are labeled by $v(k,m)$; K_m is the level of interaction between $v(k,m)$ variables; $t_{k,m}$ indicates the knot locations for H_m; and $s_{k,m}$ takes +1 for the hockey-stick basis function and -1 for its mirror image.

MARS models are developed through a two-stage forward/backward stepwise regression procedure. In the forward stage, the entire domain D is split into overlapping subregions and the model parameters are estimated by minimizing a lack-of-fit criterion. If the maximum number of subregions is not specified, an over fitted model with basically one subregion (basis function) per realization is built in this stage, and all possible interactions among the explanatory variables are considered. In the backward stage, the basis functions which no longer contribute to the accuracy of the fit will be removed. To make the MARS algorithm computationally affordable, the level of interaction between variables, as well as the maximum number of basis functions in the model are specified by the user.

A modified version of the generalized cross validation criterion (MGCV) is used in the MARS algorithm as the lack-of-fit criterion:

$$\text{MGCV} = \frac{1}{N_r} \sum_{i=1}^{N_r} [y_i - \hat{f}(X_i)]^2 / [1 - \frac{\tilde{C}(M)}{N_r}]^2 \tag{4.8}$$

where N_r is the number of realizations, $[1 - \frac{\tilde{C}(M)}{N_r}]^2$ is a penalty factor accounting for the increased variance resulting from a complex model, and $\tilde{C}(M)$ is defined as:

$$\tilde{C}(M) = C(M) + d \cdot M \tag{4.9}$$

where $C(M)$ is the number of parameters being fit, and d is another penalty factor with 3 as a typical value [25].

4.2.2 Modeling the HOEP Using MARS

In this section, the MARS technique is applied to the problem of HOEP forecasting. Thus, MARS models are developed to generate 3-hour-ahead and 24-hour-ahead forecasts for the six-week study period described in Section 3.4.1. Recall that this six-week period comprised two consecutive weeks of low demand, summer peak-demand and winter peak-demand.

Two scenarios are considered for model building: In the first scenario, denoted by SCN_1, MARS models are developed solely based on the historical HOEP values; this scenario can be imagined as a univariate adaptive non-linear autoregressive modeling paradigm. In the second scenario, denoted by SCN_2, the explanatory variable candidates selected in Section 3.2 are used as the right-hand side variables for model building. The planned outages variable is denoted by x_{11}, and is also considered here to capture any non-linear effects this variable may have on the HOEP behavior. The New England market demand variable is excluded from the set of explanatory variable candidates because this is the only *after-the-fact* explanatory variable candidate, as discussed in Section 3.5.1. The second scenario can be thought of as an adaptive non-linear multivariate dynamic regression modeling paradigm.

The models are built using 8 weeks of historical data, and the Scientific Computing Associates (SCA) statistical system [73] is used in this work for building MARS models. The final set of explanatory variable candidates for SCN_2 is presented in Table 4.1.

Table 4.1: The final explanatory variable candidates for SCN_2

x_1: 3-hour-ahead PDP	x_7: LMP_{NENY}
x_2: 3-hour-ahead PDD	x_8: LMP_{NYON}
x_3: Predicted supply cushion (PSC)	x_9: LMP_{PJMON}
x_4: The SSR Ontario demand forecast	x_{10}: The SSR capacity excess
x_6: New York market demand	x_{11}: The SSR planned outages

In order to limit the negative effects of outlier price observations on forecasting performance of the MARS models, the data is pre-processed before being used for model building. An upper limit of $200/MWh is defined for the HOEP, since prices over this amount are treated as anomalous prices by the Ontario IESO [23]. The pre-processing scheme is defined such that if the HOEP is more than $200/MWh, it will be replaced with a demand weighted average of the HOEPs of three similar previous days. This scheme is formulated below:

$$\text{HOEP}_t = \begin{cases} \text{HOEP}_t & \text{if } \text{HOEP}_t < 200 \\ \frac{\sum_{i=1}^{3}(\text{HOEP}_{t-168\cdot i}\cdot\text{Demand}_{t-168\cdot i})}{\sum_{i=1}^{3}\text{Demand}_{t-168\cdot i}} & \text{if } \text{HOEP}_t \geq 200 \end{cases} \tag{4.10}$$

SCN_1

In order to build MARS models in SCN_1, lags $l \in L^{SCN_1}$ of the HOEP (i.e., $HOEP_{t-l}$) are initially considered as explanatory variables, with L^{SCN_1} defined in (4.11). However, in this scenario, it is found that lags 335, 336 and 337 of the HOEP do not contribute to the developed models for all studied weeks. Furthermore, the contributing lags for individual weekly models are different for each week, which is consistent with the HOEP model-

Table 4.2: Significant lags for each week in SCN_1

$\text{Week}_{1,2}$	$L_{1,2}^{\text{SCN}_1} = \{1, 24, 48, 49, 120, 121, 144, 169\}$
Week_3	$L_3^{\text{SCN}_1} = \{1, 25, 48, 96, 97, 119, 120, 144, 145, 169\}$
Week_4	$L_4^{\text{SCN}_1} = \{1, 2, 23, 24, 25, 48, 72, 96, 120, 144, 145, 168, 169\}$
$\text{Week}_{5,6}$	$L_{5,6}^{\text{SCN}_1} = \{1, 2, 24, 48, 49, 72, 96, 120, 144, 168\}$

instability observed in Section 3.4.4.

$$L^{\text{SCN}_1} = \{1, 2, 3, 23, 24, 25, 47, 48, 49, 71, 72, 73, 95, 96, 97, 119, \\ 120, 121, 143, 144, 145, 167, 168, 169, 335, 336, 337\} \tag{4.11}$$

The set of final identified explanatory variables for each week in SCN_1 can then be presented as:

$$X_w^{\text{SCN}_1} = \{\text{HOEP}_{t-\eta} | \eta \in L_w^{SCN1}\} \tag{4.12}$$

where $w \in \{1, 2, 3, 4, 5, 6\}$ is the index of the studied week. The sets of the detected HOEP lags in SCN_1 are denoted by $L_w^{\text{SCN}_1}$ and presented in Table 4.2.

SCN_2

In SCN_2, MARS models are built by initially considering lags $l \in L^{\text{SCN}_1}$ of the HOEP, the explanatory variable candidates presented in Table 4.1 (i.e., $x_{i,t}$), and lags $l \in L^{\text{SCN}_2}$ of the explanatory variable candidates (i.e., $x_{i,t-l}$). Given the findings of Section 3.4.3 which showed only the recent values of the explanatory variables were informative, L^{SCN_2} is defined as follows:

$$L^{\text{SCN}_2} = \{1, 2, 3, 23, 24, 25, 47, 48, 49\} \tag{4.13}$$

Table 4.3: Significant lags of the final explanatory variables for each week in SCN_2, for both forecasting horizons

$\text{Week}_{1,2}$	$L_{1,2}^{3h,SCN2} = \{0, 1, 24, 25\}$	$L_{1,2}^{24h,SCN2} = \{0, 1, 24\}$
$\text{Week}_{3,4}$	$L_{3,4}^{3h,SCN2} = \{0, 1, 2, 24, 25\}$	$L_{3,4}^{24h,SCN2} = \{0, 1, 2, 24\}$
$\text{Week}_{5,6}$	$L_{5,6}^{3h,SCN2} = \{0, 1, 24, 25\}$	$L_{5,6}^{24h,SCN2} = \{0, 1, 2, 24, 25\}$

It is observed in this scenario that the autoregressive parts of the developed models are the same as found in SCN_1; in other words, the HOEP lags which appear in the weekly models in SCN_2 are identical to those that appeared in SCN_1. Similar to the explanatory variables found in Section 3.4.3, the 3-hour-ahead PDP and 3-hour-ahead PDD variables and their lags are identified here also in the case of 3-hour-ahead forecasting. In the case of 24-hour-ahead forecasting, the PSC and the SSR Ontario demand forecast variables and their lag are the detected informative variables here. The sets of final identified explanatory variables for each of the two forecasting horizons can be presented as:

$$X_w^{3h,\text{SCN}_2} = \{\text{HOEP}_{t-\eta}, x_{1,t-\gamma}, x_{2,t-\gamma} | \eta \in L_w^{SCN1}, \gamma \in L_w^{3h,SCN2}\} \tag{4.14}$$

$$X_w^{24h,\text{SCN}_2} = \{\text{HOEP}_{t-\eta}, x_{3,t-\beta}, x_{4,t-\beta} | \eta \in L_w^{SCN1}, \beta \in L_w^{24h,SCN2}\} \tag{4.15}$$

where, $L_w^{3h,SCN2}$ and $L_w^{24h,SCN2}$ are the sets of detected lags of the corresponding explanatory variable for the 3-hour and 24-hour forecasting horizons, respectively, for week w in SCN_2; and $x_{i,t-\gamma}$ and $x_{i,t-\beta}$ represent the explanatory variable x_i lagged by γ and β steps. Note that lag 0 in $L_w^{3h,SCN2}$ and $L_w^{24h,SCN2}$ represents the current value of the corresponding explanatory variable, i.e., $x_{i,t}$. The sets $L_w^{3h,SCN2}$ and $L_w^{24h,SCN2}$ are presented in Table 4.3.

The maximum number of basis functions in this work was initially selected to be very high, i.e., 2200, and an interaction level of 2 was examined between the input variables. However, while the models with the interaction took a much longer time to be generated,

they were not found to be more accurate overall than the models developed with no interactions considered. Hence, no interactions are assumed among the explanatory variables. It was also observed that a maximum number of 165 basis functions is enough to build the models. All final models took less than two minutes to be built on a Pentium(R) 4 CPU 2.53 GHZ, 1.0 GB RAM computer.

The final MARS models take the general form of:

$$\text{HOEP}_t = a_0 + \sum_{m=1}^{M} a_m \max(s_m(x_m - t_m), 0) + \epsilon \tag{4.16}$$

where HOEP_t is the value of the HOEP at time t to be forecasted; M is the final number of basis functions in the model; and s_m takes either +1 or -1; x_m represents the explanatory variables which contribute to the m^{th} basis function; and t_m is the knot location for the m^{th} basis function. As examples of (4.16), the developed models to forecast the HOEP during Week_1 in SCN_1 and in SCN_2 for a forecasting horizon of 3 hours are presented in Appendix A; the models developed for the other weeks are generally similar to those presented in Appendix A.

4.2.3 Numerical Results

The MARS models developed in the previous section are used to generate forecasts for the six individual weeks considered for the studies. As in Section 3.5, weekly MAEs and MAPEs are the measures used to assess the accuracy of the forecasts. The values of the weekly MAEs and MAPEs of the generated forecasts are presented in Table 4.4. For the sake of comparison, Table 4.5 summarizes the lowest MAPE values achieved by using the time series models presented in Chapter 3 and the MARS models discussed here.

Observe from the results presented in Table 4.4 that, in the case of 24-hour-ahead forecasting, the overall MAPE of the forecasts in SCN_2 is only improved by 1.1% . While this improvement is higher in the case of 3-hour-ahead forecasting, i.e., 2.4%, it is

Table 4.4: Weekly MAPEs (%) and weekly MAEs ($/MWh) of the forecasts by the MARS models

	24-hour-ahead				**3-hour-ahead**			
	SCN_1		SCN_2		SCN_1		SCN_2	
	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE
$Week_1$	17.5	7.8	16.1	7.5	14.2	5.8	12.5	6.1
$Week_2$	17.9	7.9	18.3	8.1	14.5	5.8	12.8	6.2
Average[a]	**17.7**	**7.9**	**17.2**	**7.8**	**14.3**	**5.8**	**12.7**	**6.1**
$Week_3$	14.2	7.1	13.3	7.3	10.6	5.1	9.6	4.3
$Week_4$	21.0	8.6	19.3	8.1	14.1	5.7	11.8	5.2
$Week_5$	15.2	9.5	15.2	9.3	13.1	7.8	10.4	7.1
$Week_6$	21.5	11.8	18.6	10.5	16.8	9.5	12.1	7.2
Average[b]	**18.0**	**9.2**	**16.6**	**8.8**	**13.6**	**7.0**	**11.0**	**6.0**
Grand Average	**17.9**	**8.8**	**16.8**	**8.5**	**13.9**	**6.6**	**11.5**	**6.0**

[a] Average for the low-demand period ($Week_1$ and $Week_2$).

[b] Average for the high-demand period ($Week_3$ to $Week_6$).

Table 4.5: The lowest six-weekly MAPEs (%) of the HOEP forecasts

24-hour-ahead		**3-hour-ahead**	
Transfer Function	MARS: SCN_2	Dynamic Regression	MARS: SCN_2
16.1	16.8	11.7	11.5

still not very significant. Hence, the inclusion of the just-in-time available market data in MARS models has not significantly improved HOEP forecast accuracy, especially when the forecasting horizon is long. This is consistent with the findings of Section 3.5, where the multivariate linear models also did not yield significantly improved forecasts than the univariate ARIMA models. These small improvements highlight the fact that the available market data is not informative enough to forecast the HOEP.

Also observe from the results presented in Tables 4.4, 3.7, 3.8, and 4.5 that forecast accuracy of the time series models in Chapter 3 and that of the MARS models in the present chapter are very close. In the case of 24-hour-ahead forecasting, the forecasts generated by the MARS models in both univariate and multivariate scenarios have slightly higher errors than those generated by the time series models. In this case, the overall forecast MAPE of the multivariate MARS models is 0.7% higher than that of the TF models. In the case of 3-hour-ahead forecasting, however, the overall MAPE of the HOEP forecasts by MARS models has improved only by 0.2%, compared to the overall MAPE of the forecasts by the time series models. It is worth mentioning that applying the pre-processing scheme helps improve the overall forecast accuracy just marginally.

The 24-hour-ahead HOEP forecasts generated by MARS for Week_3 and Week_4 are presented in Figure 4.2. Observe from this figure that, similar to the linear time series models in Chapter 3, the unusually low or high prices are not accurately predicted by the MARS models either.

Despite the fact that MARS models have not contributed to improving HOEP forecasting accuracy significantly, some of its modeling advantages should be highlighted. Given the automatic model building procedure in MARS, developing MARS models requires minimal intervention compared to time series models. For the time series models, many diagnostic tests and trial and error steps have to be considered in order to find the final best possible fit. In MARS, there are only a few parameters which need to be decided, such as, the interaction level and the maximum number of basis functions. Furthermore,

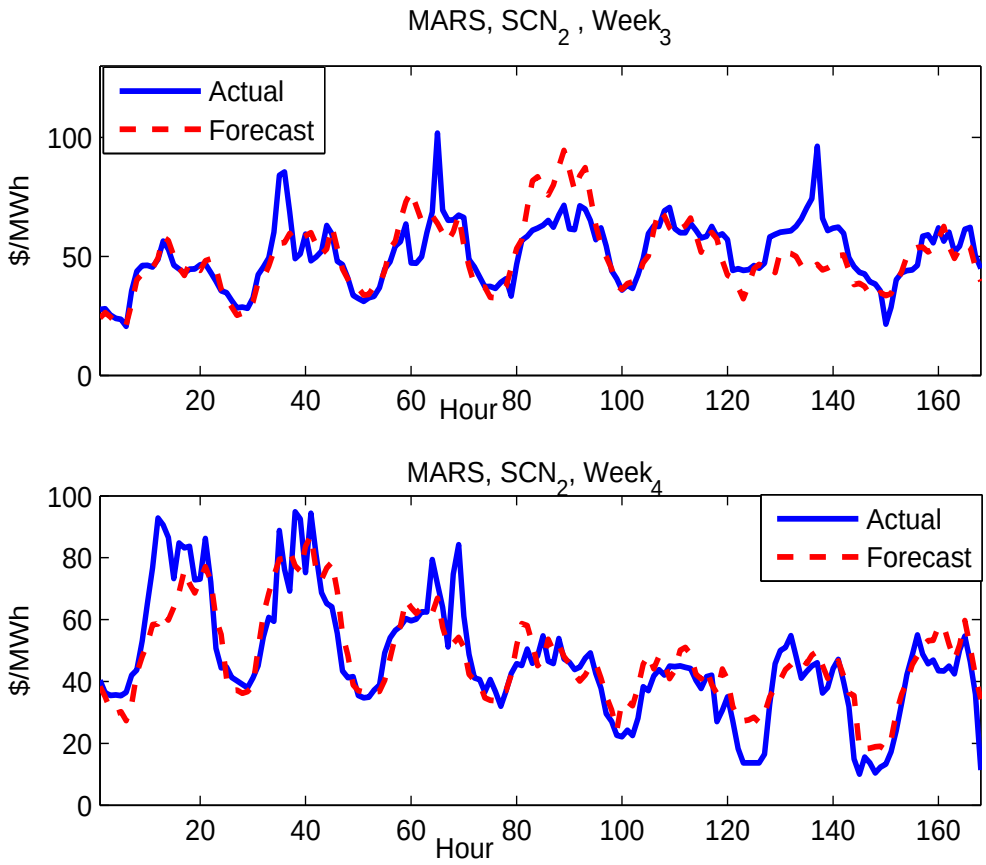

Figure 4.2: The 24-hour-ahead HOEP forecasts by MARS for Week_3 and Week_4.

comparison of the two-step model building procedure in Section 3.4.3 and the MARS model building process presented in this chapter shows that it is more cumbersome to build time series models than MARS models when a large number of explanatory variables are involved.

It should be noted that the inclusion of explanatory variable lags for building MARS models in this work is principally inspired by the development of time series models in Chapter 3. Although building linear models is more troublesome than MARS, linear models provide useful insights into the set of explanatory variable candidates lags which

are useful for building MARS models.

4.3 Artificial Neural Networks for Forecasting

ANNs are inspired by the biological structure of the human brain. The brain consists of a large number (about 10^{11}) of highly connected (approximately 10^4 connections per element) neurons. Each neuron receives electrical signals from other neurons as input, applies a transfer function to them and sends the resulting output signal to other neurons. MLP networks with Back Propagation learning algorithm are among the most common ANN architectures for function approximation and pattern recognition [74, 75, 76]. The feedforward MLP architecture is employed in this research, and is briefly explained next.

4.3.1 Feedforward MLP Networks

A typical artificial neuron model is shown in Figure 4.3. In this model, individual inputs $p_1, p_2, \ldots, p_R$ are multiplied by corresponding weights $w_{1,1}, \ldots, w_{1,R}$, and then a transfer function f is applied to the summation of weighted inputs with a bias b, typically 1. The output of the neuron can be written as $a = f(WP+b)$. An ANN usually has a few layers with each layer consisting of several neurons. The last layer is called the output layer, and other layers are referred to as hidden layers (Figure 4.5). In most cases, the transfer function for hidden layers is a log-sigmoid function, which is defined as:

$$f(n) = \frac{1}{1+e^{-n}} \tag{4.17}$$

where n is the input; a log-sigmoid function is depicted in Figure 4.4.

If the flow of signal in an ANN network is from the input factors to the output layer, without any feedback connection, the network is usually called a feedforward MLP. The number of input factors and output neurons in an MLP network are defined by the corresponding problem; however, there is no precise method to determine the number of

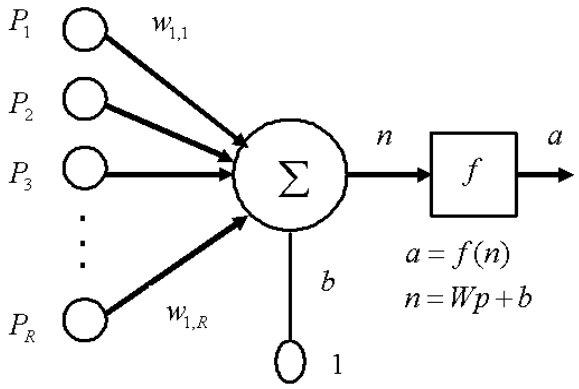

Figure 4.3: A simple artificial neuron model.

hidden layers as well as the number of neurons in each layer. MLP networks may be used for function approximation and pattern recognition purposes, and with two or three layers, they usually exhibit acceptable performance in most applications [75, 76].

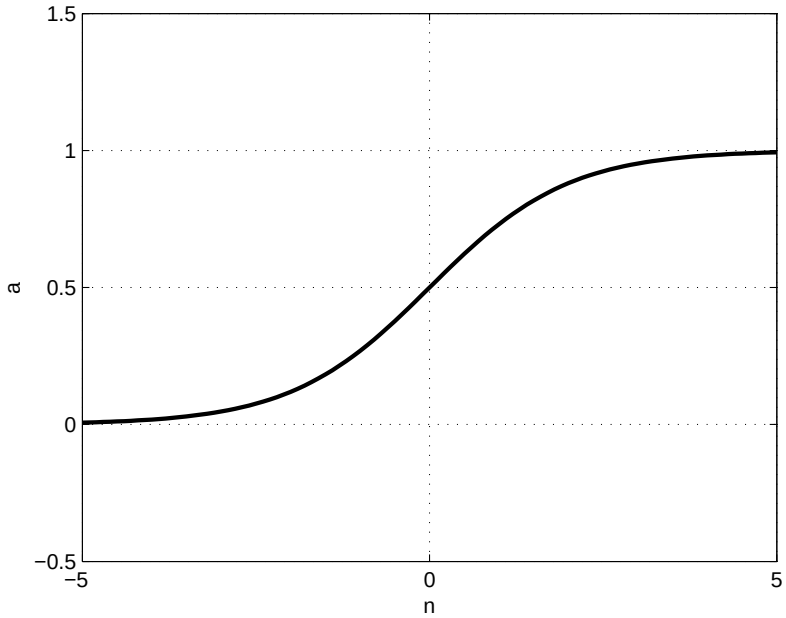

Figure 4.4: Log-sigmoid transfer function.

