12-2009

Congestion and Price Prediction in Locational Marginal Pricing Markets Considering Load Variation and Uncertainty

Rui Bo
University of Tennessee - Knoxville

To the Graduate Council:

I am submitting herewith a dissertation written by Rui Bo entitled "Congestion and Price Prediction in Locational Marginal Pricing Markets Considering Load Variation and Uncertainty." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Electrical Engineering.

Fangxing (Fran) Li, Major Professor

We have read this dissertation and recommend its acceptance:

Kevin Tomsovic, Leon M. Tolber, Tse-wei Wang

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a dissertation written by Rui Bo entitled "Congestion and Price Prediction in Locational Marginal Pricing Markets Considering Load Variation and Uncertainty." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Electrical Engineering.

Fangxing (Fran) Li, Major Professor

We have read this dissertation
and recommend its acceptance:

Kevin Tomsovic

Leon M. Tolber

Tse-wei Wang

Accepted for the Council:

Carolyn R. Hodges
Vice Provost and Dean of the Graduate School

Congestion and Price Prediction in Locational Marginal Pricing Markets Considering Load Variation and Uncertainty

A Dissertation Presented for
the Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Rui Bo
December 2009

Acknowledgements

I would like to thank my advisor, Dr. Fangxing (Fran) Li, for his continuous guidance and support, not only in my doctoral studies, but also in student life. His dedication to teaching, tutoring, and assisting me is greatly appreciated.

I am also thankful to Dr. Kevin Tomsovic, Dr. Leon M. Tolbert, and Dr. Tse-wei Wang for their valuable inputs and comments on this work. Thanks to all students in the UT Power Lab, including those already graduated. Their friendship, encouragement, and help were indeed beneficial to my study and research work.

I would also like to acknowledge the staff of the Department of Electrical Engineering and Computer Science for their nice job. They are always readily available to help students with various kinds of issues.

Last but not least, I am greatly indebted to my family, especially my wife Ying Ji, for their endless love and unconditional support, which made it possible for me to finish this work.

Abstract

This work investigates the prediction of electricity price and power transmission network congestions under load variation and uncertainty in deregulated power systems. The study is carried out in three stages.

In the first stage, the mathematical programming models, which produce the generation dispatch solution, the Locational Marginal Price (LMP), and the system statuses such as transmission congestions, are reviewed. These models are often referred to as Optimal Power Flow (OPF) models, and can be categorized into two major groups: Alternating Current OPF (ACOPF) and Direct Current OPF (DCOPF). Due to the convergence issue with the ACOPF model and the concern of inaccuracy with the DCOPF model, a new DCOPF-based algorithm is proposed, using a fictitious nodal demand (FND) model to represent power losses at each individual line. This is an improvement over the previous work that assigns losses to a few user-defined buses, and is capable of achieving a better tradeoff between computational effectiveness and the accuracy of the results.

In the second stage, the solution features are explored for each of the three OPF models to predict critical load levels where a step change of LMP occurs due to the change of binding constraints. After careful examinations of the mathematical relationship of the OPF solutions, nodal prices, and congestions, with respect to load variation, simplex-like method, quadratic interpolation method, and variable substitution method are proposed for each of the three OPF models respectively in order to predict price changes and system congestion.

In the last stage, the probabilistic feature of the forecasted LMP is investigated. Due to the step change characteristic of the LMP and uncertainty in load forecasting, the forecasted LMP represents only a certain possibility in a lossless DCOPF framework. Additional possible LMP values exist, other than the deterministically forecasted LMP. Therefore, the concept of Probabilistic LMP is introduced and a systematic approach to quantify the probability of the forecasted LMP, with respect to load variation, is proposed. Similar concepts and methodology have been applied to the ACOPF and FND-based DCOPF frameworks, which can be useful for power market participants in making financial decisions.

Table of Contents

List of Tables

List of Figures

Nomenclature

Indices:

i	Index of buses, or, index of marginal units
j	Index of marginal units, or, index of buses
k	Index of branches of a transmission system

Dimensions:

N	Number of buses
N_{MG}	Number of marginal units
M	Number of lines
M_{CL}	Number of congested lines
M_{UL}	Number of un-congested lines

Sets:

$\mathcal{N}$	Set of buses
$\mathcal{MG}$	Set of marginal units
$\mathcal{NG}$	Set of non-marginal units
$\mathcal{R}$	Set of branches
$\mathcal{R}_c$	Set of congested branches
$\mathcal{CL}$	Set of congested lines
$\mathcal{UL}$	Set of un-congested lines
$\varnothing$	Empty set

Variables:

G_i	Generation dispatch at Bus *i* (MWh)
D_i	Demand at Bus *i* (MWh)
P_i	Net injection at bus *i* (MWh)
E_i	Fictitious Nodal Demand (FND) at bus *i* (MWh)
f_i	Load growth participating factor at bus *i*
DF_i	Marginal delivery factor at bus *i*
LF_i	Marginal loss factor at bus *i*
P_{Loss}	Total system losses (MWh)
F_k	Line flow at line *k* (MWh)
LMP_i	Locational Marginal Price (LMP) at Bus *i* ($/MWh)
MG_j	Generation dispatch of j^{th} marginal unit (MWh)
NG_j	Generation dispatch of j^{th} non-marginal unit (MWh)

Constants:

c_i	Generation cost of the generator at Bus *i* ($/MWh)
G_i^{max}	The maximum generation output of the generator at Bus *i* (MWh)
G_i^{min}	The minimum generation output of the generator at Bus *i* (MWh)
$GSF_{k\text{-}i}$	Generation shift factor to line *k* from bus *i*
$Limit_k$	Transmission limit of line *k* (MWh)
R_k	Resistance at line *k* (MWh)

Operators:

$\Pr(\cdot)$	Operator of probability
$E(\cdot)$	Operator of expected value

Fonts:

bold	Vector or matrix

1 Introduction

1.1. Background

1.1.1. Deregulation

The physical structure of power systems, consisting of generation, transmission, and distribution networks, appears the same as it did decades ago, as shown in Figure 1.1. However, the ownership structure and the operation of power systems have changed fundamentally.

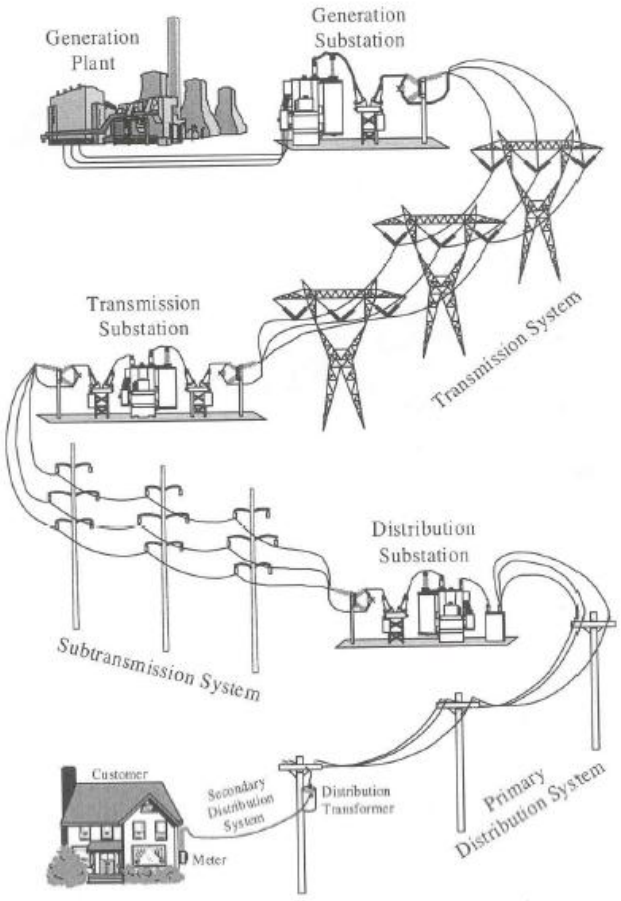


Figure 1.1. Typical electric power systems [1]

Traditionally, the power industry is vertically integrated, within which the generation, transmission, and distribution are bundled together as a utility to serve its customers. With the expansion of each individual power system, additional power grids are connected to reap the benefits of scale effect.

In the wake of the successful deregulation of other industries, such as airlines, wireless communications, and mailing services, individuals, in the late 1980's, researched the possibility of deregulating the power industry in the hopes of lowering the price of electricity. Some argued that unbundling the generation, transmission, and distribution sectors of the power industry would, naturally, create competition within each sector. Later, the power engineering community gradually agreed that the transmission sector needs to maintain the power systems as a centralized operation. This is due to the natural monopoly feature of transmission systems created by the inefficiency and high cost of duplicate investments. For similar reasons, the distribution sector should maintain a monopoly mode as well. However, the physical network should be operated as a fair and open platform for generators and loads to carry out the electricity trade. Therefore, the Federal Energy Regulatory Commission (FERC) issued an order in 2000 to create non-profit organizations, called Independent System Operator (ISO) or Regional Transmission Organization (RTO), to organize regional power systems to ensure non-discriminatory transmission services to generation companies (GENCOs) and bilateral transactions. An ISO or RTO is committed to providing open and fair transmission access, called "Open Access", and to treating all participants equally. In addition, it is responsible for operating the power grid reliably and efficiently. This is achieved through sound market rules, proper monitoring and regulation, and timely and accurate information publications, such as wholesale market prices.

The deregulation resulted in a wholesale power market, also known as an energy market or electricity market, with GENCOs, load serving entities (LSEs) or distribution companies

(DISCOs), and traders as major market players, as seen in Figure 1.2. GENCOs, DISCOs, and large consumers trade electricity directly in the wholesale market, while small consumers purchase electricity through DISCOs who operate the distribution network and serve the customers within its territory. Market players in this wholesale pool market can buy or sell electricity by submitting offers and bids. The ISO will select the offers and bids, from an economical perspective, while ensuring the security of the power systems. Market players can also execute bilateral transactions that utilize the transmission network as a wheeling service. These transactions will be submitted to the ISO to ensure feasibility.

A sufficient number of players usually ensure the effectiveness of competition in a power market. For example, in 2003 in the Pennsylvania, Jersey, Maryland Power Pool (PJM) market, there were more than 200 market buyers, sellers, and traders [18]. Besides players who are actively engaged in buying and selling transactions, other entities, such as transmission owners (TOs), legislators, and environmentalists, are also involved in the design, evolvement, and regulation of the power market.

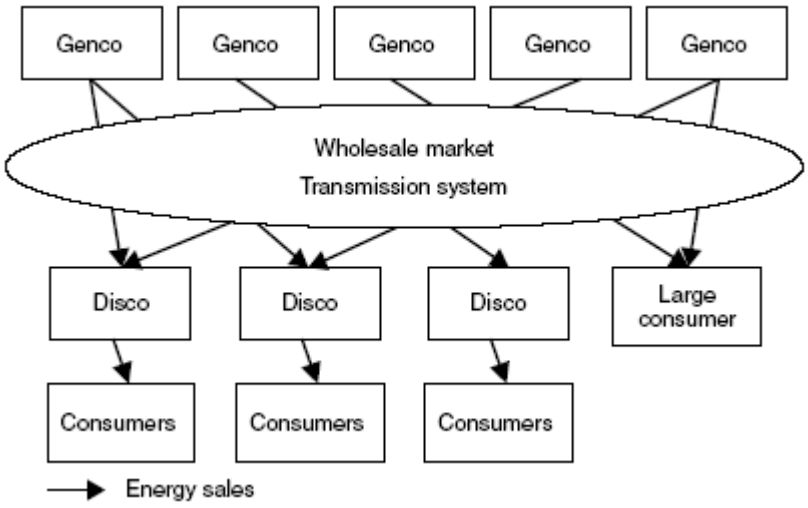


Figure 1.2. Wholesale model of electricity market [7]

It should be pointed out that the electricity market differs from regular commodity markets such as oil, bread, and even water, in the sense that no practical technology is available for large-scale storage of electricity, except for a very small portion of pumped storage units. This implies that electricity has to be procured in real-time from primary energy sources such as coal, gas, wind, and solar. In addition, the balance of supply and demand has to be maintained instantaneously, or the quality of the electricity, either the frequency or voltage level, will be compromised. A severe imbalance will lead to system instability, or even collapse, within a few minutes. Another important feature is electricity flow paths. Electricity, sent from a source to a sink, is distributed among transmission lines that are subject to physical laws such as Kirchhoff's Circuit Law; instead of taking designated paths like other commodities are delivered.

1.1.2. Day-ahead and real-time energy markets

A power market may be comprised of different types of markets. An energy market is the market where the financial or physical trading of electricity takes place. It typically consists of a day-ahead market and real-time market, while the ancillary service market is the market to procure services such as the synchronized reserve, regulation, and black start to support the reliable operation of the transmission system in the consideration of unexpected events. In order to help market participants hedge the risk of being exposed to high electricity prices due to transmission system congestions, the Financial Transmission Rights (FTR) market was created. Participants can acquire FTRs in monthly or annual FTR auctions to hedge against congestion costs.

The energy market will be presented next in further detail as it is the primary focus of this work. The day-ahead market is a forward market and runs on the day before the operating

day. Generation offers, demand bids, and bilateral transactions are submitted to the market by a certain time, regulated by the market timeline. Virtual offers and bids are also accepted in the market as their contribution increases market liquidity. The load is estimated by Load Forecasting tools. An optimization model called the Security Constrained Unit Commitment (SCUC) and Security Constrained Economic Dispatch (SCED) is then established and solved to produce the generation dispatch and electricity prices for each hour of the operating day. A SCUC problem is to obtain a unit commitment schedule at minimum generation cost with system reliability requirements enforced, such as generation resource operating constraints. It may incorporate a subset of transmission constraints, which represent the most likely binding constraints, sometimes called "watch list transmission flowgate constraints" [13]. Due to the performance limitations of the SCUC solvers, SCUC can not handle the full set of transmission constraints. These are addressed in SCED [13]. SCED is an optimization problem similar to SCUC with a given unit commitment obtained from SCUC, and is capable of modeling a full set of prevailing constraints.

The major portion of the load is matched by the generation scheduled in the day-ahead market, which significantly reduces price uncertainty and provides market players a tool to lock in an advance price as opposed to being exposed to volatile prices in real-time. This also helps to reduce incentives used to exercise market power [21].

The real-time market is a spot market and aims to balance the deviations between the forecasted load in real-time and the forecasted load in the day-ahead market through the SCUC and SCED. The SCUC may be performed on an hourly basis or as needed, while the SCED is typically run on shorter intervals. For instance, a SCED is performed on a 5-minute basis in the PJM and Midwest power markets, and the price is calculated on the same interval.

In terms of the settlement in the day-ahead market, market players will receive or make payments based on their hourly scheduled quantities and the day-ahead price. Participants in

the real-time market will receive or make payments for the deviation between their actual power generation/consumption and their scheduled quantities in the day-ahead market, based on the real-time price. This two-settlement mechanism provides incentive to market players to follow the dispatch instructions, and helps the two markets converge in terms of prices.

1.1.3. Locational Marginal Price (LMP)

When there is no transmission bottleneck and losses present during the transportation of the electricity, the cheapest power producer will be selected to serve the loads at all locations and therefore, the electricity price will be the same across the grid. This price is often called the market clearing price (MCP). In this scenario, the grid that connects all the generators and loads is similar to a single bus which has infinite transportation capability and induces no loss. When congestion occurs so that one or more transmission lines reach their thermal limit and are unable to carry additional power, a more expensive generation unit will be scheduled to serve the load since the cheaper generators could not reach the load location due to congestion. Consequently, electricity prices at this location will increase since it is served by the more expensive power producers. In addition to transmission congestion, power transmission losses also contribute to the varying prices at the different locations. For instance, a load, connected to the grid through a higher resistive transmission line, will be subject to a higher price since more electricity is lost during transportation, as opposed to the case of a lower resistive line. As a result, electricity price varies with locations. These characteristics lead to the concept of Locational Marginal Price (LMP).

LMP was firstly introduced by F.C. Schweppe in 1998 [4]. By definition, the LMP at a given bus is the incremental cost of serving an infinitesimal change of load at that bus, while respecting all physical constraints. The Locational Marginal Pricing (LMP) methodology has been the dominant approach used in the U.S. power markets to calculate the electricity price

and to manage transmission congestion. Currently, LMP has been implemented or is under implementation at a number of ISO's such as the PJM, New York ISO, ISO-New England, California ISO, and Midwest ISO [11, 14, 15].

In theory, the LMP is the by-product of the economic dispatch (actually SCED) problem. Specifically, the LMP at a given bus is the shadow price of the power balance equation associated with that bus. Furthermore, the LMP can be decomposed into three parts: marginal energy price, marginal loss price, and marginal congestion price. These three parts represent the marginal cost associated with energy, loss, and congestion, respectively. The reason that the LMP is split into three components is that the marginal congestion component is used to calculate congestion revenue and the value of the FTR [32].

In practice, the day-ahead market generates the LMP, called "ex-ante LMP", because the LMP is calculated before the event happens. In the real-time market, besides the calculation of the "ex-ante" LMP, a "post-LMP" calculation will be performed, for example, every 5 minutes, to reflect what has actually happened in the market. The calculation respects the actual system conditions and generator responses, according to State Estimation results. These prices are called "post-LMP" prices. Theoretically, the post-LMP would be the same as the ex-ante LMP, if things go exactly as expected or forecasted. In practice, the post-LMP should be close to the ex-ante LMP, in most cases. In addition, in a well-designed and operated power market, the ex-ante and ex-post LMP are expected to converge over time. In this work, we focus on the ex-ante LMP as the research context is forecasting.

The electricity price in a wholesale energy market changes constantly for a number of reasons, such as load changes, changes of generation offers and demand bids, change of transmission system in the event of outage and maintenance, and change of availability of generators due to outage. Among these factors, the load is changing the most frequently.

1.1.4. Economic Dispatch

Economic dispatch (ED) is of primary interest in this dissertation as it is the model that produces LMP. Economic dispatch in the day-ahead and real-time markets normally optimizes generation dispatch in such a way that social welfare is maximized. In the generation-side power market, where no dispatchable loads are available, the goal is equivalent to achieving minimum total generation cost without violating the safe operating range of any system component. However, in order to ensure minimum cost, several components will have to perform at their operating limits since the optimal dispatch will utilize the most cost effective components to their full capabilities without endangering them. For example, a transmission line may be scheduled to run at its thermal limit. There are three types of thermal limits for transmission lines: a normal operation rating, a long-term emergency rating (4 hours), and a short-term emergency rating (15 minutes) [65]. They represent the maximum ability of transporting power in the long-time, short-time, and very short time, respectively. The normal operation rating is used as the thermal limit in economic dispatch. This implies the dispatch should not endanger any transmission line after running for a long-time. For example, no transmission lines will be overheated due to ohmic losses and result in excessive sags, which could create a fault.

In typical power market scheduling, N-1 security has to be maintained. This means the system has to be able to survive the disturbance of losing a single component without overloading any other system component and violating any constraint. When N-1 security is modeled in the economic dispatch problem, the problem is often called the Security Constrained Economic Dispatch (SCED).

1.1.5. Optimal Power Flow (OPF) and its models

Mathematically, economic dispatch is a specific type of optimal power flow problem. Optimal power flow (OPF) normally refers to an optimization problem subject to the physical limitations of the power system [2]. The optimization objective has different forms other than the minimum generation cost, such as the minimum transmission losses or the minimum load shedding schedule. In this work, OPF will be used to refer to economic dispatch specifically for naming conventions.

The OPF model contains an objective function, equality constraints such as power balance equations, and inequality constraints such as the power flow thermal limit, generator ramp rate, and generator output limit. When contingency security is considered, it is sometimes called the Security Constrained OPF (SCOPF) to emphasize the inclusion of the security constraints model. In this dissertation, the contingency security is not explicitly modeled. However, it can be easily integrated through different sets of generation shift factors (GSF), as discussed in the following chapters. For notational convenience, the model in this work is referred to as OPF instead of SCOPF.

According to the form of the power flow model, OPF models can be categorized into the Alternating Current OPF (ACOPF) and Direct Current OPF (DCOPF). The ACOPF model establishes the power balance equations in a regular AC model. They address the real and reactive power flows at the same time. In contrast, the DCOPF model employs the DC power flow, which is a linearized, simplified power flow model. The DC power flow model is derived utilizing assumptions which are reasonable for high-voltage transmission systems. For example, it is assumed that the voltage profile is 1.0 per unit (p.u.) throughout the network, and the voltage phase angle difference between adjacent buses is minimal. Reactive power, considered a local issue, is ignored in the DC power flow model.

A full ACOPF model is normally viewed as the most representative mathematical model to the practical scheduling problem and therefore, its corresponding results are considered as accurate and could be used as benchmark data. Despite the accuracy of its results, the ACOPF is a nonlinear programming problem and requires a good initial point to help the solver to converge. Occasionally, the model suffers from convergence problems, especially for large-scale power systems. Therefore, although the ACOPF is being used in some real-world applications, such as in the California ISO [20], the DCOPF is used more often in scheduling with AC power flow verification. The DCOPF model is a much less complicated linear programming model, and could be solved with less effort and issues. The DCOPF has the advantage over the ACOPF in terms of convergence and speed. Therefore, the DCOPF is used by a number of ISOs in the U.S., such as the PJM [18], MISO [13], ISO New England [32], and NYISO [14]. The DCOPF is also often used in power system planning where numerous hypothetic cases are to be studied, and speed and convergence are two crucial factors in choosing the correct OPF model. Therefore, even though the DCOPF occasionally suffers from yielding inaccuracy results under heavily loaded systems, insufficient local reactive power supply, or transmission systems with a high r/x ratio, it is still very popular in the power industry among ISOs and market participants for operation and planning purposes.

1.2. Motivation

In power system planning and real-time operation, it is always desirable to forecast the system status. For a steady-state analysis, a crucial task is to identify whether the system is operating in a stressed condition. With more system components operating close to their capacity limits, the system becomes more vulnerable to potential disturbances. In contrast, if the system has fewer components close to its operating limits, the system has a bigger margin for disturbance and is more robust.

A stressed power system incurs economic burdens in addition to technical challenges. In a deregulated environment, transmission congestion will generate additional cost due to the re-dispatch of more expensive generators [3]. In 2006, the total congestion cost in the PJM was approximately $1.06 billion, which accounts for 8% of the total PJM billing in that year [12]. Therefore, it is of great interest to investigate the influential factors that affect congestions. One is the load change/variation. For instance, when the load grows beyond the transmission capability of a transmission line or interface, congestion will occur. Corresponding actions, such as upgrading or new line construction, should be taken if economically viable.

In fact, the successful prediction of important changes in the system's status provides information, in addition to congestion. For instance, when a cheap generator is reaching its upper-limit, a more expensive generator will be committed to serve the load increase. Then the LMPs throughout the system will change, with some changing dramatically.

To date, no ready-to-use tools exist that can predict the important changes of system status, such as congestions, with respect to perturbations like load variation under the framework of the Optimal Power Flow (OPF). In fact, the power industry views the OPF problem from a simulation perspective, and runs the OPF under each of the operating conditions of interest (for instance, different load levels), while overlooking the fact that operating conditions change continuously since the load varies continuously, particularly in the short term. An individual OPF run provides a snapshot for a specific scenario while the system conditions, such as the load levels, change continuously. Therefore, it is of interest to investigate the dynamic behavior of a deregulated power system operation. For instance, it shall be interesting to investigate the sensitivity of OPF solutions with respect to change in load. Similarly, it can also be interesting to investigate the variations pattern of transmission congestion and price with respect to load variation.

On the other hand, the optimal solution of an OPF problem is a function of a number of parameters in the model, but the parameters are not equally significant in the study as they have different characteristics with respect to time. For example, the generator bounds and transmission line/transformer thermal limits are relatively less dynamic (i.e., more or less a constant value), at least in the short term. Generator offers may change on an hourly basis and have stochastic characteristics due to bidding strategies. However, generator offers could be inferred from historical data since bidding strategies may remain unchanged for days of the week, hours of the day, or both [6]. The availability of generation and transmission network components changes occasionally, either at designated times (in the case of a forced outage and unit commitment) or are hard to predict (in the case of an unplanned outage). Among these parameters, the load is the one that changes constantly, contributing directly to price volatility. Although the load has stochastic characteristic, it exhibits strong cyclical patterns. In addition, load forecasting is performed on the bus or area level, where loads under the same bus or area are aggregated and easier to predict.

Therefore, the load and its variations are of the primary interest. This work intends to forecast potential line congestions, as well as the change of marginal units and the resulting LMP in response to load variations. Results of this work will help market participants in forecasting market prices and developing their bidding strategy. The ISO could also perform the prediction and publish the forecasted price signal to allow customers to make adjustments in advance to hedge a potential price spike, while currently, customers can only respond after the price is published or make predictions from historical data. This capability will in turn benefit the system by reducing congestion time and the need for peak generation.

The solution pattern of the OPF is dependent on the type of OPF model that is adopted. The DCOPF will be studied since it is widely used. Although the AC power flow model, which is a system of nonlinear equations, is normally used for verification purposes, the

ACOPF is not normally used in real applications due to its complex formulation and computational issues, such as convergence and speed. Nevertheless, using ACOPF is the ultimate goal for the future and serves as the source of benchmarked results. Therefore, studies will be conducted under both the DCOPF and ACOPF frameworks.

In addition to the OPF models used, the successful prediction of the system status and LMP relies on acquired model parameters. One of the more volatile parameters is the load, which is never exactly known and has to be estimated using load forecasting tools. Uncertainty in the forecasted load could result in the predicted LMP being different from the actual LMP. Therefore, this work will also investigate the confidence level of the forecasted LMP, which could be used by market participants to formulate bidding strategies and hedge risk against load variation and uncertainty.

1.3. Dissertation Outline

Literatures relevant to this work are briefly reviewed in Chapter 2.

Chapter 3 first revisits the conventional OPF models from which generation dispatch and LMP are calculated. The reviewed models include the conventional Lossless DCOPF, ACOPF, and DCOPF with loss model. The conventional Lossless DCOPF and ACOPF have their advantages and disadvantages in accuracy and convergence. The conventional DCOPF with loss model was an attempt to achieve a tradeoff between accuracy and convergence, but it suffers from problems as well. Furthermore, few studies have been carried out to differentiate results between the DCOPF and benchmark ACOPF. In this regard, first, a new DCOPF-based model based on fictitious nodal demand (FND) is proposed to achieve an improved tradeoff between accuracy and speed. In addition, a rule of thumb will be stated which assesses the accuracy and confidence level of the approximated model when compared with the ACOPF results.

In Chapter 4, congestion and price prediction will be conducted for each of the three models: the conventional lossless DCOPF model, ACOPF model and the proposed FND-based DCOPF model. For the lossless DCOPF dispatch model, a simplex-like, direct method is proposed to calculate the generation sensitivity, and identify new binding and unbinding limits in an efficient way. For the ACOPF model, an interpolation method is proposed to approximate the scheduled generation output and estimate the upcoming congestion. For the FND-based DCOPF model, a variable substitution approach will be presented.

Chapter 5 presents the concept of a probabilistic LMP which takes into consideration the effects of load forecasting errors. Methodology for evaluating the confidence level of the forecasted LMP will be proposed for the lossless DCOPF first and then extended to the ACOPF and FND-based DCOPF models, respectively.

The approaches and methodology are concluded in Chapter 6. Future work is presented as well.

Figure 1.3 depicts the structure of this dissertation, where the green blocks represent works that have been presented in this dissertation. The works are organized in a two dimensional manner to make the connections clear to readers. The first dimension presents three research stages such as the OPF and LMP calculation, congestion and price prediction, and probabilistic LMP. The second dimension presents OPF models including the lossless DCOPF, ACOPF, and DCOPF with loss model. Each work in this dissertation refers to a module and an OPF model.

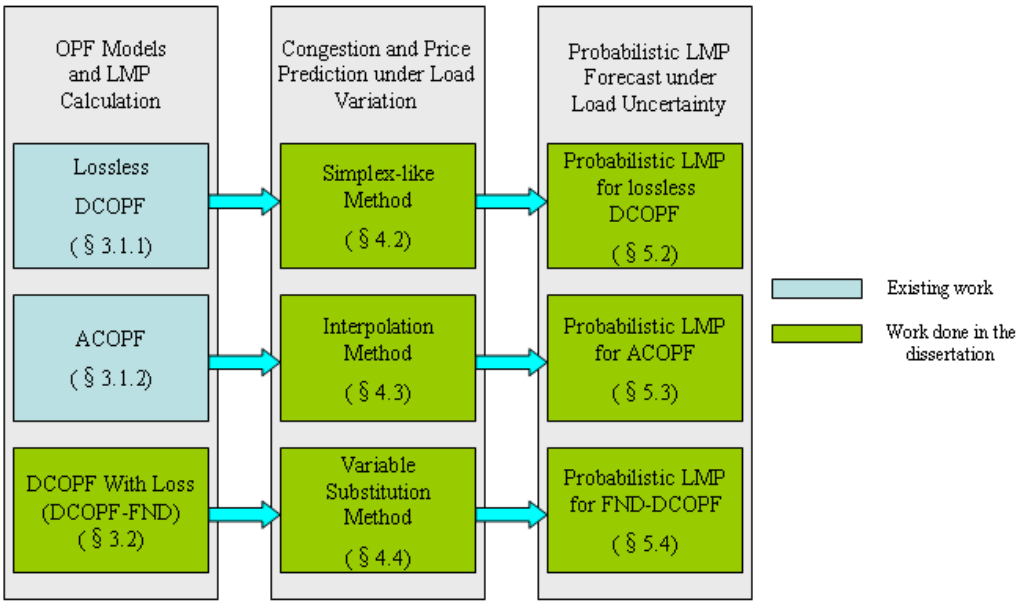


Figure 1.3. Dissertation Structure

1.4. Scope and Contribution of This Work

The problem of fast and accurate prediction of critical system status (such as congestion and marginal units) and LMP is considered in this work. Several general assumptions are taken exclusively in the work.

- The power system under study is assumed to be a three-phase synchronous power system with balanced loads and system components. Therefore, a one-line diagram will be used for better illustration.
- An aggregated network model is adopted. For example, a transmission line will be modeled as a resistance and reactance connected in series since only the voltages at both ends of the line, rather than the voltages across the line, are of concern.
- Only the steady state of the power system is under consideration.

- The voltage stability problem is not present in the system. It is assumed that there is a sufficient reactive power supply in the studied system so that each load bus maintains a reasonable power factor (such as 0.95).
- The load is modeled as constant power. It is reasonable since a steady state study is the main focus and the voltage profile is flat in the absence of the voltage stability issue.

This work first proposes a new DCOPF-based model by introducing the concept of the Fictitious Nodal Demand (FND). Compared to pre-determined loss distribution factors [32], the FND is a concept that naturally models the distribution of power losses among transmission lines. A LP-based iterative algorithm is presented to solve the FND-based DCOPF model. Numerical studies show that the new model could be solved in 4~5 LP iterations for a 30-bus system and its accuracy is superior to that of the lossless DCOPF model. In this regard, the FND-based DCOPF method achieves an acceptable balance between accuracy and speed. A LMP calculation is performed for various load levels, and the difference of the LMP obtained from the FND-based DCOPF model and benchmarking ACOPF model is studied. Conclusions are drawn that the LMP difference may be large when a different marginal unit set is identified using different DC based algorithms. This serves as a good rule of thumb for assessing the credibility of results obtained from approximated models such as the lossless DCOPF and FND-based DCOPF. In addition, it is observed that a LMP step change occurs at certain load levels, called critical load levels (CLLs), where the marginal unit set changes and a new congestion may appear.

A methodology for predicting the upcoming congestion and the change of marginal units is developed for the lossless DCOPF model, utilizing the fact that the optimal solution is

actually the solution of a system of linear equations. The generation output sensitivity and LMP sensitivity are derived as well. The proposed methodology enables fast prediction of the LMP with respect to load variations. Approaches for the FND-based DCOPF and ACOPF models will be presented afterwards.

The predicted LMP will be further assessed in terms of its probability of occurring, with consideration of the load forecast uncertainty. A framework is developed for the lossless DCOPF model, which produces the probability mass function of the LMP at project time and the alignment probability curve for the deterministic forecasted LMP. It reveals the probabilistic characteristics of the commonly-used deterministic LMP, and assists market participants in formulating a bidding strategy and making financial decisions.

2 Literature Review

This chapter will briefly review existing works relevant to this work.

2.1. OPF Models and LMP Calculation

DCOPF and ACOPF

From the viewpoint of generation and/or transmission planning, it is always crucial to simulate or forecast the LMP, which may be obtained using the traditional production (generation) cost optimization model given the data of generation, transmission, and load [2, 3]. Typically, the DC Optimal Power Flow (DCOPF) is utilized for LMP simulation or forecasting, based on the production cost model via Linear Programming (LP), due to LP's robustness and speed. The popularity of the DCOPF lies in its natural fit into the Linear Programming model. Various third-party LP solvers are readily available to plug into the DCOPF model to reduce the development effort for the vendors of LMP simulators. In industrial practice, the DCOPF has been employed by several software tools for chronological LMP simulation and forecasting, such as ABB's GridView™, GE's MAPS™, LCG UPLAN, Promod IV®, and Siemens PTI PSS™ LMP [58, 59, 60, 61, 62].

Other literature also shows the acceptability of the DC model in power flow studies if the line flow is not very high, the voltage profile is sufficiently flat, and the R/X ratio is less than 0.25 [23]. A comparison between the DCOPF and ACOPF is conducted for a 12965-bus model of the Midwest U.S. transmission grid in [24]. Results show that the DCOPF appears to do an estimable job of identifying congestion patterns and can be up to 60 times faster than its AC counterpart, which is a substantial advantage over the ACOPF. Another issue with the selection of the ACOPF or DCOPF is the algorithm robustness. A Linear Programming (LP)-

based DCOPF algorithm can always yield solutions while a Nonlinear Programming (NLP)-based ACOPF algorithm is less robust and often experiences convergence problems [63]. In addition, [63] points out that a deregulated power market imposes new challenges to the solution of the ACOPF problem. For example, non-differentiable piecewise offer and bid curve brings computational difficulty to the ACOPF solution algorithms.

Distribution Factors and Loss Factors

In DCOPF formulations, two factors, the Generation Shift Factor (GSF) and Loss Factor (LF), are often used to model power flow and power loss [32].

The Generation Shift Factor (GSF), also known as the Power Transfer Distribution Factor (PTDF), can be viewed as the quick calculation of line flow changes with respect to a small injection change at a specific bus [2]. It should be noted that the corresponding sink to absorb the injection change is the slack bus. The GSF or PTDF is widely used in operation, planning, and research. In fact, it is known empirically [26] and has been proven [27] that PTDFs are approximately constant values independent of the system operating condition, if the topology of the system does not change and bus voltage magnitudes are constants with sufficient local reactive power support. When these conditions are not satisfied, PTDFs can change significantly as the loading changes [27]. With the assumption of the DC power flow model, PTDFs associated with the DC model only depend on the topology and parameters of the transmission system, and therefore, are exactly constants. This feature of the DC PTDF facilitates sophisticated studies. For example, [28] proposes a scheme to reduce the power system to a smaller equivalent one based on zonal DC PTDFs.

The loss factor is defined as the incremental loss incurred by the unit net injection at a specific bus [2]. Reference [29] demonstrates the usefulness of the DC power flow model in calculating loss penalty factors, which may have a significant impact on generation scheduling. The authors of [29] also note it is not advisable to apply the predetermined loss

penalty factors from a typical scenario to all cases. Reference [30] presents a real-time solution, without repeating a traditional power flow analysis, to calculate loss sensitivity for any market-based slack bus from traditional Energy Management System (EMS) products, based on multiple generator slack buses.

Marginal Loss Pricing

The Locational Marginal Price (LMP) may be decomposed into three components including marginal energy price, marginal congestion price, and marginal loss price [3, 6]. Several previous works [31, 32, 30, 29, 33] have reported the modeling of the LMP, especially in the marginal loss model and related issues. Reference [31] notes the significance of the marginal loss price, which may have a variance of up to 20% among different zones in the New York Control Area, based on the actual operating data. Reference [32] presents a slack-bus-independent approach to calculate the LMP and LMP congestion component by introducing loss distribution factors to explicitly distribute the losses into buses, though it does not specifically address how to obtain the distribution factors which are crucial to the LMP calculation. Reference [33] presents marginal loss pricing algorithms, based on the DC model, by introducing a delivery factor, which is defined as one minus the loss factor, to account for losses in the energy balance equation.

LMP and its Decomposition

The characteristics of the LMP have been discussed and interesting observations are presented in [35]. First, the LMP can be larger than the highest generation offer due to congestion, in which case the expensive unit will be committed to replace the cheaper unit. Second, the LMP can be lower than the cheapest generator offer, in the case that it is cheaper to pay customers at locations where load consumption helps to relieve congested transmission lines. Similar observations are presented in [44].

In the LMP power market, Financial Transmission Rights (FTRs) are designed to hedge the risk of volatile LMPs due to congestion. The mechanism relies on the congestion component of the LMP and therefore, requires the decomposition of the LMP into components. [36] presents an integrated optimal spot pricing model and decomposes the price into different components associated with generation, loss, and ancillary services such as spinning reserve, voltage control, and security control. [37] proposes a more general decomposition method and states that the LMP can be theoretically and uniquely decomposed into independent components associated with generators and constraints. [32] presents a DCOPF-based LMP calculation model and derives explicitly three components (namely, energy, loss, and congestion), which are consistent with industrial practices [11, 14, 13]. Although the LMP and its congestion component are independent on the selection of the slack bus, the decomposed components of energy and loss are still reference bus dependent. The explicit formula of calculating three components using a distributed reference bus is derived in [38]. Another decomposition approach is presented in [39] which achieves the slack bus independent loss component. The differences in the congestion components between any two buses are also independent to the reference bus. [40] presents a general formulation for evaluating LMP components by using the concept of marginal nodes, to which an unbinding generator is connected. The comprehensive framework includes various published decomposition methods by defining the policy for marginal nodes. Note that although the decomposition method is not unique, the decomposition policy adopted for marginal nodes determines the decomposition for all non-marginal nodes.

2.2. Congestion and Price Prediction

The majority of the work associated with the prediction of the power market status has concentrated on price forecasting because of the importance of price signals to LMP market

participants and operators. In contrast, no work has yet been reported regarding the prediction of important power system statuses such as congestions and shift of marginal units in power markets. The prediction of this type of information is also crucial in delivering more detailed insights about the potential power system status in the projected time horizon.

LMP forecasting has been a hot research area as the forecasted price plays a key role in the decision making process for both market participants and market operators. [45] lists a few of the typical applications. Power suppliers and consumers use the forecasted price to optimize the profit in the day-ahead market and bilateral contracts. Facility owners rely on the forecasted price to make investment decisions. The ISO could use the forecasted price to evaluate power market indices such as the Lerner Index. [49] presents an application from a power producer's perspective to formulate an optimal bidding strategy utilizing the forecasted price.

Compared to the load, the electricity price in the power market is much more volatile. Factors contributing to price volatility include change of fuel price, load uncertainty, fluctuations in hydro and renewable power production, generation outages, transmission congestion, market participant behavior and so on [6]. A study shows the price forecasting error was 10% or more compared to a 3% error for load forecasting [6].

Price forecasting methods can be roughly classified into three groups. The first is statistical method, for instance, the time-series model, econometric model, and regression model, which fit a predefined mathematical formula to historical data. The underlying assumption is that observations closer in time tend to be more closely related than observations further apart. [48] employs the dynamic regression approach to establish a price prediction model which relates the future price with the past price, and past demand as well. Adjustments are made to the input data to minimize the effects of data outliers due to unexpected events and therefore, achieve a better prediction of performance. The authors

report the 24 hour ahead price forecasting error is 5% for the Spanish market and 3% for the California market, for the studied weeks.

The second method employs artificial intelligence (AI) techniques including the Neural Network (NN) and fuzzy systems to predict price, which normally involves a training stage based on historical data. [46] proposed an NN-based forecasting approach and uses a similar day method to select the proper input data, through which each hour of the forecast day has a separate set of similar days. The method was tested for the PJM market and reported to produce accurate results. An adaptive wavelet NN-based price forecasting method is presented in [47], which is capable of mapping the input-output space by adapting the shape of the wavelet basis function, of the hidden layer neurons, to training data. The method was tested on the Spanish market and concluded to be superior to other forecasting techniques, such as the Auto wavelet-Regression Integrated Moving Average (ARIMA), multi-layer perceptron (MLP), Radial Basis Function Neural Network (RBFNN), and Fuzzy Neural Network (FNN). Reference [45] proposed a forecasting method which combines the fuzzy inference system (FIS) and least squares estimation (LSE). Input data included temporal indices, historical price, area loads at current and previous hours, and transmission constraints of the current hour. This method claimed to have the advantage of high accuracy and explicit reasoning.

The third model of price forecasting is the simulation method, which is presented in [44]. The simulation method utilizes a transmission constrained market simulation program that mimics the actual dispatch and market clearing process while explicitly taking into account the system operating constraints. However, this method requires intensive data input/output such as transmission model, SCED, generation unit data, transaction data, and involves intensive computational efforts.

Statistical and AI methods rely heavily on the data of past events, and prediction results are less certain as the forecast lead time is longer [48]. The selection of input data is also crucial for predicting performance since the input data should demonstrate the pattern that is expected for the forecasted time. Manual picks, or techniques such as the similar day method, may be needed toward this requirement. The reason for these limitations lies in the fact these methods are basically black-box methods, which tend to discover the correlation between future price and the most significant factors such as the past price, load, and congestion index [6], while the internal model, which indeed relates these factors, is ignored. In fact, the correlation could be discovered by studying the OPF model, which explicitly models the interaction among all factors, including electricity price. Unlike the load, which is hard to establish in a model to study its behavior, electricity price (namely, LMP) is the shadow price of the OPF problem and has a definite formulation to study of price behavior. In addition, although the future price could be estimated by drawing patterns from historical data, the electricity price at any future time has no memory effect and is essentially independent of the past price and system conditions. In other words, the price is only determined by the operating conditions at that particular future time.

Therefore, instead of pursuing a black-box prediction method, this work provides a white-box method for the prediction of electricity price, as well as important system status such as congestions and change of marginal units. It should be pointed out that, although the white-box method employs OPF as a prediction tool, similar to the aforementioned simulation method in [44], it is different in the sense that it explores the solution feature of the OPF with respect to parameter variation, and therefore, saves a significant amount of computational time; while the simulation method performs intensive calculations on each presumed condition, even if the condition has similar characteristics to a solved condition, and the new solution could be easily obtained from the previously solved solution.

2.3. Generation Sensitivity and LMP Sensitivity

In order to predict electricity price and its spikes through exploring the solution features of OPF, we have to find the internal factors embedded in the OPF solutions that directly affect the price and trigger price spikes.

In doing a comparison between the DCOPF and ACOPF in revealing congestion patterns in [24], an interesting statement is made: “once a constraint becomes binding, it will have a discrete, potentially large, impact on the bus LMPs. … large LMP differences do not necessarily indicate large deviations in the power system solutions.” It is true that the OPF solution itself is not a good indicator of the LMP, while other by-products of the OPF solution, such as binding constraints, could serve the purpose of an LMP indicator. In order to find the potential binding constraints, such as transmission line congestion and the marginal unit reaching its output limit, the sensitivities of generation output and line flow need to be explored.

The PTDF, or GSF, is a type of generation sensitivity. GSF reflects the power flow change pattern with generation variation [2, 6]; however, it does not address how the generation responds, under the cost-driven economic dispatch, to system state variations like load changes. Moreover, the GSF does not take into account the network constraints and generation limits when the load changes. On the other hand, in operation and planning, it is extremely helpful for the dispatchers to know how the power flow will change with respect to the assumed load change pattern under the OPF-based economic dispatch framework.

Mathematically, generation sensitivity is defined as the incremental generator output with respect to infinitesimal load changes at the given operating point, while optimal power flow criteria are still satisfied.

Recent work in [41] presents a generalized, ACOPF-based model for LMP sensitivity with respect to the load and other variables. The approach could be easily applied to calculate other sensitivities, such as generation sensitivity. The approach applies infinitesimal perturbations to the given operation condition, or mathematically, make a derivative to the optimality condition of the system, and then obtain the sensitivity values numerically by solving the slightly perturbed system. However, sensitivities have to be numerically calculated by solving a set of equations, and there is no direct, explicit formulation for the sensitivity, which limits the use of this method in the analytical analysis context.

Ref. [39] employs a similar approach as in [41], while giving an analytical formulation of generation sensitivity. However, the formulation is written as a function of the state variables, which themselves are yet unknown variables at scenarios other than the given operating point. This implies that the sensitivity formulation also has to be evaluated individually and locally. For example, when the load grows and the operating point moves, it is not known if the new sensitivity at the new operating point will be larger or smaller when compared with the previous value, unless the sensitivity is numerically evaluated against at the new operating point. To study the sensitivity change pattern in a range of load, a sensitivity calculation has to be done at multiple sample points. This kind of brute-force approach is not computationally effective. Moreover, since the sensitivity formulation involves state variables that depend on the parameter settings (such as load level); it could not be presented as a function with only one independent variable, the studied parameter. Therefore, it does not offer practical help for prediction under parameter variation (such as load variation).

In addition, the sensitivities obtained through these sensitivity analysis approaches [41, 39] are only valid when the perturbation is small enough that no change of marginal units under the load variation is present. It does not address the issue when the load continuously grows beyond the next critical load level (CLL) where a new binding limit occurs.

2.4. Load Uncertainty Impact

The congestion and price prediction methodology proposed in this work relies on the load forecasting results. Therefore, it is necessary to investigate how the load forecasting error is passed along in the prediction model and its impact on the prediction results.

Although no existing work has been found that specifically addresses the impact of load uncertainty on the LMP, a number of works have been conducted on other economic impacts of load forecasting uncertainty. [50] studies the impact of short term load forecasting from a utility perspective. A small load forecasting error may incur significant costs if it leads to the commitment of an extra unit or purchasing power from neighboring utilities. Results suggest up to 5% error in load forecasting is acceptable since the economic benefit from further reducing the error is small. It is also shown that the greatest benefit in load forecasting error reduction is the improved peak load forecasting. [51] assesses the economic cost of inaccurate load forecasts through a detailed simulation analysis which includes the load forecasting simulation, unit commitment, and economic dispatch models. Despite dependence on the load and generation characteristics, economic value is evaluated for specific systems. Conclusions are drawn that a 1% reduction in the mean absolute percentage error (MAPE) decreases generation costs by approximately 0.1%~0.3%, when MAPE is in the range of 3%~5%. [54] employs similar approaches to assess the economic impact of load forecast errors, taking into account the energy not served due to generation outages. [52] studies the effect of temperature, generation availability, and load on the estimation of the generation production cost. It is discovered that load uncertainty accounts for a large portion of the variance of production costs. [53] points out that, for a specific reliability level, a higher capacity reserve is needed to satisfy an uncertain load than to serve a known load. Moreover, the load forecast uncertainty has a larger impact on deficient generation and transmission

systems than on strong generation and transmission networks. [55] defines a risk index based on the standard deviation of the load increment to quantify the impact of load forecasting uncertainty.

3 Optimal Power Flow Problem and LMP Calculation

3.1. Chapter Introduction

Several technical assumptions are made in this chapter:

- The ex-ante LMP calculation is explicitly considered. The ex-post LMP is not of interest in this work. In fact, the principal of the ex-post LMP is to encourage generation units to stick with the scheduled dispatch. Hence, from a forecasting or prediction viewpoint, the models used in this chapter indirectly apply to the forecast in the ex-post LMP market from the long-run viewpoint.
- The capacity reserve constraint is not modeled in the OPF formulation. The reason is that the capacity reserve issue is assumed to have already been addressed in the Unit Commitment.
- The generator ramp rate constraint is not modeled in the OPF formulation for simplicity. In fact, it is the same type of constraint as the generation output limit constraint. Therefore, it could be easily modeled by revising the generator limit constraint.
- Generation cost is assumed to have a linear model. In fact, a quadratic cost curve can be represented with piece-wise-linear curves to ensure the application of the LP, as evidenced by the industrial LMP simulators mentioned previously.

The following assumptions are made for notational convenience:

- The generator-side market is assumed in the OPF formulation. It means no demand elasticity is considered and the OPF objective is converted to minimize the total generation cost. Although the load bids are ignored, they could actually be modeled as dispatchable negative generation, which fits the generator-side OPF model.

- N-1 constraints are not explicitly modeled in the OPF formulation to provide a better illustration. These constraints have the same mathematical characteristics (normally modeled as linear algebraic inequalities) as the normal state constraints, but having different representations and illustrations.
- Each bus has one generator and one load for simplicity of discussion. The actual implementation can be more complicated since multiple generators and/or loads may be connected to an individual bus.
- Occasionally, each transmission constraint (thermal, contingency, or nomogram) is modeled as if it has a unidirectional limit, for simple formulation although it may have bidirectional limits in reality. The actual implementation should have two equations for each bidirectional transmission limit.
- A single-block generation cost (or bid) model is assumed, while in reality a monotonically-increasing multi-block model is commonly used. This assumption is also for simplicity of illustration. The multi-block model requires additional computational and modeling effort, but does not change the mathematical kernel.
- Dispatch is performed on an hourly base so capacity (MW) is numerically the same as energy (MWh).

The above assumptions and simplifications are mainly for notational convenience and do not change the mathematical kernel of this work. The actual implementation, for the tests presented in this work, employs more complicated models such as multiple generators at a bus and bi-directional limits of the transmission lines.

OPF models can be grouped into two types according to the way the power flow is modeled: Alternating Current OPF (ACOPF) and Direct Current OPF (DCOPF). The ACOPF model presents an optimization problem with a full AC power flow model, in which both the

active and reactive power balances are considered. The DCOPF model utilizes the DC power flow model, which only preserves the active power balance when certain conditions are met. Traditionally, the DCOPF does not address power loss, and therefore, is referred to as the "Lossless DCOPF". Few works have been reported on the loss modeling in DCOPF [33, 32], and this type of improved model is called the "DCOPF with loss model" in this work. After these traditional models are introduced, a new DCOPF-based model is proposed in an attempt to achieve a better tradeoff between computation speed and result accuracy.

3.2. Traditional Optimal Power Flow Models

3.2.1. Lossless DCOPF Model

The generic DCOPF model [2], without the consideration of losses, can be easily modeled as the minimization of the total production cost subject to energy balance and transmission constraints. The voltage magnitudes are assumed to be unity and reactive power is ignored. This model may be written as a Linear Programming (LP) formulation

$$\text{Min} \sum_{i=1}^{N} c_i \times G_i \tag{3.1}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} G_i = \sum_{i=1}^{N} D_i \tag{3.2}$$

$$\left| \sum_{i=1}^{N} GSF_{k-i} \times (G_i - D_i) \right| \le Limit_k \text{, for k=1, 2, …, M} \tag{3.3}$$

$$G_i^{\min} \le G_i \le G_i^{\max} \text{, for i = 1, 2, …, N} \tag{3.4}$$

where

N = number of buses;

M= number of lines;

c_i = generation cost at Bus i ($/MWh);

G_i = generation dispatch at Bus i (MWh);

G_i^{max}, G_i^{min} = the maximum and minimum generation output at Bus i;

D_i = demand at Bus i (MWh);

$GSF_{k\text{-}i}$ = generation shift factor to line k from bus i;

$Limit_k$ = transmission limit of line k.

It should be noted that the actual GSF values depend on the choice of slack bus, although the line flow in (3.3) is based on the GSF is the same with different slack buses.

Equations (3.1)~(3.4) can be represented in matrix formulation as follows,

$$\min_{\mathbf{G}} \quad \mathbf{c}^T \times \mathbf{G}$$

$$s.t. \quad \mathbf{1}^T \times \mathbf{G} = \mathbf{1}^T \times \mathbf{D}$$

$$\left|\mathbf{GSF} \times \mathbf{G} - \mathbf{GSF} \times \mathbf{D}\right| \le \mathbf{Limit}$$

$$\mathbf{G}^{\min} \le \mathbf{G} \le \mathbf{G}^{\max}$$

3.2.2. ACOPF Model

As a comparison, a model based on the ACOPF is presented. This is not a typical model for market price simulation purposes due to its relatively slow computational speed and convergence problem in a fairly large system. Rather, it is used more often for comparison and illustration.

Generally, the ACOPF model can be presented as minimizing the total generation cost, subject to nodal real power balances, nodal reactive power balances, transmission limits, generation limits, and bus voltage limits. Details may be found in [2]. The LMP at each bus from the ACOPF formulation is the Lagrange multiplier of the equality constraints of the nodal real power balance [24, 41].

In general the ACOPF model can be formulated as

$$\text{Min } \sum c_{Gi} \times P_{Gi} \ . \tag{3.5}$$

Subject to:

$$P_{Gi} - P_{Li} - P(V,\theta) = 0 \text{ (Real power balance)} \tag{3.6}$$

$$Q_{Gi} - Q_{Li} - Q(V,\theta) = 0 \text{ (Reactive power balance)} \tag{3.7}$$

$$\left|F_k\right| \le F_k^{\max} \text{ (Line flow MVA limits)} \tag{3.8}$$

$$P_{Gi}^{\min} \le P_{Gi} \le P_{Gi}^{\max} \text{ (Gen. real power limits)} \tag{3.9}$$

$$Q_{Gi}^{\min} \le Q_{Gi} \le Q_{Gi}^{\max} \text{ (Gen. reactive power limits)} \tag{3.10}$$

$$V_i^{\min} \le V_i \le V_i^{\max} \text{ (Bus voltage limits)} \tag{3.11}$$

where

c_{Gi} = cost of generator *Gi*;

P_{Gi} , Q_{Gi} = real and reactive output of generator *Gi*;

P_{Gi}^{min}, P_{Gi}^{max} = min and max limit of P_{Gi};

Q_{Gi}^{min}, Q_{Gi}^{max} = min and max limit of Q_{Gi};

P_{Li} , Q_{Li} = real and reactive demand of load *Li*;

F_k , F_k^{max} = line flow and its maximum limit at line *k*;

V_i^{min}, V_i^{max} = min and max voltage limit at bus *i*.

The LMPs from the above formulation are the Lagrange multipliers of the equality constraints as shown in equation (3.6).

3.2.3. DCOPF with Loss Model

Earlier studies of LMP calculations with the DCOPF ignores the line losses. Thus, the energy price and the congestion price follow a perfect linear model with a zero loss price. However, challenges arise if losses need to be considered to calculate the marginal loss component in the LMP, especially considering the significance of marginal loss which may be up to 20% different among the different zones in the New York Control Area, based on actual data [31]. The prmary challenge of the loss model lies in that the conventional, lossless DC model represents a linear network, but lacks the capability to calculate marginal loss pricing, an important component in the LMP methodology.

3.2.3.1. Loss Factor and Delivery Factor

The key to considering the marginal loss price is the marginal loss factor, or loss factor (LF) for simplicity, and the marginal delivery factor, or delivery factor (DF). Mathematically, it can be written as

$$DF_i = 1 - LF_i = 1 - \frac{\partial P_{Loss}}{\partial P_i} \tag{3.12}$$

where

DF_i = marginal delivery factor at bus i;

LF_i = marginal loss factor at bus i;

P_{Loss} = total loss of the system;

$P_i = G_i - D_i$ =net injection at bus i.

The loss factor and delivery factor can be calculated as follows. Based on the definition of loss factor, we have

$$P_{Loss} = \sum_{k=1}^{M} F_k^2 \times R_k \tag{3.13}$$

$$\frac{\partial P_{Loss}}{\partial P_i} = \frac{\partial}{\partial P_i}\left(\sum_{k=1}^{M} F_k^{\,2} \times R_k\right) \tag{3.14}$$

where

F_k= line flow at line k;

R_k = resistance at line k.

In a linear DC network, a line flow can be viewed as the aggregation of the contribution from all power sources (generation as positive source and load as negative source), based on the superposition theorem. The sensitivity of the contribution to a line flow from a bus is known as the Generation Shift Factor (GSF). This can be written as

$$F_k = \sum_{j=1}^{N} GSF_{k-j} \times P_j \,. \tag{3.15}$$

Equation (3.15) can be utilized to further expand the LF as below

$$\begin{aligned} \frac{\partial P_{Loss}}{\partial P_i} &= \sum_{k=1}^{M} \frac{\partial}{\partial P_i}\left(F_k^{\,2} \times R_k\right) = \sum_{k=1}^{M} R_k \times 2F_k \times \frac{\partial F_k}{\partial P_i} \\ &= \sum_{k=1}^{M} 2 \times R_k \times GSF_{k-i} \times \left(\sum_{j=1}^{N} GSF_{k-j} \times P_j\right) \end{aligned} \,. \tag{3.16}$$

Interestingly, the loss factor at a bus may be positive or negative. When it is positive, it implies that an increase of injection at the bus may increase the total system loss. If it is negative, it implies that an increase of injection at the bus may reduce the total loss. For example, Figure 3.1 shows a simple three-bus system with Bus B as the reference bus. If there is a hypothetical injection increase at Bus A, and the increased injection is absorbed by the reference bus (or the two load buses proportionally), the line flows, as well as the losses, will increase. Hence, the loss factor at Bus A is positive. If there is a hypothetical injection increase at Bus C and it is absorbed by the reference bus (or the two load buses proportionally), this will reduce the Line BC flow and thus, reduce the system loss. So, the loss factor at Bus C is negative.

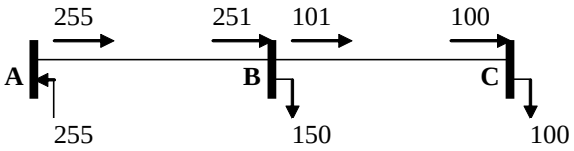


Figure 3.1. A Three-Bus System with Bus B as the reference bus

Consequently, if a loss factor is positive, the corresponding delivery factor is less than 1. If marginal loss factor is negative, the marginal delivery factor is greater than 1.

3.2.3.2. DCOPF Algorithm Considering Marginal Loss

As shown in Eq. (3.16), the loss factor depends on the net injection, P_j, which is the actual dispatch minus the load at Bus *j*. On the other hand, the generation dispatch may be affected by loss factors since different generators may be penalized differently, based on their loss factors.

Since P_j is unknown prior to performing any dispatch, one way to address this is to have an estimation of the dispatch to obtain an estimated LF at each bus. Then, the estimated loss factors will be used to obtain new dispatch results. This logical reasoning leads to the proposed iterative DCOPF approach. In other words, in the $(l+1)^{st}$ iteration, the dispatch results from the l^{th} iteration are used to update P_j and, therefore, the loss and delivery factors. Here, in each iteration, a LP-based DCOPF is solved. The iterative process is repeated until the convergence stop criteria are reached. After convergence, the LMP can be easily obtained from the final iteration. Certainly, the very first iteration is a lossless DCOPF in which the estimated loss is zero. The algorithm can be formulated as follows

$$\text{Min} \sum_{i=1}^{N} c_i \times G_i \tag{3.17}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} DF_i^{est} \times G_i - \sum_{i=1}^{N} DF_i^{est} \times D_i + P_{loss}^{est} = 0 \tag{3.18}$$

$$\left| \sum_{i=1}^{N} GSF_{k-i} \times (G_i - D_i) \right| \le Limit_k \text{, for k} \in \{\text{all lines}\} \tag{3.19}$$

$$G_i^{\min} \le G_i \le G_i^{\max} \text{, for i} \in \{\text{all generators}\} \tag{3.20}$$

where

DF_i^{est} = delivery factor at Bus *i* from the previous iteration;

$P^{est}_{Loss} = P_{loss}$ from the previous iteration.

As previously mentioned, the delivery factor, DF_i, is calculated based on the dispatch result, i.e., G_i, from the previous iteration. Therefore, the loss and delivery factors are updated iteratively since they are related to the actual generation dispatch. Once converged, the estimated DF_i and P_{loss} from the next-to-last iteration will be the same as the final values. It is not surprising that this iterative algorithm provides more accurate results with a longer running time than the lossless DCOPF. The number of iterations is acceptable. The tests in later Sections show that the iterative DCOPF (with the proposed FND model also in the later section) requires 4 iterations to converge for the PJM 5-bus system, even if a very low tolerance of 0.001 MW is applied for high accuracy. If compared with the ACOPF, the iterative DCOPF model is still much faster than ACOPF, which may be up to sixty times slower than DCOPF [24]. In addition, the ACOPF needs more careful work in preparing accurate input data to make it converge. Therefore, the iterative DCOPF model should still be of advantage to the ACOPF, especially for simulation and planning purposes.

Note that in real-time operation, delivery factors can be quickly obtained from real-time EMS/SCADA data. Unfortunately, this is not a viable option for simulation or planning study. Therefore, it is necessary to identify a feasible approach, such as the iteration method, to obtain more accurate delivery factors for simulation and planning purposes. This is consistent

with the observations in [29], which shows that it is not advisable to apply penalty factors from a typical scenario with the DC model to all other cases.

After obtaining the optimal solution for generation scheduling, the LMP, at any bus *B*, can be calculated with the Lagrangian function. This function and LMP can be written as

$$\begin{aligned}\psi = &\left(\sum_{i=1}^{N} c_i \cdot G_i\right) - \lambda\left(\sum_{i=1}^{N} DF_i \cdot G_i - \sum_{i=1}^{N} DF_i \cdot D_i + P_{loss}\right) \\ &- \sum_{k=1}^{M} \mu_k \left(\sum_{i=1}^{N} GSF_{k-i} \times (G_i - D_i) - Limit_k\right)\end{aligned} \tag{3.21}$$

$$\begin{aligned}LMP_B = \frac{\partial \psi}{\partial D_B} &= \lambda \cdot DF_B + \left(\sum_{k=1}^{M} \mu_k \times GSF_{k-B}\right) \\ &= \lambda + \left(\sum_{k=1}^{M} \mu_k \times GSF_{k-B}\right) + \lambda (DF_B - 1)\end{aligned} \tag{3.22}$$

where

LMP_B= LMP at Bus B

λ = Lagrangian multiplier of Eq. (3.18) = energy price of the system = price at the reference bus;

μ_k = Lagrangian multiplier of Eq. (3.19) = sensitivity of the k^{th} transmission constraint.

