2019

# Data-Driven Test Cases for Sustainability Assessment of Smart Grid Initiatives in Organized Electricity Markets

Venkat Durvasulu
*South Dakota State University*

# DATA-DRIVEN TEST CASES FOR SUSTAINABILITY ASSESSMENT OF SMART GRID INITIATIVES IN ORGANIZED ELECTRICITY MARKETS

BY

VENKAT DURVASULU

A dissertation submitted in partial fulfillment of the requirements for the

Doctor of Philosophy

Major in Electrical Engineering

South Dakota State University

2019

DATA-DRIVEN TEST CASES FOR SUSTAINABILITY ASSESSMENT OF SMART GRID INITIATIVES IN ORGANIZED ELECTRICITY MARKETS

Venkat Durvasulu

This dissertation is approved as a creditable and independent investigation by a candidate for the Doctor of Philosophy in Electrical Engineering degree and is acceptable for meeting the thesis requirements for this degree. Acceptance of this dissertation does not imply that the conclusions reached by the candidates are necessarily the conclusions of the major department.

Timothy M. Hansen, Ph.D.
Dissertation Advisor — Date

George Hamer, Ph.D.
Department Head — Date
Electrical Engineering and Computer Science

Dean, Graduate School — Date

I dedicate this work to my mother, Rajeswari, and my brother, Vishy. None of this may have been possible without their love and support.

*"Strength does not come from physical capacity. It comes from an indomitable will."*

Mahatma Gandhi

# ACKNOWLEDGEMENTS

I would like to express my sincere gratitude towards my research advisor Dr. Timothy M. Hansen for his continuous guidance and encouragement. I would like to acknowledge him for conceptualizing and funding this work. This work would not be possible without his support, motivation, and enthusiasm. Besides my advisor, I would like to thank my committee, Dr. Reinaldo Tonkoski, Dr. Zhen Ni, Dr. Robert Fourney and graduate faculty representative Dr. Anne Fennel for their participation and interest in my work.

Additionally, I would thank Dr. Siddharth Suryanarayana, Dr. Howard Jay Siegel, Dr. Anthony A. Maciejewski, Swagata Sharma, Berk Celik from Colorado State University (CSU) and Hendy Syahril for their contributions to this work. I would like to thank all the graduate students who were present here during my courses for their help and support in providing constructive feedback on my work.

The research in this dissertation was supported by the South Dakota Board of Regants (SDBoR) PhD funds, and National Science Foundation under grant number-ECCS-1608722.

CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ABSTRACT

DATA-DRIVEN TEST CASES FOR SUSTAINABILITY ASSESSMENT
OF SMART GRID INITIATIVES IN ORGANIZED ELECTRICITY MARKETS

VENKAT DURVASULU

2019

The primary aim of this dissertation is to deliver a technique to augment power system test cases with realistic open-source data to represent a deregulated power system. These test cases are intended to be used by power system researchers who require a test case that is capable of performing economic and environmental analysis on a bulk-power level. These test cases are capable of estimating the cost of bulk-energy for economic analysis and harmful greenhouse gas (GHG) and air polluting (AP) emissions for environmental sustainability analysis. These cases are developed for simulations that are intended to be at the transmission level where the independent system operator (ISO) has control. In the second part of this dissertation, an aggregator based demand response (DR) model is studied as-a-service to the bulk-power market, and its economic benefit is estimated using the augmented test cases.

The augmentation technique presented in this dissertation has three-layer data over the existing generator information in a test case. The first layer of augmented data replaces the cost functions of the test case generators with functions developed based on the generator offers from a real electricity market. An unsupervised learning technique had to be implemented to classify the market offer data because the identity of the generators is masked to honor a fair market policy. The offer data was converted to cost

functions and is sampled statistically such that the test cases represent a similar generator supply curve as the real power system.

In addition to the cost functions layer, the test case generator data has an augmented generator fuel-turbine data. This data in a test case will represent the energy sources and generator technology of the system that the test case is intended to emulate. The hourly energy mix of the electricity market is utilized to augment the generator fuel-turbine type of test case generators. Because the number and capacities of test case generators may not represent the real system, assigning one fuel-turbine type to one test case generator will not result in a right energy mix. The augmentation technique creates an additional layer of information for each test case generator which can represent multiple fuel-types. The third layer of augmented data on test cases contains the heat curve and emission information. With all these layers of data, the test case is capable of representing the dynamic cost nature of a deregulated power system and is able to dispatch generators similar to the real power system. PJM interconnection data was chosen to implement the proposed augmentation technique. The marginal cost result from optimal power flow (OPF) is compared with the marginal cost of energy of the PJM interconnection along with the GHG and AP emissions.

Smart-grids have opened opportunities for end customers to participate in the power system operation. DR is one of the activities that the end customers can perform to participate in the electricity market. Revenue earned from energy markets has been relatively low compared to DR used for capacity markets and ancillary services. An aggregated DR model participating in the bulk-power market as a service through a pool-based entity called demand response exchange (DRX) is proposed to improve the

benefits of DR to the market. The economic benefits to the market entities have been studied using the proposed augmented test cases.

The key contributions of this dissertation are:

- power systems test case generator data for researchers who do not have access to the real power system data,
- a technique that utilizes only open-source data to develop augmented data for any test case to represent the dispatch of a real power system in terms of cost, and emissions,
- a DR model capable of improving the revenue for DR participants in the bulk-energy market.

# CHAPTER 1 Introduction

## 1.1 Background

The power systems have evolved from a small local utility in the late 19th century to one of the largest industries in the U.S. in the 21st century. The electric power industry is adapting and evolving to keep up with the ever-growing need for electricity. The importance of power systems is only going to grow in the future with innovations in a generation to the delivery of electricity . The electric power industry is going to be the most critical industry, and the sustainability of this industry is vital for any nation.

The power system generation resources have been undergoing rapid changes over the last decade. Illustrated in Fig. 1.1 is the rapid change in U.S. electricity sector fuel mix over the last decade. The changes happened at such a rapid rate that, natural gas which was the third most used energy resource before 2008 is the primary source of energy in 2018 [1]. Renewable energy which includes wind, solar, and geothermal was the least among all generating sources before 2008 is producing more energy than hydropower as of 2018.

One of the prime reasons that allow an industry this large to change at such a rapid rate is deregulation [2]. The power system in the United States has witnessed significant regulatory changes over the past two decades. The deregulated power system has granted open access to the transmission system, which has encouraged competition between generators and distributors of electrical power [3]. Independent System Operators (ISO) is established as a central entity to organize and operate the bulk-power markets by facilitating the trade of electricity between power producers and consumers [4]. In a

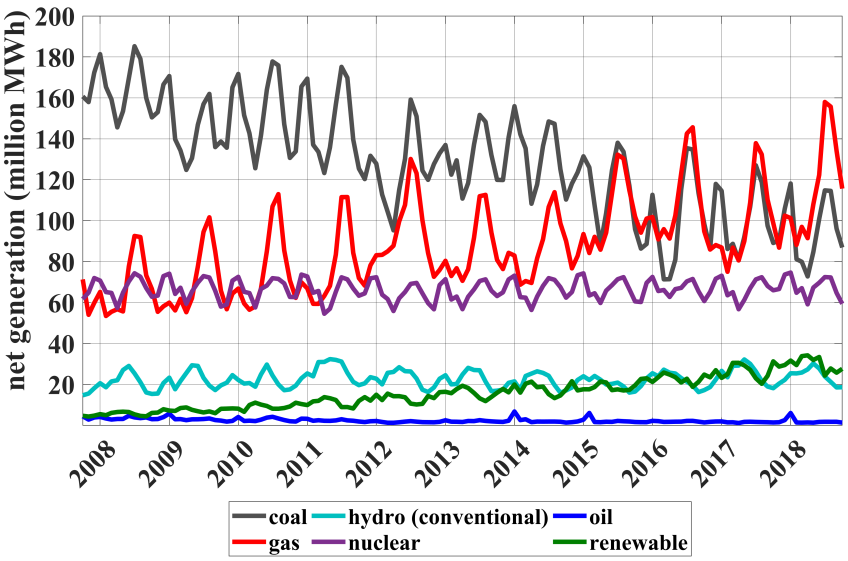



Figure 1.1. U.S. electric sector generation by major energy sources from 2008 to 2018 in one month resolution

deregulated power system, each generator and load-serving entities are independent actors, participating in the ISO power market by submitting bids/offers through a cyber network, making the cyber-physical power system (CPPS) one of the largest cyber-physical systems in operation.

The ability to operate the power system securely in the most economical way was the ultimate goal of any innovation. But with rising environmental awareness and global warming concerns, the goal of a power system operator should be to operate the system securely, economically, and also environmentally responsible. In 2016, the electric power industry generated the largest share (28%) of greenhouse gas (GHG) emissions across all the primary energy consuming sectors in the U.S. [5]. In 2016, approximately 68 % of the electricity was generated from fossil fuel combustion the majority of which is coal and natural gas [6]. Apart from GHG fossil-fueled power plants also produce harmful air pollutants (AP) such as $SO_2$ and $NO_x$ that reduce the air quality and are also the primary

cause for acid rain. All primary energy consuming sectors are taking steps to curb GHG emissions to control global warming. The electricity sector has an extensive research community that is developing technologies that would reduce GHG emissions.

The future of power systems research innovating in any sector from generation to delivery of power must improve at least one of the goals (security, economical operation, and environmental responsibility) while not affecting the others. All the physical and regulatory changes in the power system must be taken into account while making an assessment of the impact of the research. One of the recent physical change is the availability of communication technologies to power systems.

Advancement in communication technologies in power systems has opened opportunities for consumers to participate in the smart-grid operation actively. Consumers are participating either by changing their demand or by generating electricity. Consumers change their demand to improve either the physical security or economic operation of the grid in response to financial benefit to them [7]. The voluntary change in demand for financial benefits is called demand response (DR). There is a growing interest in studying the social-environmental impact of DR [8].

Advancement in solar cell technologies has made solar energy affordable to residential customers. Apart from solar, energy storage in the form of electric vehicles and batteries have formed an essential infrastructure to consumers [9]. Consumers participation in power systems either by distributed generation or demand response is being extensively studied for economic impact on the system[10]–[12]. Cheaper communication systems and availability of high computational power even at the consumer level have allowed high-level optimization between assets to improve the

revenue of power entities [13].

The power system infrastructure is a critical and expensive system to experiment on directly. All the state-of-art research for change in demand or generation are published based on simulations. The network information of a real power system interconnection is not publicly available due to security concerns. All research data are simulated on power system test cases that are a synthetic representation of the real system [14]. A realistic test case is a valuable piece of an asset for a power system researcher. A lot of research is being conducted recently to develop realistic state-of-art test cases [15]

Even though the existing state-of-art test cases are good enough to prove most of the research techniques, they are not adequate enough to make an investment decision based on the results because none of these test cases are capable of exhibiting the dynamic nature of an electricity market.

## 1.2 State-of-Art Power System Research

The rapid growth of the electric power industry is possible because of the intellectual research community of power systems. There is a strong motivation across the world to create a secure, economical and sustainable and clean electric power system [16]–[18]. Even though all through the 20$^{th}$ century the electric power industry grew as a regulated system, the cost aspects of the electricity was foreseeable as it was only based on the cost of fuel. At the dawn of 21$^{st}$ century the economic aspect has changed because of deregulation and the cost of electricity is not foreseeable [19].

### 1.2.1 Economic Analysis Research

The economic analysis, most importantly of those research in which the cost of electricity is a deciding factor will need a test that represents a deregulated power system. In one of the research works the authors propose an optimal strategy for ISO to schedule DR such that it maximizes the contribution of DR [20]. The proposed DR scheduling strategy was simulated on a IEEE-RTS test case, and the benefits in conclusion were drawn on the basis of the profit which depends on the actual cost of electricity.

The RTS-96 test case used for simulations is an older test that does not represent any electricity market-based cost functions [21]. One of the conclusions was related to emission reduction. Though the algorithm is capable of reducing the emissions, the exact amount that they claim through this simulation setup may not be the original estimate as the fuel mix of the power system had changed from 1996 when the test case was published to 2014 when this work was published.

One of the most commonly used objective function in analyzing a deregulated power system is the *surplus* minimization. One of the use cases for surplus minimization is in the study of congestion management [22] in which the author considered a deregulated operation on an IEEE 30-bus system. The proposed optimization algorithm is capable of scheduling loads and generators to reduce congestion to improve the economic reliable operation of the system. But, the actual savings shown may not reflect any real system, as the cost functions for generators used in the minimization problem were not based on any real system.

There are many other optimization techniques for demand response and

congestion management that are proposed to outperform the state-of-art and show improvement in revenue [23]–[25]. The results analysis of these works are performed based on a deregulated system, but the simulations are conducted on a system that was designed for a regulated power system. It is unclear if the revenue claims made in these publications can be transpired to the real-world revenue as there is no cost data affiliation to a real deregulated system.

### 1.2.2 Emissions Analysis Research

Power systems research to evaluate the effect smart grid initiatives like DR/DER on emissions is increasing. The pressure to reduce GHG and AP from the electric power sector is the driving point of this research. A smart city EMS work that presents the effect of residential solar power on the effect of $CO_2$ reduction simulated on a 31-bus distribution network [26]. The change in demand has to be simulated on a transmission level system to evaluate its effect on large fossil-fueled generators. Emissions are reduced only when the fossil-fueled generation is reduced.

One of the research works presents a novel technique to optimize demand based on pollutant emissions [27]. In this work, the authors utilize IEEE 14-bus, and the PJM Interconnection system to implement the proposed load management algorithm. Only the results from the IEEE test case were presented in detailed as the authors cannot publish the details of the PJM network. The results from the PJM network were only presented as proof of validation of their work. Very few researchers and grid operators have access to the real network, most of which cannot be published for security reasons.

There are multiple load optimization projects and researches that also look into

estimating and reduce emissions [28]–[30]. The test case serves the purpose of evaluating the functionality of the algorithm. A test case with real generators attributes(fuel mix, and emission factors) are required to make an assessment of the actual impact of these works on the actual emissions.

### 1.2.3 Demand Response Research

DR impact research is studied from residential EMS [31] to the market level [32]. DR is projected to be one of the most important assets in the smart-grid initiatives. Most of the economic benefits from DR are evaluated from the retailer's point of view. The economic benefits of DR are evaluated as a cost-reducing technique on the distribution side [33]–[35]. New regulations for DR participation in the market have been issued [36] to encourage more market-level participation. DR has emerged as a sizeable asset in the capacity and ancillary service market [37].

Despite the regulations, the revenue earned from the energy market is lower compared to the capacity market. Unlike the ancillary, and capacity market which improve physical security and reliability, the energy market is more of economic service. One of the DR models to participate in the market operation is by trading DR offers [25], [35], [38]. There is a strong motivation to explore the opportunities for DR in the energy market. A test case that can produce a time-series of prices similar to a deregulated power system is required to conduct an exploratory analysis of DR models in the energy market.

### 1.2.4 Sustainability Research

Sustainability of a critical industry like the electricity sector is vital to any nations' growth. According to a 1987 United Nations report, sustainable development meets our

need without compromising our future [39]. The goals of sustainable development as identified by the 2005 World Summit on Social Development are economic well-being, environmental sustainability, and social development [40].

In one of my publications, an aggregated DR model for residential customers is designed and the effectiveness of this model is presented as sustainability metrics [41][1]. The economic sustainability of this DR model has been proposed as the savings for the customer using the DR. The change of demand was simulated on a 6-bus network to evaluate environmental sustainability as a factor of reduced capacity factor of the marginal generating unit. The test case was modified to represent the installed capacity of a real power interconnection. The lack of real network generation information withheld us from validating the reduced emissions against a real system.

## 1.3 Objectives

There are many smart-grid initiatives that are aimed to make the power systems more reliable, economical and sustainable. A test case that is capable of making realistic costs and emissions assessment will be a valuable asset for the power systems research community. Given the data of a real electricity market, the core objectives of this dissertation are:

1. to develop market-based synthetic test case data that represents the dispatch of a real deregulated electricity market statistically,

2. use the market-based test cases to evaluate the sustainability of an aggregator-based DR participating in the bulk-power market.

[1]This work was performed jointly in collaboration with the full list of co-authors in [41], and was awarded the best paper award in the Smart City Conference at Bangkok, 2017

## 1.4 Contributions

Following contributions from this work are aimed to improve the existing state-of-art in power system research:

1. provide the power system research community with the power system test case data that represent the cost, and emission performance of a real power system,
2. a generalized technique to augment any standard test case using open-source data to perform the time-series simulation that can represent a real-power system costs and emissions,
3. understanding the factors affecting DR participants in energy markets to help improve their revenue,
4. ability to explore new smart-grid initiatives using the market-based synthetic test case data and evaluate its economic and environmental sustainability.

## 1.5 Dissertation Outline

In the Chapter 2 a literature review of the state-of-art test cases are provided and the need for developing the augmentation based test cases. The first data layer in the augmentation process is presented in Chapter 3 in which PJM market offer data is analyzed using unsupervised learning techniques to cluster the masked offer data. This data is sampled over to augment the test case cost function and is updated for every day in a time-series simulation. The proposed technique is tested on eight different test cases from six buses to 2000 buses. The next two layers of augmented test case data are the fuel-turbine type and the heat-curve and emission factors data for the generators. This

technique is presented in Chapter 4, where the market hourly fuel data is augmented on to the test case by co-relating the capacity factors. Finally in Chapter 5 a bulk-market level DR model is proposed, and its economic analysis is conducted using the proposed augmented test case.

# CHAPTER 2 State-of-Art Test Cases

A review of the present state-of-art test cases along with the missing information required to make market-level simulations is presented in this chapter. The framework to augment these test cases to perform similar to a real deregulated system is also presented in this chapter.

## 2.1 Power Systems Tests Cases

Power system test cases are developed to perform power flows that are used to conduct exploratory analysis on the design and performance of the system. Test cases are developed to statistically represent a real power system so that the results obtained from simulations reflect the real system performance. Power system researchers analyze the physical security such as the physical limits violation (voltage, and lines), and the economic analysis based optimal power flow/economic dispatch. Generators are dispatched based on a cost minimization problem called as economic dispatch (ED). The generator cost functions are critical information to perform optimal power flow (OPF) from which economic analysis can be performed.

There have been many efforts to develop test cases for realistic economic studies. An approximate model of the European interconnection was developed based on the real network to study the impact of cross-border trades used for transmission pricing and congestion management [42]. The developed model was adequate to study a specific problem of cross-border trade, but due to unavailability of the latest data, the authors used thermal and nuclear plant pricing data published in 1995, which is more than two decades out-of-date [43]. Recently, there have been efforts to develop large-scale synthetic test

cases that represent the complexity of today's electricity grid for energy economic studies. In [15], [44], two approaches were proposed to augment their developed synthetic test cases with generator cost data that represents the behavior of an actual power system. Unit commitment and economic dispatch for one day were simulated on the modified test case to illustrate its capabilities to perform economic studies.

Most existing literature assigns fuel-cost based cost functions to generators on test cases, whereas in the post-deregulation ISO-organized market, the market is cleared based on generator offer prices [45], [46]. There are multiple factors apart from the fuel costs that govern the offer strategy of a generator participating in a competitive electricity market [47]. There is sound research in bidding strategy for generator offers of various energy sources under different market scenarios, but there is a lacking methodology for using such bid strategies in test cases in which the generator types are unknown [47], [48]. In [49], a method is presented to calculate the average bid price for groups of generators participating in the day-ahead market, but the work does not discuss a method to use these prices in power system test cases.

## 2.2 Deregulated Power System Test Cases

Detailed generator characteristics for a realistic test case based on the IEEE-118 bus system were developed in [50]. A detailed generation model with constraints such as ramping rates, operational limits, and heat rates for the generators in the power system test case was presented. The generators in the modified IEEE-118 bus system use linear heat rates, and the generation capacity has been increased to 24,500 MW provided by 327 generators. The generator energy mix is based on the future (2025) projected installed

capacity of a South California. There is no way to validate the cost functions of this test case, nor its capability to perform similar to a market over a time-series.

There are two test cases that have been developed for simulations based on deregulated power systems with generator offers as cost functions. One of them is the ISO-NE 8-bus test system which models its 8 buses based on the 8 regions under the ISO-NE [51]. The cost functions are said to be developed from the market data but do not explain the technique to update them as the market offer costs keep changing. The other test case that is based on an ISO is the Federal Energy Regulatory Commission, Regional Transmission Operator Unit commitment Test System (FERC RTO UC) which has detailed generator information, but the network information is not open source [52].

Among these test cases, only the FERC RTO UC test system utilizes the offer data from a real electricity market (PJM) to develop detailed generator information [52]. Since the generator fuel information is masked in the PJM offer data, they utilized an unspecified statistical technique with few assumptions to match the offered size to the EIA-411 data which contains the installed capacity of a generator [53]. Even though the generator data from the test case is made open, the network information of this test case is considered as a part of Critical Energy Infrastructure Information (CEII).

## 2.3 Proposed Data-Driven Augmented Test Cases

Based on the literature presented in the Section 4.2, the test cases require generator cost functions that represent the electricity market generator offer data. Analyzing the literature presented in Section 2.2 it is clear that a fixed cost function is not capable of producing a time-series similar to a real-deregulated power market. To overcome these

issues an augmented data for any standard power system test case is proposed in this work.

The line, bus, and generator data which are developed for the state-of-art test cases are static in nature, and a do not change drastically over time. The dynamic data in a deregulated system are the cost functions of the generators. Even though the installed capacity of the generation changes, the rate at which this change occurs is not high. However, the fuel mix that contributes to this capacity is changing at a rapid rate.

The generator cost and fuel mix have to be updated dynamically for any time-series simulations that represent the dispatch of a deregulated power system. Fig. 2.1 represents the various data that are required to augment these test cases to represent the dispatch of a real power system. All the data required to develop the augmented data comes from open-sources.

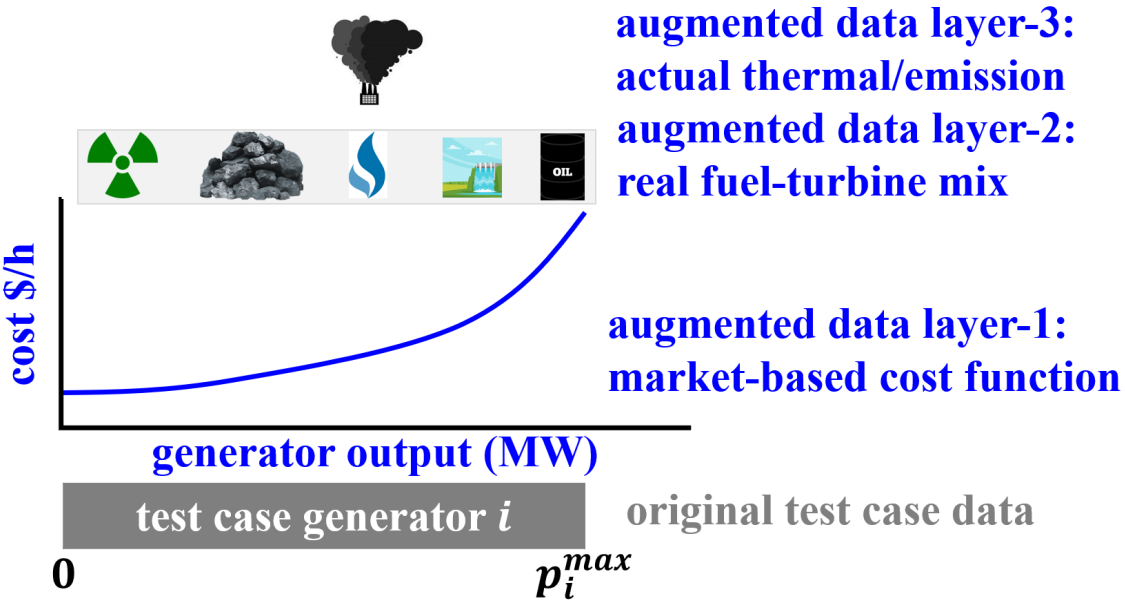



Figure 2.1. A graphical representation of the proposed three-layer augmented data for test cases to represent a real market.

The first layer of augmented data over a test case generator is assigning a cost function which is developed based on the generator offer bids submitted to an electricity

market. The method used to develop the augmented cost function data is presented in Chapter 3[2]. Based on the dispatch of the test case generators using the market-based cost functions, fuel-turbine type is assigned based on the hourly fuel mix of the real system. Over this data, a heat curve and emission factors data are assigned to the test case generator. The data required to develop the fuel, and emission data are obtained from publicly available data. The method used to develop the augmented fuel, and emission data is discussed in Chapter 4.