Preparing an MLP for function approximation has three major stages, namely, training, validation, and test stages. Accordingly, the data set is divided into three parts with the first and biggest part of the data being used for the training stage. In the training stage, the training data, comprising input patterns and the desired/measured output targets, are fed into the network and an optimization problem is solved to minimize an error function. The error function is usually defined based on the difference between the network's actual output and the desired output; mean square error is one typical error function. Many training algorithms are available for ANN training, with the Back Propagation algorithm being the most common. Standard Back Propagation is a gradient descent algorithm in which the network weights are moved along the negative of the gradient of the error function. There are several other variations of the basic algorithm that are based on other standard optimization techniques such as, Conjugate Gradient, Scaled Conjugate Gradient, and Newton methods [74, 76]. Training is continued until a defined performance goal is met. In the training stage, the neurons biases and connection weights are estimated according to a learning algorithm. The validation part of the data set is then fed to the network to finish its training. At this point, the MLP is ready to be tested by exposing it to the test part of the data. Depending on the performance of the network in the test stage, more training may be required.

4.3.2 Modeling the HOEP Using MLP Networks

The fully connected feedforward MLP networks discussed above, with one hidden layer, are trained and used for forecasting the HOEP during the six-week study period. The explanatory variables presented in Table 4.1 and their lags $l \in \{1, 2, 3, 23, 24, 25\}$ were considered for as MLP inputs. In addition to these, lags $l \in L^{\text{SCN}_1}$ of the HOEP, with L^{SCN_1} defined in (4.11), were also considered. Since there is no globally accepted rule for finding the optimal MLP structure, various combinations of MLP inputs were examined via numerous scenarios. The explanatory variables detected by the TF and DR models

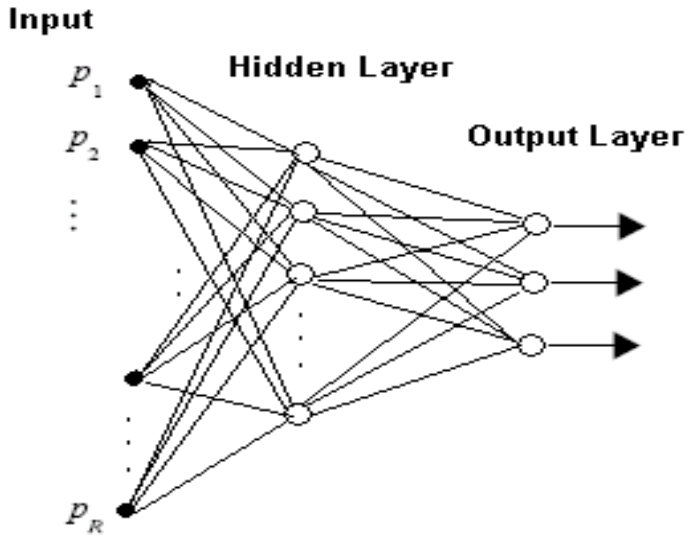

Figure 4.5: A typical feedforward ANN diagram.

in Section 3.4.3, and those detected by the MARS models in Section 4.2.2 were among the studied sets of inputs.

Separate MLP models were trained for 3-hour-ahead and 24-hour-ahead forecasting, because of the different informative explanatory variables previously found for each forecasting horizon. Note that in the case of 24-hour-ahead forecasting, x_1 and x_2 were not available at the forecasting origin, and hence were not considered for network building purposes. In the case of 3-hour-ahead forecasting, the SSR Ontario demand forecasts variable was excluded from the inputs, since the 3-hour-ahead PDD variable was considered in the set input variables.

Different numbers of neurons for the hidden layer were tested and different ranges of historical data, from 10 previous weeks to 36 previous weeks, were considered. The preprocessing scheme discussed in Section 4.2.2 was also applied to the employed data here. The MLP networks were developed using the MATLAB Neural Network Toolbox, and various training algorithms were examined. The Scaled Conjugate Gradient algorithm [76] yielded the best performance.

4.3.3 Numerical Results

Among the numerous scenarios examined, the highest forecast accuracy was achieved by the MLP networks when the explanatory variables detected by the MARS models in SCN_2 in Section 4.2.2 were used as the MLP inputs. In fact, the MARS models were used as a filter to detect MLP inputs, similar to the approach used in [71]. Seven neurons were assigned for the hidden layer in these scenarios, and 15 weeks of historical data were found to yield the most accurate results. The weekly MAEs and MAPEs of the forecasts are presented in Table 4.6.

Observe in Table 4.6 that the six-week MAPE of the 3-hour-ahead and 24-hour-ahead HOEP forecasts are 13.7% and 18.3%, respectively; these MAPEs are in the same order as those of the univariate ARIMA models reported in Chapter 3 (14.1% and 17.6%). However, these forecasts have generally higher error measures than those obtained from the proposed multivariate time series models in Chapter 3, and the MARS models in SCN_2 earlier in the present chapter (see Tables 3.7, 3.8, and 4.4).

The low accuracy of the MLP networks developed in this work is consistent with the findings reported in [14, 15], where ANNs are applied to forecasting electricity market prices. In [14], it is found that time series models out-perform ANNs for PJM market price forecasting. Furthermore, the HOEP forecasts generated by the neuro-fuzzy models in [15] is also reported to have high error levels. Considering the fact that determining the optimal structure of the ANNs is a difficult and time-consuming task, and also given the higher accuracy of the other proposed models in this research, MLP networks would not be the preferred models in the case of HOEP forecasting.

4.4 Summary

In this chapter, two non-linear modeling approaches, namely, MARS and MLP neural networks, are employed to forecast the HOEP. MARS is applied to the electricity mar-

Table 4.6: Weekly MAPEs (%) and weekly MAEs ($/MWh) of the forecasts by the MLP networks

	24-hour-ahead		**3-hour-ahead**	
	MAPE	MAE	MAPE	MAE
$Week_1$	17.8	7.8	14.2	5.8
$Week_2$	18.6	8.1	14.6	6.3
Average[a]	**18.2**	**8.0**	**14.4**	**6.0**
$Week_3$	14.5	7.7	10.1	5.2
$Week_4$	21.5	9.2	14.1	5.9
$Week_5$	15.6	9.5	13.1	7.9
$Week_6$	21.8	11.9	16.2	9.7
Average[b]	**18.3**	**9.6**	**13.4**	**7.2**
Grand Average	**18.3**	**9.1**	**13.7**	**6.8**

[a] Average for the low-demand period ($Week_1$ and $Week_2$).

[b] Average for the high-demand period ($Week_3$ to $Week_6$).

ket price forecasting problem for the first time in this work. The MARS models are developed in this study in two scenarios: a univariate scenario where only the historical HOEP data are used as explanatory variables, and a multivariate scenario where other candidates as explanatory variable are also considered. Consistent with the findings of Chapter 3, the SSR Ontario demand forecasts and the PSC variables appear in the final 24-hour-ahead MARS models. In the case of 3-hour-ahead forecasting, the 3-hour-ahead

PDP and the 3-hour-ahead PDD variables are detected by the MARS models as informative variables. It is further observed that the inclusion of market data in the non-linear MARS models does not improved forecast accuracy significantly.

Despite a slightly higher forecasting error rate for the MARS models, compared to the time series models, MARS has a more straightforward model building procedure than the time series models. Furthermore, it can easily handle a large number of explanatory variable candidates. These advantages are of practical importance given the fact that a high level of model-instability is observed in the case of Ontario market prices.

It is observed that applying MLP neural networks to forecasting the HOEP does not improve forecast accuracy. The highest forecast accuracy is attained by the MLP networks when the explanatory variables detected by the MARS models are used as the network inputs; however, this accuracy is even lower than that of the univariate ARIMA models developed in Chapter 3. In addition, developing ANNs and deciding about their architecture and inputs is a more complicated and time-consuming task than those of the time series and MARS models. Hence, MLP networks are not considered here as preferred tools for forecasting the HOEP.

Chapter 5

Price Volatility Analysis for the Ontario Electricity Market[1]

5.1 Introduction

In Chapter 3, it is observed that the accuracy of the price forecasts generated for the Ontario electricity market is lower than forecasts generated for the neighboring New England, New York, and PJM electricity markets. Furthermore, although the forecasts generated in Chapters 3 and 4 are significantly more accurate than any other available HOEP forecasts, they still have relatively high error levels. In order to explain these observations, a comparative volatility analysis is presented in this chapter and price uncertainty is compared across the New England, New York, PJM, and Ontario electricity markets.

Volatility indices in this work are developed based on the historical volatility and the price velocity concepts, previously applied to other electricity market prices. Intra-day, trans-day, and trans-week price fluctuations are the basis of the price volatility analyses

[1] Findings of this chapter have been submitted to *Journal of Energy Policy [77]*.

in this research. The analysis is carried out in two scenarios: in the first scenario, the volatility indices are determined for the price time series as a whole without splitting it into separate time series; in the second scenario, the price data is broken up into 24 time series corresponding to each of the 24 hours, and respective volatility indices are calculated separately. Considering the high HOEP forecasting errors in Chapter 3 and the volatility analysis in the present chapter, the relationship between price volatility and price predictability is discussed. The price volatility indices determined for the Ontario electricity market are also compared to the results reported for other electricity markets.

5.2 Analysis Measures and Methodology

Volatility refers to the unpredictable fluctuations of a process observed over time in everyday life. In economics and finance, volatility is basically a criterion to study the risks associated with holding assets. Volatility analysis, volatility modeling, and volatility forecasting have many applications such as risk management and option valuation in financial markets [78]. These applications can be extended to power markets [7], considering the fact that electricity market participants have several options for managing their energy needs, in the short-term, mid-term, and long-term. In addition, quantifying and comparing electricity market price volatility across different electricity markets can help market authorities in making appropriate amendments and advancements to market regulations and physical power systems.

With the advent of deregulated electricity markets in many countries, economic operation of power systems has been influenced by the volatile nature of electricity market prices [3]. In [3], volatility of the prices in the Spanish, Californian, UK, and PJM electricity markets are analyzed, concluding that the Spanish and PJM market prices were the lowest and the highest volatile, respectively. Volatility features of the Nordic day-ahead electricity market are studied in [5] for a 12-year period up until the year 2004,

and a high level of price volatility, in comparison to other financial markets, is reported. In [79], the periodic part of the price variations, for an unknown market, is separated out using a frequency-domain method and volatility of the remaining part is analyzed. Value at risk methodology is used in [80] to study volatility of the Californian market prices. Volatility analysis for 14 electricity markets, world-wide, is reported in [81], with widely varying price volatility behaviors being observed across different markets. A multivariate GARCH model is employed in [82] to study the inter-relationship among prices and price volatilities in the 5 Australian electricity markets. To the best of the author's knowledge, no price volatility analysis has been reported for the Ontario electricity market.

5.2.1 Historical Volatility

Let denote p_t as the spot price for a commodity at time t. The arithmetic return over a time period h is defined as:

$$R_{t,h} = \frac{p_t - p_{t-h}}{p_{t-h}} \tag{5.1}$$

and the logarithmic return, over the time period h, is defined as:

$$r_{t,h} = \ln(\frac{p_t}{p_{t-h}}) = \ln(p_t) - \ln(p_{t-h}) \tag{5.2}$$

When returns are small, the arithmetic and logarithmic returns are close, given the fact that:

$$r_{t,h} = \ln(\frac{p_t}{p_{t-h}}) = \ln(1 + R_{t,h}) \approx R_{t,h} \tag{5.3}$$

Most volatility analysis studies consider the logarithmic return over arithmetic return due to several reasons which are discussed in [83, 84]; hence, logarithmic return is used in the present work as well.

If the return values are i.i.d. over a time window T, one can present them as:

$$r_{t,h} = \hat{\mu}_{h,T} + \hat{\sigma}_{h,T}\, \epsilon_t \tag{5.4}$$

where $\hat{\mu}_{h,T}$ is the conditional mean return; $\hat{\sigma}_{h,T}$ is the conditional return variance; and the random variable ϵ is a mean zero, unit variance, i.i.d. innovation. $\hat{\sigma}_{h,T}$ is referred to as historical volatility over the time window T; in other words, historical volatility is defined as the standard deviation of arithmetic or logarithmic returns over a time window T. Given the return values, the estimated value of $\hat{\sigma}_{h,T}$ can be calculated as:

$$\sigma_{h,T} = \sqrt{\frac{\sum_{t=1}^{N_o}(r_{t,h} - \overline{r}_{h,T})^2}{N_o - 1}} \tag{5.5}$$

where $\sigma_{h,T}$ is the estimated value of historical volatility, N_o is the number of $r_{t,h}$ observations, and $\overline{r}_{h,T}$ is the simple $r_{t,h}$ average, all of them over the time window T.

In most volatility analysis studies, $h = 1$ is the commonly used time period. However, since electricity market prices usually follow the general trend of electricity demand, it is not surprising to encounter significant price fluctuations when moving from the off-peak hours to the on-peak hours of a day. In [5], the time period h is selected to be 24 hours and trans-day price fluctuations are analyzed. In the present study, however, in order to quantify the price uncertainty to which the market participants are exposed when moving from one week to the next week, trans-week price fluctuations are also considered. Thus, $h = 168$ hours is considered for the analyses, in addition to $h = 1$, and 24 hours.

The definition of historical volatility in (5.4) is based on the assumption that the logarithmic return observations follow an i.i.d. random variable. In other words, the returns are assumed to behave randomly, having constant mean and variance over the time window T. These assumptions are usually true for most stochastic returns in economics and finance, when ignoring small return correlations for the first few time steps [85, 83, 84]. However, electricity market prices follow daily, weekly and sometimes seasonal patterns, which are basically due to the seasonal behavior of electricity demand. As a result, electricity market price returns are highly correlated and do not behave as an i.i.d. random variable. For example, $r_{t,24}$ is calculated for the HOEP using (5.2), and their autocorrelation functions [60] for two arbitrary samples of 168 and 24 return observations ($T = 168$

and $T = 24$), are determined and displayed in Fig. 5.1. Observe that the returns are correlated for the first 7 lags in the $T = 168$ case, whereas for $T = 24$ the correlations are negligible. Accordingly, when studying electricity market price volatility, the time window T should be selected to be short enough in order to have negligible return correlations, which is the case in this study. Furthermore, selecting a short time window T allows for analyzing the original price time series without considering separation of the periodic and random parts of the price data.

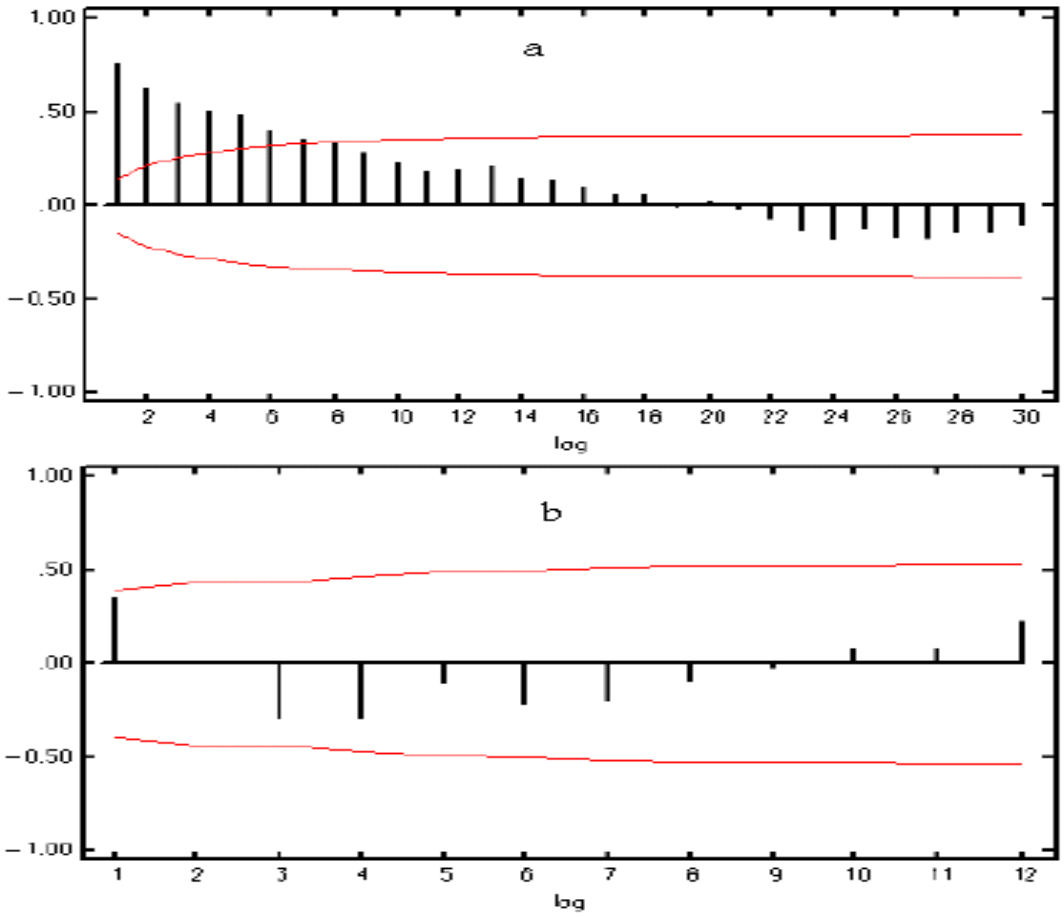

Figure 5.1: Autocorrelation functions for the logarithmic HOEP returns: a) $T = 168$; b) $T = 24$.

In order to define historical volatility indices, market price data are dealt with in two

scenarios. In the first scenario, a price time series is treated as a whole signal for all 24 hours of a day, and volatility indices are calculated as criteria for overall price behavior. In the second scenario, a price time series is broken up into 24 time series corresponding to each of the 24 hours. This scenario provides insight into the risks associated with the price at each particular hour of a day.

Scenario 1

In this scenario, the time window T is selected to be 24 hours (one full day) and the historical volatility for each studied day, i.e., $\sigma_{h,24}(d)$, is determined as:

$$\sigma_{h,24}(d) = \sqrt{\frac{\sum_{t=1+24\times(d-1)}^{24\times d}(r_{t,h} - \overline{r}_{t,h})^2}{23}} \tag{5.6}$$

where $d \in \{1, 2, 3, ..., 366\}$ is the index of studied days, and $\overline{r}_{t,h}$ is the logarithmic returns average over each day. In this scenario, hourly ($h = 1$), daily ($h = 24$), and weekly ($h = 168$) logarithmic returns are the basis of analysis and the averages of $\sigma_{h,24}(d)$ over all studied days, i.e., $\overline{\sigma}_{h,24}$, $h = 1, 24, 168$, are used as volatility indices.

Scenario 2

In this scenario, the logarithmic return for hour j, day k, over a time period h, is defined as:

$$r^j_{k,h} = \ln(\frac{p^j_k}{p^j_{k-h}}) = \ln(p^j_k) - \ln(p^j_{k-h}) \tag{5.7}$$

where p^j_k refers to price at hour j on day k. Choosing $h = 1$ in (5.7) implies that market prices at hour j for two consequent days are compared. By considering a 7-day time window (one full week), i.e., $T = 7$, and $h = 1$, historical volatility for hour j for week w can be determined as follows:

$$\sigma^j_{1,7}(w) = \sqrt{\frac{\sum_{t=1+7\times(w-1)}^{7\times w}(r^j_{k,1} - \overline{r}^j_{k,1})^2}{6}} \tag{5.8}$$

where w is the index of each studied week, $j \in \{1, 2, 3, ..., 24\}$ is the index of hour, and $\overline{r}^{j}_{k,1}$ is the return average over the 7-day time window for hour j. The average of $\sigma^{j}_{1,7}(w)$ over all studied weeks, i.e., $\overline{\sigma}^{j}_{1,7}$, is used as the volatility index for price at hour j. In fact, this scenario quantifies the fluctuations of price at a particular hour in subsequent days over a 7-day period.

5.2.2 Price Velocity

The authors in [79, 3, 5] employ the historical volatility concept in order to analyze electricity market prices volatility. On the other hand, the authors in [81], define *price velocity* for quantifying price uncertainty. Let define the absolute value of the difference between two prices which are h time period apart as:

$$\delta_{t,h} = |p_t - p_{t-h}| \tag{5.9}$$

Two volatility measures are defined in [81] for electricity market prices as "daily velocity based on overall average price" (DVOA), and "daily velocity based on daily average price" (DVDA), as follows:

$$\text{DVOA}_{d,h} = \frac{\text{Daily Average of } \delta_{t,h}}{\text{Overall Average of } p_t} \tag{5.10}$$

$$\text{DVDA}_{d,h} = \frac{\text{Daily Average of } \delta_{t,h}}{\text{Daily Average of } p_t} \tag{5.11}$$

where d is the index of studied day. The averages of $\text{DVOA}_{d,1}$ and $\text{DVDA}_{d,1}$ over the studied days, i.e., $\overline{\text{DVOA}}_1$ and $\overline{\text{DVDA}}_1$, are employed in [81] to compare price volatility in 14 electricity markets, world-wide. Observe that when $\overline{\text{DVOA}}_1$ is 0.2 for a specific market, it basically implies that average intra-day price change has been about 20% of the long-term average price for the studied period. Similarly, when $\overline{\text{DVDA}}_1$ is 0.2, it means that the average intra-day price change has been about 20% of the average daily price. It should be noted that the concept of price velocity differs from historical volatility in the

sense that it employs the daily average of price changes to quantify price uncertainty, in lieu of the standard deviation of price returns in historical volatility.

In this study, in order to analyze trans-day and trans-week price fluctuations, the price velocity concept is extended for the time periods $h = 24$ and $h = 168$ hours. By choosing $h = 24$ or 168 hours, $\overline{\text{DVOA}}_{24}$ and $\overline{\text{DVOA}}_{168}$ represent the average changes in prices in subsequent days and average changes in prices for a specific day in subsequent weeks, as a fraction of overall average price, respectively; $\overline{\text{DVDA}}_{24}$ and $\overline{\text{DVDA}}_{168}$ can be interpreted similarly. Values of $\overline{\text{DVOA}}_h$ and $\overline{\text{DVDA}}_h$ for $h = 1$, 24, and 168 hours, are also employed here as volatility indices.

Briefly, the following volatility indices are used in the present work to compare electricity market price volatility in Ontario and other markets: $\overline{\sigma}_{h,24}$, $\overline{\text{DVOA}}_h$, and $\overline{\text{DVDA}}_h$, with $h = 1, 24, 168$, and $\overline{\sigma}^j_{1,7}$.