From (3.22), the LMP can be easily decomposed into three components: the marginal energy price, marginal congestion price, and marginal loss price. The LMP formulation can be written as Eqs. (3.23)-(3.26), which are consistent with industry practices [11, 14].

$$LMP_B = LMP^{energy} + LMP_B^{cong} + LMP_B^{loss} \tag{3.23}$$

$$LMP^{energy} = \lambda \tag{3.24}$$

$$LMP_B^{cong} = \sum_{k=1}^{M} GSF_{k-B} \times \mu_k \tag{3.25}$$

$$LMP_B^{loss} = \lambda \times (DF_B - 1) \tag{3.26}$$

3.2.3.3. On the Equality Constraints of Energy Balance

Note that the P_{loss}^{est} in Eq. (3.18) is used to offset the doubled system loss caused by the (marginal) Loss Factor, *LF*, and the (marginal) Delivery Factor, *DF*. The inclusion of the P_{loss}^{est} eliminates the over-estimated loss issue reported in [33]. This is consistent with the fact that the marginal loss (injection multiplied by marginal loss factor) is twice the actual loss (also referred to as the average loss) in the DC model, since the line loss is linearly related to the square of the bus injection. The rigorous proof of the validity of Equation (3.18) is given as follows

$$
\begin{aligned}
&\sum_{i=1}^{N} DF_i \times G_i - \sum_{i=1}^{N} DF_i \times D_i \\
&= \sum_{i=1}^{N} DF_i \times (G_i - D_i) = \sum_{i=1}^{N} DF_i \times P_i = \sum_{i=1}^{N} \left(-LF_i\right) \times P_i \\
&= \sum_{i=1}^{N} \left(1 - \frac{\partial P_{loss}}{\partial P_i}\right) \times P_i = \sum_{i=1}^{N} P_i - \sum_{i=1}^{N} \left(\frac{\partial P_{loss}}{\partial P_i} \times P_i\right) \\
&= \sum_{i=1}^{N} P_i - \sum_{i=1}^{N} \left(\left(\sum_{k=1}^{M} 2R_k \times F_k \times GSF_{k-i}\right) \times P_i\right) \\
&= \sum_{i=1}^{N} P_i - \sum_{k=1}^{M} \left(2R_k \times F_k \times \sum_{i=1}^{N} (GSF_{k-i} \times P_i)\right) \\
&= \sum_{i=1}^{N} P_i - \sum_{k=1}^{M} (2R_k \times F_k \times F_k) = \sum_{i=1}^{N} P_i - 2 \cdot \sum_{k=1}^{M} (R_k \times F_k^2) \\
&= P_{loss}^{schd} - 2P_{loss}^{act} = -P_{loss}
\end{aligned}
\tag{3.27}
$$

where

$P_{loss}^{schd} = \sum_{i=1}^{N} P_i = \sum_{i=1}^{N} \left(G_i - D_i\right)$=scheduled loss;

$P_{loss}^{act} = \sum_{k=1}^{M} (R_k \times F_k^2)$ =actual loss.

In the above derivation, P_{loss}^{schd} represents the system net injection at all buses. Therefore, it is called the scheduled loss of the entire system. Meanwhile, the actual loss is represented by

P_{loss}^{act}, which is the sum of the actual losses at all lines. After the iterative approach converges, the scheduled loss should be equal to the actual loss, or $P_{loss}^{schd} = P_{loss}^{act} = P_{loss}$.

From the above derivation, it is apparent that the net injection multiplied by the loss factor, i.e., $\sum_{i=1}^{N} \frac{\partial P_{loss}}{\partial P_i} \times P_i$, provides the doubled system loss. This shows the reason that Eq. (3.18) must include an extra deduction of system loss (P_{loss}^{est}) when marginal loss factors are applied. Computationally, the actual loss value from the previous iteration is used for the current iteration to keep the linearity of the optimization formulation. The convergence criteria, i.e., the dispatch of each generator, will ensure the convergence of P_{Loss}, i.e., $P_{loss}^{schd} = P_{loss}^{act}$.

Eq. (3.18) may be verified with a sample system, shown in Figure 3.2 for illustration. The system is slightly modified from the PJM 5-bus system [11]. The generation cost at Sundance is modified from the original \$30/MWh to \$35/MWh in order to differentiate its cost from the Solitude unit for better illustration. It should be noted that the PJM 5-bus system is a realistic, yet simplified system, and is used often in several research works [19, 32].

The system may be roughly divided into two areas, a generation center consisting of Buses A and E with three low-cost generation units and a load center consisting of Buses B, C, and D with a 900MWh load and two high-cost generation units. The transmission line impedances are provided in Table 3.1, where the reactance is obtained from [11] and the resistance is assumed to be 10% of the reactance. Here, only the thermal flow limit of Line ED is considered for illustrative purpose.

Table 3.2 and Table 3.3 clearly show that the dispatch will provide doubled losses if $P^{est}{}_{loss}$ is excluded from Eq. (3.18). The result is more reasonable if $P^{est}{}_{loss}$ is included.

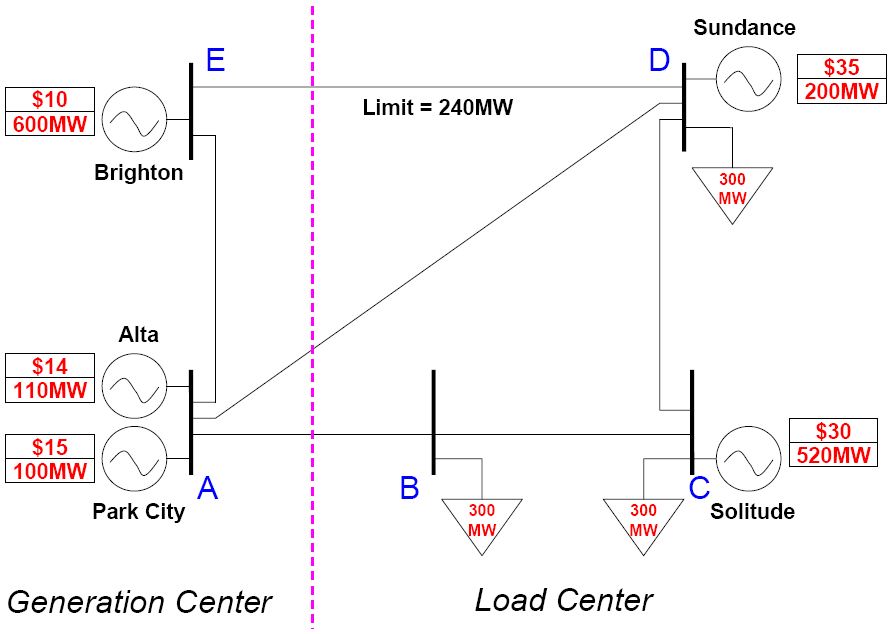


Figure 3.2. The Base Case of the PJM Five-Bus Example

Table 3.1. Line impedance and flow limits

	A-B	A-D	A-E	B-C	C-D	D-E
R (%)	0.281	0.304	0.064	0.108	0.297	0.297
X (%)	2.81	3.04	0.64	1.08	2.97	2.97
Limit (MW)	999	999	999	999	999	240

Note: Only Line D-E has a binding limit for illustrative purposes.

Table 3.2. Verification of Eq. (3.18) to avoid doubled losses caused by marginal delivery factors at the 900MW load level

	With P^{est}_{loss} in Eq. (3.18)	Without P^{est}_{loss} in Eq. (3.18)
Load	900	900
Scheduled Gen.	908.81	917.61
Scheduled losses	8.81	17.61
Actual line losses	8.81	8.81
Error	**0%**	**99.9%**

Note: (1) All units are in MWh except Error in %;

(2) Scheduled losses = Scheduled Generation – Load.

Table 3.3. Verification of Eq. (3.18) to avoid doubled losses caused by the marginal delivery factors at selected load levels

Total Load	Without P_{loss}^{est} in Eq. (3.18)			With P_{loss}^{est} in Eq. (3.18)		
	Scheduled Losses	Actual Losses	Error	Scheduled Losses	Actual Losses	Error
900	17.61	8.81	100%	8.81	8.81	0%
945	18.71	9.35	100%	9.35	9.35	0%
990	19.73	9.86	100%	9.93	9.93	0%
1035	19.02	9.51	100%	9.77	9.77	0%
1080	18.35	9.18	100%	9.42	9.42	0%
1125	17.73	8.86	100%	9.09	9.09	0%
1170	17.14	8.57	100%	8.78	8.78	0%
1215	16.60	8.30	100%	8.49	8.49	0%
1260	16.10	8.05	100%	8.22	8.22	0%
1305	15.64	7.82	100%	7.98	7.98	0%
1350	15.22	7.61	100%	7.76	7.76	0%

Note: (1) All units are in MWh except Error in percentage;

(2) Scheduled losses = $P_{loss}^{schd} = \sum_{i=1}^{N} P_i = \sum_{i=1}^{N} (G_i - D_i)$;

(3) Actual losses = $P_{loss}^{act} = \sum_{k=1}^{M} (R_k \times F_k^2)$.

3.3. FND (Fictitious Nodal Demand)-based DCOPF Model

In this section, first, an iterative DCOPF-based algorithm is presented with the Fictitious Nodal Demand (FND) model to calculate the LMP. The algorithm has three features: the iterative approach is employed to address the non-linear marginal loss; FND is proposed to eliminate the large mismatch at the reference bus if the FND is not applied; and a deduction of system loss in the energy balance equation is proved to be necessary because the net injection multiplied by the marginal delivery factors creates a doubled system loss.

Secondly, the proposed FND-based DCOPF algorithm is compared with the results from the ACOPF algorithm for accuracy of the LMP results at various load levels using the PJM 5-bus system. It is clearly shown that the FND algorithm is a good estimation of the LMP, calculated from the ACOPF algorithm, and outperforms the lossless DCOPF algorithm.

Thirdly, the FND-based DCOPF algorithm is employed to analyze the sensitivity of the LMP with respect to the system load. A simple, explicit equation of LMP sensitivity is presented and validated. A special case of infinite sensitivity under the step change of the LMP is discussed. If the operating point is close to the critical load level of the LMP step change, the sensitivity is less reliable and may not be applied to a large variation of the load.

3.3.1. Iterative DCOPF Algorithm with Fictitious Nodal Demand for Losses

3.3.1.1. Mismatch at the Reference Bus in the Traditional DCOPF with Loss Model

The above model appropriately addresses the marginal loss price through the delivery factors. However, the line flow constraints in Eq. (3.19) still assume a lossless network. Meanwhile, the system energy balance constraint in Eq. (3.18) enforces the idea that the total generation should be greater than the total demand by the average system loss. This leads to a mismatch at the reference bus because the amount of the mismatch has to be absorbed by the

system reference bus. If the amount of demand is a large amount like several GW, the system loss may be on the scale of tens to hundreds of MW. It is inaccurate to have all the loss absorbed by the reference.

Taking the PJM 5-bus sample system, the dispatch result is shown in Table 3.4 and Figure 3.3. The result shows the nodal mismatch, defined as *Nodal Generation – Nodal Demand + Line Injection* at all connected buses. Although all buses, except the reference Bus D, have zero mismatches, the mismatch at Bus D is relatively large as it absorbs the total system loss of 8.80MW. This is a centralized loss model, which means that all losses are centrally absorbed by the reference bus.

Table 3.4. Dispatch Results from the Iterative DCOPF

	G (MW)	L (MW)	Line Inj. (MW)	Mismatch (MW)
Bus A	210	0	-210	0
Bus B	0	300	300	0
Bus C	0	300	300	0
Bus D	124.88	300	183.92	8.80
Bus E	573.92	0	-573.92	0

Note: Line Injection at a bus is the sum of the flows of all connected lines. A positive sign corresponds to a net incoming flow.

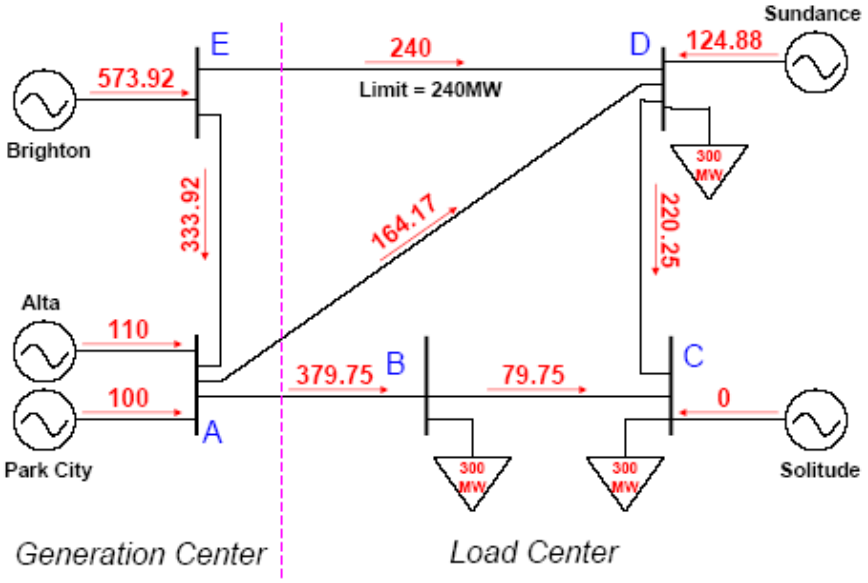


Figure 3.3. The dispatch results for the base case

3.3.1.2. FND-Based Iterative DCOPF Algorithm

To address the mismatch issue at the reference bus, it is desirable to have the line losses represented in the transmission lines. Since line flow is represented with the GSF in a LP-based DCOPF, it is challenging to include the line loss without losing the linearity of the model.

This work employs the concept of the Fictitious Nodal Demand (FND) to represent the losses of the lines connected to a bus. The FND is similar, yet different, from the fictitious load and midpoint bus model in [25]. Reference [25] uses the fictitious load and midpoint bus to partition an inter-area tie line and eventually model a multi-area OPF. This research work does not need the fictitious midpoint bus and uses a different representation of the fictitious loss model, as shown in Eq. (3.28). More important, the FND is applied here to distribute system losses into each individual line in order to eliminate a significant mismatch at the reference bus. The FND model is illustrated in Figures 3.4 and 3.5. With this approach, the

loss in each transmission line is divided into two equal halves, attached to both buses of that line. Each half is represented as if it is an extra nodal demand. For each bus, the total of all equivalent line losses is the proposed fictitious nodal demand.

Here the FND at Bus *I*, is written as E_i, and is defined as follows

$$E_i = \sum_{k=1}^{M_i} \tfrac{1}{2} \times F_k^2 \times R_k \qquad (3.28)$$

where

M_i = the number of lines connected to Bus *i*.

The line flow, $F_{k,}$ may be obtained from the FND calculation in the previous iteration. The new calculation of line flow may be formulated as

$$F_k = \sum_{j=1}^{N} GSF_{k-j} \times (G_j - D_j - E_j^{est}) . \qquad (3.29)$$

The loss factor calculation equation may remain the same, as that shown in (3.23), however, the value of F_k will be different under the new approach of the FND.

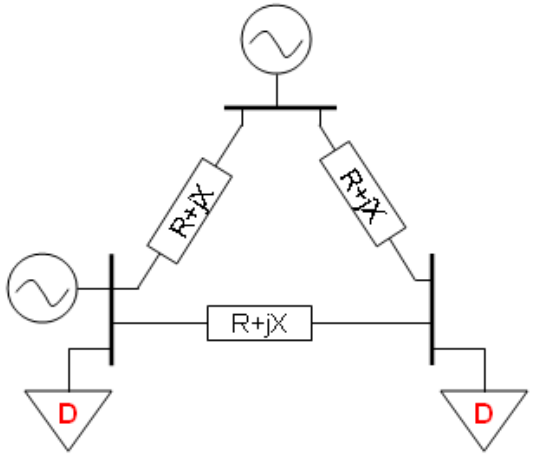


Figure 3.4. A System with line resistance

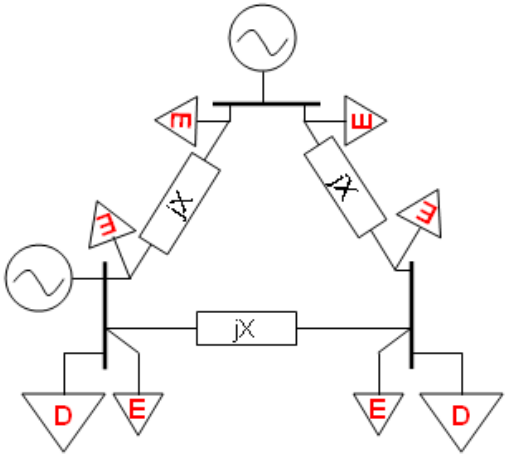


Figure 3.5. A system with the FND to represent line losses

Therefore, the new iterative DCOPF formulation, which replaces Eqs. (3.17) to (3.20), can be formulated as

$$\text{Min} \quad \sum_{i=1}^{N} c_i \times G_i \tag{3.30}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} DF_i^{est} \times G_i - \sum_{i=1}^{N} DF_i^{est} \times D_i + P_{loss}^{est} = 0 \tag{3.31}$$

$$\left| \sum_{i=1}^{N} GSF_{k-i} \times \left(G_i - D_i - E_i^{est} \right) \right| \le Limit_k \text{, for k} \in \{\text{all lines}\} \tag{3.32}$$

$$G_i^{\min} \le G_i \le G_i^{\max} \text{, for i} \in \{\text{all generators}\} \text{ .} \tag{3.33}$$

When the above formulation converges using the generation dispatch of each unit (G_i) as the convergence criterion, other parameters, like the line flows (F_k), the delivery factors (DF_i), and the system loss (P_{loss}), will converge as well. Appendix A shows a schematic proof of the convergence feature of this new algorithm.

The detailed procedure of this FND-based Iterative DCOPF algorithm is given as follows:

1. Set LF^{est}_i=0, DF^{est}_i=1, E^{est}_i=0 (for i=1, 2, …, N) and P^{est}_{loss}=0;
2. Perform generation dispatch using Eqs. (3.30) to (3.33);
3. Update LF^{est}_i, DF^{est}_i, E^{est}_i and P^{est}_{loss} using Eq. (3.19), (3.20), (3.23) and (3.28);
4. Perform another dispatch using Eqs. (3.30) to (3.33);
5. Check the results of the dispatch of each generator with that from the previous dispatch. If the difference at one or more buses is greater than the pre-defined tolerance, go to Step 3. Otherwise, go to Step 6.
6. Calculate the three LMP components using Eqs. (3.23) to (3.26).

The results of the proposed new iterative DCOPF model are shown in Figure 3.6. The total loss is distributed into each *individual* line. At each bus, the nodal generation, plus incoming flow from connected lines, and then minus nodal demand is equal to the Fictitious Nodal Demand (FND), which represents half of the losses in all connected lines. Therefore, the system loss is well distributed in each line and numerically represented by the FND at each bus. The mismatch at the reference bus, like at any other buses, is just 50% of the losses of all connected lines, not the total system loss.

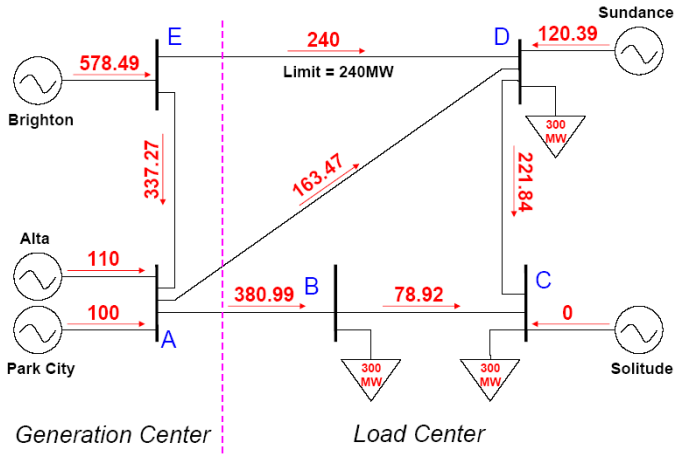


Figure 3.6. Dispatch results with the Fictitious Nodal Demand approach

Table 3.5 shows the Fictitious Nodal Demand (FND) at each bus under various load levels. The second-to-last column presents the sum of the FND at all buses, namely, the actual total loss. The last column lists the sum of net generation at all buses, namely, the scheduled total loss. Note that the actual power loss is very close, but not exactly equal, to the scheduled loss. The slight discrepancy between these two losses is due to the fact that the power flows at both ends of any transmission line are considered identical in the DCOPF model when power loss is calculated, as shown in equation (3.13).

Table 3.5. FND at each bus

Load Level	FND @A	FND @B	FND @C	FND @D	FND @E	$\sum_{i=1}^{N} E_i$	$\sum_{i=1}^{N} (G_i - D_i)$
900	2.81	2.07	0.76	1.99	1.22	8.86	8.88
945	2.98	2.23	0.87	2.09	1.25	9.42	9.43
990	3.12	2.35	0.95	2.16	1.27	9.85	9.86
1035	3.12	2.34	0.90	2.12	1.27	9.75	9.76
1080	3.11	2.32	0.84	2.07	1.27	9.62	9.63
1125	3.04	2.25	0.74	1.98	1.25	9.27	9.29
1170	2.97	2.19	0.65	1.90	1.24	8.95	8.97
1215	2.90	2.13	0.57	1.82	1.23	8.65	8.68
1260	2.83	2.07	0.50	1.75	1.22	8.37	8.40
1305	2.77	2.02	0.44	1.68	1.21	8.12	8.15
1350	2.70	1.98	0.38	1.62	1.20	7.88	7.92

Note: All units are in MWh.

3.3.2. Benchmarking the FND-based DCOPF and Lossless DCOPF Algorithms with ACOPF Algorithm

In this section, the ACOPF algorithm is briefly discussed. Then, solutions from the FND and lossless DCOPF algorithms are benchmarked against that of the ACOPF algorithm using the PJM 5-bus system.

3.3.2.1. Test Results from PJM Five-Bus System

This section provides the test results with the slightly modified PJM 5-bus system, as shown in Fig. 3.2. In the ACOPF run, all loads are assumed to have 0.95 lagging power factors. The generators are assumed to have a reactive power range between 150MVar capacitive to 150MVar inductive so that reactive power will not be a limiting issue. The ACOPF is implemented with the MATPOWER package [17].

The LMP calculations are performed using the lossless DCOPF algorithm, the FND-based Iterative DCOPF algorithm, and the ACOPF algorithm in the previous sub-section. The LMP results from the two DCOPF algorithms which are benchmarked with the ACOPF under various load levels from 1.0 per unit to 1.3 per unit of the base-case load (900MWh). Tests are performed with a step size of 0.0025 p.u. load increase. All bus loads are varied proportionally, and the same power factor is kept at each bus for the ACOPF case. Test results show that the FND algorithm quickly converges in 4-5 iterations for the PJM 5-bus case, even if a low tolerance of 0.001 MW is applied for high accuracy.

Figures 3.7 and 3.8 plot the maximum difference (MD) and the average difference (AD) of the nodal LMPs between the two models. The MD and AD of the LMP at a given load level are given as

$$MD_{LMP}(\%) = \pm \max_{i\in\{1,2,\ldots,N\}} \left\{ \left| \frac{LMP_i^{(1)} - LMP_i^{(2)}}{LMP_i^{(2)}} \right| \times 100 \right\} \tag{3.34}$$

$$AD_{LMP}(\%) = \frac{\sum_{i=1}^{N} \left| \frac{LMP_i^{(1)} - LMP_i^{(2)}}{LMP_i^{(2)}} \right| \times 100}{N} \tag{3.35}$$

where

$LMP_i^{(1)}$ = LMP from the lossless DCOPF algorithm or the FND algorithm;

$LMP_i^{(2)}$ = LMP from the ACOPF algorithm;

Sign of MD is determined by the sign of ($LMP_i^{(1)}$ - $LMP_i^{(2)}$).

Figure 3.9 depicts the Marginal Unit Difference Flag of the FND-based DCOPF algorithm and Lossless DCOPF algorithm when compared with the benchmark ACOPF algorithm. At any load level within the investigated load range, when the DC algorithm provides the same marginal unit set as the benchmark ACOPF algorithm does, the Marginal Unit Difference Flag is set to zero; and is set to one otherwise.

The MD and AD of the generation dispatch, similar to those for LMP in Eqs. 3.34-3.35, are also presented in Figs. 3.10 and 3.11.

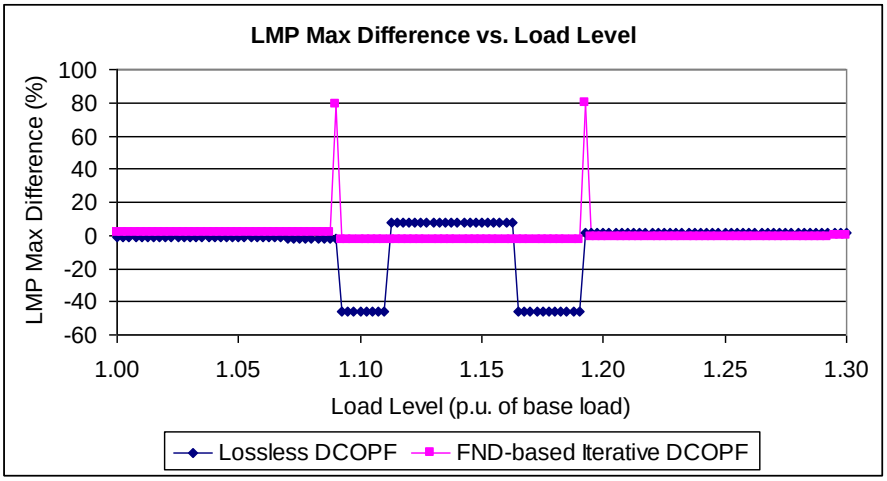


Figure 3.7. The Maximum Difference of the LMP in Percentage between each DCOPF algorithm and the ACOPF for the PJM 5-bus system

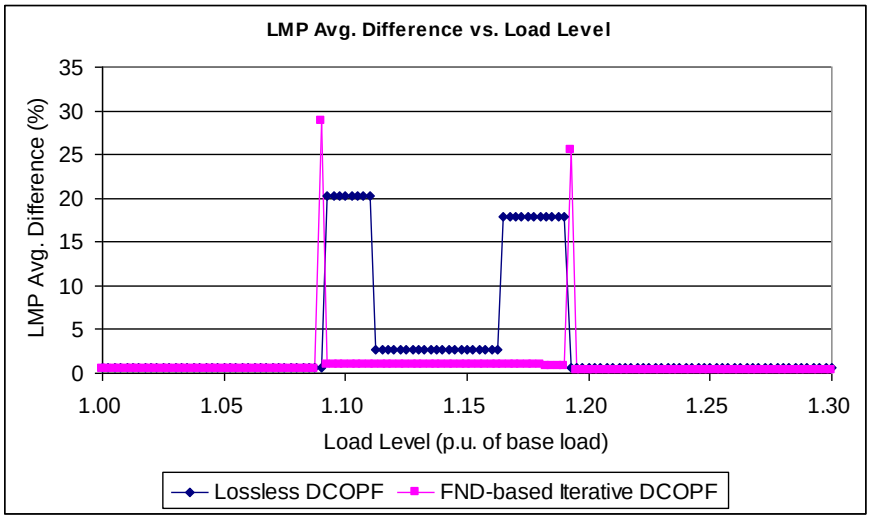


Figure 3.8. The Average Difference of the LMP in Percentage between each DCOPF algorithm and ACOPF for the PJM 5-bus system

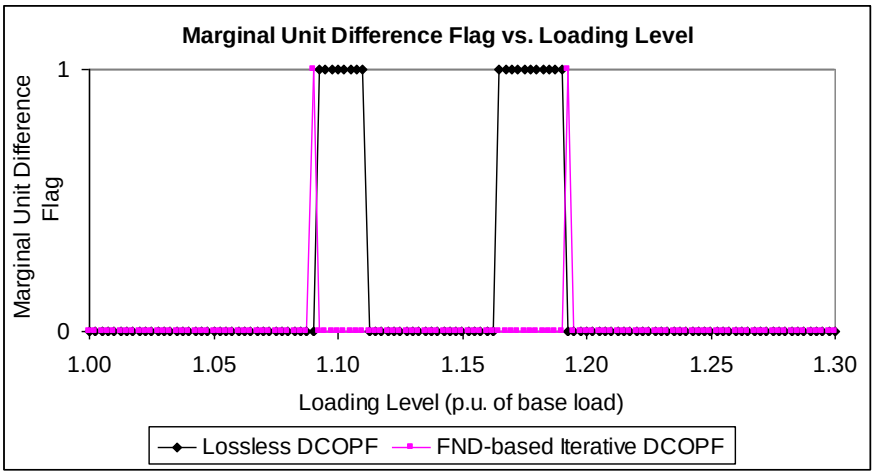


Figure 3.9. Marginal Unit Difference Flag of each DCOPF algorithm when compared with the benchmark ACOPF for the PJM 5-bus system

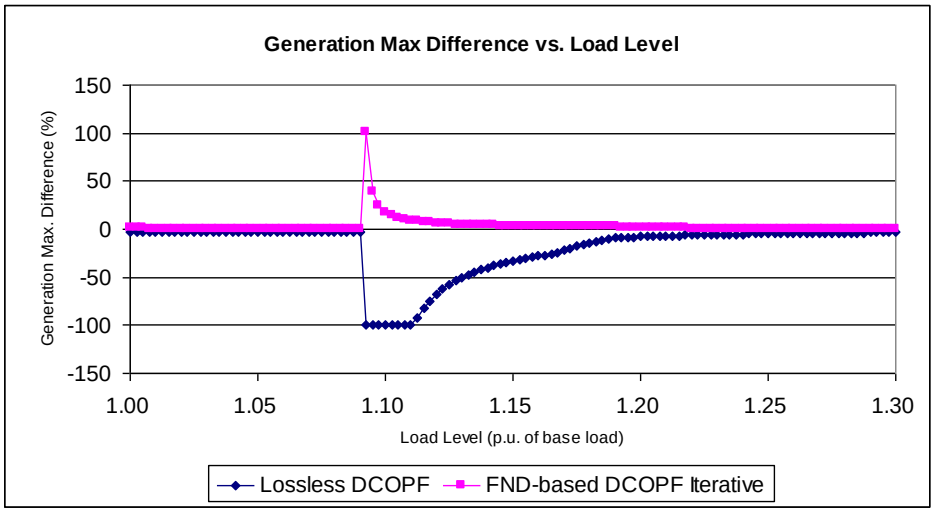


Figure 3.10. The Maximum Difference of Generation Dispatch between each DCOPF algorithm and the ACOPF for the PJM 5-bus system

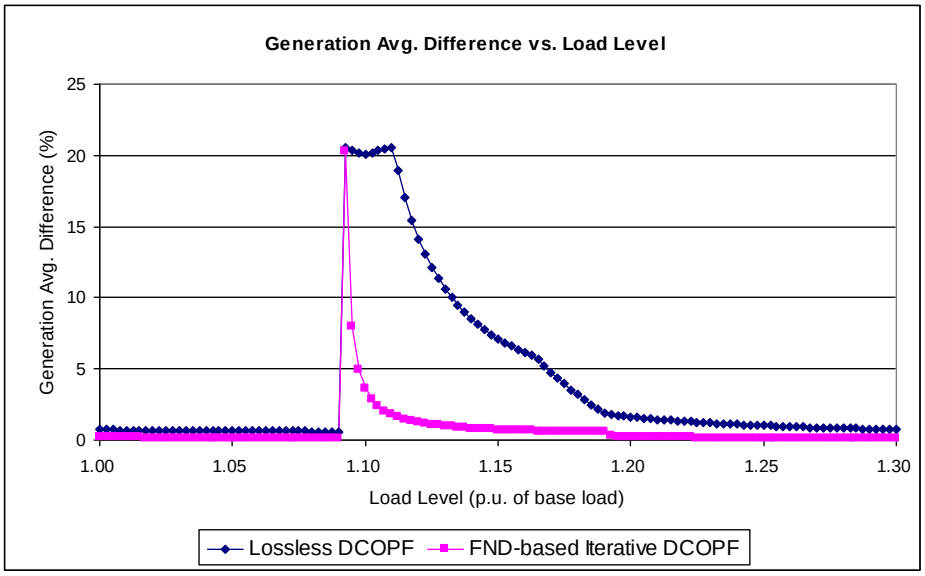


Figure 3.11. The Average Difference of Generation Dispatch between each DCOPF algorithm and ACOPF for the PJM 5-bus system

As Figs. 3.7-3.11 show, the LMP from the lossless DCOPF algorithm matches the ACOPF results for 82% of all load levels tested. This is consistent with the results reported in [24]. However, the lossless DCOPF has significant errors at 18% load levels.

The FND algorithm is superior to the lossless DCOPF algorithm when using the ACOPF as a benchmark for the LMP, as well as generation scheduling. For example, the LMP results from the FND algorithm are very close to the ACOPF LMP results with exceptions at only two particular load levels: 1.0900 and 1.1925 per unit of the base load. As a comparison, the LMP from the lossless DCOPF produces significant errors in two bands of load levels, i.e., [1.0900, 1.1125] and [1.1625, 1.1925]. Similar observations can be found in generation scheduling. Since the lossless DCOPF ignores the line loss, it is not surprising that it performs poorer than the FND-based Iterative DCOPF algorithm.

Further tests in the IEEE 30-Bus System are also performed. Observed results are very similar to the results from the PJM 5-bus system. For instance, the FND DCOPF algorithm

provides a much closer approximation than the lossless DCOPF algorithm in all four measures, MD of LMP, AD of LMP, MD of generation dispatch, and AD of generation dispatch.

3.3.2.2. Tests Results from IEEE 30 Bus System

The second test system is the IEEE 30-bus test system. The detailed system configuration and data are a revised version of the IEEE 30-bus test system [16] and available in [17]. The bidding prices of the 6 generators are assumed here to be 10, 15, 30, 35, 40, and 45, respectively, all in $/MWh. The branch susceptibilities and transformer tap ratios are all ignored for simplicity. To make the ACOPF converge beyond the load level of 1.05 per unit of the base-case load, the network data is slightly modified: 1) the load power factor is kept at 0.95 lagging as load increases; and 2) the transmission limit of Line 6-8 is increased by 10%. Test results show that the FND algorithm converges in about 5 iterations for this system, even if a low tolerance 0.001MW is applied for high accuracy.

Figs. 3.12-3.13 and 3.15-3.16 show the Maximum Difference and Average Difference of the LMP and generation dispatch between each of the two DCOPF algorithms and the ACOPF algorithm. Figure 3.14 shows the Marginal Unit Difference Flag of the two algorithms when compared with the benchmark ACOPF algorithm. Similar observations can be made that the FND algorithm performs better than lossless DCOPF and is a very good approximation of ACOPF algorithm for load levels [0.70, 1.16] and [1.22, 1.30]. Nevertheless, neither the FND nor lossless DCOPF algorithm can identify the same marginal units as the ACOPF algorithm for load levels [1.16, 1.22]. This is reasonable since DCOPF algorithms are based on the DC model assumptions and approximations. In general, the FND-based Iterative DCOPF greatly outperforms the lossless DCOPF.

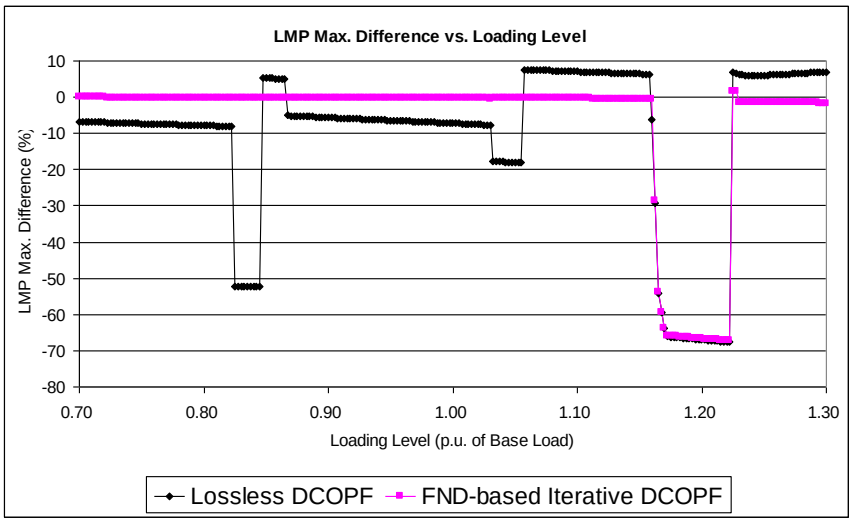


Figure 3.12. The Maximum Difference of the LMP between each DCOPF algorithm and ACOPF for the IEEE 30-bus system

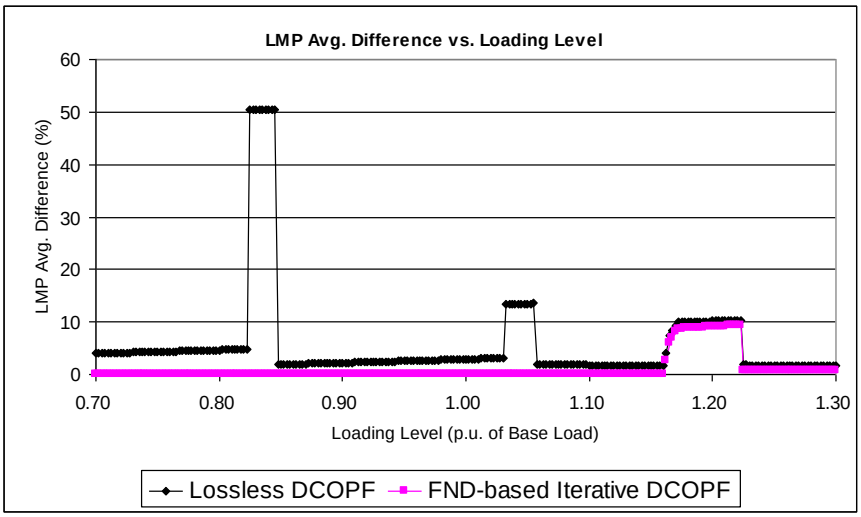


Figure 3.13. Average Difference of the LMP between each DCOPF algorithm and ACOPF for the IEEE 30-bus system

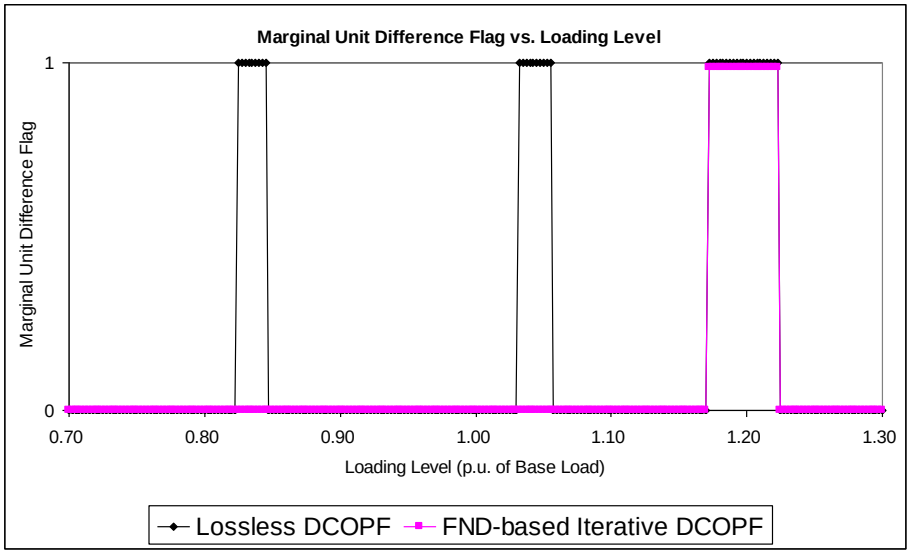


Figure 3.14. Marginal Unit Difference Flag of each DCOPF algorithm when compared with the benchmark ACOPF for the IEEE 30-bus system

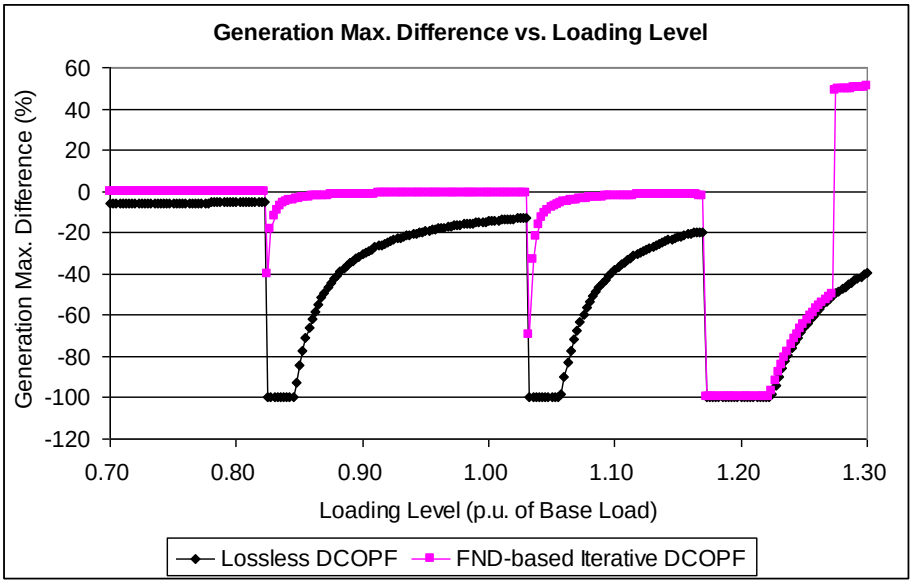


Figure 3.15. The Maximum Difference of Generation Dispatch between each DCOPF algorithm and ACOPF for the IEEE 30-bus system

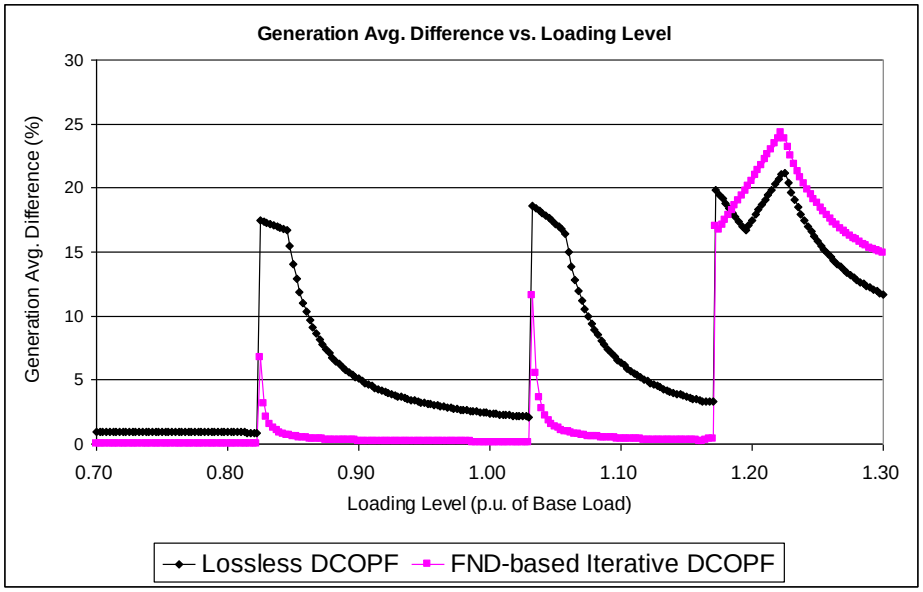


Figure 3.16. The Average Difference of the Generation Dispatch between each DCOPF algorithm and ACOPF for the IEEE 30-bus system

3.3.2.3. Discussion on the Simulation Results

A. Effects of power loss on LMP and generation dispatch

As Figs. 3.12-3.13 and 3.15-3.16 show, the LMP from the lossless DCOPF algorithm matches the ACOPF results for 82% of all studied load levels. This is consistent with the results reported in [24]. However, the lossless DCOPF causes significant errors at the other 18% load levels.

The FND algorithm outperforms the lossless DCOPF algorithm in terms of accuracy of the LMP as well as generation scheduling. For example, the LMP results from the FND algorithm are very close to the ACOPF LMP results with exceptions at only two particular load levels: 1.0900 and 1.1925 per unit of the base load. As a comparison, the LMP from the lossless DCOPF produces significant errors at two bands of load levels, i.e., [1.0900, 1.1125] and [1.1625, 1.1925]. Similar observations can be found in generation scheduling. Since the

lossless DCOPF ignores the line loss, and power loss does affect the LMP and generation scheduling, it is not surprising that it performs much more poorly than the FND-based Iterative DCOPF algorithm.

B. Occurrence of significant LMP Difference between DCOPF and ACOPF

The occurrence of significant LMP difference between the DC model and AC model is due to the different set of identified marginal units, as can be seen from Figs 3.7-3.8 and Figure 3.9 in the PJM 5-bus system and Figs 3.12-3.13 and Figure 3.14 in the IEEE 30-bus system, respectively. In the PJM 5-bus system, for instance, the load range of the significant LMP difference always lines up with the load range of different marginal unit set. On the other side, at the load level where the marginal unit set is the same between the DCOPF and ACOPF, the LMP difference is inconsiderable, since the LMP at any bus is either equal to the marginal unit price at that bus, or determined by all marginal unit cost/bidding.

For example, when the load level is 1.09 per unit of the base load, the dispatch results are as shown in Table 3.6. The base case diagram of the PJM 5-bus system is shown in Figure 3.2. With the ACOPF model, the marginal units are Sundance and Brighton. However, in the FND-based Iterative DCOPF, the Brighton unit, which is dispatched extremely close to, but not at, its maximum capacity as in the ACOPF is now dispatched at its maximum capacity. In addition, the Solitude is dispatched at a very small amount of generation. So, the marginal units for the FND algorithm are the Sundance and Solitude.

Therefore, the different marginal units lead to the LMP difference because they determine the overall trend of the LMP. In this case, the generation cost difference between the Sundance and Brighton is relatively big, i.e., ($35-$10)/MWh = $25/MWh. This leads to the considerable MD (80%) between the FND algorithm and the ACOPF algorithm at the 1.0900 load level, as shown in Fig. 3.7. However, once the DCOPF identifies the same marginal units as that of ACOPF at load levels, such as 1.0925 p.u., the LMPs will be very close.

Table 3.6. The Generation Dispatch Results from DCOPF and ACOPF at load level 1.09 p.u.

	Maximum Capacity (MW)	Cost ($/MWh)	FND-based DCOPF	ACOPF
Alta	110	14	110.00	110.00
Park City	100	15	100.00	100.00
Solitude	520	30	**0.49**	0.00
Sundance	200	35	**180.39**	**179.94**
Brighton	600	10	600.00	**599.79**
Total			990.88	989.72

Marginal Units (with Dispatch Amount in Bold Font):

FND-based DCOPF – Solitude and Sundance

ACOPF – Sundance and Brighton.

This observation has practical implication for real systems. For a system consisting of a generation center with abundant low-cost generation resources and load center with expensive generators, when the units in the low-cost, net-exporting area are approaching their maximum capacity, it is very likely that the difference between the DCOPF and ACOPF may lead to a significant price difference because the two approaches may provide different sets of marginal units. Special care, such as verification with AC model, may be necessary for system planners if the DC model is the primary approach.

Moreover, the reason for the different marginal unit sets identified by DCOPF model and ACOPF model lies in the natural difference between the DC and AC models. Since the DC model linearizes the network by setting the voltage magnitude to unity and ignoring the

reactive power, there must be some difference in the power flow calculation which causes the different marginal units and hence, affects the LMP at many buses at a particular load level.

C. Generation Dispatch Difference

As for the generation dispatch results shown in Figs. 3.10 and 3.11, the results from the FND algorithm are very close to those from the ACOPF algorithm for most cases, except for the load levels between 1.09 to 1.10. Actually, the difference is not as large as it first appears because Figs. 3.10 and 3.11 show relative difference, which can amplify the facts. For instance, when a unit is dispatched as a small value, e.g. 0.5MWh in the ACOPF, the difference percentage is as big as 100% when the FND provides 1.0MWh. In this case, the large relative difference is not as surprising as it appears.

In addition, a large dispatch difference does not necessarily correspond to a large LMP difference. For example, although the generation output is quite different at load levels [1.23, 1.30] compared to benchmark data in the IEEE 30-bus system, as observed in Figs 3.12-3.16, LMPs at these load levels are still very close. In fact, as long as the significant generation difference occurs at load levels where the marginal unit set identified by the FND algorithm is the same as that of the ACOPF, the LMP difference may not be noteworthy.

3.3.3. Sensitivity Analysis of LMP With Respect to Load

The previous section showed that the FND-based Iterative DCOPF algorithm is a trustable approximation, especially when compared with the lossless DCOPF, of the ACOPF-based LMP. This section will examine the sensitivity of the LMP with respect to load changes based on the FND-based Iterative DCOPF.

3.3.3.1. LMP Sensitivity without Loss

Based on the DCOPF formulation, sensitivity is strongly related to the loss model. In other words, if no loss is considered, the sensitivity of the LMP should be zero, if there is a very small change in demand (actually, as long as there is no change of marginal units). This is due to the linear characteristics in the DCOPF model. It can be shown as follows

$$LMP_i^{no\_loss} = LMP^{energy} + LMP_i^{cong} = \lambda + \sum_{k=1}^{M} \mu_k \times GSF_{k-i} \tag{3.36}$$

$$\frac{\Delta LMP_i^{no\_loss}}{\Delta D_j} = \frac{\Delta \lambda}{\Delta D_j} + \sum_{k=1}^{M} \frac{\Delta \mu_k}{\Delta D_j} \times GSF_{k-i} = 0 \ . \tag{3.37}$$

In the above equations, λ is independent of demand because it represents the change of the dispatch cost with respect to the change of demand. If there is a small increase of demand, the same marginal unit(s) shall provide a matching amount of power to cover the demand increase. The reason is that the DCOPF model is based on a Linear Programming model. Hence, the change of generation of each marginal unit with respect to a load change at a specific bus should also be linear. This can be written as

$$\frac{\Delta G_l}{\Delta D_j} = \alpha_{lj} \text{ for all marginal unit l.} \tag{3.38}$$

With the assumption of a small load change without new binding constraints, the energy price can be written as

$$\lambda = \frac{\Delta Cost}{\Delta D_j} = \frac{\sum_{l=1}^{M_l} (\Delta G_l \cdot c_l)}{\Delta D_j} = \sum_{l=1}^{M_l} (\alpha_{lj} \cdot c_l) \tag{3.39}$$

where M_l = number of marginal units.

Hence, λ is independent of demand and $\frac{\Delta\lambda}{\Delta D_j}$ is equal to zero. Similarly, due to the linear formulation, μ_k represents the change of cost when there is a 1MW relaxation of the k^{th} transmission constraint. As long as there is no new marginal unit, the reduced cost will remain constant or independent of D_j. Therefore, $\frac{\Delta\mu_k}{\Delta D_j}$ is equal to zero.

The LMP sensitivity is first tested on the PJM 5-bus system [11]. The base case of the system is shown in Figure 3.2. Figure 3.17 shows the nodal LMP at each bus without considering losses, with respect to different load levels, 300 MWh to 330 MWh, at Bus B. The LMPs remain constant within this small range.

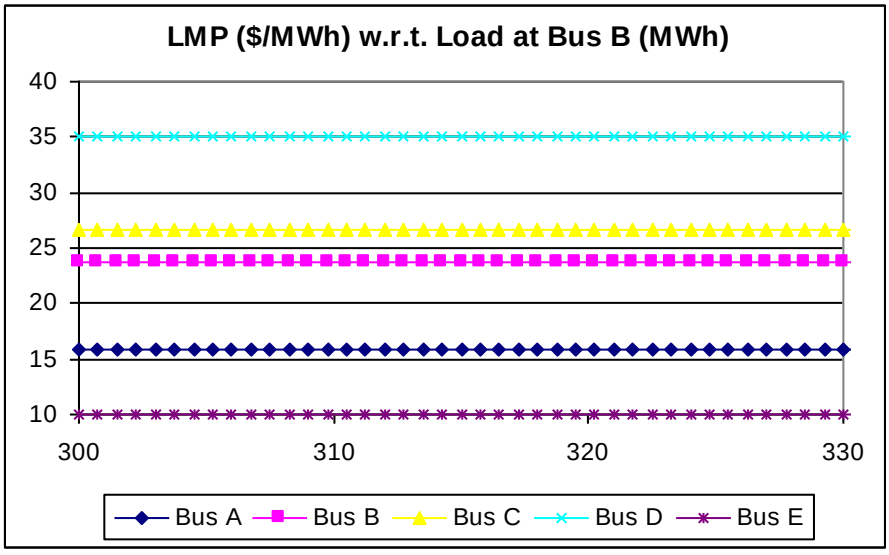


Figure 3.17. LMP from Lossless DCOPF at each bus with respect to Load at Bus B

3.3.3.2. LMP Sensitivity Considering Loss

As shown in the above analysis and test, the possible non-zero sensitivity of the LMP in the paradigm of the DCOPF must be attributed to the loss model. When the load grows, the loss grows quadratically with demand. Here, a misleading intuition is that the LMP

sensitivity to the load shall only be related to the delivery or loss factor since the LMP sensitivity is zero when there is no loss, as shown in (3.37) and Fig. 3.17.

However, a careful analysis shows that the change of load will lead to a change of not only *DF*, but λ and μ. This is because the change of the Delivery Factor in the DCOPF model shall lead to a new λ and μ, when the load is varied. In summary, the sensitivity of the LMP at bus *i* to the demand at bus *j* can be written as

$$\frac{\Delta LMP_i}{\Delta D_j} = \frac{\Delta\left(DF_i \cdot \lambda + \sum_{k=1}^{M} \mu_k \cdot GSF_{k-i} \right)}{\Delta D_j} \ .$$

Hence, we have

$$\frac{\Delta LMP_i}{\Delta D_j} = \frac{\Delta DF_i}{\Delta D_j} \cdot \lambda + \frac{\Delta \lambda}{\Delta D_j} \cdot DF_i + \sum_{k=1}^{M} \frac{\Delta \mu_k}{\Delta D_j} \cdot GSF_{k-i} \ . \tag{3.40}$$

And, $\frac{\Delta \lambda}{\Delta D_j} \neq 0$ and $\frac{\Delta \mu_k}{\Delta D_j} \neq 0$ in general. This makes the case with the loss very different from the lossless case.

Figure 3.18 shows the normalized DF at each bus for the PJM 5-bus system with respect to the Bus B Load in the range between 300 MWh and 330 MWh; Figure 3.19 shows the actual μ of Line ED with respect to the Bus B Load; and Figure 3.20 shows the normalized LMP at each bus with respect to the Bus B Load. The normalized values are used so it is easier to observe the linear growth of the DF or LMP at all buses versus the Bus B Load. So, the LMP sensitivities, with respect to load, are the slopes of the straight lines in Fig. 3.20. Fig. 3.21 plots the actual LMP sensitivity with respect to the load at Bus B.

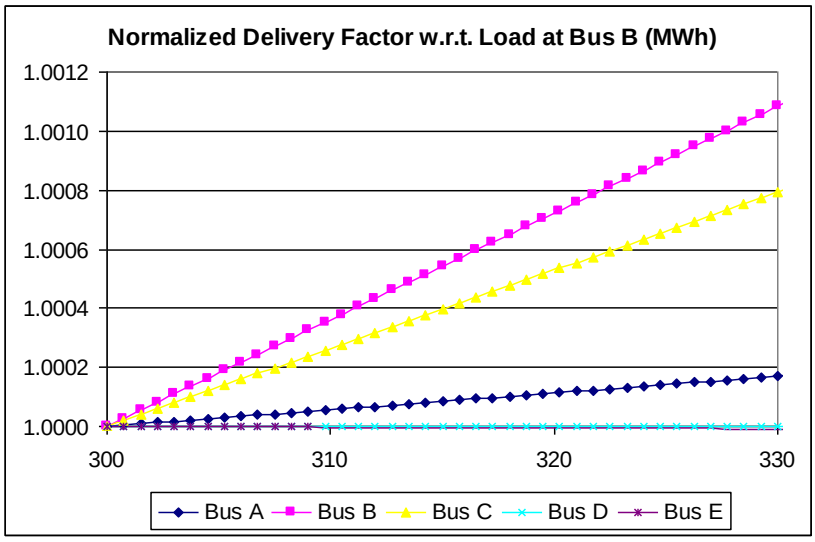


Figure 3.18. Delivery Factors normalized to base case at each bus with respect to the Load at Bus B. The DFs at Base Case for the 5 buses are 0.98992, 1.01130, 1.01304, 1.00000, and 0.98561, respectively.

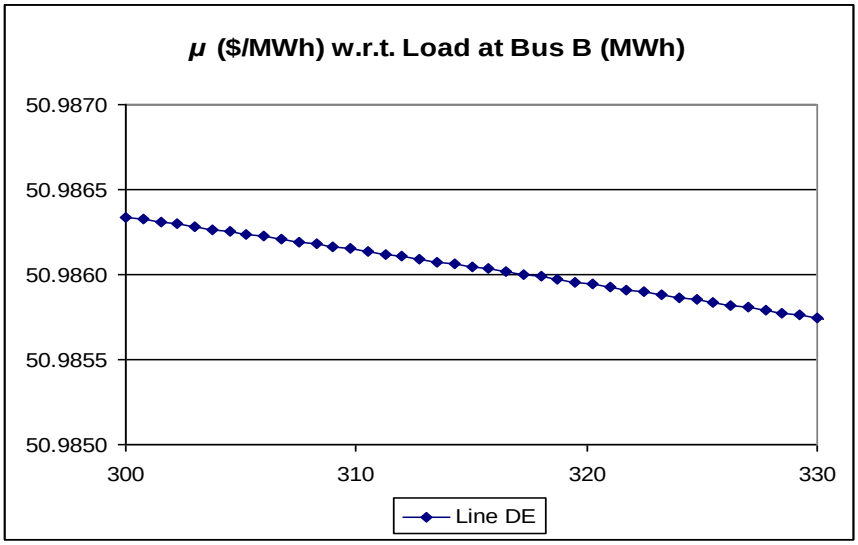


Figure 3.19. μ of the Constraint of Line ED with respect to the Bus B Load

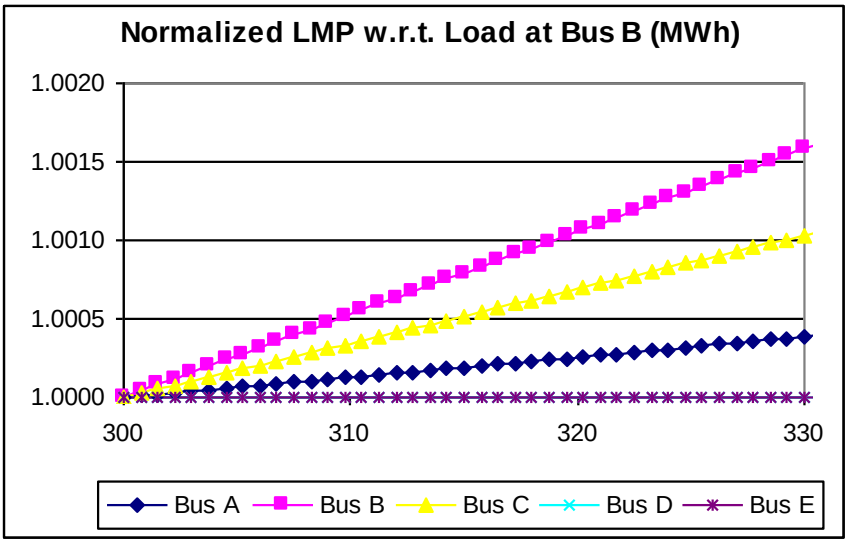


Figure 3.20. LMP normalized to base case with marginal loss at each bus with respect to the Load at Bus B. The LMPs of the base case for the 5 buses are 15.86, 24.30, 27.32, 35.0, and 10.0 $/MWh, respectively.

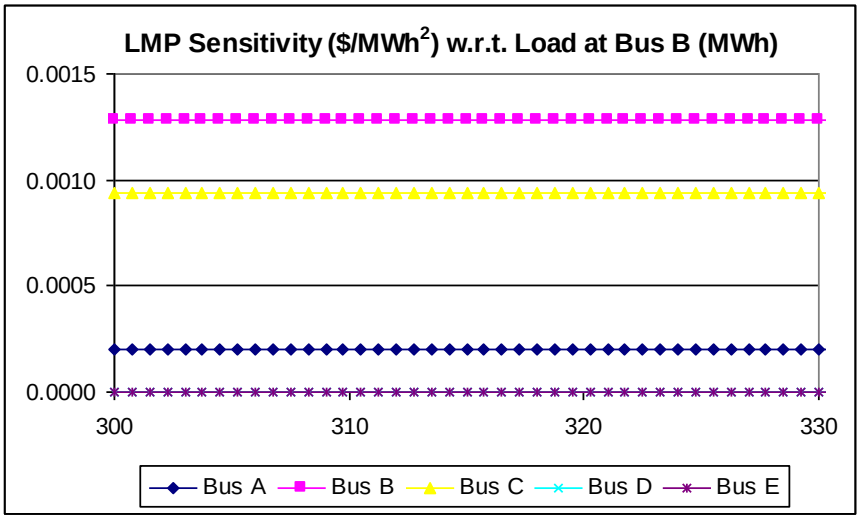


Figure 3.21. LMP Sensitivity ($/MWh^2) with respect to the Load at Bus B (MWh)

Table 3.7 shows the μ of Line ED, the DF at Bus B, the LMP at Bus B, the DF at Bus C, and the LMP at Bus C with respect to the Bus B Load from 300 MW to 330 MW. It can be easily verified that each variable is linearly related to the Bus B Load, shown in Table 3.7.