[2]This data is a part of the publication [54],and is publicly available in a open-source repository at `https://github.com/Dvenkat30`

# CHAPTER 3 Market-Based Generator Cost Functions for Power System Test Cases

## 3.1 Introduction

Every generation company participating in the bulk-power market is an independent entity trying to maximize their profit. Therefore, the fundamentals of economic analysis to determine the marginal cost of energy has to be changed from minimizing the total sum of generator *fuel costs*, to minimizing the cost of discrete generator *offers (bids)* [55]. Fuel cost is just one of the deciding factor in determining the $/MWh offer price for a generator participating in an ISO-based electricity market [49]. Because the ISOs determine the marginal energy cost from minimizing generator offers, to reflect these prices in simulations, test cases should be provided with market-based generator offers. Test cases that reflect market-offers are required for performing analysis of power system studies, such as demand response, where price of electricity is a deciding factor [56], [57]. Detailed power system test cases with realistic cost functions that represent actual electricity market wholesale prices are highly valuable to industry and academia.

Test cases with generator cost functions based on fuel costs were accurate to perform economic studies of power systems before restructuring. Simulation studies are an essential tool for power system engineers during planning, research, and operation. Simulation results for new policy or operational changes should reflect real world costs for deciding on whether or not to adopt the changes. Most power system test cases available for simulation were developed based on the Mid Western American Electrical Power System in the 1960's and 1970's [58]. The generator cost functions used in these test cases

are developed based on then fuel prices and heat curves of mostly thermal units. Studies such as contingency analysis and voltage stability can still be performed without issue on such systems, but accurate economic studies (e.g., demand response) cannot be performed without updates to the fuel-cost based generator cost functions.

In this chapter, a methodology to use publicly available market data to design market-based generator cost functions is presented. These cost functions are used to augment existing power system test cases that can accurately represent real marginal energy costs for use in economic studies. Hourly marginal energy costs and demand for most U.S. electricity markets is publicly available. Additionally, daily generator offer data is available with a four-month delay [59]. Because the generator offer data is sanitized of identifying information, a statistical pattern recognition (K-means clustering) to cluster the market generators into *types*. The three generator types ($K = 3$ clusters) identified from the PJM market (base, intermediate, and peak generators) are used to fit market-based generator cost functions. These new cost functions are used to augment existing power system test cases to more accurately represent market behavior of marginal costs.

This methodology is used to augment eight power system test cases of different sizes to represent the generator mix of a real electricity market. Optimal power flow (OPF) was executed on these test cases, and the marginal energy costs produced by the simulation are compared to the real day-ahead market prices. The major contributions of this chapter are:

1. the design of a methodology to augment existing power system test cases with market-based generator cost functions to represent the the marginal energy costs of real

power markets for use in economic studies,

2. a comparison of the proposed approach to the existing state-of-the-art fuel-cost based power system test cases for multiple year-long OPF simulations with one-hour resolution.

The following section of this chapter introduces the existing state-of-the-art fuel-cost based power system test cases and illustrates the need for market-based cost functions for use in analyzing economic impact in power system simulations. The methodology for augmenting power system test cases with market-based generator cost functions is described in Section 3.5. In Section 3.8, the eight power system test cases are presented and the simulation parameters are given. A comparative study of the marginal energy price of each augmented test case is conducted for 2014–2016 in a one-hour resolution in Section 3.9. The output of our method is compared to the actual PJM market and the existing fuel-cost based method. Concluding remarks and directions for future work are presented in Section 3.10.

## 3.2 Economic Analysis of Electric Power Systems

### 3.2.1 Overview of Existing Techniques

A brief description of the calculation of the marginal cost of energy is presented in this section. Some of the key regulatory changes have caused existing fuel-based generator cost functions inadequate for simulation studies with an economical focus (such as the impact of demand response or distributed energy resources). Economic studies in power systems aim to reduce the cost of energy and to improve economic efficiency. In most of these studies, the deciding factor in adopting a method depends on the marginal price of

energy, which can obtained from the result of OPF [60]. In the following subsection, generator cost functions and the concept of OPF are introduced. An illustrative example comparing the marginal energy costs of PJM versus the existing state-of-art fuel-based generator cost functions highlights the need for market-based generator cost functions when trying to calculate energy costs in an ISO organized market. The section concludes with changes to generator costs due to regulatory changes post market restructure.

### 3.2.2 Optimal Power Flow

OPF is the core component of economic studies, introduced in the 1960's [61], [62]. In this dissertation, all simulations on power system test cases were carried out using the OPF formulated in MATPOWER [63]. To illustrate the cost minimization problem of the power system, OPF can be formulated in its most simplified form without considering some of the constraints (e.g., bus voltage limits) in Eqs. (5.1–5.5). For generator $i$, let $C_i$ be the cost function, and $P_i$ be generation output.

$$\min_{P_i} \sum_{i=1}^{N} C_i(P_i) \tag{3.1}$$

$$C_i(P_i) = \alpha_i P_i^2 + \beta_i P_i + \gamma_i \tag{3.2}$$

$$\sum_{i=1}^{N} P_i = \sum_{j=1}^{M} D_j + P_L \tag{3.3}$$

subject to,

$$P_i^{min} \leq P_i \leq P_i^{max}, \forall i \tag{3.4}$$

$$0 \leq P_{ij} \leq P_{ij}^{max}, \forall i, j \tag{3.5}$$

The objective function Eq. (5.1) is subject to constraints Eqs. (5.3) and (5.5), where $D_j$ is the demand at location $j$ among $M$ load points on the network. At any given instant, the total generation must equal the total demand plus system transmission losses ($P_L$), as represented in Eq. (5.3). Eq. (5.4) describes the generator operational limits, and Eq. (5.5) describes the line flow limits for all lines/transformers in the network. Eq. (5.2) represents the polynomial cost of operation of generator $i$ with coefficients $\alpha_i$ (\$/MWh$^2$), $\beta_i$ (\$/MWh), and $\gamma_i$ (\$/h). The Lagrangian multipliers of this optimization problem provide the marginal energy cost of the system [61]. The marginal cost of energy is the cost incurred in the production of one additional unit (MWh). When line limits are reached, generators must be dispatched non-optimally (i.e., different marginal costs), and the marginal price of the entire system cannot be equal. The difference in marginal energy cost is called a congestion cost, and the sum of congestion cost and marginal energy cost is the location marginal price (LMP). Every generator is paid the LMP at the common point of coupling (POC) to the electric grid, and every load needs to pay the LMP at its POC. The last committed generator is called the marginal generator, and its incremental offer decides the cost of bulk electricity at every POC.

## 3.3 Fuel-Cost Based Generator Cost Functions

Generator cost functions in the state-of-the-art synthetic test cases [44] are derived based on input-output characteristics, efficiency, and fuel costs of the major energy contributors like natural gas, coal, nuclear, and hydro/renewable. The input-output characteristic of a thermal generator is the ability to convert thermal energy into electrical energy; these data may be obtained from design parameters of that generator. The cost function for a thermal generator $i$ can be represented as Eq. (3.6), where $a_{0i}$ (\$/h) is its no-load cost to operate, and $b_{1i}$ (MBtu/MWh) and $b_{2i}$ (MBtu/MWh$^2$) are the quadratic coefficients of the thermal input-output curve of that generator with fuel cost $F_i$, expressed as \$/MBtu. To obtain the quadratic cost coefficients for a generator, Eq. (3.6) relates to Eq. (5.2), where $F_i b_{2i}$ equals $\alpha_i$, $F_i b_{1i}$ equals $\beta_i$, and $a_{0i}$ equals $\gamma_i$. The cost function of hydro/renewable generators comprises only no-load cost, because they are non-thermal units and do not have fuel costs associated for generating electricity.

$$C_i(P_i) = a_{0i} + F_i(b_{1i}P_i + b_{2i}P_i^2) \tag{3.6}$$

The reliability test system of 1996 (RTS-96) is one of the most utilized test case for power flow studies. This test case is chosen to built the state of art test case by augmenting such that it represents the real-world generation mix of the PJM region for the year 2014 [64]. The cost functions are developed based on fuel costs of the major energy sources during 2014 [1] and is presented in Table 3.1. The cost functions are developed using the no-load cost and the fuel costs of the thermal generators and thermal data from

EIA [65], and only the no-load costs for the non-thermal generators.

To illustrate the capabilities of the test case, OPF was simulated for one year using MATPOWER [63] with a demand curve with a one-hour resolution. The marginal energy cost of the test case were compared to the PJM marginal energy cost for 2014 using an appropriately scaled demand curve from PJM for the same time period. In this dissertation, I will use the notation that lower case variables describe quantities used for a power system *test case*, and capital variables describe quantities of the real power network. To simulate OPF at hour $t$, the load $d_j(t)$ on any bus $j$ is obtained using Eq. (4.1), where $d'_j$ is the default demand on bus $j$ provided in the test case, $D_{mkt}(t)$ is the total demand on the real PJM network, and $D^{max}_{mkt}$ is the annual peak load of PJM for the year in consideration (i.e., 2014 for this simulation). A scaling factor, $\psi$, is used to scale the default load such that the ratio of load to generation is similar to that of the real market. For this simulation, the scaling factor was set at $\psi = 0.96$. The scaling factor is described in detail in Section 3.8. The weighted average marginal price of all nodes in the test case is compared to the marginal price of the PJM market. The weighted average marginal price is defined using Eq. (3.8), where $\Lambda(t)$ is the weighted average marginal price of electricity (in \$/MWh) of the the test case for hour $t$, and $\lambda_j(t)$ is the LMP at bus $j$ among the $s$ buses of the power system test case.

$$d_j(t) = \psi d'_j \frac{D_{mkt}(t)}{D^{max}_{mkt}}, \ \forall j = 1, ..., s \tag{3.7}$$

$$\Lambda(t) = \frac{\sum_{j=1}^{s} (\lambda_j(t) \times d_j(t))}{\sum_{j=1}^{s} d_j(t)} \tag{3.8}$$

Table 3.1. Monthly average fuel costs of coal and natural gas for the year 2014.

| **month** | **coal ($/MBtu)** | **natural gas ($/MBtu)** |
|---|---|---|
| January | 2.29 | 7.02 |
| February | 2.32 | 7.40 |
| March | 2.36 | 6.00 |
| April | 2.39 | 5.07 |
| May | 2.40 | 4.93 |
| June | 2.38 | 4.84 |
| July | 2.38 | 4.43 |
| August | 2.37 | 4.12 |
| September | 2.37 | 4.20 |
| October | 2.31 | 4.10 |
| November | 2.30 | 4.48 |
| December | 2.51 | 4.36 |

The simulation was performed on the RTS96 test case using the 2014 PJM market data by scaling the annual demand curve using Eq. (4.1). Nuclear, coal, natural gas-fired, hydro, wind, and solar are the six types of generators defined in the test case. The thermal conversion rates for the nuclear, coal, and natural gas-fired used in this simulation were based on [15]. The other three types of generation were non-thermal and had no associated fuel cost. Though the simulation is performed using one-hour resolution, the fuel cost data obtained from Energy Information Administration (EIA) [66] was the averaged fuel cost price per month. Fuel cost of nuclear was almost constant, and 0.85 $/MMBtu was used [44]. The fuel cost of the two other major energy resources of the test case, coal and natural gas, are presented in Table 3.1. The results of OPF are compared to the marginal energy cost of PJM in Fig. 3.1. The year 2014 was a very cold winter for North America, triggering a gas shortage for electricity, and resulted in a large spike in electricity prices. This effect did not appear in the simulated results, as the gas prices only varied by a factor of 1.7 throughout the year, compared to a factor of 100 in energy costs. These results establish that the cost of electricity in a market is not just based on the fuel

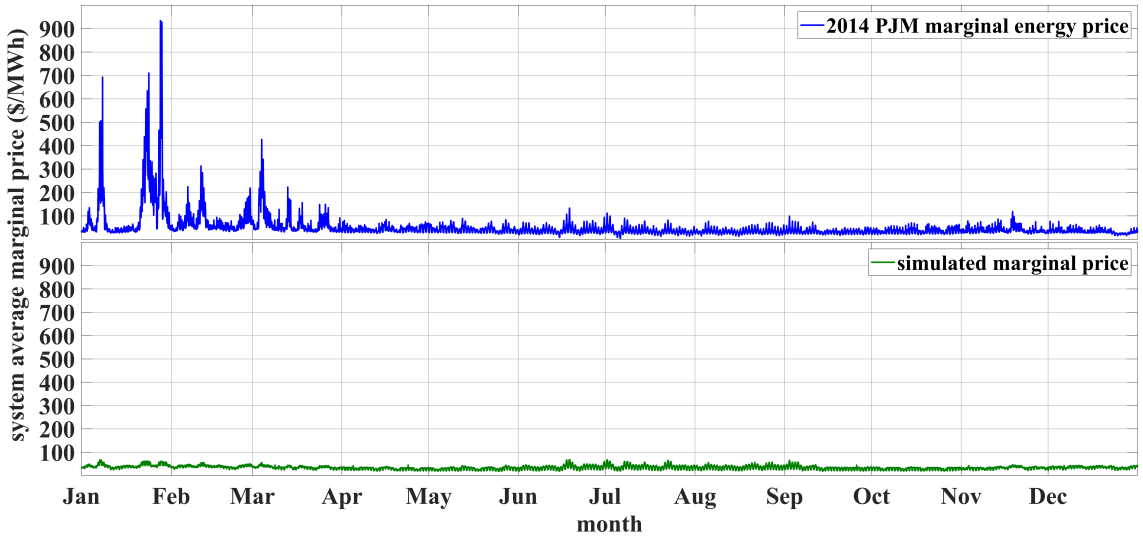



Figure 3.1. Comparison of the marginal cost of energy of PJM in 2014 (top-curve in blue) to the average energy cost simulated on RTS-96 using a fuel-cost based generator cost functions for a scaled PJM demand (bottom-curve in green).

cost model of the marginal generator. Existing methods and test cases do not adequately represent real market costs, which may lead to large errors in analyzing the economic impact of new power systems technologies as stated in research using existing test cases.

## 3.4 Generators in an Electricity Market

In this section, the role of the generator offer in an organized market is presented in comparison with the fuel-cost method used prior to market restructuring. Generators that wish to sell electricity need to participate in an electricity market in their region. The cost of electricity is decided by the offer price of the marginal generator. Eligible generating entities need to submit the size of the generator and offer price to the ISO, along with their operational constraints (e.g., runtimes and economic/operational limits). The offer price of each generator is submitted in incremental blocks of their generators capacity ($/MW). An example of a generator's offer from the PJM market is shown in Fig. 3.3 in Section 3.5. Apart from the incremental costs, the generators must also submit their

start-up and reserve costs which have to be paid to the generators if they are turned on or keep them in spinning reserve respectively.

The price of bulk power is determined in a day-ahead market under each ISO across the U.S. The electricity markets perform unit commitment and economic dispatch one day ahead of the supply day to schedule generation. The total cost of generation is minimized based on generator offers and demand bids submitted by generating entities and load-serving entities, respectively. Each generator submits a daily offer to the electricity market containing the following information:

1. incremental offer costs (energy cost per segment output range $/MWh vs. MW),
2. upper and lower limits of units economic operation in MW,
3. start-up and no-load cost,
4. time operational limits such as minimum runtime and maximum number of starts in a day,
5. maximum and minimum economic, operational limits.

This information is published publicly on the electricity market websites as per the FERC ordered lag period of four months [67]. The number of offer blocks can range from one to ten steps in increasing order of energy and cost. Some markets have a restriction on the maximum offer price, for example PJM restricts their generators to bid a maximum of 1,000 $/MW in their day-ahead energy market [68]. Some large generators who cannot shut down in a short notice, like nuclear plants, submit negative offer prices to ensure they are fully-committed [60]. Each market has to maintain some percentage of generation as a

reserve, and the market needs to pay the generator the spinning reserve cost mentioned in the offer. The identity of the generators is protected so that this data is not used for market malpractices, such as gaming the offer price to outbid other competitive generators. An electricity market uses all the information mentioned above in the offers to construct an optimization problem to minimize the total cost of electricity for the required demand, resulting in hourly marginal energy prices for each location on their network.

## 3.5 Market-Based Synthetic Cost Curves

### 3.5.1 Classification of Generator Types

Due to cyber and physical security threats to the power system, the real grid data is not available in the public domain for simulation purposes. For economic studies, generator cost functions on available simulation test cases need to be similar to those in real bulk-power markets. There are a number of test cases that provide realistic line limits and line parameters (e.g., impedance, admittance) that physically represent real power networks, but they lack realistic cost functions. In this chapter I describe a general methodology to turn real generator market offer data into market-based generator cost curves and assign them to power system test cases, expanded from our prior work in [69]. The resulting augmented test cases emulate real market marginal energy prices in simulations.

The number of generators on real power systems is much higher than the number of generators on most test cases. Generators with similar fuel types and comparable sizes submit similar offers [70]. Even though the generation mix of a region is known, it is difficult to recognize the type of generator (e.g., coal, nuclear, gas) by observing

individual offers submitted to an electricity market as this data is not revealed (due to security concerns). An unsupervised pattern recognition technique is required to obtain groups of generators with similar bidding and operational attributes (i.e., generator *types*), but may not be of the same fuel type. Applying a statistical pattern recognition (SPR) technique on the market offer data creates clusters of generators that are statistically similar, and these generator types can be used to serve as a synthetic representation of the generation mix of the electricity market. The idea of this work is to interpolate cost functions of each generator type, and apply these cost functions to the generators on the test cases in the same ratio as in the synthetic generation mix.

The generator offer data published by an electricity market contains various information about the generator. The number of dimensions (i.e., features) of the dataset is crucial for an accurate SPR. Among the various information that is published in the offer data of a generator, as described in the previous subsection, the following features were selected for SPR:

1. weighted average offer price ($/MWh),

2. generator size (MW),

3. minimum runtime per day (h), and

4. ratio of minimum economic operation limit to the maximum operational limit.

Among the various features tested, these features gave consistent clusters for the selected SPR described in Section 3.8. A generator's offer can contain anywhere from one to ten incremental blocks, with the number of blocks represented by $B_i$. For generator $i$,

the selling price of energy is decided by generator output based on the offer curve defined by offer blocks $\mathscr{W}_i = \{w_{i1}, ..., w_{iB_i}\}$ (MW) versus the corresponding offer price $\mathscr{R}_i = \{r_{i1}, ..., r_{iB_i}\}$ ($/MWh). The weighted average offer of unit $i$, $\mathscr{O}_i$ ($/MWh), is defined by Eq. (3.9), where $b$ is the block index. Along with the average price of generation, another important feature for classifying generators is the size of the generating unit in MW (obtained as the quantity of the last offer block), as it separates the generators into clearly defined types (e.g., base versus peak) via the SPR. The minimum runtime is an operational constraint specifying the minimum number of hours a generator has to be committed in a day. The runtime information is used in classifying the generator types based on system loading conditions such as base-load, peak-load, and intermediate load [71]. The last feature of the dataset describes the ratio of minimum operation limit to maximum operation limit to represent the flexibility of the generator (e.g., a base-load generator will be inflexible with a ratio near 1, and a peak-load generator will be more flexible with a higher ratio).

$$\mathscr{O}_i = \frac{\sum_{b=1}^{B_i} (r_{ib} w_{ib})}{\sum_{b=1}^{B_i} w_{ib}} \tag{3.9}$$

## 3.6 Assigning Market Generators to Test Case Generators

Using the feature set described in the previous subsection and a suitable SPR, the *market* generator offer dataset is classified into $K$ clusters, which serves as statistical generator *types*. To create similar classification in the power system test case, the *test case* generators are also classified into $K$ types, with each type having a similar percentage of

generation capacity on the test case as the generator types from the market data. The process of grouping the test case into $K$ types is described in Algorithm (1). First, the mean generation size of each cluster in the market data is calculated. The $K$ market clusters are arranged in descending order of their capacities, and their percent share of the total generation is evaluated as $y_1, ..., y_K$, where $y_1$ is the percent share of the cluster with the largest mean generator size.

A similar procedure is then carried out with the test case generators, by arranging the generators in descending order of their generation capacities. An index to point at a test case generator is initialized with one. The test case generation capacity of generators representing the generation type $k$ is determined as $\rho_k$, which is $y_k\%$ of the total generation $p^{max}$ as shown in Step 10. Each generator's capacity ($p_i$) is subtracted from the generator *type's* capacity until the total capacity is met with $n_k$ generators. This process starts from the largest test case generator, $n_1$ generators are chosen in descending order of their generation capacity to make a test case generator type that contains $y_1$ percent of the total generation on the test case. This process is repeated to determine the $K$ generator types on the test case from the $y_2, ..., y_K$ generation percent from the market clusters.

For illustrative purposes, let us consider a 6-bus, 11-generator system [72] of 240 MW capacity with market offer data that is clustered into $K = 3$, with $y_1 = 59.5\%$, $y_2 = 19.2\%$, and $y_3 = 21.4\%$ (from Table 3.2). The four largest test case generators (three 40 MW, and one 20 MW) representing 140 MW (i.e., 58.3% of the percent share of the total test case generation) would be assigned to market generator type 1. The process would continue, assigning three generators with 20 MW capacity representing 25% to type 2, and assigning four generators (one with 20 MW, one with 10 MW, and two with 5

MW) to type 3 to represent the remaining percentage. The following subsection describes the process of assigning the market generator offers to the generator cost curves for each test case generator.

**Algorithm 1** Algorithm to assign test case generators to a generator type obtained from SPR of market generator offer data.

**Input:** day-ahead market generator offer data and test case generator data
1: extract the required features from the day-ahead market generator offer data
2: perform SPR with an optimal number of clusters ($K$)
3: determine the mean size of generator capacity in each cluster
4: arrange clusters in descending order of mean generator capacity
5: determine the percent share of capacity of each cluster $(y_1, ..., y_K)$
6: arrange test case generators in descending order of their capacity
7: initialize the index for test case generator $i = 1$
8: **for** $k = 1$ to $K$ **do**
9: initialize number of generators in *type k* ($n_k = 0$)
10: determine the capacity of test case generation classified as *type k* ($\rho_k = y_k p^{max}$)
11: **while** $\rho_k \geq 0$ **do**
12: subtract the generator $i$ from the required capacity $\rho_k = \rho_k - p_i$
13: assign generator $i$ to type $k$
14: increase the generator index $i = i + 1$
15: increase the generator count $n_k = n_k + 1$
16: **end while**
17: **end for**
**Output:** test case with statistical classification of generator types

## 3.7 Offer-based Generator Cost Curve Fitting

### 3.7.1 Overview

Each generator in the test case is assigned a market-based cost function derived from the market generator offer data of the generator *type* it is assigned. OPF tools, like MATPOWER [63], require generator costs represented as piece-wise linear or polynomial functions. I investigate two techniques in this section of the chapter to design generator costs from the clustered market offer data: (a) second-order polynomial and (b) piece-wise linear functions, described in the following subsections.

### 3.7.2 Polynomial Cost Curve

Most existing test cases are provided with quadratic cost functions ($/h), as they were designed based on the quadratic heat rates of thermal units, i.e., Eq. (3.6). In this section, I replace the existing fuel-based cost functions with equivalent market generator offer-based cost functions ($/h). The offer-based cost function of generator $i$ is obtained by (i) multiplying the elements of the offer price, $r_b \in \mathscr{R}_i$, with the corresponding offer block, $w_b \in \mathscr{W}_i$, to obtain the offer rate, and then (ii) fitting a second-order polynomial to determine the coefficients $\alpha_i$, $\beta_i$, and $\gamma_i$ from Eq. (5.2). The resulting offer-based cost functions are used to augment the test case generator cost functions.

For all the generators with offers having three or more offer blocks (i.e., $B_i \geq 3$), a quadratic equation is fit to the offer-rate curve using the least-squared error method. For example, a generator offer with $B_i = 7$ is converted into an offer-rate curve as shown in Fig. 3.2 in red, and a second-order polynomial is fit to the offer-rate curve using least-squared error represented in the green curve. The resulting coefficients, $\alpha_i = 0.031$ \$/MWh$^2$, $\beta_i = 8.75$ \$/MWh, and $\gamma_i = 510.93$ \$/h, would represent the offer-based cost function for generator $i$ when placed onto a test case. The offer has a minimum limit of 270 MW at 5068 \$/h. The fitted quadratic curve is plotted between the generators operation limits between 270 MW and 800 MW.

For offers with two blocks, a linear curve is fit between the two operational points resulting in an $\alpha_i = 0$. Generators with offers of a single block have a $\gamma_i$ equivalent to the offer rate, with $\alpha_i = \beta_i = 0$.