5.3 Price Volatility Analysis for the Ontario Electricity Market

Historical HOEP data for the period of May 1, 2002, to April 30, 2005, is used to analyze price volatility in the Ontario electricity market. The data is available at www.ieso.ca. The analyzed HOEP data is depicted in Fig. 5.2, where the HOEP fluctuations over an arbitrary week (the week May 17-23, 2004), are also displayed. Observe from Fig. 5.2-a that during the first year of market operation, the prices were higher and more unstable than the next two years. For instance, from the first to the third year, the number of hours during which the HOEP exceeded \$200/MWh was 106, 10, and 3, respectively, and the average HOEP was \$58.4/MWh, \$48.2/MWh, and \$51.2/MWh, respectively. Unusual prices during the on-peak and off-peak hours are one of the features of the Ontario electricity market and happen on both weekdays and weekends, as illustrated in Fig. 5.2-b.

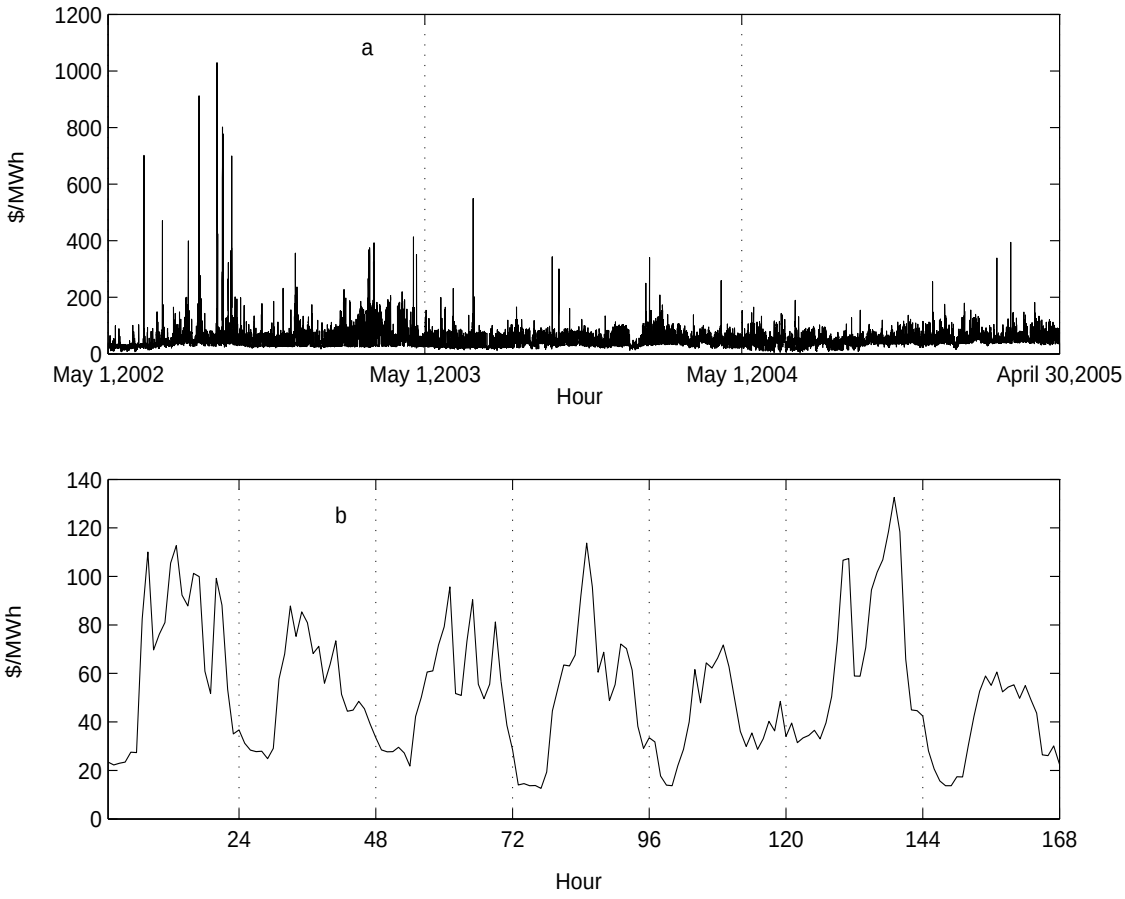

Figure 5.2: a) HOEP fluctuations over the the period of May 1, 2002, to April 30, 2005; b) HOEP over the week May 17-23, 2004.

5.3.1 Historical Volatilities

Scenario 1

Historical volatilities, i.e., $\sigma_{h,24}(d)$, are calculated for $h = 1$, 24, and 168 hours and are depicted in Fig. 5.3. Observe in this figure that the highest HOEP volatiles occurred in February and early March in year 2003, as discussed in Section 2.5. The period of January, February, and early March, 2003 was an extremely cold period and natural

gas prices were very high. As a result of this, many gas-fulled generating facilities experienced difficulties and were unavailable to the IESO. High demand and shortage of supply in this period resulted in unusually high and volatile HOEPs even in the off-peak hours [23]. It is also observed that the historical volatilities are relatively higher during the high demand periods than the low demand periods, as expected.

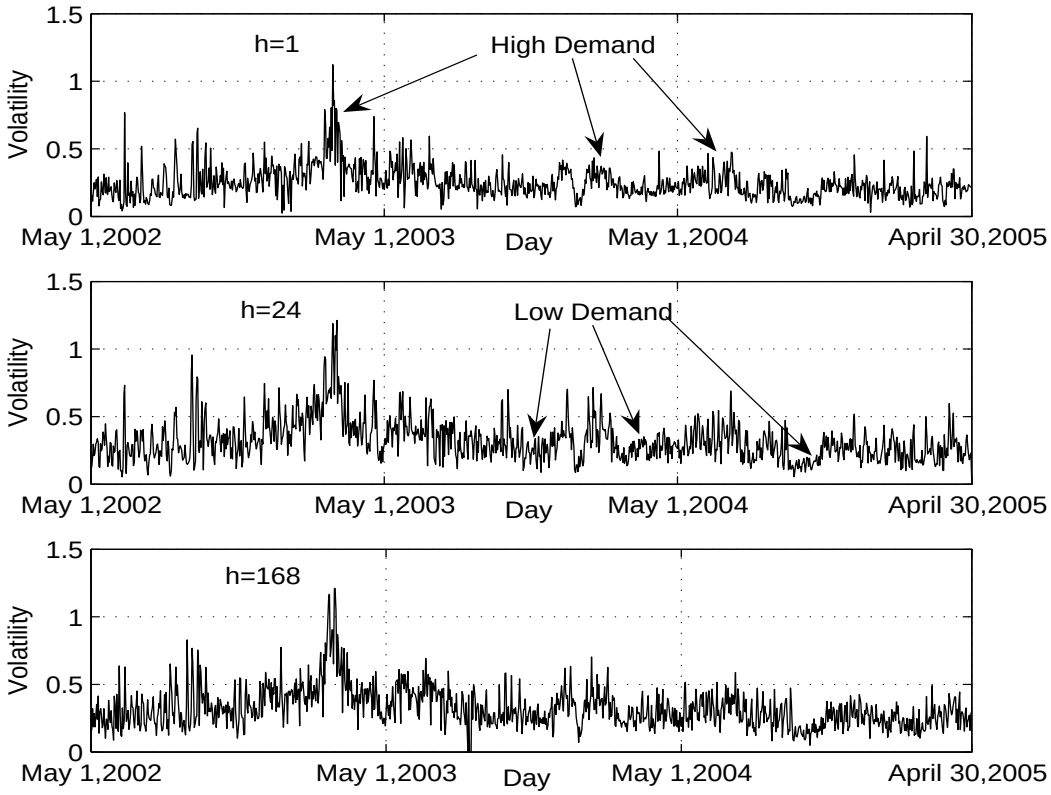

Figure 5.3: HOEP volatilities for $h = 1, 24, 168$; and $T = 24$.

Volatility indices $\overline{\sigma}_{h,24}$, $h = \{1, 24, 168\}$, are determined over the entire 3-year period and for each of the 3 years of market operation and are presented in Table 5.1. Observe that these volatility indices have declined from the first year to the third year, with the first

year showing the highest volatility. This improvement can be attributed to the amendments made to the market rules and regulations, and also is a matter of market maturity. Furthermore, the values of trans-day and trans-week volatility indices, i.e., $\overline{\sigma}_{24,24}$ and $\overline{\sigma}_{168,24}$, are higher than the value of intra-day volatility index, i.e., $\overline{\sigma}_{1,24}$; this basically implies that, on average, the trans-day and trans-week price changes fluctuate in a wider range than the intra-day price changes.

Table 5.1: Historical volatilities for Ontario market

	$\overline{\sigma}_{1,24}$	$\overline{\sigma}_{24,24}$	$\overline{\sigma}_{168,24}$
Year 1	0.2843	0.3771	0.3821
Year 2	0.2477	0.3214	0.3220
Year 3	0.2088	0.2623	0.2613
3-Year Period	0.2469	0.3203	0.3222

Historical volatilities are studied in [5] for the Nordic electricity market over a 12 years period ending May 2004. Only the time period $h = 24$ is considered and it is found that $\overline{\sigma}_{24,24} = 0.16$ (12-year average). Furthermore, in [5], the level of price volatility in the Nordic electricity market is compared with average historical volatilities in some other markets, such as stock indices with $\overline{\sigma}_{24,24} = 0.01 - 0.015$, crude oil markets with $\overline{\sigma}_{24,24} = 0.02 - 0.03$, and natural gas markets with $\overline{\sigma}_{24,24} = 0.03 - 0.05$. Comparing $\overline{\sigma}_{24,24} = 0.3203$ for the HOEP with these volatility indices reveals higher price volatility in the Ontario electricity market.

Scenario 2

In this scenario, $\sigma^j_{1,7}(w)$, $j = 1, 2, 3, ..., 24$, $w = 1, 2, 3, ..., 156$, are calculated using (5.8), and the corresponding three-year averages with respect to w, i.e., $\overline{\sigma}^j_{1,7}$, are dis-

played in Fig. 5.4. Observe in this figure that $\overline{\sigma}^{j}_{1,7}$s fluctuate across the different hours, with hours 5 and 8 have the lowest and the highest indices, respectively. Furthermore, prices at the on-peak hours, i.e., between hours 7 to 21, are the most volatile prices, while prices at off-peak hours, i.e., hours 22 to 24 and hours 1 to 6, are the least volatile, as expected. However, it should be noted that the price volatility at off-peak and on-peak hours are both significantly high.

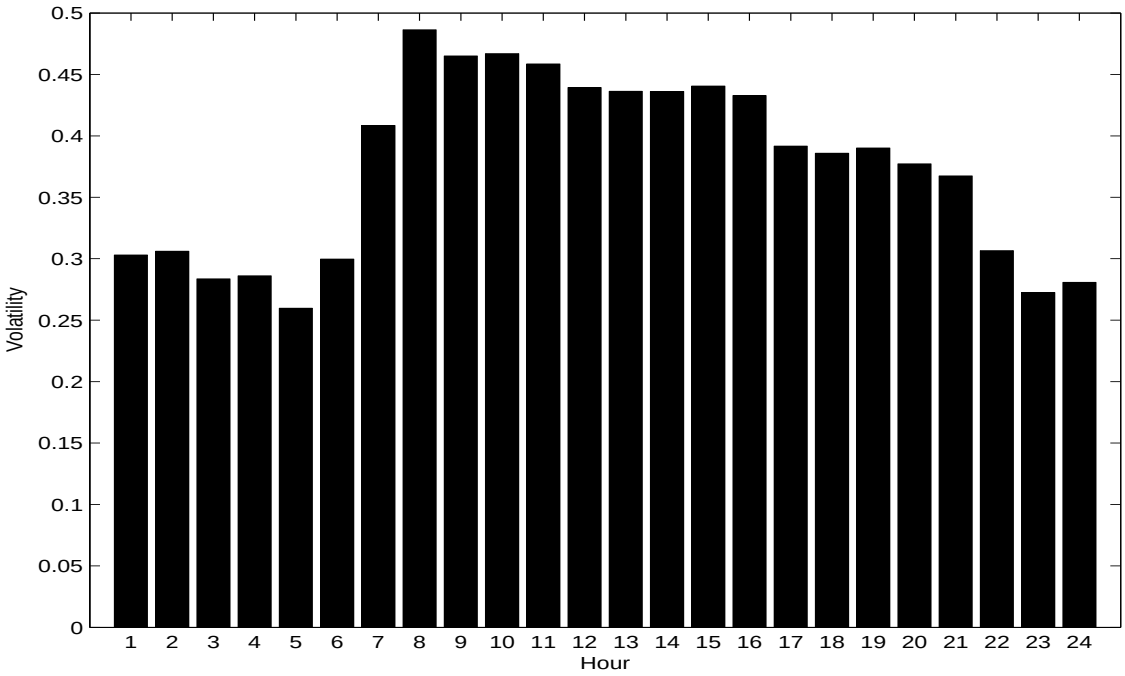

Figure 5.4: HOEP volatilities for each hour.

5.3.2 Price Velocities

The values of $\overline{\text{DVOA}}_h$ and $\overline{\text{DVDA}}_h$, for $h = 1$, 24, and 168 hours, for the HOEP are also calculated over the 3-year period, as well as for each of the three years of market operation, and are presented in Table 5.2. These results demonstrate that the $\overline{\text{DVOA}}_h$ and $\overline{\text{DVDA}}_h$ values have also declined over the years, which is consistent with the findings

of historical volatility. Furthermore, $\overline{\text{DVOA}}_h$ and $\overline{\text{DVDA}}_h$ have the highest values for $h = 168$ hours, and the lowest values for $h = 1$ hour; from this, it can be inferred that the average trans-day and trans-week changes in prices are higher than the average intra-day changes in prices.

Table 5.2: Price velocities for the Ontario market

	$\overline{\text{DVOA}}_1$	$\overline{\text{DVDA}}_1$	$\overline{\text{DVOA}}_{24}$	$\overline{\text{DVDA}}_{24}$	$\overline{\text{DVOA}}_{168}$	$\overline{\text{DVDA}}_{168}$
Year 1	0.2438	0.1890	0.4208	0.3430	0.5106	0.4379
Year 2	0.1604	0.1719	0.2694	0.2966	0.3083	0.3430
Year 3	0.1463	0.1478	0.2346	0.2421	0.2761	0.2849
3-Year Period	0.1835	0.1696	0.3083	0.2939	0.3655	0.3557

Values of $\overline{\text{DVOA}}_1$ and $\overline{\text{DVDA}}_1$ for the Scandinavia, Spain, California, New Zealand, the UK, Leipzig (Germany), New England, Australia (New South Wales, Victoria, South Australia, Queensland), Alberta (Canada), Netherlands, and PJM electricity markets are presented in [81]. Time duration of the study varies across the markets, all of them ending December 31, 2001. Six of the studied markets in [81], namely, Alberta, PJM, Netherlands, Victoria, South Australia, and Queensland, show higher $\overline{\text{DVOA}}_1$ values than that for the Ontario market. Furthermore, Alberta, South Australia, and Queensland electricity markets are reported to have $\overline{\text{DVDA}}_1$ values higher than the $\overline{\text{DVDA}}_1$ value obtained for the Ontario market. Despite the differences in time durations of the study in [81] and the present paper, it can be concluded that Ontario electricity market is among the most volatile markets from the $\overline{\text{DVOA}}_1$ and $\overline{\text{DVDA}}_1$ point of views.

5.4 Price Volatility Analysis for the Ontario's Neighboring Electricity Markets

As discussed earlier in Chapter 2, the Ontario electricity market is interconnected with New York electricity market and Quebec, as well as with the Michigan, Manitoba, and Minnesota control areas (see Fig. 2.6). The interconnected electricity markets have facilitated interjurisdiction energy trades in the region and the supply and demand sides have the option of freely playing in these markets.

The price and demand information regarding Michigan, Manitoba, and Minnesota control areas are only available after April 1, 2005, on which the Midwest market was launched, and hence is beyond the study period of this paper. Furthermore, Quebec has a regulated electricity sector and no market activity has been initiated yet. Therefore, New York, New England and PJM markets are the only studied electricity markets here.

The year 2004 historical LMP data for 9 pricing points in New England, 9 pricing points in New York electricity, and 15 pricing points in PJM electricity market are used in this study to calculate the neighboring markets' volatility indices. The data are available at www.iso-ne.com, www.nyiso.com, and www.pjm.com, respectively. These pricing points include load zones, and interfaces with other areas. Since the day-ahead market has the dominant share of energy transactions in these markets, only day-ahead LMPs are considered for the analysis. The presented quantities in this section are the average of the corresponding quantities for all studied pricing points for each market.

5.4.1 Historical Volatilities

Scenario 1

The averages of volatility indices $\overline{\sigma}_{h,24}$, $h = \{1, 24, 168\}$, for the studied pricing points in each of the three markets are presented in Table 5.3. The corresponding indices for

the HOEP over the same year are also presented in this table for the sake of comparison. These results show that Ontario market price has the highest historical volatility among the studied market prices, with the New England market showing the lowest historical volatilities. Observe that the intra-day volatilities are higher than the trans-day and trans-week volatilities for the neighboring markets, while for Ontario is the reverse.

Table 5.3: Historical volatilities for the Ontario and its neighboring markets, year 2004

	$\overline{\sigma}_{1,24}$	$\overline{\sigma}_{24,24}$	$\overline{\sigma}_{168,24}$
New England	0.0844	0.0676	0.0722
New York	0.1117	0.0837	0.0907
PJM	0.1637	0.1294	0.1343
Ontario	0.2212	0.2813	0.2805

As an illustration, let simply assume that the zero-mean $r_{t,h}$ values follow a normal distribution over the studied time windows. A historical volatility of $\overline{\sigma}_{h,T}$ implies that, on average and with a 95% confidence, one expects that:

$$-2\,\overline{\sigma}_{h,T} \leq \ln \frac{p_t}{p_{t-h}} \leq 2\,\overline{\sigma}_{h,T} \tag{5.12}$$

or

$$p_{t-h}e^{-2\,\overline{\sigma}_{h,T}} \leq p_t \leq p_{t-h}e^{2\,\overline{\sigma}_{h,T}} \tag{5.13}$$

With $\overline{\sigma}_{24,24} = 0.2813$ for the Ontario market in year 2004, prices in a given day could be up to 75.5% higher than the prices in the day before, or they could be 43% lower. These numbers for the New England market are 14.4% and 12.6%, respectively, which reflects a much narrower range for price changes in the New England market.

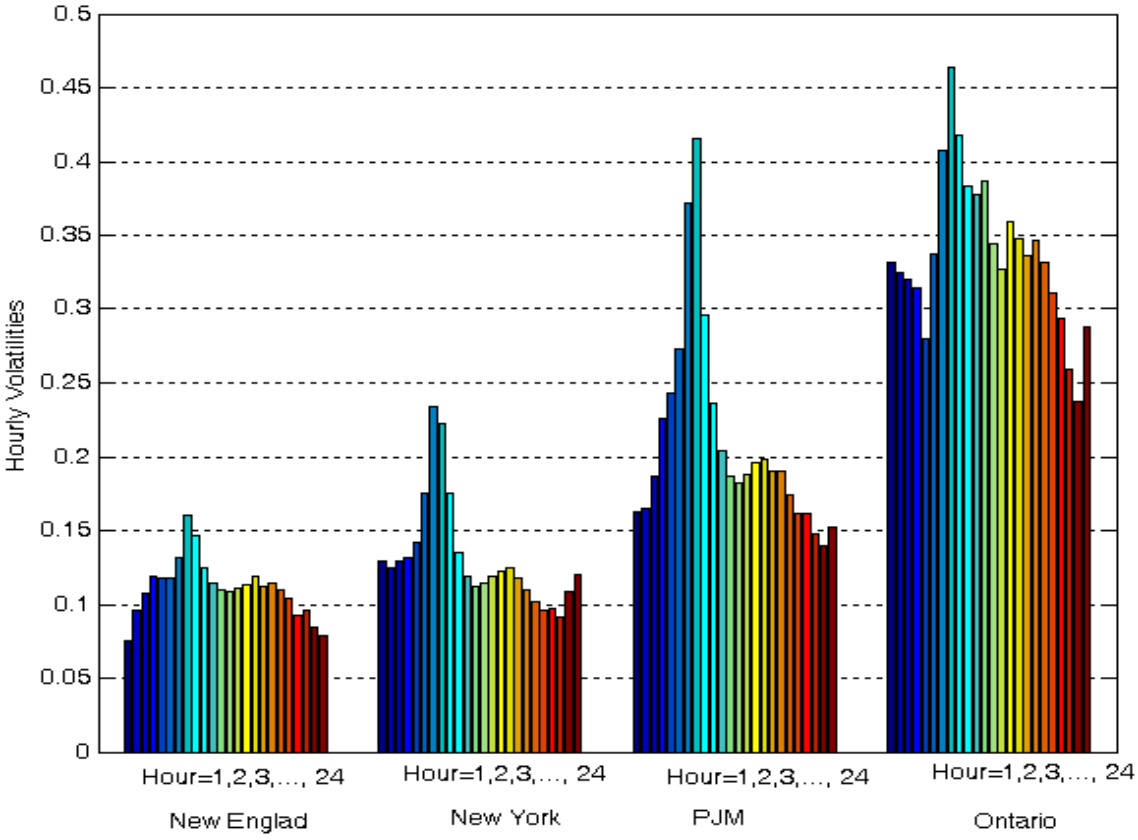

Figure 5.5: Historical volatilities for each hour: Ontario and its neighboring markets.

Scenario 2

The volatility indices for price at each specific hour in the three neighboring markets are presented in Fig.5.5, along with the corresponding results for the Ontario market. This figure illustrates the fact that for each of the 24 hours of a day, Ontario market prices are more volatile than the prices in the three neighboring markets. Observe that prices at hours 7, 8, and 9 in the morning are the most volatile prices for all markets, and the New England market has the lowest volatilities for all hours. In addition, volatility of the PJM market prices at the hours 7, 8, and 9 are somewhat close to those for Ontario; however, Ontario prices are significantly more volatile at other hours.