Table 3.7. μ, DF and LMP with respect to Different Load Levels at Bus B

Load @ B	μ of Line BD	DF @ B	LMP @ B	DF @ C	LMP @ C
300	50.98634	1.011301	24.30337	1.013040	27.32212
303	50.98628	1.011411	24.30721	1.013120	27.32494
306	50.98622	1.011520	24.31105	1.013200	27.32776
309	50.98617	1.011630	24.31490	1.013280	27.33058
312	50.98611	1.011739	24.31874	1.013361	27.33340
315	50.98605	1.011848	24.32258	1.013441	27.33621
318	50.98599	1.011958	24.32643	1.013521	27.33903
321	50.98593	1.012067	24.33027	1.013601	27.34185
324	50.98587	1.012177	24.33411	1.013682	27.34467
327	50.98581	1.012286	24.33796	1.013762	27.34749
330	50.98575	1.012396	24.34180	1.013842	27.35031

In addition to the results shown in Figs. 3.18-3.21 and Table 3.7, the energy component of the LMP, or λ, is \$35/MWh, constantly. However, this is a special case because the marginal unit happens to be the reference bus, so λ is constant. As previously stated, λ is usually not a constant in the DCOPF model with loss, i.e., $\frac{\Delta\lambda}{\Delta D_j} \neq 0$. The GSF of line DE to Bus B is -0.2176.

Eq. (3.40) can be validated with Figs. 3.18 to 3.21. Taking the LMP at Bus B as an example (i.e., *i=j=Bus B*), we have

Normalized $\frac{\Delta DF_i}{\Delta D_j} = \frac{0.001028}{30}$ (See Fig. 3.18)

Actual $\frac{\Delta DF_i}{\Delta D_j} = 1.01130 \times \frac{0.001082}{30} = 3.647 \times 10^{-5}$

Actual $\frac{\Delta DF_i}{\Delta D_j} \cdot \lambda = 3.647 \times 10^{-5} \times 35 = 1.277 \times 10^{-3} (\$/\text{MWh}^2)$.

We also have

$$\frac{\Delta \lambda}{\Delta D_j} \cdot DF_i = 0 \ (\$/\text{MWh}^2)$$

$$\frac{\Delta \mu_{DE}}{\Delta D_j} \cdot GSF_{DE-i} = \frac{(50.9857 - 50.9863)}{30} \times (-0.2176)$$

$=0.004 \times 10^{-3}$ ($/MWh^2) (See Fig. 3.19) .

Therefore, we have

$$\frac{\Delta DF_i}{\Delta D_j} \cdot \lambda + \frac{\Delta \lambda}{\Delta D_j} \cdot DF_i + \sum_{k=1}^{M} \frac{\Delta \mu_k}{\Delta D_j} \cdot GSF_{k-i}$$

$= (1.277+0+0.004) \times 10^{-3}$

$=1.281 \times 10^{-3}$ ($/MWh^2) .

From Fig. 3.20, we have

Normalized $\frac{\Delta LMP_i}{\Delta D_j} = \frac{0.001581}{30}$

Actual $\frac{\Delta LMP_i}{\Delta D_j} = 24.30 \times (\frac{0.001581}{30}) = 1.281 \times 10^{-3} \ (\$/\text{MWh}^2)$.

Hence, the LMP sensitivity $\frac{\Delta LMP_i}{\Delta D_j}$ is very close to the value computed with $\frac{\Delta DF_i}{\Delta D_j} \cdot \lambda + \frac{\Delta \lambda}{\Delta D_j} \cdot DF_i + \sum_{k=1}^{M} \frac{\Delta \mu_k}{\Delta D_j} \cdot GSF_{k-i}$. This validates Eq. (3.40) and also matches the data in Fig. 3.21, numerically. Similar validations can be made for other buses.

Although the second and the third parts in Eq. (3.40) for this case is either zero or very small compared with the first part in Eq. (3.40), this does not mean they can be generally ignored. In fact, when the load level is approximately 360-390MW, the third part will play a larger role in LMP sensitivity than the first two parts. The next section provides test results from the IEEE 30-bus system, in which all three parts of LMP sensitivity are compared numerically.

As shown in Figure 3.20, the LMP at the marginal unit buses (e.g. Bus E) is constant, and is equal to the bidding price of the local generator Brighton as this generator is always a marginal unit when the Bus B Load varies between 300MW and 330MW. So, the local load increase at Bus E will be solely picked up by the local marginal generator Brighton. Thus, the sensitivity of the LMP at Bus E is zero, as shown in Fig. 3.21. This is also the case for Bus D because the local generator Sundance is also a marginal unit.

For non-marginal-unit buses (A, B, or C in this study), the LMPs linearly increase, as the load increases. Since the loss is a quadratic function of the load, the generation is a quadratic function of the load as well. If there is no change of marginal units (i.e., due to the very small change of the load), the dispatch cost is quadratically related to the load. The LMP, defined as incremental cost over incremental load, should be a linear function of the load, as shown in Fig. 3.20. Therefore, the LMP sensitivity at a bus without any marginal unit should be a non-zero constant, as shown in Fig. 3.21.

3.3.3.3. LMP Sensitivity results from IEEE 30-bus system

The LMP sensitivity with the loss considered is also tested on the IEEE 30-bus system [17]. The network topology is shown in Figure 3.22. The system data is slightly modified for illustration purposes: 1) The bidding prices of the 6 generators are assumed as 10, 15, 30, 35,

40, and 45, respectively, all in $/MWh, and 2) the transmission limit of Line 6-8 is increased by 10%.

Figure 3.23 shows the LMP sensitivity with respect to the Load at Bus 8 between 27 and 35MWh in the system. Only the LMP sensitivities at a few buses are shown in the figure for better illustration. Again, LMP sensitivities have constant values at all these buses.

Further tests show there will be a step change of the LMP sensitivity because of a new binding constraint when the load reaches approximately 36 MW. The diagram beyond 36MW is not shown simply because it is difficult to scale into one figure.

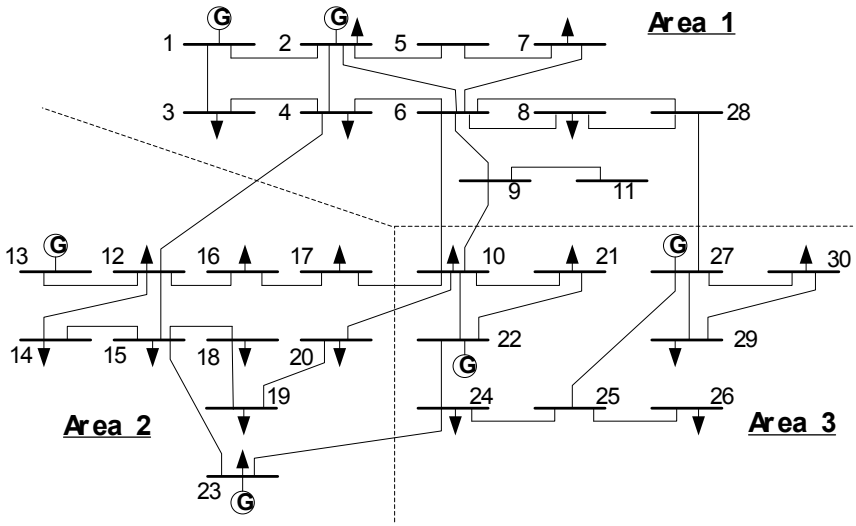


Figure 3.22. The Network Topology of the IEEE 30 Bus System

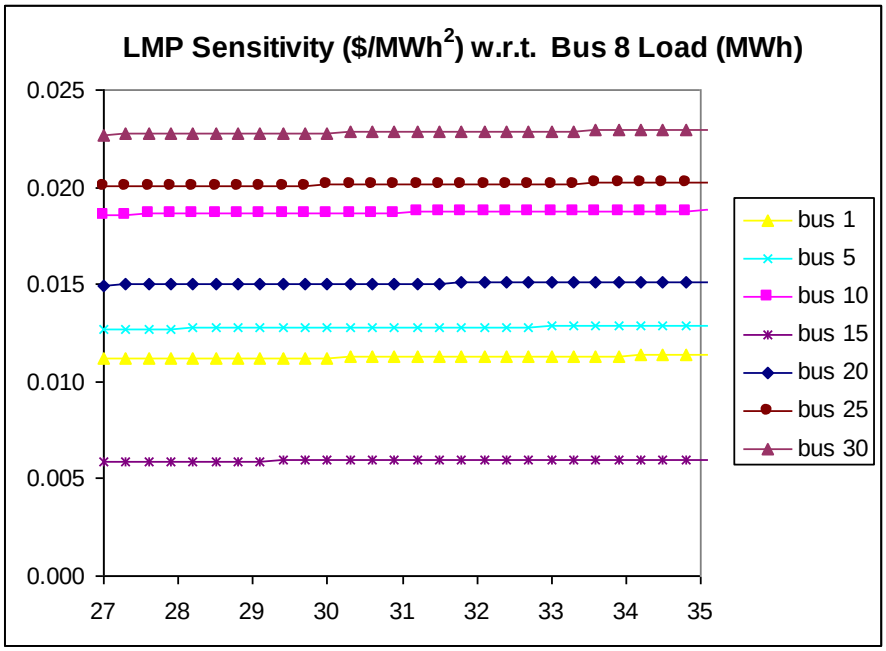


Figure 3.23. LMP Sensitivity at a few buses with respect to the Load at Bus 8 ranging from 27 MWh to 35 MWh (base case load=30MWh) in the IEEE 30-bus system

When the new (and only) binding transmission constraint appears, a non-zero μ occurs and its sensitivity is considerable. Equation (3.40) can be briefly verified using results at Bus 30 with respect to the Bus 8 Load, varied from 37.500 to 37.575 MW. This can be shown as follows (here i=30 and j=8)

$$\frac{\Delta DF_i}{\Delta D_j}\cdot\lambda = (-0.0007781/0.075) \times 26.96535 = -0.27976(\$/\text{MWh}^2)$$

$$\frac{\Delta\lambda}{\Delta D_j}\cdot DF_i = (-0.012238/0.075) \times 1.109955 = -0.18112\ (\$/\text{MWh}^2)$$

$$\frac{\Delta\mu_{Line6,8}}{\Delta D_j}\cdot GSF_{Line6,8-30} = (-0.\ 26436/0.075) \times (-0.12866)$$

$$= 0.45350(\$/\text{MWh}^2)$$

$$\frac{\Delta DF_i}{\Delta D_j}\cdot\lambda+\frac{\Delta\lambda}{\Delta D_j}\cdot DF_i+\sum_{k=1}^{M}\frac{\Delta\mu_k}{\Delta D_j}\cdot GSF_{k-i} = -0.007383(\$/\text{MWh}^2)\ .$$

We also have

$$\frac{\Delta LMP_i}{\Delta D_j} = -0.00054264/0.075 = -0.007235\ (\$/\text{MWh}^2)\ .$$

Hence, the error of Eq. (3.40) is less than 2.0%. Figure 3.24 shows the LMP at Bus 30 with respect to the Load at Bus 8 between 37.5 and 39 MWh. It can be easily verified that the slope of the LMP curve is roughly -0.007 $/MWh^2. It remains this value since there is no new binding constraint when the Load at Bus 8 is increased from 37.5 to 39 MWh.

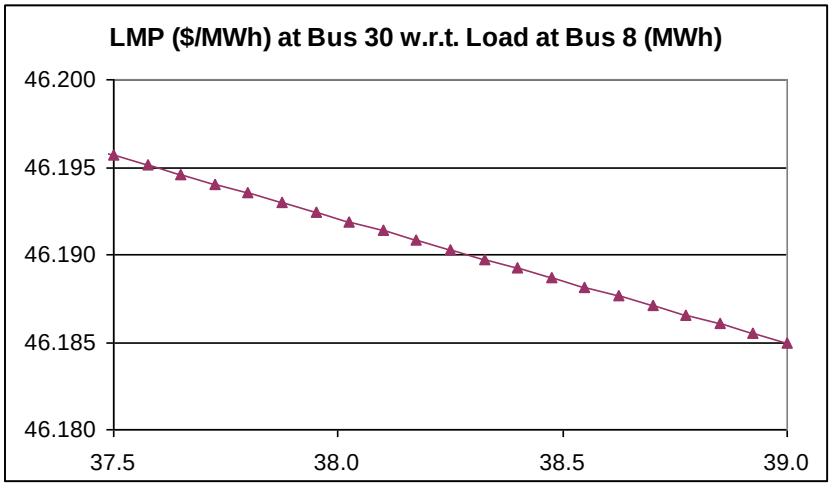


Figure 3.24. LMP at Bus 30 with respect to the Bus 8 Load from 37.5 to 39 MWh in the IEEE 30-bus system

Notes on Delivery Factor and LMP sensitivity

Note on Delivery Factor - Figure 3.18 and its caption show that the delivery factor may be greater than 1, which implies a negative penalty factor. Taking Bus C as an example, if there is a hypothetical injection increase at Bus C and it is absorbed by the reference Bus D, this will reduce the majority of the line flows such as Line EA, AB, BC, and DC, and therefore, reduce the system loss. So, the marginal loss factor is negative and the delivery factor is greater than 1. Similar observations can be obtained at other buses.

Note on LMP Sensitivity - Reference [41] presents a generalized, ACOPF-based model for LMP sensitivity with respect to the load and other variables. A matrix formulation needs to be solved to calculate the LMP sensitivity eventually, therefore, there is no direct, explicit formulation available from [41] about LMP sensitivity to load. This work does not intend to override the work in [41]; instead, this research work does present an explicit formulation, Eq. (3.40), about LMP sensitivity to load, based on the DCOPF with Delivery Factor, which is neither applicable nor necessary in the ACOPF model. Hence, with the concept of the Delivery Factor, the LMP sensitivity to load in the DC model is more straightforward and simple in such a way that it is more helpful to obtain a big picture about LMP sensitivity. This is very reasonable considering the simplifications from the AC model to the DC model. The observed results match the analytical equation (3.40) and clearly show that the LMP sensitivity is related to the loss component, linear to the sensitivity of delivery factors, and a numerical constant. Without the loss component, the LMP sensitivity is zero if the load is varied over a small range.

3.3.3.4. Sensitivity When there is a Change of Marginal Units

Figures 3.25 and 3.26 show the normalized delivery factor and the LMP sensitivity, respectively, when the load at Bus B is varied from 300 MWh to 390 MWh in the PJM 5-bus

system. Again, other loads remain unchanged for simplicity. These two figures show there is a turn of delivery factors and a sharp change of LMP sensitivity, when the Bus B Load increases from 346.50 to 347.25 MWh. This is due to a change of marginal units from Brighton and Sundance to Solitude and Sundance. This will change the LMP prices at each bus significantly. In addition, this will change the Delivery Factor (DF) growth pattern because the DF is related to the generation locations. Therefore, the delivery factor sensitivity has a sharp change as does the LMP sensitivity.

As shown in Fig. 3.25, the delivery factor at Bus C decreases after the critical load level of the step change of LMP sensitivity. At the critical load level and above, the new marginal unit Solitude will generate more power to supply its local load. So, the power flow through Lines DC and BC will be considerably reduced while the power flows through other lines remains unchanged. This will reduce the line losses and the fictitious demand at Bus C. Thus, the delivery factor decreases. Hence, the delivery factor sensitivity changes sharply as does the LMP sensitivity.

This representative case well illustrates that delivery factors may be affected by generation scheduling. Hence, this also shows the necessity to adopt the iterative DCOPF approach rather than using the DF from a pre-defined typical scenario.

The step change pattern implies the applicability of LMP sensitivities. When the present operating point is far from a change of marginal units, the LMP sensitivities (due to losses) can indicate how the LMP will change under load variations. On the other hand, when several of the current marginal units are near their generation limits or several transmission lines are congested due to a small load growth, the LMP sensitivities calculated in the present operating point are less reliable because a step change of the LMP, as well as LMP sensitivity, may occur, even with a small load growth.

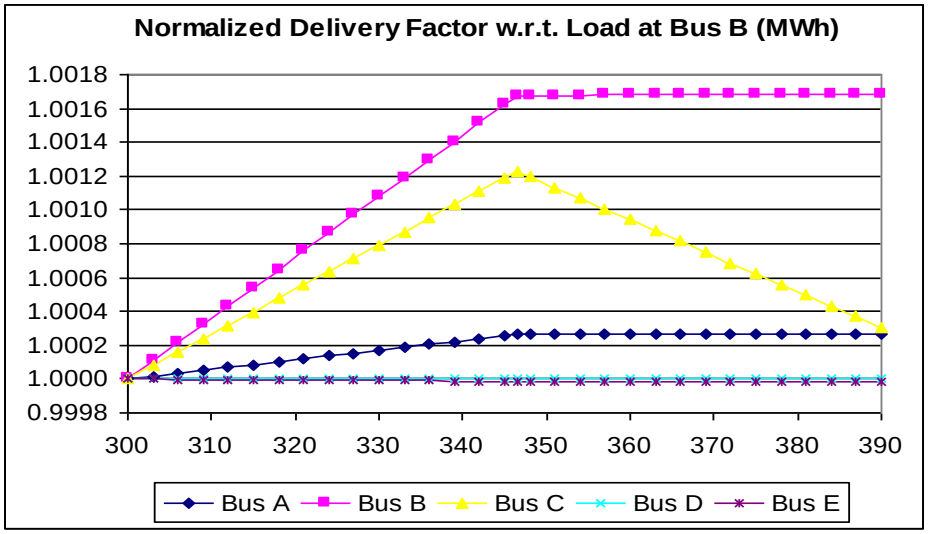


Figure 3.25. Normalized Delivery Factor at each bus with respect to the Load at Bus B ranging from 300 MWh to 390 MWh in the PJM 5-bus system

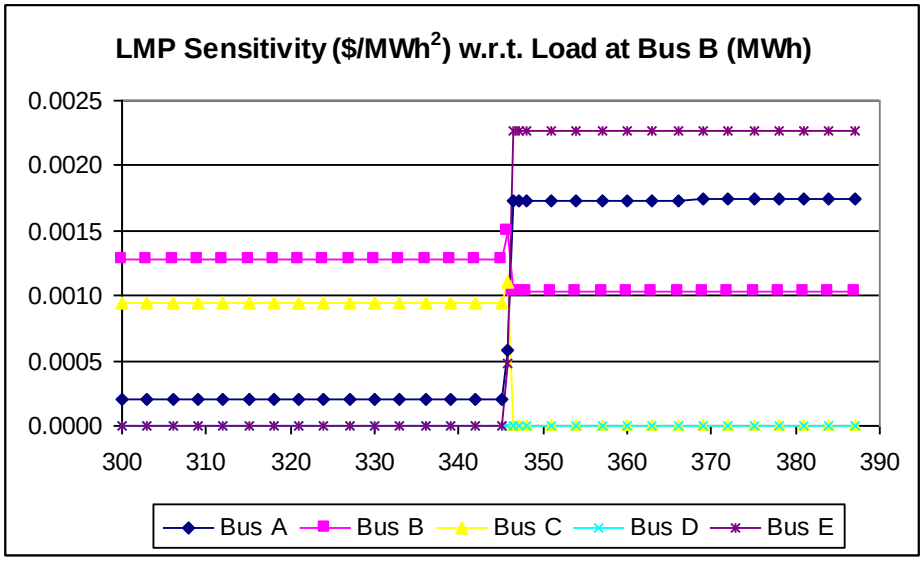


Figure 3.26. LMP Sensitivity with respect to the Load at Bus B ranging from 300 MWh to 390 MWh in the PJM 5-bus system

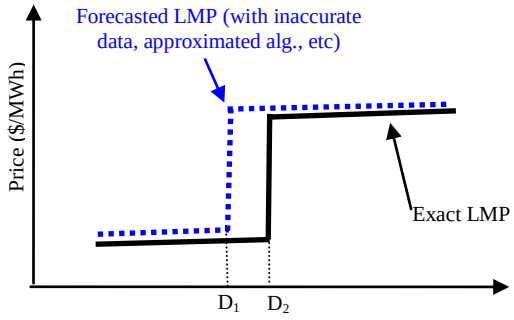


Figure 3.27. Forecasted LMP and Exact LMP

Note that this step change, or infinite sensitivity, is not an artifact of the DC model. In fact, even with the AC model, the step change still occurs due to the marginal characteristics in the LMP definition. If the load grows to a critical level so that a new congestion occurs, there will be a new marginal unit that leads to a step change in price. This has an important implication on the present LMP methodology: there is significant uncertainty or risk in LMP forecasting due to inaccurate data or an approximated LMP algorithm. This is illustrated in Fig. 3.27 in which the LMP error is relatively significant when the load is between D_1 and D_2. Hence, all approximated LMP algorithms cannot completely eliminate the relatively considerable errors in a range of load levels in which a step change of the LMP occurs. However, a better approximated algorithm should be able to narrow the range of LMP errors, as demonstrated by the proposed FND-based DCOPF as opposed to the lossless DCOPF.

3.4. Discussion and Conclusions

The proposed FND algorithm may be further simplified by executing only the first two iterations. Basically, the first iteration is essentially a lossless DCOPF run to provide an estimation of delivery factors, FND, and system loss. Then, another DCOPF is performed

based on the estimation. The reason for this simplification is that our research found that an initial estimation of delivery factors and losses have a bigger impact on the LMP and dispatch than the later iterations used to refine delivery factors, FND, and system loss. This can reduce the computational effort since it does not require the algorithm to run until convergence. Therefore, it fits a simulation or planning purpose well if the accuracy is reasonably acceptable. Our tests on the PJM 5-bus system and IEEE 30-bus system show that the two-iteration simplification of the FND algorithm produces results very close (less than 4% error in the Maximum Difference of nodal LMPs) to the fully converged FND algorithm. Nevertheless, this is a heuristic observation and needs further research to be credibly applied to much larger, real systems.

The proposed FND-based Iterative DCOPF shall be applicable to the Security (Contingency) Constrained Optimal Power Flow, i.e., SCOPF or CCOPF, because there is no mathematical difference between the SCOPF and OPF, despite a more computational complexity. In general, additional arrays of the Generation Shift Factors under contingency scenarios may be added to model contingency constraints. The security limit can be modeled similar to the line limits presented in (3.32).

This chapter first presents the loss and delivery factors based on the Generation Shift Factor (GSF). The reduction of system loss in the energy balance equality constraint is rigorously proved. Then, the challenge of a considerable nodal mismatch at the reference bus is presented. The mismatch issue is overcome with the proposed Fictitious Nodal Demand (FND) model, in which the total loss is distributed into each *individual* line and there is no nodal mismatch.

This chapter also presents a comparison of the LMP results from the lossless DCOPF, the FND-based DCOPF, and the ACOPF algorithms. The results indicate that a FND-based

Iterative DCOPF provides better results than the lossless DCOPF, and represents a better approximation of the ACOPF LMP.

In addition, this chapter presents a simple and explicit formulation of LMP sensitivity with respect to the load, based on the FND algorithm. Without the loss component, the LMP sensitivity is zero if the load is varied over a small range. The LMP sensitivity may be infinite (i.e., a step change in LMP) when the load grows to a critical level and will lead to a new marginal unit. This step-change nature presents uncertainty and risk in the LMP forecast, especially when considering the possible data inaccuracy or algorithm approximation. Therefore, future research could explore approaches for smoothing out the step changes using penalty or rebate functions on constraints, and evaluate whether such approaches would ease forecasting of prices while still preserving the correct economic signals.

4 Congestion and Price Prediction under Load Variation

4.1. Chapter Introduction

As previously mentioned the Locational Marginal Pricing (LMP) methodology has been a dominant approach in the energy market operation and planning in the identification of the nodal price and management of the transmission congestion. For system operators and planners, it is important to know the future price and possible new binding limit as the system load grows. This information can be used for congestion mitigation and load management in both the short-term and long-term. Meanwhile, for generation companies, it is also important in predicting the future price and possible congestion, as evidenced by the adoption of optimal power flow (OPF)-based market simulators incorporating full transmission models. These demands stimulate the research presented in this work: to devise algorithms to efficiently identify congestion and LMP as a function of the load levels.

Challenges arise because there is a step change of the LMP when the load grows to a certain level [42]. This is caused by the occurrence of a new binding limit, either a transmission line or a generator reaching its limit. Then, there will be a change of the marginal unit set and the sensitivity of marginal generation with respect to the load. Figure 4.1 shows a typical LMP versus load curve with a given growth pattern for a sample system slightly modified from the well-known PJM 5-bus system, as shown in Figure 3.2.

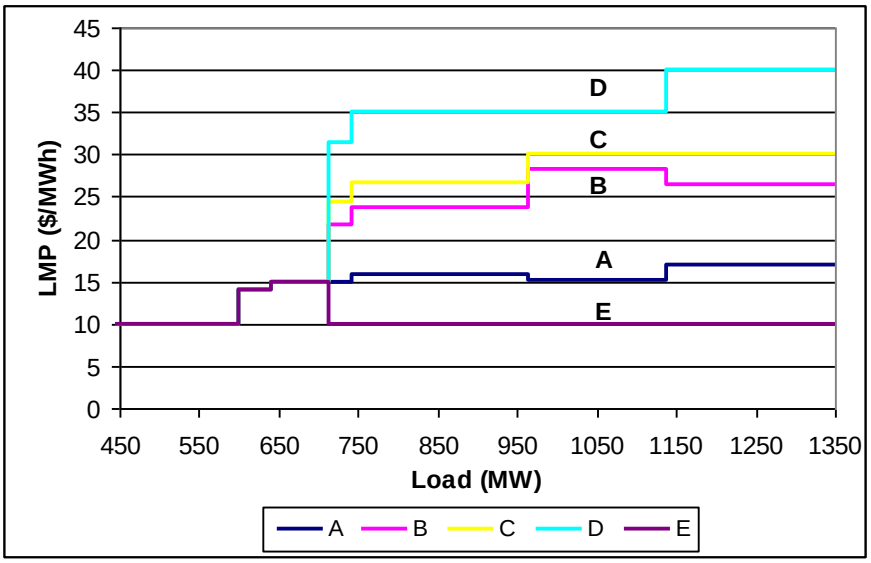


Figure 4.1. LMP at all buses with respect to different system loads

Certainly, the curve can be obtained if we repetitively run an optimization model and then perform the LMP calculation at many different load levels. This approach of repetitive optimization runs can be relatively time-consuming, especially for short-term applications. Even though it may be still fast enough in practice for a one-scenario application, such as a real-time dispatch, it will be not be fast enough if many different scenarios need to be run. For instance, a short-term market participant or system planner may want to run multiple scenarios with different load growth patterns and/or different transmission and generation maintenances. Then, multiple curves similar to that in Fig. 4.1 need to be obtained. This will make the run time of the repetitive optimization-run approach less competitive. A more efficient algorithm is highly desired.

The most important step is to *efficiently* identify the next critical load level (CLL), defined as a load level at which a step change occurs, and the corresponding new binding limit, either a congested transmission line or a marginal unit reaching its generation limit. In addition to

finding the next CLL, it is also desirable to solve the following questions related to the system status at that CLL:

- Which unit will be the next new marginal unit?
- What is the new generation sensitivity of each marginal unit? (Note: The generation sensitivity of the old units will change.)
- What is the new LMP at each bus?

Due to the fact that different OPF models have different mathematical characteristics, this study will be conducted for three major OPF models, namely, the lossless DCOPF, ACOPF, and the proposed FND-based DCOPF.

4.2. Simplex-like Method for Lossless DCOPF Framework

A conventional sensitivity analysis [39, 41] provides the sensitivity under a small perturbation, and no change of marginal units under the load variation is assumed. It does not address the issue when the load continuously grows beyond the next critical load level (CLL) where a new binding limit occurs. In contrast, this work will present a systematic approach, without a new optimization run, based on a simple matrix formulation to identify the new CLL, marginal units, congested lines, and nodal prices when the load growth leads to a new binding constraint and a step change of LMP and congestion. This is the primary mathematical significance of this work. The proposed approaches are very different from the previous works [20, 45], which solve the LMP at different hours (hence different load levels) using chronological optimization runs [20] or artificial intelligence [45]. This work starts from the present optimum and finds the solution at the next CLL by directly utilizing the unique features of the optimal dispatch model. Hence, it avoids repetitive optimization runs and should be more computationally efficient.

This section is organized as follows. First, the fundamental formulation of the proposed algorithm is presented to express marginal variables and the objective function in non-marginal variables. Then, the formulation under load variations is established. We next present the calculation of the new binding limit, critical load level, set of marginal generators, generation sensitivity of all marginal generators, and prices as the load reaches the next critical level. An easy-to-follow example is shown to illustrate the algorithm in matrix formulation based on the PJM 5-bus system. Last, the performance speedup results with the PJM 5-bus, the IEEE 30-bus, and the IEEE 118-bus systems are presented.

4.2.1. Fundamental Formulation of the Proposed Algorithm

In this section, first, slack variables are applied to all inequality transmission constraints to convert them to equality constraints. Then, a matrix formulation is presented to rewrite all constraints so that marginal variables are expressed with non-marginal variables. Note that in this work a single vector is typically denoted in bold font, while a vector composed of multiple vectors or a matrix is denoted in bold font with brackets.

Assume at the present load level, we have N_{MG} marginal units. Hence, we should have N_{MG}-1 congested lines, because the total number of marginal units is one more than the total number of congested lines [7, 42]. This can be written as

$$N_{MG} = M_{CL} + 1 = M - M_{UL} + 1 \tag{4.1}$$

where

N_{MG} = number of marginal units;

M_{CL} = number of congested lines;

M_{UL} = number of un-congested lines.

Hence, we have energy balance equality constraint as:

$$\sum_{j \in \mathcal{MG}} MG_j + \sum_{j \in \mathcal{NG}} NG_j = \sum_{i \in \mathcal{N}} D_i \tag{4.2}$$

where $\mathcal{MG}$, $\mathcal{NG}$, and $\mathcal{N}$ represent the marginal unit set, the non-marginal unit set, and the all bus set, respectively.

The transmission inequality constraints can be written as equality constraints by introducing a non-negative slack variable. If we use UL_k to represent the slack variables of un-congested lines and CL_k for congested lines, we have

$$\sum_{j \in \mathcal{MG}} GSF_{k-j} \times MG_j + \sum_{j \in \mathcal{NG}} GSF_{k-j} \times NG_j - \sum_{i \in \mathcal{N}} GSF_{k-i} \times D_i + CL_k = F_k^{\max}, \forall k \in \mathcal{CL} \tag{4.3}$$

$$\sum_{j \in \mathcal{MG}} GSF_{k-j} \times MG_j + \sum_{j \in \mathcal{NG}} GSF_{k-j} \times NG_j - \sum_{i \in \mathcal{N}} GSF_{k-i} \times D_i + UL_k = F_k^{\max}, \forall k \in \mathcal{UL} \tag{4.4}$$

where

$\mathcal{CL}$ is the set of congested lines;

$\mathcal{UL}$ is the set of un-congested lines.

Note: CL_k=0 for each congested (binding) transmission constraint.

Equations (4.2) to (4.4) can be re-written in matrix formulation as follows

$$\left[\begin{array}{c|c} \mathbf{1} & \mathbf{0} \\ \hline \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{array}\right] \times \begin{bmatrix} \mathbf{MG} \\ \mathbf{UL} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{p}_1 \\ \mathbf{p}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{1} \\ \mathbf{q}_1 \\ \mathbf{q}_2 \end{bmatrix} \times \mathbf{D} + \begin{bmatrix} \mathbf{-1} \\ \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix} \times \mathbf{NG} + \begin{bmatrix} \mathbf{0} \\ \mathbf{t}_1 \\ \mathbf{t}_2 \end{bmatrix} \times \mathbf{CL} \tag{4.5}$$

where

MG = N_{MG}×1 vector representing marginal generator output;

NG = N_{NG}×1 vector representing non-marginal generator output;

UL = M_{UL}×1 vector representing the slack variables of non-congested branches (lines);

D = $[D_1\ D_2\ \ldots\ D_N]^T$ = N×1 vector representing all loads with the assumption that each bus has a load for notational simplicity;

CL = M_{CL}×1 vector representing the slack variables of congested lines; and it is a zero vector for the base case;

1 = row vector of 1's (dimension is case dependent);

0 = row vector of 0's (dimension is case dependent);

$\mathbf{A_{11}} = \mathbf{GSF_{CL\text{-}MG}}$ = M_{CL}×N_{MG} matrix representing the GSF of M_{CL} (=N_{MG}-1) congested lines w.r.t. marginal unit buses;

$\mathbf{A_{12}} = \mathbf{0}$ = M_{CL} × M_{UL} zero matrix;

$\mathbf{A_{21}} = \mathbf{GSF_{UL\text{-}MG}}$ = M_{UL}× N_{MG} matrix representing the GSF of un-congested lines w.r.t. marginal unit buses;

$\mathbf{A_{22}} = \mathbf{I}$ = M_{UL} × M_{UL} identity matrix;

$\mathbf{p}_1 = \mathbf{F}_{\mathbf{CL}}^{\mathbf{max}}$ = M_{CL}×1 vector representing the line flow limit of congested lines;

$\mathbf{p}_2 = \mathbf{F}_{\mathbf{UL}}^{\mathbf{max}}$ = M_{UL}×1 vector representing the line flow limit of un-congested lines;

$\mathbf{L_1} = \mathbf{GSF_{CL-N}}$ = M_{CL}×N matrix representing the GSF of congested lines w.r.t. all buses (load buses);

$\mathbf{L_2} = \mathbf{GSF_{UL-N}}$ = M_{UL}×N matrix representing the GSF of non-congested lines w.r.t. all buses (load buses);

$\mathbf{L_1} = -\mathbf{GSF_{CL\text{-}NG}}$ = M_{CL}×N_{NG} matrix, representing the negative of GSF of congested lines w.r.t. non-marginal unit buses;

$\mathbf{L_2} = -\mathbf{GSF_{UL\text{-}NG}}$ = M_{UL}×N_{NG} matrix of the GSF of congested lines w.r.t. non-marginal unit buses;

$\mathbf{A}_{12} = -\mathbf{I}$ = $M_{CL} \times M_{CL}$ negative unity matrix;

$\mathbf{A}_{22} = \mathbf{0}$ = $M_{UL} \times M_{CL}$ zero matrix.

Equation (4.5) can be rewritten as

$$\mathbf{A} \times \begin{bmatrix} \mathbf{MG} \\ \mathbf{UL} \end{bmatrix} = \mathbf{p} + \mathbf{A}_{11} \times \mathbf{D} + \mathbf{A}_{12} \times \mathbf{NG} + \mathbf{A}_{12} \times \mathbf{CL} . \qquad (4.6)$$

It should be mentioned that **CL**, the vector of slack variables for present congested lines, is a vector of 0 at the present load level. However, **CL** needs to remain as a set of variables, rather than constants of 0, in the formulation. The reason is that **CL** may change as the load grows beyond the next critical load level (CLL). In other words, a congested line may become un-congested as the load varies. Hence, **CL** should be kept as a vector of variables. This is a critical step to the following analysis.

Equation (4.6) can be further simplified by first finding the inverse of the matrix **[A]**, i.e., $\begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$. This is given by:

$$\begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}'_{11} & \mathbf{0} \\ \mathbf{A}_{21} & \mathbf{I} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}'^{-1}_{11} & \mathbf{0} \\ -\mathbf{A}_{21} \times \mathbf{A}'^{-1}_{11} & \mathbf{I} \end{bmatrix} \qquad (4.7)$$

where $\mathbf{A}'_{11} = \begin{bmatrix} \mathbf{1} \\ \mathbf{A}_{11} \end{bmatrix}$, which is an $N_{MG} \times N_{MG}$ square matrix.

Therefore, only the inverse of $\mathbf{A}'_{11}$ requires computation. It should be noted that $\mathbf{A}'_{11}$ is full rank and the inverse should exist. Since only a few marginal units exist, the size of $\mathbf{A}'_{11}$ is usually very small. This will not provide a computational burden to the algorithm. With the inversion of the **[A]** matrix, we could solve equation (4.7). Hence, we have

$$\begin{bmatrix} \mathbf{MG} \\ \mathbf{UL} \end{bmatrix} = \mathbf{P} + \begin{bmatrix} \mathbf{Q}_{MG} \\ \mathbf{Q}_{UL} \end{bmatrix} \times \mathbf{D} + \begin{bmatrix} \mathbf{R}_{MG} \\ \mathbf{R}_{UL} \end{bmatrix} \times \mathbf{NG} + \begin{bmatrix} \mathbf{T}_{MG} \\ \mathbf{T}_{UL} \end{bmatrix} \times \mathbf{CL} \tag{4.8}$$

where

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{MG} \\ \mathbf{P}_{UL} \end{bmatrix} = \mathbf{A}^{-1}\mathbf{p}; \quad \begin{bmatrix} \mathbf{Q}_{MG} \\ \mathbf{Q}_{UL} \end{bmatrix} = \mathbf{A}^{-1}\mathbf{b}; \quad \begin{bmatrix} \mathbf{R}_{MG} \\ \mathbf{R}_{UL} \end{bmatrix} = \mathbf{A}^{-1}\mathbf{F};$$

$$\begin{bmatrix} \mathbf{T}_{MG} \\ \mathbf{T}_{UL} \end{bmatrix} = \mathbf{A}^{-1}\mathbf{I}.$$

As the above equation shows, Eq. (4.8) consists of three parts:

- 1 equation representing the energy balance equation;
- M_{CL} (=N_{MG}-1) equations representing the congested lines with slack variables **CL** being zeros;
- M_{UL} equations representing the un-congested lines with non-zero slack variables **UL**.

As we can see, the formulation is written in a matrix form similar to the dictionary format of a simplex method to solve linear programming problems. The reason in doing this is to re-write the non-zero variables (or basic variables) like **MG** and **UL** at the left-hand side, then any small change of load can be expressed as a corresponding change of **MG** or **UL**. Hence, the objective function can be written without **MG** or **UL**, as shown below.

Using (4.8), we can rewrite the original objective function

$$z = \sum_{j \in \mathcal{MG}} C_j \times MG_j + \sum_{j \in \mathcal{NG}} C_j \times NG_j \tag{4.9}$$

as

$$\begin{aligned} z &= \mathbf{C}_{MG}^{T} \times \mathbf{MG} + \mathbf{C}_{NG}^{T} \times \mathbf{NG} \\ &= \mathbf{C}_{MG}^{T} \times (\mathbf{P}_{MG} + \begin{bmatrix} \mathbf{Q}_{MG} \end{bmatrix} \times \mathbf{D} + \\ &\quad \begin{bmatrix} \mathbf{R}_{MG} \end{bmatrix} \times \mathbf{NG} + \begin{bmatrix} \mathbf{T}_{MG} \end{bmatrix} \times \mathbf{CL}) + \mathbf{C}_{NG}^{T} \times \mathbf{NG} \\ &= \mathbf{C}_{MG}^{T} \times \mathbf{P}_{MG} + (\mathbf{C}_{MG}^{T} \times \begin{bmatrix} \mathbf{Q}_{MG} \end{bmatrix}) \times \mathbf{D} + \\ &\quad (\mathbf{C}_{MG}^{T} \times \begin{bmatrix} \mathbf{R}_{MG} \end{bmatrix} + \mathbf{C}_{NG}^{T}) \times \mathbf{NG} + \mathbf{C}_{MG}^{T} \times \begin{bmatrix} \mathbf{T}_{MG} \end{bmatrix} \times \mathbf{CL} \end{aligned} \tag{4.10}$$

where

$\mathbf{C_{MG}}$ = column vector of marginal generator costs;

$\mathbf{C_{NG}}$ = column vector of non-marginal generator costs;

$\mathbf{[R_{MG}]}$ = the first N_{MG} rows of the **[R]** matrix;

$\mathbf{[T_{MG}]}$ = the first N_{MG} rows of the **[T]** matrix.

4.2.2. Load Variation

If there is a change of system load, with the assumption of linear participating factors, we can rewrite the load as

$$D_i = D_i^{(0)} + \Delta D_i = D_i^{(0)} + f_i \times \Delta D_\Sigma$$

$$D_\Sigma = \sum_i D_i = \sum_i \left(D_i^{(0)} + f_i \times \Delta D_\Sigma \right) = D_\Sigma^{(0)} + \Delta D_\Sigma$$

$$\mathbf{\Delta D} = \mathbf{f} \times \Delta D_\Sigma$$

where

$f_i = \dfrac{\partial D_i}{\partial D_\Sigma}$ (load growth participating factor), and $\sum_{i \in \mathcal{N}} f_i = 1$;

$\mathbf{f} = \begin{bmatrix} f_1 & f_2 & \ldots & f_N \end{bmatrix}^T$, an N×1 column vector;

$\Delta\mathbf{D}$ is a column vector and ΔD_Σ is a scalar.

With the above load variation model, the change of each bus load follows a linear participating factor with respect to the system load change. This model is reasonable because each bus load can be modeled to have its own variation pattern so that different load characteristics like industrial loads, commercial loads, and residential loads can be modeled accordingly. It is also flexible because the variation at each bus load is independent on the initial load. This model is particularly useful for short-term planning.

Considering the system load will be varied by ΔD_{Σ}, we have

$$\begin{bmatrix} \Delta \mathbf{MG} \\ \Delta \mathbf{UL} \end{bmatrix} = \begin{bmatrix} \mathbf{Q} \end{bmatrix} \times \Delta \mathbf{D} + \begin{bmatrix} \mathbf{R} \end{bmatrix} \times \Delta \mathbf{NG} + \begin{bmatrix} \mathbf{T} \end{bmatrix} \times \Delta \mathbf{CL}. \tag{4.11}$$

With the consideration of the linear participating factors of the load variation pattern, i.e., $\Delta \mathbf{D} = \mathbf{f} \times \Delta D_{\Sigma}$, we can re-write the above equations to

$$\begin{aligned} \begin{bmatrix} \Delta \mathbf{MG} \\ \Delta \mathbf{UL} \end{bmatrix} &= \begin{bmatrix} \mathbf{Q} \end{bmatrix} \mathbf{f} \times \Delta D_{\Sigma} + \begin{bmatrix} \mathbf{R} \end{bmatrix} \times \Delta \mathbf{NG} + \begin{bmatrix} \mathbf{T} \end{bmatrix} \times \Delta \mathbf{CL} \\ &= \mathbf{Q}' \times \Delta D_{\Sigma} + \begin{bmatrix} \mathbf{R} \end{bmatrix} \times \Delta \mathbf{NG} + \begin{bmatrix} \mathbf{T} \end{bmatrix} \times \Delta \mathbf{CL} \end{aligned} \tag{4.12}$$

where

$\mathbf{Q}' = \begin{bmatrix} \mathbf{Q} \end{bmatrix} \mathbf{f}$, a ($N_{MG}$+$M_{UL}$)×1 column vector.

We can further decouple the above equation into

$$\Delta \mathbf{MG} = \mathbf{Q}'_{\mathbf{MG}} \times \Delta \mathrm{D}_{\Sigma} + \begin{bmatrix} \mathbf{R}_{\mathbf{MG}} \end{bmatrix} \times \Delta \mathbf{NG} + \begin{bmatrix} \mathbf{T}_{\mathbf{MG}} \end{bmatrix} \times \Delta \mathbf{CL} \tag{4.13}$$

$$\Delta \mathbf{UL} = \mathbf{Q}'_{\mathbf{UL}} \times \Delta \mathrm{D}_{\Sigma} + \begin{bmatrix} \mathbf{R}_{\mathbf{UL}} \end{bmatrix} \times \Delta \mathbf{NG} + \begin{bmatrix} \mathbf{T}_{\mathbf{UL}} \end{bmatrix} \times \Delta \mathbf{CL}. \tag{4.14}$$

With the above two equations, we can immediately obtain the sensitivity of **MG** and **UL** with respect to load

$$\frac{\partial \mathbf{MG}}{\partial D_{\Sigma}} = \mathbf{Q}'_{\mathbf{MG}} \tag{4.15}$$

$$\frac{\partial \mathbf{UL}}{\partial D_{\Sigma}} = \mathbf{Q}'_{\mathbf{UL}}. \tag{4.16}$$

We can also write the change of the objective function *z* as follows

$$\begin{aligned} \Delta \mathrm{z} &= (\mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times \begin{bmatrix} \mathbf{Q}_{\mathbf{MG}} \end{bmatrix}) \times \mathbf{f} \times \Delta D_{\Sigma} + (\mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times \begin{bmatrix} \mathbf{R}_{\mathbf{MG}} \end{bmatrix} + \mathbf{C}_{\mathbf{NG}}^{\mathbf{T}}) \times \Delta \mathbf{NG} + \\ &\quad \mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times \begin{bmatrix} \mathbf{T}_{\mathbf{MG}} \end{bmatrix} \times \Delta \mathbf{CL} \\ &= (\mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times \mathbf{Q}'_{\mathbf{MG}}) \times \Delta D_{\Sigma} + (\mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times \begin{bmatrix} \mathbf{R}_{\mathbf{MG}} \end{bmatrix} + \mathbf{C}_{\mathbf{NG}}^{\mathbf{T}}) \times \Delta \mathbf{NG} + \\ &\quad \mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times \begin{bmatrix} \mathbf{T}_{\mathbf{MG}} \end{bmatrix} \times \Delta \mathbf{CL} \end{aligned}. \tag{4.17}$$

Therefore, we know the sensitivity of the objective function w.r.t. **NG** and **CL** are as follows

$$\frac{\partial z}{\partial NG_j} = \left(\mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times \left[\mathbf{R}_{\mathbf{MG}}\right] + \mathbf{C}_{\mathbf{NG}}^{\mathbf{T}}\right)_j \tag{4.18}$$

$$\frac{\partial z}{\partial CL_k} = \left(\mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times \left[\mathbf{F}_{\mathbf{MG}}\right]\right)_k . \tag{4.19}$$

4.2.3. Identification of New Binding Limit, New Marginal Unit and LMP

The above formulations can be applied to perform three important tasks: 1) identifying the next new binding limit, either the generation limit or transmission limit, and the next CLL; 2) identifying the next unbinding limit such as a new marginal unit; 3) finding the new generation sensitivity of all marginal units and the new LMP. These steps provide important information such as generation dispatch sensitivity or transmission congestion prediction, which is aligned with the main goal of this work, i.e., to find congestion and LMP versus the load, starting from any initial load level, say $D_\Sigma^{(0)}$, to any load level without running the OPF repetitively.

The first section identifies the next CLL, $D_\Sigma^{(1)}$, as shown in Fig. 4.2 below, where a new binding limit will appear. The algorithm utilizes the feature that marginal units and congested lines will remain the same when the load variation ΔD_Σ does not push the load level beyond $D_\Sigma^{(1)}$. Then, the next section will identify the change of binding limits and marginal unit if the load grows to the immediate right side of $D_\Sigma^{(1)}$. The algorithm is based on finding the least incremental cost among all possible changes of the non-marginal units or slack variables of the congested lines. The algorithm has a simple final formulation and is very efficient. Essentially, these steps provide new dispatches and congested lines at the new CLL, $D_\Sigma^{(1)}$. The LMP can then be easily calculated at $D_\Sigma^{(1)}$.

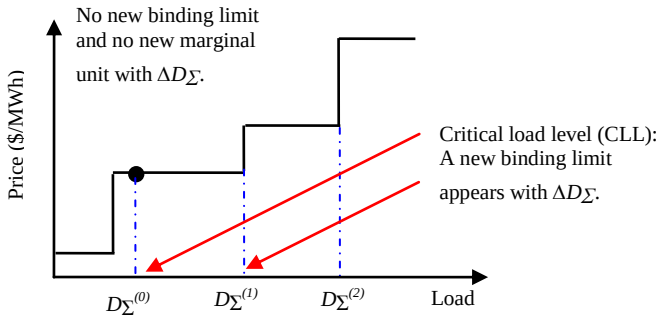


Figure 4.2. LMP versus Load Curves

Similarly, starting from $D_{\Sigma}^{(1)}$, we can repeat the above process to find the congestion and LMP at the "next-next" CLL, $D_{\Sigma}^{(2)}$.

4.2.3.1. Identification of new binding limit and new critical load level

When the load grows and this growth does not lead to any change of the marginal units, a non-marginal unit output should remain at its minimum or maximum, and the slack variable of a congested line should remain at zero. In other words, **ΔNG=ΔCL=0.** Meanwhile, marginal generators and unbinding transmission lines should change and may approach their respective limits gradually. The one reaching its limit first will be the next binding limit. To analyze this, we have

$$\begin{bmatrix} \Delta\mathbf{MG} \\ \Delta\mathbf{UL} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}'_{\mathbf{MG}} \\ \mathbf{Q}'_{\mathbf{UL}} \end{bmatrix} \times \Delta D_{\Sigma} \ . \tag{4.20}$$

Since all generation output and line flows at the present load level are provided from the initial OPF, it is not difficult to obtain the present values of the slack variables. In fact, many optimization solvers will give these values as output.

For un-congested transmission lines, the slack variable is **UL**. Since the sensitivity of **UL,** with respect to load change $\Delta D_{\Sigma,}$ is given by $\frac{\partial \mathbf{UL}}{\partial D_{\Sigma}} = \mathbf{Q}'_{\mathbf{UL}}$, we can obtain the allowed load

growth before a line, say, the k^{th} line, reaches its limit (i.e., UL_k reaching zero). This is given by

$$\Delta D_k^{allowed} = \frac{UL_k}{Q'_{UL_k}} \quad . \tag{4.21}$$

For generators, the slack variable is not given explicitly in the previous formulation. However, it can be viewed as

$$\mathbf{MG} + \mathbf{s_{MG}} = \mathbf{MG^{max}} \tag{4.22}$$

$$s_{MG_i} = MG_i^{\max} - MG_i \tag{4.23}$$

where s_{MG_i} is the slack variable of the i^{th} marginal generator.

Hence, the allowed load growth corresponding to the i^{th} marginal generator as the load grows can be given as

$$\Delta D_i^{allowed} = \frac{s_{MG_i}}{Q'_{MG_i}} = \frac{MG_i^{\max} - MG_i}{Q'_{MG_i}} \quad . \tag{4.24}$$

Then, the minimum allowed load growth can be obtained by finding the minimum value among all $\Delta D^{allowed}$ given by (4.21) and (4.24). Hence, the new critical load is equal to $D_{\Sigma}^{(0)} + \Delta D_{\min}^{allowed}$.

4.2.3.2. Identification of new unbinding constraint

When the load increases or decreases, a new binding constraint will occur at the CLL, together with the appearance of an unbinding constraint, which could be from the generation or transmission. From the previous subsection, it is known that the new binding constraint can be transmission or generation. These two scenarios will be discussed below.

1) Assume the l^{th} marginal unit becomes non-marginal (binding).

Since the l^{th} marginal unit is binding, it cannot grow as the load increase beyond $D_{\Sigma}^{(1)}$ in Fig. 4.2. Therefore, the change of the load ΔD_{Σ}, must be offset by either a previously non-marginal (binding) unit output by ΔNG_j or a previously binding slack variable by ΔCL_k. From (4.12), we have

$$0 = \mathbf{Q}'_{\mathbf{MG}_l} \times \Delta D_{\Sigma} + \mathbf{R}_{\mathbf{MG}_l} \times \mathbf{\Delta NG} + \mathbf{T}_{\mathbf{MG}_l} \times \mathbf{\Delta CL} \tag{4.25}$$

where

$\mathbf{R}_{\mathbf{MG}_l}$ and $\mathbf{T}_{\mathbf{MG}_l}$ are the l^{th} row vectors of $[\mathbf{R}_{\mathbf{MG}}]$ and $[\mathbf{T}_{\mathbf{MG}}]$, respectively.

It is very important to note that when the load is slightly more than the new CLL, there should be one and only one non-zero variables among all ΔNG_j and ΔCL_k. This is determined by the characteristics of linear programming, because the solution shall move from one vertex to an adjacent vertex (even though the vertices or boundaries themselves of the polytope should also change because of the change of D_{Σ}). Then, the determination of which ΔNG_j or ΔCL_k should be chosen as the next non-zero variable is based on the change of the objective function.

If the j^{th} non-marginal unit will become marginal, then we know from (4.25) that

$$\frac{\partial NG_j}{\partial D_{\Sigma}} = -\frac{Q'_{MG_l}}{R_{MG_{lj}}}, \; j \in \mathcal{NG} \tag{4.26}$$

It should be noted that the above sensitivity must give NG_j a possible change that will not violate its limit. For instance, for the case of a load increase, if NG_j is already at its maximum, the above sensitivity should be considered only if it is negative. And, if NG_j is at its minimum, the above sensitivity should be considered only if it is positive.

If the k^{th} congested line will become un-congested, then we know from (4.25) that

$$\frac{\partial CL_k}{\partial D_\Sigma} = -\frac{Q'_{MG_l}}{T_{MG_{lk}}}, k \in \mathcal{L} \quad . \tag{4.27}$$

Again, the above sensitivity is considered only if it does not push CL_k to negative values. At the CLL, CL_k should be zero.

Then, taking Eq. (4.18-27), we can easily calculate the expected incremental cost vector $\begin{bmatrix} \frac{\partial z}{\partial NG_j} \times \frac{\partial NG_j}{\partial D_\Sigma}, j \in \mathcal{NG} \\ \frac{\partial z}{\partial CL_k} \times \frac{\partial CL_k}{\partial D_\Sigma}, k \in \mathcal{L} \end{bmatrix}$ that can be expanded as:

$$\begin{bmatrix} \mathbf{C}^{\mathbf{T}}_{\mathbf{MG}} \times \left[\mathbf{R}_{\mathbf{MG}} \right]_j + \mathbf{C}^{\mathbf{T}}_{\mathbf{NG}} \times \left(-\frac{Q'_{MG_l}}{R_{MG_{lj}}} \right), j \in \mathcal{NG} \\ \mathbf{C}^{\mathbf{T}}_{\mathbf{MG}} \times \left[\mathbf{T}_{\mathbf{MG}} \right]_k \times \left(-\frac{Q'_{MG_l}}{T_{MG_{lk}}} \right), k \in \mathcal{L} \end{bmatrix} . \tag{4.28}$$

Finally, we can choose the smallest positive one in the load growth case (or largest negative one in the load drop case), and the corresponding j (or k) will be the new marginal unit (or new un-congested line)

2) Assume the r^{th} non-congested line becomes congested (binding).

Similar to (4.25), we have

$$0 = Q'_{UL_r} \times \Delta D_\Sigma + \mathbf{R}_{\mathbf{UL}_r} \times \mathbf{\Delta NG} + \mathbf{T}_{\mathbf{UL}_r} \times \mathbf{\Delta CL} \quad . \tag{4.29}$$

If the j^{th} non-marginal unit will become marginal, then we know from (4.29) that

$$\frac{\partial NG_j}{\partial D_\Sigma} = -\frac{Q'_{UL_r}}{R_{UL_{rj}}}, j \in \mathcal{NG} \quad . \tag{4.30}$$

If the k^{th} congested line will become un-congested, then we know from (4.29) that

$$\frac{\partial CL_k}{\partial D_\Sigma} = -\frac{Q'_{UL_r}}{T_{UL_{rk}}}, k \in \mathcal{L} \quad . \tag{4.31}$$

Similar to the discussions below (4.26) and (4.27), the sensitivity in (4.30) and (4.31) should be considered if and only if it presents a move away from the present binding limit.

Next, we can calculate the incremental cost vector

$$\begin{bmatrix} \left(\mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times \left[\mathbf{R}_{\mathbf{MG}}\right]_{\cdot j} + \mathbf{C}_{\mathbf{NG}}^{\mathbf{T}}\right)_j \times \left(-\dfrac{Q'_{UL_r}}{R_{UL_{rj}}}\right), j \in \{NG\} \\ \left(\mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times \left[\mathbf{T}_{\mathbf{MG}}\right]_{\cdot k}\right) \times \left(-\dfrac{Q'_{UL_r}}{T_{UL_{rk}}}\right), k \in \{CL\} \end{bmatrix}. \tag{4.32}$$

Similarly, we should choose the smallest positive one in the load growth case (or, largest negative one in the load drop case), and the corresponding j (or k) will be the new marginal unit (or new un-congested line).

Note on the new sensitivity of marginal units when load is beyond the next CLL, $D_{\Sigma}^{(1)}$:

It should be mentioned that the sensitivity of all existing marginal units will also change after the introduction of a new marginal unit at $D_{\Sigma}^{(1)}$. This can be quantitatively calculated as:

$$\left[\frac{\partial \mathbf{MG}}{\partial D_{\Sigma}}\right] = \mathbf{Q}'_{\mathbf{MG}} + \left[\mathbf{R}_{\mathbf{MG}}\right] \times \left[\frac{\partial \mathbf{NG}}{\partial D_{\Sigma}}\right] + \left[\mathbf{T}_{\mathbf{MG}}\right] \times \left[\frac{\partial \mathbf{CL}}{\partial D_{\Sigma}}\right]. \tag{4.33}$$

If there is no new marginal unit at CLL (such as the load decreases to have a congested line become unbinding), then we have $\left[\dfrac{\partial \mathbf{NG}}{\partial D_{\Sigma}}\right] = [\mathbf{0}]$ while one and only one variable in $\left[\dfrac{\partial \mathbf{CL}}{\partial D_{\Sigma}}\right]$ is not zero. Similarly, if there is a new marginal unit, then we have $\left[\dfrac{\partial \mathbf{CL}}{\partial D_{\Sigma}}\right] = [\mathbf{0}]$ while one and only one variable in $\left[\dfrac{\partial \mathbf{NG}}{\partial D_{\Sigma}}\right]$ is not zero.

A more straightforward approach is to re-formulate (4.5) using the new **MG**, **NG**, **UL,** and **CL** vectors. The change to these vectors should be very little, because only one marginal variable in either **MG** or **UL** will be switched into **NG** or **CL**. In addition, only one non-marginal variable in either **NG** or **CL** will be switched into **MG** or **UL**. Then, we can apply (4.12) to obtain new generation sensitivity. By doing so, we can repeat the previous process and eventually identify the "next-next" CLL ($D_{\Sigma}^{(2)}$), the "next-next" binding limit, etc.

4.2.3.3. LMP at the new critical load level

The above process can identify the new congestion, CLL, and marginal unit as the load grows, but has not addressed the price calculation. A similar approach can be taken since generation sensitivity is the key to calculate LMP. However, there is a little difference between the previous steps and this step. In the previous step, the load variation is a "global" scope variation where all load buses are assumed to vary together following some pattern. However, the LMP is calculated as the change of cost to supply a "local" scope change of the load at a single bus, after the generation dispatch has been addressed for the "global" change of the load. Nevertheless, the LMP calculation can be performed with essentially the same approach, in particular, Eq. (4.12). The only difference is that we use a different participating factor, expressed as **f**=[0 0 …1…0 0], since the LMP is locational dependent.

When the load is beyond $D_{\Sigma}^{(1)}$, we can first formulate the new **MG**, **NG**, **UL,** and **CL** vectors due to the change of marginal units and so on. Then, we can use (4.12) to calculate the marginal unit sensitivity with respect to a single bus load change using the "local" **f**=[0 0 …1…0 0]$^{\mathrm{T}}$. Therefore, LMP at a particular bus can be easily calculated as

$$LMP_i = \frac{\partial z}{\partial D_i} = \mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times \frac{\partial \mathbf{MG}}{\partial D_i} + \mathbf{C}_{\mathbf{NG}}^{\mathbf{T}} \times \frac{\partial \mathbf{NG}}{\partial D_i} = \mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times \frac{\partial \mathbf{MG}}{\partial D_i} \tag{4.34}$$

where

∂D_i = load change at a single bus i.

4.2.4. Case Study with the PJM 5-bus System

In this section, the PJM 5-Bus system with a slight modification will be employed to illustrate steps to identify the new binding limit, unbinding limit, generation sensitivity, and LMP as the load grows. The modifications to the original PJM 5-bus system [11] are as follows:

- The output limit of the Alta unit is reduced from 110 MW to 40 MW, while the output limit of the Park City unit is increased from 100 MW to 170 MW;
- The cost of Sundance unit at Bus D is changed from $30/MWh to $35/MWh to differentiate its cost from the Solitude unit;
- Line AB is assumed to have 400MW limit.

These changes are made so that there will be reasonably more binding limits within the investigated range of loading levels. Two binding limits will not occur at very close loading levels. Hence, a better illustration will be achieved when the price curves versus the loading levels are drawn.

We assume that the system load change is distributed to each nodal load proportional to its base case load for simplicity. Therefore, the load change is equally distributed at Buses B, C, and D since each has a 300 MW load in the base case. Fig. 4.3 shows the configuration of the system, Table 4.1 shows the line reactances and flow limits, and Table 4.2 shows Generation Shift Factors of Lines AB and ED with respect to all buses.