Algorithm 2 is used to augment the generator's on the existing test cases with the

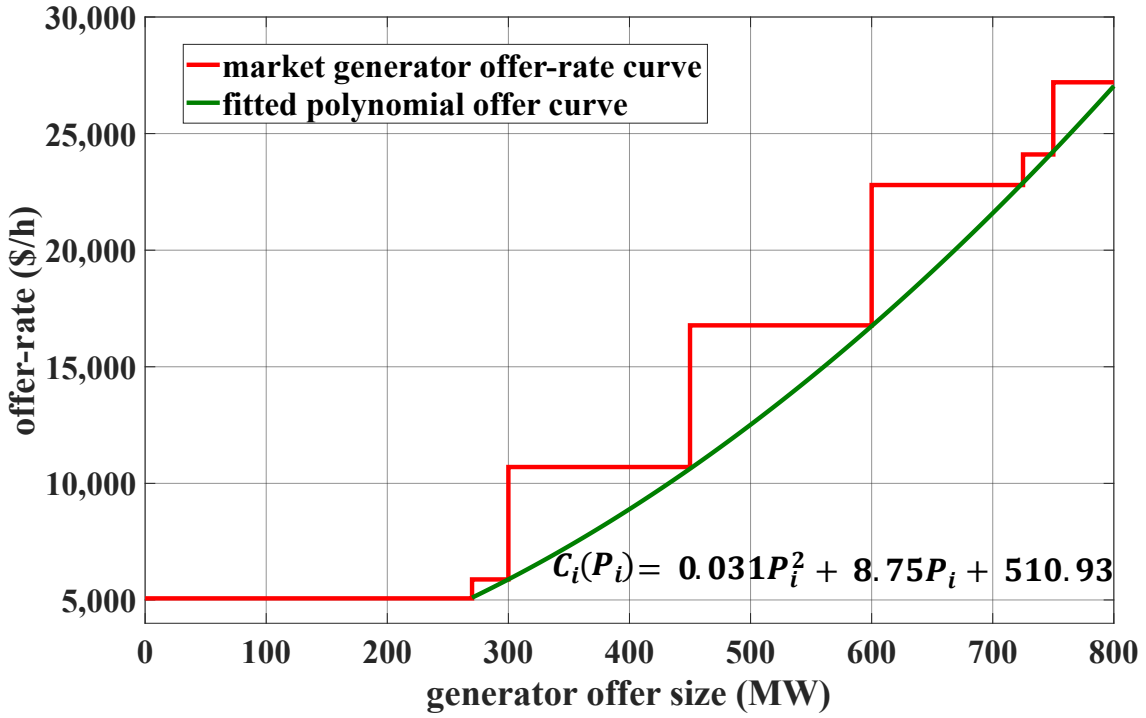



Figure 3.2. Fitting a second-order polynomial as the generator market offer-based cost function (green curve) using least-squared error onto an 800 MW generator's offer-rate curve with seven blocks (red curve).

market offer-based cost functions. The proposed approach starts by converting the day-ahead market offer of each generator to a second-order polynomial market offer-based cost function, as described above. The number of test case generators ($n_k$) for each generator *type* are known, and $n_k$ offer-based cost functions are randomly selected from the corresponding cluster $k$. Each test case generator that represents a generator type is assigned a market-based cost polynomial that is randomly chosen from the cluster it represents. This approach ensures similar generation mix and costs in the test case as the market it represents.

As shown in Fig. 3.1, the market LMP varies daily, because the generator offer data changes daily (with hourly differences due to changing demand). This algorithm

should be re-run for each day of interest to obtain accurate market simulations through time. Some generator offers do not have a good fit for a second-order polynomial. Such polynomials are eliminated from the dataset before starting the assignment process. The following subsection discuses an approach to overcome this issue.

**Algorithm 2** Algorithm to augment existing test cases with second-order polynomial market offer-based cost functions.

**Input:** market-based classified test case, clustered generator market data
1: convert every offer into offer-rate curve
2: fit a second-order polynomial to each offer-rate curve
3: **for** $k = 1$ to $K$ **do**
4: select $n_k$ cost functions randomly from cluster $k$
5: assign a cost function to every generator in the $n_k$ sized test case cluster
6: **end for**
**Output:** updated test case generator cost functions

### 3.7.3 Piecewise Linear Cost Curve

Not all generators in the bulk-power market submit offers that can be assigned a quadratic cost curve with a good fit. Such offers that result in a poor fit may not accurately represent the actual generator offer price. Most power system simulation software are capable of performing OPF using piecewise linear cost functions. For this approach, instead of fitting a second-order polynomial to the offer-rate curve, the market generator offers are directly used by scaling the offer quantity to fit the test case generator size. With this approach, the offer price of a generator remains the same on a test case as in the market, even though the offer has characteristics that makes a second-order polynomial a poor fit (i.e., a higher-order polynomial would be required for a good fit).

The process of developing the clusters for the piecewise linear approach is similar to that of polynomial approach. Instead of developing offer-rate curves as in the previous

approach, this approach takes the incremental offers $\mathscr{R}_i$ directly from the market data for fitting a piecewise linear cost function for the test cases. To utilize these offers to develop cost functions for test cases using Algorithm 3, the market offers must be extrapolated from 0 MW to the generator's maximum operational limit. The market offer is partitioned into $s$ equi-capacity blocks equal to $1/s$ of the generator's maximum output. These partitioned offer prices are used as offers at $s$ equi-capacity blocks for the test case generator. In this study, the lower limit of operation is neglected for all generators in the test cases because MATPOWER OPF does not consider unit commitment, hence even in the partition of the market offer the lower operational limit of a generator is ignored. The cost function assignment process is similar to that of the polynomial approach, where the test case generators are grouped into generator *types* and cost functions are chosen from their respective market generator offer dataset clusters.

Fig. 3.3 illustrates an example of converting the same 800 MW generator market offer shown in Fig. 3.2 to a 200 MW test case generator's offer. The original offer submitted to the electricity market on Jan. 28, 2014, had submitted an operational limit between 270 MW and 800 MW. If this generator offer is selected to augment a 200 MW test case generator with $s = 5$, the lower limit is higher than $1/s$ of the generator capacity, the lower limit is extrapolated with the first submitted offer price of 18.7 \$/MWh. The market offer cost function shown in red in Fig. 3.3 is partitioned into five equi-capacity blocks shown as the dashed blue curve. The five offer prices of the partioned curve is used as the five offer blocks for the test case offer curve shown in green. Based on the software that is used for the OPF, the offer blocks of the test case can be used as incremental step offers or piecewise linear offers as shown as the green dashed lines linking each

incremental block in Fig. 3.3.

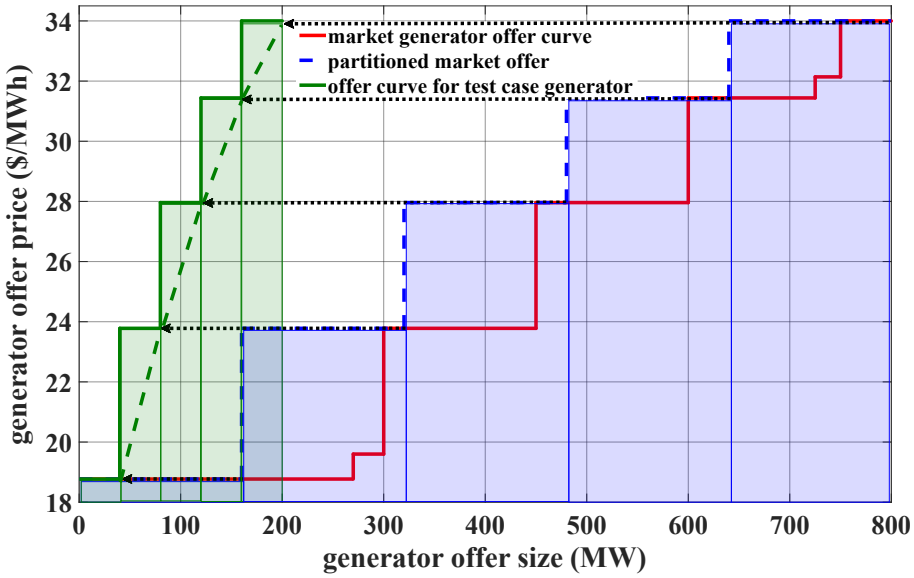



Figure 3.3. Developing generator offer curve for a 200 MW test case generator (shown in green blocks) by partitioning an 800 MW generator offer submitted to the market (partitioned curve in blue dashed-blocks and the offer curve submitted to the market in red). The piecewise linear curve of the developed generator offer is shown in green dashed line.

**Algorithm 3** Algorithm to augment existing test case with market-based piecewise linear cost function.

**Input:** market-based classified test case, clustered generator market data

1: determine the $s$ equi-capacity extrapolated market offer price intercepts for every generator
2: **for** $k = 1$ to $K$ **do**
3: divide each test case generator capacity into $s$ equi-capacity incremental offers
4: select $n_k$ piecewise market offer intercepts randomly from cluster $k$
5: assign the randomly selected incremental offer price to each test case generator
6: **end for**

**Output:** updated test case generator cost functions

## 3.8 Simulation Setup

### 3.8.1 Simulation Overview

The methodology of assigning market-based generator cost functions is tested using the two different techniques to assign offer data to generator cost functions in eight

standard test cases of varying size (six buses to 2,000 buses). The SPR from Section 3.5.1 was implemented using K-means clustering. Any unsupervised learning technique can be used to implement Algorithm 1. The choice of SPR is not explored in this work, as the main contribution of the work is the design of the methodology to implement realistic market behavior on power system test cases. Any electricity market that publishes the generator offer data can be used for developing the cost functions. At the point of writing this dissertation, I am aware of at least two markets that publish the generator offer data publicly (a) PJM [67], and (b) ISO-NE [73]. I chose PJM generator offer data to develop generator cost functions for the eight test cases because of the nature of wide range of data availability that would support other projects in this dissertation. Marginal energy prices obtained from OPF on these test cases are statistically compared to the real PJM market marginal energy price. Negative price offers were eliminated as these small number of offers ($\sim$1%) did not impact the marginal cost of energy. The test cases were assigned scaled PJM hourly demand to each test case for the years 2014–2016. This section will describe in detail the K-means clustering method (including our choice of K), statistics of each cluster formed by K-means on PJM generator offer data, PJM market demand and generation statistics, and an introduction to the power system test cases used in this study.

### 3.8.2 K-means Clustering

In this work, I chose to implement K-means as an initial baseline implementation of the SPR. K-means is a commonly applied unsupervised learning technique for pattern recognition that is computationally efficient and produces well-separated clusters for well-defined data [74]. The "$K$" value for the market generator-offer data is chosen such

that the clusters produced have maximum separation and are well-defined. Well-defined clusters are clusters that have a maximum distance between cluster centroids, and least distance between each element within a cluster. I use the Calinski-Harabasz criterion (CHC) to determine the optimal number of clusters $K$ for our offer dataset [75]. I chose CHC for this problem as CHC provides a K that gives maximum separation between cluster groups so that the generator *types* can be properly identified. With $N$ observations (i.e., generator offers) and $K$ clusters, CHC can be determined by Eq. (3.10). CHC is a maximization criterion that is directly proportional to the inter-cluster variance, defined in Eq. (3.11), and is inversely proportional to the intra-cluster variance, as defined in Eq. (3.12). For cluster $k$, let $n_k$ be the number of observations, $m_k$ be the centroid, and $x$ be a multi-dimension data-point. Additionally, let $m$ be the centroid of the dataset.

$$\mathrm{CHC}_K = \frac{SS_B}{SS_W} \times \frac{(N-K)}{(K-1)} \tag{3.10}$$

$$SS_B = \sum_{k=1}^{K} n_k \|m_k - m\|^2 \tag{3.11}$$

$$SS_W = \sum_{k=1}^{K} \sum_{x \in k} \|x - m_k\|^2 \tag{3.12}$$

The CHC was determined for the electricity market data for each day of the year for the years 2014–2016 using the four-feature dataset. Fig. 3.4 is the histogram for the $K$ value corresponding to the maximum CHC for each day's PJM generators offer data for the year 2014. The optimal number of clusters based on CHC was found to be three for 358 days out of 365 for the year 2014. Similar observations were found for the years 2015

and 2016, where all 365 daily datasets had maximum CHC with 3 clusters. For that reason, I use $K = 3$ for all further simulations.

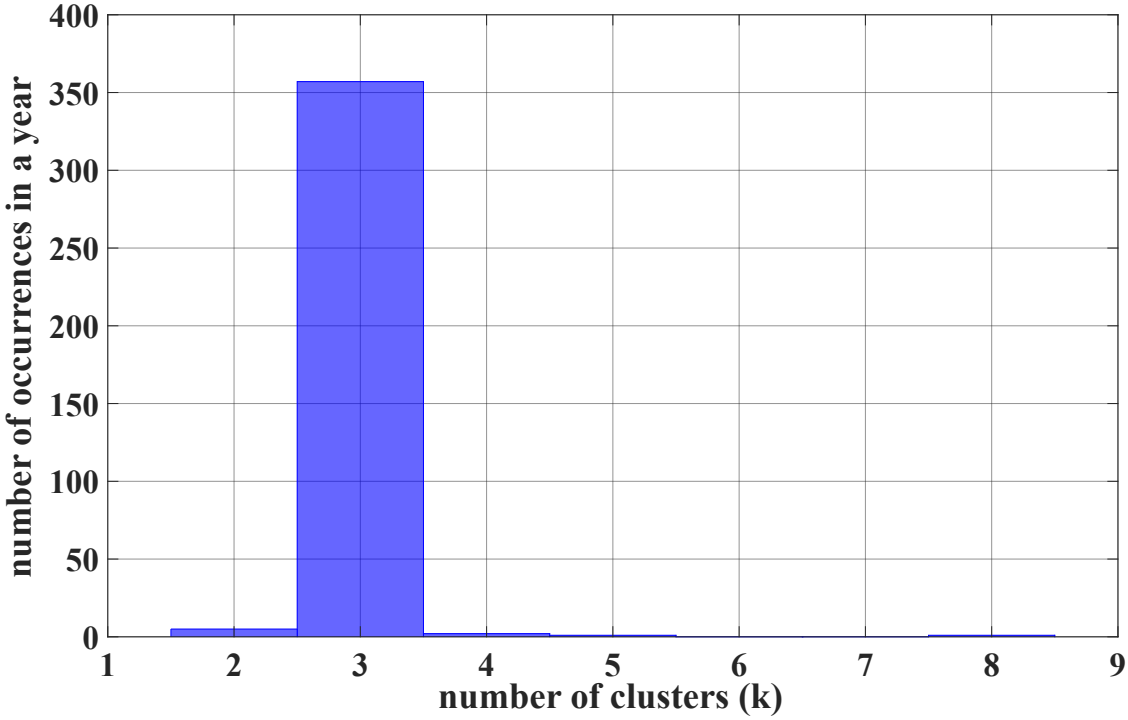



Figure 3.4. Histogram of the optimal number of clusters (K) for the PJM generator's offer data for the year 2014 using Calinski-Harabasz criterion.

After clusters of generator types were determined via K-means, the generator types were labeled based on the analysis of each feature of the cluster data using power system domain knowledge. Base-load units, intermediate-load units, and peak-load units are the labels given to the three clusters. A base-load generator is defined as a generator that operates 24 hours per day, intermediate units are those that meet the daily peaks (operating a few hours daily), and peak-load generators are operated only during annual peaks (a few hours annually) [71]. Similar minimum runtimes of generator types can be observed in Fig. 3.5. In this cumulative density plot of PJM generator offer data on Jan. 28, 2014, the

cluster that has more than 90% of the generators with 24 hour minimum runtime has been labeled as the base-load units, and the generators with lower minimum run-times as peak-load and intermediate based on their cluster average weighted offer price.

Fig. 3.6 shows the cluster of the generators based on the offer size and offer cost of the generator for the same day in January, as those are the features that are used to augment the test cases. The generators in each cluster appear to be overlapping as the data contains more than two plotted dimensions (i.e., four dimensions from the four features selected for the SPR). Table 3.2 presents the market offer data statistics based on the clusters shown in Fig. 3.6. The mean generator size of the base-load units is the largest, and the peak-load units are the smallest. The generation share of these three clusters would serve as the share of generators in test cases for any simulation based on this day (i.e., $y_1, y_2, y_3$).

Table 3.2. K-means cluster summary for PJM bid offer data on Jan. 28, 2014.

| **unit type** | **generation share (%)** | **mean offer quantity (MW)** | **mean offer price ($/MWh)** |
|---|---|---|---|
| **peak-load** | 21.4 | 83.1 | 687.9 |
| **intermediate** | 19.2 | 122.9 | 177.6 |
| **base-load** | 59.5 | 311.8 | 180.8 |

Jan. 28, 2014, was the peak price day of the year with the marginal energy cost in the PJM region reaching 965 $/MWh at 16:00. The generator distribution on July 6, 2014, which was the peak *demand* day of the year of 139,571 MW, is presented in Fig. 3.7. Comparing the distributions of the two extreme days show that the offer price of the generators does not depend only on the demand of the system. On a given day, the price of electricity follows the demand as the offer price of a generator is fixed for that day, but can

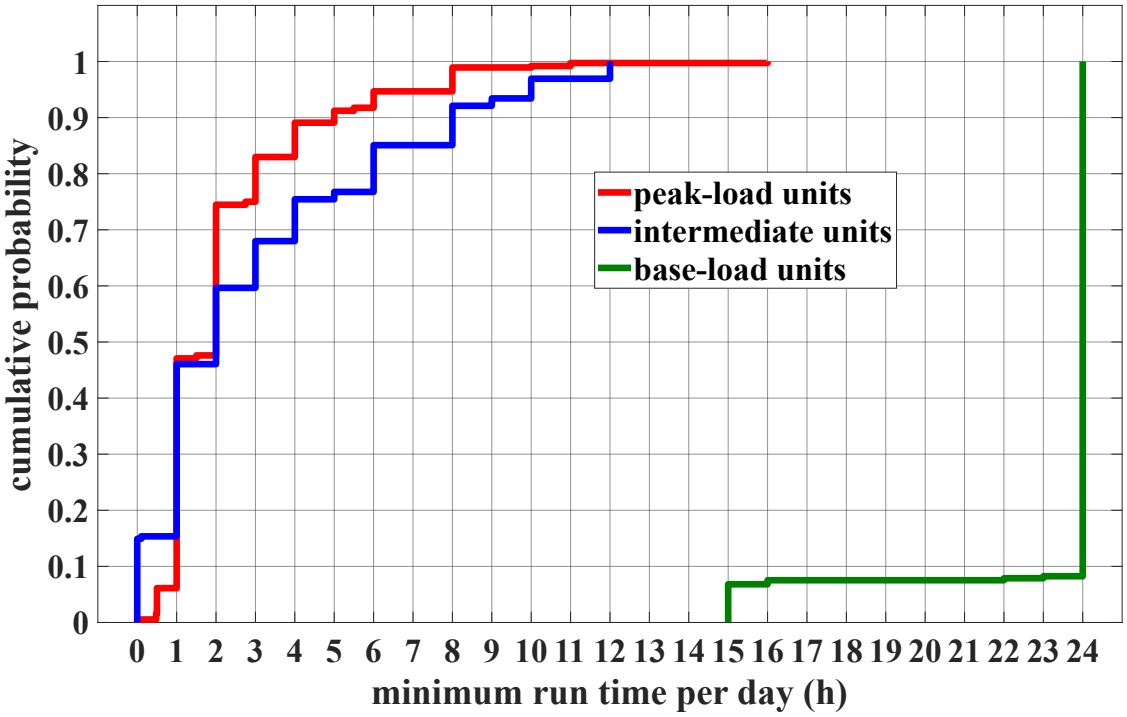



Figure 3.5. Cumulative density function of the minimum runtime of generators in each cluster for the market offer data of PJM on Jan. 28, 2014.

change throughout the year as fuel prices and other expenses keep fluctuating. Table 3.3 shows the offer data statistics of July 6, 2014, where the share of each cluster does not change by a large margin, but the mean offer price is approximately six times less than that of Jan. 28. The lower generator offer price of July 6 would result in lower electricity prices, despite greater demand than that of Jan. 28, which would not be apparent in existing test case cost functions.

Table 3.3. K-means cluster summary for PJM bid offer data on July 6, 2014.

| **unit type** | **generation share (%)** | **mean offer quantity (MW)** | **mean offer price ($/MWh)** |
|---|---|---|---|
| **peak-load** | 18.3 | 60.0 | 213.6 |
| **intermediate** | 23.3 | 156.8 | 103.3 |
| **base-load** | 58.4 | 359.5 | 37.2 |

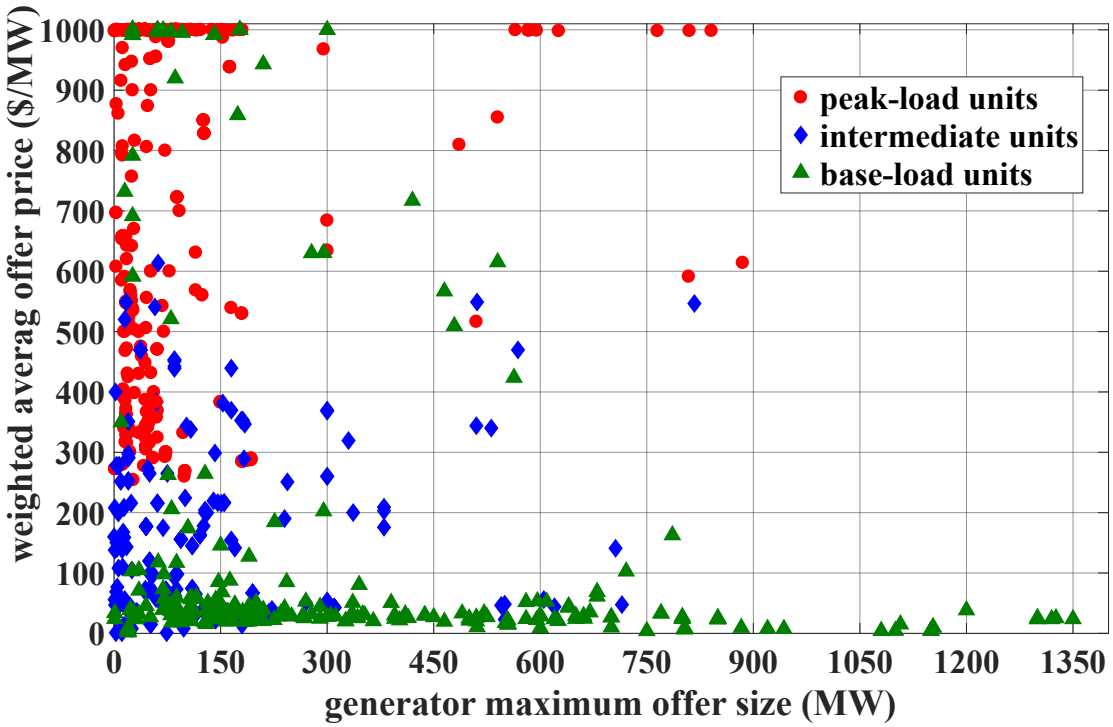



Figure 3.6. Distribution of all the generators by weighted average offer price to their offer size, clustered into three partitions using K-means for the generators participating in the PJM market on Jan. 28, 2014.

### 3.8.3 Principal Components Analysis Based Clustering

A separate clustered generator data-set is developed for each day for the entire time-series intended for analysis. Generator cost functions are developed by polynomial curve fitting from the market offer data which is used as cost functions for the test cases. From the knowledge of power system I know that base-load generators are usually large efficient generators and peak-load are usually small fast ramping generators, this can be also visualized from the Fig. 3.8. The cost functions for the test cases are also assigned in such a way that the largest generators have the curves from the base-load cluster and in descending order through the intermediate units to the smallest units assigned from peak-load cluster. As the number of generators in a test case are much fewer than than the

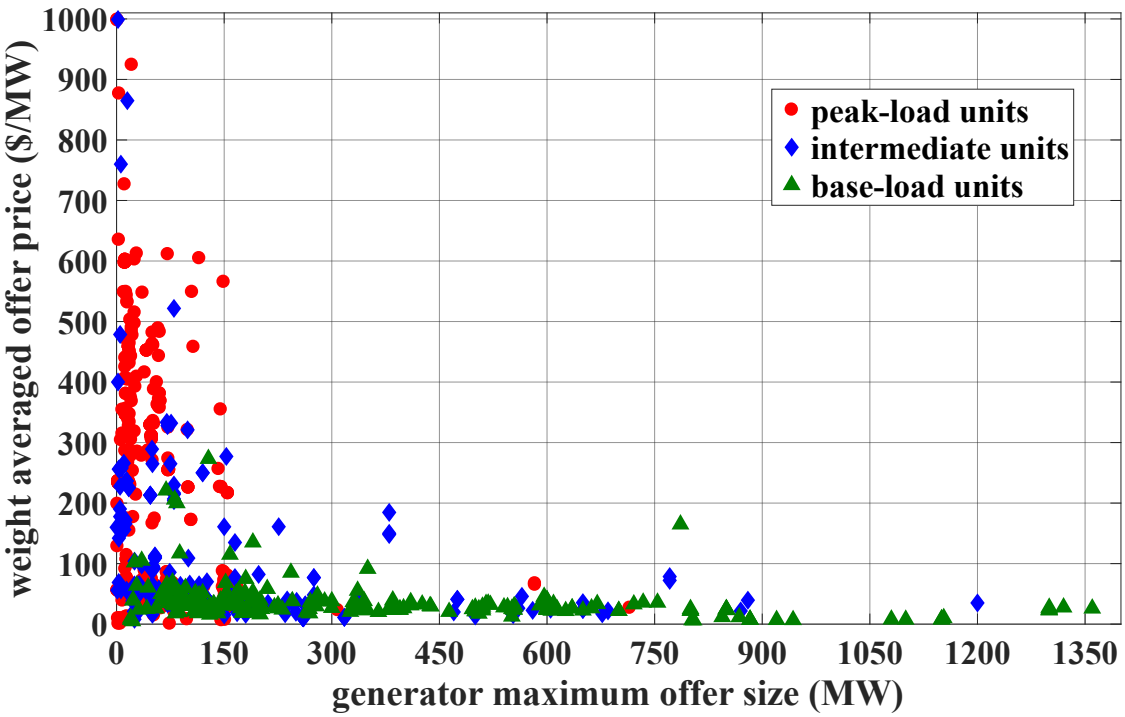



Figure 3.7. Distribution of all the generators by weighted average offer price to their offer size clustered into three partitions using K-means for the generators participating in the PJM market on peak demand day of the year, July 6, 2014.

market generators cluster, cost functions are randomly sampled from the data and assigned to test case generators. Each cluster is samples in such a way that the test case also has similar percentage share of each cluster group.