5.4.2 Price Velocities

The values of $\overline{\text{DVOA}}_h$ and $\overline{\text{DVDA}}_h$, $h = \{1, 24, 168\}$, for the mentioned LMPs were also calculated, and their averages over the studied pricing points are presented in Table 5.4. Observe from these results that the price velocity indices are also higher for the Ontario market than the other three markets.

As a simple illustration, the $\overline{\text{DVOA}}_{24} = 0.2541$ implies that the changes in prices over subsequent days was 25.41% of the 2004 HOEP average; with the 2004 HOEP average being \$49.9/MWh, the average change in HOEP in subsequent days could be up to \$12.7/MWh. The year 2004 average LMP for the New England is US\$52.83/MWh; hence, with a $\overline{\text{DVOA}}_{24} = 0.0976$, the average change in New England day-ahead market LMPs in subsequent days could be up to US\$5.1/MWh, which is less than half of that obtained for Ontario.

Table 5.4: Price velocities for the Ontario and its neighboring markets, year 2004

	$\overline{\text{DVOA}}_1$	$\overline{\text{DVDA}}_1$	$\overline{\text{DVOA}}_{24}$	$\overline{\text{DVDA}}_{24}$	$\overline{\text{DVOA}}_{168}$	$\overline{\text{DVDA}}_{168}$
New England	0.0603	0.0508	0.0976	0.0802	0.1452	0.1282
New York	0.0730	0.0726	0.0932	0.0929	0.1327	0.1283
PJM	0.1129	0.1133	0.1448	0.1489	0.1955	0.1957
Ontario	0.1586	0.1551	0.2541	0.2573	0.2933	0.3005

5.5 Discussion

5.5.1 Market Structure and Price Volatility: The Case of New England

According to the volatility analysis presented in the previous sections, price volatility in the Ontario electricity market is significantly higher than price volatility in the New England, New York, and PJM electricity markets. It should be noted that despite many differences in the detailed market rules and regulations, and in physical characteristics of the supply and demand sides, the New England, New York, and PJM electricity markets are all based on Standard Market Design (SMD) structure [86], which is basically a two-settlement market with nodal pricing. Also, observe that the volatility indices obtained for the New England market are the lowest, very close to those for the New York market. Furthermore, the PJM electricity market has gone through various market expansions during 2004. Finally, the New England market was a real-time market with a region-wide uniform price, similar to the current Ontario market structure, before the implementation of the SMD structure. Therefore, only the New England market is selected for the discussions presented in this section.

The New England wholesale electricity market was launched on May 1, 1999, as a single-settlement real-time market. On March 1, 2003, the SMD structure was implemented, which converted the market structure into a new LMP-based market comprising a day-ahead market and a real-time market. The New England market before the implementation of the SMD structure is referred to as the New England Interim Market. More than 31,000 MW of generation capacity along with imports from Canada and New York State serve the New England market demand with a peak demand of 28,127 MW (2006). From the addition of more than 9,000 MW of new generation capacity comprising gas-fired generation units from 2000 to 2004, cleaner power has been made available with the prices declining through this period by 5.7%. Natural gas-fired generators (about 43%

of the total capacity), and nuclear generators (about 28% of the total capacity), are the major source of power in this market, and the gas-fired units are the most frequent price setters [87].

The current structure of the Ontario electricity market, which is a single-settlement real-time market with a province-wide uniform price, is similar to the New England Interim Market structure. More than 31,000 MW of generation capacity, along with imports from neighboring regions, serve the Ontario demand with a peak demand of 27,005 MW (2006). Coal-fired generators are the most frequent Ontario market price setters, while expensive gas-fired units are the main price setters during extreme demand hours [23]. The total installed generation capacity and peak load in Ontario and New England electricity markets are in the same order.

In order to provide a more detail insight into price volatility in the New England market, the employed volatility indices are calculated for the first three years of the operation of the New England Interim Market, and are presented in Tables 5.5 and 5.6. The three-year averages of the respective volatility indices for Ontario are also presented in these tables for comparison purposes. Observe from these volatility indices that price volatility has been high in the New England Interim Market, and fairly close to the volatility indices obtained for Ontario. On the other hand, and as expected, after implementation of the SMD structure in New England, price volatility indices have declined significantly, as demonstrated by the results presented in Tables 5.4 and 5.3, for the current New England market. These results imply that the real-time nature of the market in Ontario is directly linked with the high levels of electricity price volatility.

5.5.2 Influential Parameters on Price Volatility in Ontario

To explain the high price volatility in Ontario market, consider the following events which frequently happen in this market and are discussed in more detail in [23]:

Table 5.5: Historical volatilities for New England's Interim Market

	$\overline{\sigma}_{1,24}$	$\overline{\sigma}_{24,24}$	$\overline{\sigma}_{168,24}$
New England	0.2261	0.2866	0.2972
Ontario	0.2469	0.3203	0.3222

Table 5.6: Price velocities for New England's Interim Market

	$\overline{\text{DVOA}}_1$	$\overline{\text{DVDA}}_1$	$\overline{\text{DVOA}}_{24}$	$\overline{\text{DVDA}}_{24}$	$\overline{\text{DVOA}}_{168}$	$\overline{\text{DVDA}}_{168}$
New England	0.1483	0.1372	0.2854	0.2563	0.3307	0.3260
Ontario	0.1835	0.1696	0.3083	0.2939	0.3655	0.3557

- Demand underforecast: A demand underforecast error during the peak hours, even in the acceptable range of 1-2%, may force the market operator to dispatch some of the expensive units and thus causing unpredictable price spikes.

- Export/Import transactions failure: Exports and imports are scheduled 1 hour before the dispatch hour in the Ontario market and are considered as fixed load and supply, respectively, in real-time [24]. Any failure in import transactions may force the market operator to instantly dispatch expensive units, which also may cause unusual price spikes. In addition, any failure in export transactions may force the Ontario market operator not to dispatch some of the marginal units, which in turn may cause unusually low prices.

- Error in non-dispatchable generators energy output forecast: In the Ontario market, price-taking self-scheduling generators (e.g., small hydro units) and intermittent generators (e.g., wind farms) forecast their hourly energy output and submit it to the IESO. Analysis of the Ontario market data shows that their real-time available capacity deviates from their forecasted values, sometimes up to more than

250 MW. Similar to the demand forecast error or export/import failure situations, dealing with the discrepancy between the forecasted and actual available capacity of the self-scheduling generators in real-time may cause unusually high or low market prices.

Dealing with such unpredictable, and most of the time unavoidable, events in real-time puts upward or downward pressure on market prices, and hence leads to high market price volatility. Moving toward a two-settlement market similar to the SMD structure is clearly necessary for Ontario, since in a single settlement real-time market, the price volatility resulting from such events affects all market participants. However, in a two-settlement market, most of the eventual real-time demand is cleared in the day-ahead market (on average 97% for the New England market [87] and 90% for New York market [88] in 2004), where no physical transactions take place. With the major part of the market demand cleared 24 hours before real-time, market participants have enough time to arrange for their supply and demand obligations, and in case of unpredictable events, only real-time prices may become volatile with a small group of market players who participate in the real-time market being affected.

5.5.3 Price Volatility and Price Predictability

The findings of Chapter 3 show that HOEP forecasts have a significantly lower level of accuracy than the price forecasts generated for the neighboring markets. Moreover, it is demonstrated in Chapters 3 and 4 that employing various forecasting methods cannot improve the HOEP forecast accuracy significantly. On the other hand, it is shown in the present chapter that the HOEP is significantly more volatile than the other studied electricity market prices. Keeping these findings in mind, the relationship between the volatility of a time series and the accuracy of the forecasts generated for the time series is investigated in this section.

Consider an ARIMA model representing a zero-mean stationary process z_t as:

$$\phi(B)(1-B)^d z_t = \theta(B)\epsilon_t \tag{5.14}$$

where $\phi(B)$, $\theta(B)$, B, and ϵ_t are defined in Section 3.3.1. The model (5.14) can be expressed as an infinite weighted sum of current and previous random shocks ϵ_j, as follows [60]:

$$\begin{aligned} z_t &= \epsilon_t + \psi_1\epsilon_{t-1} + \psi_2\epsilon_{t-2} + ... \\ &= \epsilon_t + \sum_{j=1}^{\infty} \psi_j \epsilon_{t-j} \end{aligned} \tag{5.15}$$

Thus, the relationship between the variance of z_t, i.e., σ_z^2, and the variance of the random shock ϵ_t, i.e., σ_ϵ^2, can be written as:

$$\sigma_\epsilon^2 = \frac{\sigma_z^2}{1 + \sum_{j=1}^{\infty} \psi_j^2} \tag{5.16}$$

On the other hand, the variance of the l-step-ahead forecast error generated by model (5.14), denoted by $\sigma_{e,l}^2$, can be presented as:

$$\sigma_{e,l}^2 = \sigma_\epsilon^2 (1 + \sum_{j=1}^{l-1} \psi_j^2) \tag{5.17}$$

From (5.16) and (5.17), $\sigma_{e,l}^2$ can be written as:

$$\sigma_{e,l}^2 = \frac{1 + \sum_{j=1}^{l-1} \psi_j^2}{1 + \sum_{j=1}^{\infty} \psi_j^2} \sigma_z^2 = \xi(\psi)\sigma_z^2 \tag{5.18}$$

where $\xi(\psi) < 1$ for a finite forecasting horizon l. For a given ARIMA model, $\xi(\psi)$ is a constant depending on the estimated parameters of the model. Hence, high σ_z^2 results in high $\sigma_{e,l}^2$, which means forecast errors can potentially be high. Note that for a stationary time series, the variance is constant and can be represented by the historical volatility measures defined earlier. Similar reasoning can be applied to TF and DR models with similar conclusions. Considering the high volatility of the HOEP time series discussed

in the previous sections and the reasoning presented above, one would expect the HOEP forecasts to have high errors compared to the other neighboring marker prices.

The 4-weekly MAPEs of the forecasts generated for the Ontario and its neighboring electricity markets discussed in Section 3.5.4, along with the intra-day volatilities from Table 5.3 are depicted in % in Fig. 5.6. Observe from this figure that the HOEP volatilities and the HOEP forecasts MAPEs are the highest among the four, as expected.

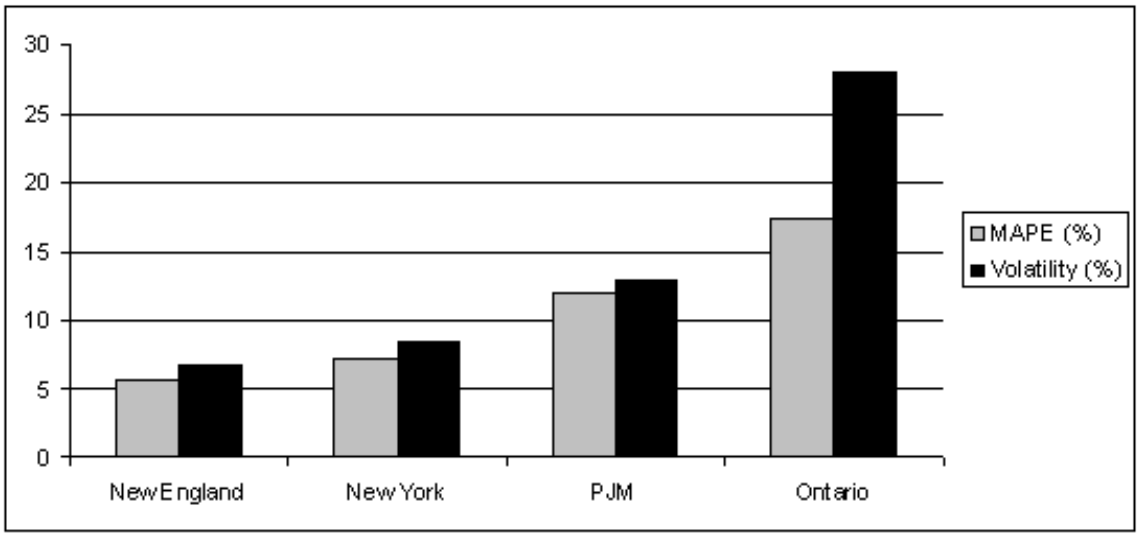

Figure 5.6: Forecast MAPEs and price volatilities ($\overline{\sigma}_{24,24}$) .

5.6 Summary

In this chapter, various volatility indices are developed based on historical volatility and price velocity concepts, and price volatility in the Ontario electricity market is quantified accordingly. The employed volatility indices are also applied to several pricing points in three of Ontario's neighboring electricity markets, namely, the New England, New York, and the PJM markets. Intra-day, trans-day and trans-week market price fluctuations are considered in calculating the volatility indices. The concept of price velocity is extended

for trans-day and trans-week analyses in this study, and trans-week historical returns are also proposed for volatility analysis in this research. It is also formally shown that out-of-sample forecast accuracy of a time series model is affected by the volatility of the time series under study.

The findings of this chapter show that price volatility in the Ontario electricity market is significantly higher than that observed in the neighboring electricity markets. Comparing the volatility indices obtained for Ontario with previously published results for other electricity markets reveals that the HOEP is among the most volatile electricity market prices worldwide. The generally higher error level in HOEP forecasts compared to other market price forecasts can be attributed to the high level of HOEP volatility. It is further argued that the highly volatile nature of the HOEP is a direct result of the real-time operating environment of the Ontario electricity market, and the lack of a hedging mechanism such as the day-ahead market in the SMD structure implemented in Mew England's electricity market.

Chapter 6

Application of Price Forecasts to Short-term Planning of BEMCs[1]

6.1 Introduction

In Chapters 2, 3 and 4, an effort is made to understand the operation of the Ontario electricity market and forecast its future price behavior as accurately as possible. However, the volatility analysis presented in Chapter 5 shows that price forecasting error is inevitable, especially in volatile markets like Ontario's. Moreover, as can be observed from the literature review presented in Chapter 1, most of the work on electricity market price forecasting is focused on improving forecast accuracy, rather than the effects of price forecast inaccuracy on market participants. Hence, there is a need to examine how erroneous market price forecasts affect the participants' planning activities, particularly in the short-term. In view of this, the present chapter addresses the economic impact of using electricity market price forecasts in the operation scheduling of two typical BEMCs.

[1] Based on the findings of this chapter, a paper is in preparation to be submitted to the *IEEE Transactions on Power Systems*.

The first case-study BEMC is a process-industry load with access to an on-site co-generation facility. The second case-study BEMC is a municipal water plant having controllable demand. These two BEMCs represent a considerably significant class of wholesale market customers and have different load management capabilities. Next-day operation of the two BEMCs is formulated as optimization problems which are solved to minimize their expected energy costs.

The Ontario electricity market is the case-market used again, and two sets of HOEP forecasts are considered as the expected future electricity market prices. The first set is the 24-hour-ahead HOEP forecasts generated by the TF models in Chapter 3, which yield the lowest error measures. The second set is the IESO-generated 24-hour-ahead PDPs which yield the highest error measures. For the sake of comparison, the *ex-post* HOEPs are also used in a complementary scenario to determine the "ideal" schedules and their associated costs. The economic impact of using price forecasts for scheduling is analyzed by comparing cost that the BEMC would have incurred if the "ideal" schedules were available and implemented in reality, and the cost associate with implementing the schedules derived based on price forecasts.

6.2 Problem Formulation

6.2.1 Optimization Under Uncertainty

From the demand-side point of view, 'optimal' operation in a competitive electricity market environment constitutes the minimization of total electricity costs. The problem of minimizing electricity costs over a specific planning period (e.g., a day) for a BEMC can be generally formulated as:

$$\begin{aligned} \min \quad & \text{Cost} = \textstyle\sum_k \rho_k \cdot P_k \\ \text{subject to} \quad & \xi(P_k) \end{aligned} \tag{6.1}$$

where k is the index of planning interval, ρ_k is the electricity market price at interval k, P_k is the net power purchased from the market in interval k, and $\xi(P_k)$ is the set of technical constraints. This optimization problem has to be solved before the start of the planning period. However, ρ_ks are unknown random variables which will be cleared after real-time, and hence, (6.1) represent an "uncertain" optimization problem.

The approach to deal with this uncertain problem in the present work is to minimize the *conditional expectation* of total costs; where *conditional* refers to the entire available information with regard to the electricity market price behavior at the optimization origin. This approach has also been used in [46] to derive optimal bidding strategies of a supply-side market participant.

The problem in (6.1) can be reformulated as follows to minimize the *expected* cost of electricity:

$$\begin{aligned} \min \quad & E[\text{Cost}|I] = E[\textstyle\sum_k \rho_k \cdot P_k|I] \\ \text{subject to} \quad & \xi(P_k) \end{aligned} \tag{6.2}$$

where E denotes the mathematical expectation, and I is the available information about electricity market price behavior. It should be noted that the random variables ρ_k only affect the objective function. Hence, by swapping the expectation and summation operators, (6.2) can be written as:

$$\begin{aligned} \min \quad & E[\text{Cost}|I] = \textstyle\sum_k E[\rho_k \cdot P_k|I] \\ & \qquad\quad = \textstyle\sum_k E[\rho_k|I] \cdot P_k \\ \text{subject to} \quad & \xi(P_k) \end{aligned} \tag{6.3}$$

Recalling from the fundamentals of time series analysis and forecasting [57,60], price forecasts generated by a time series model are the conditional expectations of actual prices. In other words, given a time series model for price, $E[\rho_k|I]$=$\hat{\rho}_k$, where $\hat{\rho}_k$ is the forecasted value for ρ_k. Therefore, (6.3) can be solved as an ordinary optimization problem as follows:

$$\begin{aligned} \min \quad & E[\text{Cost}|I] = \textstyle\sum_k \hat{\rho}_k \cdot P_k \\ \text{subject to} \quad & \xi(P_k) \end{aligned} \tag{6.4}$$

The solution of this problem provides the BEMC with the optimal operational schedules which minimize the total *expected* electricity costs.

6.2.2 Forecast Inaccuracy Economic Impact

In order to analyze the economic impact of using price forecast for short-term scheduling, let us consider two price scenarios. In the first scenario, which is a fictitious scenario, it is assumed that the market price forecasts are "exact"; in other words, actual prices are available before real-time. Let denote the solution of the optimization problem (6.4) under this price scenario by P_k^{a}. Hence, if P_k^{a} were available and implemented in reality, the BEMC would have incurred a cost of:

$$\text{C} = \sum_k \rho_k^{\text{a}} \cdot P_k^{\text{a}} \tag{6.5}$$

where ρ_k^a is the actual market price cleared for hour k, and C is the cost associated with implementing the schedule P_k^{a}.

In the second price scenario, it is assumed that a set of price forecasts $\hat{\rho}_k$ is available before real-time and is used to solve the optimization problem (6.4). Let denote the solution of (6.4) when using $\hat{\rho}_k$s as price forecasts by $\hat{P}_k$. The actual electricity cost that the BEMC will pay after implementing $\hat{P}_k$ schedule can be formulated as:

$$\hat{\text{C}} = \sum_k \rho_k^{\text{a}} \cdot \hat{P}_k \tag{6.6}$$

where $\hat{\text{C}}$ is the actual costs incurred by the BEMC when implementing the $\hat{P}_k$ schedule. Note that although $\hat{P}_k$ is determined based on price forecasts, the actual electricity charges are determined based on actual prices ρ_k^{a}.

The Forecast Inaccuracy Economic Loss (FIEI) index is proposed here as:

$$\text{FIEI (\%)} = 100 \times \frac{\hat{\text{C}} - \text{C}}{\hat{\text{C}}} \tag{6.7}$$

Since P_k^{a} and $\hat{P}_k$ are two particular solutions of the optimization problem (6.4), obtained by using two sets of electricity market prices, the FIEI can be positive, negative, or zero. A positive value of FIEI indicates the percentage of the final costs which is attributable to price forecasting errors. In other words, the final incurred electricity cost could be lower by FIEI percent if the price forecasts were "exact". When the FIEI is negative, it basically means that the actual incurred cost are *unexpectedly* lower than the "ideal" cost. A value of zero for FIEI indicates that the actual cost is the same as the "ideal" cost, despite the presence of price forecasting errors.

The value of FIEI can only be found after real-time when the actual market prices are available. Therefore, it can only be used as an *after-the-fact* index to evaluate the overall usefulness of employing electricity market price forecasts for short-term planning.

6.3 Operation of the Case-study BEMCs

The first case-study BEMC considered in this work is a process-industry load having both thermal and electrical energy demand and owning on-site generation facilities (see Figure 6.1). The on-site generation facility is a gas engine, equipped with a heat recovery boiler for Combined Heat and Power (CHP) production. A traditional oil-boiler is also installed to cover the thermal demand.

The gas engine can be employed to generate electricity in cases the electricity market prices are expected to be high. Thermal energy is a by-product of the gas engine that increases its overall energy efficiency and makes on-site generation a viable option. Having an understanding of the future prices, a combination of the gas engine, oil-boiler, and grid electricity can be "optimally" scheduled to minimize total energy costs.