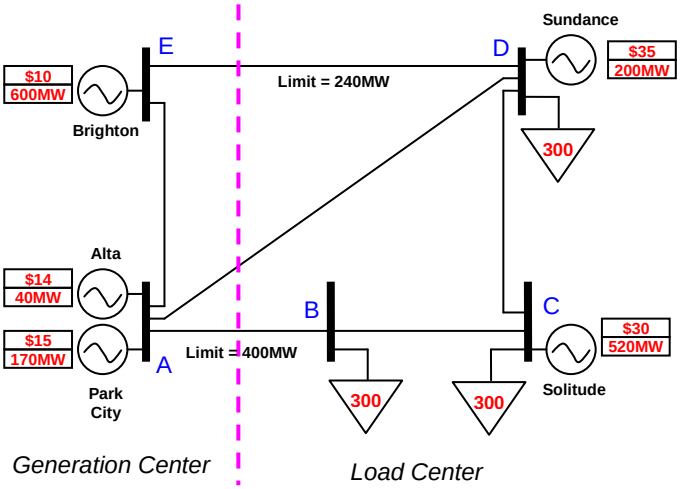


Figure 4.3. The Base Case Modified from the PJM Five-Bus Example

Table 4.1. Line impedance and flow limits

Line	AB	AD	AE	BC	CD	DE
X (%)	2.81	3.04	0.64	1.08	2.97	2.97
Limit (MW)	400	999	999	999	999	240

Table 4.2. GSF of Line AB and ED

	A	B	C	D	E
Line AB	0.1939	-0.4759	-0.349	0	0.1595
Line DE	0.3685	0.2176	0.1595	0	0.4805

The basic OPF model for economic dispatch can be written as

$$\min \quad C_1 \cdot G_1 + C_2 \cdot G_2 + C_3 \cdot G_3 + C_4 \cdot G_4 + C_5 \cdot G_5$$

$$s.t. \quad \begin{cases} G_1 + G_2 + G_3 + G_4 + G_5 = D_\Sigma \\ \sum GSF_{1i} \times G_i + CL_1 = F_1^{\max} + \sum GSF_{1i} \times D_i \\ \sum GSF_{2i} \times G_i + UL_1 = F_2^{\max} + \sum GSF_{2i} \times D_i \\ \sum GSF_{3i} \times G_i + UL_2 = F_3^{\max} + \sum GSF_{3i} \times D_i \\ \sum GSF_{4i} \times G_i + UL_3 = F_4^{\max} + \sum GSF_{4i} \times D_i \\ \sum GSF_{5i} \times G_i + UL_4 = F_5^{\max} + \sum GSF_{5i} \times D_i \\ \sum GSF_{6i} \times G_i + UL_5 = F_6^{\max} + \sum GSF_{6i} \times D_i \end{cases}.$$

After solving the initial case OPF (load = 900MW), there are 2 marginal units and 3 non-marginal units as well as 5 un-congested lines and 1 congested line. Hence, we have 7 non-zero basic variables, 2 for the marginal units and 5 for the un-congested lines. Then, we can rewrite the above equations in matrix formulation as follows

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.4805 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.1595 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0.3600 & 0 & 1 & 0 & 0 & 0 \\ 0 & -0.5195 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0.1595 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0.1595 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} MG_1 \\ MG_2 \\ UL_1 \\ UL_2 \\ UL_3 \\ UL_4 \\ UL_5 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ -240 \\ 400 \\ 999 \\ -999 \\ 999 \\ -999 \end{bmatrix} + \begin{bmatrix} 1.0000 \\ -0.1257 \\ -0.2750 \\ 0.1493 \\ 0.1257 \\ 0.0584 \\ 0.3917 \end{bmatrix} D_\Sigma + \begin{bmatrix} -1.0000 & -1.0000 & -1.0000 \\ 0.3685 & 0.3685 & 0.1595 \\ -0.1939 & -0.1939 & 0.3490 \\ -0.4376 & -0.4376 & -0.1895 \\ -0.3685 & -0.3685 & -0.1595 \\ -0.1939 & -0.1939 & 0.3490 \\ -0.1939 & -0.1939 & -0.6510 \end{bmatrix} \begin{bmatrix} NG_1 \\ NG_2 \\ NG_3 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} CL_1$$

where MG_1 and MG_2 represent unit Sundance and Brighton respectively; NG_1, NG_2 and NG_3 represent the units at Alta, Park City, and Solitude, respectively; CL_1 represents line DE, UL_1 to UL_5 represent the remaining Lines, namely, AB, AD, AE, BC, and CD.

The above equation can be re-written as

$$\begin{bmatrix} MG_1 \\ MG_2 \\ UL_1 \\ UL_2 \\ UL_3 \\ UL_4 \\ UL_5 \end{bmatrix} = \begin{bmatrix} -499.5298 \\ 499.5298 \\ 320.3060 \\ 819.1642 \\ -739.4702 \\ 919.3060 \\ -1078.6940 \end{bmatrix} + \begin{bmatrix} 0.7384 \\ 0.2616 \\ -0.3167 \\ 0.0551 \\ 0.2616 \\ 0.0166 \\ 0.3500 \end{bmatrix} D_\Sigma + \begin{bmatrix} -0.2330 & -0.2330 & -0.6679 \\ -0.7670 & -0.7670 & -0.3321 \\ -0.0716 & -0.0716 & 0.4020 \\ -0.1615 & -0.1615 & -0.0699 \\ -0.7670 & -0.7670 & -0.3321 \\ -0.0716 & -0.0716 & 0.4020 \\ -0.0716 & -0.0716 & -0.5980 \end{bmatrix} \begin{bmatrix} NG_1 \\ NG_2 \\ NG_3 \end{bmatrix} + \begin{bmatrix} -2.0814 \\ 2.0814 \\ -0.3321 \\ -0.7493 \\ 1.0814 \\ -0.3321 \\ -0.3321 \end{bmatrix} CL_1 \quad . \tag{4.35}$$

At the present operating point ($D_\Sigma = 900$), we have the following results from the initial OPF

$$\begin{bmatrix} NG_1 \\ NG_2 \\ NG_3 \end{bmatrix} = \begin{bmatrix} 40 \\ 170 \\ 0 \end{bmatrix}.$$

$$\begin{bmatrix} CL_1 \end{bmatrix} = \begin{bmatrix} \end{bmatrix}$$

Therefore, we have

$$\begin{bmatrix} MG_1 \\ MG_2 \\ UL_1 \\ UL_2 \\ UL_3 \\ UL_4 \\ UL_5 \end{bmatrix} = \begin{bmatrix} 116.0757 \\ 573.9243 \\ 20.2495 \\ 834.8262 \\ -665.0757 \\ 919.2495 \\ -778.7505 \end{bmatrix}.$$

If the first two equations in (4.35) are put into the objective function, as shown in (4.10), we have

$$\begin{aligned} z &= \mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times \mathbf{P}_{\mathbf{MG}} + \left(\mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times \left[\mathbf{Q}_{\mathbf{MG}}\right]\right) \times \mathbf{D} \\ &\quad + \left(\mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times \left[\mathbf{R}_{\mathbf{MG}}\right] + \mathbf{C}_{\mathbf{NG}}^{\mathbf{T}}\right) \times \mathbf{NG} + \mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times \left[\mathbf{T}_{\mathbf{MG}}\right] \times \mathbf{CL} \\ &= -12488.2459 + 28.4595 \times D_{\Sigma} + \begin{bmatrix} 1.8256 & -0.8256 & 3.3015 \end{bmatrix} \times \begin{bmatrix} NG_1 \\ NG_2 \\ NG_3 \end{bmatrix} \\ &\quad + (-52.0344) \times CL_1 \end{aligned}.$$

Next, the load increase case will be taken to illustrate the process.

4.2.4.1. Calculate the next binding limit

Assuming a load variation ΔD_{Σ}, we have the following equation

$$\begin{bmatrix} \Delta MG_1 \\ \Delta MG_2 \\ \Delta UL_1 \\ \Delta UL_2 \\ \Delta UL_3 \\ \Delta UL_4 \\ \Delta UL_5 \end{bmatrix} = \begin{bmatrix} 0.7384 \\ 0.2616 \\ -0.3167 \\ 0.0551 \\ 0.2616 \\ 0.0166 \\ 0.3500 \end{bmatrix} \Delta D_{\Sigma} \ .$$

The allowed load growth corresponding to each un-congested line is given by (4.21)

$$\Delta D_k^{allowed} = \frac{-UL_k}{Q'_{UL_k}} \Rightarrow -\begin{bmatrix} \frac{20.2495}{-0.3167} \\ \frac{834.8262}{0.0551} \\ \frac{-665.0757}{0.2616} \\ \frac{919.2495}{0.0166} \\ \frac{-778.7505}{0.3500} \end{bmatrix} = \begin{bmatrix} 63.9391 \\ -15157.0897 \\ 2542.1269 \\ -55265.5401 \\ 2225.2134 \end{bmatrix}.$$

Considering the maximum capacity case in (4.24), the allowed load growth of each marginal generator is given by

$$\Delta D_i^{allowed} = \frac{MG_i^{\max} - MG_i}{Q'_{MG_i}} \Rightarrow \begin{bmatrix} \frac{200-116.0757}{0.7384} \\ \frac{600-573.9243}{0.2616} \end{bmatrix} = \begin{bmatrix} 113.6603 \\ 99.6694 \end{bmatrix}.$$

Similarly, for the minimum capacity, the allowed load growth of each marginal generator is given by

$$\Delta D_i^{allowed} = \frac{MG_i^{\min} - MG_i}{Q'_{MG_i}} \Rightarrow \begin{bmatrix} \frac{0-116.0757}{0.7384} \\ \frac{0-573.9243}{0.2616} \end{bmatrix} = \begin{bmatrix} -157.2035 \\ -2193.7179 \end{bmatrix}.$$

Therefore, the minimum positive value of the allowed load growth is 63.94MW, which corresponds to the congestion of the line flow AB in the positive direction. So, the next binding limit will be the line flow AB at the load 963.94MW.

4.2.4.2. Find the new marginal unit at load = 963.94 MW

When the system load grows to 963.94 MW at which a new binding transmission limit occurs (Line AB), the sensitivity of the new non-marginal generator sensitivity is given by (4.30)

$$\frac{\partial NG_j}{\partial D_\Sigma} = -\frac{Q'_{UL_r}}{R_{UL_{rj}}} \Rightarrow -\begin{bmatrix} \frac{-0.3167}{-0.0716} \\ \frac{-0.3167}{-0.0716} \\ \frac{-0.3167}{0.4020} \end{bmatrix} = \begin{bmatrix} -4.4260 \\ -4.4260 \\ 0.7879 \end{bmatrix}.$$

We also have

$$\begin{aligned} & \mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times [\mathbf{R}_{\mathbf{MG}}] + \mathbf{C}_{\mathbf{NG}}^{\mathbf{T}} \\ &= [5.0000 \quad 10.0000] \times \begin{bmatrix} -0.2330 & -0.2330 & -0.6679 \\ -0.7670 & -0.7670 & -0.3321 \end{bmatrix}. \\ &\quad + [4.0000 \quad 15.0000 \quad 30.0000] \\ &= [1.8256 \quad -0.8256 \quad 3.3015] \end{aligned}$$

The incremental cost vector for **NG** is

$$\begin{aligned} & (\mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times [\mathbf{R}_{\mathbf{MG}}] + \mathbf{C}_{\mathbf{NG}}^{\mathbf{T}})_j \times \left(-\frac{Q'_{MG_l}}{R_{MG_{lj}}} \right) \\ &\Rightarrow \begin{bmatrix} -1.8256 \times (-4.4260) \\ -0.8256 \times (-4.4260) \\ 3.3015 \times 0.7879 \end{bmatrix} = \begin{bmatrix} 8.0800 \\ 3.6540 \\ 2.6011 \end{bmatrix}. \end{aligned}$$

If we examine the sensitivity of un-congested lines based on (4.31), we have

$$\frac{\partial CL_k}{\partial D_\Sigma} = -\frac{Q'_{UL_r}}{T_{UL_{rk}}} \Rightarrow -\left[\frac{-0.3167}{-0.3321} \right] = [0.9537]$$

$$\mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times [\mathbf{T}_{\mathbf{MG}}] = [5.0000 \quad 10.0000] \times \begin{bmatrix} -2.0814 \\ 2.0814 \end{bmatrix} = -52.0344$$

$$(\mathbf{C}_{\mathbf{MG}}^{\mathbf{T}} \times [\mathbf{T}_{\mathbf{MG}}])_L \times \left(-\frac{Q'_{UL_r}}{T_{UL_{rk}}} \right) \Rightarrow [52.0344 \times (-0.9537)] = [9.6277].$$

If the incremental cost for **NG** and **UL** is the smallest positive value of 2.6011, which corresponds to $\mathbf{NG}_3$, the Solitude unit is at Bus C. So, the new marginal unit will be Solitude and there is no new congested line in this case.

4.2.4.3. Calculate the new LMP at load level 963.94 MW

For the new marginal unit set, apply (4.13) with $\mathbf{f}=[0\ 0\ \ldots 1 \ldots 0\ 0]^{\mathrm{T}}$, to calculate $\frac{\partial \mathbf{MG}}{\partial D_i}$ for the load variation at each single bus. Again, the load variation occurs at a specific bus only. Since we have $\left[\frac{\partial \mathbf{MG}}{\partial D_1} \quad \frac{\partial \mathbf{MG}}{\partial D_2} \quad \frac{\partial \mathbf{MG}}{\partial D_3} \quad \frac{\partial \mathbf{MG}}{\partial D_4} \quad \frac{\partial \mathbf{MG}}{\partial D_5}\right] =$

$$\begin{bmatrix} 0.3519 & -0.3636 & 0.0000 & 1.0000 & 0.0000 \\ 0.8261 & 0.0000 & 0.0000 & 0.0000 & 1.0000 \\ -0.1780 & 1.3636 & 1.0000 & 0.0000 & 0.0000 \end{bmatrix},$$

the LMP at Bus 1 can be calculated as

$$LMP_1 = \mathbf{C}_{MG}{}^{\mathrm{T}} \times \frac{\partial \mathbf{MG}}{\partial D_1} = \begin{bmatrix} 5 & 10 & 30 \end{bmatrix} \times \begin{bmatrix} 0.3519 \\ 0.8261 \\ -0.1780 \end{bmatrix} = 15.2379 .$$

Similarly, we can obtain the LMP for all buses from 963.94 MW to the next CLL as

$$\mathbf{LMP} = \begin{bmatrix} 15.2379 \\ 28.1818 \\ 30.0000 \\ 35.0000 \\ 10.0000 \end{bmatrix} (\$/\mathrm{MWh}).$$

4.2.5. Performance Speedup

As previously mentioned, this approach is particularly suitable for a short-term or online application. Hence, performance is very important. The advantage of this approach is to start from the present optimal state to directly evaluate the new CLL, and the associated

congestion and price step changes. This approach avoids repetitive optimization runs by taking advantage of features unique to the optimal dispatch model. Not surprisingly, this approach is computationally more efficient than the approach of repetitive optimization runs.

Here it is assumed that the range for trial-and-error is in 1000 intervals, such as from D_Σ to D_Σ+1000 MW with 1 MW as the acceptable accuracy or from D_Σ to D_Σ+100 MW with 0.1 MW as the acceptable accuracy. With the most optimistic assumption that there is only one step change during these intervals, we need to execute $\log_2 1000$ (≈ 10) DCOPF runs on average with a binary search, which is the most efficient searching algorithm in this case. With this estimated number of DCOPF runs, Table 4.3 shows that the speedup can be up to 51.6 for the IEEE 118-bus system. Here, speedup is defined as the average running time of the repetitive DCOPF-run approach divided by the average running time of the proposed algorithm, which provides the same output as the repetitive DCOPF runs such as the CLL, marginal units, congested lines, and LMPs. It is more encouraging to observe that the speedup increases with larger systems. This makes the direct approach highly promising for an online application, compared with the trial-and-error approach of repetitive OPF runs.

Table 4.3. Speedup of the proposed algorithm compared with the common practices of repetitive DCOPF runs

System	**Speedup compared with multiple (~10) DCOPF runs**
PJM 5-bus	15.2
IEEE 30-bus	30.0
IEEE 118-bus	51.6

The test of the DCOPF algorithm is implemented with Matlab packages using a linear programming function, *linprog()*. A sparse matrix technique is applied for both approaches for the larger systems, the IEEE 30-bus and 118-bus cases. It should be noted that although the kernel of some commercial LP packages may have the capability to perform a speeded follow-up LP run if it starts from the results of a previous case with careful data re-preparation, the repetitive OPF-run approach would still be more time-consuming. The reasons are: 1) there is a need to run the OPF multiple times to find the next CLL; and 2) each OPF, even if speeded up, should be still slower than the direct algorithm presented in this work, at least due to the overhead such as data re-preparation before each OPF. Moreover, if a higher resolution is needed for the CLL, the number of runs will increase beyond the assumed 10 times in the test presented here.

4.2.6. Discussion and Conclusions

Discussion

The above test illustrates that we can quickly obtain the congestion or binding constraints at the next CLL without repetitively running OPFs at many different load levels. In fact, if we start from zero loads, we can also efficiently calculate all binding constraints and prices at different load levels. Table 4.4 shows the marginal units and congested lines corresponding to the different CLLs for the PJM 5-bus case, calculated from the proposed approach. The price versus load curve can be easily plotted as well. It is ignored here since it is exactly the same as in Fig. 4.1. This proposed direct approach requires only six runs of the proposed algorithm because there are only six CLLs (i.e., step changes). As a comparison, to obtain Fig. 4.1 with a similar resolution of CLLs, hundreds of DCOPF runs are needed. This also shows the high efficiency and great potential of the proposed algorithm.

Table 4.4. Marginal units and congestion versus load growth

Load Range(MW)	**Marginal Unit(s)**	**Congested Line(s)**
0~600	Brighton	None
600~640	Alta	None
640~ 711.8083	Park City	None
711.8083~ 742.7965	Park City Brighton	ED
742.7965~ 963.9391	Sundance Brighton	ED
963.9391~ 1137.0152	Solitude Sundance Brighton	AB ED

Another note is that the proposed algorithm can be applied to the Continuous LMP (CLMP) methodology in [42] such a way that it is not necessary to re-run another optimization to obtain the LMP at the next CLL, $D_{\Sigma}^{(1)}$, after we calculate ΔD_{Σ} from $D_{\Sigma}^{(0)}$. Nevertheless, it is more important to emphasize the application of the algorithm in the presently dominant LMP paradigm, because the immediate and important application in congestion and price prediction versus load growth is apparent.

The best application of this work is for short-term operation and planning, when the load change in each bus or area should be close to linear and proportional, and the impact from other factors, like unit commitment, may not be a significant factor. If applied for long-term planning, the proposed model will be less accurate, if compared with the real-time operation. However, there is no existing model that can perform the same work easily and it is

sometimes unnecessary to obtain a high accuracy for long-term planning. So this work should be still valuable for long-term planning. Nevertheless, it is certain that the impact of the unit commitment is an area for future research.

In addition, the generation ramping rate is another factor to consider in the future, especially for short-term applications in which the ramping rate is a possible constraint. Other possible future works may lie in the different inputs such as the non-linear load variation pattern, generation uncertainty, transmission outage, and so on. If these are coupled with the loss and ACOPF models, it will become more complicated.

It is true that the running time of the proposed approach, after some modifications to address the above modeling details, will be slower than its present version. However, even with all these complications, the corresponding trial-and-error approach with repetitive optimization runs should be slower as well. Therefore, it would not be surprising if the relative speedup will be on the similar scale as that shown in Table 4.3.

The proposed algorithm is named the ‘Simplex-like’ method in this work because like the Simplex method in Linear Programming, it utilizes the concept of basic and non-basic variables and explores the sensitivity of constraints with respect to basic variables in order to determine the new basic set in the process of the algorithm. The primary difference between the simplex method and the proposed algorithm lies in the properties of the problem. For the Simplex method, it solves a linear programming problem, which is a fixed polyhedron. It starts from an initial point and jumps to the “best” adjacent extreme point until an optimal solution is found. However, for the proposed algorithm, it deals with an ever-changing polyhedron with respect to the load since most constraints change with the load. The algorithm starts from an optimal solution and finds the new optimal solution in one step when the load changes.

Conclusions

It is very useful to market-based operation and planning, especially in the short term, if information regarding congestion and price versus load can be easily obtained. The proposed algorithm helps the system operators and planners to easily identify possible congestion as the system load grows. It also provides useful information to generation companies to identify possible congestion and price changes as the system load grows, since many of these companies use the OPF model for congestion and price forecasting to achieve better economic benefits. Technical challenges arise if the load variation leads to a change of the binding constraint, which will lead to a change of the marginal unit set and a step change in the LMP. The previous work on the sensitivity of the LMP and other variables with respect to the load works only for a small load variation without a change of the binding constraints and cannot work when there is a large variation of the system load leading to a new congestion and a step change of the LMP.

This section presents a systematic approach to povide a global view of congestion and price versus load, from any given load level to another level, without multiple optimization runs. As shown in the mathematical derivation and case study, this approach is performed in the following steps:

- It first expresses marginal variables as a function of other non-marginal variables.
- Then, it identifies the next binding limit and the next critical load level (CLL).
- Next, the next unbinding limit such as a new marginal unit can be selected.
- Finally, the new generation output sensitivity at the CLL can be obtained because the objective function is expressed as non-marginal variables. Therefore, the new LMP can be obtained when the load is greater than the CLL.
- The same procedure can be repeated to run through another CLL.

In conclusion, this approach has great potential in market-based system operation and planning, especially for the short term, for congestion management and price prediction. Future works may lie in the impact of unit commitment, generation ramping rate, different load variation model, uncertainties, inclusion of loss model, and using ACOPF models.

4.3. Interpolation Method for ACOPF Framework

In the previous chapter, a computationally effective method for fast identification of CLLs is proposed for a fully linearized, DCOPF-based dispatch model. However, so far there are no available methods and tools that can predict the CLLs with respect to load variations within the framework of an ACOPF. It should be noted that an ACOPF is a closer representation of the actual operating model, which is the so-called successive LP OPF requiring iterations of solving the DCOPF and verifying transmission constraints by a full AC power flow.

Several researchers have utilized the sensitivity of system state variables to predict changes of the LMP under the ACOPF framework when the load changes. Reference [41] applies the perturbation method to calculate the LMP sensitivity. A similar approach is presented in [39]. However, these methods identify the sensitivities by essentially linearizing the optimality condition of an ACOPF model at a particular operating point, and therefore, the calculated results are only valid for a small change around that specific operating point. As shown in Figure 4.4, due to the nonlinearity of the AC model, the sensitivity at the present load level, D0, shown as the slope of the tangent line in Figure 4.4, should not be applied over a wide range. Hence, it cannot be used to predict the previous and next CLLs, i.e., Points A and B in Figure 4.4, which have different sensitivities from the present operating point in the non-linear AC model.

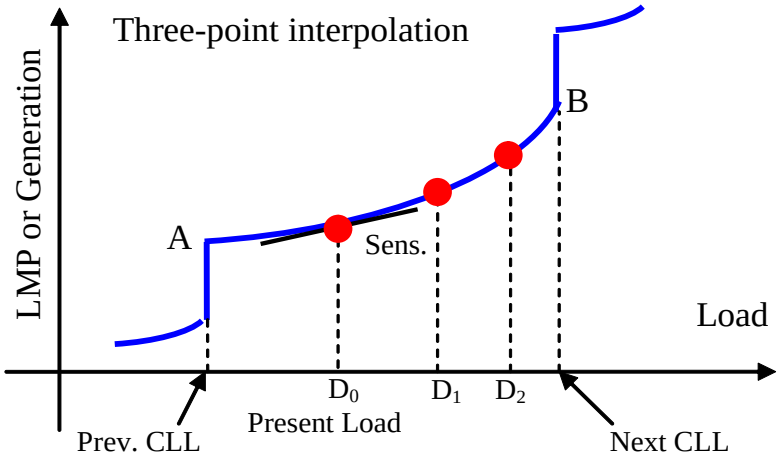


Figure 4.4. Illustration of the non-linear relation between the LMP or generation versus the system load level

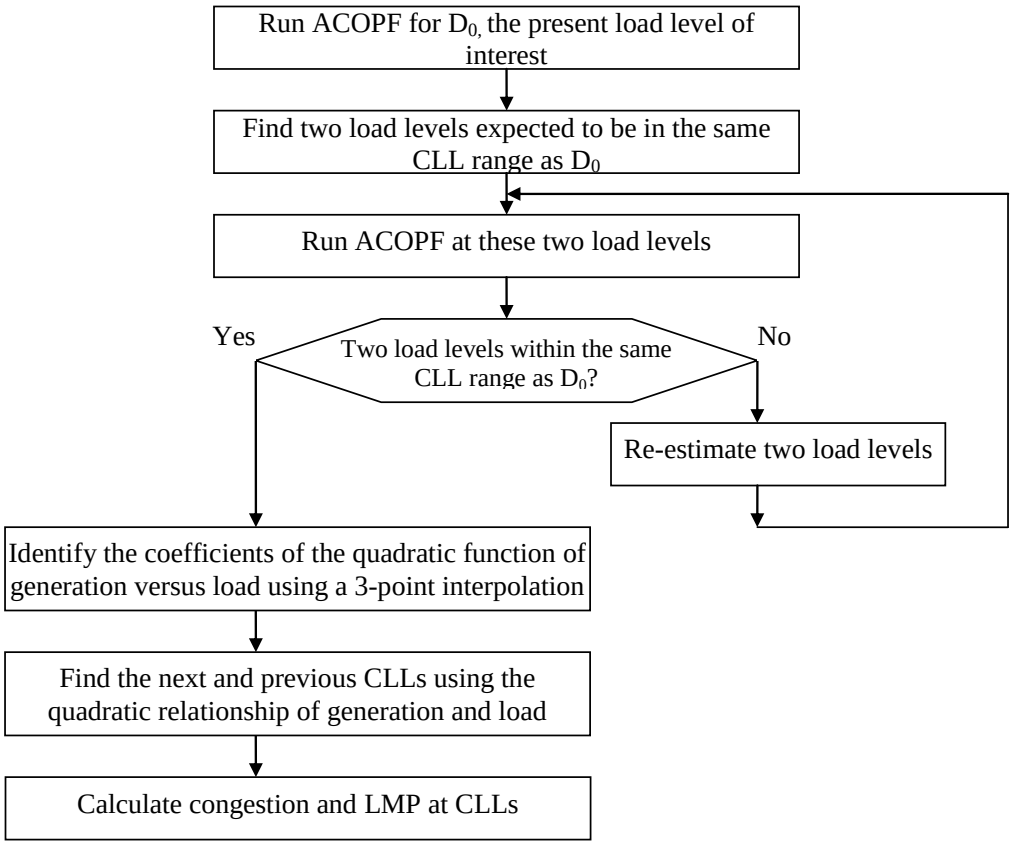


Figure 4.5. High-level illustration of the proposed method

To find CLLs under the non-linear model, a straightforward, brute-force approach is to repetitively run the ACOPF at different load levels. This is certainly not desirable, especially for short-term applications. Even for long term applications, this will be computationally problematic if multiple load variation patterns under different possibilities need to be considered. This essentially adds another dimension of complexity because multiple repetitive-ACOPF runs are needed. Hence, a more efficient way is of high interest. This is also the motivation of the proposed approach, schematically illustrated in Figure 4.5. As shown in the figure, an initial ACOPF is run at the present load level of interest. Then, the ACOPF runs can be performed at two load levels (such as D1 and D2 in Figure 4.4), which are expected to be within the same two adjacent CLLs as the present load level D0. If not, an adjustment will be made to D1 and D2. Then, the marginal unit generation and LMPs can be expressed as some analytical function (shown as quadratic in this chapter) of the system load through interpolation, using results at D0, D1, and D2. Two noteworthy points are listed below:

1. The LMP sensitivity versus system load is directly related to the derivative of the marginal generation sensitivity with respect to the system load, because the LMP can be viewed as a weighted summation of the marginal unit costs with the marginal generation sensitivity as the weights.

2. The analytical function of the marginal unit generation output versus system load may follow a complicated, high-order polynomial given that the system load variation is between two neighboring CLLs. Fortunately, the numerical study via curve-fitting technique in this chapter shows that a quadratic relation is sufficiently accurate. Then, only three points are needed to interpolate the generation output versus load level, as shown in Figure 4.4.

This chapter is organized as follows. First, to address the challenges of predicting the system status (prices, transmission congestions, and generation dispatch,) under the ACOPF framework, this paper applies polynomial curve-fitting to discover the quadratic variation pattern of the system statuses such as generator dispatches and line flows, with respect to load changes. Second, in order to minimize the computational efforts brought by polynomial curve-fitting approaches, an algorithm based on the quadratic interpolation is proposed to effectively identify the coefficients of the quadratic pattern and correspondingly predict the CLLs of the system.

4.3.1. Polynomial Curve-fitting for Marginal Unit Generation and Line Flow

4.3.1.1. Variation Pattern of Marginal Unit Generation and Line Flow with Load Changes

In Section 4.2, it is rigorously proved that for a lossless DCOPF simulation model, generations of all the marginal units follow a linear pattern with respect to load variation. However, for a more accurate ACOPF framework, losses are not negligible and introduce the challenge of nonlinearity. It is natural to bring up the following question: *What type of nonlinear pattern do the marginal unit generations and line flows follow, with respect to load changes, under an ACOPF framework?*

It is hard to address this question analytically due to the nonlinearity of the ACOPF model. A sensitivity analysis might be one option. It is easy to calculate the sensitivities of the generation or line flow, with respect to load changes, for power flow problems; however, it is much more difficult to calculate the sensitivities within the OPF framework.

Reference [41] employs a perturbation method to obtain the sensitivity of the LMP with respect to load. The same idea can be applied to numerically calculate the sensitivity of the marginal unit generation and line flow with respect to load, but an analytical formulation is absent. Reference [39] derives a symbolic formulation for generation sensitivity to the load; however, the sensitivity formula contains system state variables (such as voltage magnitudes and angles), which themselves are unknowns at a new operating point, and therefore, cannot be used for identifying the variation pattern outside a certain load range.

Here, these questions are studied through numerical methods based on polynomial curve-fitting. While more sophisticated pattern matching approaches could be used, the results will show this approach to be highly accurate.

4.3.1.2. Application of Polynomial Curve-fitting for Marginal Unit Generation and Line Flow

A typical ACOPF model can be found in Section 3.2.2. It should be noted that the objective function is the total cost of generation, which is assumed to be linear. The MATPOWER package is employed to solve the ACOPF problem [17]. When the solved ACOPF runs at sampling load levels, we obtain the data of the marginal unit generation and line flow. They are viewed as the benchmark data and serve as the input for the polynomial curve-fitting.

Assume both the marginal unit generation and line flow data are fitted by the polynomial functions as follows

$$MG_j = a_{n,j}D^n + a_{n-1,j}D^{n-1} + \cdots + a_{1,j}D + a_{0,j}, \forall j \in \mathcal{MG} \tag{4.36}$$

$$F_k = b_{n,k}D^n + b_{n-1,k}D^{n-1} + \cdots + b_{1,k}D + b_{0,k}, \forall k \in \{\mathcal{B} \setminus \mathcal{Bc}\} \tag{4.37}$$

where

MG_j is the generation of the marginal unit j, which is obtained from an ACOPF run;

$a_{i,j}$ represents the i^{th} degree coefficient of the polynomial function of the marginal unit j;

D is the total system load;

n is the degree of the polynomial function;

$\mathcal{MG}$ is the marginal unit set;

F_k is the line flow through line k, which is obtained from the ACOPF run;

$b_{i,k}$ represents the i^{th} degree coefficient of the polynomial function of the line flow through line k;

$\mathcal{B}$ represents the set of all lines;

$\mathcal{Bc}$ represents congested line set;

$\{\mathcal{B} \setminus \mathcal{Bc}\}$ represents the non-congested line set.

For the j^{th} marginal unit, a set of the generation data at m different load levels are available from he ACOPF runs. The corresponding curve-fitting formulation is given as

$$\begin{bmatrix} MG_j^{(0)} \\ MG_j^{(1)} \\ \vdots \\ \vdots \\ MG_j^{(m)} \end{bmatrix} = \begin{bmatrix} (D^{(0)})^n & (D^{(0)})^{n-1} & \cdots & D^{(0)} & 1 \\ (D^{(1)})^n & (D^{(1)})^{n-1} & \cdots & D^{(1)} & 1 \\ & & \vdots & & \\ & & \vdots & & \\ (D^{(m)})^n & (D^{(m)})^{n-1} & \cdots & D^{(m)} & 1 \end{bmatrix} \times \begin{bmatrix} a_{n,j} \\ a_{n-1,j} \\ \vdots \\ a_{0,j} \end{bmatrix} \tag{4.38}$$

where the superscript in parenthesis, and *(i)*, represents the i^{th} sampling load level, i=1, 2, …, m.

In matrix form, this can be written as

$$\mathbf{MG}_j = \mathbf{A} \times \mathbf{a}_j \tag{4.39}$$

where

$\mathbf{MG_j}$ is an $m\times 1$ vector;

$\mathbf{A}$ is an $m\times(n+1)$ matrix;

$\mathbf{a_j}$ is an $(n+1)\times 1$ vector (Normally n is much less than m.).

The problem formulated in (4.39) implies more known variables, $\mathbf{MG_j}$ than unknowns, $\mathbf{a_j}$. So, there are redundant equations. The curve-fitting problem for line flows can be formulated and solved in a similar way, and therefore, is not repeated here.

Typically, equation (4.39) can be solved using the least-square algorithms. It should be noted that the condition number of matrix $\mathbf{A}$ in equation (4.39) could be high due to its construction in polynomial pattern, and therefore the solution $\mathbf{a_j}$ may be highly sensitive to small changes in $\mathbf{MG_j}$. Typically in this case, we should perform some process to the original data such as scaling and dropping rank, and then re-compute. This process is skipped in this work because the solution of (4.39) is only used in computer simulation, instead of in constructing physical systems. Therefore, the instability of solution will not incur significant additional costs.

4.3.2. Numerical Study of Polynomial Curve-fitting

This section presents the numerical study for the polynomial curve-fitting of benchmark data of the marginal unit generations and line flows for a modified PJM 5-bus system and the IEEE 30-bus system. Results show that the benchmark data can be well approximated by polynomial curve-fitting, with a quadratic curve-fitting having the least computational effort, but still maintaining a high accuracy. Therefore, only the quadratic curve-fitting results are presented. For simplicity, the load is assumed to follow a variation pattern where the load increases proportionally to the base load at each load bus. Other load change patterns can be defined and easily employed.

4.3.2.1. Results for PJM 5-Bus System

The first test system is the small, yet informative PJM 5-bus system [11] with modifications, as detailed in section 4.2. The base case diagram of the system is shown in Figure 4.3. In the ACOPF runs, all loads are assumed to have a 0.95 lagging power factor. The generators are assumed to have a reactive power limit of 150 MVar capacitive to 150 MVar inductive. This is selected so that the system has sufficient reactive power resources and system voltage profile is not a major concern. The R/X ratios of the transmission lines are set at 10%.

Figure 4.6 compares the benchmark data from the ACOPF results and quadratic curve-fitting results of the generation of the marginal unit Sundance with the load variation from 900 MW to 922.5 MW, a 2.5% increase of the base case load. During this load range, the Line ED is always congested, and the marginal units are Sundance and Brighton. The differences between the benchmark data and curve-fitting results are shown in Figure 4.6.

The benchmark data and the quadratic curve-fitting results of the flow on the Line AB, and their differences in percentages are shown in Fig. 4.7. The flow on Line AB is steadily increasing toward its thermal limit. Note that the line flow is in MVA not MW.

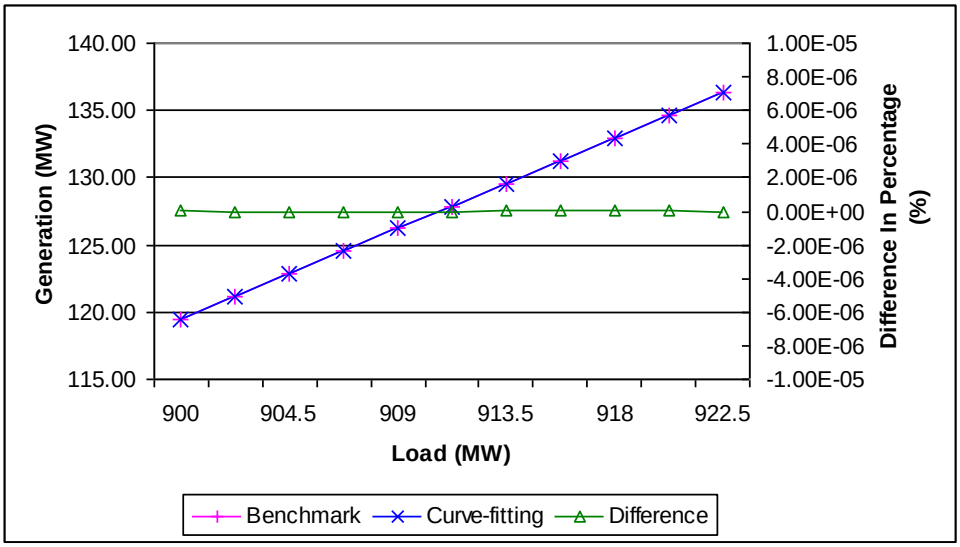


Figure 4.6. Quadratic curve-fitting results, the benchmark data, and their differences of the generation of the marginal unit Sundance for the PJM 5-bus system

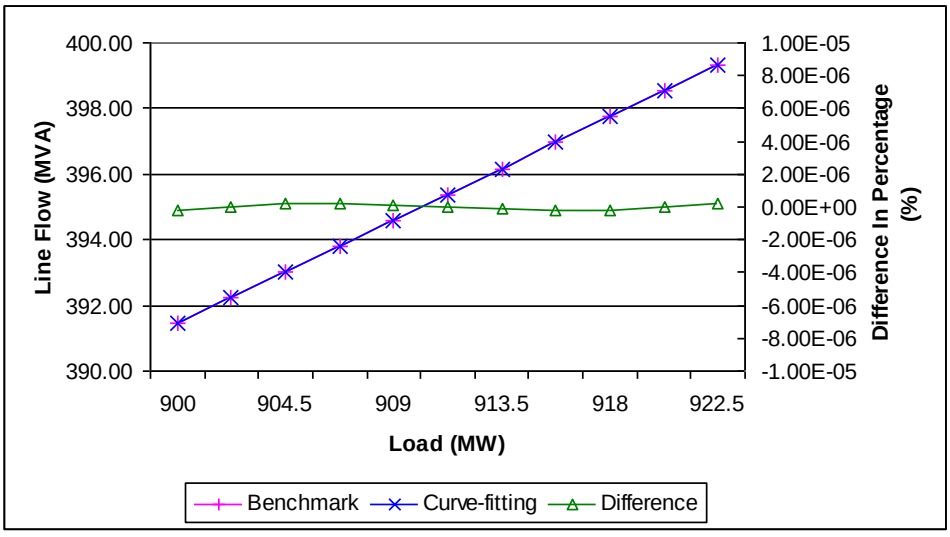


Figure 4.7. Quadratic curve-fitting results, the benchmark data, and their differences of the line flow through the Line AB for the PJM 5-bus system

Figures 4.6-4.7 demonstrate that the quadratically fitted curves are a good approximation to the benchmark data for both the marginal unit generation and line flow. It is also true for all other marginal unit generations and line flows. In fact, similar patterns are observed for simulations at other load intervals which do not contain a CLL.

It should be noted that in Figs. 4.6-4.7, the difference percentage between the benchmark data and quadratic curve-fitting results is less than 0.00001%. This is a small number given that the simulations span a 2.5% load variation around the base case. In addition, the high accuracy of the approximation achieved by the quadratic curve-fitting can be maintained for larger load range as long as there is no change of binding constraints during the load variation window, namely, the load range does not contain a CLL.

When the polynomial curve-fitting of a higher degree is applied to fit the marginal unit generation and line flow, a high accuracy of the fit is also expected. However, the quadratic curve-fitting is accurate and recommended since it leads to less computational efforts. In addition, the linear curve-fitting is reasonable in the modified PJM 5-bus system; however, this is expected to have larger errors for systems demonstrating greater nonlinearity, as exemplified in the next section.

4.3.2.2. Results for IEEE 30-Bus System

The second test system is the IEEE 30-bus system. The detailed system configuration and data are available in [16]. The bidding prices of the 6 generators are assumed here to be 10, 15, 30, 35, 40, and 45, respectively, all in $/MWh. The branch susceptances and the transformer tap ratios are all ignored for simplicity. To create a scenario with more than one path of congestion and help the ACOPF converge over a wider range from the base-case, the network data is slightly modified: 1) load power factor is kept at a 0.95 lagging as load

changes; 2) the thermal limit of lines 6-8, 21-22, 25-27 is increased by 10%, 10%, and 30%, respectively; and 3) the thermal limit of Line 12-13 is reduced by 50%.

The studied load range is from 189.20 MW to 193.93 MW, namely, 1.0 p.u. to 1.025 p.u.. A comparison of the benchmark data and quadratic curve-fitting results is shown in Figs. 4.8-4.9. For simplicity, only the marginal unit generation at Bus 22 and the line flow through line 24-25 are depicted. It can be clearly seen that the quadratic curves fit the benchmark data well. Again, this is true only when the studied load range is within two adjacent critical load levels.

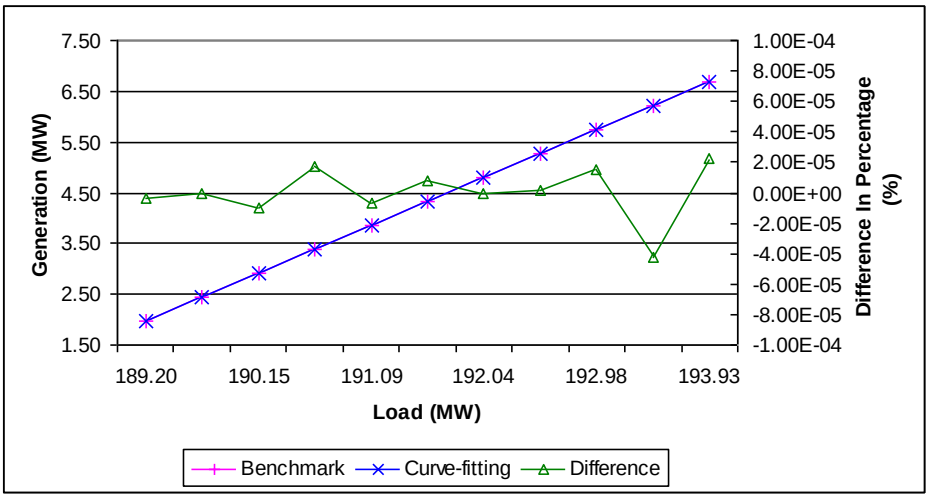


Figure 4.8. Quadratic curve-fitting results, the benchmark data, and their differences of the generation of the marginal unit at Bus 22 for the IEEE 30-bus system

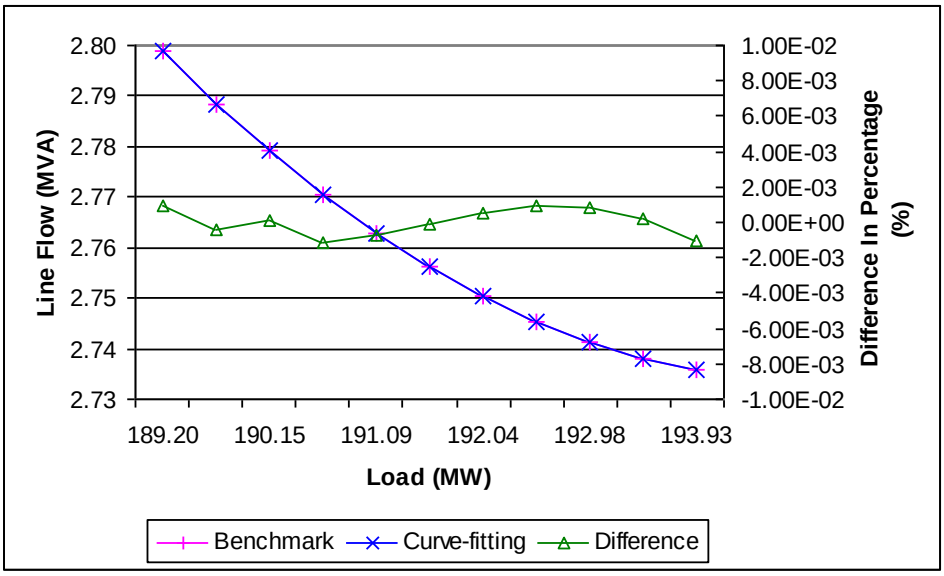


Figure 4.9. Quadratic curve-fitting results, the benchmark data, and their differences of the line flow through Line 24-25 for the IEEE 30-bus system

Table 4.5. Polynomial coefficients of the quadratic curve-fitting results for the generation of marginal unit at Bus 22 and the line flow through Line 24-25 for the IEEE 30-bus system

Polynomial Coefficients	Generation of Marginal Unit at Bus 22	Line Flow through Line 24-25
a_2	3.20×10^{-4}	0.001996
a_1	0.8699	-0.7782
a_0	-174.07	78.58

The polynomial coefficients of the quadratic curve-fitting for the generation of the marginal unit at Bus 22 and the line flow on Line 24-25 are shown in Table 4.5.

Figure 4.9 and Table 4.5 demonstrate a clear quadratic pattern of the line flow through Line 24-25. From Fig. 4.9 it can be easily seen, even by visual inspection, that the line flow follows a non-linear curve, and the linear curve-fitting should have considerable errors even in the studied small range of the load variation. Hence, it is not advisable to use a linear approximation under the ACOPF framework. In fact, the sum of the squares, due to the error (SSE) and R-square for quadratic curve-fitting are 4.64×10^{-9} and 1.0, respectively, which is far superior to 1.71×10^{-4} and 0.96 for a linear model.

4.3.3. Quadratic Interpolation Method

The curve-fitting results in the previous section are encouraging since the variations of the marginal generation and line flow with respect to load changes suggest a nearly perfect quadratic pattern within two adjacent CLLs and hence, facilitate prediction of CLLs. However, it is not practically useful because it involves numerous ACOPF studies at different load levels to get the benchmark data for the curve-fitting. Therefore, a practical approach requiring less computational efforts is needed. In this section a quadratic interpolation approach is proposed.

The basic idea is to solve an ACOPF at three different load levels and apply the quadratic interpolation using the benchmark data at these three load levels. The crucial problem here is to ensure that all the three load levels are between adjacent CLLs. To locate three load levels satisfying this requirement, an empirical setting or a DCOPF-based approach may offer assistance.

4.3.3.1. Three-Points Pattern for Quadratic Interpolation

The detailed procedure of obtaining the three load levels are presented as follows:

1) The load level of the initial operating point is taken as the first load level, denoted by $D^{(0)}$;

2) Obtain an initial estimate for the second load level $D^{(1)}_{guess}$. It could be empirically determined, for example, $D^{(0)}$ plus 0.025 p.u.. $D^{(1)}_{guess}$ could also be set as the estimated critical load level in load growth direction, denoted $D^{critical}_{+}$, by solving a DCOPF-based congestion prediction at the initial operating point as in section 4.2. Some downscaling of $D^{critical}_{+}$ can be applied to make it more possible to be less than the next CLL;

3) Run ACOPF at $D^{(1)}_{guess}$, and examine the marginal unit and congested line set. If they are the same as those at the first load level $D^{(0)}$, then $D^{(1)}_{guess}$ is selected as the second load level, denoted by $D^{(1)}$, and go to 5); otherwise, go to 4);

4) Set $D^{(0)} + (D^{(1)}_{guess} - D^{(0)})/2$ as the new $D^{(1)}_{guess}$, go to 3);

5) Take $(D^{(0)} + D^{(1)})/2$ as the third load level, denoted by $D^{(2)}$.

In many cases, the $D^{(1)}_{guess}$ obtained in step 2 will qualify for the second load level. Hence, in step 3, only one additional ACOPF run is performed for verification purposes. In case the $D^{(1)}_{guess}$ obtained in step 2 lies beyond the next (or previous) critical load level, $D^{(1)}_{guess}$ will be

updated iteratively towards $D^{(0)}$. Therefore, normally, in step 3, only a few additional ACOPF runs are needed, even in the worst case scenario.

In step 2, it is not advisable to set $D^{(1)}_{guess}$ near the initial operating point. The reason is that the three load levels are expected to cover as many operating points as possible to be representative and avoid numerical errors in the computation. $D^{(1)}_{guess}$ can also be set to a value close to the CLL calculated for the DCOPF models, as presented in Section 4.2, since the DCOPF model may produce the same marginal unit and binding constraint sets as its ACOPF counterpart for a large portion of the load levels, as reported in Section 3.3 and Reference [24].

4.3.3.2. Quadratic Interpolation for Marginal Unit Generation and Line Flow

The ACOPF results for the first load level is intended to be an input to interpolation, and the ACOPF run for the second load level is done during the search for the three load levels. Therefore, one more ACOPF run needs to be performed at the third load level.

With ACOPF results at all three load levels, a quadratic interpolation can be performed on each marginal unit and line flow. Consider the generation of marginal unit *j* as an example, (4.39) can be rewritten as

$$\overline{\mathbf{MG}}_j = \overline{\mathbf{A}} \times \bar{\mathbf{a}}_j \qquad (4.40)$$

where $\overline{\mathbf{MG}}_j$ is a *3*×*1* vector; $\overline{\mathbf{A}}$ is a *3*×3 matrix; and $\bar{\mathbf{a}}_j$ is a 3×*1* vector. It is apparent that the coefficients $\bar{\mathbf{a}}_j$ can be uniquely determined.

It should be noted that with a good initial guess, the quadratic interpolation requires only two additional ACOPF runs and can be solved very efficiently. In contrast, the quadratic curve-fitting requires numerous, additional ACOPF runs.

4.3.3.3. Prediction of Critical Load Levels

With the knowledge of how the marginal unit generation and line flow will change with respect to th load variation, which is shown to follow quadratic patterns, it is easy to forecast the critical load levels as the load increases or decreases. Let $\Delta D_{\Sigma}\big|_j^{ub}$ ($\Delta D_{\Sigma}\big|_j^{lb}$) represent the minimum system load change from the initial operating point until the upper (lower) limit of marginal generator *j* is reached. Similarly, let $\Delta D_{\Sigma}\big|_k^{ub}$ ($\Delta D_{\Sigma}\big|_k^{lb}$) represent the minimum load change from initial operating point until the k^{th} transmission line reaches its limit in the positive (negative) direction.

Then, these load variations $\Delta D_{\Sigma}\big|_j^{ub}$, $\Delta D_{\Sigma}\big|_j^{lb}$, $\Delta D_{\Sigma}\big|_k^{ub}$ and $\Delta D_{\Sigma}\big|_k^{lb}$ can be obtained by solving the following quadratic equations

$$a_{2,j}(D_{\Sigma}^{(0)} + \Delta D_{\Sigma}\big|_j^{ub})^2 + a_{1,j}(D_{\Sigma}^{(0)} + \Delta D_{\Sigma}\big|_j^{ub}) + a_{0,j} = MG_j^{\max}, \forall j \in \mathcal{MG} \quad (4.41)$$

$$a_{2,j}(D_{\Sigma}^{(0)} + \Delta D_{\Sigma}\big|_j^{lb})^2 + a_{1,j}(D_{\Sigma}^{(0)} + \Delta D_{\Sigma}\big|_j^{lb}) + a_{0,j} = MG_j^{\min}, \forall j \in \mathcal{MG} \quad (4.42)$$

$$b_{2,k}(D_{\Sigma}^{(0)} + \Delta D_{\Sigma}\big|_k^{ub})^2 + b_{1,k}(D_{\Sigma}^{(0)} + \Delta D_{\Sigma}\big|_k^{ub}) + b_{0,k} = F_k^{\max}, \forall k \in \{\mathcal{B} \setminus \mathcal{B}c\} \quad (4.43)$$

$$b_{2,k}(D_{\Sigma}^{(0)} + \Delta D_{\Sigma}\big|_k^{lb})^2 + b_{1,k}(D_{\Sigma}^{(0)} + \Delta D_{\Sigma}\big|_k^{lb}) + b_{0,k} = -F_k^{\max}, \forall k \in \{\mathcal{B} \setminus \mathcal{B}c\} \quad (4.44)$$

where

$MG_j^{\max}$ and $MG_j^{\min}$ are the maximum and minimum generation capacity of marginal unit *j*; and

$F_k^{\max}$ is the thermal limit of line *k*.

Further, these load variations will determine the margin from the present load level to the nearest load level where there is a change of the binding constraints. The load variation's positive and negative directions are defined, respectively, as

$$\Delta D_{\Sigma+}^{margin} = \min_{j\in\mathcal{MG},k\in\mathcal{B},\Delta D_\Sigma\geq 0}\{\Delta D_\Sigma\big|_j^{ub}, \Delta D_\Sigma\big|_j^{lb}, \Delta D_\Sigma\big|_k^{ub}, \Delta D_\Sigma\big|_k^{lb}\} \tag{4.45}$$

$$\Delta D_{\Sigma-}^{margin} = \max_{j\in\mathcal{MG},k\in\mathcal{B},\Delta D_\Sigma< 0}\{\Delta D_\Sigma\big|_j^{ub}, \Delta D_\Sigma\big|_j^{lb}, \Delta D_\Sigma\big|_k^{ub}, \Delta D_\Sigma\big|_k^{lb}\} \ . \tag{4.46}$$

Once the margin is determined, the new binding constraint, either generation or transmission, is simultaneously identified. For instance, when a transmission line constraint becomes binding, a new congestion is identified. Thus, this important information can be easily obtained without doing exhaustive simulations on all load levels.

Finally, the previous and next critical load levels are determined by

$$D_{\Sigma-} = D_\Sigma^{(0)} + \Delta D_{\Sigma-}^{margin} \tag{4.47}$$

$$D_{\Sigma+} = D_\Sigma^{(0)} + \Delta D_{\Sigma+}^{margin} \tag{4.48}$$

where

$D_{\Sigma-}$ and $D_{\Sigma+}$ are previous and next critical load levels respectively; and

$D_\Sigma^{(0)}$ is the present load level.

4.3.4. Case Study of Prediction of Critical Load Levels

In this section, the proposed approach of predicting critical load levels, which employs a quadratic interpolation, will be tested on the PJM 5-bus system and IEEE 30-bus system. Prediction results will be compared with those utilizing a quadratic curve-fitting. In addition, the predicted previous and next critical load levels will be compared with the benchmark data obtained from the enumerative simulation.

4.3.4.1. Results for the PJM 5-Bus System

For notational convenience, the generators Alta, Park City, Solitude, Sundance, and Brighton are numbered from 1 through 5, respectively. The congestion prediction study is performed at the 900 MW load level, namely, 1.0 p.u. of the base-case. In this case, there are two marginal units, Sundance at Bus 4 and Brighton at Bus 5, and Line ED is congested. For the quadratic curve-fitting, the ACOPF simulations are performed on eleven load levels evenly distributed between 900 MW and 922.5 MW.

Table 4.6 shows the load variation distances calculated by the quadratic curve-fitting approach as introduced in Section 4.3.1, and by the quadratic interpolation approach as proposed in Section 4.3.3, respectively. The numbers in bold font in Table 4.6 are actually the load variation margins from the present load level. The predicted previous and next critical load levels are compared with actual values obtained from the enumerative ACOPF simulation, as shown in Table 4.7.

Table 4.6. Load margins from the present operating point for the PJM 5-bus system

Load Variation (MW)	Quadratic Curve-fitting Approach	Quadratic Interpolation Approach
$\Delta D_{\Sigma}\big\vert_{4}^{ub}$	107.87	107.87
$\Delta D_{\Sigma}\big\vert_{5}^{ub}$	81.79	81.79
$\Delta D_{\Sigma}\big\vert_{4}^{lb}$	**-160.37**	**-160.37**
$\Delta D_{\Sigma}\big\vert_{5}^{lb}$	-2,247.91	-2247.85
$\Delta D_{\Sigma}\big\vert_{lineAB}^{ub}$	**24.41**	**24.41**
$\Delta D_{\Sigma}\big\vert_{lineAD}^{ub}$	-13,986.71	-13,992.13
$\Delta D_{\Sigma}\big\vert_{lineAE}^{ub}$	2195.52	2194.63
$\Delta D_{\Sigma}\big\vert_{lineBC}^{ub}$	-49377.00	-49445.66
$\Delta D_{\Sigma}\big\vert_{lineCD}^{ub}$	1951.48	1949.30
$\Delta D_{\Sigma}\big\vert_{lineAB}^{lb}$	-2825.97	-2826.52
$\Delta D_{\Sigma}\big\vert_{lineAD}^{lb}$	28,623.66	28,571.64
$\Delta D_{\Sigma}\big\vert_{lineAE}^{lb}$	N/A	N/A
$\Delta D_{\Sigma}\big\vert_{lineBC}^{lb}$	67,849.08	67,682.47
$\Delta D_{\Sigma}\big\vert_{lineCD}^{lb}$	-5287.72	-5357.34

Note: N/A represents no solution for Equations (4.41)-(4.44).

Table 4.7. Previous and next critical load levels for the PJM 5-bus system

CLL (MW)	Quadratic Curve-Fitting Approach	Quadratic Interpolation Approach	From Actual Enumerative Simulation	New Binding Constraint
$D_{\Sigma+}$	924.41	924.41	924.40	Line AB
$D_{\Sigma-}$	739.63	739.63	739.56	Generator Sundance

From Tables 4.6 and 4.7, it can be seen that Line AB will reach its limit in the positive direction if the system load increases by 24.41 MW from the present load level. In other words, the constraint for Line AB will be binding, and the system will have one additional point of congestion. This is the first change of the binding constraints with a load increase. In the case of a load decrease, the first change of binding constraints will occur when the load decreases by 160.37 MW, at which point the marginal unit Sundance at Bus 4 will reach its lower limit and is no longer a marginal unit. At this load level, the Park City unit will become a marginal unit, but Line ED remains congested.

Tables 4.6-4.7 show that the predication results obtained from the quadratic curve-fitting and quadratic interpolation methods are almost identical and also match the benchmark results precisely. This demonstrates that the quadratic interpolation method successfully achieves the desired results while greatly reducing the computational effort.

The reason for the good results in Table 4.7 is that the calculated polynomial coefficients from both approximation approaches are numerically very close. As an example, Table 4.8 shows the nearly identical coefficients of the polynomial function for the marginal unit Sundance for both approaches.

Table 4.8. Polynomial coefficients of the generation of the marginal unit Sundance from the quadratic curve-fitting and quadratic interpolation approaches for the PJM 5-bus system

Polynomial Coefficients	Quadratic Curve-fitting Approach	Quadratic Interpolation Approach
a_2 (MW^{-1})	4.1041×10^{-6}	4.1022×10^{-6}
a_1	0.7384	0.7384
a_0 (MW)	-548.4064	-548.4079

4.3.4.2. Results for the IEEE 30-Bus System

The study on the IEEE 30-bus system is performed at the load level of 189.20 MW, namely, 1.0 p.u. of the base case load. In this case, there are two marginal units and one congested line. For a quadratic curve-fitting, ACOPF simulations are conducted on eleven load levels evenly distributed between 189.20 MW and 193.93 MW.

For simplicity, the detailed results of the load variation distances for each marginal unit and non-congested line will not be detailed in this paper, and only the load variation margins are presented, as shown in Table 4.9.

Table 4.9. Previous and next critical load levels from the present operating point for the IEEE 30-bus system

CLL (MW)	Quadratic Curve-Fitting Approach	Quadratic Interpolation Approach	From Actual Enumerative Simulation	New Binding Constraint
$D_{\Sigma+}$	219.81	219.81	219.82	Line 8-6
$D_{\Sigma-}$	187.20	187.20	187.20	Generator @ Bus 22

Once again, Table 4.9 demonstrates that both the quadratic curve-fitting and quadratic interpolation approach provide highly accurate results when compared with the benchmark data obtained from the enumerative simulation.

In the studied case, 11 load levels are chosen to render curve-fitting. The quadratic curve-fitting approach therefore requires at least 10 ACOPF runs at selected load levels other than the initial operating point, while the quadratic interpolation approach typically requires only 2 additional ACOPF runs. Hence, the quadratic interpolation method is a computationally efficient approach and produces highly reliable results, and therefore, has the potential to be employed in real applications to predict critical load levels. The applications can be for short-term planning for market participants, as well as long-term planning when multiple load variation patterns under different possibilities are considered, leading to multiple repetitive-ACOPF runs.

4.3.5. Discussion and Conclusions

This section applies polynomial curve-fitting to identify the variation patterns of the marginal unit output and line flow with respect to load changes. Numerical studies on the PJM 5-bus system and the IEEE 30-bus system show that the marginal unit generation and line flow follow a nearly perfect polynomial pattern. In particular, a quadratic polynomial is recommended as it provides adequate accuracy and requires fewer computations than a higher order curve-fitting.

Next, a quadratic-interpolation-based approach is proposed in order to further reduce computational efforts. The approach requires ACOPF data at three load levels to perform the calculation. A heuristic algorithm for seeking these three load levels within two adjacent CLLs is presented, in which an estimated critical load level obtained from a DCOPF may serve as the initial guess for the search. Then, the approach of predicting CLLs which employs the quadratic pattern of the marginal unit generation and line flow is presented.

The proposed approach using quadratic interpolation is compared with the approach employing quadratic curve-fitting. Both approaches are tested on the PJM 5-bus system and IEEE 30-bus system. Results show that the polynomial coefficients calculated from the quadratic interpolation are very close to those obtained from the quadratic curve-fitting. The results of the predicted critical load levels are also verified and are sufficiently close to the actual values obtained from the enumerative ACOPF simulations. In addition to the prediction of CLLs, a new binding constraint, such as a new congestion, is simultaneously identified.

The application of the proposed method can be for short-term planning, as well as long-term planning when multiple repetitive-ACOPF runs are needed to evaluate possible different

load variation scenarios. Future work may include the impact of insufficient reactive power support.

4.4. Variable Substitution Method for FND-based DCOPF Framework

As presented in Section 3.3, the FND-based DCOPF results are shown to be reasonably close to the ACOPF for the majority of the studied load levels, in terms of the dispatched generation and LMP calculation. Certainly, the interpolation method proposed in Section 4.3 could easily be applied to the FND-based DCOPF framework to predict CLLs, congestions, and LMPs. However, in contrast to the ACOPF model, the FND-based DCOPF model is a simplified OPF model, which makes it possible to propose methods involving even less computational efforts than the interpolation method, which requires three OPF runs.

One approach is to utilize the quadratic characteristics of the marginal unit generation, which is verified in Section 4.3. Together with the solution features of the FND-based DCOPF model for load levels within two adjacent CLLs, a set of equations will be formulated to solve for the coefficients of the quadratic functions.