Under-sampling issues can arise when randomly selecting few generator offers from a cluster with large number of generators. For example, in the peak-load cluster as shown Fig. 3.8 the generators spread from a small generator with its capacity in few 10's of MW offered at nearly 900$/MWh to large 700 MW generators with offer price lower than the base-load generators. Because the identity of all these generators are masked, there is no technical knowledge to filter out any outliers. When a cost function is randomly sampled from these clusters the probability of picking up any generator within

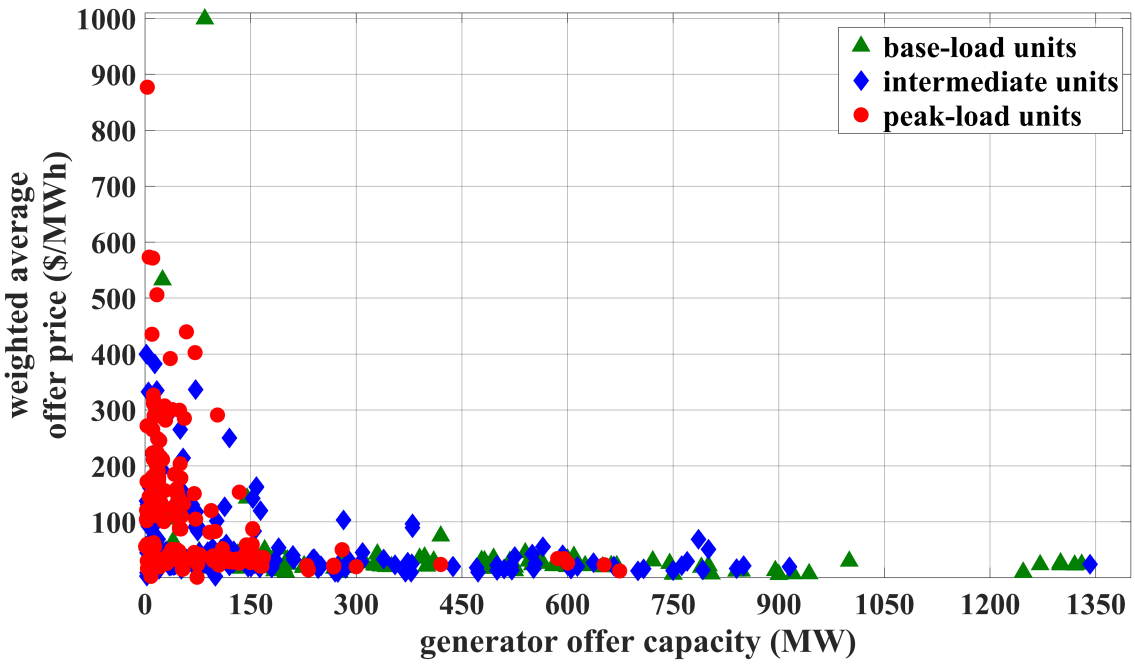



Figure 3.8. Distribution of all the PJM generators participating on July 25, 2016 day-ahead market along the weighted average offer price and offer capacity, clustered into three partitions using k-Means.

the cluster is equal, which sometimes can reduce the accuracy of this method.

To reduce the under-sampling issue the generator cost functions must be selected based on the density of the generators in a cluster [76]. One of the well established methods of estimating the maximum likelihood or density of a multi-dimensional data is by principal component analysis (PCA) [77]. PCA is a dimensionality reduction technique that projects the data on axes where the variance under projection is maximal. Using PCA the multi-dimensional generator offer data was reduced to two dimensions along the major axes of principal components. A bi-variate histogram is determined for each of the cluster along the two principal component (PC) axes with the number of bins determined by the Scott formula presented in [78]. Fig. 3.9 represents the probability distribution of the peak-load generators of July 25, 2016 along the PC1 and PC2 axes. Because the PCA is

just transformation and rotation of data to maximize the variance along an axis, the original information of the data can be traced back to extract the required information.

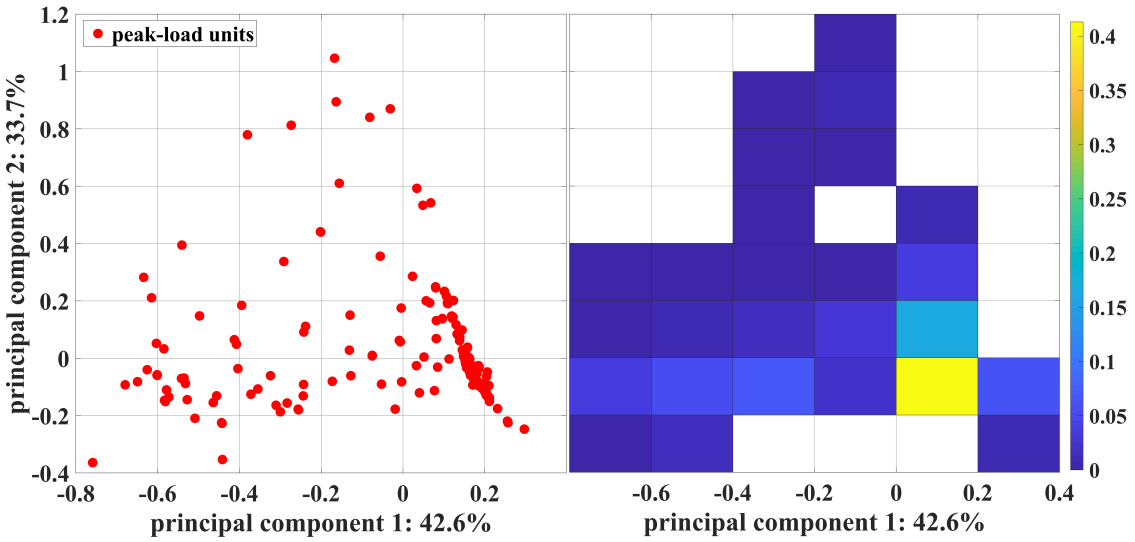



Figure 3.9. Normalized PCA of the peak load cluster along the PC1 and PC2 on the reght, and two dimensional histogram along the two major principal components of the peak-load generators cluster of PJM on July 25, 2016 on the left. The histogram represents the density based on number of generators in each bin.

The generator offers are sampled based on the probability of the bi-variate distribution of generators in each cluster. This process will reduce the chance of outlier generators being selected for the test cases. This density-based selection will increase the chances of a test case generator being classified and perform similar to other test case generators of the cluster. For example, I would expect a large test case generator to behave as base load unit, and the cost function of such a unit must be reflect the characteristics of a base-load unit consistently over the time-series. This consistency of being assigned the cost functions from the most likelihood region of a cluster is important to this work as a particular test case generator will be associated with the same fuel-generator type over the time-series.

### 3.8.4 PJM Market

This chapter aims to compare the marginal energy prices from the OPF results of the augmented test cases to the marginal energy prices in the actual PJM electricity market annually with a one hour resolution from 2014–2016. For a fair comparison, the test cases and the PJM market are statistically matched regarding demand and generation, described in Table 3.4. The average demand is calculated as the mean hourly demand in the calendar year. The peak demand is the largest hourly demand during the calendar year, and the generation capacity is the sum of the maximum operating limits of the generator offers on that particular day. The market demand factor (MDF) is obtained by the ratio of peak annual demand of the PJM market ($D^{max}$) to the generation capacity ($P^{D^{max}}$), given in Eq.( 3.13). This ratio gives the maximum percentage of generation utilized during that year. This information is used to scale the PJM demand curve to match the test cases. For this study, MDF$=80\%$ was chosen for each of the three years.

$$\text{MDF} = \frac{D^{max}}{P^{D^{max}}} \tag{3.13}$$

Table 3.4. PJM annual demand and generation summary.

| **year** | **average demand (MW)** | **peak demand (MW)** | **generation capacity (MW)** | **MDF (%)** |
|---|---|---|---|---|
| 2014 | 89,317 | 139,571 | 176,126 | 79 |
| 2015 | 87,712 | 136,510 | 181,674 | 75 |
| 2016 | 87,793 | 142,928 | 178,064 | 80 |

### 3.8.5 Power System Test Cases for Simulation

Eight power systems test cases ranging from six to 2,000 buses were tested using the two proposed offer-based generator cost curve fitting techniques. The eight test cases used are Roy Billinton test case (RBTS89) [72], IEEE 14-bus system (IEEE14) [79], IEEE Reliability test case 1979 (RTS79) [14], IEEE 39-bus New England test case (NE-93) [80], IEEE Reliability test case 1996 (RTS96) [21], ACTIVSg200 Synthetic Illinois 200-bus power system model (IL200), ACTIVSg500 Synthetic South Carolina 500-bus power system model (SC500), and ACTIVSg2000 Synthetic Texas 2000-bus power system model (TX2000) [81]. Table 3.5 presents the details of the size and capacities of each test case considered for this study. The test case demand factor (TCDF) is the ratio of the sum of all test case loads to the sum of the maximum generation of the test case generators. To match the PJM market MDF of 80%, the default loads on the test case were scaled using the scaling factor $\psi$, defined in Eq. (3.14), where $p'$ is the total generation capacity of the test case, and $d'$ is the default load on the test case.

Due to OPF convergence issues during some peak demand and demand valley times when using the scaled PJM load on some of the test cases, the scaling factor $\psi$ was adjusted. In Table 3.5, the rows marked with a '*' have been adjusted to ensure convergence. The exact methodology for determining $\psi$ for these test cases is described in the following subsections. The minimum generation limit of all generators in the RTS79/96, IL200, SC500, and TX2000 test case were set to 0 MW, as described in Section 3.7.3 (all other test cases have default minimum limits of 0 MW). To maintain load model accuracy, the same scaling factor used for active power as described in

Eq. (3.14) has been used to scale the reactive power demand at each bus.

$$\psi = \frac{p'}{d'}\text{MDF} = \frac{\text{MDF}}{\text{TCDF}} \tag{3.14}$$

Table 3.5. Power system test cases with their default setup and their scaling factors considered for this simulation.

| **test cases** | **number of buses** | **installed capacity (MW)** | **default load (MW)** | **TCDF (%)** | $\psi$ |
|---|---|---|---|---|---|
| **RBTS89** | 6 | 240 | 185 | 77 | 1.04 |
| **IEEE14** | 14 | 772 | 259 | 34 | 2.20* |
| **RTS79** | 24 | 3,405 | 2,850 | 84 | 0.96 |
| **NE39** | 39 | 7,367 | 6,254 | 84 | 0.88* |
| **RTS96** | 73 | 10,215 | 8,868 | 84 | 0.96 |
| **IL200** | 200 | 3,379 | 1,475 | 44 | 1.60* |
| **SC500** | 500 | 12,189 | 7,750 | 64 | 1.25 |
| **TX2000** | 2000 | 94,791 | 67,109 | 71 | 1.10* |

### 3.8.6 RBTS89

The RBTS89 test case is a small system with relatively high line flow limits that cannot be overloaded; this test case converges for all loading conditions. The largest generator of the RBTS89 system is 40 MW, which is smaller than the average peak-load generator of the PJM market. The cost function for a peak-load generator of the test case is developed based on PJM generator offers that are on an average ten times larger. The relative difference in size of generators creates cost functions that have steep variation due to the reduced size of the generator when compared to the market. There are eleven generators in this test case which represent each of the three generator types. Due to a small number of generators, a small variation in demand can sometimes create a large jump in price, as the chance of the marginal generator changing from one type to another

is high.

### 3.8.7 IEEE test cases

The IEEE14 test case has no line limits, and the only operational constraint is the generator maximum limits and voltage limits of the buses. This test case has only five generators varying between 100 MW and 334 MW. This small mix of generators and results in large step changes in hourly price over annual simulations because the smallest generator of the test case is twice as the mean generator size of the peak-load generator of the PJM market. With just five generators, it is difficult to accurately group generators to resemble similar generation share as the market clusters. Even though the test case can converge for a scaling factor that would result in 80% loading, the resulting marginal prices are very high as the small changes in load (inter-day demand changes) does not result in marginal generator choice.

This is similar to the NE39 test case, where there are ten generators varying between 508 MW and 1100 MW. The smallest generator on the test case is at least fifteen times larger than the average peak-load generator of the PJM market. Just like the 14-bus system, an accurate percentage share of clusters cannot be formed. These two test cases represent a part of the power system which does not need to have the entire diversity of generation profile. The scaling factor on this test case has been lowered for the same reason as the IEEE14, where a higher scaling factor would result in much higher median marginal energy price.

### 3.8.8 IEEE Reliability test case

The RTS79 test case has realistic line limits, voltage limits, and generation limits. This test case has 32 generators varying between 12 MW and 400 MW. Even though this test case was proposed for performing reliability studies, it is still a good test case to conduct economic studies because of the realistic generation mix and line limits. The RTS96 test case is comprised of three RTS79 systems, interconnected by five lines, making a total of 96 generators and 73 buses. The generation and demand mix on each of the three subsystems of the RTS96 test case remains unchanged from RTS79.

### 3.8.9 Synthetic test cases

Three synthetic test cases (IL200, SC500, and TX2000) were developed using statistical techniques [15]. They are designed to represent the electrical grid of a geographical region by capacity. The IL200 test case represents a hypothetical grid in southern Illinois, U.S., with 200 buses and 49 generators that range from 4 MW to 569 MW. This test case has a default load of 44% of its maximum generation, which was scaled to 70% using $\psi = 1.6$. This value is 10% less than would be calculated from Eq. (3.14), as beyond that percentage the upper voltage limits on a few buses would be violated. On some buses, there is little-to-no load which causes a voltage rise as the generation increases to meet the load on other buses. The SC500 test case represents a region of South Carolina, U.S., with 500 buses and 90 generators ranging between 1 MW and 772 MW. This test case was scaled by a factor of $\psi = 1.25$, which is equal to the value calculated from Eq. (3.14), because this case converges for all loading conditions. The TX2000 test case is a synthetic network that covers the entire Texas region with 2000

buses and 544 generators ranging between 1 MW to 1354 MW, which is very close to the original PJM market. Even though the upper limit of scaling is achievable, this was the only system that did not converge for any load that was scaled 84% below its default load. The simulation for this test case was set up such that the loads on this test case are scaled so that the minimum demand on the system is 84% of the default load of the test case, and the peak load is 110% of the default load. The scaling for the TX2000 is achieved using Eq. (3.15), to ensure convergence. With this scaling the ratio of the valley load to the peak load over the the annual simulation would be 77% (i.e., 0.84/1.1). The ratio of minimum demand to the maximum demand on PJM network for the year 2014 was 39%.

$$d_j(t) = \psi \times d_j \times \frac{(1.1 - 0.84) \times (D_{mkt}(t) - D_{mkt}^{min})}{D_{mkt}^{max} - D_{mkt}^{min}} + 0.84 \tag{3.15}$$

## 3.9 Simulation and Results

### 3.9.1 Simulation Overview

The capability of the proposed approaches is illustrated by comparing the energy prices from the OPF with the real PJM market marginal energy price. Simulations were performed on all eight test cases with the scaled PJM demand using Eq. (4.1) with the scaling factor $\psi$ as mentioned in Table 3.5. All simulations were carried out using MATPOWER 6.0 in MATLAB (R2017a)[3]. While the simulation and analysis were conducted for 2014–2016, only the results from 2014 are presented and analyzed in detail

[3]All data from this work has been made publicly available in a GitHub repository with an open source license at `https://goo.gl/rSGeBX`

for brevity. The energy prices of 2014 are interesting as that year witnessed high variation in energy prices across the year due to its very cold winter temperatures in North America and unavailability of natural-gas for electricity.

### 3.9.2 OPF Simulation

The market offer-based polynomial and piecewise linear approach simulations were performed using the generator cost functions derived using the proposed methods. The cost functions for the test cases were updated in one-day resolution. The same load curve was used for simulating all three setups (default, polynomial, and piecewise linear) derived using Eq. (4.1) for each test case, except the TX2000 which uses Eq. (3.15). All the marginal energy prices presented in the results are the weighted average energy price using Eq. (3.8).

To compare the hourly energy prices over the one-year time-series simulations with the PJM marginal energy prices, both visual and statistical techniques were investigated. One such statistical method is to compare the distribution of the hourly marginal energy prices from the augmented test case OPF with that of PJM over the same year. In this case, the distribution is represented as violin plots, as shown in Fig. 3.10. Violin plots are similar to boxplots, but the probability density of the values can be observed, represented by the width of the violin plot. The dashed black lines in each of the violin plots represent the three quartiles of the distribution with the middle line representing the median energy price. The two ends of the violin plot represent the extremes of the observed data. In Fig. 3.10 the first violin plot is the energy price distribution of PJM market during 2014 represented in the red plot. The distribution of

marginal energy price of each test case using the polynomial approach is represented in the left half of each violin plot in blue, and on the right in green using the piecewise linear approach. The distribution of marginal energy price of each test case using the two proposed techniques is compared with the PJM market marginal energy distribution. Though the cost functions for all the test cases were drawn from the same pool of cluster data, the performance of each test case is different, as the physical structure and generator configuration are different.

From the Fig. 3.10 I can observe that the test cases having realistic line limits, no convergence issues, realistic generation profile in terms of numbers and sizes (i.e, both the RTS system, and SC500) produced energy price distribution similar to that of the PJM market. Though the TX2000 test case has all the characteristics of a real power system, due to its lower bound convergence issues the energy prices have a higher median. However, the TX2000 price distribution shape closely resembles the PJM market distribution. The IL200 has the least matching distribution due to the upper limit convergence issues.

The marginal energy prices produced by OPF using the two approaches have comparable distributions. Regarding the density and mean of the marginal electricity price, both the RTS cases and the SC500 test case are similar to the PJM market. One reason these test cases perform well is the availability of various sized generators similar to that of the real market. The distribution of energy prices from the simulation of RBTS89, IEEE14, and NE39 are similar, and have a narrow violin plot when compared to the PJM market. A narrow violin plot represents less density of observations around the mean, indicating more variation in marginal energy costs. The IL200 test case produced

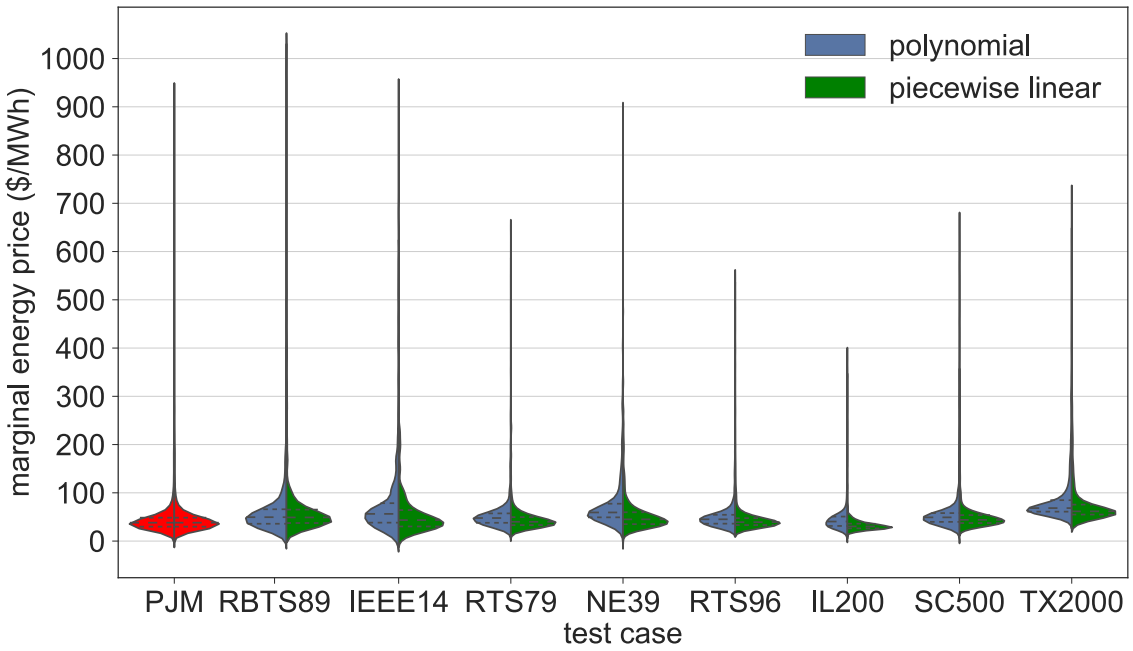



Figure 3.10. Violin plots representing the distribution of marginal energy price of the PJM market for the year 2014 in red and by simulation on test cases using polynomial approach presented in blue on the left half, and using the piecewise linear approach presented in green on the right side of the violin plot.

the lowest mean, as the upper scaling of this test case was reduced due to the voltage constraint convergence issues.

Another statistical metric to compare the performance of the proposed technique is the goodness of fit (GOF), given in Eq. (3.16), where $\Lambda_{PJM}$ is the marginal energy cost of the PJM market, and $\Lambda$ is the marginal cost resulting from the augmented test case. GOF provides a metric between -∞ to 1, where 1 is an exact fit and -∞ indicates a very poor fit.

$$GOF = 1 - \frac{\|\Lambda_{PJM} - \Lambda\|}{\|\Lambda_{PJM} - \Lambda_{PJM}^{mean}\|} \tag{3.16}$$

Table 3.6 presents the GOF, and the distribution of marginal prices produced by

each test case using the two proposed approaches compared to the original PJM market for 2014. The table compares the OPF simulation results using the mean and standard deviation of marginal energy cost for one year with a one-hour resolution. The first row shows the results from one year OPF using the single set of default cost functions that are provided with the test cases, except for results of SC500 and TX2000 which are simulated using the heat rates of the generators and monthly fuel costs (as described in Section 3.3. The distribution of the SC500 marginal price for default cost curve is derived from the plot presented in Fig. 3.1.

The mean and standard deviation of the SC500 and the two RTS systems are closest to the real PJM distribution. Even though the TX2000 system has physical dimensions comparable to the real system, due to the simulation limitations in scaling the load, a much higher mean is observed. The simulation results using the scaled 2015, and 2016 showed similar trends where the RTS and SC500 had the best marginal energy price distribution and GOF among the other test cases. The SC500 had obtained a GOF of 0.54 using the polynomial approach for the year 2015 and the least performed test case was IL200. The year 2016 had low fluctuation in marginal price, and a low peak, because of this the TX2000 could only achieve a 0.04 GOF for the year 2016 as the median price from the simulation was higher than the first quartile price of the PJM market.