The second case-study BEMC is a municipal water plant with an obligation to meet its hourly water demand day-by-day. The water plant is composed of an inexhaustible potable water source, five pumps, an elevated reservoir, and a main pipeline to convey

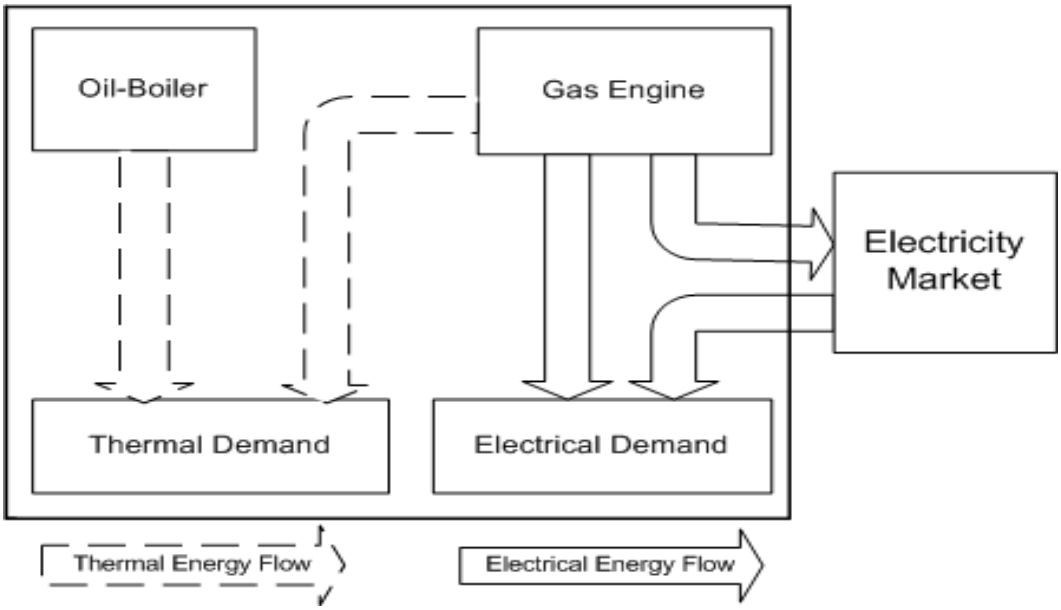

Figure 6.1: The process-industry load.

water from the pumping station to the elevated reservoir, [45] (see Figure 6.2). The water plant is modeled in this research using a simplified mass-balance model in which the nodal pressure requirements are assumed to be satisfied if the water level in the elevated reservoir remains in the desired range [45]. The five constant velocity centrifugal pumps work in parallel and their pumping capacity is assumed to be constant during each hour. Thirty-two possible pump combinations can be considered, from the state in which all pumps are off-line to the state in which all pumps are in service. For this BEMC, pumping operation can be optimally scheduled at low price hours, over a 24-hour planning period, in order to minimize electricity costs.

Operation of the above two case studies is modeled over a 24-hour planning period. A 24-hour planning period is equally divided into 24 planning intervals, denoted here by $k \in \{1, 2, ..., 24\}$. All the scheduled variables are assumed to be constant over the 1-hour planning intervals. The system characteristics and model formulations of the water plant and the process-industry load are taken from [45] and [43], respectively, and are

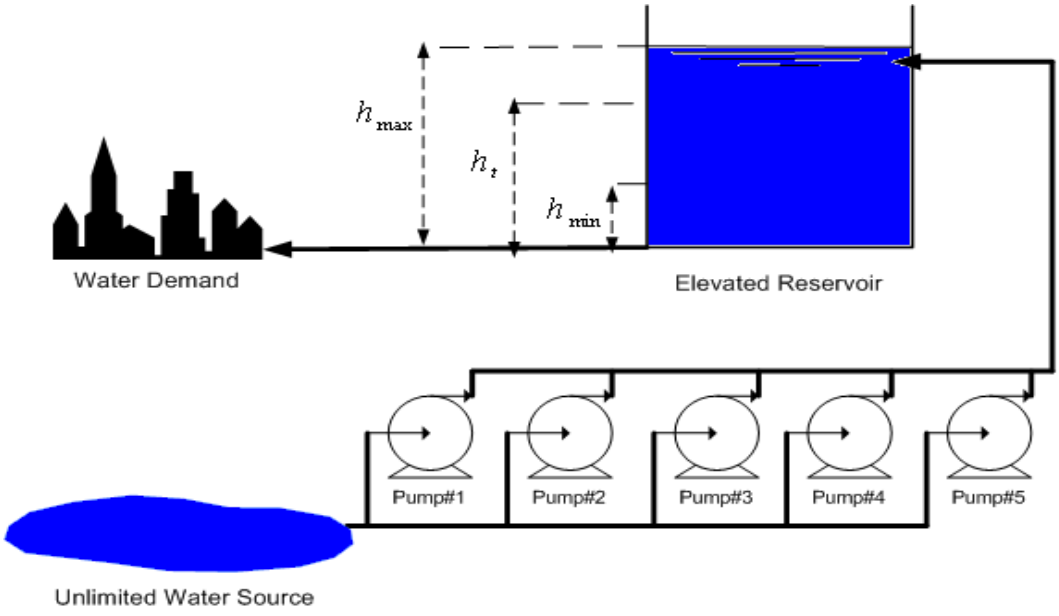

Figure 6.2: The water plant.

provided in detail in the following sections. It is important to highlight the fact that, however, the objective functions for minimizing total expected costs of the BEMCs, and applying constraints to impose emission limitations on the operation of the cogeneration system are novel contributions of the present work.

The following assumptions are made for the analysis:

- The BEMCs carry out their own forecast of the next-day electricity market prices, right before the start of the new day, and plan their operation accordingly.
- No rescheduling or revision of the initially obtained schedules are considered during the planning period.
- The BEMCs are price-taker customers which means that they cannot affect market clearing prices by any strategic behavior.
- The BEMCs' bids to purchase electricity from the market are always cleared.

6.3.1 Formulating the Process-industry Load

Oil Boiler Formulation

The oil-boiler fuel consumption during interval k, denoted with F_k (ton/hour), is represented by a linear function of the thermal energy produced during the hour [43], as follows:

$$F_k = A_1 V_k + B_1 \cdot TP_k^{\text{ob}} \tag{6.8}$$

$$F_{\text{min}} \cdot V_k \leq F_k \leq F_{\text{max}} \cdot V_k \tag{6.9}$$

$$\text{Cost}^{\text{ob}} = (M^{\text{ob}} + \rho^{\text{oil}}) \sum_k F_k \tag{6.10}$$

where A_1 and B_1 are the coefficients of the linear functions which are obtained from the oil-boiler technical performance data; TP_k^{ob} (MW) is the oil-boiler thermal power during interval k; F_{min} and F_{max} are fuel consumption limits of the oil-boiler, Cost^{ob} is the total operation cost of the oil-boiler over the 24-hour planning period; M^{ob} is the operation and maintenance (O&M) costs of the oil-boiler; and ρ^{oil} is the contracted price of oil which remains constant over the planning period. V_k is a binary variable representing oil-boiler status at the planning interval k that can be defined as follows:

$$V_k = \begin{cases} 1 & \text{if the oil-boiler is on.} \\ 0 & \text{if the oil-boiler is off.} \end{cases} \tag{6.11}$$

Gas Engine Formulation

Similarly, the gas engine fuel consumption during interval k, denoted by G_k ($\text{km}^3\text{(N)/h}$)[2], is represented by a linear function of the thermal and electrical energy produced during

[2]Kilo normal cubic meter of natural gas per hour.

that interval [43], as follows:

$$G_k = A_2 \cdot W_k + B_2 \cdot TP_k^{\text{ge}} \tag{6.12}$$

$$G_k = A_3 \cdot W_k + B_3 \cdot EP_k^{\text{ge}} \tag{6.13}$$

$$G_{\text{min}} \cdot W_k \leq G_k \leq G_{\text{max}} \cdot W_k \tag{6.14}$$

$$\text{Cost}^{\text{ge}} = M^{\text{ge}} \sum_k EP_k^{\text{ge}} + \rho^{\text{ng}} \sum_k G_k \tag{6.15}$$

where A_2, B_2, A_3, and B_3 are the linear coefficients which are obtained from the gas engine technical performance data; TP_k^{ge} (MW) is the gas engine thermal power during interval k, EP_k^{ge} (MW) is the gas engine electrical power during interval k; G_{min} and G_{max} are fuel consumption limits of the gas engine; M^{ge} represents the O&M costs of the gas engine; and ρ^{ng} is the fixed natural gas price. The total costs associated with the gas engine operation, i.e., Cost^{ge}, is formulated as a function of the electrical energy produced and the amount of fuel consumed. W_k is a binary variable representing the status of the gas engine during the interval k as follows:

$$W_k = \begin{cases} 1 & \text{if the gas engine is on.} \\ 0 & \text{if the gas engine is off.} \end{cases} \tag{6.16}$$

Gas Engine Carbon Dioxide Emissions

It is assumed that the process-industry load has to maintain its CO_2 emissions, resulting from electricity generation, below a certain specified limit of $EmCap$ (ton/day), as follows:

$$Em^{\text{ge}} \cdot \sum_k EP_k^{\text{ge}} \leq EmCap \tag{6.17}$$

where Em^{ge} is the CO_2 emission of the gas engine per MWh of electrical energy generated.

Electricity Market Transaction Cost

The cost of the net electricity transaction with the market can be presented as:

$$\text{Cost}^{\text{mar}} = (1+\alpha)\sum \rho_k \cdot E_k^{\text{imp}} - \sum \rho_k \cdot E_k^{\text{exp}} \tag{6.18}$$

where E_k^{imp} (MW) and E_k^{exp} (MW) are the electrical energy imported and exported from/to the market during the planning interval k, respectively, and Cost^{mar} is the net cost of electricity transaction with the market. Observe that there is an extra uplift charge, represented by α, associated with the energy imported from the market to account for the network charges and other regulated fees; for example, a 30% uplift charge on top of total electricity costs typically applies to Ontario electricity consumers.

Energy Balance

The thermal demand must be met at all hours by the thermal energy produced either by the oil-boiler or by the gas engine or both. The electrical demand must also be met either by the electricity purchased from the market or the electricity produced by the gas engine or a combination of the two. Hence, the energy balance constraints can be written as follows:

$$TP_k^{\text{ob}} + TP_k^{\text{ge}} = TD_k \tag{6.19}$$

$$E_k^{\text{imp}} + E_k^{l} = ED_k \tag{6.20}$$

$$EP_k^{\text{ge}} = E_k^{l} + E_k^{\text{exp}} \tag{6.21}$$

where TD_k (MW) and ED_k (MW) are the hourly thermal and electrical loads, respectively. E_k^l (MW) is the electric power from the gas engine supplying the local demand during hour k.

Objective Function

The optimization objective is to minimize the total expected energy cost over a 24-hour planning period while meeting all the system constraints defined in (6.8) to (6.21). Thus:

$$\min_{k} \quad E[\text{Cost}^{\text{pil}}|I] = \text{Cost}^{\text{ob}} + \text{Cost}^{\text{ge}} + (1+\alpha)\sum \hat{\rho}_k \cdot E_k^{\text{imp}} - \sum \hat{\rho}_k \cdot E_k^{\text{exp}} \quad (6.22)$$

where $E[\text{Cost}^{\text{pil}}|I]$ is the expected total energy cost of the process-industry load.

The above optimization model is a Linear Mixed-Integer Programming (LMIP) problem and is solved using the well-known CPLEX solver in the GAMS programming environment [89].

After the actual electricity market prices are released, the final energy cost of the process-industry load, denoted by $\hat{\text{Cost}}^{\text{pil}}$ here, can be found as:

$$\hat{\text{Cost}}^{\text{pil}} = \hat{\text{Cost}}^{\text{ob}} + \hat{\text{Cost}}^{\text{ge}} + (1+\alpha)\sum \rho_k^a \cdot \hat{E}_k^{\text{imp}} - \sum \rho_k^a \cdot \hat{E}_k^{\text{exp}} \quad (6.23)$$

where $\hat{E}_k^{\text{imp}}$ and $\hat{E}_k^{\text{exp}}$ are the scheduled energy import and export from and to the market by solving (6.22), and $\hat{\text{Cost}}^{\text{ob}}$ and $\hat{\text{Cost}}^{\text{ge}}$ are the costs of the oil boiler and the gas engine associated with the solution of (6.22). Note that $\hat{\text{Cost}}^{\text{ob}}$ and $\hat{\text{Cost}}^{\text{ge}}$ do not depend on the electricity market prices, and hence will take the same values as those found from solving (6.22).

6.3.2 Formulating the Water Plant

While an alternative source of electrical energy is available in the case of the process-industry load, electricity from the grid is the only source of energy in the case of the water plant. However, the existence of a water reservoir enables the water plant to pump more water during low price hours than during high price hours. Hence, the objective function here is to shift pumping operation such that the expected cost of consumed electricity over a planning period is minimized.

Let us denote the pump index by $i \in \{1, 2, 3, 4, 5\}$; water level in the reservoir at the 1-hour planning interval k by h_k (m); initial water level by h_{ini} (m); water demand at interval k by D_k^w (m^3); the rated power of pump i by P_{max}^i (MW); and the water discharge of pump i by Q^i (m^3). The optimization model can then be formulated as follows:

$$\min \qquad E[\text{Cost}^{\text{wp}}|I] = \sum_k \hat{\rho}_k \sum_i P_{\text{max}}^i U_k^i \tag{6.24}$$

subject to

$$h_k = h_{t-1} + \frac{1}{A}(\sum_i (U_k^i Q^i) - D_k^w) \tag{6.25}$$

$$h_{\text{min}} \leq h_k \leq h_{\text{max}} \tag{6.26}$$

$$h_0 = h_{\text{ini}} \tag{6.27}$$

$$h_{24} \geq h_{\text{ini}} \tag{6.28}$$

where $E[\text{Cost}^{\text{wp}}|I]$ is the total expected electricity cost of the water plant, and U_k^i is a binary variable representing the status of pump i during the planning interval k as:

$$U_k^i = \begin{cases} 1 & \text{if pump } i \text{ is on.} \\ 0 & \text{if pump } i \text{ is off.} \end{cases} \tag{6.29}$$

Equations (6.25) and (6.26) ensure the water level remains in the desired range, and equations (6.27) and (6.28) define the initial level and the desired level at the end of each planning period. Note that for a 1-hour planning interval, the water discharge and electric power consumption of an assigned pump combination is fixed, and the nonlinearities in the combination of the pumps are neglected.

The above optimization model is also a LMIP problem and is solved using the CPLEX solver in the GAMS programming environment [89].

The cost of implementing the above schedules can be calculated when the actual electricity market price data are available, as follows:

$$\hat{\text{Cost}}^{\text{wp}} = \textstyle\sum_k \rho_k^a \sum_i P_{\text{max}}^i \hat{U^i}_k \tag{6.30}$$

where $\hat{\text{Cost}}^{\text{wp}}$ is the water plant's actual electricity cost, and $\hat{U^i}_k$ is the solution of the optimization problem (6.24)-(6.28).

6.4 Numerical Results and Discussion

6.4.1 Data Sets

Typical chronological hourly thermal and electrical demand profiles have been considered here for the process-industry load and are shown in Figure 6.3; the corresponding hour-wise data are also presented in Table B.1 in Appendix B. Various parameter representing the process-industry load are presented in Table 6.1. The operation and maintenance (O&M) costs of the gas engine and the oil-boiler are selected as per available typical values [32, 43]. The average 2004 natural gas and oil prices are obtained from [90], and converted to appropriate units in Canadian dollar[3], and are also presented in Table 6.1. The value of α is considered to be 30%, in line with what has been observed in Ontario.

The water plant is assumed to have access to a reliable hour-by-hour forecast of water demand for the next 24-hour planning period, and actual water demand values do not deviate from the forecasted values significantly. A total daily demand of 54,788 m^3 is considered here, and a typical chronological water demand curve is shown in Figure 6.4. The area of the elevated reservoir is $A = 2600$ m^2, and the maximum and the minimum water levels are $h_{\text{max}} = 7$ m, and $h_{\text{min}} = 1$ m, respectively. The water level at the beginning of a planning period is $h_{\text{ini}} = 3$ m, and it has to be maintained at the end of the planning period, i.e., $h_{24} \geq 3$. The power and discharge capacity of the pumps are provided in Table 6.2.

[3]An average exchange rate of US$1=Can$1.40 is used for 2004.

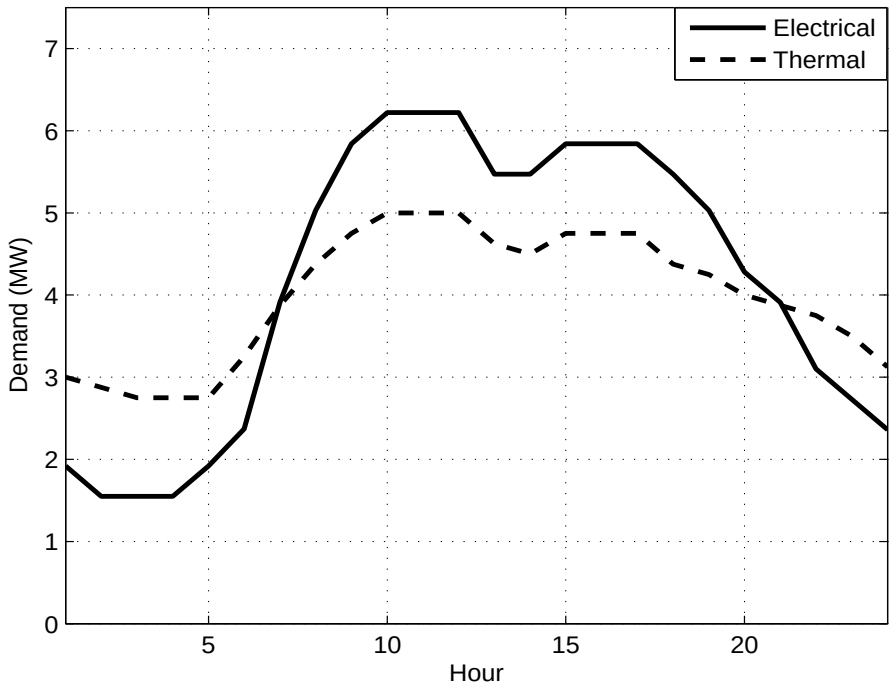

Figure 6.3: Thermal and electrical demand of the process-industry load. Note that thermal demand has been converted to equivalent MW units for the sake of uniformity [43].

6.4.2 Short-term Planning

The optimization problems, developed in Section 6.3, are individually solved for each day of the 6-week study period. Two sets of price forecasts are considered for the simulations: the 24-hour-ahead HOEP forecasts generated by the TF models which were found to yield the lowest errors, and the IESO-generated 24-hour-ahead PDPs which were found to yield the highest error measures. A fictitious scenario is also analyzed here when the corresponding *ex-post* HOEP values are used for scheduling, so that daily FIEI indices may be calculated according to (6.7) to allow for comparisons. The overall economic impact of using price forecasts for scheduling on the case studies is also

Table 6.1: System parameters for the process-industry load (all costs and prices are in average 2004 Canadian Dollar equivalent)

$EP^{ge}_{min} = 0.4$ MW	$EP^{ge}_{max} = 7.0$ MW	$TP^{ge}_{min} = 1.06$ MW
$TP^{ge}_{max} = 5.15$ MW	$TP^{ob}_{min} = 0.4$ MW	$TP^{ob}_{max} = 6.0$ MW
$EmCap = 55$ ton/day	$Em^{ge} = 560$ kg/MWh	ρ^{oil} = \$491/ton
ρ^{ng} = \$410/km^3(N)	M^{ge} = \$10/MWh	M^{ob} = \$2.5/ton
$G_{min} = 0.1604$ km^3(N)/h	$G_{max} = 1.652$ km^3(N)/h	$F_{min} = 0.0406$ ton/h
$F_{max} = 0.6062$ ton/h	$A_1 = 0.0002$ ton/h	$B_1 = 0.101$ ton/MWh
$A_2 = -0.228$ km^3(N)/h	$B_2 = 0.365$ km^3(N)/MWh	$A_3 = 0.07$ km^3(N)/h
$B_3 = 0.226$ km^3(N)/MWh	$TD_{max} = 5$ MW	$ED_{max} = 6.22$ MW

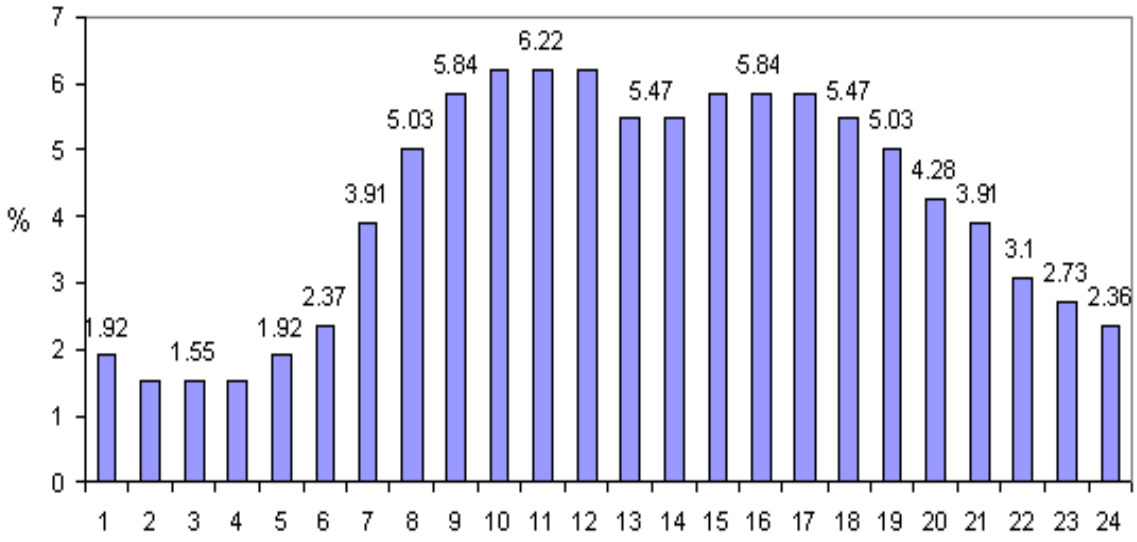

Figure 6.4: Water demand curve of the water plant.

Table 6.2: Technical characteristics of the pumps.

	Discharge (m^3/h)	Power (MW)
Pump 1	1800	0.595
Pump 2	1440	0.445
Pump 3	828	0.260
Pump 4	828	0.260
Pump 5	1800	0.595

Table 6.3: Summary of total operating cost during the six-week period based on the three different price scenarios for the two case-study BEMCs

	The process-industry Load	The Water Plant
Ex-post HOEPs	\$445,500	\$34,172
HOEP forecasts by the TF models	\$454,670	\$35,936
PDPs	\$463,930	\$36,177

assessed using the following six-weekly FIEI index:

$$\text{FIEI (\%)} = 100 \times \frac{\sum_{\text{day}=1}^{42}(\hat{\text{C}} - \text{C})}{\sum_{\text{day}=1}^{42}\hat{\text{C}}} \tag{6.31}$$

The total operation costs under each of the price scenarios are presented in Table 6.3. Observe from this table that the overall cost of operation for both case studies decreases with more accurate price forecasts, as expected. However, the extent of this decrease varies for the case studies, which is discussed later.