4.4.1. Characteristic Constraints of the FND-based DCOPF model

4.4.1.1. Revisit of the FND-based DCOPF Model

The FND-based DCOPF model is represented by equations (3.30)-(3.33) in Section 3.3. According to the classification of the generation units: marginal unit and non-marginal unit generation, the model is rewritten as follows

$$\min_{\mathbf{MG}} \quad \sum_{j \in \mathcal{MG}} c_j \times MG_j + \sum_{j \in \mathcal{NG}} c_j \times NG_j \tag{4.49}$$

$$s.t.\ \sum_{j\in\mathcal{MG}} MG_j \times DF_j + \sum_{j\in\mathcal{NG}} NG_j \times DF_j - \sum_{i\in\mathcal{N}} D_i \times DF_i + P_{loss} = 0 \tag{4.50}$$

$$\sum_{j\in\mathcal{MG}} GSF_{k-j} \times MG_j + \sum_{j\in\mathcal{NG}} GSF_{k-j} \times NG_j - \sum_{i\in\mathcal{N}} GSF_{k-i} \times (D_i + E_i) \le F_k^{\max}, \forall k \in \mathcal{B} \tag{4.51}$$

$$MG_j^{\min} \le MG_j \le MG_j^{\max}, \forall j \in \mathcal{MG}. \tag{4.52}$$

It should be stressed that the staircase bidding price, namely, the piece-wise linear cost function, is assumed in the model since it is a common practice in power markets.

4.4.1.2. Characteristic Constraints of the Model

Among the constraints (4.50)-(4.52), the energy balance equation (4.50) is an equality constraint while the line limits in (4.51) and generation output limits in (4.52) are inequality constraints. At the operating point of the CLL, several of the inequalities become equalities. In other words, some unbinding constraints become binding, such as the line limit at the congested lines and the generation output limits for the non-marginal units. Furthermore, when the system load varies within the same interval, or between two adjacent CLLs, as the original operating point, all binding constraints will remain binding. For instance, the non-marginal unit set will remain unchanged, and the congested lines will remain congested. The binding constraints are called “Characteristic Constraints” in this section. The corresponding equalities are written as follows

$$\sum_{j\in\mathcal{MG}} MG_j \times DF_j + \sum_{j\in\mathcal{NG}} NG_j \times DF_j - \sum_{i\in\mathcal{N}} D_i \times DF_i + P_{loss} = 0 \tag{4.53}$$

$$\sum_{j\in\mathcal{MG}} GSF_{k-j} \times MG_j + \sum_{j\in\mathcal{NG}} GSF_{k-j} \times NG_j - \sum_{i\in\mathcal{N}} GSF_{k-i} \times (D_i + E_i) = F_k^{\max}, \forall k \in \mathcal{Bc} \tag{4.54}$$

It should be pointed out that when losses are ignored (i.e., $DF_i=1$, $E_i = 0$), the FND-based DCOPF becomes a lossless DCOPF model, which is a linear programming problem. For a lossless DCOPF model, it can be easily proved that the number of equations in (4.53)-(4.54) equals the number of variables (MG_j, $j \in \mathcal{MG}$), since the number of marginal units is equal to the number of congested lines plus 1, i.e., $N_{MG}=M_{CL}+1$ [42]. Therefore, only one solution exists for equations (4.53)-(4.54) for a lossless DCOPF, as previously discussed in Section 4.2. It implies that the optimal solution can be effectively determined by only the binding constraints as the feasible region constrained by all the binding constraints shrinks down to a single point. Therefore, for the lossless DCOPF, equations (4.53)-(4.54) are essentially equivalent to the lossless DCOPF model in (4.49)-(4.52) with $DF_i=1$ and $E_i = 0$. It should be noted that the cost function does not affect the generation dispatch once the binding constraint set is determined because a staircase constant bidding price is assumed.

The above statements are expected to be true for a FND-based DCOPF in most cases, since the FND-based DCOPF is a variant of the lossless DCOPF and not highly nonlinear due to a low loss percentage in high voltage transmission networks. In addition, equations (4.53)-(4.54) hold true not only at a specific operating point, but also for a range of the system load, as long as there is no change of the binding/unbinding constraint set.

Like in Section 4.2, the participation factor $\boldsymbol{f}$ is used to define the load variation pattern. The equations are correspondingly rewritten as

$$\sum_{j\in\mathcal{MG}} MG_j \times DF_j + \sum_{j\in\mathcal{NG}} NG_j \times DF_j - \sum_{i\in\mathcal{N}} (f_i \times D_{\Sigma} + D_i^{(0)} - f_i \times D_{\Sigma}^{(0)}) \times DF_i + P_{loss} = 0 \tag{4.55}$$

$$\sum_{j\in\mathcal{MG}} GSF_{k-j}\times MG_j + \sum_{j\in\mathcal{NG}} GSF_{k-j}\times NG_j - \sum_{i\in\mathcal{N}} GSF_{k-i}(f_i\times D_\Sigma + D_i^{(0)} - f_i\times D_\Sigma^{(0)} + E_i) = F_k^{\max}, \forall k\in\mathcal{Bc} \quad . \quad (4.56)$$

In operation, the load variation percentage of each bus could be treated as a fixed value in the short term. In planning, the bus load is normally assumed to vary conformingly, which is also assumed to be the case for other studies, such as Continuation Power Flow [8]. Therefore, in this section, f_i is assumed to be a constant. In fact, a more complicated load variation pattern could be modeled by assigning piece-wise constants to the factors, and the following proposed method would still be applicable.

4.4.2. Variable Substitution Method

In equations (4.55)-(4.56), the only independent variable is D_Σ, and dependent variables are ***MG***. It is desirable to solve this set of algebraic equations for the marginal unit generation as a function of the system load. However, it is hard to derive a closed form because P_{loss} and E_i are nonlinear functions of ***MG***.

An alternative approach is to obtain an approximated solution for ***MG***. Since power loss is basically a quadratic function of the load, and is balanced by the marginal unit generation, as seen in the power balance equation, it is reasonable to assume the marginal unit generation follows a quadratic function pattern. In fact, a quadratic pattern is reported to be as a good approximation of the marginal unit generation under the ACOPF dispatch framework in Section 4.3. Therefore, we define a quadratic polynomial function $h_{MG,j}$ to approximate MG_j as follows

$$h_{MG,j}(a_{2,j}, a_{1,j}, a_{0,j}, D_\Sigma) = a_{2,j}\times D_\Sigma^{\,2} + a_{1,j}\times D_\Sigma + a_{0,j}, \forall j\in\mathcal{MG} \quad (4.57)$$

where $a_{2,j}, a_{1,j}, a_{0,j}$ denote the 2nd degree, 1st degree coefficients, and constant part of the j^{th} marginal unit generation, respectively.

Substituting (4.57) into the left-hand side of (4.55) yields a function with respect to D_Σ and $a_{2,j}, a_{1,j}, a_{0,j}$, which is defined as $h_e(a_{2,j}, a_{1,j}, a_{0,j}, \forall j \in \mathcal{MG}, D_\Sigma)$. Likewise, the function found by substituting (4.57) into the left-hand side of (4.56) is defined as $h_{l,k}(a_{2,j}, a_{1,j}, a_{0,j}, \forall j \in \mathcal{MG}, D_\Sigma), \forall k \in \mathcal{Bc}$. By completing this substitution, the variables ***MG*** are replaced with a new set of variables $a_{2,j}, a_{1,j}, a_{0,j}, \forall j \in \mathcal{MG}$. We assume that there exists such a set of $a_{2,j}, a_{1,j}, a_{0,j}$ so that the corresponding approximation of ***MG*** satisfies equations (4.55)-(4.56). It implies

$$h_e(a_{2,j}, a_{1,j}, a_{0,j}, \forall j \in \mathcal{MG}, D_\Sigma) = 0 \tag{4.58}$$

$$h_{l,k}(a_{2,j}, a_{1,j}, a_{0,j}, \forall j \in \mathcal{MG}, D_\Sigma) = F_k^{\max}, \forall k \in \mathcal{Bc}. \tag{4.59}$$

Equations (4.58)-(4.59) are expected to hold true at the given operating point $D_\Sigma^{(0)}$ and any point $D_\Sigma^{(1)}$ which is close to $D_\Sigma^{(0)}$ so that no change of binding constraints occurs. $D_\Sigma^{(1)}$ could be empirically determined. For example, $D_\Sigma^{(1)} = D_\Sigma^{(0)} \times (1 + 0.0025)$. Combined with the marginal generation $\boldsymbol{MG^{(0)}}$ at the given operating point, a set of nonlinear equations are established as follows

$$h_{MG,j}(a_{2,j}, a_{1,j}, a_{0,j}, D_\Sigma^{(0)}) = MG_j^{(0)}, \forall j \in \mathcal{MG} \tag{4.60}$$

$$h_e(a_{2,j}, a_{1,j}, a_{0,j}, \forall j \in \mathcal{MG}, D_\Sigma^{(0)}) = 0 \tag{4.61}$$

$$h_{l,k}(a_{2,j}, a_{1,j}, a_{0,j}, \forall j \in \mathcal{MG}, D_\Sigma^{(0)}) = F_k^{\max}, \forall k \in \mathcal{Bc} \tag{4.62}$$

$$h_e(a_{2,j}, a_{1,j}, a_{0,j}, \forall j \in \mathcal{MG}, D_\Sigma^{(1)}) = 0 \tag{4.63}$$

$$h_{l,k}(a_{2,j}, a_{1,j}, a_{0,j}, \forall j \in \mathcal{MG}, D_\Sigma^{(1)}) = F_k^{\max}, \forall k \in \mathcal{Bc}. \tag{4.64}$$

Equations (4.60)-(4.64) are nonlinear functions of the new variables $a_{2,j}, a_{1,j}, a_{0,j}, \forall j \in \mathcal{MG}$. The number of equations is 3 M_{CL} +3 (=M_{CL} +1+1+ M_{CL} +1+ M_{CL}), which is equal to the number of variables (=3*(M_{CL}+1)). Therefore, the set of equations could be solved by standard nonlinear equation algorithms. It should be noted that M_{CL} is typically a small integer number, and therefore, solving the above nonlinear equations is a small-scale problem.

The complete process of using the variation substitution method to predict congestions and CLLs is as follows:

1) Establish equations (4.60)-(4.64) and solve for coefficients of the approximated marginal unit generation, $a_{2,j}, a_{1,j}, a_{0,j}, \forall j \in \mathcal{MG}$;

2) Solve equations (4.41)-(4.44) for load variations $\Delta D_{\Sigma}\big|_j^{ub}$, $\Delta D_{\Sigma}\big|_j^{lb}$, $\Delta D_{\Sigma}\big|_k^{ub}$ and $\Delta D_{\Sigma}\big|_k^{lb}$;

3) Solve equations (4.45)-(4.48) for the previous CLL ($D_{\Sigma-}$) and the next CLL ($D_{\Sigma+}$);

4.4.3. Case Study of Prediction of Critical Load Level

The proposed method is tested on the PJM 5-bus system [11] with modifications as detailed in Section 4.2. The base case diagram of the system is shown in Figure 4.3. The R/X ratios of the transmission lines are set at 10%.

Figure 4.10 compares the benchmark data of the generation of the marginal unit Sundance from repetitive FND-based DCOPF runs and the corresponding quadratic approximation

results from the variable substitution method, with a load variation of 747 MW to 792 MW and a 5% increase of the base case load. During this load range, Line ED is always congested, and the marginal units are Sundance and Brighton. The differences between the benchmark data and approximation results are shown in Figure 4.10. The benchmark data and the quadratic approximation results of the generation of the marginal unit Brighton, and their differences in percentage are shown in Fig. 4.11. Similar patterns are observed for simulations at other load intervals which do not contain a CLL.

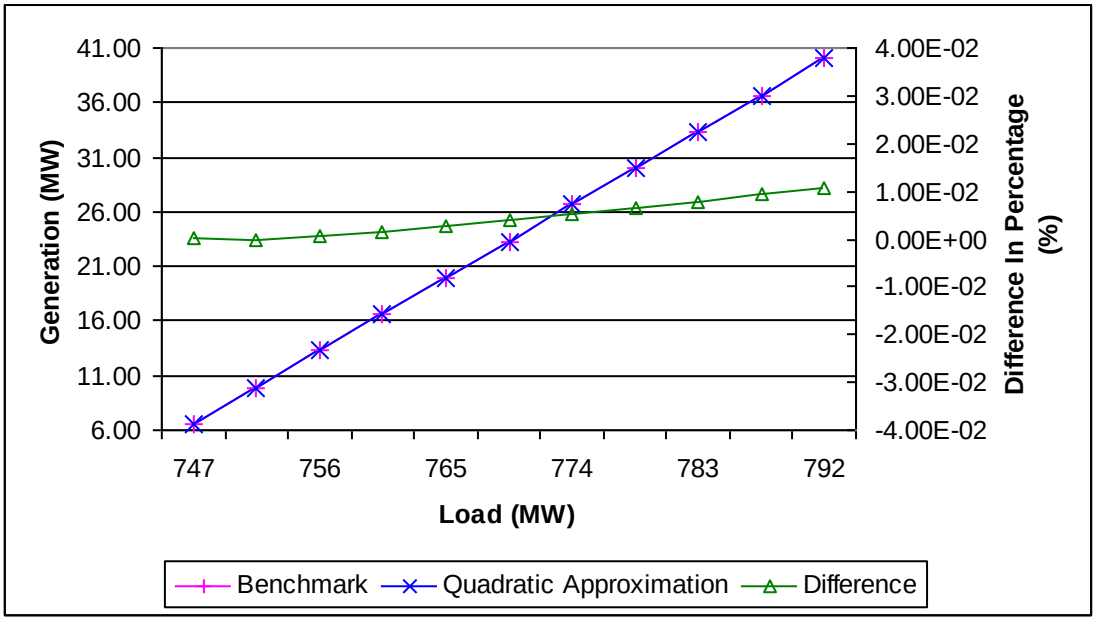


Figure 4.10. Quadratic approximation, the benchmark data, and their differences of the generation of the marginal unit Sundance for the PJM 5-bus system

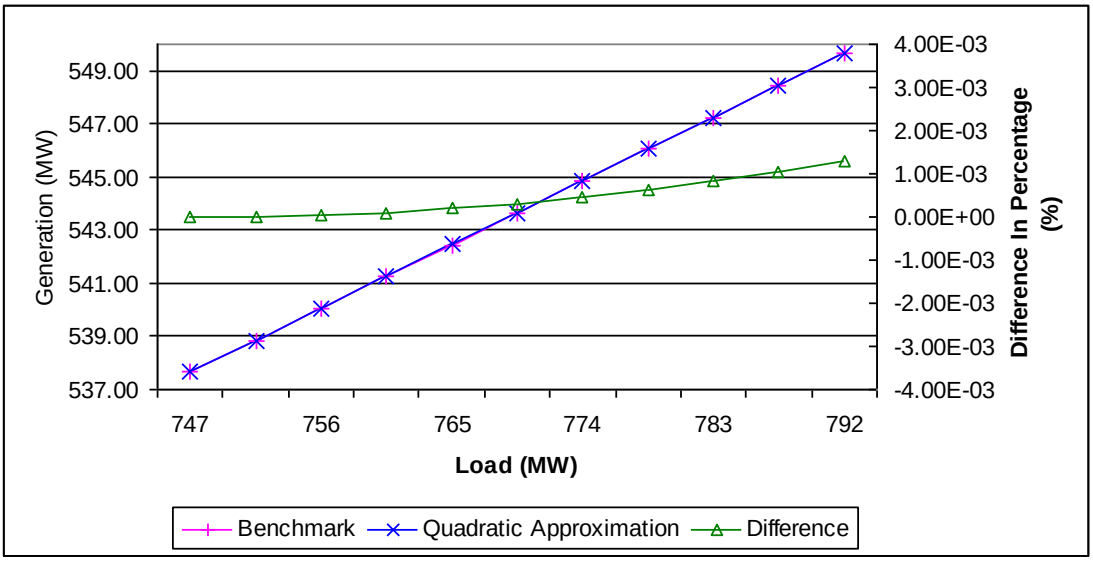


Figure 4.11. Quadratic approximation, the benchmark data, and their differences of the generation of the marginal unit Brighton for the PJM 5-bus system

In Figs. 4.10-4.11, the differences appear to be increasing. Nevertheless, the quadratic approximation is engineering-acceptable, given that the difference between the benchmark data and quadratic approximation results is less than 0.01%, while the load variation spans a 5% difference of the base case load. It should also be noted that although linear approximation is also a choice, it will generate greater differences in general.

The congestion prediction study is performed at the 747 MW load level, namely, 0.83 p.u. of the base-case load. For notational convenience, the generators Alta, Park City, Solitude, Sundance, and Brighton are numbered from 1 through 5, respectively.

Table 4.10 shows the load variation distances calculated using the quadratic coefficients obtained from the variable substitution method. The numbers in bold font in Table 4.10 are actually the load variation margins from the given load level. The predicted previous and next

critical load levels are compared with the actual values obtained from the enumerative FND-based DCOPF simulation, as shown in Table 4.11.

Table 4.10. Load margins from the given operating point 747 MW for the PJM 5-bus system

Load Variation (MW)	Variable Substitution Method
$\Delta D_{\Sigma}\big\vert_{4}^{ub}$	259.65
$\Delta D_{\Sigma}\big\vert_{5}^{ub}$	232.65
$\Delta D_{\Sigma}\big\vert_{4}^{lb}$	**-8.85**
$\Delta D_{\Sigma}\big\vert_{5}^{lb}$	-2127.84
$\Delta D_{\Sigma}\big\vert_{lineAB}^{ub}$	**214.37**
$\Delta D_{\Sigma}\big\vert_{lineAD}^{ub}$	-10573.63
$\Delta D_{\Sigma}\big\vert_{lineAE}^{ub}$	-5254.35
$\Delta D_{\Sigma}\big\vert_{lineBC}^{ub}$	-23059.04
$\Delta D_{\Sigma}\big\vert_{lineCD}^{ub}$	-3304.22
$\Delta D_{\Sigma}\big\vert_{lineAB}^{lb}$	-2322.95
$\Delta D_{\Sigma}\big\vert_{lineAD}^{lb}$	N/A
$\Delta D_{\Sigma}\big\vert_{lineAE}^{lb}$	2577.21
$\Delta D_{\Sigma}\big\vert_{lineBC}^{lb}$	N/A
$\Delta D_{\Sigma}\big\vert_{lineCD}^{lb}$	2402.32

Note: N/A represents no solution for Equations (4.41)-(4.44).

Table 4.11. Previous and next critical load levels from the given operating point 747 MW for the PJM 5-bus system

CLL (MW)	Variable Substitution Method	From Actual Enumerative Simulation	New Binding Constraint
$D_{\Sigma+}$	214.37	212.49	Line AB
$D_{\Sigma-}$	-8.85	-8.14	Generator Sundance

The results in Tables 4.10-4.11 indicate that Line AB will reach its limit in the positive direction if the system load increases from the given load level by 214.37 MW, which is very close to the actual value of 212.49MW. If this occurs, the constraint of Line AB will become binding and the system will contain one additional congestion point. This is the first change of binding constraints with a load increase. In the case of a load decrease, the first change of binding constraints is expected to occur when the load decreases by 8.85 MW, as opposed to the actual value 8.14MW, at which point the marginal unit Sundance at Bus 4 will reach its lower limit and is no longer a marginal unit. At this load level, the Park City unit will become the marginal unit and Line ED will remain congested.

Tables 4.10-4.11 show that the predicated results obtained from the proposed variable substitution method are acceptable when compared with the benchmark results obtained from the enumerative FND-based DCOPF runs, while the proposed method involves a lower computational effort and complexity.

For any given operating point, we can obtain an estimated previous CLL and next CLL by applying the variable substitution method. Table 4.12 shows the estimated previous and next CLLs for different given load levels. For each CLL, it can be estimated from the left and from the right, and the results may differ. In this regard, the estimated CLLs in Table 4.12 are rearranged, as shown in Table 4.13, such that each actual CLL can be conveniently compared with the estimated values from the left and right. Consequently, the differences between the estimated CLLs and the actual CLLs are shown in Table 4.14. The maximum error is approximately 0.008 p.u. of the base-case load, or 7.2 MW, when estimating the CLL as 0.7873 p.u. from the left. The results in Table 4.13 and Table 4.14 suggest that the proposed variable substitution method can provide results with acceptable accuracy.

Table 4.12. Previous and next critical load levels from various of given operating points for the PJM 5-bus system

Given Load Level (MW)	$D_{\Sigma+}$ (MW)	$D_{\Sigma-}$ (MW)
450	594.36	N/A
603	634.14	594.72
648	701.37	634.14
711	738.18	708.57
765	961.38	738.18
990	1126.44	959.49
1260	N/A	1122.66

Note: N/A represents undefined or meaningless CLL.

Table 4.13. Actual CLLs and estimated CLLs for the PJM 5-bus system

Actual CLLs (MW)	Estimated CLLs (MW)	
	Estimated From Left	Estimated From Right
594.63	594.36	594.72
634.14	634.14	634.14
708.57	701.37	708.57
738.09	738.18	738.18
959.49	961.38	959.49
1126.35	1126.44	1122.66

Table 4.14. Differences between the estimated CLLs and actual CLLs for the PJM 5-bus system

Actual CLLs (MW)	Error of Estimated CLLs (MW)	
	Estimated From Left	Estimated From Right
594.63	-0.27	0.09
634.14	0	0
708.57	-7.11	0.09
738.09	0.09	0.09
959.49	1.89	0
1126.35	0.09	-3.69

4.4.4. Discussion and Conclusions

When the load is in the range of the two adjacent critical load levels (CLLs), the FND-based DCOPF model could be essentially represented by the binding constraints. By introducing the load variation participation factor, the equations for the binding constraints can be written as a set of implicit functions of the marginal unit generation with respect to only one variable, namely, the system load. With the intuition and impression gained from Section 4.3 that shows the marginal unit generations can be well approximated by quadratic polynomials, we substitute the quadratic polynomials for the marginal unit generation in those functions; which are correspondingly transformed into functions of the quadratic coefficients with respect to the load and can be easily solved due to their small scale. The quadratic coefficients can be consequently used to perform congestion prediction and estimate the previous and next CLLs. A case study on the PJM 5-bus system demonstrates the applicability of the proposed method.

As indicated by the case study results, the variable substitution method does not yield results as accurate as the interpolation method. The reason for this is that the interpolation method utilizes OPF solutions at a couple of other load levels in addition to the given operating point; whereas for the variable substitution method only the OPF solution at the given operating point is available. Hence, the input to the interpolation method contains more information about the future when the load varies. Despite providing less accuracy, the variable substitution method can produce engineering-acceptable results with less computations and an easier implementation.

4.5. Conclusions

In previous chapters we have observed the step change phenomenon of the LMP versus load curve, the next step is to effectively and efficiently locate the step changes, which are called critical load levels (CLLs).

The study is firstly performed on lossless DCOPF model. The linear characteristics of the lossless DCOPF enable the use of a simplex-like analytical method to quickly calculate the CLLs. No additional OPF runs are needed, and the performance is shown to be much superior to a binary search method.

ACOPF incorporates power losses and is a nonlinear programming model, which makes it hard to perform analytical study. Therefore, a simulation based approach is adopted. The marginal generation and line flow are numerically shown to follow perfectly quadratic polynomial patterns. Then, a quadratic interpolation method is proposed to help reduce the computation efforts with repetitive ACOPF runs. The proposed approach typically requires only two additional ACOPF runs and gives highly accurate estimation of CLLs.

In order to further reduce the computation with the additional OPF runs, a variable substitution method is presented for FND-based DCOPF model. The characteristic constraints are defined and used to form a set of nonlinear equations, which can be easily solved with much less computation. Nevertheless, the drawback is the loss of high accuracy due to limited information on load levels other than current operating point.

The proposed methods for the various OPF models present efficient calculations for the LMP versus Load curve. The curve may be used to predict price spikes, given the forecasted Load versus time curve. Another application is for the quick estimation of a new dispatch and LMP. When the forecasted load for the next interval and the current load level are within the same two adjacent critical load levels, the new dispatch and LMP at the forecasted load could

be quickly obtained by looking up the LMP versus Load curve/table and applying the generation sensitivities. There is no need to re-run the optimization solver repetitively when the load keeps changing back and force within the two adjacent CLLs. This efficient application could be useful in both operation and planning.

5 Probabilistic LMP Forecasting Under Load Uncertainty

5.1. Chapter Introduction

The prediction of the load, especially a short-term load consumption, has long been an important topic in academia and industrial research and practices [43]. With the deregulation of the power industry and the adoption of the locational marginal pricing (LMP) methodology, LMP forecasting has garnered attention because of the significance of the LMP in delivering market price signals and its use for settlements [44, 45, 46].

It is known that the LMP can be decomposed into three components, with each representing the marginal energy price, loss price, and congestion price, respectively [36, 20]. The decisive factors of the LMP include supply bids, demand offers, load forecasting, and network topology. In the day-ahead power market, once the market is closed (for example at 12:00 noon before the operating day), the offers and bids are fixed and a transmission network model will be used for day-ahead market scheduling. Nevertheless, the load remains uncertain as there is essentially no way to discover the exact load of each hour of the next operating day. Load forecasting is applied to address this issue, but performance varies with models, algorithms, and the nature of the problem. It is apparent that the uncertainty associated with the load directly leads to the uncertainty of the LMP. Therefore, as equally important as the study of the other economic impacts of load forecasting [50, 52, 53, 54], it is necessary to investigate how the LMP will be affected by the uncertainty of load, or, the uncertainty of load forecasting results in practice.

The Optimal Power Flow (OPF) problem has been discussed in [63] with special attention to the computational issues created by deregulation. A methodology of computing LMP sensitivities with respect to the load in the AC Optimal Power Flow (ACOPF) framework has

been presented in [41]. Approaches for the DC Optimal Power Flow (DCOPF) have been applied in Section 4.2. In Section 4.2, a perturbation-based algorithm is proposed to identify the next critical load level (CLL), defined as the load level at which a LMP step change, as well as the change of binding and unbinding limits, occurs. This algorithm is demonstrated to be very computationally efficient since it does not require multiple optimization runs. It essentially enables the efficient study of the LMP over any range of load variations.

Figure 4.1 is redrawn as Figure 5.1 for quick reference. It shows a typical LMP versus load curve for a sample system slightly modified from the original PJM 5-bus system defined in [11]. Losses are ignored in these studies so that this research is concentrated on the overall behavior of the LMP due to congestion. In theory, the horizontal axis denotes the actual load. In practice, the axis represents the forecasted load, since the forecasted load is utilized to perform the dispatch and LMP calculations.

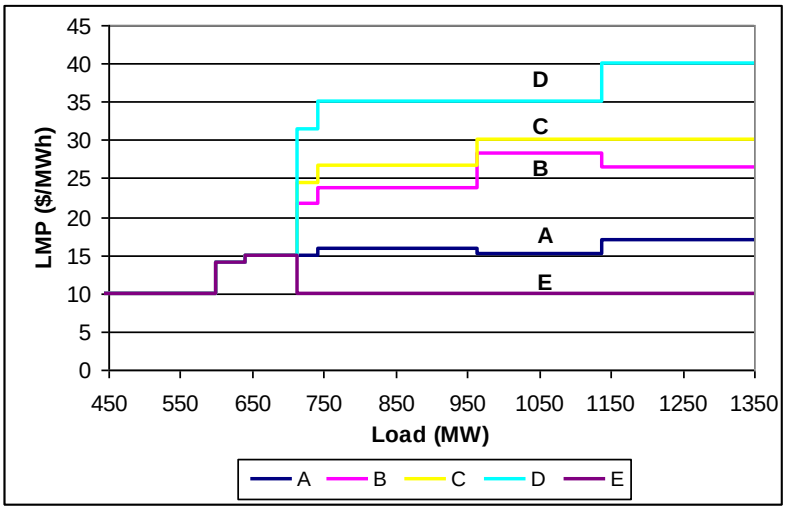


Figure 5.1. LMP at all buses with respect to the different system loads for the modified PJM 5-bus system

It can be seen that there is a step change of the LMP when the load increases to a CLL, e.g., the load level at 600 MW, 640MW, 711.81MW, etc. At each new CLL, a new binding limit, either a transmission line thermal limit or a generator capacity limit, occurs. Meanwhile, there is a change of the marginal unit set and marginal generation sensitivity with respect to the load, which results in the LMP step change.

At the CLLs, the LMP is highly sensitive. The sensitivity of the LMP with respect to load is evaluated as mathematically infinite. This step change characteristic of the LMP leads to the ambiguity of the LMP evaluation at the CLLs. For example, when the forecasted load happens to be 711.81MW, there are at least two choices to set the price at Bus D: \$15.00/MWh or \$31.46/MWh. However, which price should be chosen is not justified. By taking into account the load variation direction, either an increase or decrease, an option for the price may be produced.

A more important impact of this step change is that the slight difference in the forecasted load may result in a dramatic difference in the LMP. For example, the LMP at Bus D is \$15.00/MWh when the forecasted load is 711.80MW, whereas the price soars to \$31.46/MWh when the forecasted load is slightly off by 0.1MW, making it 711.90MW. On the other hand, it is very likely that a load forecasting tool, even well-tuned, will produce a result with an error greater than 0.1MW for a target load at approximately the 700MW level. Therefore, the load forecasting uncertainty may significantly affect LMP forecasting and consequently, market participants' financial or bidding decisions.

A few reasons exist that lead to the uncertainty of the load forecasting result. First, the future load is a random variable indeed and cannot be accurately predicted. Each load forecasting method has its own theoretical foundations and will likely produce results that differ from the other method's results. Each method may excel in certain applications, but no one method can achieve 100% accuracy. There is always a certain error range associated with

the forecasted results as well. Second, even for the same method, the results may be different by using different settings, tunings, and assumptions. Last, the majority of methods suffers from missing data and relies heavily on the accuracy of the input data.

Although the uncertainty in the load forecast is unavoidable, load forecasting errors are often described by certain probability distributions, which enable the study of the correlation between the forecasted load and LMP in a probabilistic sense. The LMP study, considering probabilistic factors, has been presented in [56, 57] by modeling the generator biddings as stochastic variables. However, no existing research work exists to specifically investigate the impact of the load forecast uncertainty on the LMP simulation results for a price forecast purpose. We intend to reveal the probabilistic aspect of the traditional LMP with respect to load uncertainty and present useful information such as the likelihood that a forecasted deterministic LMP will occur. This assists generation companies or load serving entities to formulate their bidding strategies, risk hedging policies, and even long-term contract negotiations. More importantly, this work systematically presents the concept of a probabilistic LMP from the viewpoint of forecasting, and indicates that the forecasted probabilistic LMP should be a set of discrete values with the associated probabilities at different load intervals. These two aspects are the motivation and significance of this work.

As in Chapter 4, the study will be conducted for the lossless DCOPF, ACOPF, and the proposed FND-based DCOPF, respectively, due to the different price patterns for these models.

5.2. Probabilistic LMP Forecasting for Lossless DCOPF Framework

5.2.1. Probabilistic LMP and its Probability Mass Function

5.2.1.1. Assumptions

The actual load, or load forecasting error, can be assumed to be a random variable and will follow a certain probability distribution. However, it is difficult to determine the distribution type due to insufficient historical data [53]. A normal distribution is frequently used and has been employed to model the actual load in a number of research works [50, 52, 53, 54], and therefore, will be used in this work to describe the actual load at hour *t*. Then, we have

$$D_t \sim N(\mu_t, \sigma_t^2) \tag{5.1}$$

$$\varphi(x) = \frac{1}{\sigma_t \sqrt{2\pi}} e^{-\frac{(x-\mu_t)^2}{2\sigma_t^2}} \tag{5.2}$$

$$\Phi(x) = \int_{-\infty}^{x} \varphi(u) du \tag{5.3}$$

where

D_t = a random variable for the actual load at hour *t*;

N denotes the normal distribution;

μ_t = mean of D_t;

σ_t^2 = variance of D_t;

$\varphi(x)$ = probability density function of D_t;

$\Phi(x)$ = cumulative density function of D_t.

It should be emphasized that for a well-tuned load forecasting model, the forecasted load at hour *t*, D_t^F, should be very close to the mean value of D_t, i.e., μ_t. However, D_t^F is also not precisely equal to μ_t due to the error of the load forecasting. Nevertheless, it is a common

practice in a market simulation or forecast to use a single forecasted value of load (D_t^F) to perform a deterministic market simulation to forecast the LMP, congestion, etc [5]. In addition, there is no reported research work to incorporate the load forecasting uncertainty into a LMP simulation and price forecasting and to study its effect. This explains the novelty of this research work.

5.2.1.2. Models for the LMP-Load Curve

As the value of the random variable D_t is represented by the horizontal axis, the LMP-Load curve (similar to the curves in Fig. 5.1) is illustratively redrawn in Fig. 5.2 to facilitate the following study.

As shown in Fig. 5.2, the load axis is divided into *n-1* segments by a sequence of critical load levels (CLLs), $\{D_i\}_{i=1}^{n}$. Here D_1 represents the no-load case (i.e., D_1=0), and D_n represents the maximum load that the system can supply due to the limits of total generation resources and transmission capabilities. Associated with each load segment *i*, is a corresponding actual LMP value, p_i , which is considered a constant in this study, as we ignore the loss model for simplicity. The model to calculate the LMP without losses was previously discussed in Section 3.2.1.

The LMP-Load curve in Figure 5.2 includes two extra segments. One is for the load from D_0 to D_1 , where D_0 denotes a negative infinite load, and the associated price, p_0 , is zero. The second additional segment is defined as the load range from D_n to D_{n+1} , where D_{n+1} represents the positive infinite load. In this segment, the price is set to be the Value of the Lost Load (VOLL) to reflect the demand response to the load shedding. Although the VOLL varies with customer groups and load interruption time and duration, it is a common practice

to produce an aggregated VOLL to represent the average loss for an area. Therefore, the VOLL is assumed to be a constant value for simplicity in this work.

The two extra segments are added for mathematical completeness. In fact, they have a minimal, if any, impact on the study because the typical load range under study (for instance, from 0.8 p.u. to 1.2 p.u. of an average case load) is far from these two extreme segments; and the possibility of having the forecasted load close to zero or greater than D_n , the maximum load that the system can supply, is extremely rare and numerically zero.

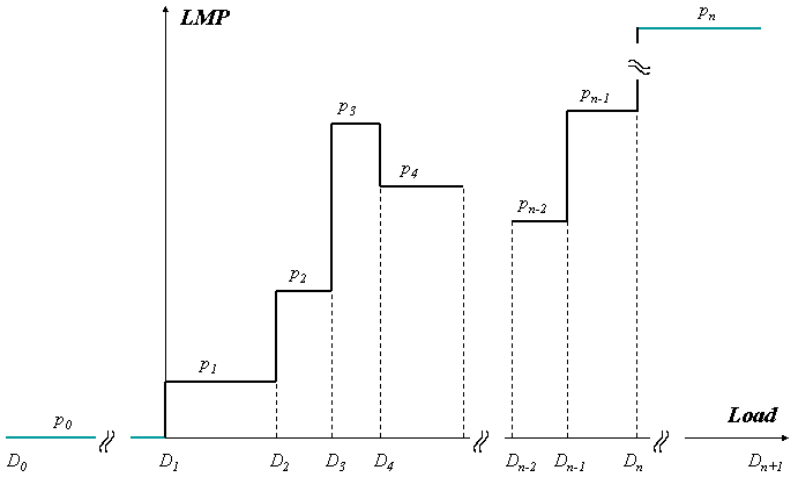


Figure 5.2. Extended LMP versus Load Curve

The curve can be formulated as

$$LMP(D) = \begin{cases} p_0, & D_0 < D \le D_1 \\ p_1, & D_1 < D \le D_2 \\ \vdots & \\ p_{n-1}, & D_{n-1} < D \le D_n \\ p_n, & D_n < D \le D_{n+1} \end{cases} \tag{5.4}$$

where

$p_0 = 0$

$p_n = VOLL$

$D_0 = -\infty$

$D_1 = 0$

$D_{n+1} = \infty$.

The compact representation is given as follows

$$LMP(D) = \{ p_i \mid i \in \{0,1,\cdots,n\}, D_i < D \le D_{i+1} \} \ . \tag{5.5}$$

Apparently, $\frac{\partial LMP}{\partial D}$, the LMP sensitivity with respect to the load, is infinite at the critical load levels (CLLs), $\{D_i\}_{i=1}^{n-1}$.

5.2.1.3. Probabilistic LMP and its Probability Mass Function

Here it is assumed that the economic dispatch and LMP calculation are performed on an hourly basis. The LMP at hour *t*, denoted by LMP_t, is a function of D_t which is a random variable, from the viewpoint of forecasting. Therefore, at the forecasting or planning stage, LMP_t should also be viewed as a random variable. This characteristic is inherited from forecasted load. Fig. 5.3 shows the LMP-Load curve and probability distribution function

(PDF) of D_t. It can be inferred from Fig. 5.3 that LMP_t should be a discrete random variable, with *n+1* possible values denoted by the sequence $\{p_i\}_{i=0}^{n}$. Certainly, the probability that the actual price is aligned with a different p_i will vary. The random variable LMP_t is named the *Probabilistic LMP* in this work in order to differentiate it from the traditionally deterministic LMP.

Furthermore, the probability that the LMP_t has an actual value of p_i can be expressed as

$$\Pr(LMP_t = p_i) = \int_{D_i}^{D_{i+1}} \varphi(u)du = \Phi(D_{i+1}) - \Phi(D_i)\,. \qquad (5.6)$$

The cumulative density function $\Phi(x)$ can be well estimated by the available approximation methods to ease the computation [9]. A schematic graph of the probability mass function (PMF) of LMP_t is shown in Fig. 5.4. Note that mathematically the PMF graph is usually presented in a way so that the possible values are sorted in ascending order, and the probability of identical prices (for example when $p_i = p_j$ where $i \neq j$) are merged together. However, this is not done in Fig. 5.4 for the purpose of easy presentation. That is, Fig. 5.4 shows the PMFs in the order of the occurrence of the associated price, p_j, as the load increases. It is apparent the probabilities of all possible prices should add up to 1.0.

Eqs. (5.5)-(5.6) and Fig. 5.4 show the important characteristics of the concept of the probabilistic LMP proposed in this work:

The Probabilistic LMP at a specific (mean) load level is not a single deterministic value. Instead, it represents a set of discrete values at a number of load intervals. Each value has an associated probability.

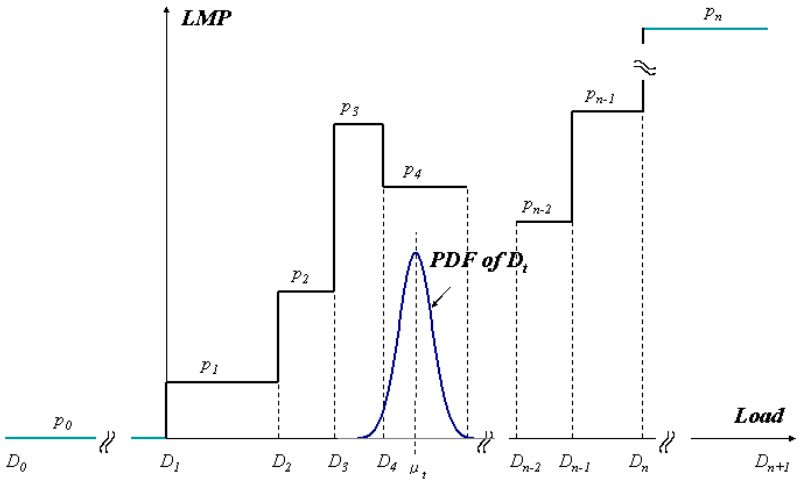


Figure 5.3. LMP-Load curve and probability distribution of D_t

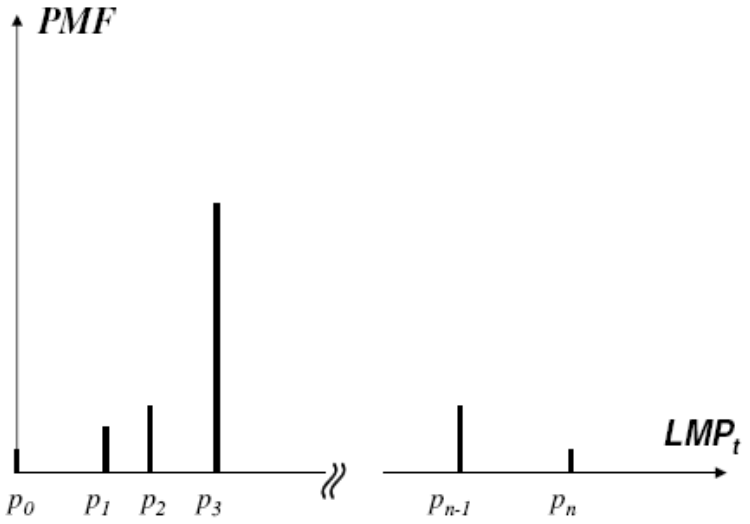


Figure 5.4. Probability Mass Function of the Probabilistic LMP at hour *t*

5.2.1.4. Alignment Probability of Deterministic LMP Forecasting versus Forecasted Load Curve

At hour *t*, if a single value of the forecasted load D_t^F is used for LMP forecasting, the calculated LMP can be deterministically identified by looking up the LMP-Load curve, as

shown in Fig. 5.2. Suppose D_t^F is between D_j and D_{j+1}, then the corresponding $LMP(D_t^F)$ is equal to p_j. This can be written as

$$LMP(D_t^F) = p_j, \qquad D_j < D_t^F \le D_{j+1} \tag{5.7}$$

where

$LMP(D_t^F)$ is the LMP corresponding to the forecasted load D_t^F. This can be called the deterministic LMP forecast.

Then, the probability that the actual price LMP_t is the same as $LMP(D_t^F)$, i.e., p_j, can be obtained from the probability mass function (PMF), as shown in Fig. 5.4. It should be noted that the actual load may not be D_t^F, or even in the range of [D_j, D_{j+1}]. Hence, the actual price is not always the same as the forecasted price. In a rigorous way, we define an alignment probability that the actual price is the same as forecasted price in deterministic LMP forecasting. This can be written as

$$\begin{aligned} \text{AP} &= \Pr\left(LMP_t = LMP(D_t^F)\right) = \Pr\left(LMP_t = p_j\right) \\ &= \int_{D_j}^{D_{j+1}} \varphi(u)du = \Phi(D_{j+1}) - \Phi(D_j) \end{aligned} \tag{5.8}$$

where AP or $\Pr(LMP_t = p_j)$ is the alignment probability in deterministic LMP forecasting. Apparently, the probability that the actual price is not the same as the forecasted price in a deterministic approach is equal to $1 - \Pr(LMP_t = p_j)$, generally.

When the above equation is evaluated for every D_t^F in the entire interval [D_1, D_n], an alignment probability versus D_t^F curve can be obtained. Each point of the curve represents the alignment probability that the projected $LMP(D_t^F)$ is the actual price when the load is D_t^F. When combined with the LMP-Load curve, this LMP alignment probability versus

forecasted load curve delivers very useful information such as how likely the projected LMP, $LMP(D_t^F)$ at the forecasting stage, is the same as the actual LMP at hour *t*, LMP_t.

The alignment probability defined in Eq. (5.8) provides the probability that the deterministically forecasted LMP and the actual LMP are exactly the same. This may not be appropriate if the actual LMP has a good chance, such as 70%, to provide a price that is very close to, but not exactly the same as, the deterministic LMP result. With Eq. (5.8), the two are not considered aligned, and the alignment probability will be as low as 30%. Hence, the tolerance level can be applied to address this. For instance, if we choose 10% as the price tolerance level, then the result of the deterministically forecasted LMP is considered aligned with the actual LMP, if the actual LMP is within [90%, 110%] of the deterministic LMP. Namely, we can define the alignment probability with tolerance, AP_α as

$$\mathbf{AP}_\alpha = \mathbf{Pr}\left\{p_j \times (1-\alpha\%) \le LMP_t \le p_j \times (1+\alpha\%)\right\} \tag{5.9}$$

where α is the tolerance percentage. This gives the confidence of having the LMP forecast within an acceptable range.

The above discussion will be further detailed in the numerical studies.

5.2.2. Expected Value of the Probabilistic LMP

5.2.2.1. Expected Value of Probabilistic LMP

Since LMP at hour *t*, LMP_t, is a random variable, it is interesting to see the expected value of LMP at hour *t*

$$
\begin{aligned}
E(LMP_t) &= \sum_{i=0}^{n} \Pr(LMP_t = p_i) \times p_i \\
&= \sum_{i=0}^{n} \left(\int_{D_i}^{D_{i+1}} \varphi(u)du \right) \times p_i \\
&= \sum_{i=0}^{n} p_i \times \int_{D_i}^{D_{i+1}} \frac{1}{\sigma_t \sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du
\end{aligned}
\tag{5.10}
$$

where $E(\cdot)$ is the expected value operator.

It can be seen that $E(LMP_t)$ is a function of μ_t and σ_t. The function is defined as

$$
E_{LMP_t}(\mu_t, \sigma_t) \overset{\Delta}{=} E(LMP_t) \ . \tag{5.11}
$$

5.2.2.2. Expected Value of Probabilistic LMP versus Forecasted Load Curve

If $E_{LMP_t}(\mu_t, \sigma_t)$ is evaluated for every μ_t in the interval $[D_1, D_n]$, the expected value of the probabilistic LMP versus μ_t curve will be obtained.

In practice, it is sometimes more interesting to see the expected value of probabilistic LMP with respect to the forecasted load (D_t^F) curve since D_t^F, instead of μ_t, is actually available. If there is a constant deviation C_{dev} of D_t^F from μ_t (for instance, due to model calibration error), namely, $D_t^F = \mu_t - C_{dev}$, then $E_{LMP_t}(D_t^F) = E_{LMP_t}(\mu_t - C_{dev})$. This implies that the expected value of the probabilistic LMP versus D_t^F curve can be obtained by left-shifting the expected value of the probabilistic LMP versus μ_t curve by C_{dev}. Likewise, if the deviation of D_t^F from μ_t is the constant portion r of μ_t, namely, $D_t^F = \mu_t(1-r)$, then $E_{LMP_t}(D_t^F) = E_{LMP_t}(\mu_t(1-r))$. This indicates that the expected value of the probabilistic LMP versus D_t^F curve can be obtained by laterally scaling the expected value of the probabilistic LMP versus μ_t curve by a factor of $(1-r)$ along the μ_t axis. Therefore, in both cases, the shape of the expected value of the probabilistic LMP versus D_t^F curve is similar to, if not the

same as, that of the expected value of the probabilistic LMP versus μ_t curve. In general, one curve can be obtained by performing a simple geometrical operation on the other curve.

Furthermore, despite the possible differences between D_t^F and μ_t, it can be normally assumed that the forecasted load D_t^F is equal to μ_t as previously mentioned. Consequently, we have $E_{LMP_t}(D_t^F) = E_{LMP_t}(\mu_t)$. For notational convenience, this assumption is used in the following study. Therefore, $E_{LMP_t}(\mu_t)$ and $E_{LMP_t}(D_t^F)$ are interchangeable, and so are μ_t and D_t^F.

It should be noted that $E_{LMP_t}(\mu_t, \sigma_t)$ is continuously differentiable at μ_t. Therefore, the sensitivity of the expected value of the probabilistic LMP with respect to μ_t can be derived using the theory of parametric derivative of integration, which is stated as follows:

If $f(x,y), f_x(x,y)$ are continuous on $[a,b]\times[c,d]$, then the derivative of $I(x) = \int_c^d f(x,y)dy$ is continuous, and $I'(x) = \int_c^d f_x(x,y)dy$.

By this theory, the sensitivity of probabilistic LMP is derived as

$$\begin{aligned}
&\frac{\partial E_{LMP_t}(\mu_t, \sigma_t)}{\partial \mu_t} \\
&= \sum_{i=0}^{n} p_i \times \int_{D_i}^{D_{i+1}} \left(\frac{1}{\sigma_t \sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} \right) \times \frac{2(u-\mu_t)}{2\sigma_t^2} du \quad . \qquad (5.12) \\
&= \sum_{i=0}^{n} p_i \times \int_{D_i}^{D_{i+1}} \frac{u-\mu_t}{\sigma_t^3 \sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du
\end{aligned}$$

Eq. (5.12) can be further derived as (5.13). Details of the derivation are included in Appendix B.

$$\frac{\partial E_{LMP_t}(\mu_t,\sigma_t)}{\partial \mu_t} = \frac{1}{\sigma_t\sqrt{2\pi}} \times \left[\sum_{i=1}^{n-1} p_i \times \left(e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}} - e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}} \right) + p_n \times e^{-\frac{(D_n-\mu_t)^2}{2\sigma_t^2}} \right] . \quad (5.13)$$

5.2.2.3. Lower and Upper Bound of the Sensitivity of the Expected Value of Probabilistic LMP

The absolute value of the sensitivity of the expected value of the probabilistic LMP has an upper bound as follows. The derivation is given in Appendix C.

$$\left| \frac{\partial E_{LMP_t}(\mu_t,\sigma_t)}{\partial \mu_t} \right| \le \frac{1}{\sigma_t\sqrt{2\pi}} \sum_{i=1}^{n} |p_i| . \quad (5.14)$$

Therefore, $\frac{\partial E_{LMP_t}(\mu_t,\sigma_t)}{\partial \mu_t}$ has a finite lower and upper bound for a given non-zero, σ_t

$$-\frac{1}{\sigma_t\sqrt{2\pi}} \sum_{i=1}^{n} |p_i| \le \frac{\partial E_{LMP_t}(\mu_t,\sigma_t)}{\partial \mu_t} \le \frac{1}{\sigma_t\sqrt{2\pi}} \sum_{i=1}^{n} |p_i| . \quad (5.15)$$

Equation (5.15) implies that the upper bound increases when the standard deviation becomes smaller. An extreme case is when the load forecast is completely accurate, namely, when σ_t is zero then the upper bound is infinite. This means that the step change may occur in this particular situation. This pattern will be exemplified in the numerical studies.

5.2.2.4. Approximate Calculation of the Expected Value of the Probabilistic LMP

The calculation of the expected value of the probabilistic LMP involves complicated mathematical integration; however, it can be simplified by applying certain approximations for particular cases.

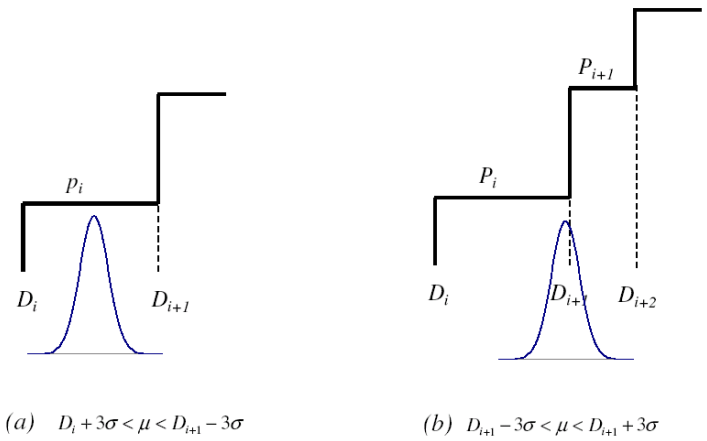


Figure 5.5. Two cases of the approximated calculation of the expected value of the probabilistic LMP

For a normal distribution with mean μ and standard deviation σ, the ratio of the probability in $[\mu - 3\sigma, \mu + 3\sigma]$ and the probability outside this interval is

$$\frac{\Pr(\mu - 3\sigma < x < \mu + 3\sigma)}{\Pr(x < \mu - 3\sigma) + \Pr(x > \mu + 3\sigma)} \approx 384 \,. \tag{5.16}$$

Therefore, if it is satisfied that $\frac{\max(p_i)}{\min(p_i) \mid p_i > 0} < 38.4$, then the calculation of the expected value of the probabilistic LMP under certain conditions, as shown in Fig. 5.5, can be approximated as follows

(a) if $D_i + 3\sigma_t < \mu_t < D_{i+1} - 3\sigma_t$, then

$$E(LMP_t) \approx p_i$$

(b) if $D_i + 3\sigma_t < \mu_t < D_{i+2} - 3\sigma_t$, then

$$E(LMP_t) \approx a \times p_i + b \times p_{i+1} \approx a \times p_i + (1 - a) p_{i+1}$$

where $a = \Phi(D_{i+1}) - \Phi(D_i) \approx \Phi(D_{i+1})$

$b = \Phi(D_{i+2}) - \Phi(D_{i+1}) \approx 1 - \Phi(D_{i+1}) = 1 - a$.

These approximations will be close enough to the actual values and help to lower computational efforts for the cases with a broad range between CLLs. For example in Fig. 5.1, the load in the range of (0MW, 570MW) falls into case (a) category, and the expected value of the Probabilistic LMP could be easily calculated, which is essentially $10/MWh throughout this interval.

Moreover, considering the fact that well-tuned commercial load forecasting tools are capable of generating only a fairly small amount of error, an analyst could focus on a narrow range around the forecasted load, such as 2 or 3 standard deviations. This could overcome the increases in computation for large market areas where many more LMP-load segments will appear due to multiple bid segments, more generators, and potentially more congested transmission lines.

5.2.3. Numerical Study of a Modified PJM 5-Bus System

In this section, a numeric study will be performed on the PJM 5-Bus system [11], with slight modifications. The modifications are for illustrative purposes and were detailed in Section 4.2.4. The configuration of the system is shown in Figure 4.3, and redrawn here in Figure 5.6 for a quick reference.

To calculate the LMP versus load curve as shown in Fig. 5.1, it is assumed, for simplicity, that the system load change is distributed to each bus load proportional to its base case load. Therefore, the load change is equally distributed at Buses B, C, and D since each has a 300 MW load in the base case. This is approximately reasonable because the proportional distribution from the area load to the bus load is commonly used in industrial practices in planning, at least for conforming loads. The proportional distribution is used in the continuation power flow for voltage stability studies. Note that the distribution pattern of the system load variation could be modeled in a more sophisticated way (see Section 4.2). Since

this work aims to illustrate the concept of a probabilistic LMP considering load uncertainty, we use the proportional variation pattern for simplicity. A more complicated model, such as when considering conforming and non-conforming loads, can be addressed in future works.

The critical load levels (CLLs) and the corresponding LMPs at each bus are shown in Table 5.1. This data is the data source for Fig. 5.1 and is calculated by the efficient solver presented in Section 4.2.

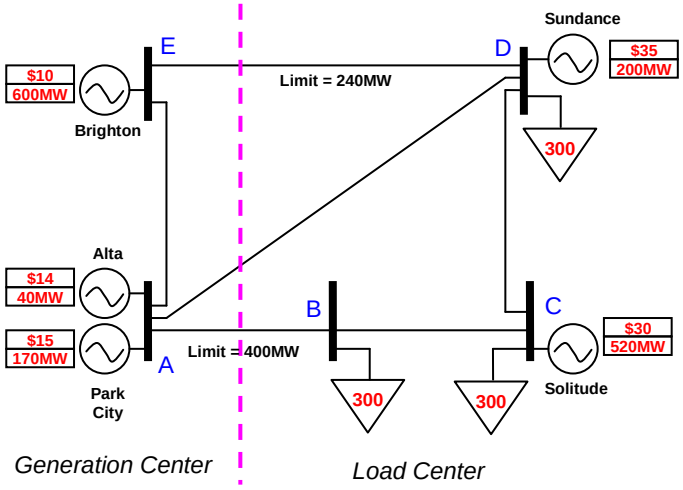


Figure 5.6. The Base Case Modified from the PJM Five-Bus System

Table 5.1. CLLs and LMPs

CLL(MW)	LMP@A	LMP@B	LMP@C	LMP@D	LMP@E
0.00	10.00	10.00	10.00	10.00	10.00
600.00	14.00	14.00	14.00	14.00	14.00
640.00	15.00	15.00	15.00	15.00	15.00
711.81	15.00	21.74	24.33	31.46	10.00
742.80	15.83	23.68	26.70	35.00	10.00
963.94	15.24	28.18	30.00	35.00	10.00
1137.02	16.98	26.38	30.00	39.94	10.00
1484.06	16.98	26.38	30.00	39.94	10.00

Note: LMPs are all in units of $/MWh; Prices in the gray boxes show the LMP at Bus B decreases when the load increases.

For simplicity and better illustration, it is assumed that μ_t is always equal to the forecasted load D_t^F, and the standard deviation σ_t is taken as 5% of the mean μ_t. The VOLL is set at $2000/MWh, which is reasonable, as the typical range of the VOLL is between $2000/MWh and $50,000/MWh [64].

5.2.3.1. Probability Mass Function of Probabilistic LMP

The probability mass function of LMP_t at Bus B at two representing forecasted load levels, 730MW and 900MW, is calculated and shown in Table 5.2. The same results are presented as a pie chart in Fig. 5.7. From the results it was discovered that the deterministic LMP with respect to D_t^F may or may not be the price with the highest probability. For example, when the forecasted load is 900MW, the corresponding deterministic LMP is $23.68/MWh and has the highest probability of 92.21%. However, the deterministic LMP $21.74/MWh for the

forecasted load 730MW has only the second highest probability of 32.80%, less than the probability of 36.29% for $23.68/MWh. This shows that the deterministic LMP associated with the mean value of the actual load does not necessarily bear the biggest probability. It should be noted that the price $28.18/MWh is listed before $26.38/MWh simply because this is the trend of price at Bus B when the load grows. This is also shown within the gray boxes in Table 5.1.

Table 5.2. PMF of the LMP_t for Bus B

LMP($/MWh)	Probability(%) when D^F_t=730MW	Probability(%) when D^F_t =900MW
0	0.00	0.00
10	0.02	0.00
14	0.67	0.00
15	30.23	0.00
21.74	32.80	0.02
23.68	36.29	92.21
28.18	0.00	7.77
26.38	0.00	0.00
2000	0.00	0.00
Total	100	100

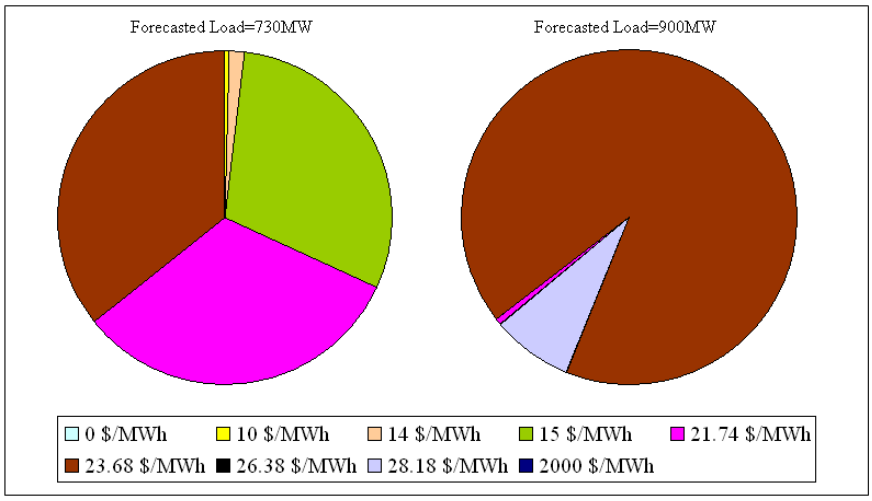


Figure 5.7. PMF of the LMP_t at Bus B.

Table 5.2 and Fig. 5.7 can be very useful for buyer and sellers to develop bidding strategies, demand response offers, and even create long-term contracts, since the results reveal the likelihood of realizing the forecasted LMP, considering the fact that there is always certain error in the load forecast results.

Fig. 5.7 may look messy when quite a few price candidates with considerable probability exist. In this case, it would be convenient to classify the prices into groups. Depending on the strategies, planners or decision makers may also care more about the probability of a range of LMP, instead of any individual LMP. For example, a planner may group all the possible prices into 3 categories, $0 \leq LMP \leq 15$, $15 < LMP < 30$, and $30 \leq LMP$. The corresponding probabilities for each of the groups are 30.92%, 69.09%, and 0% for the forecasted load of 730MW, and 0%, 100%, and 0% for the forecasted load of 900MW.

5.2.3.2. Alignment Probability of Deterministic LMP

Fig. 5.8 shows the curve of the alignment probability of the deterministic LMP at Bus B versus the forecasted load. By comparing Fig. 5.8 with Fig. 5.1, we can see that the low probabilities occur near the CLLs, and the lowest probability is approximately 30%,

indicating little confidence in the occurrence of the deterministic forecasted LMP. When the forecasted load is over 1300MW, the probability continues to decrease as the forecasted load approaches the maximum level (i.e., price of the VOLL) that the system can afford, namely, 1484.06MW.

Fig. 5.9 shows the alignment probability with a 10% price tolerance. Taking Table 5.2 as an example, without tolerance, the alignment probability at 730MW, where the deterministic LMP is $21.74/MWh, is 32.80% using (5.8). As a comparison, if a 10% price tolerance is adopted, the alignment probability will be 69.09% (=32.80% + 36.29%) using (5.9).

As shown in Figs. 5.8 and 5.9, the alignment probability curve will be higher with a 10% price tolerance, and the worst-case probability increases from 30% to 50%. Especially, the valley at around 1137MW in Fig. 5.8 disappears in Fig. 5.9. The alignment probability at the load of 1137MW is approximately 54% in Fig. 5.8, while it increases to nearly 99% with a 10% price tolerance considered; because in this case the difference of the deterministic LMP at the CLL of 1137.02MW is within 10%.