None of the results from simulations using default cost functions could generate a positive GOF for the year 2014, but managed to get a positive GOF of 0.11 for SC500 and 0.04 for RTS96 using 2016 demand. The RTS79 using the piecewise linear cost functions produced the closest marginal price distribution in terms of mean and standard deviation to the PJM marginal price for 2014. The GOF of RTS79 is also the highest among all the

Table 3.6. Statistical comparison of marginal energy price by simulation with PJM market.

| cost function | parameter | PJM | RBTS89 | RTS79 | RTS96 | IL200 | SC500 | TX2000 |
|---|---|---|---|---|---|---|---|---|
| default | mean $/MWh | 49.2 | 0.7 | 17.31 | 24.4 | 18.51 | 37.8 | 158.42 |
| | SD | 51.9 | 0.07 | 7.7 | 13.7 | 2.62 | 43.3 | 6 |
| | GoF | n/a | -1084 | -6.5 | -2.87 | -21.8 | -0.13 | -18 |
| market-poly | mean $/MWh | n/a | 65.5 | 55.6 | 64.7 | 46.3 | 57.66 | 87.2 |
| | SD | n/a | 71.3 | 43.7 | 47.9 | 32.66 | 46.32 | 57.2 |
| | GoF | n/a | 0.36 | 0.36 | 0.36 | 0.15 | 0.44 | 0.16 |
| market-pwl | mean $/MWh | n/a | 66.6 | 49.76 | 48.3 | 45.85 | 53.5 | 81.5 |
| | SD | n/a | 71.2 | 43.77 | 40.3 | 35 | 38.1 | 61.4 |
| | GoF | n/a | 0.39 | 0.45 | 0.41 | 0.14 | 0.15 | 0.24 |

test cases with 0.45, closely followed by polynomial cost function based SC500 with 0.44. The highest GOF of 0.49 was obtained among all the simulations for RTS96 during 2015 with market-based polynomial cost functions. The IL200 test case did not perform as well as the other test cases as the scaling of this particular system was 10% lower than that of the system that is being compared. Similarly, the case with TX2000 which suffers from lower scaling produced higher rate when compared to other test cases. It is to be noted that these values of GOF are far from the best fit value, because the comparison is between OPF results from a test case that is physically different from the real market in terms of size and complexity. Secondly, the results presented in this chapter are marginal energy prices of test cases from OPF, which only resembles the economic dispatch problem of an electricity market, but not the unit commitment. Considering these two approximations, the positive GOF obtained are significant improvements when compared to the existing

default test case results.

Fig. 3.11 presents the weighted average marginal energy cost from OPF simulation on RTS96 (green) using the market-based polynomial generator cost function for the scaled PJM demand of 2014. The marginal energy cost has been compared to the marginal energy of PJM during the year 2014. The hours during winter (January-March) show a higher marginal energy cost similar to that of the market. The hours during the summer (June-August), where the system demand is higher compared to other seasons of the year, the price was relatively low. When the marginal energy cost time-series of the proposed technique is compared to the fuel-cost based approach in Fig. 3.1, the time-series in the proposed technique could represent most of the valleys and peaks similar to that of the real PJM market.

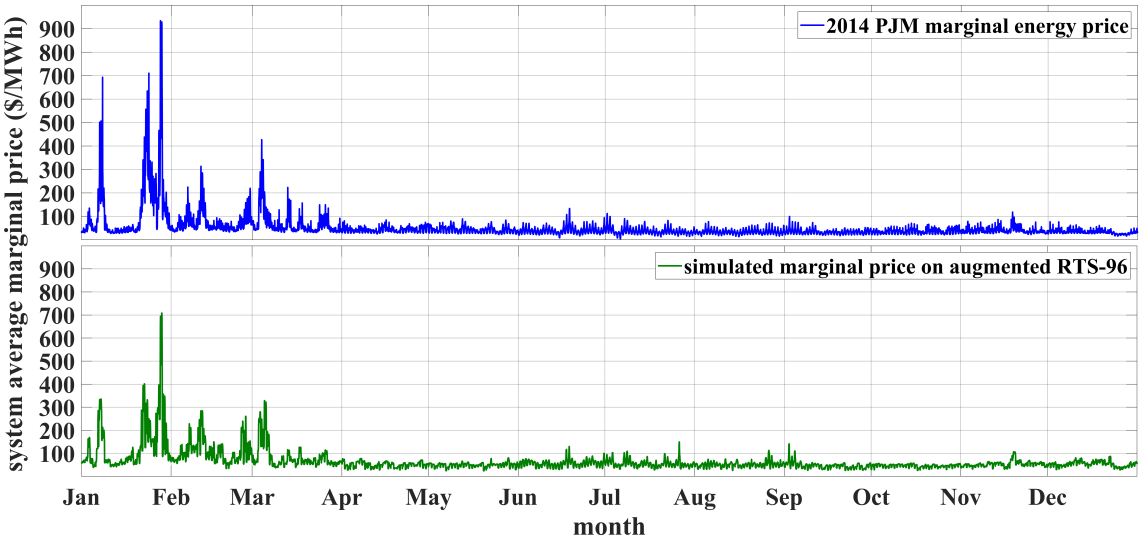


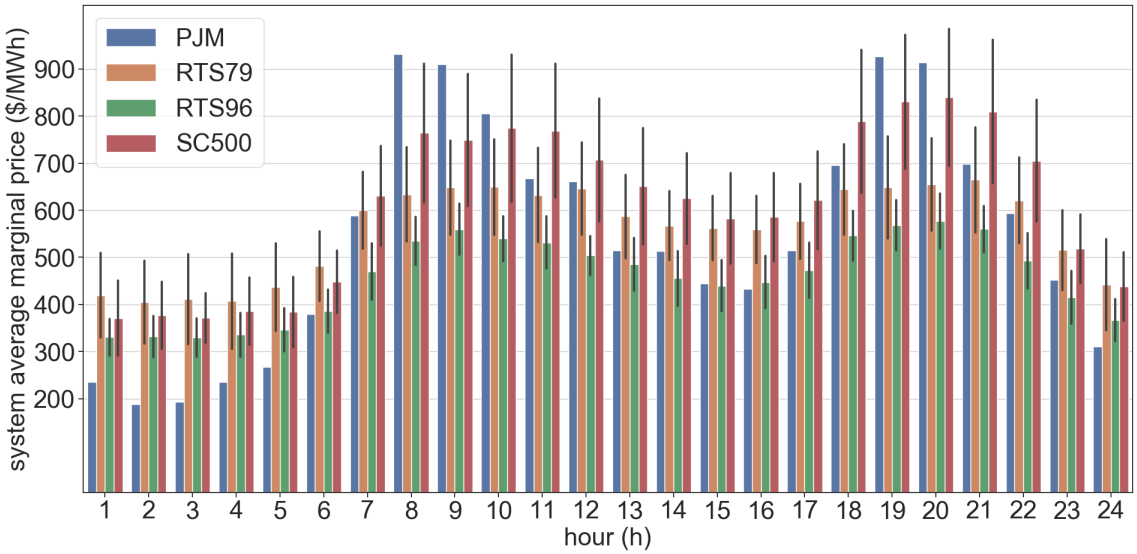

Figure 3.11. Annual marginal energy price of PJM for 2014 with one-hour resolution (top blue curve) and the simulated marginal energy price using the market offer-based polynomial cost functions on the RTS-96 test case with scaled PJM demand (bottom green curve).

### 3.9.3 Monte Carlo Simulation

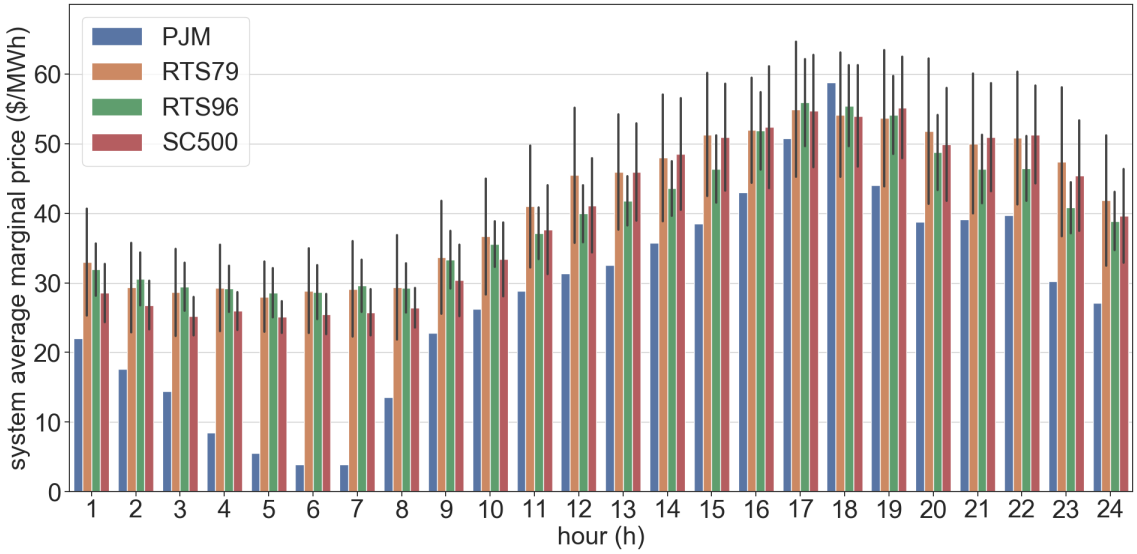
There is randomness in the selection process of cost functions from the PJM market cluster data to generators on the test cases, as the number of generators in the PJM system are much greater than the number of generators on any test case (i.e., 853 in PJM, versus 544 in TX2000, the largest test case available in this study). The proposed algorithm is a viable solution to create test case cost functions only if its solution is stable in the random selection of generators. Monte Carlo simulation was performed to show the consistency and stability of the proposed technique. OPF was performed for one day with a one-hour resolution for 100 Monte Carlo trials. Each trial, the selection of generators to the test case from the PJM market was re-sampled. Though the Monte Carlo simulation was performed on all test cases for both curve-fitting techniques, only results of RTS79, RTS96, and SC500 using market-based polynomial curves were presented for brevity, as they had the highest GOF. The other test cases showed similar stability.

Fig. 3.12 compares the stability of the three test cases versus the PJM market using bootstrap plots of the weighted average marginal cost. Each thin colored curve represents one Monte Carlo trial of the test case, with the dark curve representing the mean. Two days were chosen for the Monte Carlo simulation, as they are the two extreme priced days (the peak and the valley) of the year.

Fig. 3.12a presents the Monte Carlo simulation of the peak priced day of the year (Jan. 28, 2014). The SC500 test case performed the best among the three for this day, and could reproduce the peak hour of the day at 8:00. Both the peak and valley of that day in the simulation performed on the test cases appear at the same hour they occur on the PJM




(a)




(b)

Figure 3.12. Monte Carlo simulation of OPF with one-hour resolution of RTS79, RTS96, and SC500 for 100 trials. (a) is the simulation for the peak price day of the year in PJM Jan. 28, 2014, and (b) is for the least price day of the year June 18, 2014.

system for most of the trails. The valley on the simulations performed on RTS96 has a one hour offset at 15:00, instead of at 16:00 for the actual market valley. Fig. 3.12b represents the Monte Carlo simulation results for the lowest price day of the year, June 18, 2014. All three systems performed equally, but none could reproduce the valley effect of PJM, even on a single trial. The inability of the proposed technique to match the valley can be attributed to the approximation of the test case simulations explained previously in Section 3.7.3. All test cases could reproduce the peak of the day and the sudden change in slope between 20:00 and 22:00. Simulations are expected to reflect such variations, as services like demand response and battery charging strategies are most likely to respond to similar pricing signals.

## 3.10 Conclusions

In this chapter a general methodology for augmenting power system test cases with generator cost functions that represent real electricity market energy prices was presented. This technique is compared to the current state-of-the-art fuel-cost and heat-rate based generator cost functions. Eight test cases were tested using the two proposed techniques by comparing their OPF results with the real PJM market day-ahead marginal energy prices. The OPF was executed using a scaled annual demand curve of PJM with one-hour resolution and compared to the marginal energy cost of the PJM market using a goodness of fit metric. Simulation with our proposed technique on all the test cases resulted in better energy price estimates than the default cost functions of the system. A Monte Carlo simulation was presented to show the stability of the proposed techniques, even with the randomness in the algorithms. The proposed technique is

intended to be used to augment test cases based on ISO-based market which publish their bid/offer data. This technique should not be used for economic analysis of other market models and on non-deregulated power systems.

Even though all test cases produced energy price closer to the realistic costs than their results using the default cost functions, not all test cases performed the same when compared to the real market price. The SC500 and both the RTS test cases had their marginal energy price close to PJM marginal energy price. These test cases performed better when compared to the other test cases as the SC500 and both the RTS test cases as they had no scaling issues and had a wide range of generator sizes (i.e., range of generators from a large units in 100's of MW to small units in range of 10 MW) as they closely represent a real power system generation profile of a large area. Such test case achieved a goodness of fit as high as 0.5, and the test cases that had generation profile that only represent a part of a power system produced lower goodness of fit. The proposed technique performed better than the existing cost functions and also the state-of-the-art synthetic test cases that use generator heat models and fuel costs to develop the cost functions. These test cases would allow the research community to perform more accurate economical analysis of CPPS with more realistic energy prices.

To produce more accurate energy prices, the proposed augmented test cases can be used along with unit commitment (UC). The formulation of the UC problem does not change, but the OPF following the UC using the proposed market-based generator cost functions will result in a realistic energy prices when compared to the fuel cost-based generator cost functions. There are other statistical pattern recognition techniques that may result in different clusters and different distributions. In this chapter, I only used one

such unsupervised technique, as the main contribution of this work is to present the method of using real market data to augment test cases for economic studies. Exploring various pattern recognition technique for this problem is not in the scope of this chapter, but could be explored in future studies. As the proposed technique has resulted in realistic energy prices in OPF simulation than the default cost functions on all eight test cases, it is fair to conclude this technique can be applied to any other power system test case for economic studies. Economic studies performed on such augmented test cases would result in realistic energy costs, and conclusions drawn from such studies would better represent the behavior of new technologies on the real-world power system.

# CHAPTER 4 Data Driven Approach to Estimate Emissions from Market-Based Power System Test Cases

## 4.1 Introduction

There is a growing interest in power systems research to reduce the emissions as a part of their results in the recent times [16]–[18]. These works are a part of the smart-grids initiative to make the power system more sustainable [26]. There has been active research to reduce the harmful emissions from the electric power industry, one of such works presents a novel technique to optimize demand based on pollutant emissions [27]. There have been published research to estimate and reduce emissions by various optimization techniques performed on open source test cases but have used the generation profile on those test cases which cannot be validated with a real network [28]–[30].

To realistically evaluate the impact of new generation or demand optimizing techniques on emissions, the test cases used for simulations should represent a real power systems. In this work a technique to further augment the test cases from the Chapter 4.2 is proposed by adding fuel-turbine type and thermal curves to the test case generators to represent a real power system. There was no fuel-generator data added to the test cases because the objective of the work in the Chapter 3 was to develop test case cost functions based on a market offer data. In this chapter a technique is presented to further augment the test case with fuel-turbine data of a real region of a power system to the augmented test case. The major contributions of this work are:

1. a data driven technique to develop time-series fuel mix data for test cases to represent the generation profile of a real system,

2. an open source test cases that have 19 different combinations of fuel, and turbine data on practically any size test system,

3. a method to develop generator thermal efficiency data for test cases based on real generators thermal efficiency to accurately estimate GHG and AP emissions.

To validate the effectiveness of the proposed technique, hourly fuel mix is compared to the real world fuel mix as well as the GHG, and AP emissions from the simulations are compared to the GHG and AP emissions of a real electric interconnection.

The rest of the chapter is organized as follows: The following section presents a brief discussion on the state-of-art of the test cases and discusses the limitations of these test cases to perform accurate emission studies. The Section 4.3 provides a brief discussion on emissions estimation in power systems and the limitations from the existing test cases. The need for using an augmented test cases to represent a deregulated power system along with the proposed technique to augment existing test cases is presented in Section 4.4. The simulation setup and the data sources for developing the augmented test case in presented in Section 4.5. The results are analyzed and discussed Section 4.6 followed by the conclusions in Section 5.7 .

## 4.2 Emission Studies on State-of-art Test Case

To establish a reference state-of-art test case for this study, I modified a moderately large test case (RTS-96) to represent the PJM interconnection. Based on the Fig. 1.1 it is clear that the generation profile has changed by a large margin from 2010 to the 2018. To match the generation profile of PJM, the test case generators have been updated with the fuel types to reflect the installed capacity of 2016 as described in the Table. 4.1 [82]. The

percentage of the PJM capacity does not sum up to 100% because the renewable generation capacity was not presented. Since the renewable energy is non-dispatchable the percentage of renewable was directly subtracted from the hourly load. Rest of the generator characteristics such as the prime mover, fuel type, heat rates along with emission information were derived from the FERC RTO UC test system generator information.

Table 4.1. Modified generation capacity of the RTS-96 test case to represent the fuel mix of PJM interconnection in 2016.

| **energy source** | **2016 PJM summer ICAP (%)** | **test case capacity (%)** |
|---|---|---|
| **coal** | 36.6 | 37.1 |
| **gas** | 35.5 | 34.5 |
| **hydro** | 4.9 | 5.2 |
| **nuclear** | 18.2 | 19.2 |
| **oil** | 3.7 | 3.5 |
| **multi-fuel** | 0.4 | 0.5 |

To establish the performance of the test case, OPF was performed on the test case with scaled PJM demand of 2016 to determine the annual fuel mix by energy based on the test case. To simulate OPF at hour $t$, the load $d_j(t)$ on any bus $j$ is obtained using Eq. (4.1), where $d'_j$ is the default demand on bus $j$ provided in the test case, $D_{mkt}(t)$ is the total demand on the real PJM network, and $D^{max}_{mkt}$ is the annual peak load of PJM for the year in consideration (i.e., 2016 for this simulation). A scaling factor, $\psi$, is used to scale the default load such that the ratio of peak demand to installed generation capacity is similar to that of the real market. For this simulation, the scaling factor was set at $\psi = 0.96$ which would make the ratio of peak demand to the installed capacity 80%.

$$d_j(t) = \psi d'_j \frac{D_{mkt}(t)}{D^{max}_{mkt}}, \ \forall j = 1, ..., s \tag{4.1}$$

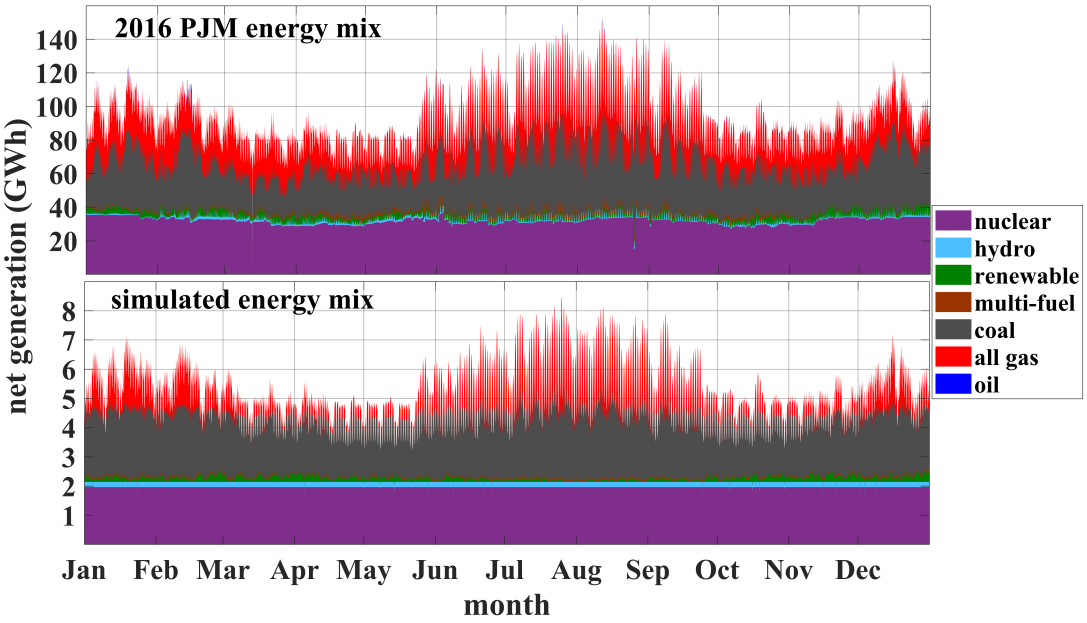



Figure 4.1. Stack plots comparing hourly generator output per fuel type for 2016 of PJM interconnection (top) with the state-of-art modified RTS-96 test case (bottom).

OPF was evaluated for every hour for the year 2016 and the dispatch of each fuel-type was observed for all the 8784 hours of the year 2016 and presented as stack plot in Fig. 4.1. Since the same cost function was used for all the days of the year, the dispatch of the generators did not change throughout the year. For example, when the nuclear generation between the PJM and simulation is compared, the output of the nuclear plant remains constant in the simulation as the cost function of a nuclear plant is the least. This mismatch in dispatch of generators can be observed in the annual energy mix pie charts in Fig. 4.2. The nuclear units in the test case resulted in a 100% capacity factor, when

compared to the real system it is close to 91%. It is clear that even though the installed capacity matches the real system, the dispatch does not behave like the real-system.

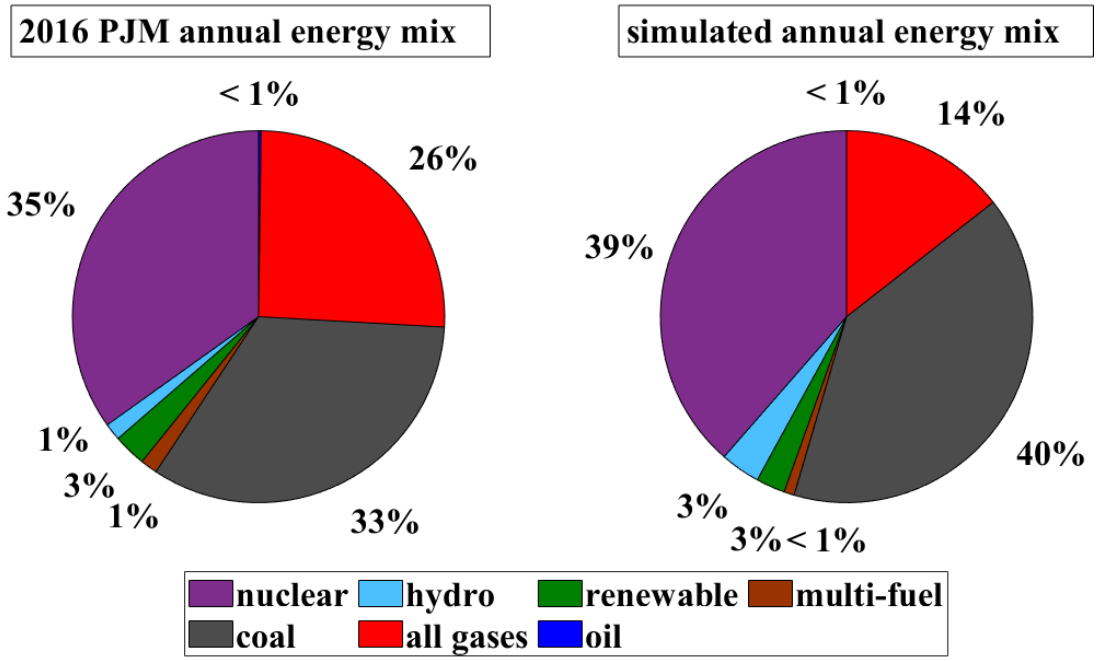



Figure 4.2. Comparing annual energy percentages per fuel type from PJM on the left pie chart to the simulated annual energy produced per fuel on the state-of-art modified RTS-96 test case with PJM 2016 ICAP on the right pie chart.

## 4.3 Emission Assessment in Power Systems

Majority of the generation in the U.S. power network are based on converting heat energy to electricity (thermal units) [1]. Heat is generated by burning fossil fuels, and this heat is converted into kinetic energy to drive the turbine-generator shaft. In the process of generating electricity during the combustion of fossil fuels, GHG along AP are emitted. Based on the thermal efficiency of the turbine-generator, the quantity of heat can be determined to generate an unit of electricity. To estimate the quantity of GHG and other AP emitted the following data is required:

1. electrical energy produced by each fuel-generator type,

2. efficiency *(heat curve)* of each fuel-generator type,

3. heat value of each fuel-type,

4. emission factor, and emission control factor of each fuel-type.

### 4.3.1 Generator Fuel Type

The most important data required to estimate the emissions from a system is the fuel quantity estimation. There are number of politico-economic and environmental reasons for the changes in energy-fuel mix. Old coal based generation are being replaced by more efficient combine-cycle gas fired units, and renewable generation adoption at an exponential rate to improved economic and environmental sustainability [83]. In PJM natural gas units are fast replacing coal units as both the base-load and the marginal generators. Combine-cycle natural gas power plants are quickly becoming the base-load generators of PJM, with its capacity factor rising from 50% in 2013 to 63% in 2016 during which the capacity factor of coal-based units reduced from 54% to 49% [84], [85]. Even though the modified RTS-96 test cases was updated with the latest fuel-type by installed capacity, it is clear from the discussion in Section 4.2 that dispatch simulation will not result in fuel mix of the actual fuel used over the time-series.

### 4.3.2 Generator Heat Curves

Generator heat curve data is the ratio of total thermal input to the useful electrical output which indicates the efficiency of a thermal power plant. The thermal efficiency of an unit depends on the operating point, usually the efficiency is the least at lower loading points as most of the heat is taken up in maintaining the minimum temperature of the

thermal system. Thermal power stations incorporate multiple heat recovery techniques such as combine-cycle installation, air-preheaters, and water-preheaters to improve the efficiency of the system. Along with the technologies used, the age of the power plants also influences the thermal efficiency of the system.