The Process-industry Load

Operation of the process-industry load during three typical days, namely, day 5, day 11, and day 39, of the 42-day study period are the basis of the discussions presented in this section. These days are selected for the folowing reasons:

- On day 5, the TF and PDP forecasts have significantly different MAPEs and so are the associated daily FIEI indices;
- on day 11, the PDPs have a significantly higher MAPE than the TF forecasts, but the FIEI index associated with the PDPs is unexpectedly lower than that with the TF forecasts;
- on day 39, the MAPE of the TF forecasts is about 10%, and the associated FIEI index is zero.

The forecast MAPEs and the daily FIEI indices for days 5, 11, and 39 are presented in Table 6.4. Observe from this table that on day 5, the MAPE of the TF forecasts has improved by 12.2%, when compared to that of the PDPs. The daily FIEI index has also significantly improved from 12.6% when using the PDPs to 0.64% when using the TF forecasts. On day 11 on the other hand, while the MAPEs of the both sets of forecasts are on the same order as for day 5, the daily FIEI associated with the TF forecasts (2.81%) is no longer better than the one obtained with the PDPs (1.7%). To explain this point, the operation of the process-industry load should be considered, as discussed next.

The process-industry load purchases electricity from the market if the market price is lower than a certain threshold. This threshold price can simply be found by gradually increasing the electricity market prices in the optimization problem from zero and determining that value beyond which no electricity is purchased from the market. This threshold for the particular studied system is found to be \$53.7/MWh, which means that when the electricity market price is below \$53.7/MWh, all electricity is purchased

Table 6.4: The forecasts MAPEs for days 5, 11, and 39, and the associated daily FIEIs for the process-industry load

	TF Forecasts		Pre-Dispatch Prices	
	FIEI (%)	MAPE (%)	FIEI (%)	MAPE (%)
Day 5	0.6	16.3	12.6	28.5
Day 11	2.81	16.9	1.7	28.7
Day 39	0	9.4	5.18	15.6

from the market. Thus, if the electricity market prices are forecasted to be higher than the threshold price, it would be economical to produce the required electricity locally. Therefore, forecasting the future electricity market prices with respect to the threshold is a crucial factor for the process-industry load.

The energy import schedules on day 5 and 11 by the process-industry load when using the PDPs and the TF forecasts, are shown in Figures 6.5 and 6.6, and Figures 6.7 and 6.8. The actual and forecasted HOEPs, and the "ideal" energy import schedules are also presented in these figures for comparison purposes. Observe in Figure 6.5 that the relative direction of the future prices with respect to the threshold is not correctly predicted by the PDP forecasts at several hours such as hours 10 and 17. At hour 10, the price forecast is lower than \$53.7/MWh, while the actual price is higher than \$53.7/MWh. In contrast, at hour 17, the price forecast is higher than \$53.7/MWh, while the actual HOEP is lower. The direction of actual prices is similarly predicted incorrectly at nine other hours by the PDPs. On the other hand, the price forecasts by the TF models missed the direction of the prices at only four such hours (see Figure 6.6). Therefore on day 5, the daily FIEI index associated with using the TF forecasts is significantly lower than that with the PDPs.

On day 11, the TF forecasts wrongly predicted the market price direction with respect to the threshold for five hours. The PDPs, however, did so for only three hours (see

Figure 6.7 and Figure 6.8). In other words, despite the higher MAPE of the PDPs than the TF forecasts, PDPs have better predicted the relative direction of prices with respect to the threshold; hence, the resulting economic loss of using the PDPs is lower on this day. Note that both sets of forecasts could predict the relative direction of price with respect to the threshold quite well on this day, when compared to day 5, resulting in generally low economic losses and hence low FIEI indices.

For day 39, the daily MAPEs of the forecasts and the resulting FIEI indices are presented in Table 6.4. Observe from this table that despite the presence of 9.4% forecast MAPE for the TF forecasts, the resulting daily FIEI index is 0%. This basically implies that the relative direction of the electricity market prices with respect to the threshold is correctly predicted by the TF forecasts, as demonstrated in Figure 6.9.

The six-weekly FIEI indices for the process-industry load are calculated, using the operation costs given earlier in Table 6.3, and are presented in Table 6.5. The forecast MAPEs are also presented in Table Table 6.5. Observe from this table that the overall six-weekly FIEI index resulting from using the TF forecasts is 2%, half the FIEI value associated with using the PDPs (4%); this is somehow consistent with the difference between the MAPEs of the two sets of price forecasts (16.1% and 40%). This improvement basically implies that, on average, the economic loss associated with using the TF price forecasts is half the loss associated with using the PDPs. For the studied process-industry load, this is equivalent to savings of $77,220 per year.

The Water Plant

In the case of the water plant, two typical days, namely, day 5 and day 39, are selected for discussion. These days are specifically selected to show how using an identical set of price forecasts can result in different economic losses for different customers.

The daily FIEI indices associated with using the TF forecasts and the PDPs for pump scheduling on day 5 and day 39 are presented in Table 6.6. As per these results, the cor-

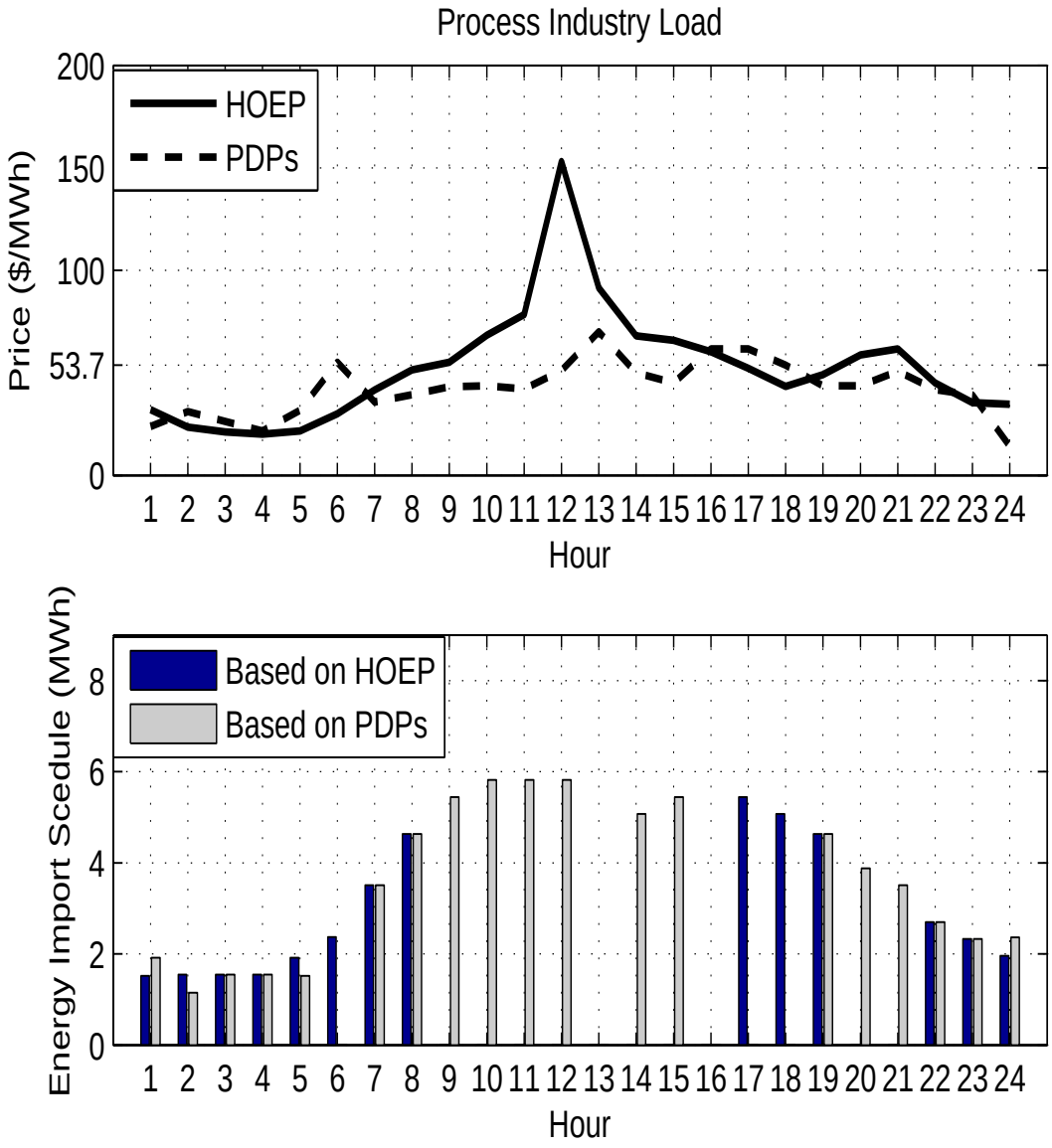

Figure 6.5: Energy imported from the market by the process-industry load during day 5 based on PDP forecasts.

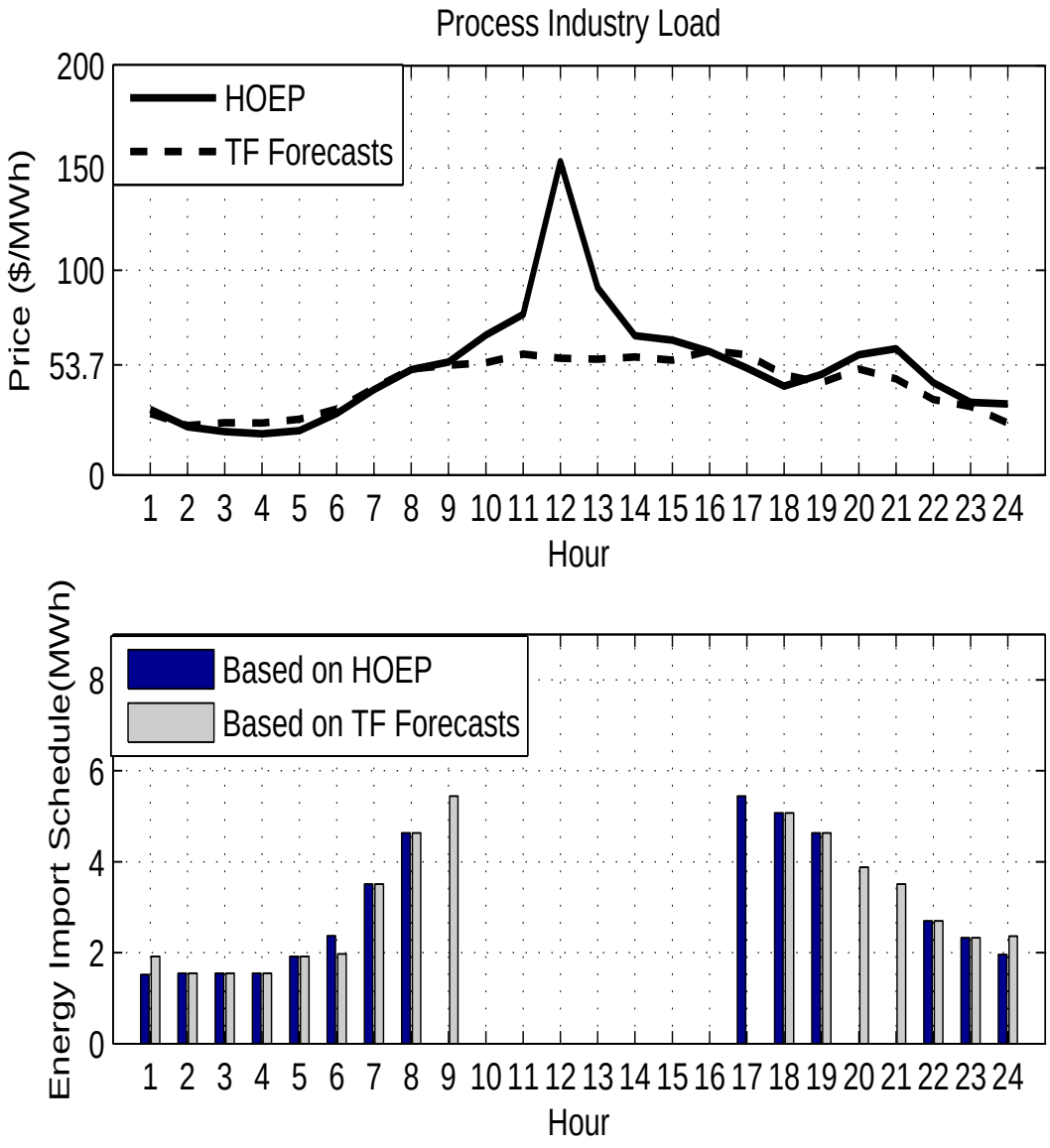

Figure 6.6: Energy imported from the market by the process-industry load during day 5 based on TF forecasts.

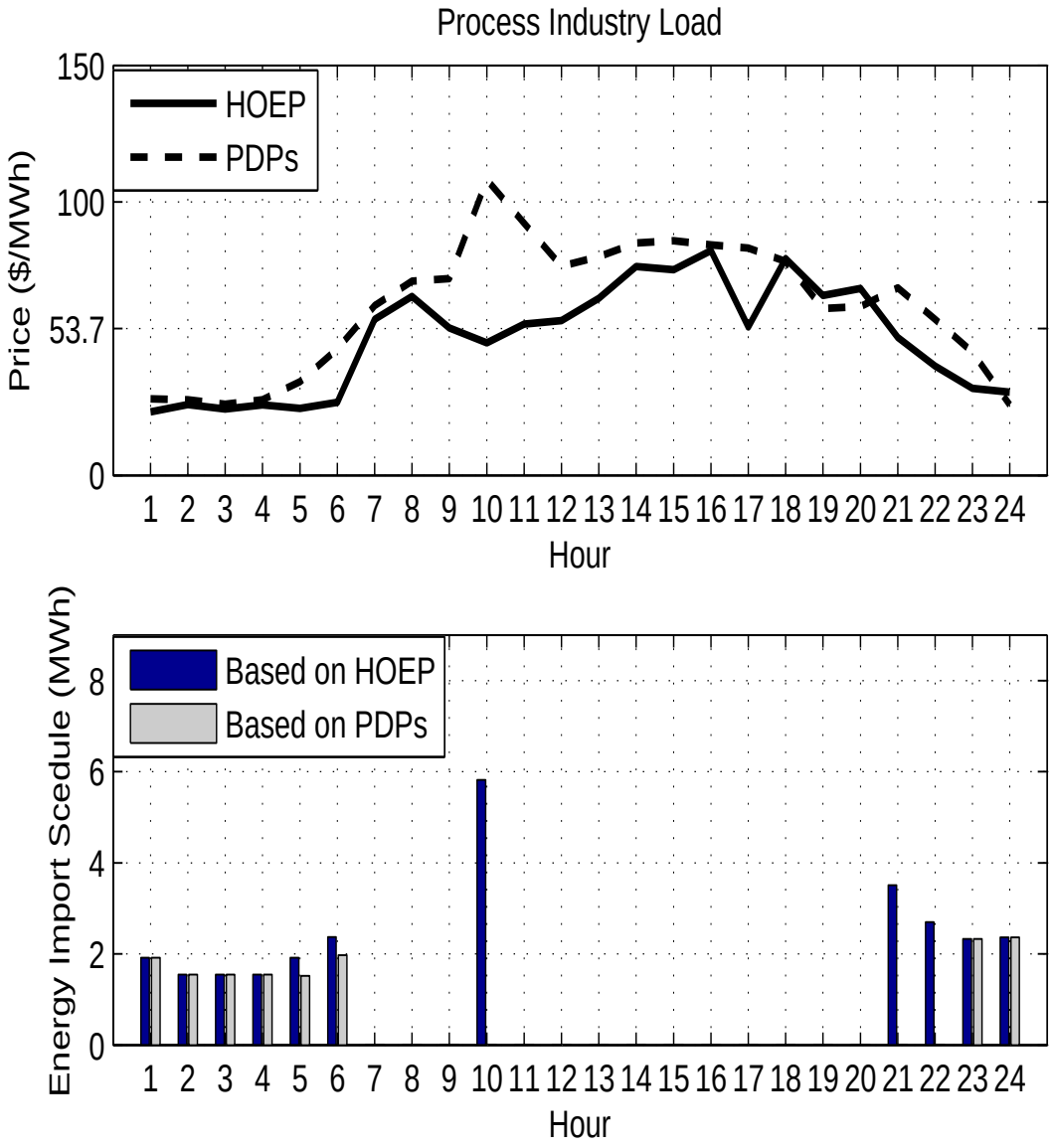

Figure 6.7: Energy imported from the market by the process-industry load during day 11 based on PDPs.

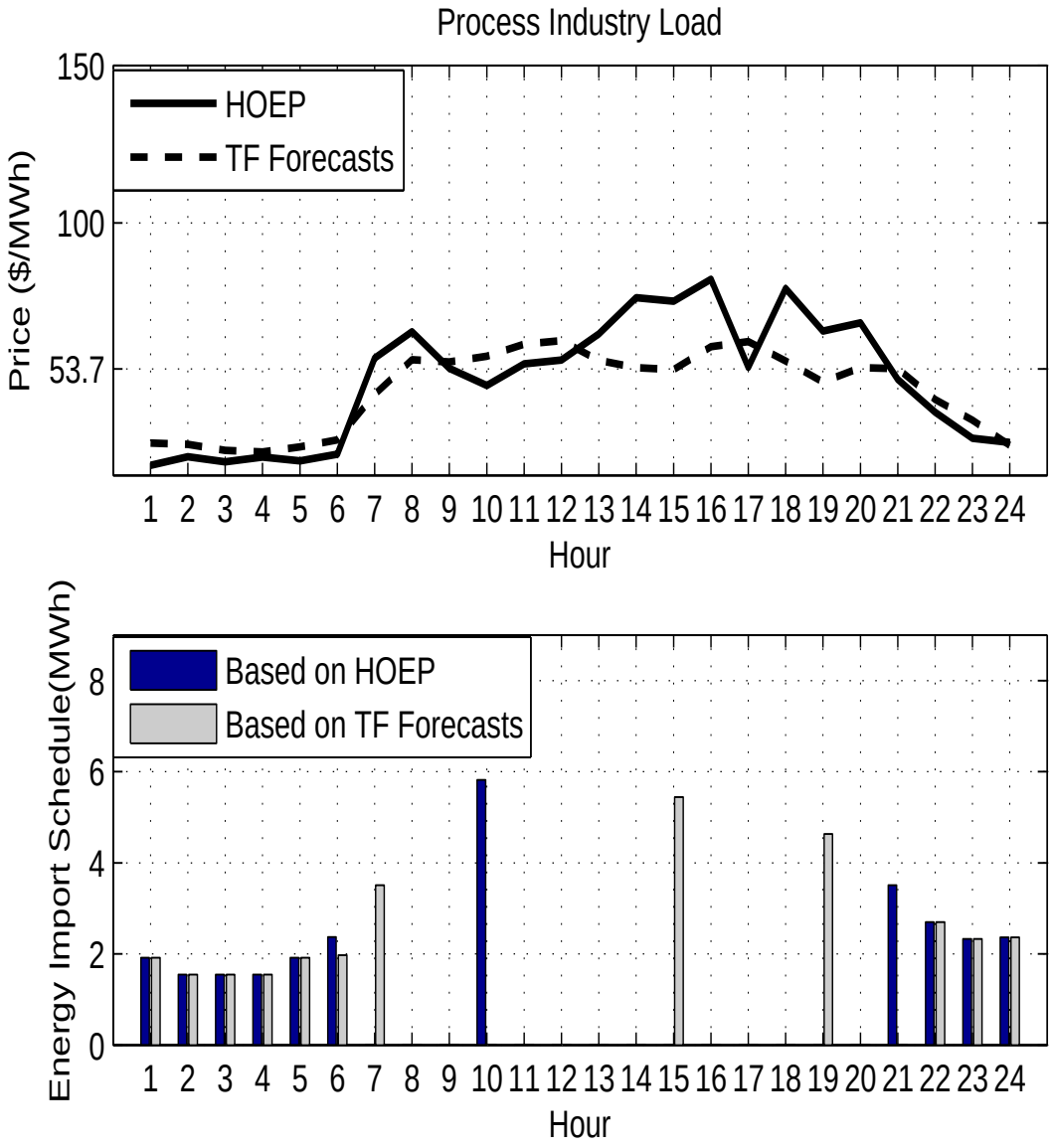

Figure 6.8: Energy imported from the market by the process-industry load during day 11 based on TF forecasts.