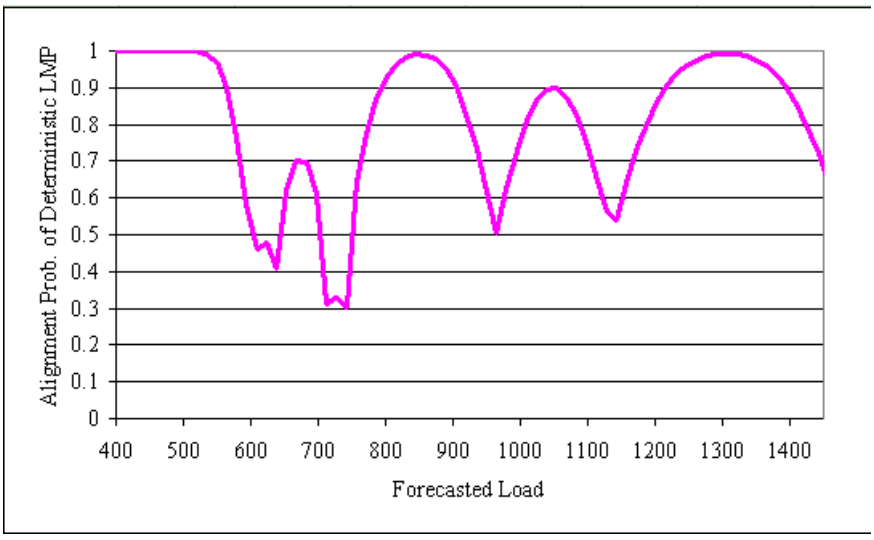


Figure 5.8. Alignment probability of deterministic LMP at Bus B versus the forecasted load

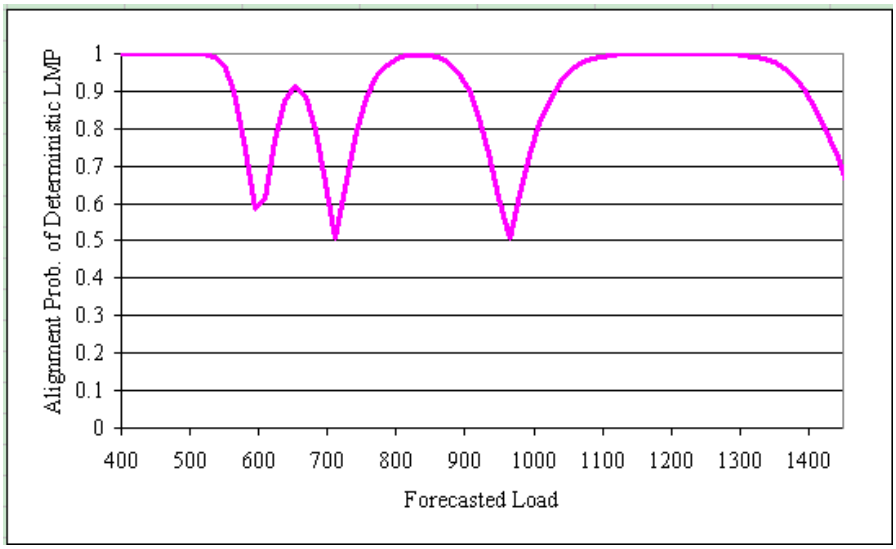


Figure 5.9. Alignment probability of the deterministic LMP at Bus B versus the forecasted load (with 10% price tolerance)

5.2.3.3. Expected Value of Probabilistic LMP

The expected value of the probabilistic LMP for the above case is compared with the deterministic LMPs, $LMP(D_t^F)$, which are shown in Table 5.3. It shows that the expected value of the probabilistic LMP can differ from the deterministic LMP for a specific forecasted load.

Table 5.3. Expected value of the probabilistic LMP in comparison with the Deterministic LMP for Bus B

$D^F{}_t$(MW)	Expected Value of Probabilistic LMP($/MWh)	Deterministic LMP($/MWh)
730	20.35	21.74
900	24.03	23.68

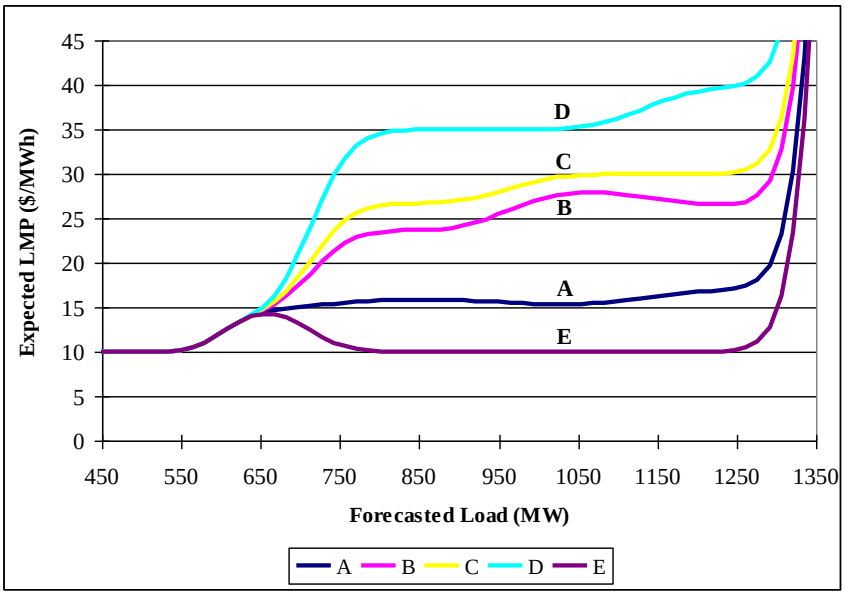


Figure 5.10. Expected value of the probabilistic LMP versus the forecasted load

The expected value of the probabilistic LMP versus the forecasted load curve is shown in Fig. 5.10. A load range beyond 1350MW is not shown simply because the high VOLL will make the curve hard to scale illustratively. It should be noted that the expected LMP will escalate sharply when the load is close to the maximum load level the system can afford, and will eventually reach 2000 $/MWh.

In the deterministic LMP-Load curve in Fig. 5.1, the sensitivity for Bus E at the 600MW load level is mathematically infinite since a step change occurs at 600MW. In the probability-based LMP-Load curve, we know that $\sum_{i=1}^{n}|p_i| = 2089\$/\text{MWh}$, $\mu_t = 600\text{MW}$, and $\sigma_t = 5\% \times \mu_t = 30\text{MW}$, therefore, the u pper bound of sensitivity is

$$\left|\frac{\partial E_{LMP_t}(\mu_t, \sigma_t)}{\partial \mu_t}\right| \le \frac{2089}{30\sqrt{2\pi}} = 27.78\$/\text{MWh}^2 .$$

Contrasted with the deterministic LMP-Load curve in Fig. 5.1, the curve of the expected value of the probabilistic LMP in Fig. 5.10 demonstrates the same overall trend. However,

Fig. 5.10 shows a much smoother curve without any step changes. This nice characteristic indicates that if the price simulation is based on the probabilistic approach described in this work, the error or uncertainty, with respect to the actual LMP in operation, will be reduced because of the elimination of step changes as shown in Fig. 5.10. Hence, the planners will not face the 1-or-0 type questions in their decision-making process when the loads are around the CLLs. This continuous function, as well as the PMF function of the Probabilistic LMP shown in Table 5.2 and the alignment probability shown in Figs. 8 and 9, gives market-participating planners, forecasters, or decision-makers a better idea regarding the potential risk due to the uncertainty in load forecasting so they can better evaluate bidding strategies, demand offers, and forward contracts.

Also shown in this probabilistic LMP forecasting figure is that when the load is closer to the CLLs, price uncertainty, i.e., the uncertainty associated with the forecasted deterministic LMP, will be higher. This matches the overall trend in the deterministic LMP in Fig. 5.1.

5.2.3.4. Impact of Load Forecasting Accuracy

In this section, three different levels of the standard deviation of load forecasting are examined, 5%, 3%, and 1%. Fig. 5.11 shows the probabilities of all possible values of LMP_t at Bus B for these three levels of standard deviation when the system load is 730MW. It can be seen from Fig. 5.11 that the probability of realizing 21.74 $/MWh, the deterministic LMP at 730MW load level, increases considerably with a smaller standard deviation. This is reasonable as a more accurate load forecast should lead to less deviation in the forecasted price.

Fig. 5.12 compares the expected value of the probabilistic LMP curves at the same bus. When the forecasted load is at a distance from any CLL, for example at 850MW, the three curves overlap very well. This suggests that different levels of the standard deviation make

trivial differences on the expected LMP at this load level. In addition, the sensitivity of the expected LMP at this load level is small, which indicates the expected LMP remains nearly constant when the forecasted load varies slightly around this level. In contrast, when the forecasted load is close to a CLL, for example at 600MW, the lower the standard deviation is and the closer the curve is to a step change curve shape. Furthermore, the inset in Fig. 5.12 shows that when the load level is closer to a CLL, the absolute value of the sensitivity of the expected LMP grows rapidly and the expected LMP becomes more sensitive to variations of the forecasted load.

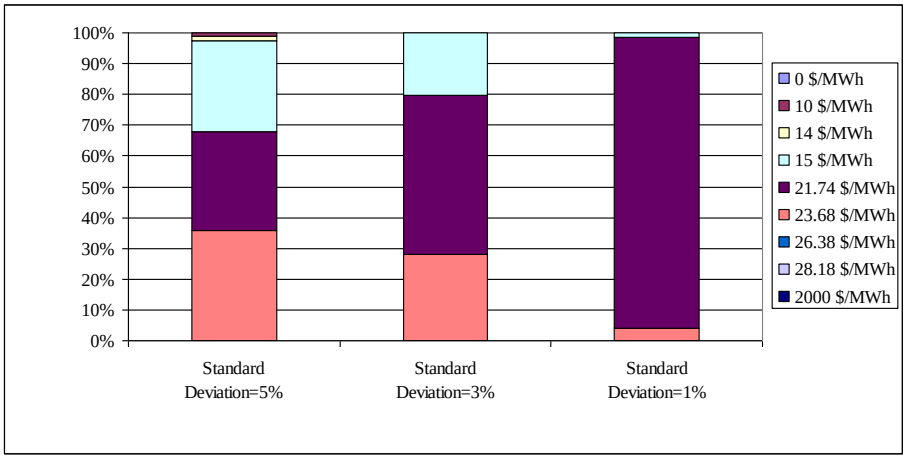


Figure 5.11. PMF of LMP_t at Bus B for three levels of standard deviation when the system load is 730MW

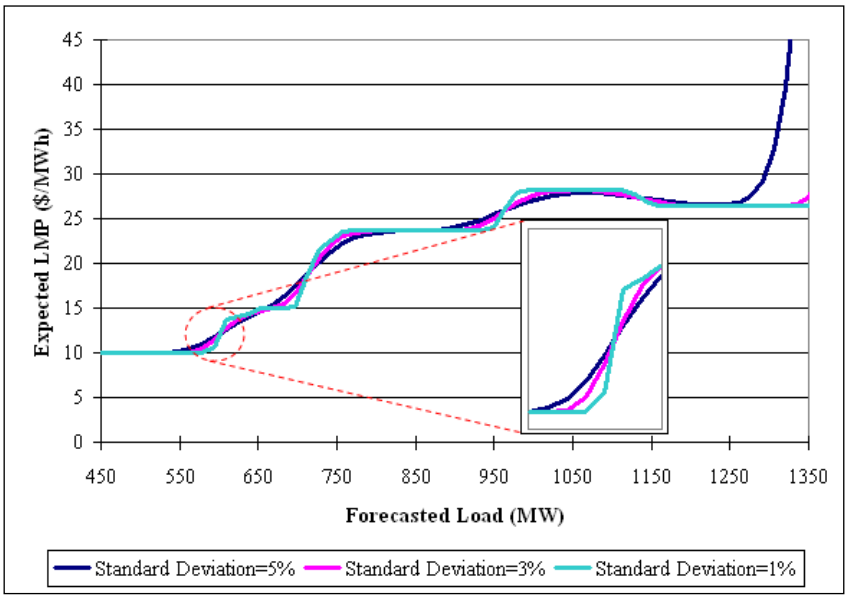


Figure 5.12. Expected value of the probabilistic LMP at Bus B versus the forecasted load for three levels of standard deviation

5.2.4. Numerical Study of the IEEE 118-Bus System

The study results on the IEEE 118-bus system [16] are briefly presented in this section to demonstrate the applicability of the proposed concepts and methods to larger systems. Conclusions similar to those for the PJM 5-bus system can be made. The system consists of 118 buses, 54 generators, and 186 branches. The system total load is 4242MW with a 9966.2MW total generation capacity. The detailed system data and diagram can be found in [16].

In the original IEEE 118-bus system, there is no generator bidding data and branch thermal limit data, which are indispensable in the performance of this study. Therefore, the generator bidding data is assumed as follows for illustrative purpose: 20 cheap generators with bidding data from \$10 to \$19.5 with \$0.5 increments; 20 expensive generators with bidding data from \$30 to \$49 with \$1 increments; and 14 of the most expensive generators with bidding data from \$70 to \$83 with \$1 increments. Five thermal limits are added into the transmission

system: 345MW for Line 69-77, 630MW for Line 68-81, 106MW for Line 83-85, and Line 94-100, 230MW for Line 80-98. The VOLL is set at 2000 $/MWh for all loads.

The deterministic LMP versus load curve for the IEEE 118-bus system is shown in Fig. 5.13. For better illustration, the curves are drawn only for a few selected buses and in the broad neighborhood of the base case load, namely, from 3550MW to 5820MW. Once again, the step change characteristic of the deterministic LMP curves is observed in Fig. 5.13.

Fig. 5.14 shows the curve of alignment probability of the deterministic LMP at two selected buses versus the forecasted load. The locations where low probabilities occur are aligned with the CLLs very well because the step changes, which the deterministic LMP contributes to the price uncertainty when there load forecast errors, are present. In Fig. 5.14, the majority of the alignment probability is less than 70%, and only a small range of the load (around 4920MW~5230MW) carries an 80% or more probability for realizing the deterministic forecasted LMP. Compared to the smaller system results in Fig. 5.8, larger systems tend to have a lower overall alignment probability since there are more CLLs or narrower ranges among two adjacent CLLs. This due to the involvement of more generators, and potentially, more congested lines.

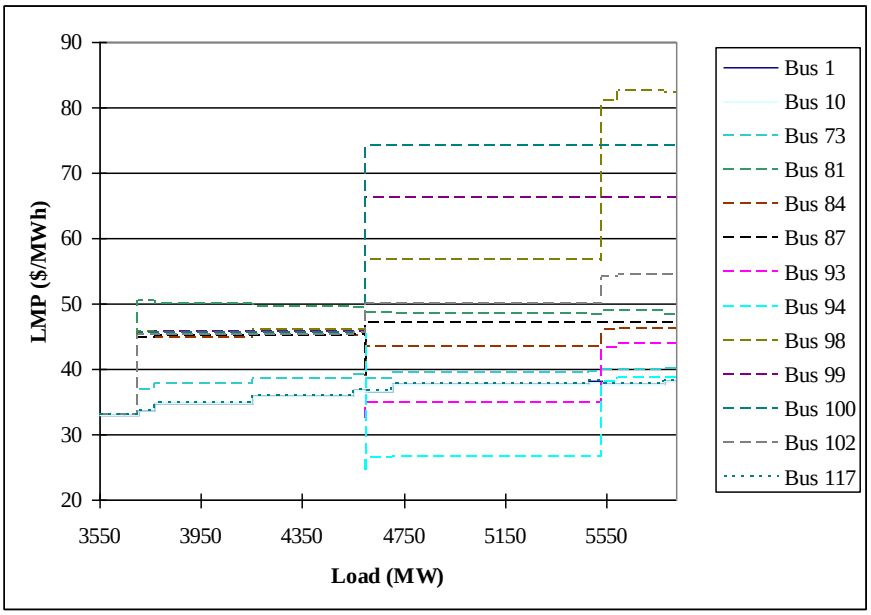


Figure 5.13. Deterministic LMP curve at selected buses with respect to different system loads for the IEEE 118-bus system

It can be seen from Fig. 5.14 that the alignment probabilities for Bus 81 and Bus 94 are almost identical for a vast load range (3550MW~ 5538MW). This is because the price changes at these two buses synchronize well with the load changes. This is a common pattern because the LMP at any specific bus will change at the CLLs unless there is a marginal unit at that bus to keep the LMP unchanged.

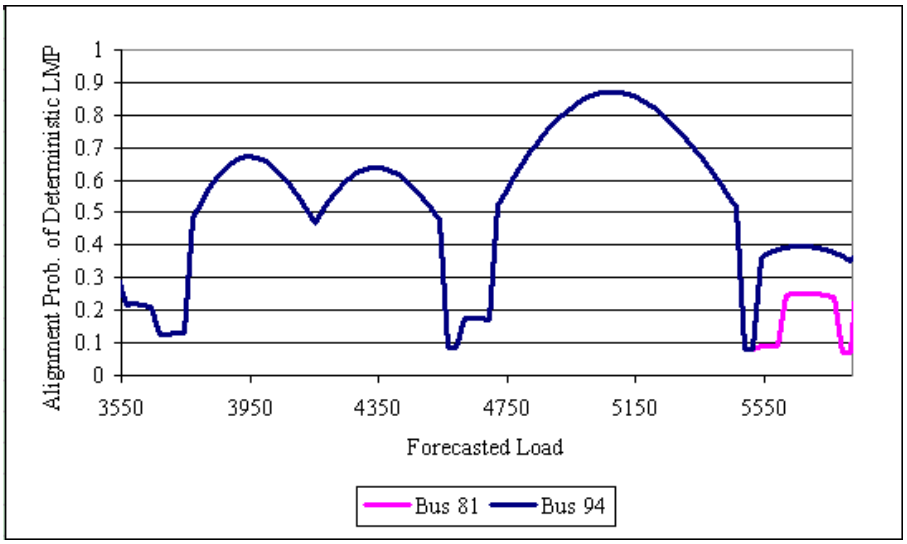


Figure 5.14. Alignment probability of the deterministic LMP at Bus 81 and Bus 94 versus the forecasted load for the IEEE 118-bus system

Fig. 5.15 presents the expected value of the probabilistic LMP versus the forecasted load curve at the same selected buses as in Fig. 5.13. The curves are observed to be highly smooth and there is no step change. Meanwhile, the curves track the overall trend of their deterministic counterparts. It should be noted that skyrocketing pattern in the right part of Fig. 5.10 for the PJM 5-bus system is not present in Fig. 5.15, because the load window shown in Fig. 5.15 is a large distant from the maximum affordable load. Therefore, the VOLL does not have any impact on this load range of interest.

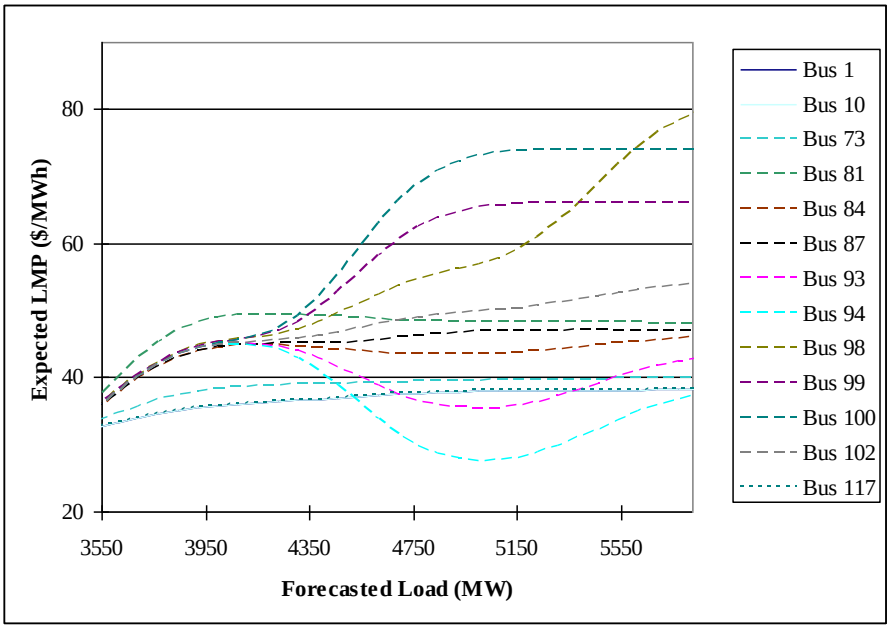


Figure 5.15. Expected value of the probabilistic LMP at selected buses versus the forecasted load for the IEEE 118-bus system

5.2.5. Conclusions

Load uncertainty exists due to a variety of reasons. Meanwhile, the LMP-Load curve has step changes at critical load levels where a new binding limit occurs. These are the major reasons of the LMP uncertainty. This work studies the LMP uncertainty with respect to the load in a probabilistic sense. The contribution can be summarized as follows:

- With the assumption of a normal distribution of the actual load, the concept of the probabilistic LMP is proposed and its probability mass function at hour t is presented. The probabilistic LMP does not correspond to a single deterministic value. Instead, it represents a set of discrete values (p_i) at a number of load intervals, and each value p_i has an associated probability.
- The alignment probability is proposed and formulated to define the likelihood that the deterministic LMP calculated based on a single value of the forecasted load is the same

as (or within a tolerance level of) the result from the probabilistic LMP. The alignment probability curve delivers the information of how likely the actual result from the probabilistic LMP is acceptable if compared with the deterministic LMP.

- The expected value of probabilistic LMP is derived and its curve with respect to the forecasted load is presented. The sensitivity of the curve is derived and shown to be bounded by finite values. In addition, the expected value of the probabilistic LMP versus the forecasted load curve is smooth and has no step changes. This avoids the 0-or-1 type of step changes if the deterministic LMP forecast is performed, and helps market participants make less risky decisions in generation bidding, demand offers, and/or forward contract negotiations.

The proposed concept and method are illustrated on a modified PJM 5-bus system as well as the IEEE 118-bus system. The results provide additional and useful information for understanding the LMP-Load curve from a probabilistic perspective.

5.3. Probabilistic LMP Forecasting for ACOPF Framework

ACOPF is deemed as the most representative mathematical model to the power generation scheduling problem and has gained some real-world applications [20]. Therefore, the impact of load forecasting uncertainty on LMP forecasting will be studied for the ACOPF framework in this section. The effect of power loss will be also examined since loss is well modeled in ACOPF while it is absent in lossless DCOPF.

The ACOPF is a much more complex model than the lossless DCOPF and contains a number of nonlinear constraints, which makes it very difficult, if not impossible, to perform analytical studies on the ACOPF solutions and by-products, such as LMPs. Apparently, the useful features of the lossless DCOPF, such as the linear marginal unit generation pattern and constant LMPs when the load varies within two adjacent CLLs, will not be valid for the

ACOPF due to existence of losses. Specifically, the LMP versus load curve will be different, though the step change phenomenon is still expected. For the lossless DCOPF framework, the LMP is constant between two adjacent critical load levels, while in the ACOPF framework, the LMP will steadily, but slightly, increase or decrease with a load variation within two adjacent critical levels. Consequently, the random variable, LMP at hour *t*, is no longer a discrete random variable; rather, it is a continuous random variable. In addition, the step change characteristic makes it a piece-wise continuous random variable. Therefore, we do not expect to produce the same representation of the probabilistic LMP for the ACOPF, even though the methodology will likely be the same. As such, we will examine the probability density function of this random variable and apply a methodology similar to that introduced in Section 5.2 to reveal its probabilistic features, such as probability density function, expected value, and its sensitivity.

It should be pointed out that same assumptions on the load made in Section 5.2 are used hereafter. That is, the actual load at hour *t*, i.e., D_t, is assumed to be a random variable and follows a normal distribution with mean μ_t and standard deviation σ_t. Its PDF and CDF functions are defined in equations (5.2)-(5.3).

5.3.1. Numeric Approach and Its Limitation

A straightforward approach is to numerically compute the Cumulative Density Function (CDF) and Probability Density Function (PDF), as well as the expected value of the random variable LMP_t.

5.3.1.1. Calculation of CDF of LMP_t

Assume the LMP versus load curve is composed of K (load, price) pairs, denoted by $(D_k, p_k), k \in \{1,2,...,K\}$, where K is a sufficiently large number. The entire load range is evenly divided by $\{D_k\}_{k=1}^{K}$.

By definition, the CDF of LMP_t at price p is formulated as follows

$$F_{LMP_t}(p) = \Pr\{LMP_t \le p\}. \tag{5.17}$$

If K is sufficiently large, the probability can be approximated as follows

$$F_{LMP_t}(p) = \Pr\{LMP_t \le p\} \approx \sum_{k \in \{p_k \le p, p_{k+1} \le p\}} \Pr\{D_k < D_t \le D_{k+1}\}. \tag{5.18}$$

Apparently it is very computational expensive when K is a large number due to the large number of evaluations of the CDF for a normal distribution. To reduce computational efforts, we can filter out every segment [D_k, D_{k+1}] which is outside $[\mu - 5\times\sigma, \mu + 5\times\sigma]$ since the probability outside this interval is numerically 0.

5.3.1.2. Calculation of PDF of LMP_t

The PDF of LMP_t at price p is defined as

$$f_{LMP_t}(p) = \lim_{\Delta p \to 0^+} \frac{\Pr\{p < LMP_t \le p + \Delta p\}}{\Delta p} = \lim_{\Delta p \to 0^+} \frac{F_{LMP_t}(p + \Delta p) - F_{LMP_t}(p)}{\Delta p}. \tag{5.19}$$

When Δp is sufficiently small, the PDF of LMP_t can be computed using CDF as follows

$$f_{LMP_t}(p) \approx \frac{F_{LMP_t}(p + \Delta p) - F_{LMP_t}(p)}{\Delta p}. \tag{5.20}$$

5.3.1.3. Calculation of Expected Value of LMP_t

When the PDF of LMP_t is computed, the expected value of LMP_t can be approximated by

$$E(LMP_t) = \int_{p_{\min}}^{p_{\max}} p \times f_{LMP_t}(p)dp \approx \sum_{j=1}^{n} p_j \times f_{LMP_t}(p_j) \times \Delta p \qquad (5.21)$$

where p_{min}, p_{max} are the minimum and maximum price of LMP_t respectively; n is a sufficiently large number; the interval [p_{min}, p_{max}] is evenly divided into n smaller intervals and $\Delta p = \frac{p_{\max} - p_{\min}}{n}$.

5.3.1.4. Discussions on the Numerical Approach

Although the numerical approach for calculating the CDF, PDF, and expected value of LMP_t is easy to implement, the computation is extremely expensive because of discretization and the large number of evaluations of complex functions such as the CDF of the normal distribution. In general, the higher the accuracy, the more computationally expensive the method will be. In addition, the results suffer significantly from an insufficient sampling resolution of discretization. Figure 5.16 illustrates how an insufficient sampling resolution may cause a change of the PDF shape and the miss of a spike area in discretizing a PDF curve. In fact, it happens for LMP_t in the ACOPF. As an example, for the PJM 5-bus system, when the load varies between two adjacent CLLs, 739.80MW and 924.75MW, the LMP at Bus A changes from 15.7977 $/MWh to 15.7951 $/MWh. When the forecasted load is 900MW, the actual PDF curve contains a pulse portion in a very narrow price range [15.7951$/MWh, 15.7977$/MWh]. The area is so narrow (about 0.0026$/MWh) that a sample resolution of 0.01$/MWh can hardly catch the pulse. This issue, worsened by the computational burden arising from the numerical method, is hard to overcome.

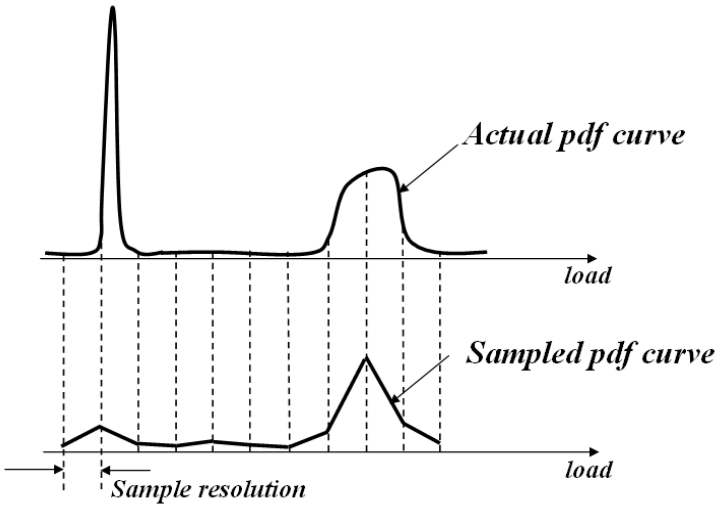


Figure 5.16. Illustration of insufficient sampling resolution in discretizing a PDF curve

5.3.2. Probabilistic LMP and its Probability Density Function

In order to reduce computational efforts and obtain more accurate results, an analytical approach will be adopted, which first establishes a mathematical model for the LMP versus the Load curve, and then derives, analytically, the formulations for CDF, PDF, expected value of LMP_t, etc.

By definition, the LMP is the partial derivative of the total generation cost with respect to the load change, and total generation cost is a linear combination of generation due to the adoption of the linear generation cost function. In addition, marginal unit generations are demonstrated to follow a perfect quadratic pattern when the load varies within two adjacent CLLs in the ACOPF framework. Hence, roughly speaking, the LMP should follow a linear pattern between any two adjacent CLLs, which can be observed in Figure 5.17. Figure 5.17 shows a typical LMP versus load curve for the modified PJM 5-bus system, defined in Section 4.3. It can be seen that step changes still exist at a few load levels, which are the

CLLs. The LMPs between the two adjacent CLLs suggest a linear pattern, as can be seen in the inset of Figure 5.17.

5.3.2.1. Models for the LMP-Load Curve

Figure 5.18 shows an illustrative picture of the LMP versus load curve in the ACOPF framework. The load axis is divided into *n-1* segments by a sequence of critical load levels (CLLs), $\{D_i\}_{i=1}^{n}$. Here, D_1 represents the no-load case (i.e., D_1=0), and D_n represents the maximum load that the system can supply due to the limits of total generation resources and transmission capabilities. The actual LMP in each load segment *i* is considered to be a straight (linear) line, with slope a_i and intercept b_i. The model to calculate LMP in the ACOPF was introduced in Section 3.2.2.

The LMP-Load curve is extended to include two extra segments in Figure 5.18. One is for the load from D_0 to D_1, where D_0 denotes a negative infinite load, and the associated price is zero. The second additional segment is defined as the load range from D_n to D_{n+1}, where D_{n+1} represents a positive infinite load. In this segment, the price is set as the Value of the Lost Load (VOLL) to reflect demand response to load shedding. Although the VOLL varies with customer groups and load interruption time and duration, it is a common practice to produce an aggregated VOLL to represent the average loss for an area. Therefore, the VOLL is assumed to be a constant value for simplicity in this work.

The two extra segments are added for mathematical completeness. In fact, they have a minimal, if any, impact on the study because the typical load range under study (for instance, from 0.8 p.u. to 1.2 p.u. of an average case load) is far from these two extreme segments, and the possibility of having the forecasted load close to zero or greater than D_n, the maximum load that the system can supply, is extremely rare and numerically zero.

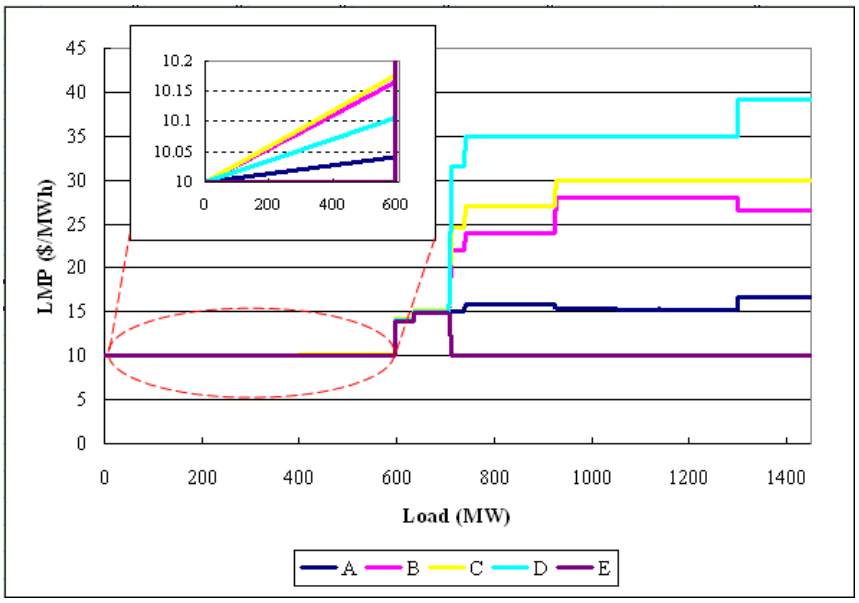


Figure 5.17. LMP at all buses with respect to different system loads for the modified PJM five-bus system

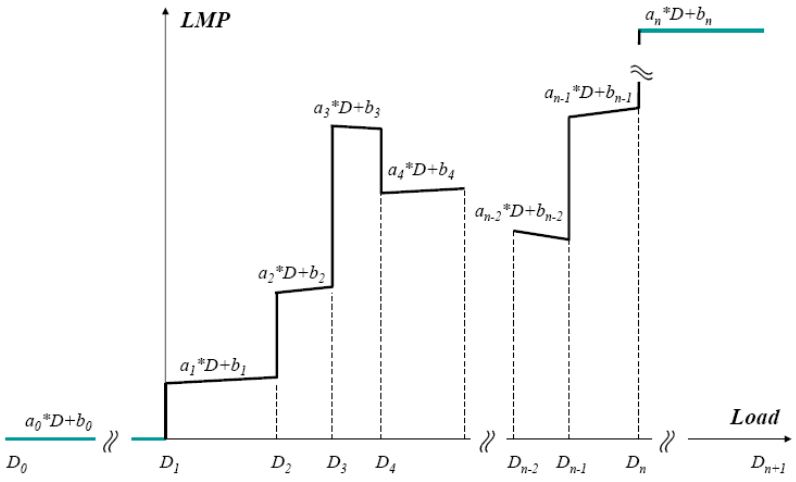


Figure 5.18. Extended LMP versus load curve

The LMP versus load curve can be formulated as

$$y(D)=\begin{cases} a_0 \times D + b_0, & D_0 < D \le D_1 \\ a_1 \times D + b_1, & D_1 < D \le D_2 \\ \vdots & \\ a_{n-1} \times D + b_{n-1}, & D_{n-1} < D \le D_n \\ a_n \times D + b_n, & D_n < D \le D_{n+1} \end{cases} \tag{5.22}$$

where

$a_0 = b_0 = a_n = 0$

$b_n = VOLL$

$D_0 = -\infty$

$D_1 = 0$

$D_{n+1} = \infty$.

The compact representation is given as follows

$$y(D) = \{a_i \times D + b_i \mid i \in \{1, \cdots, n\}, D_i < D \le D_{i+1}\}. \tag{5.23}$$

Apparently, $\frac{\partial LMP}{\partial D}$, the LMP sensitivity with respect to the load, is infinite at the critical load levels (CLLs), $\{D_i\}_{i=1}^{n-1}$.

5.3.2.2. Probabilistic LMP

Here it is assumed that the economic dispatch and LMP calculations are performed on an hourly basis. The LMP at hour *t*, denoted by LMP_t, is a function of D_t which is a random variable from the viewpoint of forecasting. Namely

$LMP_t = y(D_t)$.

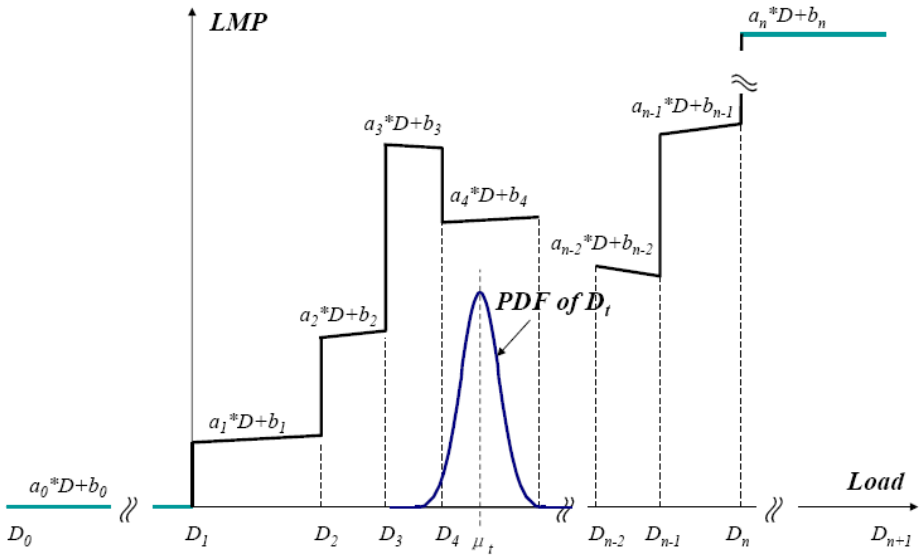


Figure 5.19. LMP-Load curve and probability distribution of D_t

Therefore, at the forecasting or planning stage, LMP_t should also be viewed as a random variable. This characteristic is inherited from the forecasted load. Fig. 5.19 shows the overlapping picture of the LMP-Load curve and the probability density function (PDF) of D_t.

Three types of curve segments exist in terms of the value of the price slope a_i. For example, for a curve segment with a positive price slope, the corresponding price range is $[a_i \times D_i + b_i, a_i \times D_{i+1} + b_i], i \in \{1,2,\cdots,n \mid a_i > 0\}$. In theory, the actual value of LMP_t could be any number in this price interval, and therefore, LMP_t is a continuous random variable. For a curve segment with a negative price slope, the price range is $[a_i \times D_{i+1} + b_i, a_i \times D_i + b_i], i \in \{1,2,\cdots,n \mid a_i < 0\}$; for a curve segment with a zero price slope, the price will be a constant value $b_i, i \in \{1,2,\cdots,n \mid a_i = 0\}$ throughout the load interval. Due to the step change phenomenon of the LMP versus load curve, there may or

may not exist intersections among these price intervals. Therefore, it can be inferred that LMP_t should be a piece-wise continuous random variable.

For an arbitrary price p, we can look up the LMP versus load curve to locate the corresponding load level(s). If there is no corresponding load level, for instance, p is not in any of the price intervals, the probability density value associated with p will be zero. If the corresponding load level(s) does exist, the probability density value associated with p will depend on the distance from the corresponding load level to the mean value of D_t, namely, μ_t . Intuitively, the shorter the distance, the higher the probability density value is. Furthermore, the probability density function is continuous on prices within any one of the price intervals, as will be shown in a later section. A special case is with the curve segment with a zero slope. The probability density value for the constant value b_i for the curve segment will be infinite because the CDF function has a step change at b_i. Figure 5.20 shows a schematic graph of a PDF curve of the piece-wise continuous random variable LMP_t. The vertical arrow represents the infinite probability density value.

It should be pointed out that the probability distributions of the price intervals are amplified for illustration purposes. In fact, the price intervals are typically so narrow that they will be displayed as single vertical bars when the PDF curve is drawn for the entire price range. However, if only one price interval is shown and well scaled in the graph, the corresponding PDF should manifest a continuous curve over the interval, excluding the two end points of the interval, as illustrated in Figure 5.20. This characteristic of the PDF curve will be exemplified in the case study section.

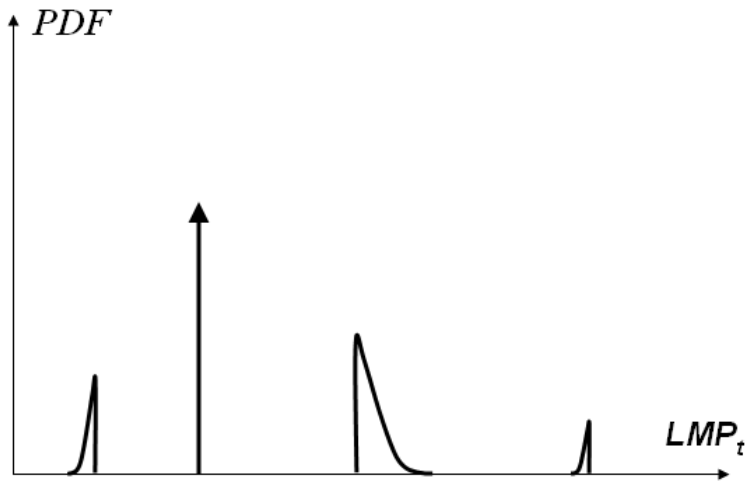


Figure 5.20. Probability Density Function of the Probabilistic LMP at hour *t*

Fig. 5.20 shows an important characteristic of the concept of the probabilistic LMP in the ACOPF framework:

The Probabilistic LMP at a specific (mean) load level is not a single deterministic value. Instead, it is a piece-wise continuous random variable with a piece-wise continuous probability density function and may contain infinite probability density values at certain price(s).

5.3.2.3. Cumulative Density Function of Probabilistic LMP

In order to obtain the formula of PDF of LMP_t, we need to firstly derive the CDF of LMP_t. Using probability theory, the cumulative density function of LMP_t can be derived as follows

$$\begin{aligned} F_{LMP_t}(p) &= \Pr(LMP_t \le p) \\ &= \Pr(y(D_t) \le p) \\ &= \Pr(D_t \in \Omega) \end{aligned} \tag{5.24}$$

Where $\Omega = \{x \mid y(x) \le p\}$.

Define $y_i(D_t) = a_i \times D_t + b_i, D_i < D_t \le D_{i+1}$ and $\Omega_i = \{x | y_i(x) \le p\}$, then we have

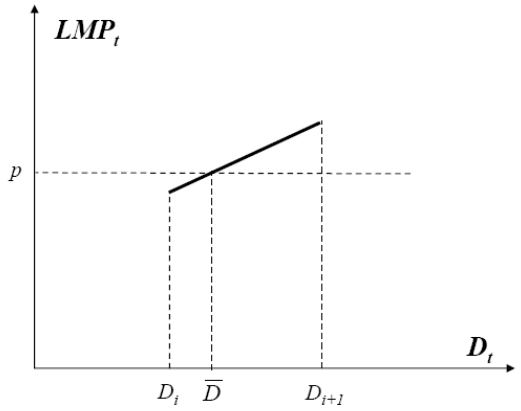
$$\Omega = \bigcup_{i=0}^{n} \Omega_i \tag{5.25}$$

$$\Omega_i \bigcap \Omega_j = \varnothing, \forall i \ne j . \tag{5.26}$$

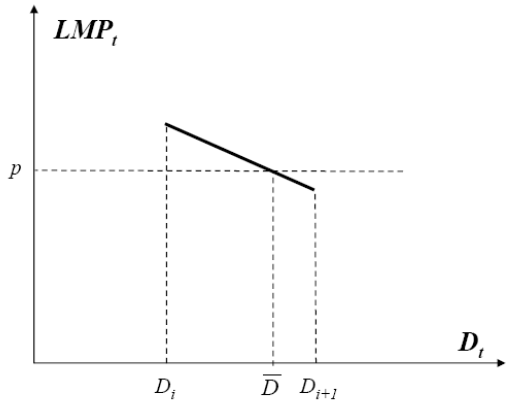
Therefore, the CDF function can be further derived as

$$\begin{aligned} F_{LMP_t}(p) &= \Pr(D_t \in \Omega) \\ &= \sum_{i=0}^{n} \Pr(D_t \in \Omega_i) \\ &= \sum_{i=0}^{n} \Pr(D_t \in \{x | y_i(x) \le p\}) \\ &= \sum_{i=0}^{n} \Pr(D_t \in \{x | a_i x + b_i \le p, D_i < x \le D_{i+1}\}) \end{aligned} . \tag{5.27}$$

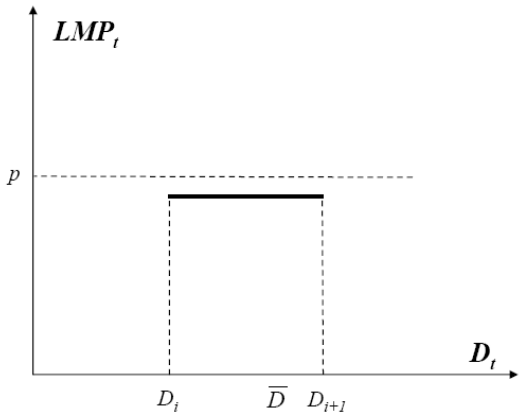
In order to calculate $\Pr(D_t \in \{x | a_i x + b_i \le p, D_i < x \le D_{i+1}\})$, three cases need to be considered respectively, i.e., (a) a_i>0., (b) a_i<0., and (c) a_i=0, as shown in Figure 5.21.



(a)



(b)



(c)

Figure 5.21. Three cases in computing the CDF of LMP_t. (a) $a_i>0$. (b) $a_i<0$. (c) $a_i=0$

(1) Case I: $a_i>0$

$$\Pr\{D_t \in x \mid a_i x + b_i \le p, D_i < x \le D_{i+1}\} = \int_{D_i}^{\overline{D}_i} \varphi(u)du \tag{5.28}$$

$$\text{where } \overline{D}_i = \begin{cases} \dfrac{p-b_i}{a_i}, & p \in [y(D_i), y(D_{i+1})] \\ D_i, & p < y(D_i) \\ D_{i+1}, & p > y(D_{i+1}) \end{cases}.$$

(2) Case II: $a_i<0$

$$\Pr\{D_t \in x \mid a_i x + b_i \le p, D_i < x \le D_{i+1}\} = \int_{\overline{D}_i}^{D_{i+1}} \varphi(u)du \tag{5.29}$$

$$\text{where } \overline{D}_i = \begin{cases} \dfrac{p-b_i}{a_i}, & p \in [y(D_{i+1}), y(D_i)] \\ D_{i+1}, & p < y(D_{i+1}) \\ D_i, & p > y(D_i) \end{cases}.$$

(3) Case III: $a_i=0$

$$\Pr\{D_t \in x \mid a_i x + b_i \le p, D_i < x \le D_{i+1}\} = \int_{D_i}^{\overline{D}_i} \varphi(u)du \tag{5.30}$$

$$\text{where } \overline{D}_i = \begin{cases} D_{i+1}, & p = b_i \\ D_i, & p < b_i \\ D_{i+1}, & p > b_i \end{cases}.$$

Summarizing the three cases, we have

$$F_{LMP_t}(p) = \sum_{i \in \{1,\cdots,n\} | a_i>0} \int_{D_i}^{\overline{D}_i} \varphi(u)du + \sum_{i \in \{1,\cdots,n\} | a_i<0} \int_{\overline{D}_i}^{D_{i+1}} \varphi(u)du + \sum_{i \in \{1,\cdots,n\} | a_i=0} \int_{D_i}^{\overline{D}_i} \varphi(u)du \tag{5.31}$$

5.3.2.4. Probability Density Function of Probabilistic LMP

To derive the formula for the PDF of LMP_t, we need to study the differentiability of the CDF function first.

Define $F_{1,i}(p) = \int_{D_i}^{\bar{D}_i} \varphi(u)du, i \in \{1,\cdots,n\}, a_i > 0$, then we have

$$F_{1,i}(p) = \begin{cases} 0, & p < y(D_i) \\ \Phi(\dfrac{p-b_i}{a_i}) - \Phi(D_i), & y(D_i) \le p \le y(D_{i+1}) \\ \Phi(D_{i+1}) - \Phi(D_i), & p > y(D_{i+1}) \end{cases}. \tag{5.32}$$

Figure 5.22 shows the schematic graph of the function $F_{1,i}(p)$. It is a continuous function, yet not differentiable at both $y(D_i)$ and $y(D_{i+1})$. Therefore, we have

$$F_{1,i}'(p) = \begin{cases} 0, & p < y(D_i) \\ \dfrac{1}{a_i}\varphi(\dfrac{p-b_i}{a_i}), & y(D_i) < p < y(D_{i+1}) \\ 0, & p > y(D_{i+1}) \end{cases}. \tag{5.33}$$

It should be noted that $F_{1,i}'(p)$ does not exist at both $y(D_i)$ and $y(D_{i+1})$.

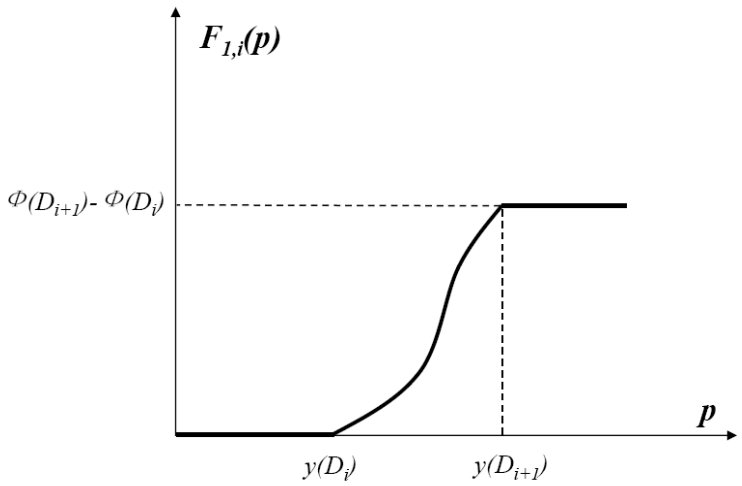


Figure 5.22. Function $F_{1,i}(p)$

Likewise, define $F_{2,i}(p) = \int_{D_i}^{D_{i+1}} \varphi(u)du, i \in \{1, \cdots, n\} | a_i < 0$, then we have

$$F_{2,i}(p) = \begin{cases} 0, & p < y(D_{i+1}) \\ \Phi(D_{i+1}) - \Phi(\frac{p-b_i}{a_i}), & y(D_{i+1}) \le p \le y(D_i) \\ \Phi(D_{i+1}) - \Phi(D_i), & p > y(D_i) \end{cases} . \quad (5.34)$$

Figure 5.23 shows the schematic graph of the function $F_{2,i}(p)$. It is a continuous function, yet not differentiable at both $y(D_i)$ and $y(D_{i+1})$. Therefore, we have

$$F_{2,i}'(p) = \begin{cases} 0, & p < y(D_{i+1}) \\ -\frac{1}{a_i}\varphi(\frac{p-b_i}{a_i}), & y(D_{i+1}) < p < y(D_i) \\ 0, & p > y(D_i) \end{cases} . \quad (5.35)$$

It should be noted that $F_{2,i}'(p)$ does not exist at both $y(D_i)$ and $y(D_{i+1})$.

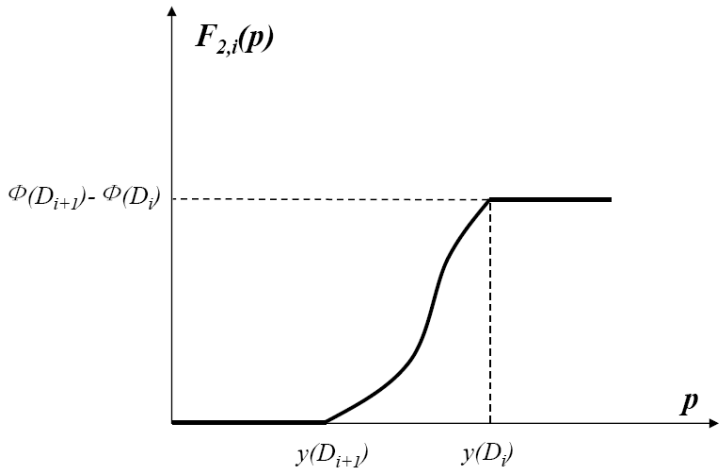


Figure 5.23. Function $F_{2,i}(p)$

Similarly, define $F_{3,i}(p) = \int_{D_i}^{\bar{D}_i} \varphi(u)du, i \in \{1, \cdots, n\} \& a_i = 0$, and then we have

$$F_{3,i}(p) = \begin{cases} 0, & p < b \\ \Phi(D_{i+1}) - \Phi(D_i), & p = b \\ \Phi(D_{i+1}) - \Phi(D_i), & p > b \end{cases} . \tag{5.36}$$

Figure 5.24 shows the schematic graph of the function $F_{3,i}(p)$. It is not a continuous function; rather, it is a step change function, which has an infinite derivative at *b*, namely, $y(D_i)$. Therefore, we have

$$F_{3,i}'(p) = \begin{cases} 0, p < y(D_i) \\ \infty, p = y(D_i) \\ 0, p > y(D_i) \end{cases} . \tag{5.37}$$

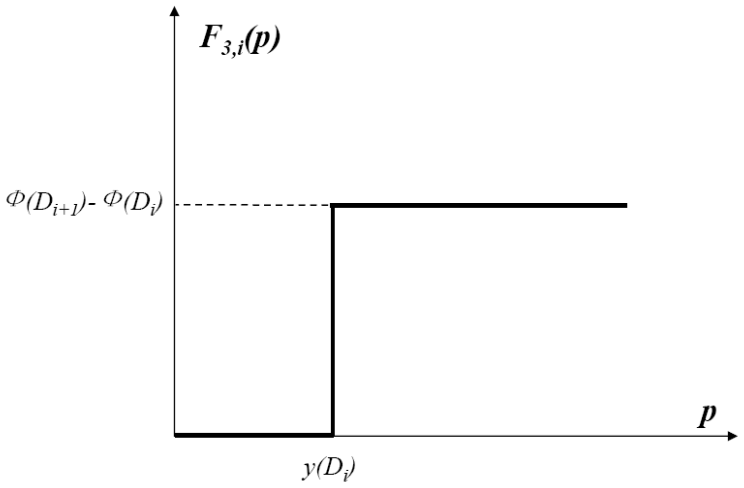


Figure 5.24. Function $F_{3,i}(p)$

With the definition of $F_{1,i}(p)$, $F_{2,i}(p)$, and $F_{3,i}(p)$, the CDF function of LMP_t can be rewritten as

$$\begin{aligned} F_{LMP_t}(p) &= \sum_{i\in\{1,\cdots,n\}|a_i>0} \int_{D_i}^{\bar{D}_i} \varphi(u)du + \sum_{i\in\{1,\cdots,n\}|a_i<0} \int_{\bar{D}_i}^{D_{i+1}} \varphi(u)du + \sum_{i\in\{1,\cdots,n\}|a_i=0} \int_{D_i}^{\bar{D}_i} \varphi(u)du \\ &= \sum_{i\in\{1,\cdots,n\}|a_i>0} F_{1,i}(p) + \sum_{i\in\{1,\cdots,n\}|a_i<0} F_{2,i}(p) + \sum_{i\in\{1,\cdots,n\}|a_i=0} F_{3,i}(p) \end{aligned} \tag{5.38}$$

Therefore, $F_{LMP_t}(p)$ is differentiable almost everywhere, except for prices at the boundaries of each interval, namely, $\{y(D_i)\}_{i=1}^n$.

In fact, we can assign arbitrary finite numbers as the derivative at those non-differentiable points. One option is to use the value of right derivative as the derivatives at those points, which is consistent with the calculations used in the numerical method. Therefore, we have

$$F_{1,i}'(p) = \begin{cases} 0, & p < y(D_i) \\ \frac{1}{a_i}\varphi(\frac{p-b_i}{a_i}), & y(D_i) \le p < y(D_{i+1}) \\ 0, & p \ge y(D_{i+1}) \end{cases} \tag{5.39}$$

$$F_{2,i}'(p) = \begin{cases} 0, & p \le y(D_{i+1}) \\ -\frac{1}{a_i}\varphi(\frac{p-b_i}{a_i}), & y(D_{i+1}) < p \le y(D_i) \\ 0, & p > y(D_i) \end{cases}. \tag{5.40}$$

The formula of the CDF function of LMP_t can be broken down into additional parts

$$\begin{aligned} F_{LMP_t}(p) &= \sum_{i\in\{1,\cdots,n\}|a_i>0} F_{1,i}(p) + \sum_{i\in\{1,\cdots,n\}|a_i<0} F_{2,i}(p) + \sum_{i\in\{1,\cdots,n\}|a_i=0} F_{3,i}(p) \\ &= \sum_{i\in\{1,\cdots,n\}|a_i>0, p\in[y(D_i),y(D_{i+1}))} F_{1,i}(p) + \sum_{i\in\{1,\cdots,n\}|a_i>0, p\notin[y(D_i),y(D_{i+1}))} F_{1,i}(p) \\ &\quad + \sum_{i\in\{1,\cdots,n\}|a_i<0, p\in(y(D_{i+1}),y(D_i)]} F_{2,i}(p) + \sum_{i\in\{1,\cdots,n\}|a_i<0, p\notin(y(D_{i+1}),y(D_i)]} F_{2,i}(p) \\ &\quad + \sum_{i\in\{1,\cdots,n\}|a_i=0, p=y(D_i)} F_{3,i}(p) + \sum_{i\in\{1,\cdots,n\}|a_i=0, p\neq y(D_i)} F_{3,i}(p) \end{aligned}. \tag{5.41}$$

Then, the probability density function of the LMP_t is derived as follows

$$\begin{aligned} f_{LMP_t}(p) &= F'_{LMP_t}(p) \\ &= \sum_{i\in\{1,\cdots,n \mid a_i>0,\, p\in[y(D_i),y(D_{i+1}))\}} \frac{1}{a_i}\varphi\left(\frac{p-b_i}{a_i}\right) + \sum_{i\in\{1,\cdots,n \mid a_i>0,\, p\notin[y(D_i),y(D_{i+1}))\}} 0 \\ &+ \sum_{i\in\{1,\cdots,n \mid a_i<0,\, p\in(y(D_{i+1}),y(D_i))\}} -\frac{1}{a_i}\varphi\left(\frac{p-b_i}{a_i}\right) + \sum_{i\in\{1,\cdots,n \mid a_i<0,\, p\notin(y(D_{i+1}),y(D_i))\}} 0 \\ &+ \sum_{i\in\{1,\cdots,n \mid a_i=0,\, p=y(D_i)\}} \infty + \sum_{i\in\{1,\cdots,n \mid a_i=0,\, p\neq y(D_i)\}} 0 \end{aligned} \quad . \quad (5.42)$$

Specifically, we have

$$f_{LMP_t}(p) = \infty, \quad p = y(D_i), i \in \{1,\cdots,n\} \mid a_i = 0 \}.$$

5.3.2.5. Alignment Probability of Probabilistic LMP

At hour *t*, if a single value of the forecasted load D_t^F is used for the LMP forecasting, the calculated LMP can be deterministically identified by looking up the LMP-Load curve, as shown in Fig. 5.17. This can be written as

$$\tilde{p} = y(D_t^F)$$

where $\tilde{p}$ is the LMP corresponding to the forecasted load D_t^F. This is called the deterministic LMP forecast.

It should be noted that the actual load may not be D_t^F, and correspondingly, the actual price is not always the same as the forecasted price.

In practice, it is interesting to know the probability associated with the deterministically forecasted LMP $\tilde{p}$. On the other hand, we have shown the LMP_t is a piece-wise continuous random variable in the ACOPF framework, and therefore, the probability of realizing any single value is zero in theory. In this regard, similar to the alignment probability concept presented in Section 5.2, we define the alignment probability in the ACOPF framework as the probability that the actual price is in a close neighborhood around the deterministic LMP

$$\begin{aligned} \mathbf{AP}_{\alpha} &= \Pr\left(\tilde{p}\times(1-\alpha\%) \le LMP_t \le \tilde{p}\times(1+\alpha\%)\right) \\ &= F_{LMP_t}\left(\tilde{p}\times(1+\alpha\%)\right) - F_{LMP_t}\left(\tilde{p}\times(1-\alpha\%)\right) \end{aligned} \tag{5.43}$$

where α is the tolerance percentage which gives the confidence of having the LMP forecast in an acceptable range. $\tilde{p}$ is the deterministically forecasted LMP.

For instance, if we choose 10% as the price tolerance level, then the results of the deterministically forecasted LMP is considered aligned with the actual LMP, if the actual LMP is within [90%, 110%] of the deterministic LMP $\tilde{p}$.

When the above equation is evaluated for every D_t^F in the entire interval $[D_1, D_n]$, an alignment probability versus D_t^F curve will be obtained. Each point of the curve represents the alignment probability that the actual price and the projected price $y(D_t^F)$ are in close vicinity when the load is D_t^F. When combined with the LMP-Load curve, this LMP alignment probability versus the forecasted load curve delivers very useful information such as how likely the projected price, at the forecasting stage, is close enough to the actual LMP at hour *t*, LMP_t.