The Fig. 4.3 shows the weighted thermal efficiency of the major generators in PJM region. The weighted thermal efficiency $H$ of a generator $k$ is evaluated using the Eq. 4.2 where $Q_{kt}$ is the heat in one million British Thermal Units (MMBtu) required to generate $E_{kt}$ MWh of energy by generator $k$ for time step $t$. For preparing the Fig. 4.3 the thermal and energy data of the generators was derived from *EIA, Power Plants Operations Report* which publishes with a monthly time resolution for a year ($T$ = 12) [85]. Though the report has the filings from all the generators in the U.S., I used the data from the generators under the PJM interconnection. From the distribution of weighted heat curve data it can be concluded that not all generators of a fuel-generator type have same efficiency, and this distribution of efficiencies must be considered when developing test cases, so that the quantity of fuel is estimated accurately based on simulations.

$$H_k = \frac{\sum_{t=1}^{T} Q_{kt} E_{kt}}{\sum_{t=1}^{T} E_{kt}} \tag{4.2}$$

### 4.3.3 Emission Factor

Emission factors are the representative value that relate the physical quantity of GHG/AP emissions to the quantity of fuel. These factors are often provided as a ratio of lbs of emission produced when producing a MMBtu of heat. The emission factors are

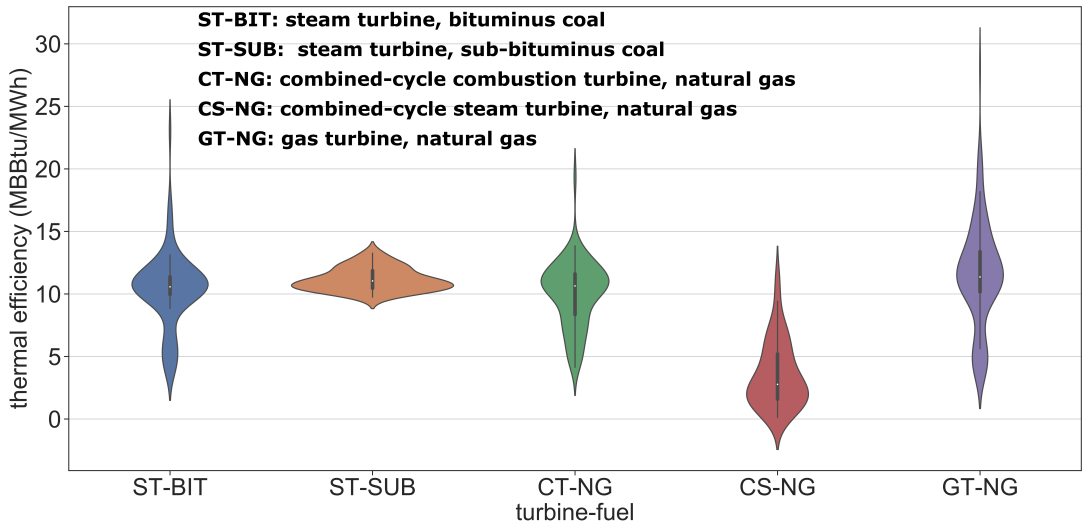



Figure 4.3. Distribution of the weighted thermal efficiency of the major fossil fuel-generator types operating in the PJM Interconnection during 2016.

used to build emissions inventories that can motivate the authorities and researchers to develop emission control strategies. In the U.S. the Environmental Protection Agency (EPA) is responsible to provide the emission factors for GHG and AP for all the fuel-types used across all the industries [86], [87]. To evaluate the emissions of a generator $k$, the total heat activity $Q_k$ is multiplied with the emission factor $\theta_k$ of the fuel used. There are multiple systems in place to reduce the emissions from a power plant, most of them are installed to reduce the AP. To evaluate the exact emissions released into atmosphere the missions have to be scaled according to the emission reduction efficiency ($\eta_k$) as shown in Eq. 4.3.

$$\mathscr{M}_k = Q_k \times \theta_k \times (1 - \frac{\eta_k}{100}) \tag{4.3}$$

The modified RTS-96 test system has all the required data to estimate emissions which was sampled from the FERC RTO test system, which was developed from the EIA, and EPA database of 2010. As described in Section 4.2 the RTS-96 was modified to represent the 2016 PJM installed capacity and the the fuel-generator data along with their corresponding emission factors were sampled to statistically represent the PJM system. Based on the OPF simulation results presented in the Section 4.2, $CO_2$ emissions was estimated and compared to the 2016 PJM system average emissions as shown in Fig. 4.4. The estimated emissions is much higher than the real emissions as the estimated fuel mix was inaccurate in representing the real system. The main reason for the higher emissions in Fig. 4.4 is because of the overestimated coal generation.

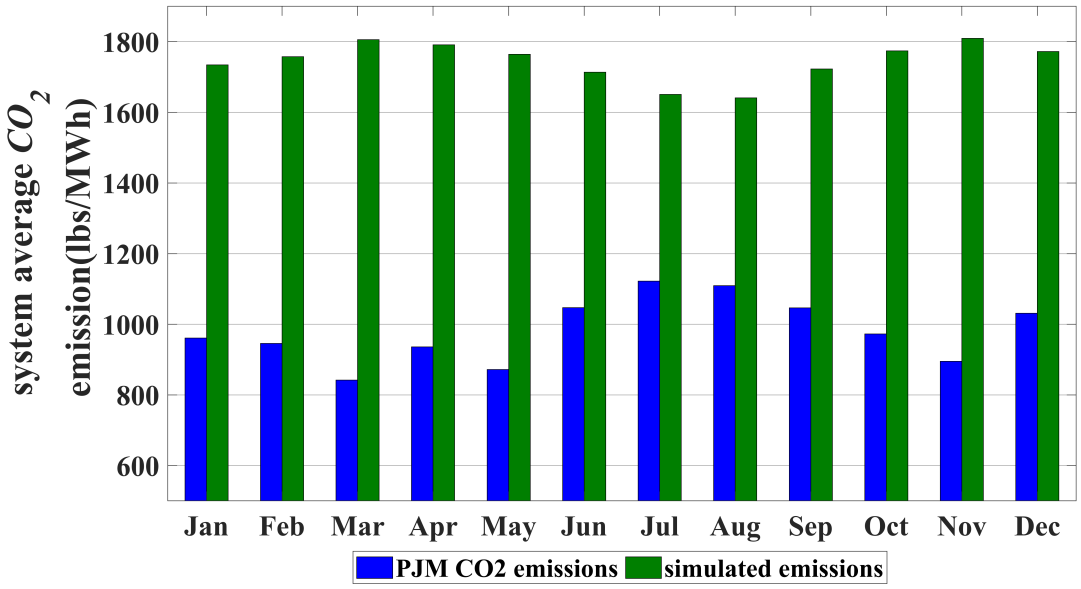



Figure 4.4. Comparing the system average $CO_2$ emissions from PJM to the estimated emissions from the modified RTS-96 test system in monthly resolution for 2016.

## 4.4 Augmented Test Case

In an electricity market the generators get to submit a fresh set of offers every-day, which can change the supply curve of the system. Daily generation fuel-mix profile not only depends on the changing load but also changes with the change in the offer cost. To replicate this dynamic characteristics of the generator supply curve in a simulation, the cost functions of the test case generator need to be dynamically updated with the latest offers of the real market-generators. At this point it is clear that developing a test case with a fixed cost functions cannot result in representing the ED of a deregulated power system for longer time-series.

### 4.4.1 Augmented Fuel-Generator Data

To represent the energy mix of the fuel sources of a real interconnection, the OPF solution of the test case is augmented with the fuel mix information from a real system. The hourly fuel mix data from electricity markets provide the energy produced from the major sources to meet the demand and exports. At the time of writing this dissertation, there are at least two markets (*PJM, and ISO-NE*) that publicly host this data [59], [88]. The data only contains the hourly energy, but does not specify any generator related parameters such as cost, size or location. Without a co-relating parameter, the hourly energy mix data cannot be attributed to a test case generator.

Capacity factor of the fuel-types in the real-system is used to co-relate with the capacity factors of the test case generators from OPF. Capacity factor is the ratio of total energy produced for a time period over the maximum energy that could be produced during that period. A generator with higher capacity factor is generally related to that of a

base-load unit and a generator with lower capacity factor is a peak load unit. For fossil fuel thermal units the capacity factor can be related to the generator cost (high capacity factor-low price, low capacity factor-expensive offer price).

The proposed Algorithm 4 statistically augments the test case generators with fuel-turbine type data by co-relating capacity factors. The data required to perform this augmentation are (a) the hourly fuel mix ($H_f^t$) (b) detailed annual fuel-turbine energy and capacity information from the EIA form: 923 [85] (c) OPF result of the test case $p_i^t$. The capacity factors $(\mathscr{K}_{f,g})$ and the share of each fuel-turbine type from each fuel type is evaluated using the Eq. 4.4, and Eq. 4.5 where $E_{f,g}^t$ is the energy produced by a fuel-turbine type at time $t$. The energy data is obtained from the EIA form:923 which is in monthly resolution in detailed for each fuel and turbine type [85] . The capacity factor is evaluated for annual ($T = 12$ months) duration and $P_{f,g}^{ICAP}$ is the installed capacity of a particular fuel-turbine type [1]. The evaluated capacity factors and their corresponding percentage share is arranged in descending order of the capacity factor.

The augmented cost functions of a test case are valid for a day, the same time resolution to assign the fuel-turbine type. There are multiple fuel-turbine sources with small percentage share which sometimes might not be possible to assign a full test case for each fuel-turbine type. For a given day $(d)$ the capacity factor $\kappa_i^d$ of a test case generator $i$ is evaluated for that day based on Eq. 4.6, where $p_i^{d,t}$ is the output of the generator at time $t$, and the $p_i^{max}$ is the peak-hour output of the generator. To evaluate the peak-capacity of each fuel-turbine type $b_{f,g}^{dmax}$ for the day, the energy share during the peak-hour $H_f^{dmax}$ of a fuel type $f$ is multiplied with the fuel-turbine share $\mathscr{S}_{f,g}$ of turbine $g$ and the total power for the peak hour as in Eq. 4.7.

**Algorithm 4** Algorithm to assign test case generators a fuel-turbine type based on the market-hourly fuel data and capacity factor.

**Input:** hourly market energy mix, hourly OPF data from market-based test case, EIA-923 data for the market region

evaluate the capacity factors $(\mathscr{K}_{f,g})$ for all fuel-turbine type
evaluate the share of each fuel-turbine type $(\mathscr{S}_{f,g})$ as a percentage of energy by the fuel ans sort them in descending order of $(\mathscr{K}_{f,g})$
arrange the energy share of each fuel-turbine type in descending order of its capacity factor
**for** $d = 1$ **to** $T$ **do**
    determine the dispatch of each generator during the peak hour of the day $(p_1^{dmax}, ..., p_n^{dmax})$
    evaluate the daily capacity factors $(\kappa_i^t)$ for all test case generators $(n)$
    arrange test case generators in descending order of their daily capacity factor $(\kappa_i^d)$
    initialize the index for test case generator $i = 1$
    **for** all $f, g$ combinations **do**
        evaluate the test case power by each fuel-turbine type at peak hour $(b_{f,g}^{dmax})$
        **while** $b_{f,g}^{dmax} > 0$ **do**
            **if** $p_i^{dmax} \leq b_{f,g}^{dmax}$ **then**
                subtract the generator $i$ capacity from the fuel-turbine type capacity $b_{f,g}^{dmax} = b_{f,g}^{dmax} - p_i^{dmax}$
                assign augmented generator capacity $a_{i,f,g}^{dmax} = P_i^{dmax}$
                increase the generator index $i = i + 1$
            **else**
                subtract the remaining fuel-turbine type $f$ capacity from the test case generator $i$ capacity $p_i^{dmax} = p_i^{dmax} - b_{f,g}^{dmax}$
                assign augmented generator capacity $a_{i,f,g}^{dmax} = b_{f,g}^{dmax}$
                set $b_f^{dmax} = 0$
            **end if**
        **end while**
    **end for**
**end for**

**Output:** augmented test case generator capacity for the day $(a_{n,F,G}^{dmax})$

$$\mathscr{K}_{f,g} = \frac{\sum_{t=1}^{T} E_{f,g}^{t}}{P_{f,g}^{ICAP} \times T} \tag{4.4}$$

$$\mathscr{S}_{f,g} = \frac{\sum_{t=1}^{T} E_{f,g}^{t}}{\sum_{t=1}^{T} E_{f}^{t}} \tag{4.5}$$

$$\kappa_{i}^{d} = \frac{\sum_{t=1}^{24} p_{i}^{d,t}}{p_{i}^{max} \times 24} \tag{4.6}$$

$$b_{f,g}^{dmax} = H_{f}^{dmax} \mathscr{S}_{f,g} \sum_{i=1}^{n} p_{i}^{dmax} \tag{4.7}$$

Test case generators are augmented until the entire capacity of each fuel-turbine type $b_{f,g}^{dmax}$ is assigned completely. The capacity of each fuel-turbine type on a test case generator $a_{i,f,g}$ is less than or equal to the capacity of the test case generator $i$. A test case generator can have anywhere from one to all the fuel-turbine types augmented over it. A representation of the augmented test case generator is presented in Fig. 4.5 in which all the fuel-turbine types are augmented over a test case generator $i$ with peak capacity for the day $p_{i}^{dmax}$. The rightmost fuel-turbine type $(F,G)$ on the cost function is the one with least capacity factor (peak-unit) and the augmented generator $a_{ifg1}$ is the one with highest capacity factor (base-unit).

In addition to the fuel-type information, the test case is augmented with the generator/turbine type and also the emission factors for GHG and AP to estimate the emissions.

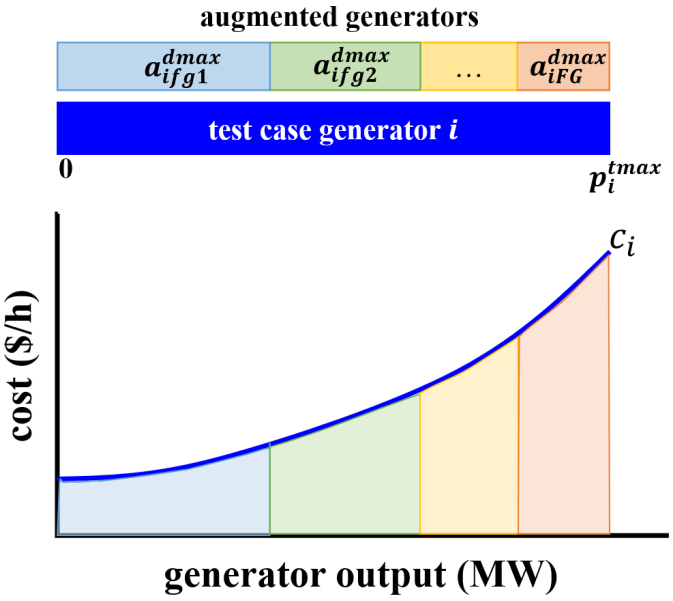



Figure 4.5. Representation of the augmented generator data on to a test case generator

### 4.4.2 Augmented Heat Curves

To estimate emissions from simulations, each of the augmented generator is assigned a heat curve. The thermal information such of each generator is provided in the EIA form-923 data [85]. The generators that participate in PJM are filtered by co-relating the company/utility names and location with the generators member information on PJM website [89]. The data contains monthly heat input (mmBtu) and the total electrical output (MWh) of each generator. This data is used to obtain the per-unit efficiency of each generator by dividing the heat input to the electrical output (mmBtu/MWh). A piecewise linear heat curve is fit to each generator based on the different operating points that a generator might have during an year.

There are multiple generator facilities in the real-system with different ages and efficiencies for each fuel-turbine type. The test cases have much fewer generators for a fuel-turbine type, to statistically sample the test case with heat curves, the probabilities of

the weighted heat was used for sampling. This thermal curve information constitutes the third augmented data layer on the test case. Along with the heat curves, emissions factors (EF) are also assigned to each augmented test case generator. EF represents the value of the quantity of a pollutant released to the atmosphere with an activity associated with the release of that pollutant [86], [87].

## 4.5 Simulation Setup

The proposed augmenting algorithm is implemented using the PJM interconnection fuel mix data for the year 2016 posted in hourly resolution [59]. Data from [82], [85] was used to develop the CF of the generators by fuel-type and turbine-type for PJM which is presented in the Table 4.2. The CF of each fuel-turbine type and each generator-type are not directly present in either of the data sources and had to be derived for this work. The real-system contains many more fuel-turbine types which have less than 0.1% of the total energy share, and were not included for this study because of their minuscule share.

The proposed technique is implemented on three test cases (a) the RTS-79, (b) RTS-96, and (c) the synthetic test case of South Carolina 500 bus system [15]. All these test cases along with the augmented fuel-turbine and heat curves data will be available on an open-source repository. The OPF was performed using the OPF solver of MATPOWER with the scaled demand of PJM as described in cost functions augmented test cases paper [90]. The time-series load curve used for the proposed technique is exactly the same load curve used in Section 4.2. Because of the nature of renewable sources being non-dispatchable, the renewable energy share is directly subtracted from the

load. The EF of each fuel-turbine type is provided in the Table 4.3.

Table 4.2. 2016 PJM generation by fuel source and unit type along with their respective capacity factors

| fuel type | % | actual fuel | turbine | $\mathscr{S}_{f,g}\%$ | capacity factor |
|---|---|---|---|---|---|
| nuclear | 34.4 | uranium | ST | 100 | 91.3 |
| coal | 33.4 | bitumi-nous | ST | 87.5 | 46.2 |
| | | sub-bituminous | ST | 10.5 | 44.6 |
| | | waste coal | ST | 2 | 42.8 |
| natural gas | 23.1 | natural gas | CC | 82 | 62 |
| | | | GT | 12.5 | 6.9 |
| | | | ST | 4.5 | 12.3 |
| | | | IC | 1 | 10 |
| hydro | 1.5 | hydro | run of | 60 | 31.4 |
| | | | pumped | 40 | 13.7 |
| multi-fuel | 1.5 | multiple | multiple | 100 | 50 |
| oil | 0.2 | distillate oil | IC,GT | 81 | 3.2 |
| | | residual oil | IC,GT | 18 | 0.7 |
| other gas | 2.7 | landfill gas | multiple | 43.5 | 50.3 |
| | | other bio | multiple | 20 | 64 |
| | | other gases | multiple | 36.5 | 8 |
| renewable | 2.8 | multiple | n.a | 100 | 27 |

## 4.6 Simulation and Results

The fuel mix and emissions from the proposed technique is compared to the PJM capacity-based generator assignment of the RTS-96 that was describes in Section 4.2. The Fig. 4.6 is the fuel mix comparison of PJM system with the simulated fuel mix in hourly resolution. The order at which the fuel-types are stacked is just one of the different orders of representing the fuel mix and has no resemblance with the order at which these

Table 4.3. Detailed classification of the fossil-fueled generators used under PJM and their corresponding emission factors

| fuel class | actual fuel | turbine | $CO_2$ EF (lb/mmBtu) | $SO_2$ EF (lb/mmBtu) | $NO_x$ EF (lb/mmBtu) |
|---|---|---|---|---|---|
| coal | bitumi-nous | ST | 210.6 | 1.62 | 0.5 |
| | refined coal | ST | 210.6 | 1.62 | 0.5 |
| | sub-bituminous | ST | 214.2 | 1.94 | 0.41 |
| | waste coal | ST | 250.6 | 1.94 | 0.41 |
| gas | natural gas | CC | 117 | 0.0006 | 0.137 |
| | | ST | 117 | 0.0006 | 0.137 |
| | | GT, IC | 117 | 0.0006 | 0.137 |
| other gas | blast furnace gas | ST | 604.7 | n.a | n.a |
| | landfill gas | IC | 117 | 0 | 0.2 |
| | other bio gas | IC | 117 | 0.0006 | 0.2 |
| | synthetic gas | CC | 130.1 | 0.0009 | 0.15 |
| | propane gas | ST | 135.5 | 0.0009 | 0.15 |
| multi-fuel | municipal-waste | ST | 200 | 0.025 | 0.49 |
| | waste wood | ST | 206.8 | 0.025 | 0.49 |
| | tire-waste | ST | 189.5 | 5.34 | 0.166 |
| oil | distillate-fuel | IC, GT | 163.1 | 1 | 0.3 |
| | residual fuel oil | IC, GT, ST | 163.1 | 1 | 0.3 |
| | kerosene | IC | 165.8 | 1 | 0.3 |

fuel-types are dispatched. When compared to the capacity-based fixed fuel type assignment in Fig. 4.1 the hourly fuel mix closely represents the actual fuel mix of the PJM interconnection. The proposed technique could represent the coal and gas mix variation across the year, which the capacity-based fuel mix could not achieve. It is clear that the proposed technique is capable of representing the PJM interconnection fuel mix in both fuel-type and the chronological usage of these fuels.

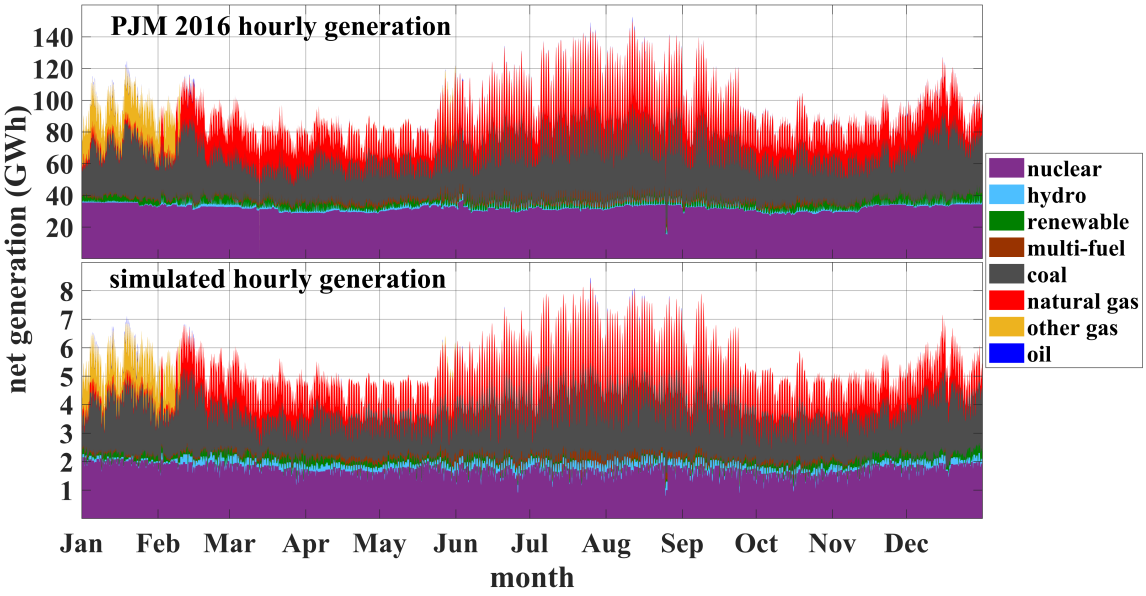



Figure 4.6. Comparing hourly generator output per fuel type of PJM interconnection with the proposed capacity-factor based dynamic fuel-type assessment on RTS-96 test case based on OPF with scaled demand of 2016 using stack plots.

To accurately evaluate the average emission from a system, the simulation has to produce similar generator dispatch as a real system. Pie charts comparing the annual energy produced by each fuel-type in PJM and by the proposed technique is presented in Fig 4.7. It can be observed that the proposed technique is capable of producing fuel mix which more accurately than the capacity-based simulation as seen in Fig. 4.2. As both the

simulations used the same load curve, the estimated energy from coal is much higher as the capacity-based fuel assignment test case used the cost functions of FERC RTO test case based on 2010 had much cheaper coal generators than the natural gas.