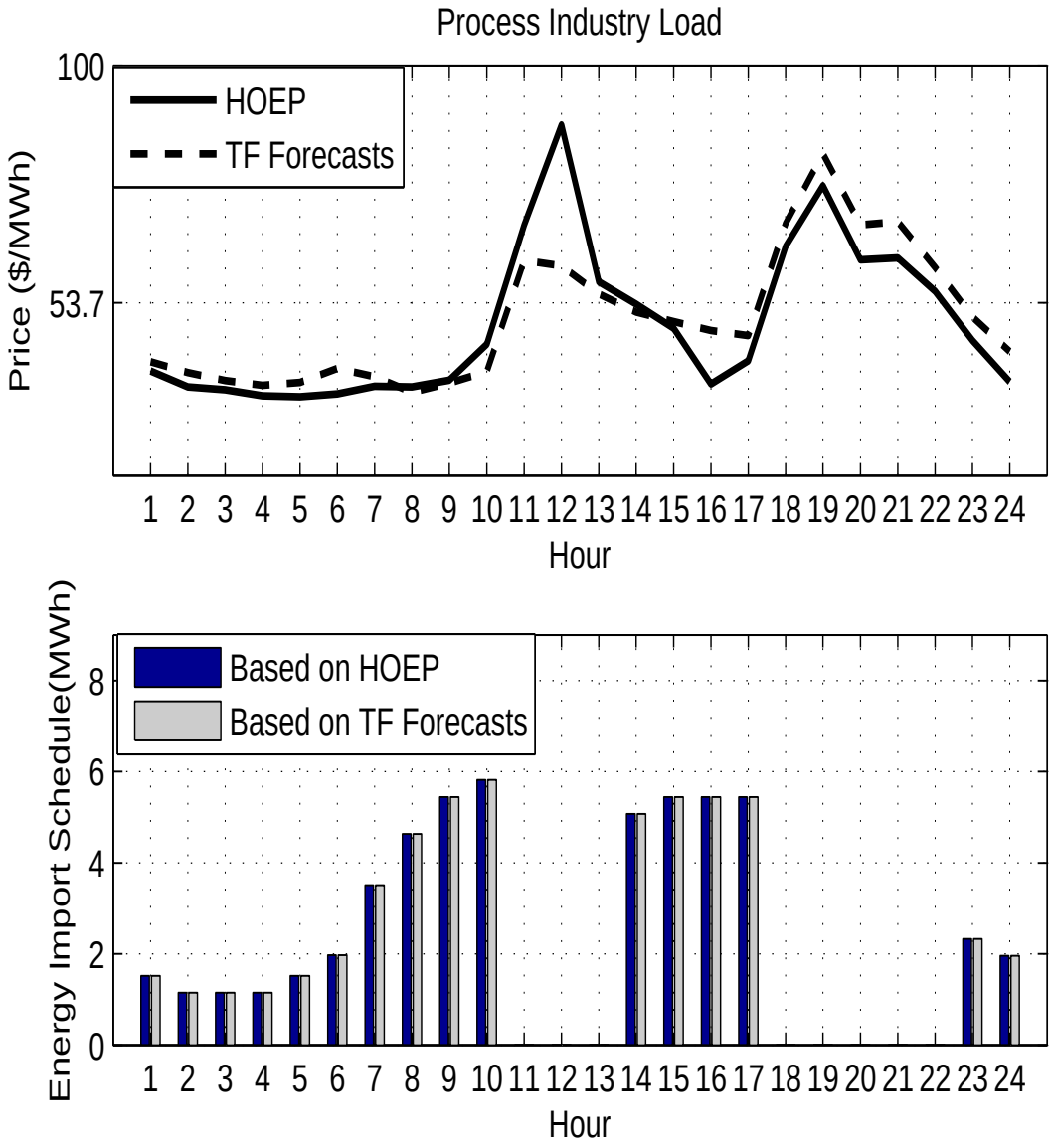

Figure 6.9: Energy imported from the market by the process-industry load during day 39 based on TF forecasts.

Table 6.5: The six-weekly forecasts MAPEs, and the associated FIEI indices for the process-industry load

TF Forecasts		Pre-Dispatch Prices	
FIEI (%)	MAPE (%)	FIEI (%)	MAPE (%)
2.0	16.1	4.0	40.0

responding daily FIEI indices on day 5, are found to be 10.5%, and 11.5%, respectively. These results imply that there is only one percentage point savings if the TF forecasts are used for water plant scheduling, rather than the PDPs. Note in Table 6.4 that the FIEI index improvement by using the TF forecasts is much more significant in the case of process-industry load. Furthermore on day 39, the FIEI index associated with the TF forecasts improves by 12.9% compared to that with the PDPs for the water plant; this improvement was 5.18% in the case of process-industry load (see Tables 6.6 and 6.4). These results highlight the fact that using the generated forecasts on these days has caused different levels of economic loss on the water plant and the process-industry load. In order to explain this difference, the nature of the water plant optimization problem is examined next.

Table 6.6: The forecasts MAPEs for days 5 and 39, and the associated FIEIs for the water plant

	TF Forecasts		Pre-Dispatch Prices	
	FIEI (%)	MAPE (%)	FIEI (%)	MAPE (%)
Day 5	10.5	16.3	11.5	28.5
Day 39	0.7	9.4	13.6	15.6

The solution of the water plant optimization problem is such that the pumping operation is mainly scheduled during the low price hours rather than the high price hours. For example, for a hypothetical two-hour planning period, assume that the actual hourly prices are 10 $/MWh and 80 $/MWh, respectively. Assume further that pumping of water during a single hour can satisfy the total water demand. Regardless of how precise these two prices are forecasted, as long as the forecasted price for the first hour is lower than the forecasted price for the second hour, the pumping will be scheduled at the first hour. Therefore, ability of a forecasting model to predict the general trend of price fluctuations is an important and valuable feature in the case of the water plant, although it might be highly inaccurate in point price forecasting.

The importance of trend forecasting for the water plant is further examined by producing a fictitious set of price forecasts by adding $10/MWh to all actual hourly prices over the 42-day period. The general trend of the fictitious forecasts and the actual hourly prices will hence be exactly identical, although the fictitious forecasts will have a 6-week MAPE of 23.4%. Using this fictitious set of forecasts for pump scheduling over the six-week study period results in a 6-week FIEI index of as low as 0.1%. This low FIEI index supports the aformentioned argument that forecasting the general trend is more important in the case of water plant scheduling.

The hourly electricity consumptions associated with the pumping schedules obtained for day 5 by using the TF forecasts and the PDPs are presented in Figure 6.10 and Figure 6.11, respectively. Observe from these plots that, although the TF forecasts and the PDPs are different in terms of point price forecasting, they both have predicted the general trend of the HOEP fairly well. This in turn has resulted in the close FIEI indices as discussed above (10.5%, and 11.5%). Similar explanation applies to day 39.

The six-weekly FIEI values associated with the TF forecasts and the PDPs and the corresponding forecasts MAPEs for the water plant are presented in Table 6.7. Observe from this table that the six-week FIEI index associated with using the TF forecasts for

Table 6.7: The bi-weekly and six-weekly forecasts MAPEs, and the associated FIEIs for the water plant

TF Forecasts		Pre-Disptch Prices	
FIEI (%)	MAPE (%)	FIEI (%)	MAPE (%)
4.9	16.1	5.5	40.0

pump scheduling is 4.9%, versus the FIEI of 5.5% associated with using the PDP forecasts; an improvement of only 0.6% achieved. These findings further support the argument that a particular set of price forecasts may result in significantly different economic impacts when used by different BEMCs.

The economic impacts of load forecasting inaccuracy have been previously addressed for various supply-side entities [47, 48, 49]. These studies have been conducted in the context of the vertically integrated power systems, under which energy utilities are required to optimally schedule their own resources to meet their demand. The economic impact in these studies was defined as the difference between the system costs had the actual hourly loads been known, and the system costs when using the load forecasts. The reported economic impact analyses show that while the savings associated with load forecasting accuracy vary across the studied systems, a high linear correlation exists between the load forecasting error and the economic impact. On average, improving the load forecast accuracy by 1% has led to savings of about 0.3% of the total incurred costs [48]. Considering the discussions presented in this section, such correlations is not observed in the case of price-forecasting problem.

It is worth mentioning that if the price forecasts generated for the New England electricity market in Section 3.5.4 were used for scheduling the case studies BEMCs, the overall 4-weekly FIEI indices found for the process-industry load and the water plant

would be 0.12% and 0.28%, respectively. This highlights the fact that operation planning based price forecasts is very efficient in markets where there is a high level of price predictability.

6.5 Summary

In this chapter, scheduling the short-term operation of BEMCs by using electricity market price forecasts is discussed. Next-day operation of two different case studies, i.e., a process-industry load and a water plant, is formulated and optimal schedules are derived to minimize their expected total energy costs. These two case studies have different load management capabilities and represent a considerable number of BEMCs. The 24-hour-ahead HOEP forecasts generated by the TF models in Chapter 3, which yielded the lowest error measures, and the IESO-generated 24-hour-ahead PDPs, which yielded the highest error measures are used as the expected future electricity market prices. The economic impact of price forecast inaccuracy is quantified by introducing the novel FIEI index.

The findings of this chapter demonstrate that electricity market price forecasts can be effectively employed for short-term scheduling. However, sensitivity to price forecast inaccuracy varies depending on the characteristics of the system under study. In other words, economic impact of using a particular set of price forecasts for short-term scheduling can significantly differ across various market customers; it can be high for one customer and at the same time, low for another. This implies that "accurate" price forecasting has different meanings for different market participants. Unlike the load-forecasting problem, a linear correlation between the traditional error measures and the economic impact of forecast inaccuracy is not found to exist for the price-forecasting problem.

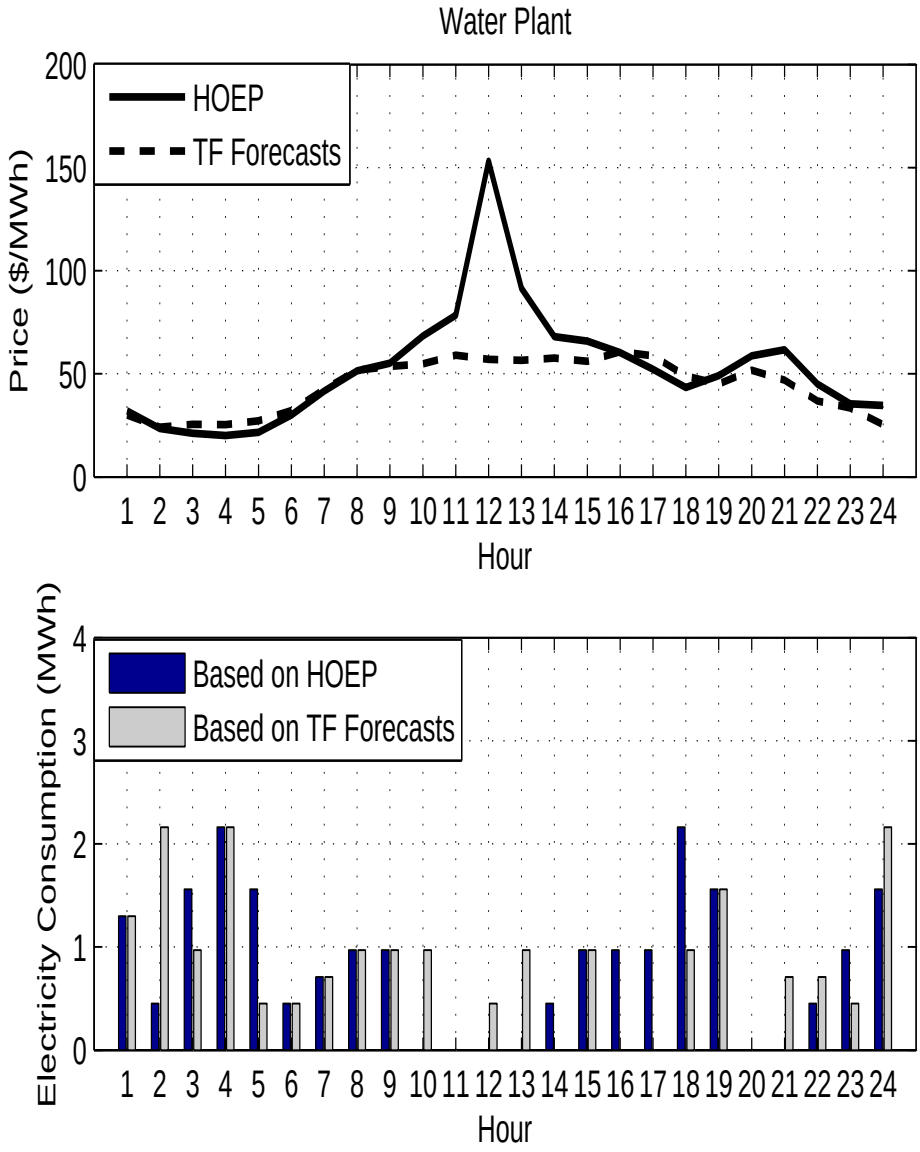

Figure 6.10: Energy consumed by the water plant during day 5 based on the TF forecasts.

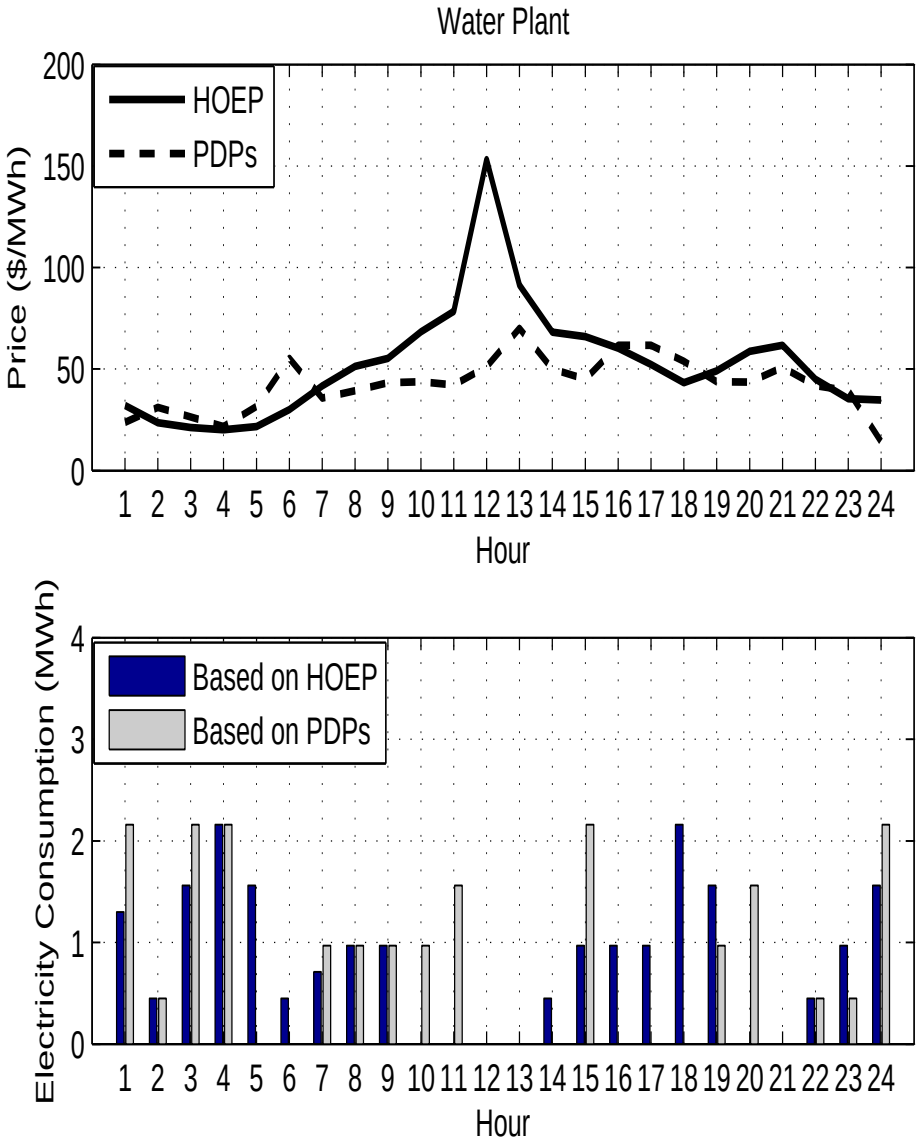

Figure 6.11: Energy consumed by the water plant during day 5 based on the PDPs.

Chapter 7

Conclusions

7.1 Summary and Conclusions

This thesis concentrates on forecasting of electricity market prices and applying the price forecasts for short-term operation planning of demand-side Bulk Electricity Market Customers (BEMCs). The structure of a case market, which is chosen to be the Ontario electricity market in this thesis, is studied in detail, and a set of explanatory variable candidates that may explain price behavior in this market is selected. Various linear and non-linear models are developed to relate electricity market price behavior in Ontario to these explanatory variable candidates. Forecasting models are also developed for the day-ahead prices in three neighboring electricity markets, namely, New York, New England, and PJM, and price predictability is compared across the studied markets. The observed differences in the accuracy of the models developed for the market prices are explained by conducting a comprehensive volatility analysis. The generated price forecasts are used for short-term scheduling of two typical case study BEMCs. Economic impact of price forecast inaccuracy on the case studies is analyzed by devising the novel Forecast Inaccuracy Economic Impact (FIEI) index.

In Chapter 2, a detailed overview of the operation of the Ontario electricity market is presented. Further in this chapter, the Ontario market outcomes, such as market prices for energy and operating reserves and demand, are analyzed for the first four years of market operation. The programs implemented by the Ontario Independent Electricity System Operator (IESO) to improve market efficiency are also described in this chapter and their effectiveness is discussed. This chapter provides a comprehensive insight into the operational aspects of the Ontario electricity market.

In Chapter 3, a wide range of market data from Ontario and its neighboring electricity markets are investigated and a final set of explanatory variable candidates which may explain electricity market price behavior in Ontario is selected. Direct and indirect effects of these variables on the Hourly Ontario Energy Price (HOEP) have been taken into account in the process of variable selection. In addition, issues such as, availability of explanatory variables before real-time and the choice of appropriate forecasting horizon, that are of practical significance, are considered. The multivariate Transfer Function (TF) and Dynamic Regression (DR) time series models are used to relate HOEP behavior to the selected explanatory variable candidates. Two forecasting horizons, i.e., 3 hours and 24 hours, are considered for building the multivariate HOEP models, taking into account the ability of market participants to react to price forecasts. Univariate Auto Regressive Integrated Moving Average (ARIMA) models are also developed for the HOEP. The novel concept of Predicted Supply Cushion (PSC) is introduced and employed as an explanatory variable candidate in this study, and the problem of multicollinearity among the explanatory variable candidates is addressed by a two-step model building procedure. In this chapter, the generated HOEP forecasts have significantly lower error measures than any other available forecast, and several issues with significant practical importance are discussed.

In Chapter 4, two non-linear modeling approaches, i.e., Multivariate Adaptive Regression Splines (MARS) and Multi-Layer Perceptron (MLP) neural networks, are examined for HOEP forecasting. MARS models are developed for the HOEP consider-

ing a variety of explanatory variables, and the 3-hour and 24-hour forecasting horizons. MARS models are also developed based on the historical HOEP behavior in a univariate scenario. Multivariate MLP networks are developed by considering numerous scenarios in terms of MLP structure and inputs. The informative explanatory variables detected by the TF, DR, and MARS models are examined as MLP inputs in identical scenarios. In this chapter, the application and advantages of using non-linear MARS approach for price forecasting are presented, and it is found that the other employed modeling approaches outperform the MLP networks for HOEP forecasting.

In Chapter 5, a comprehensive price volatility analysis is carried out to explain the high level of error in the HOEP forecasts and the differences observed in price predictability across the Ontario and its three neighboring electricity markets. Previously reported volatility measures are extended and new indices are formulated based on historical volatility and price velocity concepts. These indices are applied to the intra-day, trans-day, and trans-week market price fluctuations to analyze the volatility of the studied electricity market prices. The relatively high error level of the HOEP forecasts obtained in previous chapters, compared to the price forecasts obtained for other markets, is explained in this chapter by showing that the HOEP is significantly more volatile than the other studied prices.

In Chapter 6, the application of price forecasts to short-term scheduling of BEMCs is presented. Next-day operation of two case study BEMCs, i.e., a process-industry load and a municipal water plant, is formulated and optimal operation schedules are generated so as to minimize their expected energy costs. The case studies have different load management capabilities and represent a significant number of small-sized demand-side market participants. The HOEP forecasts generated by the TF models, which have the lowest error measures, and the IESO-generated Pre-Dispatch Prices (PDPs), which have the highest error measures, are used in the optimization models as the expected future electricity market prices. Economic impact of using uncertain price forecasts for short-term scheduling is analyzed by devising the novel FIEI index. It is demonstrated

in this chapter that a particular set of price forecasts may lead to significantly different economic impacts for different BEMCs, and "accurate" price-forecasting has different meanings depending on BEMCs' characteristics.

7.2 Contributions

The focus of this thesis is on two main issues:

- Developing the most accurate electricity market price forecasting models that are practically realizable and feasible from the participants' viewpoint.
- Examining the application of price forecasts to short-term operation scheduling of BEMCs, and studying the associated economic impacts.

This research is novel because it deals with the practical implementation of price forecasting, and addresses the problem of short-term operation scheduling of price-responsive electricity consumers within the context of competitive electricity markets.

The following are the highlights and main contributions of this thesis:

1. A unique and comprehensive overview of the Ontario electricity market is presented, providing a clear picture of the operational aspects and performance of this market.
2. Forecasting models are developed by considering a wide range of publicly available data from the case market and its neighboring electricity markets. Various important issues are taken into account when building the price models; these include market time-line and data availability before real-time, and the choice of forecasting horizon based on the participants' ability to react to the price forecasts.

3. The linear multivariate TF and DR time series models are applied to the problem of HOEP forecasting, significantly improving the accuracy of HOEP forecasts compared to other reported forecasts.

4. Well-established multivariate approaches, i.e., the linear TF and DR, and the non-linear MARS and ANNs, are employed to relate HOEP behavior to a broad set of explanatory variable candidates, demonstrating the high level of uncertainty and the lack of reliable information content in the large set of public data available on the Ontario electricity market. Furthermore, it is shown that the final set of informative explanatory variables, as well as model accuracy, varies by forecasting horizon.

5. The problem of model-instability for the HOEP is highlighted, underlining the difficulties of developing practical HOEP-forecasting tools.

6. The non-linear MARS approach is applied to electricity market price forecasting, demonstrating its modeling advantages which are of significant importance in case of model-instability.

7. A volatility analysis is conducted, demonstrating that the HOEP is significantly more volatile than the prices in the neighboring day-ahead markets and showing that it is among the most volatile electricity market prices worldwide. Furthermore, it is proved that out-of-sample forecast accuracy of a time series is affected by the time series' volatility. These observations explain the high level of HOEP forecast inaccuracy and the differences observed in price predictability across the Ontario and its neighboring electricity markets.

8. By considering two typical BEMCs with different load management capabilities, it is shown that short-term electricity market price forecasts can be effectively used for operation planning. Furthermore, by highlighting the technical differences between the two studied BEMCs, it is demonstrated that sensitivity to price forecast

inaccuracy varies significantly for different market participants. The novel FIEI index is also introduced in order to quantify the economic impact of price forecast inaccuracy on the BEMCs.