5.3.3. Expected Value of Probabilistic LMP

5.3.3.1. Expected Value of Probabilistic LMP

By the Conditional Expectation theory [10], the expected value of LMP_t is derived as

$$\begin{aligned} & E(LMP_t) \\ &= \sum_{i=0}^{n} \left[E(LMP_t \mid D_i < D_t \le D_{i+1}) \times \Pr(D_i < D_t \le D_{i+1}) \right] \\ &= \sum_{i=0}^{n} \left[E(y(D_t) \mid D_i < D_t \le D_{i+1}) \times \Pr(D_i < D_t \le D_{i+1}) \right] \\ &= \sum_{i=0}^{n} \left[E(a_i \times D_t + b_i \mid D_i < D_t \le D_{i+1}) \times \Pr(D_i < D_t \le D_{i+1}) \right. \end{aligned} \tag{5.44}$$

If a random variable *Y* is a linear function of a random variable *X*, namely, *Y=l(X)*, where *l(.)* denotes the linear function, then the expected value of *Y* is $E(Y)=E(l(X))=l(E(X))$. By this theory, we have

$$\begin{aligned} E(LMP_t) &= \sum_{i=0}^{n} \left[E(a_i \times D_t + b_i \mid D_i < D_t \le D_{i+1}) \times \Pr(D_i < D_t \le D_{i+1}) \right] \\ &= \sum_{i=0}^{n} \left[(a_i \times E(D_t \mid D_i < D_t \le D_{i+1}) + b_i) \times \Pr(D_i < D_t \le D_{i+1}) \right] \end{aligned} . \quad (5.45)$$

By the definition of the conditional probability density function [10], the conditional density function of D_t, given any event $D_i < D_t \le D_{i+1}$ is

$$f_{D_t|D_i<D_t\le D_{i+1}}(u) = \begin{cases} \dfrac{f(u)}{\Pr(D_i < D_t \le D_{i+1})}, & x \in (D_i, D_{i+1}] \\ 0, & x \notin (D_i, D_{i+1}] \end{cases} . \quad (5.46)$$

Therefore, the expected value of D_t given any event $D_i < D_t \le D_{i+1}$ is

$$\begin{aligned} E(D_t \mid D_i < D_t \le D_{i+1}) &= \int_{-\infty}^{\infty} u \times f_{D_t|D_i<D_t\le D_{i+1}}(u) \times du \\ &= \int_{D_i}^{D_{i+1}} u \times \frac{f(u)}{\Pr(D_i < D_t \le D_{i+1})} \times du \\ &= \frac{1}{\Pr(D_i < D_t \le D_{i+1})} \int_{D_i}^{D_{i+1}} u \times f(u) \times du \\ &= \frac{1}{\Phi(D_{i+1}) - \Phi(D_i)} \int_{D_i}^{D_{i+1}} u \times \varphi(u) \times du \end{aligned} . \quad (5.47)$$

Then, the expected value of the LMP_t is further derived as follows

$$\begin{aligned} E(LMP_t) &= \sum_{i=0}^{n} \left[(a_i \times E(D_t \mid D_i < D_t \le D_{i+1}) + b_i) \times \Pr(D_i < D_t \le D_{i+1}) \right] \\ &= \sum_{i=0}^{n} \left[\left(a_i \times \left(\frac{1}{\Phi(D_{i+1}) - \Phi(D_i)} \int_{D_i}^{D_{i+1}} u \times \varphi(u) \times du \right) + b_i \right) \times (\Phi(D_{i+1}) - \Phi(D_i)) \right] \\ &= \sum_{i=0}^{n} \left[\left(a_i \times \left(\int_{D_i}^{D_{i+1}} u \times \varphi(u) \times du \right) + b_i \times (\Phi(D_{i+1}) - \Phi(D_i)) \right) \right] \end{aligned} \quad (5.48)$$

$$\int_{D_i}^{D_{i+1}} u\varphi(u)du = \int_{D_i}^{D_{i+1}} u \times \frac{1}{\sigma_t\sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du$$

$$= -\frac{\sigma_t}{\sqrt{2\pi}} \int_{D_i}^{D_{i+1}} \left[-\frac{1}{\sigma_t^2}(u-\mu_t) \times e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} - \frac{\mu_t}{\sigma_t^2} \times e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} \right] du$$

$$= -\frac{\sigma_t}{\sqrt{2\pi}} \int_{D_i}^{D_{i+1}} \left[d\left(e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} \right) - \frac{\mu_t}{\sigma_t^2} \times e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du \right] \quad . (5.49)$$

$$= -\frac{\sigma_t}{\sqrt{2\pi}} \left[e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} \Big|_{D_i}^{D_{i+1}} - \frac{\mu_t\sqrt{2\pi}}{\sigma_t} \int_{D_i}^{D_{i+1}} \frac{1}{\sigma_t\sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du \right]$$

$$= \sigma_t^2 \left(\varphi(D_i) - \varphi(D_{i+1}) \right) + \mu_t \left(\Phi(D_{i+1}) - \Phi(D_i) \right)$$

The ultimate formula for the expected value of LMP_t is derived as

$$E(LMP_t) = \sum_{i=0}^{n} \left[a_i \times \left(\sigma_t^2 \left(\varphi(D_i) - \varphi(D_{i+1}) \right) + \mu_t \left(\Phi(D_{i+1}) - \Phi(D_i) \right) \right) + b_i \times \left(\Phi(D_{i+1}) - \Phi(D_i) \right) \right]$$

$$= \sum_{i=0}^{n} \left[a_i \times \sigma_t^2 \left(\varphi(D_i) - \varphi(D_{i+1}) \right) + \left(a_i \times \mu_t + b_i \right) \times \left(\Phi(D_{i+1}) - \Phi(D_i) \right) \right] \quad (5.50)$$

5.3.3.2. Sensitivity of Expected Value of Probabilistic LMP

Define $E_{LMP_t}(\mu_t, \sigma_t) \overset{\Delta}{=} E(LMP_t)$

$$E_{LMP_t}(\mu_t, \sigma_t) = \sum_{i=0}^{n} \left[\frac{a_i \times \sigma_t}{\sqrt{2\pi}} \left(e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}} - e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}} \right) + \left(a_i \times \mu_t + b_i \right) \times \left(\int_{D_i}^{D_{i+1}} \frac{1}{\sigma_t\sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du \right) \right] \quad (5.51)$$

Taking the partial derivative with respect to μ_t gives

$$\frac{\partial E_{LMP_t}(\mu_t,\sigma_t)}{\partial \mu_t}$$
$$= \sum_{i=0}^{n}\left[\frac{a_i\times\sigma_t}{\sqrt{2\pi}}\left(\frac{D_i-\mu_t}{\sigma_t^2}e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}}-\frac{D_{i+1}-\mu_t}{\sigma_t^2}e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}}\right)\right]$$
$$+\sum_{i=0}^{n}\left[a_i\times\left(\int_{D_i}^{D_{i+1}}\frac{1}{\sigma_t\sqrt{2\pi}}e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}}du\right)+(a_i\times\mu_t+b_i)\times\left(\int_{D_i}^{D_{i+1}}\frac{u-\mu_t}{\sigma_t^3\sqrt{2\pi}}e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}}du\right)\right] \quad (5.52)$$

In Appendix B, we have already derived that

$$\int_{D_i}^{D_{i+1}}\frac{u-\mu_t}{\sigma_t^3\sqrt{2\pi}}e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}}du=\frac{1}{\sigma_t\sqrt{2\pi}}\left(e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}}-e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}}\right). \quad (5.53)$$

Therefore, the sensitivity of the expected value of LMP_t is

$$\frac{\partial E_{LMP_t}(\mu_t,\sigma_t)}{\partial \mu_t}$$
$$= \sum_{i=0}^{n}\left[\frac{a_i\times\sigma_t}{\sqrt{2\pi}}\left(\frac{D_i-\mu_t}{\sigma_t^2}e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}}-\frac{D_{i+1}-\mu_t}{\sigma_t^2}e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}}\right)\right]$$
$$+\sum_{i=0}^{n}\left[a_i\times\left(\int_{D_i}^{D_{i+1}}\frac{1}{\sigma_t\sqrt{2\pi}}e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}}du\right)+(a_i\times\mu_t+b_i)\times\frac{1}{\sigma_t\sqrt{2\pi}}\left(e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}}-e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}}\right)\right]$$
$$=\sum_{i=0}^{n}\left[a_i\times(D_i-\mu_t)\times\varphi(D_i)-(D_{i+1}-\mu_t)\times\varphi(D_{i+1})\right]$$
$$+\sum_{i=0}^{n}\left[a_i\times\Phi(D_{i+1})-\Phi(D_i)+(a_i\times\mu_t+b_i)\times\varphi(D_i)-\varphi(D_{i+1})\right] \quad (5.54)$$

5.3.3.3. Lower and Upper Bounds of the Sensitivity of the Expected Value of Probabilistic LMP

In order to study the lower and upper bounds of the sensitivity of the expected value of LMP_t, we need to first study the bounds of the item $\frac{D_i-\mu_t}{\sigma_t^2}e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}}-\frac{D_{i+1}-\mu_t}{\sigma_t^2}e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}}$ in the formula.

Define $g(x)=x\times e^{-x^2}$, then we have

$$g'(x)=e^{-x^2}+x\times e^{-x^2}\times(-2x)=(1-2x^2)e^{-x^2} \tag{5.55}$$

$$\begin{cases} g'(x)<0, x<-\frac{1}{\sqrt{2}} \\ g'(x)\geq 0, -\frac{1}{\sqrt{2}}\leq x\leq \frac{1}{\sqrt{2}} \\ g'(x)<0, x>\frac{1}{\sqrt{2}} \end{cases} \tag{5.56}$$

So the maximum and minimum values of $g(x)$ are $g(\frac{1}{\sqrt{2}})$ and $g(-\frac{1}{\sqrt{2}})$ respectively, i.e., $g(-\frac{1}{\sqrt{2}})\leq g(x)\leq g(\frac{1}{\sqrt{2}})$. This function is depicted in Figure 5.25.

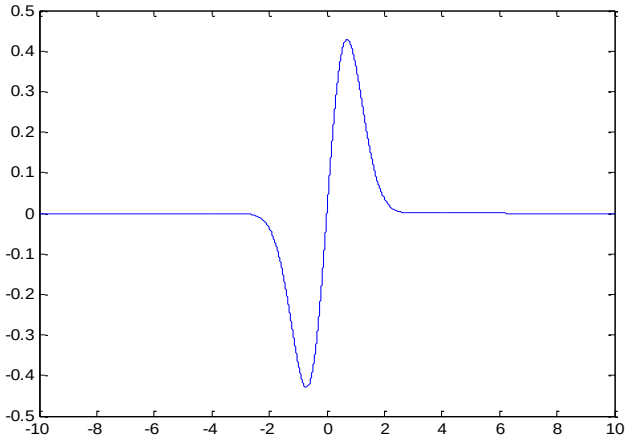


Figure 5.25. Function *g(x)*

Therefore, the upper bound of $\frac{D_i-\mu_t}{\sigma_t^2}e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}}-\frac{D_{i+1}-\mu_t}{\sigma_t^2}e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}}$ is derived as

$$\begin{aligned}&\frac{D_i-\mu_t}{\sigma_t^2}e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}}-\frac{D_{i+1}-\mu_t}{\sigma_t^2}e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}}\\&\quad=\frac{\sqrt{2}}{\sigma_t}\left(\frac{D_i-\mu_t}{\sqrt{2}\sigma_t}e^{-\left(\frac{D_i-\mu_t}{\sqrt{2}\sigma_t}\right)^2}-\frac{D_{i+1}-\mu_t}{\sqrt{2}\sigma_t}e^{-\left(\frac{D_{i+1}-\mu_t}{\sqrt{2}\sigma_t}\right)^2}\right).\qquad(5.57)\\&\quad\le\frac{\sqrt{2}}{\sigma_t}\left(g(\frac{1}{\sqrt{2}})-g(-\frac{1}{\sqrt{2}})\right)=\frac{2e^{-\frac{1}{2}}}{\sigma_t}\end{aligned}$$

Similarly, the lower bound of $\frac{D_i-\mu_t}{\sigma_t^2}e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}}-\frac{D_{i+1}-\mu_t}{\sigma_t^2}e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}}$ is as follows

$$\begin{aligned}&\frac{D_i-\mu_t}{\sigma_t^2}e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}}-\frac{D_{i+1}-\mu_t}{\sigma_t^2}e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}}\\&\quad=\frac{\sqrt{2}}{\sigma_t}\left(\frac{D_i-\mu_t}{\sqrt{2}\sigma_t}e^{-\left(\frac{D_i-\mu_t}{\sqrt{2}\sigma_t}\right)^2}-\frac{D_{i+1}-\mu_t}{\sqrt{2}\sigma_t}e^{-\left(\frac{D_{i+1}-\mu_t}{\sqrt{2}\sigma_t}\right)^2}\right).\qquad(5.58)\\&\quad\ge\frac{\sqrt{2}}{\sigma_t}\left(g(-\frac{1}{\sqrt{2}})-g(\frac{1}{\sqrt{2}})\right)=-\frac{2e^{-\frac{1}{2}}}{\sigma_t}\end{aligned}$$

In summary, the lower and upper bound of $\frac{D_i-\mu_t}{\sigma_t^2}e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}}-\frac{D_{i+1}-\mu_t}{\sigma_t^2}e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}}$ is

$$-\frac{2e^{-\frac{1}{2}}}{\sigma_t}\le\frac{D_i-\mu_t}{\sigma_t^2}e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}}-\frac{D_{i+1}-\mu_t}{\sigma_t^2}e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}}\le\frac{2e^{-\frac{1}{2}}}{\sigma_t}.\qquad(5.59)$$

In addition, we have

$$0\le\int_{D_i}^{D_{i+1}}\frac{1}{\sigma_t\sqrt{2\pi}}e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}}du\le 1\qquad(5.60)$$

$$-1\le e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}}-e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}}\le 1.\qquad(5.61)$$

Hence, the upper bound of the sensitivity of the expected value of the LMP_t is derived as follows

$$
\begin{aligned}
&\frac{\partial E_{LMP_t}(\mu_t,\sigma_t)}{\partial \mu_t} \\
&= \sum_{i=0}^{n}\left[\frac{a_i\times\sigma_t}{\sqrt{2\pi}}\left(\frac{D_i-\mu_t}{\sigma_t^2}e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}}-\frac{D_{i+1}-\mu_t}{\sigma_t^2}e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}}\right)\right] \\
&\quad+\sum_{i=0}^{n}\left[a_i\times\left(\int_{D_i}^{D_{i+1}}\frac{1}{\sigma_t\sqrt{2\pi}}e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}}du\right)+\left(a_i\times\mu_t+b_i\right)\times\frac{1}{\sigma_t\sqrt{2\pi}}\left(e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}}-e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}}\right)\right] \\
&\le \sum_{i\in\{1,\cdots,n\}\,|\,a_i>0}\left(\frac{a_i\times\sigma_t}{\sqrt{2\pi}}\times\frac{2e^{-\frac{1}{2}}}{\sigma_t}\right)+\sum_{i\in\{1,\cdots,n\}\,|\,a_i<0}\left(\frac{a_i\times\sigma_t}{\sqrt{2\pi}}\times\left(-\frac{2e^{-\frac{1}{2}}}{\sigma_t}\right)\right)+\sum_{i\in\{1,\cdots,n\}\,|\,a_i>0}\left(a_i\times1\right)+\sum_{i\in\{1,\cdots,n\}\,|\,a_i<0}\left(a_i\times0\right) \\
&\quad+\sum_{i\in\{1,\cdots,n\}\,|\,a_i\times\mu_t+b_i>0}\left(\frac{a_i\times\mu_t+b_i}{\sigma_t\sqrt{2\pi}}\times1\right)+\sum_{i\in\{1,\cdots,n\}\,|\,a_i\times\mu_t+b_i<0}\left(\frac{a_i\times\mu_t+b_i}{\sigma_t\sqrt{2\pi}}\times\left(-1\right)\right) \\
&= \sum_{i\in\{1,\cdots,n\}\,|\,a_i>0}\left(\frac{2a_i\times e^{-\frac{1}{2}}}{\sqrt{2\pi}}+a_i\right)+\sum_{i\in\{1,\cdots,n\}\,|\,a_i<0}\left(-\frac{2a_i\times e^{-\frac{1}{2}}}{\sqrt{2\pi}}\right) \\
&\quad+\sum_{i\in\{1,\cdots,n\}\,|\,a_i\times\mu_t+b_i>0}\left(\frac{a_i\times\mu_t+b_i}{\sigma_t\sqrt{2\pi}}\right)+\sum_{i\in\{1,\cdots,n\}\,|\,a_i\times\mu_t+b_i<0}\left(-\frac{a_i\times\mu_t+b_i}{\sigma_t\sqrt{2\pi}}\right)
\end{aligned}
\tag{5.62}
$$

Similarly, the lower bound of sensitivity of expected value of LMP_t is presented as follows.

$$
\begin{aligned}
&\frac{\partial E_{LMP_t}(\mu_t,\sigma_t)}{\partial \mu_t} \\
&= \sum_{i=0}^{n}\left[\frac{a_i\times\sigma_t}{\sqrt{2\pi}}\left(\frac{D_i-\mu_t}{\sigma_t^2}e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}}-\frac{D_{i+1}-\mu_t}{\sigma_t^2}e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}}\right)\right] \\
&\quad+\sum_{i=0}^{n}\left[a_i\times\left(\int_{D_i}^{D_{i+1}}\frac{1}{\sigma_t\sqrt{2\pi}}e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}}du\right)+(a_i\times\mu_t+b_i)\times\frac{1}{\sigma_t\sqrt{2\pi}}\left(e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}}-e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}}\right)\right] \\
&\geq \sum_{i\in\{1,\cdots,n\,|\,a_i>0\}}\left(\frac{a_i\times\sigma_t}{\sqrt{2\pi}}\times\left(-\frac{2e^{-\frac{1}{2}}}{\sigma_t}\right)\right)+\sum_{i\in\{1,\cdots,n\,|\,a_i<0\}}\left(\frac{a_i\times\sigma_t}{\sqrt{2\pi}}\times\frac{2e^{-\frac{1}{2}}}{\sigma_t}\right)+\sum_{i\in\{1,\cdots,n\,|\,a_i>0\}}(a_i\times 0)+\sum_{i\in\{1,\cdots,n\,|\,a_i<0\}}(a_i\times 1) \\
&\quad+\sum_{i\in\{1,\cdots,n\,|\,a_i\times\mu_t+b_i>0\}}\left(\frac{a_i\times\mu_t+b_i}{\sigma_t\sqrt{2\pi}}\times(-1)\right)+\sum_{i\in\{1,\cdots,n\,|\,a_i\times\mu_t+b_i<0\}}\left(\frac{a_i\times\mu_t+b_i}{\sigma_t\sqrt{2\pi}}\times 1\right) \\
&= \sum_{i\in\{1,\cdots,n\,|\,a_i>0\}}\left(-\frac{2a_i\times e^{-\frac{1}{2}}}{\sqrt{2\pi}}\right)+\sum_{i\in\{1,\cdots,n\,|\,a_i<0\}}\left(\frac{2a_i\times e^{-\frac{1}{2}}}{\sqrt{2\pi}}+a_i\right) \\
&\quad+\sum_{i\in\{1,\cdots,n\,|\,a_i\times\mu_t+b_i>0\}}\left(-\frac{a_i\times\mu_t+b_i}{\sigma_t\sqrt{2\pi}}\right)+\sum_{i\in\{1,\cdots,n\,|\,a_i\times\mu_t+b_i<0\}}\left(\frac{a_i\times\mu_t+b_i}{\sigma_t\sqrt{2\pi}}\right)
\end{aligned}
\tag{5.63}
$$

5.3.4. Numerical Study of a Modified PJM 5-Bus System

In this section, a numeric study will be performed on the PJM 5-Bus system [11], with slight modifications. The modifications are for illustration purposes and are detailed in Section 4.2.4. The configuration of the system is shown in Figure 4.3.

To calculate the LMP versus load curve as shown in Fig. 5.17, it is assumed that the system load change is distributed to each bus load proportional to its base case load for simplicity. For better illustration, it is assumed that μ_t is always equal to the forecasted load D_t^F, and the standard deviation σ_t is taken as 5% of the mean μ_t unless otherwise stated. The VOLL is set at \$2000/MWh.

5.3.4.1. Approximation of LMP Curve

A linear polynomial curve-fitting is employed to approximate the actual LMP between every two adjacent CLLs, and the coefficients are used to establish the mathematical model for the LMP versus load curve. Table 5.4 shows the curve-fitting coefficients for the LMP curves at all buses when the load is within [0, 590] MW. It implies that the LMPs at all buses, except Bus E, increase slightly, while the LMP at Bus E remains 10$/MWh for the entire load interval.

With the linear polynomial coefficients obtained through the curve-fitting, the mathematical LMP versus load model is established, which is a piece-wise linear curve with step changes at the CLLs. The mathematical representation of the curve is a very good approximation to the actual LMP versus load curve, and therefore, can be used to facilitate an analytical study on topics such as CDF, PDF, etc. In fact, the curve represented by the mathematical formula looks almost identical to the actual curve shown in Figure 5.17. The largest difference is less than 0.07$/MWh, approximately 0.7% of the lowest LMP, $10/MWh. Therefore, the curve is not redrawn here.

Table 5.4. Curve-fitting coefficients for the LMP curves at all buses when the load is within [0, 590] MW

	a ($/(MWh*MW))	b ($/MWh)
LMP@A	0.0001	9.9999
LMP@B	0.0003	9.9993
LMP@C	0.0003	9.9993
LMP@D	0.0002	9.9997
LMP@E	0	10

5.3.4.2. Cumulative Density Function of Probabilistic LMP

Figure 5.26 shows the CDF curve of the probabilistic LMP at Bus B for the forecasted load at 730 MW and 900 MW, respectively. The figure suggests the staircase pattern of the CDF curve. Combined with the LMP versus the Load curve such as in Figure 5.17, it can be seen that the prices at which a step change occurs coincide with the price intervals near the forecasted load level for Bus B. The corresponding PDF values for these prices are expected to be higher than the PDF values of other prices, as will be verified in the next section.

A careful study reveals that the majority of the step changes observed in the CDF curve are not really step changes. Figure 5.27 redraws the same curves in Figure 5.26 in a narrow range around \$24/MWh, where a step change appears. It can be seen that both CDF curves move smoothly from \$23.95/MWh to \$24.02/MWh. It is actually consistent with the aforementioned theoretical part in Section 5.3.2.3, where the CDF function is shown to be differentiable almost everywhere except at the price boundaries of each interval of the LMP versus the Load curve, namely, $\{y(D_i)\}_{i=1}^{n}$. Nevertheless, the change of the CDF values happens in such narrow price intervals that it looks just like a step change when plotted for a broader range of prices.

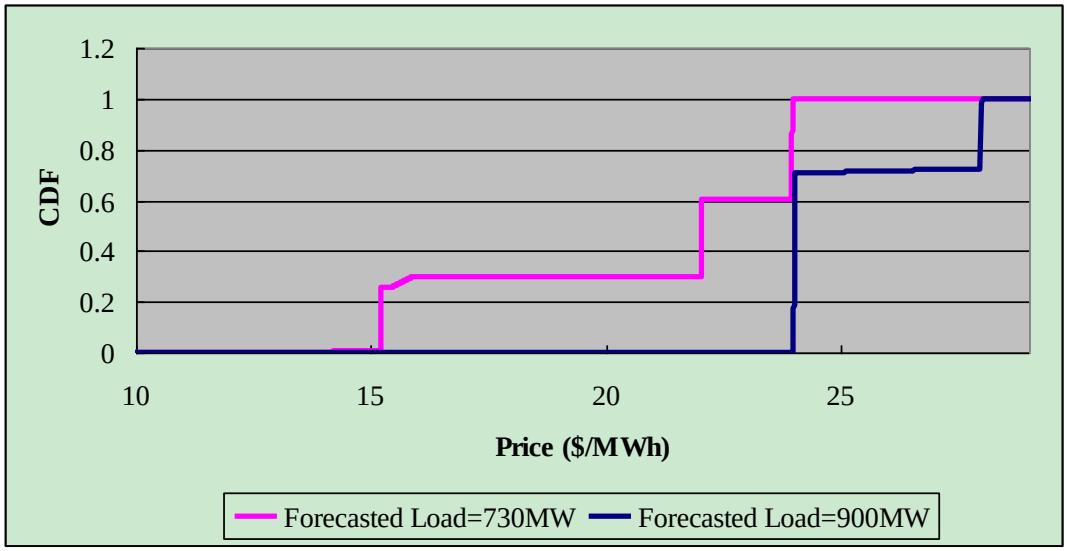


Figure 5.26. Cumulative density function of the probabilistic LMP at Bus B for two forecasted load levels

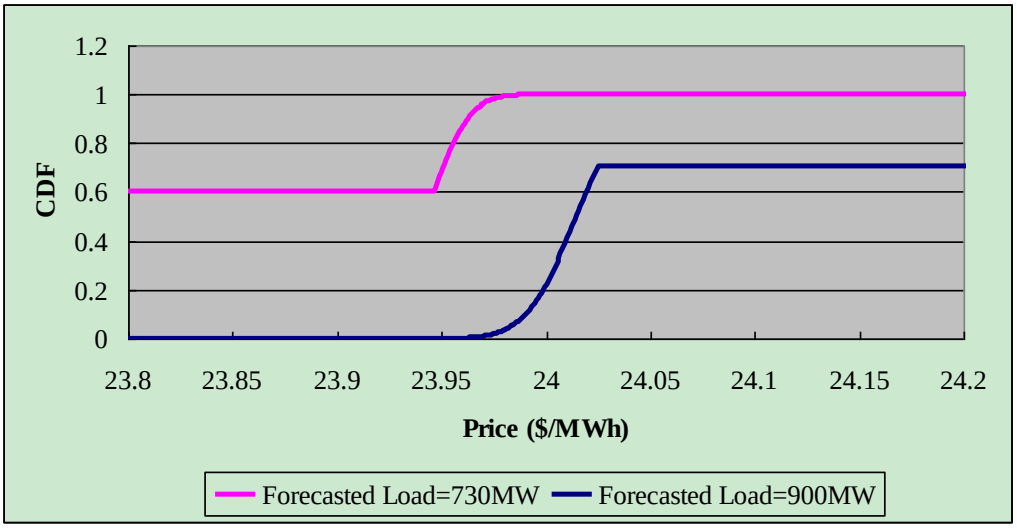


Figure 5.27. Cumulative density function of the probabilistic LMP at Bus B for two forecasted load levels in the price interval 23.95~24.02 $/MWh

5.3.4.3. Probability Density Function of Probabilistic LMP

The PDF curve of the probabilistic LMP at Bus B is shown in Figure 5.28 for the same two forecasted load levels. When the forecasted load is 730 MW, the probability density function of the probabilistic LMP is mainly scattered in three price intervals: 15.19~15.23 $/MWh, 22.01~22.05 $/MWh, and 23.95~24.02 $/MWh while 23.95~24.02 $/MWh and 27.94~28.02 $/MWh are the two price intervals with a high probability density for the forecasted load at 900MW. The probability density is numerically zero for almost everywhere else outside these price intervals. Furthermore, these price intervals are consistent with those where the CDF values have a jump, as seen in Figure 5.26.

It should be noted that the vertical bars in Figure 5.28 are actually smooth curves which are not legible due to scaling issues. A well scaled graph is shown in Figure 5.29. It can be seen from Figure 5.29 that the probability density functions are continuous, and differentiable, curves in the 23.95~24.02 $/MWh range.

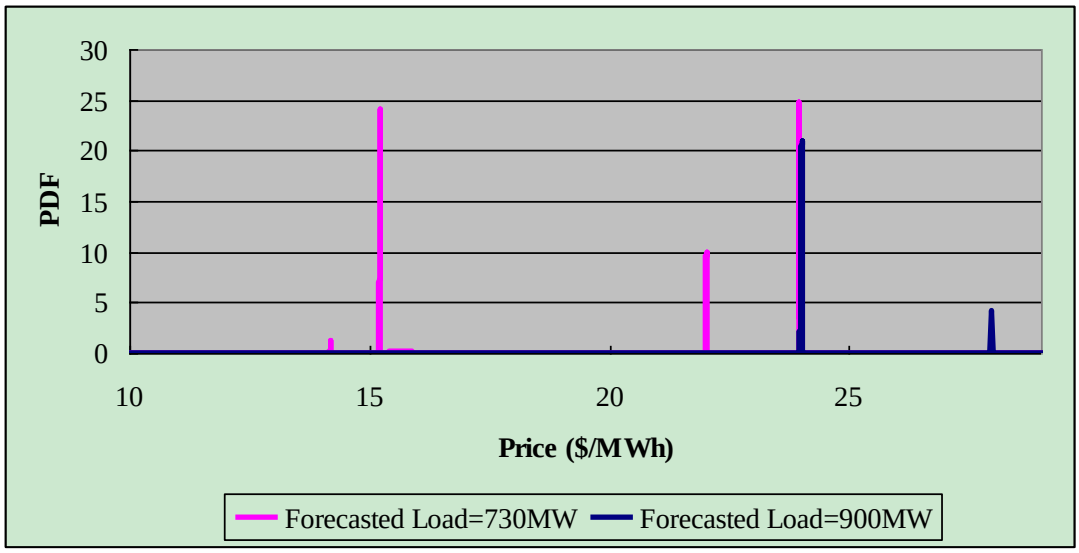


Figure 5.28. Probability density function of the probabilistic LMP at Bus B for two forecasted load levels

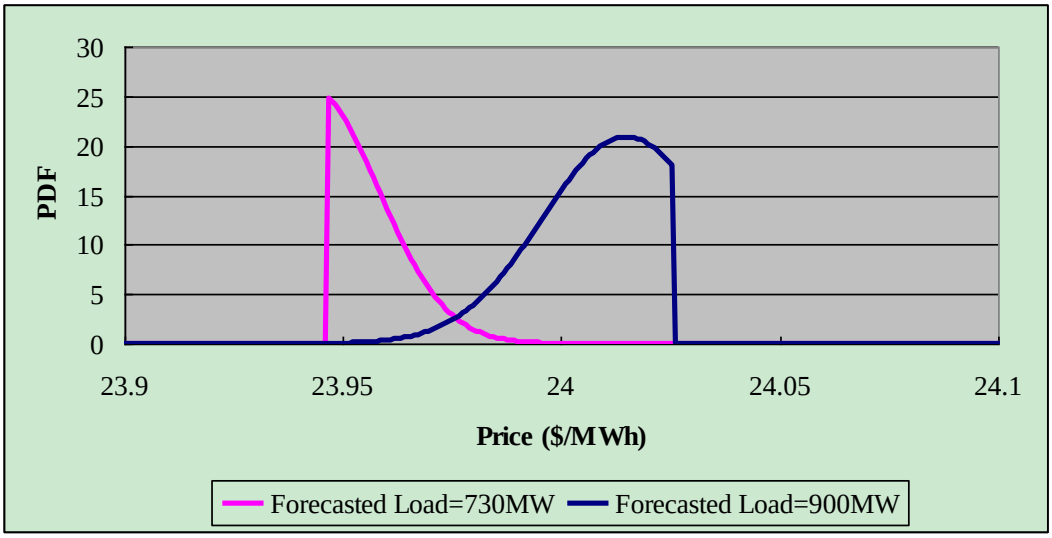


Figure 5.29. Probability density function of the probabilistic LMP at Bus B for two forecasted load levels in the price interval 23.95~24.02 $/MWh

Since the probability of any single price is zero, it is more useful to divide the entire price range into a few intervals and investigate the probability of an actual LMP falling into each interval. The vertical bars observed in the PDF curves, such as in Figure 5.28, can be used to help make this classification. In practice, the categorization is at the discretion of the decision maker and can vary with cases and purposes.

The probability of the LMP_t at Bus B falling into the selected price intervals is calculated and shown in Table 5.5 for two representing forecasted load levels, 730MW and 900MW. The same results are presented as a pie chart in Fig. 5.30. The results discover the fact that the deterministic LMP with respect to D_t^F may or may not fall into the price interval with the highest probability. For example, when the forecasted load is 900MW, the corresponding deterministic LMP is $24.01/MWh and its close neighborhood $23.9~24.1/MWh has the highest probability of 70.87%. However, the close neighborhood $22.0~22.1/MWh of the deterministic LMP $22.03/MWh for the forecasted load 730MW has only the second highest probability of 30.88%, less than the probability of 39.42% for the price interval

$23.9~24.1/MWh. It shows that the deterministic LMP associated with the mean value of the actual load does not necessarily bear the largest probability.

Table 5.5. Probability of LMP_t at Bus B in the Selected Price Intervals

LMP Range	Probability(%) when D^F_t=730MW	Probability(%) when D^F_t =900MW
0~15.0 $/MWh	0.48	0.00
15.0~15.3 $/MWh	25.11	0.00
15.3~22.0 $/MWh	4.12	0.00
22.0~22.1 $/MWh	30.88	0.02
22.1~23.9 $/MWh	0.00	0.00
23.9~24.1 $/MWh	39.42	70.87
24.1~27.9 $/MWh	0.00	1.52
27.9~28.0 $/MWh	0.00	27.53
28.0~2000 $/MWh	0.00	0.07
Total	100	100

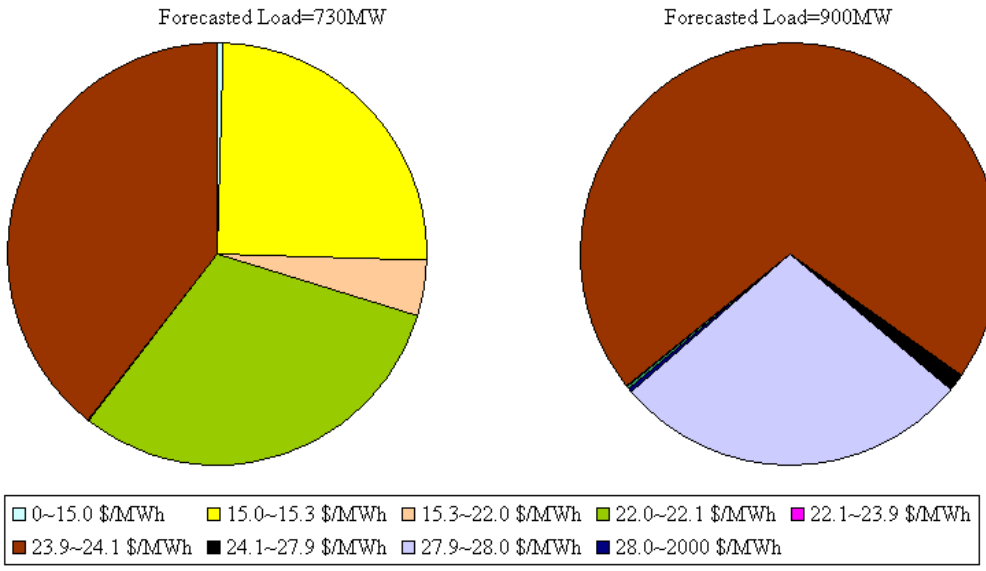


Figure 5.30. Probability of LMP_t at Bus B for the selected price intervals

Table 5.5 and Fig. 5.30 reveal the likelihood of realizing the forecasted LMP and its close vicinity, and therefore, can be very useful for buyer and sellers in making their financial decisions, such as developing bidding strategies.

5.3.4.4. Alignment Probability of Probabilistic LMP

Fig. 5.31 shows the curve of the alignment probability of the deterministic LMP at Bus B versus the forecasted load, with a 10% and 20% price tolerance, respectively. By making a comparison between Fig. 5.31 and Fig. 5.17, we can see that the low probabilities occur near the CLLs. For instance, the lowest probability is 51.44% at the forecasted load level 924.21MW, which is very close to the CLL at 924.75MW. When the forecasted load is over 1300MW, the probability keeps decreasing as the forecasted load is approaching the maximum level (i.e., price of the VOLL) that the system can afford, namely, 1467MW.

As shown in Figure 5.31, the alignment probability curve is closer to 1.0 with a higher price tolerance, and the valley at around 924.21MW disappears when the price tolerance is

20%. For example, the deterministic LMP is $24.03/MWh at the load level 924.21MW, and the prices in the close neighborhood of 924.21MW differ by up to $4.01/MWh, which is less than 20% of the deterministic LMP. Therefore, the alignment probability at the load level 924.21MW increases from 51.44% with a 10% price tolerance, to nearly 100% with a doubled price tolerance.

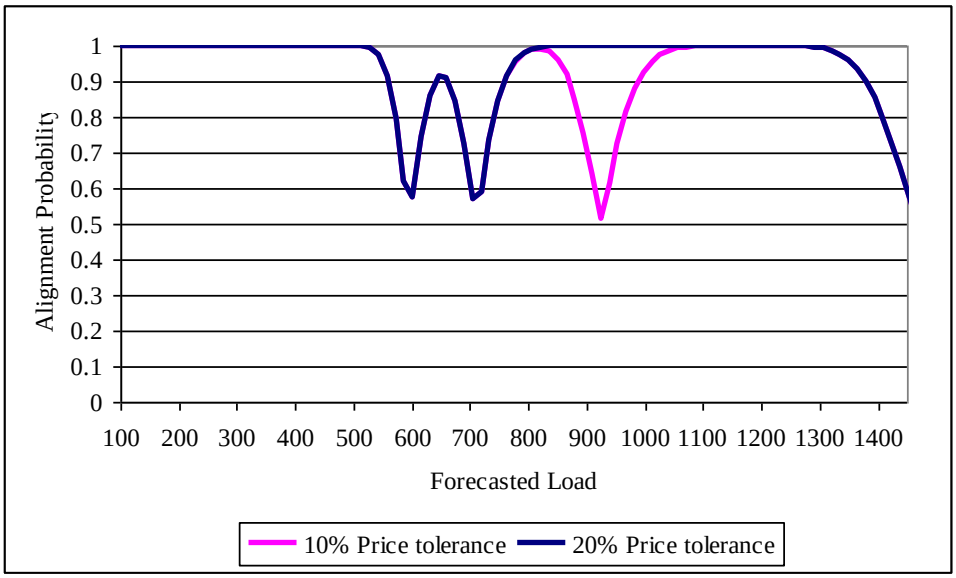


Figure 5.31. Alignment probability of the deterministic LMP at Bus B versus the forecasted load, with a 10% and 20% price tolerance, respectively

5.3.4.5. Expected Value of Probabilistic LMP

The expected value of the probabilistic LMP for the above case is compared with the deterministic LMP, $y(D_t^F)$, which is shown in Table 5.6. It shows that the expected value of probabilistic LMP may differ from the deterministic LMP for a specific forecasted load.

The expected value of the probabilistic LMP versus the forecasted load curve is shown in Fig. 5.32. A load range beyond 1300MW is not shown simply because the high VOLL will make the curve hard to scale for a good illustration. It should be noted that the expected LMP will escalate sharply when the load is close to the maximum load level the system can afford, and will eventually reach 2000 $/MWh.

Table 5.6. Expected value of the probabilistic LMP in comparison with the Deterministic LMP for Bus B

D^F_t(MW)	Expected Value of Probabilistic LMP($/MWh)	Deterministic LMP($/MWh)
730	20.78	22.03
900	25.13	24.01

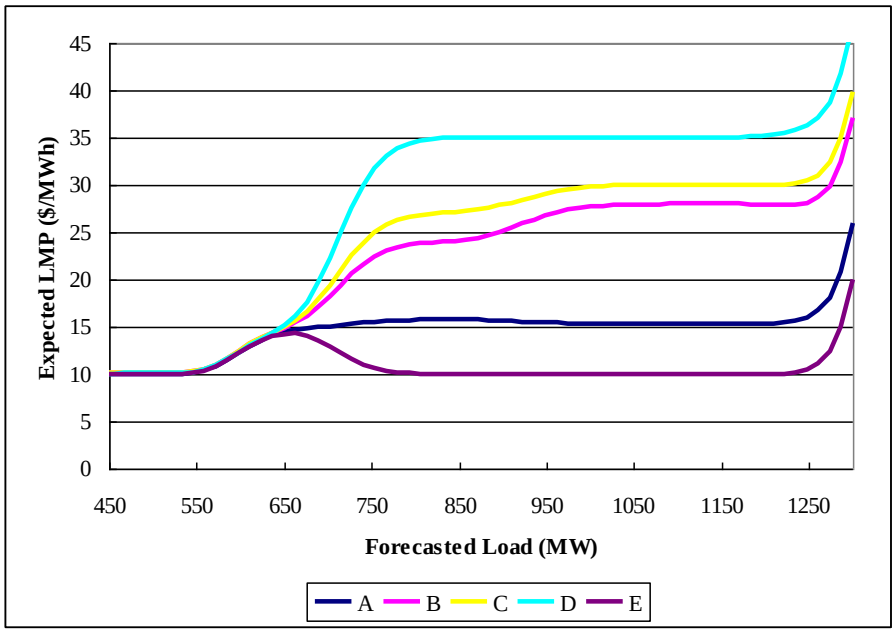


Figure 5.32. Expected value of probabilistic LMP versus forecasted load

In the deterministic LMP-Load curve in Fig. 5.17, the sensitivity for Bus B at 924.75MW is mathematically infinite since a step change occurs at this load level. In the probability-based LMP-Load curve, the upper and lower bounds of the sensitivity can be estimated using (5.62) and (5.63)

$$-18.70 \le \frac{\partial E_{LMP_t}(\mu_t, \sigma_t)}{\partial \mu_t} \le 18.50\$/MWh^2 .$$

When contrasted with the deterministic LMP-Load curve in Fig. 5.17, the curve of the expected value of the probabilistic LMP in Fig. 5.32 demonstrates the same overall trend. However, Fig. 5.32 shows a much smoother curve without any step changes, which contributes to the reduction of price uncertainty, especially around the CLLs.

Also shown in this probabilistic LMP forecasting figure is that when the load is closer to the CLLs, price uncertainty, i.e., the uncertainty associated with the forecasted deterministic LMP, will be higher. This matches the overall trend in the deterministic LMP in Fig. 5.17.

5.3.4.6. Impact of Load Forecasting Accuracy

In this section, three different levels of standard deviation of load forecasting are examined, 5%, 3%, and 1%. Fig. 5.33 shows the probabilities of the random variable LMP_t at Bus B falling into a few price ranges for these three levels of standard deviation when the system load is 730MW. It can be seen from Fig. 5.33 that the probability of realizing the actual price in the range of 22.0~22.1 $/MWh where the deterministic LMP $22.03/MWh falls into, increases considerably with a smaller standard deviation. This is reasonable because more a accurate load forecast should lead to less deviation in the forecasted price.

Fig. 5.34 compares the expected value of the probabilistic LMP curves at the same bus. When the forecasted load is at a large distance from any CLL, for example at 850MW, the three curves overlap very well. This suggests that different levels of the standard deviation make minimal differences on the expected LMP at this load level. In addition, the sensitivity of the expected LMP at this load level is small, which indicates the expected LMP remains nearly constant when the forecasted load varies slightly around this level. In contrast, when the forecasted load is close to a CLL, for example at 595.80MW, the lower the standard deviation is, the closer the curve is to a step change curve shape. Furthermore, the inset in Fig. 5.34 shows that when the load level is closer to a CLL, the absolute value of the sensitivity of the expected LMP grows rapidly, and the expected LMP becomes more sensitive to variations of the forecasted load.

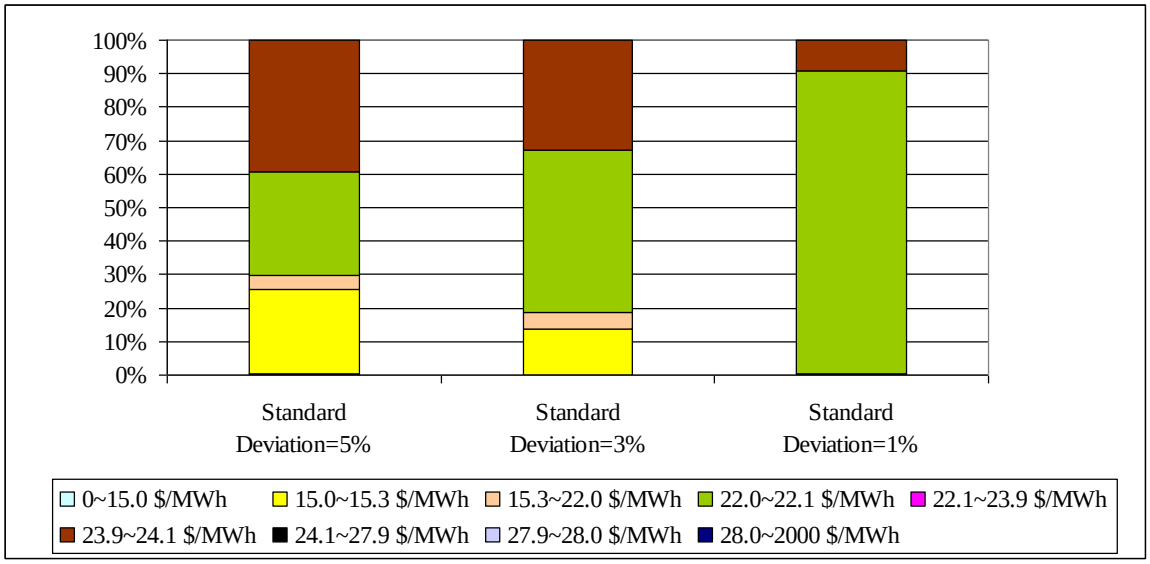


Figure 5.33. Probability of LMP_t at Bus B in a few price ranges for three levels of standard deviation when the system load is 730MW

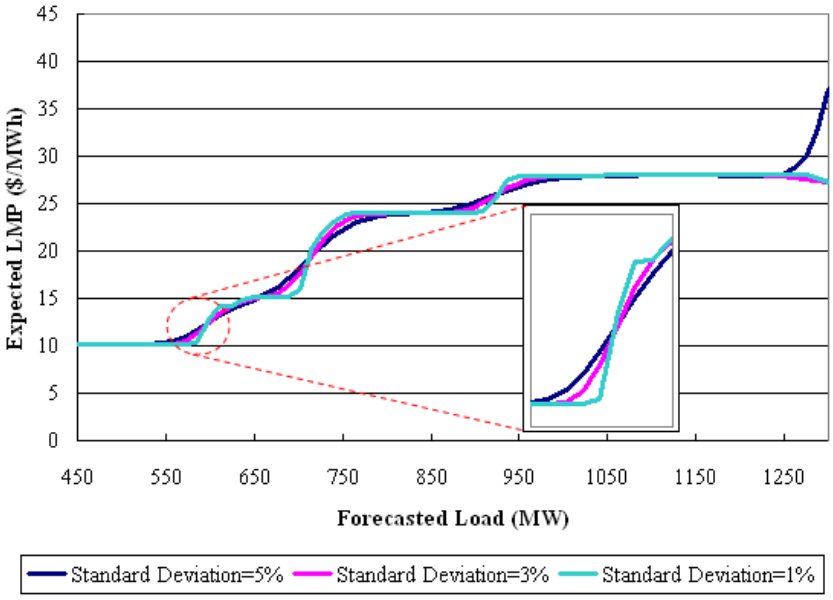


Figure 5.34. Expected value of the probabilistic LMP at Bus B versus the forecasted load for three levels of standard deviation

5.3.5. Discussions and Conclusions

A methodology similar to that in the lossless DCOPF framework has been employed for studying probabilistic LMP in the ACOPF framework. Similar to what has been done in Section 5.2, concepts such as the alignment probability and the expected value of probabilistic LMP and its sensitivity are presented. And, mathematical formulas are derived, including the upper and lower bounds of the sensitivity of the expected value of the probabilistic LMP. The proposed concepts and formulas are exemplified and verified with a case study on a modified PJM 5-bus system.

5.3.5.1. Differences between Probabilistic LMP in the ACOPF framework and that in the Lossless DCOPF framework

Although the LMP versus load curves for ACOPF and lossless DCOPF look alike as shown in Figure 5.1 and 5.17, they differ in a few aspects.

First, the locations where LMP exhibits a step change may be significantly different. Table 5.7 compares the CLLs for ACOPF and lossless DCOPF for the modified PJM 5-bus system studied in Section 5.2.4. Most of the CLLs for ACOPF have been identified by lossless DCOPF with an acceptable accuracy, however, the CLL at 1299.6MW is quite distant from the estimated CLL 1137.02MW in the lossless DCOPF framework.

Table 5.7. CLLs for ACOPF and Lossless DCOPF for a modified PJM 5-bus system

CLL for ACOPF (MW)	**CLL for Lossless DCOPF (MW)**
595.8	600.00
635.4	640.00
706.05	N/A
710.55	711.81
739.8	742.80
924.75	963.94
927	N/A
1299.6	**1137.02**
1467	1484.06

Note: N/A represents no CLL

Second, the LMPs for ACOPF framework and lossless DCOPF framework can be quite close in most cases, yet may be quite different at load levels close to a CLL. For instance, when load is 900MW, the corresponding LMP at Bus B is $24.01/MWh on LMP versus load curve for ACOPF framework while the LMP is $23.68/MWh for the lossless DCOPF framework, close to $24.01/MWh. In contrast, when load level is 930MW, the LMPs for the two frameworks are considerably different, which are $27.96/MWh and $24.01/MWh, respectively.

Third, the price between two adjacent CLLs in the LMP versus the Load curve for the ACOPF framework is not a constant value; rather, it is typically a steadily and slightly increasing or decreasing curve. In the lossless DCOPF framework, LMP always remains as a constant within each segment of LMP versus load curve.

The aforementioned differences in the LMP versus load curves result in the following major differences in the probabilistic LMP studies.

- The price between two adjacent CLLs in the LMP versus Load curve can be modeled by a linear polynomial, instead of a constant value.
- The Probabilistic LMP in the ACOPF framework is a piece-wise continuous random variable, rather than a discrete random variable.
- The cumulative density function and probability density function are derived and shown to be differentiable at almost everywhere except for the prices at the CLLs, namely, the boundary prices of each segment of the LMP versus Load curve.
- The probability associated with the deterministic LMP is not meaningful due to the continuous feature of the probabilistic LMP. Instead, the probability is used to reflect the likelihood of the actual price falling into a range of prices. This is why it is more reasonable to define alignment probability with an acceptable tolerance.

In addition, some noteworthy differences between numeric results of probabilistic LMP in the lossless DCOPF and ACOPF frameworks could be recognized with a careful examination.

As shown in Tables 5.2-5.3 in the case study section for lossless DCOPF on the modified PJM 5-bus system, when forecasted load is 900MW, the resultant deterministic forecasted LMP is $23.68/MWh, and its corresponding probability is 92.21%. The expected value of LMP is $24.03/MWh. In contrast, as shown in Tables 5.5-5.6 for ACOPF, the deterministic LMP is $24.01/MWh, which is close to that in the lossless DCOPF framework; however, the probability of the actual LMP falling into a close vicinity of the deterministic LMP is 70.87%, considerably less than 92.21%, and the expected value of LMP is $25.13/MWh, about 4% greater than its counterpart for lossless DCOPF.

These differences result from the different CLLs identified in both frameworks. For instance, in the lossless DCOPF framework, as shown in Tables 5.2 and 5.7, LMP remains

constant as $23.68/MWh when load varies from 900MW to 963.94MW and jumps to $28.18/MWh at the CLL of 963.94MW. In the ACOPF framework, the LMP step change occurs at the CLL of 924.75MW, much closer to the forecasted load 900MW than 963.94MW, which leads to less probability for price range $23.9/MWh ~ $24.1/MWh and greater expected value of LMP at the deterministic LMP $24.01/MWh. In fact, the differences will become more significant when the forecasted load is closer to the CLL at 924.75MW.

For each CLL identified in the ACOPF framework, its counterpart in lossless DCOPF framework could not be accurate because the lossless DCOPF is a simplified model with losses ignored. Therefore, the differences in probabilistic LMP as illustrated above will always be expected, especially when the forecasted load is close to CLLs of which the lossless DCOPF framework fails to produce a good estimation.

5.3.5.2. Connections between Probabilistic LMP in the ACOPF framework and that in the Lossless DCOPF framework

Despite the differences discussed in the previous section, probabilistic LMP in the ACOPF framework does have connections with that in the lossless DCOPF framework, as implied in the similarities observed in case study sections 5.2.4 and 5.3.4.

As shown in Table 5.2 for the lossless DCOPF framework, the three most significant probabilities are 30.23%, 32.80% and 36.29%, corresponding to LMPs at $15/MWh, $21.74/MWh and $23.68/MWh, respectively, when forecasted load is 730MW. The results for the ACOPF framework have been shown in Table 5.5, where the three most significant probabilities are 25.11%, 30.88%, and 39.42% for price ranges $15.0/MWh ~ $15.3/MWh, $22.0/MWh ~ $22.1/MWh, and $23.9/MWh ~$24.1/MWh, respectively. It suggests that

probabilistic LMP may demonstrate similar patterns in the ACOPF framework as those in the lossless DCOPF framework, which will be analyzed carefully as follows.

The LMP versus load curve for the ACOPF framework can be considered as a piece-wise linear curve and the LMP between two consecutive CLLs slightly increases or decreases. In most of the cases, the non-constant characteristic of each LMP segment results from the power losses. The slope of the LMP slight change is typically so small for a high voltage power transmission system that it is hard to be visually perceived in the LMP versus load curve, unless specifically scaled as shown in Figure 5.17. With the consideration of step changes at CLLs, it implies the possible values of random variable LMP_t reside in a few narrow price ranges, denoted by $\{[y(D_i), y(D_{i+1})), i=1,\ldots,n-1\}$, where it is assumed that $y(D_i) \le y(D_{i+1})$ for notational convenience.

The probability that actual LMP falls into $[y(D_i), y(D_{i+1}))$ in the ACOPF framework is defined as

$$\Pr(LMP_t \in [y(D_i), y(D_{i+1}))) = F_{LMP_t}(y(D_{i+1})) - F_{LMP_t}(y(D_i)) \tag{5.64}$$

For simplicity, we assume that the price ranges $\{[y(D_i), y(D_{i+1})), i=1,\ldots,n-1\}$ are mutually exclusive, namely,

$$[y(D_i), y(D_{i+1})) \cap [y(D_j), y(D_{j+1})) = \varnothing, \quad i \neq j \tag{5.65}$$

Therefore, substituting equation (5.31) into (5.64) yields

$$\Pr(LMP_t \in [y(D_i), y(D_{i+1}))) = \int_{D_i}^{D_{i+1}} \varphi(u)du = \Phi(D_{i+1}) - \Phi(D_i) \tag{5.66}$$

For easy comparison, equation (5.6) for lossless DCOPF framework is rewritten as follows.

$$\Pr(LMP_t = p_i) = \int_{\tilde{D}_i}^{\tilde{D}_{i+1}} \varphi(u)du = \Phi(\tilde{D}_{i+1}) - \Phi(\tilde{D}_i) \tag{5.67}$$

where $\tilde{D}_i$ and $\tilde{D}_{i+1}$ represent the load levels at the two sides of the LMP segment that has the value p_i. In fact, $\tilde{D}_i$ and $\tilde{D}_{i+1}$ are two CLLs in the lossless DCOPF framework.

It should be noted that equation (5.66) represents the probability of the probabilistic LMP for ACOPF while equation (5.67) denotes the probability of the probabilistic LMP for the lossless DCOPF framework. It should also be noted that the deterministic LMP p_i in lossless DCOPF is typically close to the price range $(y(D_i), y(D_{i+1}))$, as we have seen earlier in this section, except when the forecasted load is close to a CLL where lossless DCOPF generates a significantly different estimation.

If the two CLLs (D_i, D_{i+1}) identified in ACOPF are exactly the same as those $(\tilde{D}_i, \tilde{D}_{i+1})$ in lossless DCOPF, we can conclude

$$\Pr(LMP_t \in (y(D_i), y(D_{i+1}))) = \Pr(LMP_t = p_i) \quad (5.68)$$

Normally, the CLLs identified in ACOPF and lossless DCOPF are not exactly the same, but are typically close as illustrated in Table 5.7. Therefore, in most cases we expect

$$\Pr(LMP_t \in (y(D_i), y(D_{i+1}))) \approx \Pr(LMP_t = p_i) \quad (5.69)$$

When one of the CLLs (D_i, D_{i+1}) is significantly different from its counterpart in $(\tilde{D}_i, \tilde{D}_{i+1})$, as has been illustrated and discussed in the previous section, equation (5.69) will not hold true. An exceptional case is when the forecasted load is adequately distant from the CLLs (for instance, larger than $3*\delta_t$), and consequently the inaccurate CLLs may lead to trivial differences in the probabilities calculated by (5.66) and (5.67).

Furthermore, price range $(y(D_i), y(D_{i+1}))$ is normally so narrow that the price difference $y(D_{i+1}) - y(D_i)$ is not comparable to $y(D_i)$ in magnitude. Suppose $\frac{y(D_{i+1}) - y(D_i)}{y(D_i)} = \alpha\%$ and therefore we have

$$\left[\tilde{p}\times(1-\alpha\%),\tilde{p}\times(1+\alpha\%)\right]\supseteq\left[y(D_i),y(D_{i+1})\right] \tag{5.70}$$

where $\tilde{p}$ is the deterministically forecasted LMP.

For simplicity, we assume the price interval $\left[y(D_i),y(D_{i+1})\right]$ has no intersection with any other price intervals $\left[y(D_j),y(D_{j+1})\right]_{j=0,j\neq i}^{n}$. Therefore, the alignment probability can be derived as follows

$$\begin{aligned}\mathbf{AP}_{\alpha} &= \Pr\left[\tilde{p}\times(1-\alpha\%)\le LMP_t\le\tilde{p}\times(1+\alpha\%)\right]\\ &= \Pr\left[y(D_i)\le LMP_t\le y(D_{i+1})\right]\\ &= \int_{D_i}^{D_{i+1}}\varphi(u)du\end{aligned} \tag{5.71}$$

Equation (5.71) implies that the alignment probability from the ACOPF framework is the same as that from the lossless DCOPF framework when the CLLs identified in both frameworks are exactly the same. In practice, the CLLs can be close, and consequently the alignment probabilities will be close. It may still be true when the assumption of no intersection of price ranges is relaxed, as long as the probabilities on the same price ranges are lumped together, which matches the observations in Figures 5.8 and 5.31.

In summary, the probabilistic LMP by the lossless DCOPF framework could serve as a good estimation of the probabilistic LMP by the ACOPF framework as long as the CLLs identified in the lossless DCOPF framework is sufficiently close to those in the ACOPF framework.

5.4. Probabilistic LMP Forecasting for FND-based DCOPF

As discussed in Chapter 3, the FND-based DCOPF model is superior to the lossless DCOPF model mainly due to its modeling of power losses. With the loss considered, the corresponding LMP between the two adjacent CLLs normally is not a constant value; instead, as shown in the LMP sensitivity study in Section 3.3.3, the LMP demonstrates a clear linear

pattern between the adjacent CLLs. Therefore, the LMP versus Load curve of the FND-based DCOPF framework has the same characteristics as those of the ACOPF framework. Hence, exactly same methodology, including the concepts, formulas, and conclusions, for the probabilistic LMP in the ACOPF framework, as presented in Section 5.3, can be directly applied to study the probabilistic LMP forecasting for the FND-based DCOPF framework. It is, therefore, not repeated in this section. Only a case study on a modified PJM 5-bus system will be presented.

5.4.1. Numerical Study of a Modified PJM 5-Bus System

In this section, a numeric study will be performed on the PJM 5-Bus system [11], with slight modifications. The modifications are for illustration purposes and are detailed in Section 4.2.4. The configuration of the system is shown in Figure 4.3. Again, it is assumed that the system load change is distributed to each bus load proportional to its base case load for simplicity. The resulting LMP versus load curve is shown in Figure 5.35.

For study on probabilistic LMP, it is assumed that μ_t is always equal to the forecasted load D_t^F, and the standard deviation σ_t is taken as 5% of the mean μ_t unless otherwise stated. The VOLL is set at \$2000/MWh.

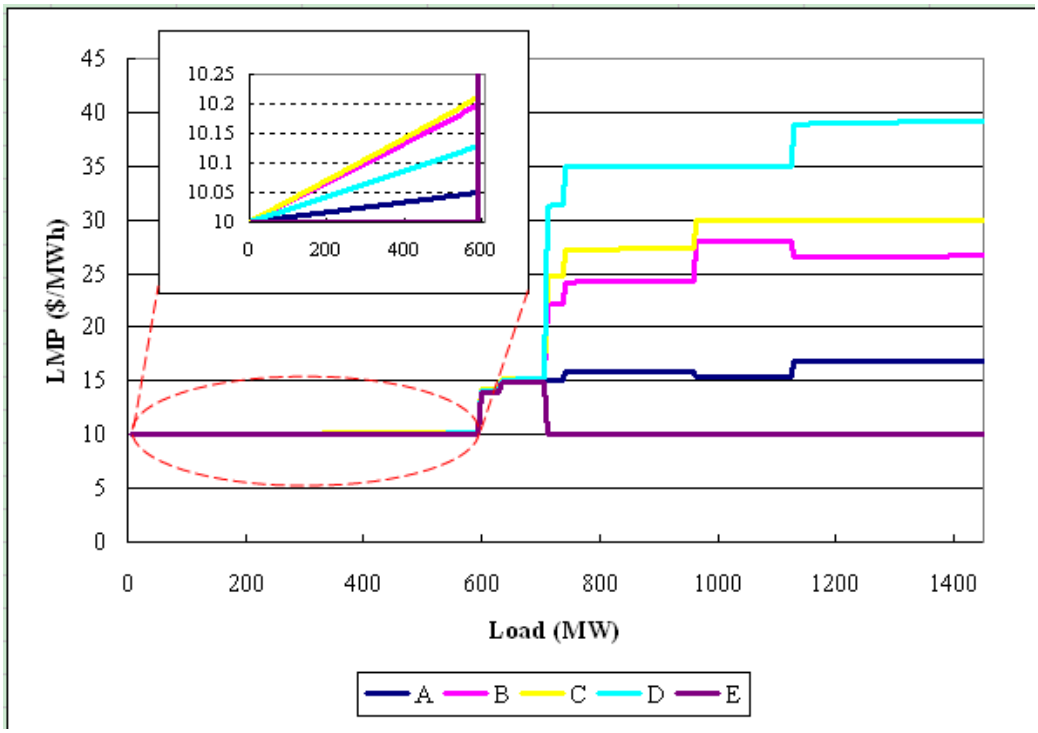


Figure 5.35. LMP versus load curve for FND-based DCOPF model

5.4.1.1. Approximation of LMP Curve

A linear polynomial curve-fitting is employed to approximate the actual LMP between every two adjacent CLLs, and the coefficients are used to establish the mathematical model for the LMP versus load curve. Table 5.8 shows the curve-fitting coefficients for the LMP curves at all buses when the load is within [0, 590] MW. It implies that the LMPs at all buses, except Bus E, increase slightly, while the LMP at Bus E remain 10$/MWh for the entire load interval.

Table 5.8. Curve-fitting coefficients for the LMP curves at all buses when the load is within [0, 590] MW

	a ($/(MWh*MW))	b ($/MWh)
LMP@A	0.0001	9.9998
LMP@B	0.0003	9.9994
LMP@C	0.0004	9.9994
LMP@D	0.0002	9.9995
LMP@E	0.0000	10.0000

With the linear polynomial coefficients obtained through the curve-fitting, the mathematical LMP versus load model is established, which is a piece-wise linear curve with step changes at the CLLs. The mathematical representation of the curve is a very good approximation to the actual LMP versus load curve, and therefore, can be used to facilitate an analytical study on topics such as CDF, PDF, etc. The curve represented by the mathematical formula looks almost identical to the actual curve shown in Figure 5.35. For example, Figure 5.36 shows the actual LMP curve at Bus B and its approximation curve through linear polynomial curve-fitting overlap very well. In fact, when the two curves are compared at the sampled load levels with a step of 0.005 p.u. load, the largest difference is less than 0.01$/MWh, approximately 0.1% of the lowest LMP, $10/MWh.