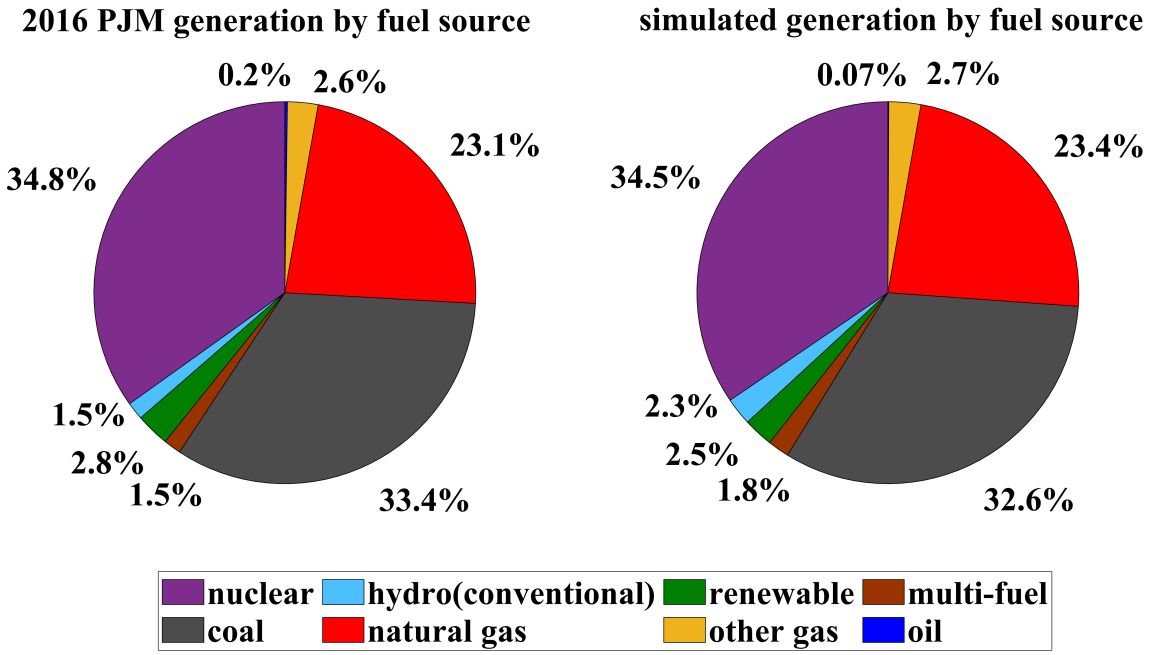



Figure 4.7. Pie chart comparing the annual energy produced per fuel source in the PJM innterconnection on the left and by simulation on RTS-96 on the right.

The system average $CO_2$ emission is evaluated as a ratio of the total emission over a time period in lbs to the the total energy produced during the same time period in MWh. PJM system publishes this system average emissions in monthly interval and is presented in comparison to the emissions produced from simulation using the proposed method in Fig. 4.8. The PJM system average emissions for $CO_2$ and air pollutants $SO_2$ ,and $NO_x$ are provided in the PJM environmental information services website [91]. As the proposed technique has some degree of randomness in selecting the heat curves the same results might not reproduced. The Fig. 4.8 is the best goodness of fit (93%) I achieved when I

simulated the proposed method 50 times.

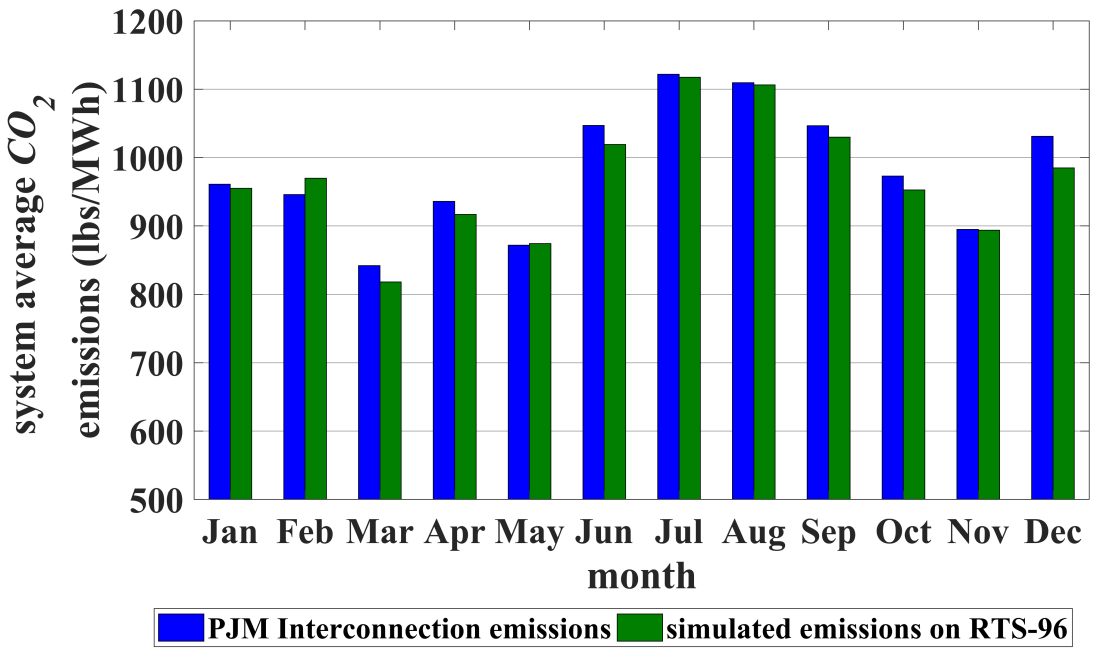



Figure 4.8. Bar graph representing the PJM monthly system-wide emissions per MWh of generation for 2016 and the bar graph of simulated emissions on RTS-96 test case.

The Fig. 4.9 presents the system average $CO_2$ emissions from the 50 trails performed on three three test cases mentioned in Section 4.5. The each bar except for the one representing PJM represents the mean of emissions from the 50 trails simulated on the corresponding test case. The thick black lines over the bars represent the standard deviation for the 50 trails. The January and February could not perform accurately as the majority of gas generation had been classified as other gasses which is unknown to us. There are many gaseous sources for generating electricity and the exact source could not be accurately traced. Nevertheless, I only require one set of heat curves for a test case, and that set can be saved for multiple other simulations when this test case is used.

As mentioned in Section 4.1 $SO_2$ and $NO_x$ are the main contributors for AP and

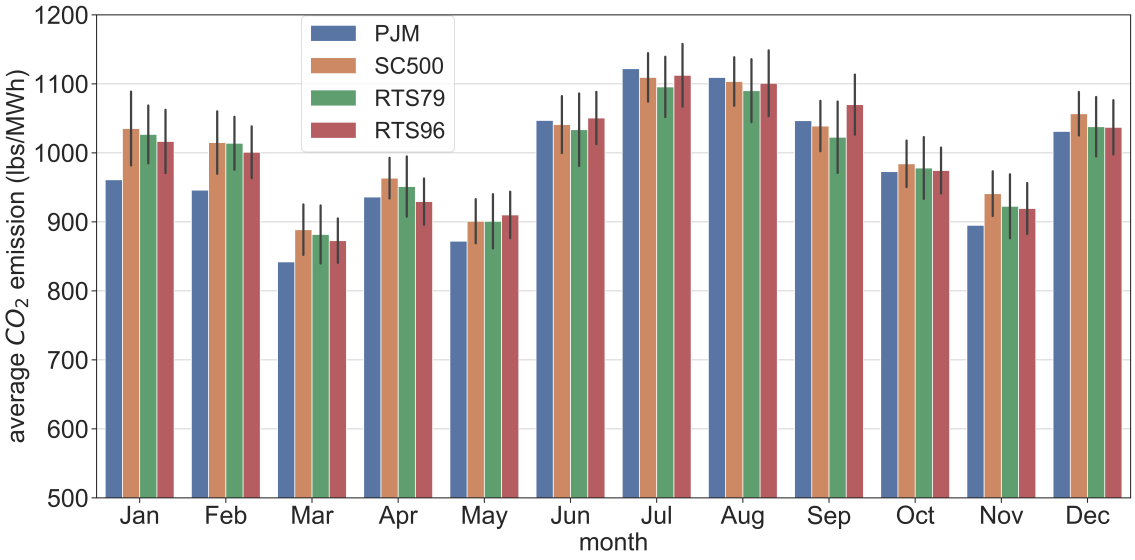



Figure 4.9. System average $CO_2$ emissions from three test cases (RTS-79, RTS-96, and the SC-500) based on 50 trails Monte Carlo simulation. The bars represent the mean of the trails and the black lines over the bars represent the standard deviation. The blue curve represents the system average emissions form the PJM interconnection.

acid rain. The proposed augmented test case are provided with emission factors along with the emission reduction efficiency for $SO_2$ and $NO_x$. A detailed description of the emission factors can be found in the Table 4.3. The system average AP activity based on simulation is compared to the PJM system average AP activity as shown in Fig. ]4.10. These emissions represent the controlled emissions of these pollutants. Each generator have to implement techniques to curb the air pollutants.

## 4.7 Conclusion

In this paper a novel approach based on real market data has been proposed to augment open-source test case generators to represent a real power system. The technique has been developed so that any power system researcher can augment a test case to perform accurate environmental emissions assessment for their research. The technique

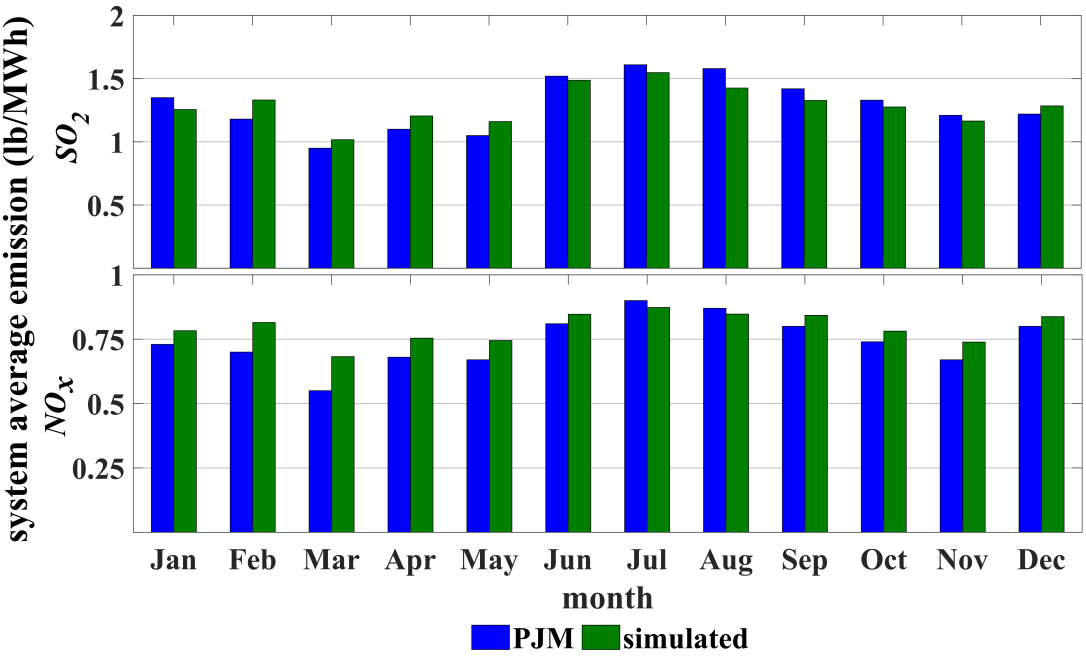



Figure 4.10. Comparison of simulated system average $SO_2$ and $NO_x$ emissions on RTS-96 in green to system average emissions form the PJM interconnection in blue for the year 2016.

presented in this paper requires test cases with augmented cost functions which are derived from a real electricity market. In combination with the cost functions and fuel-turbine data, these test cases can represent the dispatch of a real deregulated power market.

The proposed technique has been implemented on three test cases of different sizes, and the results from each test case has been compared to the real power system. The test cases showed fuel-mix and emissions similar to the real power system. These augmented test cases can be used for simulating transmission level dispatch based on a deregulated power system, and evaluate the economic and environmental impact of changing load and generation. The accuracy of the proposed technique can be improved if the fuel types labeled *other* can be determined accurately. The AP emissions can also be improved if accurate pollution control techniques are known.

# CHAPTER 5 Benefits of Aggregated Demand Response Participating in Bulk-Power Market

## 5.1 Introduction

The U.S. Federal Energy Regulatory Commission (FERC) issued Order No. 888 to deregulate the U.S. electric power system [92], which granted open access of transmission to all generators. The forethought of deregulation was to encourage investments to provide cheaper electric power generation by competing independent power producers. Under deregulation, increasing electric demand, combined with the physical constraints of the electric power network, can give unlimited market power to a few generators, resulting in relatively large locational marginal prices (LMP). These generators are typically fossil-fueled generators that have a fast start and ramp time.

In such cases of LMP spikes, demand response (DR) can be used to intentionally change normal power consumption in response to the electricity price (LMPs), or in exchange for financial incentives [7]. Technical and economic benefits of DR have been identified by law makers [7] and the research community [93]. Across the literature, the most common advantages of DR are:

1. provide financial benefits for DR participants (end-consumers) and electricity retailers (load serving entities (LSE)) [93]–[95];

2. improve reliability for independent system operators (ISOs)—also providing benefits from deferred infrastructure investment [95], [96];

3. increase market economic efficiency by reducing price fluctuations and

congestion [24], [97]; and

4. reduce green house gas emissions (GHG) [41], [98].

The literature broadly classifies DR programs into two categories: (a) price-based (PBDR), and (b) incentive-based (IBDR) [93]. Under PBDR, customers are exposed to time-varying rates to which they are expected to adapt their demand. PBDR is administered by the electricity retailers or LSEs, and the participants are mostly residential and small commercial organizations. Among the two classes of DR, PBDR has a smaller contribution ($\sim$7% in 2010) [99], but this number is increasing in the past decade as more utilities are offering dynamic pricing, and customers have gained access to advanced metering [100].

As per the latest report period, IBDR has greater enrollment than PBDR [100]. Under IBDR, customers are paid incentives for modifying their demand during requested time periods [93]. While both classes of DR have their own advantages and impediments, some of the challenges are as follows [101]:

1. price volatility is increased using real-time pricing (RTP) [102];
2. additional investment is required for advanced metering to communicate RTP to customers [103];
3. RTP requires customers' prompt response in consumption to reduce billing [9]; and
4. a new peak may be formed during off-peak hours, commonly known as the *rebound-effect* [104].

The main source of revenue in PBDR comes from the *energy market*, as the demand is modified in response to the price of retail electric energy [34]. FERC Order No. 719 was issued to allow non-generating resources to participate in the bulk-power market, which enabled entities to bid load reduction directly into the electricity market [105]. Bidding load reductions are only possible with IBDR, where the response is determined by the difference between the customer baseline (CBL) and the actual electricity consumption. With the capability to participate in the organized power market, IBDR can generate revenue from *capacity, energy,* and *ancillary markets* [99]. The ability to control loads in IBDR generates interest among the power research community, as the impact of load variation on the power system can be studied [33], [106]. There are a growing number of power system researchers that consider IBDR as an effective solution to increase economic efficiency [102], [107]. Increased sustainability in electric power systems is also commonly listed as a benefit of DR; in [41], the authors describe the economic and environmental sustainability of an IBDR in a smart grid. In this work, a pool-based IBDR model is used to reduce utility payments, improving the economic efficiency of the system.

As IBDR provides the opportunity to control loads, regulatory bodies (FERC Order No. 745) provided opportunity to participate in the market as a resource [36]. Some early work was conducted to study the impact of DR participating in the energy market [32], [108], where DR was modeled as a price responsive load to maximize social welfare. Because the demand is elastic to price, the load shifted to a lower price point to minimize system operation cost. The DR exchange (DRX) is a conceptual pool-based market to trade DR offers [38]. DRX provides the power system operator and other

market entities with additional flexibility [38]. There has been increasing interest in DRX-based DR because of the flexibility in modeling the objective function, and the ability to maximize the social welfare of the DR provider and other market entities. There has been considerable research in forming market clearing mechanisms for DRX with various objective functions [109]–[111]. In this work, the DRX is implemented to minimize electric utility payments. The DR providers are paid by utilities for the service, as the utility is benefited via reduced LMPs and lower payments. The DRX model in this work ensures that the rebound-effect does not significantly increase utility payments during off-peak hours.

Residential resource optimization to maximize profit of DR aggregators (DRA) participating in a bulk-power market through a DRX is presented in [35], but there was no model presented to integrate DRX into existing electricity markets. The authors in [112] propose a market clearing algorithm for DR offers, and illustrate the proposed technique on a test system. Though the authors provided a market clearing technique, there is no evidence of interaction with a bulk-power market, and market entity surplus is not included. In [110], a DRX is presented that operates in the day-ahead, intra-day, and balancing markets, where wind power plants participate in DRX to maximize their profit. The paper mainly quantifies the DR-market interaction to maximize the profit of the wind power plant, but the interaction of bulk-power entities and surplus is missing. A similarly themed research publication can be found in [111], where the DRX is used by virtual power plants to maximize their profits. The work presented in [113] describes DR bid/offer modeling based on CBL attributes. This is a significant work in this field, as customer willingness/behavior is important in determining the cost of DR. However, a

market model was not considered for the ISO–DRX interaction.

One aspect of DRX-related research that has not been well-studied is models for integrating DRX into the existing bulk-power market. A significant work related to DRX interaction with an existing day-ahead market model is proposed in [109]. The DR offers are modeled using customer willingness, and are cleared in the day-ahead market along with renewable energy sources (RES). The proposed model maximizes DR seller profit and market welfare based on a two-step market clearing process. The source of generator cost functions used in the test cases have not been mentioned, which is important for making analyses on simulation results based on restructured power systems, as generator cost functions should represent market offer prices rather than fuel cost-based functions [114].

The literature proves there is a growing interest in IBDR and DRX, but there is missing work that consolidates the advantages and challenges for integrating a DRX into an existing electricity market structure. In addition to DRX-market integration, the interaction of DR offers with the electricity market is investigated to study the main factors that influence the profitability of DR and DR-as-a-service. In this chapter, the DR offers are modeled such that they are a multi-period, incremental offer block structure. The DR offer not only contains curtailment, but additionally contains load shift information to mitigate rebounds. The updated DR offer structure and the multi-period market clearing is simulated on an augmented test case that statistically represents the prices of an actual power market. The scope of DR opportunities is investigated in existing bulk-power markets, and some of its benefits are presented in this chapter. Additionally this chapter provides details of how DRX can overcome some of the impediments DR

faces for power market adoption. The unique contributions of this project are:

1. an extended review of integrating DR-as-a-service to the bulk-power market, describing the benefits and challenges to bulk-power market entities;
2. an investigative study to determine the factors that influence the impact of DR on reducing utility payments in a day-ahead market; and
3. the design of a multi-period DR market clearing technique that considers the monetary impact of demand rebound.

This chapter is organized in a way that the reader can understand the structure, operation, and challenges of DRX. The electricity market structure with the day-ahead energy market operation and market sponsored DR programs are introduced in Section 5.2. In Section 5.3, an extended literature review of DR-as-service is presented, including the advantages and challenges to bulk-power market entities. The proposed model of a DRX integrated with a day-ahead market is described in Section 5.4, where the multi-period market clearing algorithm for the DRX is presented. The simulation setup is described in Section 5.5, including the augmented test case for representing actual power market prices. The results and discussion are presented in Section 5.6, with Section 5.7 concluding and providing possible areas of future work for DR-as-a-service.

## 5.2 Electricity Markets

Each entity in a fully deregulated market is responsible for the operation, maintenance, and expansion of its business. In most parts of the U.S., there are organized bulk-power markets to trade electricity through an ISO/regional transmission organization

(RTO). This section describes the operation of an organized ISO market, and the opportunities bulk-power markets created for DR.

### 5.2.1 Day-Ahead Market

In general, a market will have two groups: sellers and buyers. In an electricity market, the sellers are the generators, and buyers are the electric utilities/retailers (LSEs), with the ISO performing the role of market operator. In a day-ahead market, the ISO chooses the most economic generation available, without violating any physical limits of the power system. This optimized system supply is achieved by running unit commitment, economic dispatch, and optimal power flow (OPF). A day-ahead market begins with the generating companies offering their available generating capacities and price, and the retail electric companies bidding demand based on load forecast.

The *objective* of the ISO is a cost minimization problem using Equation (5.1). In its simplest form, OPF can be formulated using Equations (5.1)–(5.5), where, for a generator $i$, $C_i$ is the offer cost function, and $P_i^t$ is the power output at time $t$, bounded by the minimum ($P_i^{min}$) and maximum ($P_i^{max}$) generator output, given by Equation (5.4). The objective function, Equation (5.1), is subject to constraints Equations (5.3) and (5.5), where $D_j$ is the demand at location $j$ among $M$ load points on the network. At any given instant, the total generation must equal the total demand plus system transmission losses ($P_L$) in Equation (5.3), and line capacity limits must not be violated, given by

Equation (5.5):

$$\min_{P_i} \sum_{t=1}^{24} \sum_{i=1}^{N} C_i(P_i^t), \tag{5.1}$$

$$C_i(P_i^t) = \alpha_i (P_i^t)^2 + \beta_i P_i^t + \gamma_i \ , \tag{5.2}$$

subject to

$$\sum_{i=1}^{N} P_i^t = \sum_{j=1}^{M} D_j^t + P_L^t, \tag{5.3}$$

$$P_i^{min} \leq P_i \leq P_i^{max}, \forall i, \tag{5.4}$$

$$0 \leq P_{ij} \leq P_{ij}^{max}, \forall i, j. \tag{5.5}$$

The outcome of OPF is the optimal dispatch for each generator and the the cost of electricity at each bus (LMP), which is set by the offer price of the marginal generator—the generator that can meet the next megawatt of load. Each generator is paid the LMP of the bus they are injecting power. Similarly, every LSE pays the LMP for every unit of electricity purchased. The LMP can spike during peak load times, or during network congestion when an expensive local generator is committed. DR can reduce demand in these situations, offsetting the expensive marginal generator and hence reducing the LMP.

### 5.2.2 DR in Electricity Markets

Most ISO organized markets in the U.S. have DR programs to improve system reliability, which provides opportunities for non-generating entities to earn revenue from

the bulk-power market. The markets that DR can currently participate are capacity, energy, and regulation. Although capacity charges make up 11% of the bulk-power cost, the DR revenue earned from capacity markets is the majority among all markets. Let us consider the PJM market; during the year 2016, the total revenue from *economic load response* (a DR program organized in the *energy* market) was $3,550,535, whereas the *emergency DR* program (a DR program organized in the *capacity* market) had a total revenue of $648,997,257 [115].

Energy charges comprise more than half of the bulk-power cost, but DR participation in energy markets has been relatively low. The main reason for poor performance of DR in energy markets is due to the poor rate at which the resource is paid. In 2016, the average rate earned for participating in the energy market was 43 $/MWh, which is far less than the $69,000 per MW/year of capacity during emergency DR. These emergency DR programs have heavy penalties for non-performers, so only large entities with sophisticated direct load control operators participate in such markets (e.g., industrial arc furnaces).

FERC Order No. 745 was a landmark order for DR that opened the energy market for DR participating as a resource. In the energy market, DR providers earn revenue based on the LMP of the bus they are operating, like other generators in the day-ahead market. Figure 5.1 represents the cumulative generator offer curve of the PJM day-ahead market for 27 January 2014. Each blue circle represents a generator and its offer price, the size of each generator given by the gap between each circle in the positive direction of the $x$-axis. Without considering any line or voltage limits, the marginal energy cost will be determined by the intersection of total demand and this supply curve. A DR offer in the

day-ahead market should be offered at an intra-marginal price to be selected by the ISO. Analyzing the recent trends in annual system marginal price of PJM, which was 53.7 $/MWh, 38.4 $/MWh, and 31.8$/MWh for 2014–2016, respectively, DR participating as a resource in the energy market may not be economically sustainable.

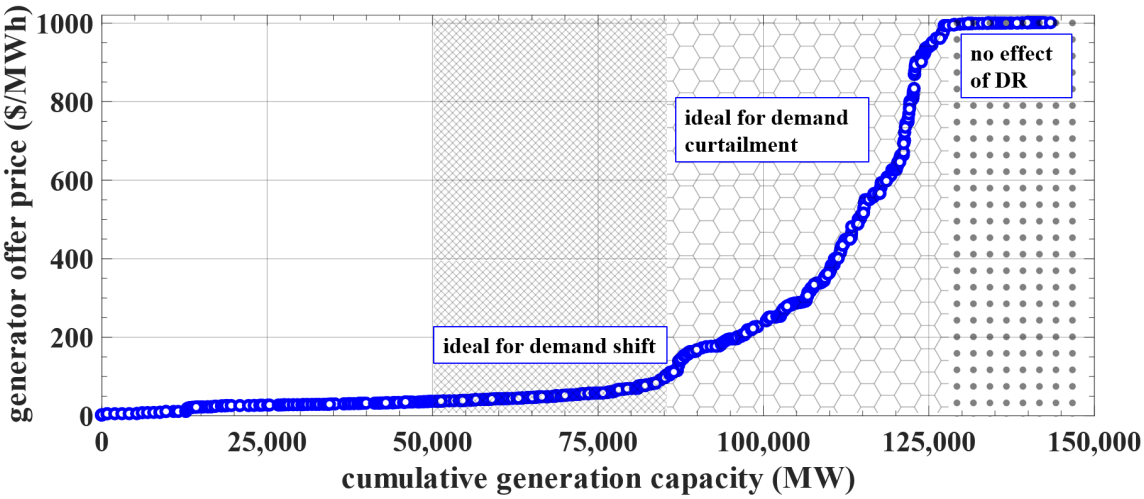



Figure 5.1. Cumulative generator supply curve of all participating in PJM day-ahead market for 27 January 2014. Each generator is represented as a blue dot, and the shaded patterns represent regions favorable for shift (grid), curtail (honeycomb), and no activity (dots) from left to right, respectively.