9. While the focus of the mainstream price forecasting research is on improving electricity price forecasting accuracy in terms of traditional error measures, this work examines how such improvements can help market participants with improving their operation planning. It is argued that improving the accuracy of price forecasts in terms of traditional error measures (e.g., MAPE) does not necessarily always guarantee economic benefits for all market participants.

7.3 Directions for Future Work

Based on the research work presented and discussed in this thesis, further research may be pursued on the following subjects:

- Designing new customer-specific forecasting models, and defining new error measures to assess and compare their forecast accuracy.
- Studying the application of price forecasts for short-term operation scheduling of other types of market participants, such as non-dispatchable small-sized hydro generators with limited hydro resources.
- Application of price forecasts for operation planning of price-setter market participants.
- Scheduling market participants based on price forecasts in a two-settlement market environment, considering different level of risks associated with each of the day-ahead and real-time market prices.

- Operation scheduling while revising the schedules based on the forecasts obtained by models with short forecasting horizons.

Appendix A

Sample MARS Models

The developed MARS model to forecast HOEP values during Week_1 in SCN_1:

$$\begin{aligned}
\text{HOEP}_t = {} & 22.69 - 0.4950 \max(\text{HOEP}_{t-1} - 90.82, 0) \\
& - 0.1269 \max(104.91 - \text{HOEP}_{t-169}, 0) \\
& + \max(\text{HOEP}_{t-24} - 23.49, 0) \\
& - 0.1870 \max(110.49 - \text{HOEP}_{t-49}, 0) \\
& - 0.1245 \max(\text{HOEP}_{t-48} - 25.192, 0) \\
& + \max(\text{HOEP}_{t-144} - 25.481, 0) \\
& - 1.1766 \max(\text{HOEP}_{t-1} - 49.23, 0) \\
& + \max(\text{HOEP}_{t-1} - 30.14, 0) \\
& + 1.12 \max(55.21 - \text{HOEP}_{t-120}, 0) \\
& + \max(\text{HOEP}_{t-121} - 33.21, 0) + \epsilon
\end{aligned} \tag{A.1}$$

The developed MARS model to forecast HOEP values during Week_1 with a forecasting horizon of 3 hours in SCN_2:

$$
\begin{aligned}
\text{HOEP}_t = {} & 61.6 - 0.5456 \max(92.1 - \text{HOEP}_{t-1}, 0) + \max(x_{1,t} - 21.22, 0) \\
& + \max(\text{HOEP}_{t-120} - 97.11, 0) - 0.2651 \max(94.73 - \text{HOEP}_{t-120}, 0) \\
& + \max(48.11 - \text{HOEP}_{t-121}, 0) + \max(x_{2,t} - 20702.0, 0) \\
& - 0.0012 \max(21434.0 - x_{2,t-1}, 0) - 0.0070 \max(x_{2,t-25} - 19700.0, 0) \\
& + \max(19501.0 - x_{2,t-25}, 0) + \max(x_{2,t-24} - 18201.0, 0) \\
& - 1.8860 \max(\text{HOEP}_{t-169} - 89.13, 0) + 1.2536 \max(83.24 - x_{1,t-25}, 0) \\
& + \max(\text{HOEP}_{t-48} - 19.13, 0) + \max(x_{1,t-24} - 56.2345, 0) \\
& - 0.7416 \max(\text{HOEP}_{t-144} - 69.7, 0) - 0.2110 \max(\text{HOEP}_{t-120} - 57.41, 0) \\
& + \max(\text{HOEP}_{t-144} - 58.5, 0) - 0.1010 \max(\text{HOEP}_{t-49} - 16.86, 0) \\
& - 0.4512 \max(x_{1,t} - 94.1, 0) + \max(97.0 - x_{1,t-1}, 0) \\
& + \max(\text{HOEP}_{t-1} - 71.21, 0) - 0.4520 \max(\text{HOEP}_{t-24} - 47.50, 0) \\
& + \max(\text{HOEP}_{t-24} - 41.210, 0) + \epsilon
\end{aligned}
\tag{A.2}
$$

Appendix B

Data

Table B.1: The thermal and electrical demand values for the process industry load

Hour	Electrical Demand (MW)	Thermal Demand (MW)
1	1.92	3.0
2	1.55	2.875
3	1.55	2.75
4	1.55	2.75
5	1.92	2.75
6	2.37	3.25
7	3.91	3.875
8	5.03	4.375
9	5.84	4.75
10	6.22	5.0
11	6.22	5.0
12	6.22	5.0

Continued on next page

Table B.1 – continued from previous page

Hour	Electrical Demand (MW)	Thermal Demand (MW)
13	5.47	4.625
14	5.47	4.5
15	5.84	4.75
16	5.84	4.75
17	5.84	4.75
18	5.47	4.375
19	5.03	4.25
20	4.28	4.0
21	3.91	3.875
22	3.1	3.75
23	2.73	3.5

Bibliography

[1] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: A review and evaluation," *IEEE Transactions on Power Systems*, vol. 16, no. 1, pp. 44–55, Feb. 2001.

[2] S. Vucetic, K. Tomsovic, and Z. Obradovic, "Discovering price-load relationships in California's electricity market," *IEEE Transactions on Power Systems*, vol. 16, no. 2, pp. 280–286, May 2001.

[3] M. Benini, M. Marracci, P. Pelacchi, and A. Venturini, "Day-ahead market price volatility analysis in deregulated electricity markets," in *IEEE Proc. PES Summer Meeting*, vol. 3, 21-25 July 2002, pp. 1354–1359.

[4] A. Breipohl, "Electricity price forecasting models," in *IEEE Proc. PES Winter Meeting*, vol. 2, 27-31 Jan. 2002, pp. 963–966.

[5] I. Simonsen, "Volatility of power markets," *Physica A: Statistical Mechanics and its Applications*, vol. 335, no. 1, pp. 10–20, September 2005.

[6] J. Bastian, J. Zhu, V. Banunarayanan, and R. Mukerji, "Forecasting energy prices in a competitive market," *Computer Applications in Power*, vol. 12, no. 3, pp. 40–45, July 1999.

[7] M. Shahidehpour, H. Yamin, and Z. Li, *Market Operations in Electric Power Systems*. New York: Wiley Interscience, 2002.

[8] A. Wang and B. Ramsay, "Prediction of system marginal price in the UK Power Pool using neural networks," in *Proc. Intl. Conf. on Neural Networks*, vol. 4, 9-12 June 1997, pp. 2116–2120.

[9] B. R. Szkuta, L. A. Sanabria, and T. S. Dillon, "Electricity price short-term forecasting using artificial neural networks," *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 851–857, Aug. 1999.

[10] J. Guo and P. B. Luh, "Selecting input factors for clusters of Gaussian radial basis function networks to improve market clearing price prediction," *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 665–672, May 2003.

[11] L. Zhang, P. B. Luh, and K. Kasiviswavathan, "Energy clearing price prediction and confidence interval estimation with cascaded neural networks," *IEEE Transactions on Power Systems*, vol. 18, no. 1, pp. 99–105, Feb 2003.

[12] J. Guo and P. Luh, "Improving market clearing price prediction by using a committee machine of neural networks," *IEEE Transactions on Power Systems*, vol. 19, no. 4, pp. 1867–1876, November 2004.

[13] L. Zhang and P. Luh, "Neural network-based market clearing price prediction and confidence interval estimation with an improved extended Kalman filter method," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 59–66, February 2005.

[14] A. J. Conejo, J. Contreras, R. Esanola, and M. A. Plazas, "Forecasting electricity prices for a day-ahead pool-based electric energy market," *Int. Jour. of Forecasting*, vol. 21, no. 3, pp. 435–462, July-May 2005.

[15] C. P. Rodriguez and G. J. Anders, "Energy price forecasting in the Ontario competitive power system market," *IEEE Transactions on Power Systems*, vol. 19, no. 1, pp. 366–374, Feb 2004.

[16] J. Contreras, R. Espinola, F. Nogales, and A. Conejo, "ARIMA models to predict next-day electricity prices," *IEEE Transactions on Power Systems*, vol. 18, no. 3, pp. 1014–1020, Aug. 2003.

[17] A. Conejo, M. Plazas, R. Espinola, and A. Molina, "Day-ahead electricity price forecasting using the wavelet transform and ARIMA models," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 1035–1042, May 2005.

[18] J. C. Cuaresma, J. Hlouskova, S. Kossmeier, and M. Obersteiner, "Forecasting electricity spot-prices using linear univariate time-series models," *Applied Energy*, vol. 77, no. 1, pp. 87–106, January 2004.

[19] F. Nogales, J. Contreras, A. Conejo, and R. Espinola, "Forecasting next-day electricity prices by time series models," *IEEE Transactions on Power Systems*, vol. 17, no. 2, pp. 342–348, May 2002.

[20] F. J. Nogales and A. J. Conejo, "Electricity price forecasting through transfer function models," *Journal of the Operational Research Society*, pp. 1–7, 2005.

[21] A. Gonzalez, A. Roque, and J. Garcia-Gonzalez, "Modeling and forecasting electricity prices with input/output hidden Markov models," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 13–24, Feb. 2005.

[22] R. Garcia, J. Contreras, M. van Akkeren, and J. Garcia, "A GARCH forecasting model to predict day-ahead electricity prices," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 867–874, May 2005.

[23] MSP, *Monitoring reports on the IESO-administrated electricity markets*. The Market Surveillance Panel, Ontario, 2002-2005, available [Online] at www.ieso.ca/imoweb /marketSurveil/mspReports.asp.

[24] H. Zareipour, C. Canizares, and K. Bhattacharya, "An overview of the operation of Ontario's electricity market," in *Proc. IEEE PES Annual General Meeting*, June 2005, pp. 2463 – 2470.

[25] J. H. Friedman, "Multivariate adaptive regression splines," *The Annual of Statistics*, vol. 19, no. 1, pp. 1–67, March 1991.

[26] P. A. W. Lewis and J. G. Stevens, "Nonlinear modeling of time series using multivariate adaptive regression splines," *Journal of American Statistical Association*, vol. 86, no. 416, pp. 864–877, Dec. 1991.

[27] H. Haas and G. Kubin, "A multi-band nonlinear oscillator model for speech," in *Conference Record of the Thirty-Second Asilomar Conference on Signals, Systems & Computers*, vol. 1, 1-4 Nov 1998, pp. 338 – 342.

[28] T. E. ad G. Kubin, "Nonlinear prediction of mobile radio channels: measurements and mars model designs," in *IEEE Proc. International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, 15-19 March 1999, pp. 2667 – 2670.

[29] S. Talukdar and C. W. Gellings, *Load Management*. New York: IEEE Press, 1986.

[30] S. Ashok and R. Banerjee, "An optimization mode for industrial load management," *IEEE Transactions on Power Systems*, vol. 16, no. 4, pp. 879–884, Nov 2001.

[31] ——, "Optimal operation of industrial cogeneration for load management," *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 931–937, May 2003.

[32] H. Zareipour, K. Bhattacharya, and C. Canizares, "Distributed generation: Current status and challenges," in *Proc. 36th Annual North American Power Symposium (NAPS)*, August 9-10 2004.

[33] *Distributed Generation in Liberalized Electricity Markets*. International Energy Agency, 2002.

[34] A. M. Borbely and J. F. Kreider, *Distributed Generation: The Power Paradigm for the New Millennium*. CRC Press, 2001.

[35] A. Baueu, D. Hart, and A. Chase, "Fuel cells for distributed generation in developing countries-an analysis," *International Journal of Hydrogen Energy*, vol. 28, pp. 695–701, 2003.

[36] M. McNeely, "Power generation order survey, the coaster ride continues downhill," *Diesel and Gas Turbine Worldwide*, Oct. 2003.

[37] K. Huhn, "Market opportunities for distributed generation systems," in *Proc. of the Seminar on Distributed Energy Systems*, vol. 1. Frost & Sullivan Inc., Jan. 2003.

[38] J. H. Horlock, *Cogeneration: Combined Heat and Power, Thermodynamics and Economics*. Pergamon Press, 1987.

[39] H. Asano, S. Sagai, E. Imamura, and R. Ito, K.; Yokoyama, "Impacts of time-of-use rates on the optimal sizing and operation of cogeneration systems," *IEEE Transactions on Power Systems*, vol. 7, no. 4, pp. 1444–1450, Nov 1992.

[40] I. Bae, J. Kim, J. Kim, and C. Singh, "Optimal operating strategy for distributed generation considering hourly reliability worth," *IEEE Transactions on Power Systems*, vol. 19, no. 1, pp. 287–292, Feb 2004.

[41] S. Illerhaus and J. Verstege, "Assessing industrial load management in liberalized energy markets," in *IEEE Proc. PES Summer Meeting*, vol. 4, 16-20 July 2000, pp. 2303–2308.

[42] B. Pribicevic, B. Krasenbrink, and H.-J. Haubrich, "Co-generation in a competitive market," in *IEEE Proc. PES Summer Meeting*, vol. 1, 21-25 July 2002, pp. 422–426.

[43] E. Gomez-Villalva and A. Ramos, "Optimal energy management of an industrial consumer in liberalized markets," *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 716–723, May 2003.

[44] J. N. Sheen, "Fuzzy financial decision-making: load management programs case study," *IEEE Transactions on Power Systems*, vol. 20, no. 4, pp. 1808–1817, Nov 2005.

[45] B. Baran, C. von Lucken, and A. Sotelo, "Multi-objective pump scheduling optimisation using evolutionary strategies," *Advances in Engineering Software*, vol. 36, no. 1, pp. 39–47, Jan. 2005.

[46] A. J. Conejo, F. J. Nogales, and J. M. Arroyo, "Price-taker bidding strategy under price uncertainty," *IEEE Transactions on Power Systems*, vol. 17, no. 4, pp. 1081 – 1088, November 2002.

[47] D. Ranaweera, G. Karady, and R. Farmer, "Economic impact analysis of load forecasting," *IEEE Transactions on Power Systems*, vol. 12, no. 3, pp. 1388 – 1392, Aug. 1997.

[48] B. Hobbs, S. Jitprapaikulsarn, S. Konda, V. Chankong, and K. L. andD.J. Maratukulam, "Analysis of the value for unit commitment of improved load forecasts," *IEEE Transactions on Power Systems*, vol. 14, no. 4, pp. 1342 – 1348, Nov. 1999.

[49] M. A. Ortega-Vazquez and D. S. Kirschen, "Economic impact assessment of load forecast errors considering the cost of interruptions," in *Proc. the IEEE PES Annual General Meeting*, June 2006, p. 8 pages.

[50] H. Zareipour, C. Canizares, and K. Bhattacharya, "The ontario competitive electricity market: Overview, experiences, and lessons," *IEEE Transactions on Power Systems*, p. 11, 2006, submitted.

[51] IESO, *Market Rules for the Ontario Electricity Market*, 2005, available [online] at: http://www.ieso.ca/imoweb/ manuals/marketdocs.asp.

[52] B. Danai, J. Kim, A. Cohen, V. Brandwajn, and S.-K. Chang;, "Scheduling energy and ancillary service in the new ontario electricity market," in *IEEE Proc. PES International Conference on Power Industry Computer Applications, PICA 2001*, 20-24 May 2001, pp. 161–165.

[53] A. I. Cohen, V. Brandwajn, and S.-K. Chang, "Security constrained unit commitment for open markets," in *IEEE Proc. PES International Conference on Power Industry Computer Applications, PICA 1999*, 16-21 May, 1999, pp. 39–44.

[54] K. Bhattacharya, M. H. Bollen, and J. E. Dallder, *Operation of Restructured Power Systems*. Boston: Kluwer Academic Publishers, 2001.

[55] Y. Chien-Ning, A. Cohen, and B. Danai, "Multi-interval optimization for real-time power system scheduling in the ontario electricity market," in *IEEE Proc. Power Engineering Society General Meeting, 2005.*, 12-16 June 2005, pp. 2717 – 2723.

[56] H. Zareipour, C. Canizares, and K. Bhattacharya, "Application of public-domain pre-dispatch information to forecast Ontario electricity market prices," *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp. 11 pages,, November 2006.

[57] A. Pankratz, *Forecasting with Dynamic Regression Models*. John Wiley and Sons, Inc., 1991.

[58] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis, Forecasting and Control*. Prentice Hall, Inc., 1976.

[59] T. C. Mills, *Time series techniques for economists*. Cambridge University Press, 1990.

[60] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis, Forecasting and Control*. USA: Prentice Hall Inc., 1994.

[61] L.-M. Lio, *Time series analysis and forecasting*. Scientific Computing Associates Corp., 2005.

[62] C. Chen and L.-M. Liu, "Forecasting time series with outliers," *Journal of Forecasting*, vol. 12, no. 1, pp. 13–23, 1993.

[63] V. M. Guerrero, "Time series analysis supported by power transformations," *Journal of Forecasting*, vol. 12, no. 1, pp. 37–48, 1993.

[64] S. Lipovetsky and W. M. Conklin, "Ridge regression in two-parameter solution," *Applied Stochastic Models in Business and Industry*, vol. 21, no. 6, pp. 525 – 540, November/December 2005.

[65] A. M. Aguilera, M. Escabias, and M. J. Valderrama, "Using principal components for estimating logistic regression with high-dimensional multicollinear data," *Computational Statistics and Data Analysis*, vol. 50, no. 8, pp. 1905–1924, April 2006.

[66] H. Zareipour, K. Bhattacharya, and C. Canizares, "Forecasting the hourly ontario energy price by multivariate adaptive regression splines," in *Proc. the IEEE PES Annual General Meeting*, June 2006, p. 7 pages.

[67] S. Mukkamala and A. H. Sung, "A comparative study of techniques for intrusion detection," in *IEEE Proc. 15th International Conference on Tools with Artificial Intelligence*, 3-5 Nov. 2003, pp. 570 – 577.

[68] Q.-S. Xu, D. Massart, Y.-Z. Liang, and K.-T. Fang, "Two-step multivariate adaptive regression splines for modeling a quantitative relationship between gas chromatography retention indices and molecular descriptors," *Journal of Chromatography*, vol. 998, no. 1-2, pp. 155–167, May 23 2003.

[69] C.-C. Yang, S. O. Prasher, R. Lacroix, and S. H. Kim, "A multivariate adaptive regression splines model for simulation of pesticide transport in soils," *Biosystems Engineering*, vol. 86, no. 1, pp. 9–15, Biosystems Engineering 2003.

[70] A. Abraham, "Analysis of hybrid soft and hard computing techniques for forex monitoring systems," in *IEEE Proc. International Conference on Fuzzy Systems*, vol. 2, 12-17 May 2002, pp. 1616 – 1622.

[71] T.-S. Lee and I.-F. Chen, "A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines," *Expert Systems With Applications*, vol. 28, no. 4, pp. 743–752, May 2005.

[72] S.-M. Chou, T.-S. Lee, Y. E. Shao, and I.-F. Chen, "Mining the breast cancer pattern using artificial neural networks and multivariate adaptive regression splines," *Expert Systems With Applications*, vol. 27, no. 1, pp. 133–142, July 2004.

[73] H. H. Stokes and W. J. Lattyak, *Multivariate Adaptive Regression Spline (MARS) Modeling Using the B34S ProSeries Econometric System and SCA WorkBench*. Scientific Computing Associates Corp., 2005.

[74] M. T. Hagan, H. B. Demuth, and M. Beale, *Neural Network Design*. PWS Publishing Company., 1996.

[75] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.

[76] H. Demuth and M. Beale, *Neural Network Toolbox Users Guide*. The Mathework Inc., 2003.

[77] H. Zareipour, K. Bhattacharya, and C. Canizares, "Electricity market price volatility: the case of ontario," *Journal of Energy Policy*, pp. 19 pages,, 2006, submitted.

[78] T. G. Andersen, T. Bollerslev, P. F. Christoffersen, and F. X. Diebold, "Volatility forecasting," *National Bureau of Economic Research*, 2005, working Paper 11188.

[79] F. Alvarado and R. Rajaraman, "Understanding price volatility in electricity markets," in *Proc. of the 33rd Annual Hawaii International Conference on System Sciences*, 4-7 Jan. 2000, pp. 1–5.

[80] R. W. Dahlgren, C. C. Liu, and J. Lawarree, "Volatility in the california power market: source, methodology and recommendations," *IEE Proceedings on Generation, Transmission and Distribution*, vol. 148, no. 2, pp. 189–193, March 2001.

[81] Y. Li and P. C. Flynn, "Deregulated power prices: comparison of volatility," *Energy Policy*, vol. 32, no. 14, pp. 1591–1601, September 2004.

[82] A. Worthington, A. Kay-Spratley, and H. Higgs, "Transmission of prices and price volatility in australian electricity spot markets: a multivariate garch analysis," *Energy Economics*, vol. 27, no. 2, pp. 337–350, March 2005.

[83] P. Jorion, *Value at Risk*. McGraw-Hill, 2001.

[84] P. F. Christoffersen, *Elements of Financial Risk Management*. USA: Academic Press, 2003.

[85] J. P. Bouchaud and M. Potters, *Theory of Financial Risks: From Statistical Physics to Risk Management*. Cambridge: Cambridge University Press, 2000.

[86] K. W. Cheung, "Standard market design for ISO New England wholesale electricity market: An overview," in *proc. of 2004 IEEE International Conference on Electric Utility Deregulation, Restructuring and Power Technologies (DRPT2004)*, 2004, pp. 38–43.

[87] NEISO, *2004 Annual Markets Report*. New England ISO Inc., 2005, available [online] at: http://www.iso-ne.com/markets/mkt_anlys_rpts/annl_mkt_rpts/2004/2004_annual_markets_report_.pdf.

[88] NYISO. The New York Independent System Operator (NYISO), Energy Market information web page, 2005, available [online] at: http://www.nyiso.com/public/products/energy_market/index.jsp?display=0.

[89] The General Algebraic Modeling System (GAMS), [online] at: http://www.gams.com/, 2004.

[90] EIA, *Energy Information Administration, Official Energy Statistics from the US Government*, 2006, available [online] at: http://www.eia.doe.gov/.