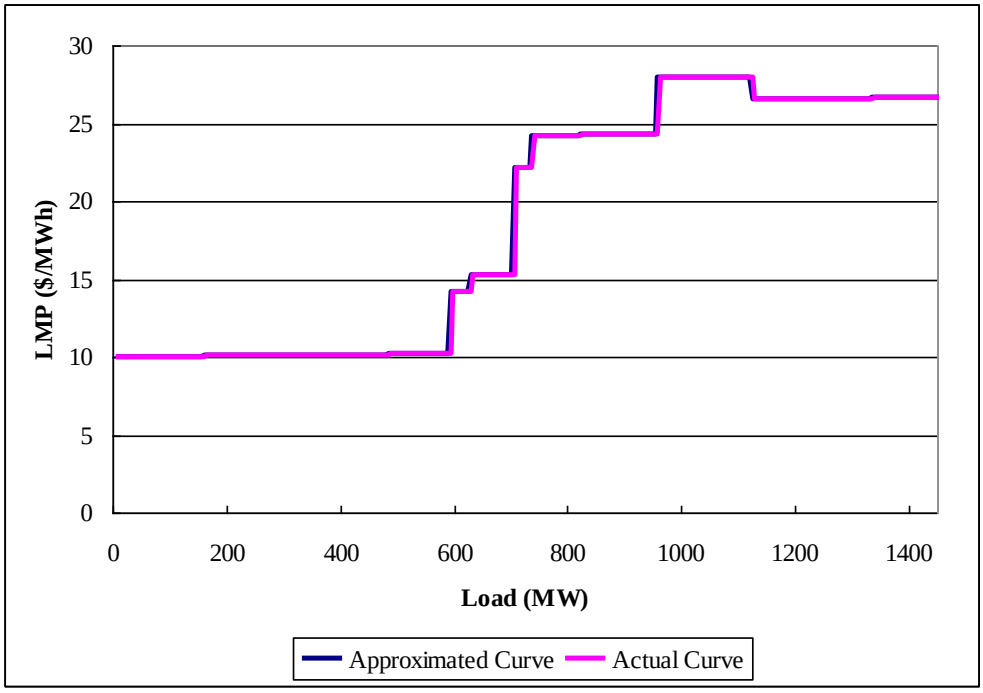


Figure 5.36. Comparison of actual LMP versus load curve and its approximation through linear polynomial curve-fitting

5.4.1.2. Cumulative Density Function of Probabilistic LMP

Figure 5.37 shows the CDF curve of the probabilistic LMP at Bus B for the forecasted load at 730 MW and 900 MW, respectively. The figure suggests the staircase pattern of the CDF curve. Combined with the LMP versus Load curve such as in Figure 5.35, it can be seen that the prices at which a step change occurs coincide with the price intervals near the forecasted load level for Bus B. The corresponding PDF values for these prices are expected to be higher than the PDF values of other prices, as will be verified in the next section.

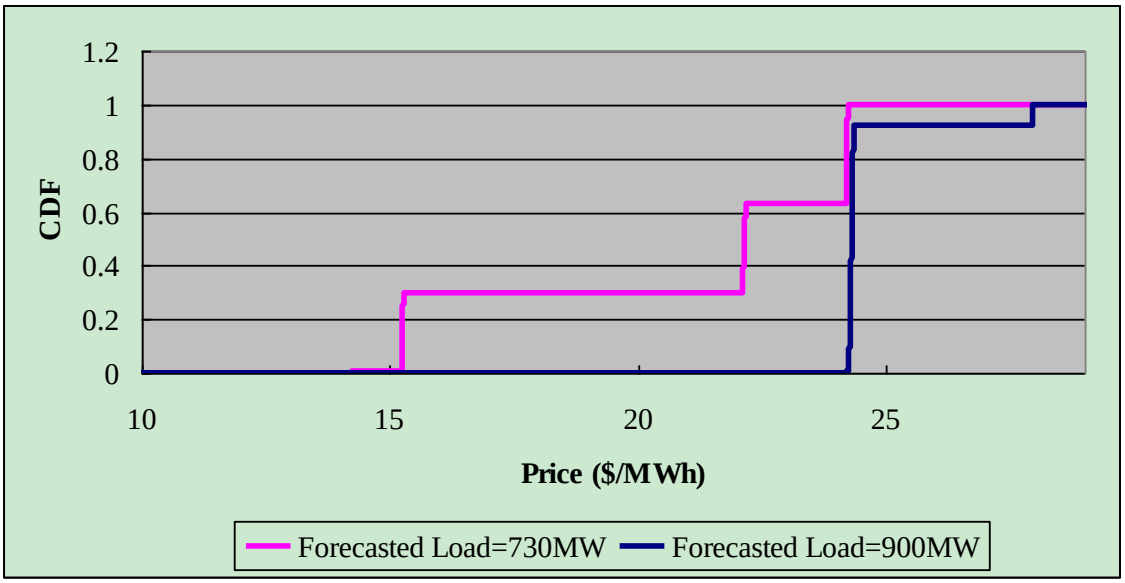


Figure 5.37. Cumulative density function of the probabilistic LMP at Bus B for two forecasted load levels

A careful study reveals that the majority of the step changes observed in the CDF curve is not really step changes. Figure 5.38 redraws the same curves in Figure 5.37 in a narrow range around \$24/MWh, where a step change appears. It can be seen that both CDF curves move smoothly from \$24.10/MWh to \$24.40/MWh. It is actually consistent with the aforementioned theoretical part in Section 5.3.2.3, where the CDF function is shown to be differentiable almost everywhere except at the price boundaries of each interval of the LMP versus the Load curve, namely, $\{y(D_i)\}_{i=1}^{n}$. Nevertheless, the change of the CDF values happens in such narrow price intervals that it looks just like a step change when plotted for a broader range of prices.

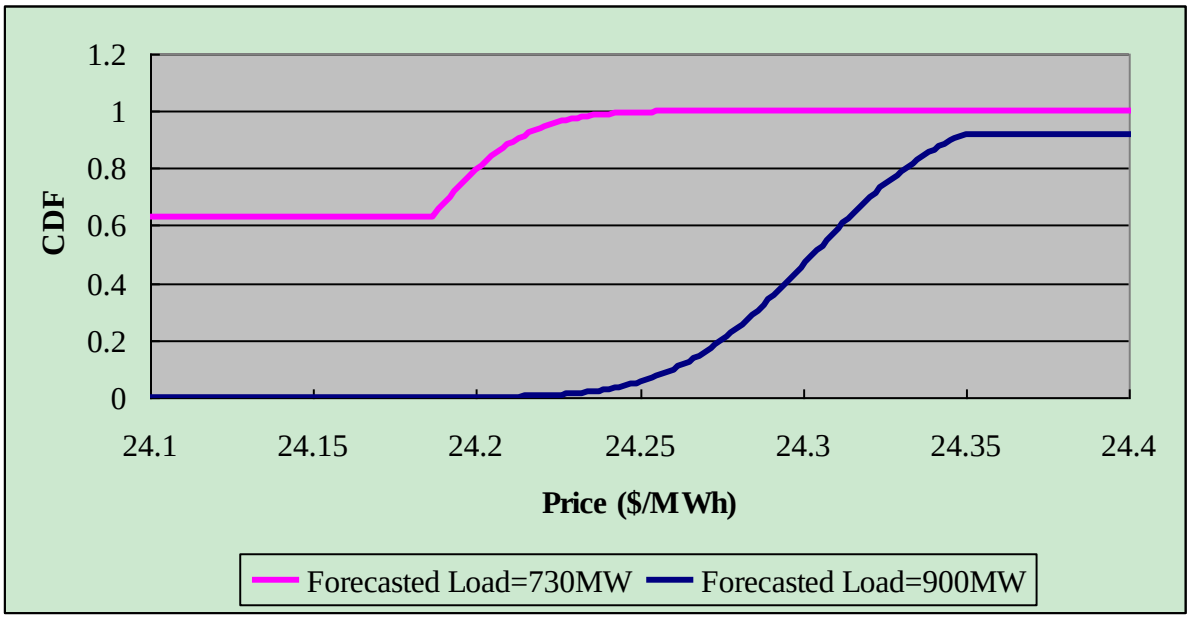


Figure 5.38. Cumulative density function of the probabilistic LMP at Bus B for two forecasted load levels in the price interval 24.10~24.40 $/MWh

5.4.1.3. Probability Density Function of Probabilistic LMP

The PDF curve of the probabilistic LMP at Bus B is shown in Figure 5.39 for the same two forecasted load levels. When the forecasted load is 730 MW, the probability density of the probabilistic LMP is mainly scattered in three price intervals: 15.23~15.28 $/MWh, 22.10~22.20 $/MWh, and 24.10~24.40 $/MWh, while 24.10~24.40 $/MWh and 27.94~27.98 $/MWh are the two price intervals with a high probability density for the forecasted load at 900MW. The probability density is numerically zero for almost everywhere else outside these price intervals. Furthermore, these price intervals are consistent with those where the CDF values have a jump, as seen in Figure 5.37.

It should be noted that the vertical bars in Figure 5.39 are actually smooth curves which are not legible due to scaling issues. A well scaled graph is shown in Figure 5.40. It can be

seen from Figure 5.40 that the probability density functions are continuous and differentiable curves in the 24.18~24.34 $/MWh range.

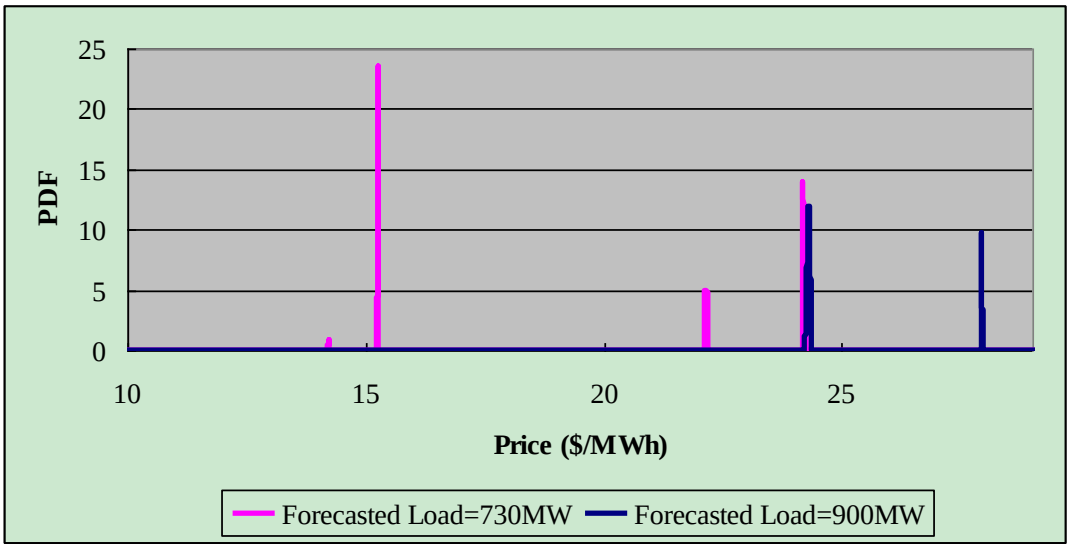


Figure 5.39. Probability density function of the probabilistic LMP at Bus B for two forecasted load levels

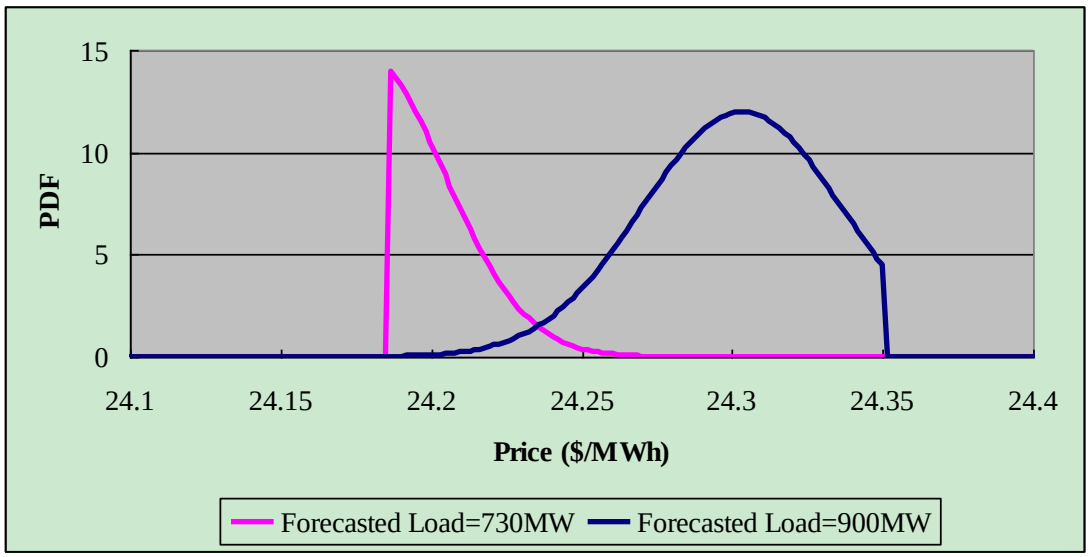


Figure 5.40. Probability density function of the probabilistic LMP at Bus B for two forecasted load levels in the price interval 24.10~24.40 $/MWh

Since the probability of any single price is zero, it is more useful to divide the entire price range into a few intervals and investigate the probability of an actual LMP falling into each interval. The vertical bars observed in the PDF curves, such as in Figure 5.39, can be used to help make this classification. In practice, the categorization is at the discretion of the decision maker and can vary with cases and purposes.

The probability of the LMP_t at Bus B falling into the selected price intervals is calculated and shown in Table 5.9 for two representing forecasted load levels, 730MW and 900MW. The same results are presented as a pie chart in Fig. 5.41. The results discover the fact that the deterministic LMP with respect to D_t^F may or may not fall into the price interval with the highest probability. For example, when the forecasted load is 900MW, the corresponding deterministic LMP is \$24.30/MWh and its close neighborhood \$24.1~24.4/MWh has the highest probability of 91.90%. However, the close neighborhood \$22.1~22.2/MWh of the deterministic LMP \$22.14/MWh for the forecasted load 730MW has only the second highest probability of 33.27%, less than the probability of 36.60% for the price interval \$24.10~24.40/MWh. It shows that the deterministic LMP associated with the mean value of the actual load does not necessarily bear the largest probability.

Table 5.9. Probability of LMP_t at Bus B in the Selected Price Intervals

LMP Range	Probability(%) when D^F_t=730MW	Probability(%) when D^F_t =900MW
0~15.0 $/MWh	0.44	0.00
15.0~15.3 $/MWh	29.69	0.00
15.3~22.1 $/MWh	0.00	0.00
22.1~22.2 $/MWh	33.27	0.02
22.2~24.1 $/MWh	0.00	0.00
24.1~24.4 $/MWh	36.60	91.90
24.4~27.9 $/MWh	0.00	0.00
27.9~28.0 $/MWh	0.00	8.08
28.0~2000 $/MWh	0.00	0.00
Total	100	100

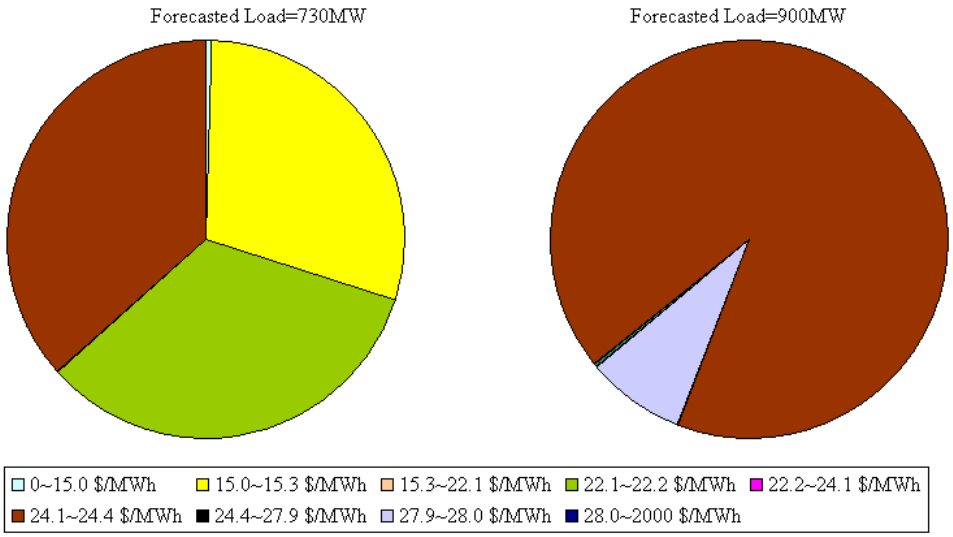


Figure 5.41. Probability of LMP_t at Bus B for the selected price intervals

Table 5.9 and Fig. 5.41 reveal the likelihood of realizing the forecasted LMP and its close vicinity, and therefore, can be very useful for buyers and sellers in making their financial decisions, such as developing bidding strategies.

5.4.1.4. Alignment Probability of Probabilistic LMP

Fig. 5.42 shows the curve of the alignment probability of the deterministic LMP at Bus B versus forecasted the load, with a 10% and 20% price tolerance, respectively. By making a comparison between Fig. 5.42 and Fig. 5.35, we can see that the low probabilities occur near the CLLs. For instance, the lowest probability is 54.29% at the forecasted load level 968.22MW, which is very close to the CLL at 963MW. When the forecasted load is over 1300MW, the probability keeps decreasing as the forecasted load is approaching the maximum level (i.e., price of the VOLL) that the system can afford, namely, 1467MW.

As shown in Figure 5.42, the alignment probability curve is closer to 1.0 with a higher price tolerance, and the valley at around 968.22MW disappears when the price tolerance is 20%. For example, the deterministic LMP is $27.94/MWh at the load level 968.22MW, and the prices in the close neighborhood of 968.22MW differ by up to $3.80/MWh, which is less than 20% of the deterministic LMP. Therefore, the alignment probability at the load level 968.22MW increases from 54.29% with a 10% price tolerance, to nearly 100% with a doubled price tolerance.

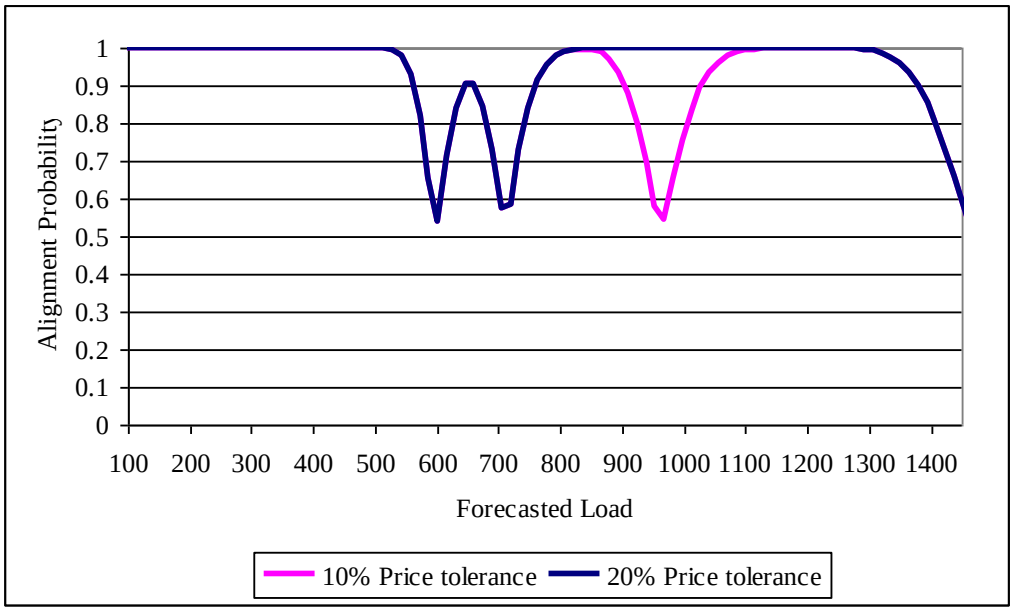


Figure 5.42. Alignment probability of the deterministic LMP at Bus B versus the forecasted load, with a 10% and 20% price tolerance, respectively

5.4.1.5. Expected Value of Probabilistic LMP

The expected value of the probabilistic LMP for the above case is compared with the deterministic LMP, $y(D_t^F)$, which is shown in Table 5.10. It shows that the expected value of probabilistic LMP can differ from the deterministic LMP for a specific forecasted load.

The expected value of the probabilistic LMP versus the forecasted load curve is shown in Fig. 5.43. A load range beyond 1300MW is not shown simply because the high VOLL will make the curve hard to scale illustratively. It should be noted that the expected LMP will escalate sharply when the load is close to the maximum load level the system can afford, and will eventually reach 2000 $/MWh.

Table 5.10. Expected value of the probabilistic LMP in comparison with the Deterministic LMP for Bus B

D^F_t(MW)	Expected Value of Probabilistic LMP($/MWh)	Deterministic LMP($/MWh)
730	20.82	22.14
900	24.59	24.30

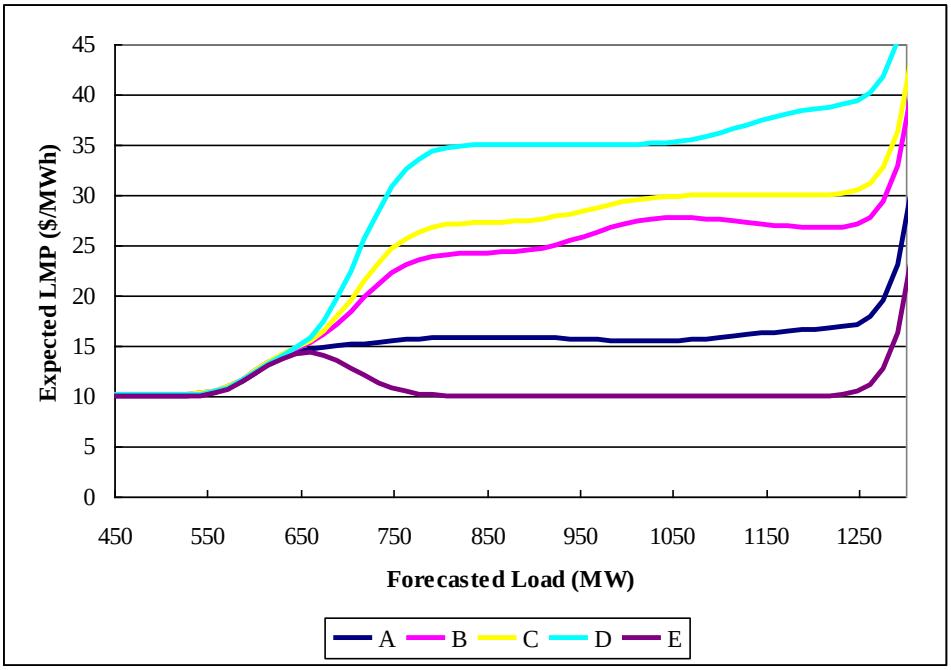


Figure 5.43. Expected value of probabilistic LMP versus forecasted load

In the deterministic LMP-Load curve in Fig. 5.35, the sensitivity for Bus B at 963MW is mathematically infinite since a step change occurs at this load level. In the probability-based LMP-Load curve, the upper and lower bounds of the sensitivity can be estimated using (5.62) and (5.63) shown as follows:

$$-17.41 \le \frac{\partial E_{LMP_t}(\mu_t, \sigma_t)}{\partial \mu_t} \le 17.41 \$/MWh^2.$$

When contrasted with the deterministic LMP-Load curve in Fig. 5.35, the curve of the expected value of the probabilistic LMP in Fig. 5.43 demonstrates the same overall trend. However, Fig. 5.43 shows a much smoother curve without any step changes, which contributes to the reduction of price uncertainty, especially around the CLLs.

Also shown in this probabilistic LMP forecasting figure is that when the load is closer to the CLLs, price uncertainty, i.e., the uncertainty associated with the forecasted deterministic LMP, will be higher. This matches the overall trend in the deterministic LMP in Fig. 5.35.

5.4.1.6. Impact of Load Forecasting Accuracy

In this section, three different levels of standard deviation of load forecasting are examined, 5%, 3%, and 1%. Fig. 5.44 shows the probabilities of the random variable LMP_t at Bus B falling into a few price ranges for these three levels of standard deviation when the system load is 730MW. It can be seen from Fig. 5.44 that the probability of realizing the actual price in the range of 22.10~22.2 $/MWh where the deterministic LMP $22.14/MWh falls into, increases considerably with a smaller standard deviation. This is reasonable because more a accurate load forecast should lead to less deviation in the forecasted price.

Fig. 5.45 compares the expected value of the probabilistic LMP curves at the same bus. When the forecasted load is a large distance from any CLL, for example at 850MW, the three curves overlap very well. This suggests that different levels of the standard deviation make minimal differences on the expected LMP at this load level. In addition, the sensitivity of the expected LMP at this load level is small, which indicates the expected LMP remains nearly constant when the forecasted load varies slightly around this level. In contrast, when the

forecasted load is close to a CLL, for example at 598.50MW, the lower the standard deviation is, the closer the curve is to a step change curve shape. Furthermore, the inset in Fig. 5.45 shows that when the load level is closer to a CLL, the absolute value of the sensitivity of the expected LMP grows rapidly, and the expected LMP becomes more sensitive to variations of the forecasted load.

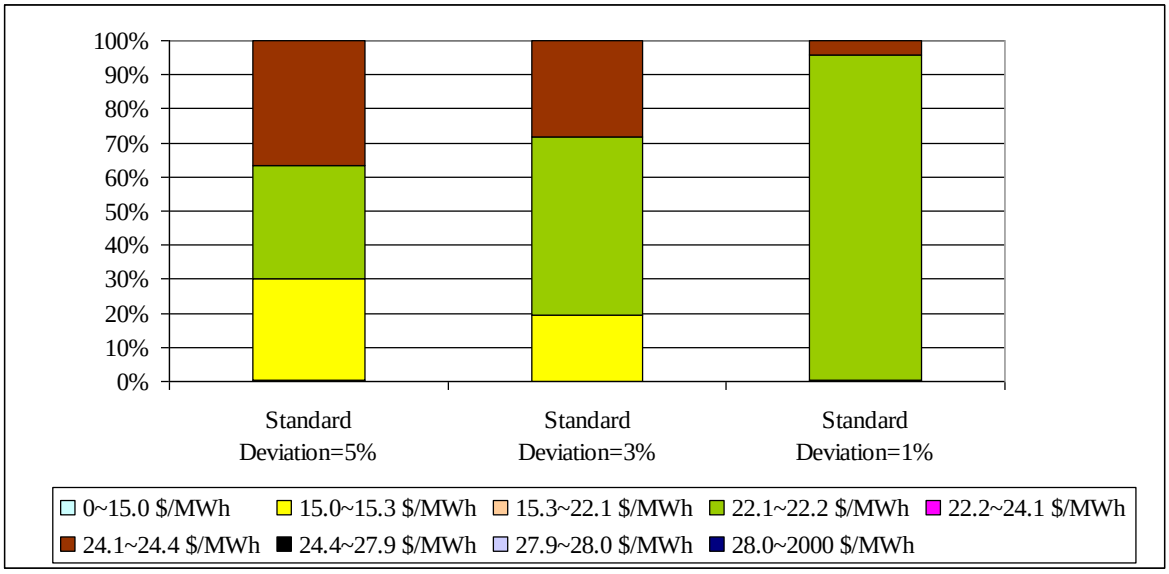


Figure 5.44. Probability of LMP_t at Bus B in a few price ranges for three levels of standard deviation when the system load is 730MW

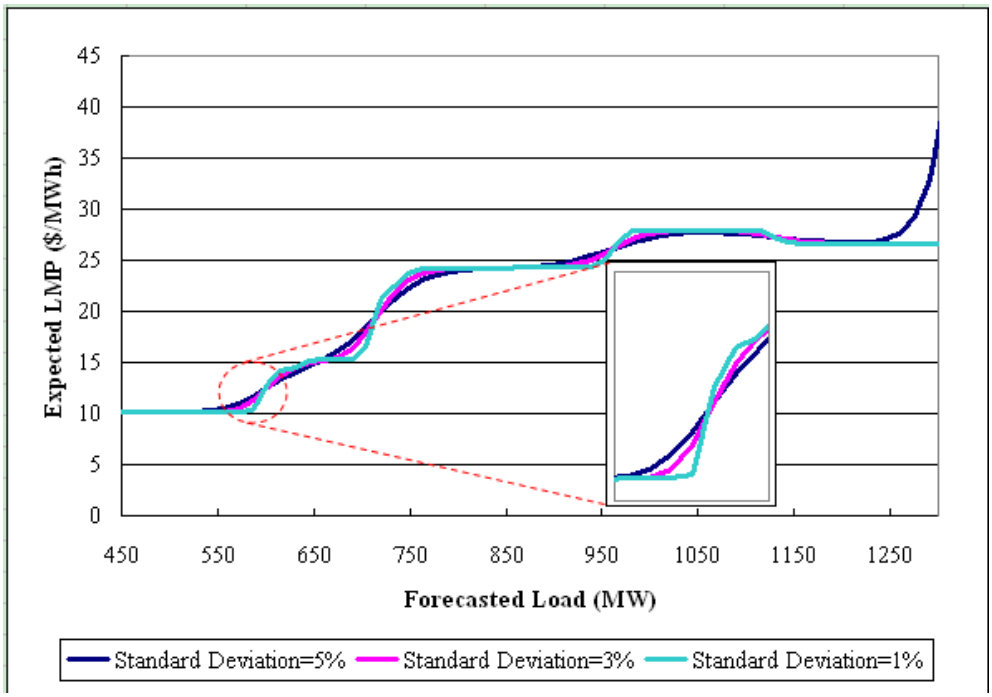


Figure 5.45. Expected value of the probabilistic LMP at Bus B versus the forecasted load for three levels of standard deviation

5.5. Conclusions

Based on the step change phenomenon of the LMP versus load curve observed in previous chapters, the load forecasting errors are expected to have considerable impacts on the forecasted LMP. This chapter therefore aims to investigate and quantify the impact.

The lossless DCOPF framework is firstly studied, with assumptions of normal distribution taken on the probability distribution of the system load. The concept of probabilistic LMP is proposed to reflect the fact that the actual LMP is not a deterministic value; rather, it is a discrete random variable. Consequently, the probability mass function and the expected value of the random variable are derived. The sensitivity of the expected value of the probabilistic LMP has been carefully studied with the proof that it is bounded by finite numbers, which matches the observations that the expected value of the probabilistic LMP versus load curve is highly smooth. In addition, the concept of alignment probability is presented to define the

probability of the actual LMP falling into a price range, which is practically useful for those who are interested in performing price forecasting with some accuracy tolerance.

As lossless DCOPF is a simplified dispatch model and does not address power losses, the probabilistic LMP for ACOPF framework has been examined. The probabilistic LMP is shown to be a piece-wise continuous random variable, which brings much more complexity into the study. The cumulative density function and probability density function are derived, and used to derive formulas for alignment probability, the expected value of probabilistic LMP, as well as its sensitivity. Interestingly, it has been shown that the CDF and PDF are differentiable almost everywhere except for the prices at two sides of each segment of the LMP versus load curve. Again, the sensitivity of the expected value of the probabilistic LMP has been proved to be bounded by finite numbers. The similarities and disparities between the probabilistic LMP in the ACOPF framework and that in the lossless DCOPF are presented.

The FND-based DCOPF framework produces a similar LMP versus load curve to that of ACOPF, and therefore the entire methodology, as well as formulas, can be applied directly to the FND-based DCOPF framework. Therefore, only a case study is shown for conciseness.

6 Conclusions

6.1. Summary of contributions

In Chapter 3, the reduction of system loss in the energy balance equality constraint in the DCOPF-with-loss model is rigorously proved. Then, the challenge of a considerable nodal mismatch at the reference bus is presented. The mismatch issue is tackled with the proposed Fictitious Nodal Demand (FND) model, in which the total loss is distributed into each *individual* line and thus, resulting in no nodal mismatches.

Chapter 3 also presents a comparison of the LMP results from the lossless DCOPF, the FND-based DCOPF, and the ACOPF algorithms. Results indicate that the FND-based Iterative DCOPF provides much better results than the lossless DCOPF and represents a better approximation of the ACOPF LMP model.

In addition, Chapter 3 presents a simple and explicit formulation of the LMP sensitivity w.r.t. load, based on the FND algorithm. Without a loss component, the LMP sensitivity is zero if the load is varied within a small range. The LMP sensitivity may be infinite (i.e., a step change in LMP) when the load grows to a critical level leading to a new marginal unit. This step-change nature presents uncertainty and risk in the LMP forecast, especially considering the possible data inaccuracy or algorithm approximation.

In Chapter 4, the problem of predicting price and congestion is conducted for each of the three major OPF models, namely, the lossless DCOPF, ACOPF, and the proposed FND-based DCOPF.

For the lossless DCOPF model, through the exploration of the characteristics of the optimal solution, a systematic approach is proposed to give a global view of congestion and price versus load, from any given load level to another, without multiple optimization runs. It first expresses marginal variables as a function of other non-marginal variables. Then, it identifies the next binding limit and the next critical load level (CLL). Next, the next unbinding limit such as a new marginal unit can be selected. Finally, the new generation output sensitivity at the CLL can be obtained because the objective function is expressed as non-marginal variables. Therefore, the new LMP can be obtained when the load is greater than the CLL. In this way, the LMP versus load curve is quickly obtained, and the curve could be used to predict price spikes, given the forecasted load versus time curve. This approach has great potential in market-based system operation and planning, especially in the short term, for congestion management and price prediction.

The nice linear feature associated with the lossless DCOPF is not valid when power loss is considered. Therefore, different methods need to be investigated for the OPF models which will address the loss issues. For the ACOPF model, the marginal unit generation and line flow are numerically verified to follow a nearly perfect quadratic polynomial pattern through polynomial curve-fitting as higher order terms are negligible. Then, a quadratic interpolation method is proposed to reduce the computational efforts arising from the polynomial curve-fitting. However, the interpolation method generally still requires at least two additional OPF runs at load levels other than the given operating point. In this regard, a variable substitution method is proposed to further reduce the computational efforts for the FND-based DCOPF model. It takes the assumption that marginal unit generations can be approximated by quadratic polynomials, and substitutes the polynomials into the characteristic constraints formed by binding constraints. This method does not require any additional OPF runs, and

involves a very limited computation involved in solving a small-scale nonlinear equation problem. However, the computation efforts are saved at the expenses of losing accuracy.

Price prediction relies on the results of load forecasting. Therefore, it is highly interesting to investigate how the load forecasting uncertainty affects the price forecasting. This study is conducted in Chapter 5 for each of the three OPF models.

For the lossless DCOPF model, the concept of a probabilistic LMP is proposed and its probability mass function at hour t is presented. The probabilistic LMP does not correspond to a single deterministic value. In stead, it represents a set of discrete values (p_i) at a number of load intervals, and each value p_i has an associated probability. The alignment probability is proposed and formulated to define the likelihood that the deterministic LMP will be realized. The expected value of the probabilistic LMP and its sensitivity is derived and shown to be bounded by finite values. In addition, the expected value of the probabilistic LMP versus the forecasted load curve is smooth and has no step changes. This avoids the 0-or-1 type of step changes if a deterministic LMP forecast is performed, and helps market participants make wise decisions in generation bidding, demand offers, and/or forward contract negotiation.

For the ACOPF model, the probabilistic LMP is also proposed and carefully studied. The probabilistic LMP for the ACOPF framework is a piece-wise continuous random variable, whose cumulative density function and probability density function are derived and shown to be differentiable at almost everywhere except a finite number of prices. The alignment probability is proposed to reflect the probability that the actual price falls into certain price range. The expected value of the probabilistic LMP is shown to be highly smooth and its sensitivities are shown to be bounded by finite values, just as those for the lossless DCOPF framework. The same methodology is applied to FND-based DCOPF framework due to the

similarity between the LMP versus Load curves of the ACOPF model and the FND-based DCOPF model.

6.2. Future Works

The following issues may be considered as future works of this dissertation.

- **On OPF problem and LMP Calculation**:
  - To investigate the convergence issue of the iterative algorithm for the FND-based DCOPF model with respect to system sizes and parameters like R/X ratios.
  - To Study the accuracy of congestion component of the LMP obtained from the FND-based DCOPF by comparing it with that of the ACOPF.
- **On Congestion and Price Prediction under Load Variation**
  - To investigate the case with insufficient reactive power support for the ACOPF model
  - To obtain the LMP versus time curve by combining the LMP versus Load curve and load forecasting data
  - To study the impact of the unit commitment on congestion and price prediction.
- **On Probabilistic LMP under Load Uncertainty**
  - To employ more sophisticated load models to consider the randomness as well as correlation among the different load areas and among consecutive hours.

List of References

Books

[1]. Richard E. Brown, *Electric Power Distribution Reliability*, Marcel Dekker, New York, 2002

[2]. A. J. Wood, B. F. Wollenberg, *Power Generation, Operation and Control* (second edition), John Wiley&Sons, New York, 1996

[3]. Steven Stoft, *Power System Economics*, IEEE/Wiley, 2002

[4]. F. C. Schweppe, M. C. Caramanis, R. D. Tabors, R. E. Bohn, *Spot Pricing of Electricity*, Kluwer Academic Publishers, 1988

[5]. Marija Ilic, Francisco Galiana, Lester Fink, *Power Systems Restructuring: Engineering and Economics*, Kluwer Academic Publishers, 1998

[6]. Mohammad Shahidehpour, Hatim Yamin, Zuyi Li, *Market Operations in Electric Power Systems: Forecasting, Scheduling and Risk Management*, John Wiley&Sons, New York, 2002

[7]. D. Kirschen and G. Strbac, *Fundamentals of Power System Economics*, John Wiley and Sons, 2004.

[8]. Mariesa Crow, *Computational Methods for Electric Power Systems*, CRC Press, 2003

[9]. Milton Abramowitz and Irene Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, 1964.

[10]. Charles M. Grinstead and J. Laurie Snell, *Introduction to Probability*, American Mathematical Society, 1997

Regulations and Tools

[11]. *PJM Training Materials (LMP101)*, PJM, http://www.pjm.com/services/training/train-materials.html.

[12]. *2006 State of the Market Report Section 7 - Congestion*, PJM.

[13]. Business Practices Manual: Energy and Operating Reserve Markets, http://www.midwestiso.org

[14]. *NYISO Transmission & Dispatch Operations Manual*, NYISO, 1999.

[15]. ISO-New England, *FERC Electric Rate Schedule*, July 2001.

[16]. *Power System Test Case Archive*, University of Washington, http://www.ee.washington.edu/research/pstca/.

[17]. R. D. Zimmerman, C. E. Murillo-Sánchez, D. Gan, *MatPower – A Matlab Power System Simulation Package*, School of Electrical Engineering, Cornell University, http://www.pserc.cornell.edu/matpower/matpower.html.

Market Design and ISO Practices

[18]. Andrew L. Ott, "Experience with PJM Market Operation, System Design, and Implementation," *IEEE Trans. on Power Systems*, vol. 18, no. 2, pp. 528-534, May 2003.

[19]. Tongxin Zheng, Eugene Litvinov, "Ex Post Pricing in the Co-Optimized Energy and Reserve Market,"*IEEE Trans. on Power Systems*, vol. 21, no. 4, pp. 1528-1538, Nov. 2006.

[20]. J. E. Price, "Market-Based Price Differentials in Zonal and LMP Market Designs," *IEEE Trans. on Power Systems*, vol. 22, no. 4, pp. 1486-1494, Nov. 2007.

[21]. Kwok W. Cheung, "Standard Market Design for ISO New England Wholesale Electricity Market: An Overview," *Proceedings of the 2004 IEEE International Conference on Deregulation, Restructuring and Power Technologies (DRPT)*, vol. 1, pp. 38-43.

[22]. Xingwang Ma, David I. Sun, Kwok W. Cheung, "Evolution Toward Standardized Market Design," *IEEE Trans. on Power Systems*, vol. 18, no. 2, pp. 460-469, May 2003.

DC/AC Power Flow/Optimal Power Flow

[23]. K. Purchala, L. Meeus, D. Van Dommelen, and R. Belmans, "Usefulness of DC Power Flow for Active Power Flow Analysis," *Proc. of IEEE PES Annual Meeting 2005*, pp. 454-459, June 2005.

[24]. T. Overbye, X. Cheng, and Y. Sun, "A Comparison of the AC and DC Power Flow Models for LMP Calculations," *Proceedings of the 37th Hawaii International Conference on System Sciences*, 2004.

[25]. A. J. Conejo and J.A. Aguado, "Multi-Area Coordinated Decentralized DC Optimal Power Flow," *IEEE Trans. on Power Systems*, vol. 13, no. 4, pp. 503-509, Nov. 1998.

Distribution Factors and Loss Factors

[26]. M. Liu, G. Gross, "Effectiveness of the Distribution Factor Approximations Used in Congestion Modeling," Proceedings of the 14th Power Systems Computation Conference, Seville, Spain, June 2002, pp. 24–28.

[27]. Ross Baldick, "Variation of Distribution Factors With Loading," *IEEE Trans. on Power Systems*, vol. 18, no. 4, pp. 1316-1323, November 2003.

[28]. Xu Cheng, Thomas J. Overbye, "PTDF-Based Power System Equivalents," *IEEE Trans. on Power Systems*, vol. 20, no. 4, pp. 1868-1876, Nov. 2005.

[29]. J. L. M. Ramos, A. G. Exposito, F. J. C. Moron and S. N. Becerra, "On the Use of Loss Penalty Factors for Generation Scheduling", *Proceedings of IEEE PES Annual Meeting 2003*, vol. 2, pp. 926-931.

[30]. J. Zhu, D. Hwang and A. Sadjadpour, "Real time loss sensitivity calculation in power systems operation," *Electric Power Systems Research*, vol. 73, no. 1, pp. 53-60, January 2005.

Marginal Loss Pricing

[31]. L. Liu and A. Zobian, "The Importance of Marginal Loss Pricing in an RTO Environment," *The Electricity Journal*, vol. 15, no. 8, pp. 40-45, October 2002.

[32]. E. Litvinov, T. Zheng, G. Rosenwald, and P. Shamsollahi, "Marginal Loss Modeling in LMP Calculation," *IEEE Trans. on Power Systems*, vol. 19, no. 2, pp. 880-888, May 2004.

[33]. Fangxing Li, Jiuping Pan and Henry Chao, "Marginal Loss Calculation in Competitive Spot Market," *Proceedings of the 2004 IEEE International Conference on Deregulation, Restructuring and Power Technologies (DRPT)*, vol. 1, pp. 205-209.

[34]. J. B. Cardell, "Marginal Loss Pricing for Hours With Transmission Congestion," *IEEE Trans. on Power Systems*, vol. 22, no. 4, pp. 1466-1474, November 2007.

LMP and Its Decomposition

[35]. Zuyi Li, Hossein Daneshi, "Some observations on market clearing price and locational marginal price," *Proceedings of IEEE PES General Meeting 2005*, pp. 2702-2709, June 2005.

[36]. Kai Xie, Yong-Hua Song, John Stonham, Erkeng Yu, and Guangyi Liu, "Decomposition Model and Interior Point Methods for Optimal Spot Pricing of Electricity in Deregulation Environments," *IEEE Trans. on Power Systems*, vol. 15, no. 1, pp. 39-50, Feb. 2000.

[37]. Luonan Chen, Hideki Suzuki, Tsunehisa Wachi, Yukihiro Shimura, "Components of Nodal Prices for Electric Power Systems," *IEEE Trans. on Power Systems*, vol. 17, no. 1, pp. 41-49, Feb. 2002.

[38]. Tong Wu, Z. Alaywan, and A. D. Papalexopoulos, "Locational Marginal Price Calculations Using the Distributed-Slack Power-Flow Formulation," *IEEE Trans. on Power Systems*, vol. 20, no. 2, pp. 1188-1190, May 2005.

[39]. Xu Cheng and T. J. Overbye, "An energy reference bus independent LMP decomposition algorithm," *IEEE Trans. on Power Systems*, vol. 21, no. 3, pp. 1041-1049, Aug. 2006.

[40]. T. Orfanogianni and G. Gross, "A General Formulation for LMP," *IEEE Trans. on Power Systems*, vol. 22, no. 3, pp. 1163-1173, Aug. 2007.

LMP Sensitivities

[41]. A. J. Conejo, E. Castillo, R. Minguez, and F. Milano, "Locational Marginal Price Sensitivities," *IEEE Trans. on Power Systems*, vol. 20, no. 4, pp. 2026-2033, November 2005.

[42]. Fangxing Li, "Continuous Locational Marginal Pricing (CLMP)," *IEEE Trans. on Power Systems*, vol. 22, no. 4, pp. 1638-1646, November 2007.

Load Forecasting

[43]. G. Gross and F. D. Galiana, "Short-term Load Forecasting," *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1558-1573, Dec 1987.

Price Forecasting

[44]. J. Bastian, J. Zhu, V. Banunarayanan, and R. Mukerji, "Forecasting Energy Prices in a Competitive Market," *IEEE Computer Applications in Power Magazine*, vol. 12, no. 3, pp. 40-45, July 1999.

[45]. G. Li, C.-C. Liu, C. Mattson and J. Lawarree, "Day-Ahead Electricity Price Forecasting in a Grid Environment," *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 266-274, Feb. 2007.

[46]. P. Mandal, T. Senjyu, N. Urasaki, T. Funabashi and A. K. Srivastava, "A Novel Approach to Forecast Electricity Price for PJM Using Neural Network and Similar Days Method," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 2058-2065, Nov. 2007.

[47]. N. M. Pindoriya, S. N. Singh, S. K. Singh, "An Adaptive Wavelet Neural Network-Based Energy Price Forecasting in Electricity Markets," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1423-1432, Nov. 2007.

[48]. Francisco J. Nogales, Javier Contreras, Antonio J. Conejo, Rosario Espinola, "Forecasting Next-Day Electricity Prices by Time Series Models," *IEEE Transactions on Power Systems*, vol. 17, no. 2, pp. 342-348, May. 2002.

Bidding Strategy

[49]. Antonio J. Conejo, Francisco J. Nogales, Jose M. Arroyo, "Price-Taker Bidding Strategy Under Price Uncertainty," *IEEE Transactions on Power Systems*, vol. 17, no. 4, pp. 1081-1088, Nov. 2002.

Uncertainty Impact

[50]. D. K. Ranaweera, G. G. Karady, and R. G. Farmer, "Economic Impact Analysis of Load Forecasting," *IEEE Transactions on Power Systems*, vol. 12, no. 3, pp. 1388-1392, Aug. 1997.

[51]. Benjamin F. Hobbs, Suradet Jitprapaikulsarn, Sreenivas Konda, Vira Chankong, Kenneth A. Loparo, Dominic 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.

[52]. J. Valenzuela, M. Mazumdar, and A. Kapoor, "Influence of Temperature and Load Forecast Uncertainty on Estimation of Power Generation Production Costs," *IEEE Transactions on Power Systems*, vol. 15, no. 2, pp. 668-674, May 2000.

[53]. Roy Billinton and Dange Huang, "Effects of Load Forecast Uncertainty on Bulk Electric System Reliability Evaluation," *IEEE Transactions on Power Systems*, vol. 23, no. 2, pp. 418-425, May 2008.

[54]. M. A. Ortega-Vazquez and D. S. Kirschen, "Economic impact assessment of load forecast errors considering the cost of interruptions," *Proceedings of 2006 IEEE Power Engineering Society General Meeting*, Montréal, Canada, 2006.

[55]. K. L. Lo, Y. K. Wu, "Risk Assessment due to local demand forecast uncertainty in the competitive supply industry", *IEE Proc. Gener. Transm. Distrib.*, vol.150, no. 5, pp. 573-582, September 2003.

[56]. Y. He and Y. H. Song, "Integrated biding strategies by optimal response to probabilistic locational marginal prices," *IEE Proc-Gener. Transm. Distrib.*, vol. 149, no. 6, pp. 633-639, November 2002.

[57]. X. Chen, Y. He, Y. H. Song, et al, "Study of impacts of physical contracts and financial contracts on bidding strategies of GENCOs," *Electrical Power and Energy Systems*, 26, 2004, pp. 715-723.

Misc

[58]. Jian Yang, Fangxing Li and L.A.A. Freeman, "A Market Simulation Program for the Standard Market Design and Generation/Transmission Planning," *Proc. of IEEE PES General Meeting*, 2003, pp. 442-446.

[59]. R. E. Clayton and R. Mukerji, "System Planning Tools for the Competitive Market," *IEEE Computer Applications in Power*, July 1996, pp. 50-55.

[60]. Rajat K. Deb and Keith D. White, "Valuing Transmission Investments: The Big Picture and the Details Matter – and Benefits Might Exceed Expectations," *The Electricity Journal*, Volume 18, Issue 7, pp. 33-42, August-September 2005.

[61]. Promod, Website: http://www1.ventyx.com/analytics/promod.asp, accessed on Feb. 10, 2009.

[62]. Xiaokang Xu, D. W. Lane, and M. J. S. Edmonds, "A simulation tool for calculating energy prices in competitive electricity markets," *Proc. 3rd International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT)*, April 6-9, 2008.

[63]. H. Wang, C. E. Murillo-Sanchez, R. D. Zimmerman, and R. J. Thomas, "On Computational Issues of Market-Based Optimal Power Flow," *IEEE Trans. on Power Systems*, vol. 22, no. 3, pp. 1185-1193, Aug. 2007.

[64]. Jian Yang, "Resource adequacy: economic and engineering challenges and proposed solutions," *IEEE Power and Energy Magazine*, vol. 4, no. 2, pp. 59-65, March-April 2006.

[65]. Jose R. Daconti, Daniel C. Lawry, "Increasing Power Transfer Capability of Existing Transmission Lines," *Proceedings of 2003 IEEE PES Transmission and Distribution Conference and Exposition*, Dallas, USA, vol. 3, pp.1004-1009, Sept. 2003.

Appendices

Appendix A

Schematic proof of the convergence feature of FND-based DCOPF algorithm

Proposition: If G_i converges after the $(l+1)^{st}$ iteration, F_k, E_i, DF_i, *and* P_{loss} all converge for the FND-based DCOPF algorithm.

Schematic Proof: Sequence $\{G_i\}$ converges $\Leftrightarrow$ for any $\varepsilon > 0$, there exists a positive integer N such that for any integer number n, $m>N$, $\left|G_i^{\,n} - G_i^{\,m}\right| < \varepsilon$.

For simplicity, let $m=l$, $n=l+1$. Then, $\left|G_i^{\,l+1} - G_i^{\,l}\right| < \varepsilon$.

(1) Since line loss in an individual line is a small portion of the line flow, we have

$$F_k = \sum_{i=1}^{N} GSF_{k-i} \times (G_i - D_i - E_i) = (1 + a_k)\left[\sum_{i=1}^{N} GSF_{k-i} \times (G_i - D_i)\right]$$

where a_k = the ratio of line loss to line flow of the k^{th} line. Typically, a_k is a small positive number less than 10%.

$$\because D_i^{l+1} = D_i^l = D_i$$
$$\therefore \left|F_k^{l+1} - F_k^l\right| = \left|(1 + a_k) \times \sum_{i=1}^{N} (GSF_{k-i} \times (G_i^{\,l+1} - G_i^{\,l}))\right|$$
$$\therefore \left|F_k^{l+1} - F_k^l\right| < (1 + a_k) \times N \times \max\{\left|GSF_{k-i}\right|\} \times \varepsilon = b_k \times \varepsilon$$

where $b_k = (1 + a_k) \times N \times \max\{\left|GSF_{k-i}\right|\}$

(2) E_i is the fictitious nodal demand, so we have

$$G_i + F_{i\Sigma} - D_i - E_i = 0 \text{ or } E_i = G_i + F_{i\Sigma} - D_i$$

where $F_{i\Sigma}$ = all line flows injecting into Bus i.

Since $\left|G_i^{\,l+1} - G_i^{\,l}\right| < \varepsilon$, $\left|F_k^{l+1} - F_k^l\right| < b_k \times \varepsilon$, and $D_i^{l+1} = D_i^l = D_i$, we have

$$\left|E_i^{\,l+1} - E_i^{\,l}\right| < \varepsilon + M_i \times \max\{b_k\} \times \varepsilon = (1 + M_i \times b) \times \varepsilon$$

where M_i = number of lines connected to Bus i; b = the maximum b_k of all lines connected to Bus i.

(3) $\because DF_i = 1 - \sum_{k=1}^{M} 2 \times R_k \times GSF_{k-i} \times F_k$

$$\therefore \left| DF_i^{\,l+1} - DF_i^{\,l} \right| = \left| \sum_{k=1}^{M} \left(\times R_k \times GSF_{k-i} \times (F_k^{\,l+1} - F_k^{\,l}) \right) \right|$$

$$\therefore \left| DF_i^{\,l+1} - DF_i^{\,l} \right| < 2M \times \max\{ \left| R_k \times GSF_{k-i} \times b_k \right| \} \times \varepsilon = c_i \times \varepsilon$$

where $c_i = 2M \times \max\{ \left| R_k \times GSF_{k-i} \times b_k \right| \}$

(4) $\because P_{loss} = \sum_{i=1}^{N} \left(G_i - D_i \right)$

$$\therefore \left| P_{loss}^{\,l+1} - P_{loss}^{\,l} \right| = \left| \sum_{i=1}^{N} \left(G_i^{l+1} - G_i^{l} \right) \right| < N \times \varepsilon \,.$$

Appendix B

Derivation of equation (5.13)

Equation (5.12) can be further decomposed into several parts.

$$\frac{\partial E_{LMP_t}(\mu_t,\sigma_t)}{\partial \mu_t}$$

$$= \sum_{i=0}^{n} p_i \times \int_{D_i}^{D_{i+1}} \frac{u-\mu_t}{\sigma_t^3\sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du$$

$$= p_0 \times \int_{D_0}^{D_1} \frac{u-\mu_t}{\sigma_t^3\sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du + \sum_{i=1}^{n-1} p_i \times \int_{D_i}^{D_{i+1}} \frac{u-\mu_t}{\sigma_t^3\sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du$$

$$+ p_n \times \int_{D_n}^{D_{n+1}} \frac{u-\mu_t}{\sigma_t^3\sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du$$

$$= \sum_{i=1}^{n-1} p_i \times \int_{D_i}^{D_{i+1}} \frac{u-\mu_t}{\sigma_t^3\sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du + p_n \times \int_{D_n}^{D_{n+1}} \frac{u-\mu_t}{\sigma_t^3\sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du \qquad \text{(B.1)}$$

Each integration of equation (B.1) can be solved as follows:

$$\int_{D_i}^{D_{i+1}} \frac{u-\mu_t}{\sigma_t^3\sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du$$

$$= \frac{1}{\sigma_t^3\sqrt{2\pi}} \int_{D_i}^{D_{i+1}} (u-\mu_t) e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du$$

$$= \frac{1}{\sigma_t^3\sqrt{2\pi}} \int_{D_i}^{D_{i+1}} \left(-\frac{1}{2}\right) \times e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} \times d\left((u-\mu_t)^2\right)$$

$$= \frac{1}{\sigma_t^3\sqrt{2\pi}} \int_{D_i}^{D_{i+1}} \left(-\frac{1}{2}\right) \times 2\sigma_t^2 \times e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} d\left(-\frac{(u-\mu_t)^2}{2\sigma_t^2}\right)$$

$$= -\frac{1}{\sigma_t\sqrt{2\pi}} \int_{D_i}^{D_{i+1}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} d\left(-\frac{(u-\mu_t)^2}{2\sigma_t^2}\right)$$

$$= -\frac{1}{\sigma_t\sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} \Bigg|_{D_i}^{D_{i+1}}$$

$$= \frac{1}{\sigma_t \sqrt{2\pi}} \left(e^{-\frac{(D_i - \mu_t)^2}{2\sigma_t^2}} - e^{-\frac{(D_{i+1} - \mu_t)^2}{2\sigma_t^2}} \right) \tag{B.2}$$

Specifically, when $i = n$, (B.2) can be further derived as

$$\begin{aligned} \int_{D_i}^{D_{i+1}} \frac{u - \mu_t}{\sigma_t^3 \sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du &= \int_{D_n}^{D_{n+1}} \frac{u - \mu_t}{\sigma_t^3 \sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du \\ &= \int_{D_n}^{\infty} \frac{u - \mu_t}{\sigma_t^3 \sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du \\ &= \frac{1}{\sigma_t \sqrt{2\pi}} e^{-\frac{(D_n - \mu_t)^2}{2\sigma_t^2}} \end{aligned} \tag{B.3}$$

Plugging (B.2) and (B.3) into (B.1), we have the formula of sensitivity of expected value of probabilistic LMP

$$\begin{aligned} &\frac{\partial E_{LMP_t}(\mu_t, \sigma_t)}{\partial \mu_t} \\ &= \sum_{i=1}^{n-1} p_i \times \int_{D_i}^{D_{i+1}} \frac{u - \mu_t}{\sigma_t^3 \sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du + p_n \times \int_{D_n}^{D_{n+1}} \frac{u - \mu_t}{\sigma_t^3 \sqrt{2\pi}} e^{-\frac{(u-\mu_t)^2}{2\sigma_t^2}} du \\ &= \sum_{i=1}^{n-1} p_i \times \frac{1}{\sigma_t \sqrt{2\pi}} \left(e^{-\frac{(D_i - \mu_t)^2}{2\sigma_t^2}} - e^{-\frac{(D_{i+1} - \mu_t)^2}{2\sigma_t^2}} \right) \\ &\quad + p_n \times \frac{1}{\sigma_t \sqrt{2\pi}} e^{-\frac{(D_n - \mu_t)^2}{2\sigma_t^2}} \\ &= \frac{1}{\sigma_t \sqrt{2\pi}} \times \left[\sum_{i=1}^{n-1} p_i \times \left(e^{-\frac{(D_i - \mu_t)^2}{2\sigma_t^2}} - e^{-\frac{(D_{i+1} - \mu_t)^2}{2\sigma_t^2}} \right) + p_n \times e^{-\frac{(D_n - \mu_t)^2}{2\sigma_t^2}} \right] \end{aligned} \tag{B.4}$$

The last step of equation (B.4) gives equation (5.13).

Appendix C

Derivation of equation (5.14)

Based on equation (5.13), the bounds of the sensitivity of expected value of probabilistic LMP can be obtained as follows,

$$\begin{aligned}
&\left|\frac{\partial E_{LMP_t}(\mu_t,\sigma_t)}{\partial \mu_t}\right| \\
&= \left|\frac{1}{\sigma_t\sqrt{2\pi}} \times \left(\sum_{i=1}^{n-1} p_i \times \left(e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}} - e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}}\right) + p_n \times e^{-\frac{(D_n-\mu_t)^2}{2\sigma_t^2}}\right)\right| \\
&\le \frac{1}{\sigma_t\sqrt{2\pi}} \times \left(\left|\sum_{i=1}^{n-1} p_i \times \left(e^{-\frac{(D_i-\mu_t)^2}{2\sigma_t^2}} - e^{-\frac{(D_{i+1}-\mu_t)^2}{2\sigma_t^2}}\right)\right| + \left|p_n \times e^{-\frac{(D_n-\mu_t)^2}{2\sigma_t^2}}\right|\right) \\
&\le \frac{1}{\sigma_t\sqrt{2\pi}} \times \left(\sum_{i=1}^{n-1} |p_i| \times 1 + |p_n| \times 1\right) \\
&\le \frac{1}{\sigma_t\sqrt{2\pi}} \sum_{i=1}^{n} |p_i| \qquad \text{(C.1)}
\end{aligned}$$

The last step in equation (C.1) gives equation (5.14).

Appendix D

Publications

Published journal and conference papers during Ph.D. study are listed as follows:

[1]. **Rui Bo** and Fangxing Li, "Probabilistic LMP Forecasting Considering Load Uncertainty," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1279-1289, August 2009.

[2]. Fangxing Li and **Rui Bo**, "Congestion and Price Prediction under Load Variation," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 911-922, May 2009.

[3]. Fangxing Li and **Rui Bo**, "DCOPF-based LMP Simulation: Algorithm, Comparison with ACOPF, and Sensitivity," *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp.1475-1485, November 2007.

[4]. **Rui Bo**, Fangxing Li, "Impact of Load Forecast Uncertainty on LMP," *Proceedings of 2009 IEEE PES Power Systems Conference & Exposition*, Seattle, Washington, USA, 2009.

[5]. **Rui Bo**, Fangxing Li, "Power Flow Studies Using Principle Component Analysis," *Proceedings of the North American Power Symposium 2008*, Calgary, Canada, 2008.

[6]. **Rui Bo**, Fangxing Li and Chaoming Wang, "Congestion Prediction for ACOPF Framework Using Quadratic Interpolation," *Proceedings of the IEEE Power Engineering Society General Meeting 2008*, Pittsburgh, USA, 2008.

[7]. **Rui Bo** and Fangxing Li, "Sensitivity of LMP Using an Iterative DCOPF Model," *Proceeding of the 3rd IEEE International Conference on Deregulation, Restructuring, and Power Technology (DRPT2008)*, Nanjing, China, 2008.

[8]. **Rui Bo** and Fangxing Li, "Comparison of LMP Simulation Using Two DCOPF Algorithms and the ACOPF Algorithm," *Proceeding of the 3rd IEEE International Conference on Deregulation, Restructuring, and Power Technology (DRPT2008)*, Nanjing, China, 2008.

[9]. Fangxing Li, **Rui Bo**, Wenjuan Zhang, "Comparison of Different LMP Calculations in Power Market Simulation," *Proceeding of 2006 International Conference on Power System Technology*, Chongqing, China, 2006.

Appendix E

Awards

Awards received during Ph.D. study are listed as follows:

[1]. Second Place Prize Award at Student Poster Contest at 2009 IEEE PES Power System Conference and Exposition, Seattle, WA, March 2009.

[2]. UT Citation Award in Extraordinary Professional Promise, April 2009.

Vita

Rui Bo received his B.S. and M.S. degrees, both in electric power engineering, from Southeast University (China) in 2000 and 2003, respectively. From 2003 to 2005, he worked at ZTE Corporation and Shenzhen Cermate Inc., respectively. He started his Ph.D. studies at The University of Tennessee, Knoxville, in January 2006. His current interests include power system operation and planning, power system economics, and market simulation. He is the recipient of the second place price award at the Student Poster Contest at 2009 IEEE Power System Conference and Exposition at Seattle, Washington, and 2009 UT Citation Award in Extraordinary Professional Promise.