Instead, if the DR is offered as a *service* for the energy market, the cost of the service can be treated independently from the LMP, and the DR would be compensated based on the benefit it provides to the market. The benefiting entity pays for the DR service, where the most common entities benefited are electric utilities, as DR causes a reduction in LMP, which reduces utility purchases from the bulk-market. In Fig. 5.1, if the demand curve intersects the supply curve in the region shaded by the *dotted-pattern*, the marginal price will be high, but demand reduction would not cause a reduction in price; DR-as-a-service would not earn revenue when the system is operating in this region.

When the demand intersects the supply in the region shaded by the *honeycomb-pattern*, a small change in demand can greatly impact the marginal price, which is the ideal region for DR in an energy market. If the DR operation can have a controllable *rebound*, then it would be ideal to shift the demand to the region shaded by the *diamond-pattern*, as additional demand will not significantly increase the marginal price. The specific values of offer price and demand shown in Figure 5.1 should not be considered as the absolute delimiter for each region, as this will change daily in the market. Rather, the demarcation of the regions is dependent on the shape of the daily supply curve.

## 5.3 DR-as-a-Service

Regulatory bodies have opened opportunities for non-generating resources, such as DR, to participate in various market operations. DR is being used in ancillary services for regulation and reserves. This section presents the idea of using DR-as-a-service (as opposed to a resource) to improve economic benefits to other bulk-market entities. This section presents a review of the benefits for (i) bulk-power entities willing to pay for DR-as-a-service for their financial benefits, and (ii) DR service providers and end-customers willing to trade DR-as-a-service. As providing DR-as-a-service is still in its nascent state, this chapter has some of the state-of-the-art research in this field. Like any new service, DR will have challenges to be solved for integration into bulk-power markets. In this section, a review of such challenges as well as literature that addresses these challenges is presented.

### 5.3.1 DR Service Buyers

Generators: FERC Order No. 745 was challenged by Electric Power Supply Association in court, as they did not want non-generating resources to be able to set the market LMP. If used as a resource in the energy market, DR does not provide benefits to generators, and in fact competes against them. Additionally, as discussed in Section 5.2.2, DR as a resource might not be economically sustainable for the DR providers; DR-as-a-service can be engineered to benefit both the DR providers and the generators.

Reducing GHG emissions has been an important topic of research in power systems. In many states where renewable portfolio standards (RPS) are enforced, traditional generating entities maintain the RPS via a credit-trading mechanism [116]; DR can be made available as a service to reduce GHG emissions [41], [98], [117] and can be used to trade carbon-credits. The DR service can be structured such that DR providers sell both DR (as a curtailment service) and carbon credits, and LSEs invest in the curtailment service to reduce their payments, while generating entities procure carbon-credits.

There has been a change in the generation profile over the past decade, as significant investments have been made in RES. An ISO that operates under a fair market policy must dispatch the available resources economically and securely. In general, demand is inelastic to changing market LMPs under fixed rate tariff and time-of-use (ToU) plans. Under lower demand conditions, the ISO sometimes needs to curtail or sell additional RES to adjacent load areas at negative revenue [118], [119]. A consolidated demand side entity (e.g., DRA) that owns multiple resources, like distributed RES and energy storage systems (ESS), can optimize the available resources and modify the

demand curve to better utilize RES while maximizing their profit through the bulk-power market [120]. A DR-as-a-service model proposed in [109] utilizes a DRX to offer DR to renewable generation and ESS entities, where the objective of the DRX is to maximize the social welfare of DR sellers and DR buyers, thus improving RES utilization.

Utilities and REtailers: Utilities/retailers (LSEs) are the main entities that benefit from DR [93], [117], [121]. In 2018, most U.S. retail markets used the fixed tariff model. DR-as-service will be beneficial for utilities that use flat rate pricing because the price at which the utility buys the bulk-power is variable, but is sold to the end-user at a fixed rate. DR can bring large changes in the bulk-power price, which may allow utilities to sell at a higher profit (purchase less peak generation) [117], [121], invest in additional infrastructure, or reduce retail rates according to the public utility commission.

DR-as-a-service would also benefit those utilities that offer ToU pricing, which is fixed for certain periods of the day and season, as market prices are volatile and peak prices may occur at non-peak price ToU hours [102]. There are a few utilities that offer real-time pricing (RTP) to their customers (sell energy to the end-user at the same cost as the bulk-power market), but IBDR may not be impactful for such utilities, as the consumers pay the same energy price of the bulk-power market.

Significant charges for a utility are distribution and wheeling charges, especially if the entity is a retailer that does not own the physical infrastructure. Transmission charges are generally calculated based on the utility contribution to the coincidental peak of the system [117], [122]. Using DR-as-a-service to reduce demand during such peak hours will ensure lower transmission charges for the utilities. These savings on transmission/wheeling charges are beneficial for utilities with any type of billing (flat,

ToU, and RTP), as these are fixed charges based on the behavior of a few peak hours per year, irrespective of the energy price during that hour.

ISO/RTO: In a fair market, the ISO should not profit from the power transaction (independent entity). Market surplus—the additional revenue generated due to congestion—is returned to entities that own financial transmission rights (FTRs). System security is ensured through the use of ancillary services; DR is one of the cheapest services, as there is no additional ISO investment [37]. While other models of DR can be successful in reducing LMP spikes, DR-as-a-service can be used for other ancillary services to keep the system secure.

The transmission system operator (TSO) is an under-discussed entity in the power market, usually a monopoly and operated by the ISO. The ISO charges a fee for open access of transmission from every participating power market entity. To cater to the needs of the power producers and buyers, the TSO needs to invest in network expansion. These expensive investments can be temporarily deferred through DR, as extra transmission capacity is mostly needed during system peaks [24], [117].

Curtailment Service Providers: The benefits for the DR providers under various PBDR and IBDR programs have been presented across the literature [34], [107], [123]. In this section, additional benefits of DR *providers* are discussed that are gained if DR is treated as a service.

Demand REsponse Aggregators: DRAs are for-profit entities that have a set of customers willing to modify their demand for incentives. The DRA provides incentives to their customers by generating revenue from selling DR to bulk-power entities. When DR is offered as a resource, the DRA must ensure the DR offer price submitted to the

bulk-power market is competitive with other generator offers. The marginal generator offers have been reducing annually, as mentioned in Section 5.2.2, which makes it difficult for the DRA to operate in the energy market.

In the DR-as-a-service model, the DRA can submit offers to the market based on their demand portfolio, but the offer does not participate in the energy market clearing process. Instead, the DR is compensated based on the quality of *service* provided, so the DRA is capable of generating enough revenue to earn profit and pay incentives to participating customers [38], [112], [117]. Even though the DRA submits an offer price greater than the marginal energy price, the benefiting entity is able to pay for the service as the total benefit received is greater than the small quantity of DR required (which will be shown in detail in the simulation study). There are more potential buyers for DR when offered as a service, creating a more sustainable business model for the DRAs.

Consumers: Electricity customers that do not have advanced metering, or do not live in the right geographical location, cannot enroll in ToU or RTP-based billing models (PBDR) [103]. IBDR programs give DRAs direct control over a few electric loads to provide DR-as-a-service; the customer need not pay attention to the DR events. By informing the DRA their willingness towards the event and required incentive, the customer earns financial benefits from successful DR events. The customer can change their willingness and/or opt-out of individual DR events, unlike PBDR. Additionally, proactive customers can provide different willingness over the DR period, where they can increase DR activity during high DR price offers, and reduce when earnings are low [109]. RTP customers not participating in DR that reside in the same load area as the DR-as-a-service program will also receive benefits of the DR activity, resulting from

reduced energy prices (although they will not receive the additional aggregator incentives).

Large consumers (industrial, large commercial) pay additional charges for capacity and distribution based on their contribution to the coincidental peak of the utility. By reducing their contribution during these coincidental peaks, they can save heavily on distribution and capacity charges. DR-as-a-service can be used to reduce this coincidental peak contribution.

### 5.3.2 Challenges Integrating DR-as-a-Service

There are challenges integrating any new model/service into existing deregulated bulk-power markets. This section presents the state-of-the-art industrial and research practices, and the unmet challenges of integrating DR-as-a-service into the bulk-power market.

Compensation, Incentives, and Penalties: From the discussion in Section 5.3.1, it is inferred that the DR service providers can gain sustainable revenue by providing DR-as-a-service. Most literature discusses the total revenue gained by the DR service providers, and the total benefit to the power market entity [109], [110], [124]. The total benefit and cost of DR-as-a-service discussed in prior work considers a single entity on the system (single utility). Free-riders are entities that enjoy the benefits of a service without paying. The authors in [38], [112] show an estimate of the free-riders, and how much each entity benefits, but only one participant of each entity is considered. The problem is more complicated when there are multiple participants of each entity located at multiple nodes on the network, and compensation techniques for DR-as-a-service need to be designed based on multiple competing entities at different network locations.

Once the DRA receives market compensation, they must distribute incentives to the participating customers. Accurate evaluation of CBL is essential in determining incentives, and also impacts customer willingness. Liberal baseline estimation will result in losses to the DRA, while a conservative estimate punishes customers. Authors have used real data of both residential and non-residential loads from California to estimate CBL using regression models. In [125], the CBL was estimated using various time-series analysis, regression, and exponential moving average techniques for different customers based on their flexibility. The authors performed profit analysis based on each customer for each electricity market. A CBL estimation technique was discussed in [126], which attributes CBL changes for customers that already reduced their energy consumption using energy efficiency techniques.

The planning of DR happens day-ahead, but the transactions occur based on real-time market values. There are chances that a DRA offer is cleared, but fails to perform the DR. During such events, the ISO commits reserves, which can cause a spike in LMP. During such situations, the DRA and end-users that failed to perform may be penalized. During partial participation of DR, where few customer loads are curtailed with no market efficiency improvement, the DRA does not receive incentives as no entity has benefited. During such events, the DRA is obligated to pay the customers who performed DR, in-spite of paying penalties to the market, which incurs losses to the DRA.

DR Offer Structure: To develop realistic DR offers, the DRA must know the demand-elasticity of customers with respect to price. The elasticity depends on the type of load (e.g., residential appliances), climate (e.g.,heating and cooling loads), and type of customer [127]. The bid offer should reflect the customers elasticity towards the incentive

offered, which requires large surveys. One such work to determine the elasticity of residential consumers is presented in [128], but the responses are local to a particular region, and these responses may change over time. After obtaining such elasticity data, they must be translated to incentives based on the expected revenue to the DRA.

Residential customers comprise 40% of the total electric demand, and as such are a great potential for DR. To model a residential DR offer, the DRA must know the demand curve to the appliance level. Such information is sensitive, and is not widely available for the research community. The authors in [13] designed a technique to synthetically generate load profiles that can be used to develop DR offers. In real-world scenarios, there would be multiple DRAs competing to provide DR-as-a-service. In addition to incentives, the DRA must also consider competition in determining offer strategy.

DR Trigger: If DR is provided as a service, there must be a triggering mechanism to call for this service. DR currently participates in ancillary services, which are triggered by certain power system operational limits. If DR-as-a-service is used for economic purposes, a triggering mechanism must be provided. In one of my previous work [129][4], a statistical pattern recognition is implemented to analyze day-ahead market LMP, load, and climate data to trigger the DR service during market inefficiencies. This classifier only triggers during high LMP-based inefficiencies, where load curtail bids can be evaluated. An investigative study needs to be performed to evaluate the factors that influence the impact of DR on the electricity market, and proper DR-as-a-service triggers need to be designed.

[4]This work was presented in North American Power Symposium 2016 in Denver, CO

## 5.4 DRX in the Day-Ahead Market

In this chapter, DRX—a non-profit entity working as a pool-based market under the ISO is used to integrate DR as-a -service. As this entity is modeled as a market, it resembles the operation of the ISO. The sellers in this market are the DRAs that offer DR-as-a-service, and the buyers are those bulk-power market entities that benefit from the DR service (i.e., electric utilities in this work).

The DRX is integrated into the day-ahead energy market operation as shown in Fig. 5.2. The day-ahead market begins its operation by accepting generator offers and utility load forecasts. The DRX operates in parallel to the ISO, and is *triggered* when economic inefficiency is incipient, and DR-as-a-service is required. The DRX accepts day-ahead DR offers from DRAs, and clears those DR offers that benefit the bulk-power market. The bulk-power market is settled by the ISO using DRX-cleared DR offers, and the energy market LMPs are posted in the day-ahead. The benefiting entities from DR-as-a-service compensate the DRA offers at the marginal DR cost. Part of the DRA revenue is then passed on to participating customers as incentives (i.e., IBDR).

The DRAs in this study are for-profit entities that have a set of customers that are willing to curtail and shift their scheduled loads. A DRA prepares a cumulative DR supply curve by aggregating all the DR resources within their purview. The objective of the DRX in this study is utility payment cost minimization [114]. The OPF evaluated at the DRX will have the same formulation as in Equation (5.1), but will now be subject to a slightly different equality constraint, as shown in Equation (5.6), where $R_j^t$ is the DR at load bus $j$ at hour $t$, and $\lambda_j^t$ is the LMP:

$$\sum_{i=1}^{N} P_i^t = \sum_{j=1}^{M} (D_j^t - R_j^t) + P_L^t, \tag{5.6}$$

$$\min_{R_j} \sum_{t=1}^{24} \sum_{j=1}^{M} (\lambda_j^t (D_j^t - R_j^t) + c_j^t (R_j^t)), \tag{5.7}$$

$$\text{utility payment} = \sum_{t=1}^{24} \sum_{j=1}^{M} \lambda_j^t (D_j^t - R_j^t), \tag{5.8}$$

$$\text{service cost} = \sum_{t=1}^{24} \sum_{j=1}^{M} c_j^t (R_j^t). \tag{5.9}$$

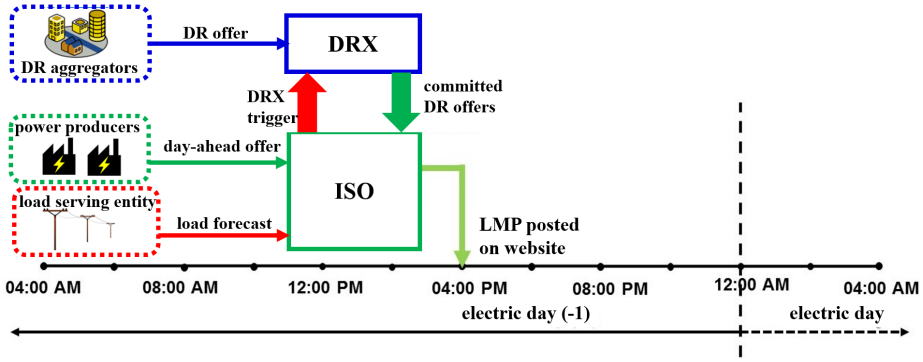



Figure 5.2. Day-ahead bulk-power market operation time-line with an integrated DRX.

The DR offer consists of five fields: (i) bus number (location on the network where the aggregated DR is connected); (ii) the hour of DR; (iii) the offer quantity blocks and (iv) corresponding offer prices; and (v) the shift window. DR offers are constructed by the DRA based on customer curtailment elasticity. The DRA arranges the DR blocks in incremental order of price, as shown in Figure 5.3. The size of each block may not necessarily be equal, as a DRA may have various classes of customers with different capacities. The maximum number of blocks per DR offer is capped at an integer $k$, similar to generator offers in PJM being capped at 10 blocks [130].

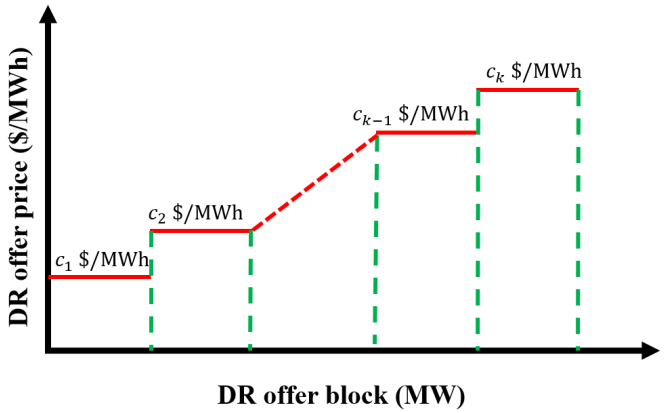



Figure 5.3. Representation of a DR offer block structure with $k$ offer segments, where each block $i = 1, \ldots, k$ is comprised of an amount (MW) and offer price ($/MWh).

Algorithm 5, which is used for selecting near-optimal DR offers, is based on the GENITOR version of the genetic algorithm (GA) [131]. The objective of this algorithm is to minimize the total utility payments for the ISO market, as described in Equation (5.7). The GA chromosome represents a solution to the optimization problem; for the DRX, the chromosome is shown in Figure 5.4. Each chromosome has an associated *fitness value*, which corresponds to the objective function of the problem (better solutions will have better fitness values). There are $p$ chromosomes that comprise the GA population.

| | | | |
|---|---|---|---|
| $u_1$ | $u_2$ | … | $u_n$ |
| $b_1$ | $b_2$ | … | $b_n$ |
| $s_1$ | $s_2$ | … | $s_n$ |

Figure 5.4. The structure of the GA chromosome with $n$ DR bids. Each gene (representing one DR offer) is comprised of (i) the DR offer selection (top blue row), where $u_i = 1$ indicates the offer is selected, (ii) the DR block selection (middle green row), where, if selected, the DR offer would use bid block $b_i$, and (iii) the shift hour (bottom red row), where, if selected, the DRA would shift the demand to this hour.

The chromosome is made up of $n$ genes (equal to the number of DR offers submitted to the DRX for that day), where each gene in the DRX GA represents one DR

offer (as described by Figure 5.3). Each gene is comprised of three parts: (i) the DR offer selection (shown as the top blue row); (ii) the DR block selection (middle green row); and (iii) the shift window (bottom red row). For a given gene $i$, if $u_i = 1$, the DR offer represented by that gene is selected by the DRX. If selected, the DR offer size and cost is provided by block $b_i = 1, \ldots, k_i$, where $k_i$ is the number of blocks in the DR offer. The demand of the selected DR offer block is shifted to hour $s_i = 0, \ldots, w_i - 1$, where $w_i$ is the number of hours in the shift window, and $s_i$ is the number of hours to shift from the beginning of the shift window (e.g., $s_i = 0$ moves the demand to the start of the shift window).

**Algorithm 5** Algorithm to select day-ahead DR offers using GA

**Input:** test case, DR offers, population size,

load all day-ahead DR offers
**for** $j = 1$ to $p$ **do**
  **for** $i = 1$ to $n$ **do**
    generate random Boolean value for $u_i$
    generate random integer between 1 and $k_i$ for $b_i$
    generate random integer between 0 and $w_i - 1$ for $s_i$
    add gene $i$ to chromosome $j$
  **end for**
  evaluate fitness of chromosome $j$
  insert chromosome $j$ into population based on fitness
**end for**
**repeat**
  select two chromosomes from the population using linear bias
  select two random indices and perform two-point crossover
  randomly mutate selected genes from child chromosomes
  evaluate fitness of child chromosomes
  insert child chromosomes into population based on fitness
  remove two chromosomes with worst fitness (highest value) from population
**until** stopping criteria
DRX select DR offers based on best chromosome (least fitness value)

**Output:** selected DR offers

The first step of the GA is to generate an initial population that describes the DRX selection of DR offers. For each gene in the chromosome, $u_i$ is randomly generated as a Boolean value (with equal probability), and $b_i$ and $s_i$ are uniformly distributed in the ranges $[1, k_i]$ and $[0, w_i - 1]$, respectively, $b_i, s_i \in \mathscr{Z}$. The initial population is comprised of $p$ such randomly generated chromosomes, each with $n$ DR offers (genes). The fitness of each chromosome in the population is calculated, and put in descending order (lower

fitness is better).

To evaluate the fitness as mentioned in the Steps 9 and 16 of Algorithm 5, all selected offers ($u_i = 1$) of a chromosome are applied to the system. Let us consider gene $i$ with $u_i = 1$, meaning the DR offer is selected. The block number for DR offer $i$ is obtained from $b_i$, and the DR quantity $R_j^t$ and cost $c_j^t$ can be determined. These DR and shift quantities are applied to the original demand forecast of the system at hour $t$ on load bus $j$. OPF is evaluated using Equation (5.1), subject to Equation (5.6), to obtain the new $\lambda_j^t$. Similarly, $R_j^t$, $c_j^t$, and $\lambda_j^t$ are obtained $\forall t$. This information is used to evaluate Equations (5.8) and (5.9), where the fitness of a chromosome is the sum of these two values. If more than one DR offer is cleared at the same bus and hour, the highest offer cost of the DR offers is used as the *marginal DR offer price*, and all accepted DR offers are paid the same marginal price irrespective of their offer block.

After the initial population is generated, the algorithm begins its search by iterating through a selection-crossover-mutation phase (Steps 13–18 in Algorithm 5) to create two new child chromosomes (solutions). In each iteration ("generation" of the GA), two chromosomes are randomly selected using linear bias [131], shown in Equation (5.10), where "random()" returns a sample of a uniform random variable in the range $[0,1)$. The function biases the choice of crossover to those chromosomes that are performing better according to a linear bias parameter, $l \in (1,2]$:

$$\text{index} = p \times \left( l - \frac{\sqrt{l^2 - 4(l-1)\cdot \text{random}()}}{2(l-1)} \right). \tag{5.10}$$

After selecting the two "parent" chromosomes, two points are uniformly randomly chosen between $[1, n]$. The chromosomes swap genes within these points (two-point crossover) to create two new "child" chromosomes. Each gene in the child chromosomes has probability $p_m$ to mutate. If gene $i$ mutates, the genes for $u_i, b_i$, and $s_i$ are randomly chosen using the same process as in the initial population.

The new chromosomes are evaluated for their fitness and inserted back into the population. The best $p$ solutions are kept at every generation (elitism). The GA iterates until the maximum number of generations is reached, or there is no significant change in the fitness value in a certain fixed number of iterations. The final DR offers are selected at Step 20 for each gene with $u_i = 1$ of the chromosome corresponding to the least fitness value (minimum utility payments). Both the new demand curve and the adjusted marginal energy prices are used to evaluate the LMP of each location, and are posted by the ISO as the final settlement of the market.

The GA was chosen as a proof-of-concept of the DRX clearing algorithm for its scalability in cost minimization problems [132], and has been shown to work well in many power system applications [31], [35], [133]. In our preliminary work, I also explored a greedy search algorithm [124]. There are multiple metaheuristic (e.g., particle swarm optimization) or classic optimization techniques for implementing cost minimization of the DRX; however, the choice of optimization is not the main contribution of this work, and does not impact the discussion of the results or the conclusions of the work.

## 5.5 Experimental Setup

### 5.5.1 Power System Test Case Setup

The augmented RTS-79 test case is used to evaluate the economic benefits of the proposed DRX market clearing technique. Two dates were chosen for the simulation study to illustrate the DRX market clearing technique. One date had a PJM DR event (23 January 2014), and the other (27 January 2014) did not have a DR event, but had a high peak LMP [64]. The non-DR event day was chosen to verify the expected advantages of using DR-as-a-service. The PJM demand curves for the specified dates were scaled to the test case using the same technique followed in Chapter 3. MATPOWER OPF was used for running the market simulations [63] to produce LMPs at every bus. The OPF was evaluated at the same frequency as the day-ahead and real-time electricity markets (i.e., one hour resolution).

### 5.5.2 DRX Offer Data

To set up the DR offers for the two days mentioned above, the size of DR with respect to the demand for that day was determined to represent the actual DR capacity of the PJM market. On 23 January 2014, as per the annual report of PJM [64], the total committed DR for the DR event was 4405 MW at hour 2:00 p.m. Thirteen of the twenty load zones of PJM participated in this DR event, which had a total demand forecast of 71,946 MW for the same hour. Thus, the percent of demand committed for DR was $\sim 6.1\%$ of the total demand forecast.

Even though the proposed market clearing technique is capable of selecting multiple DR offers from a single bus, in this case study, at most a single offer for each bus

at a given time was generated. The maximum curtailment per offer was restricted to 4–7% of the demand on that bus (similar to the PJM case), and only eight (half of the load buses) were allowed to have DR offers. Even though the number of blocks per DR offer can vary, for this simulation, the number of blocks per offer was fixed at $k_i = 5, \forall i = 1, \ldots, n$. The five offer blocks were considered to be equally sized (i.e., each offer block was $1/5$ of 4-7% of $d_j^t$, where $d_j^t$ is the demand of the bus $j$ at time $t$ for the test case with PJM scaled demand).

The literature is scarce for designing optimal incentive prices for consumers. Determining the optimal incentive is a sociology research topic, which is not in the scope of this paper. To determine the price of the DR offer supply curves, the DR offers are developed based on the range of retail electricity prices of the PJM region, with the offer prices for each block randomly selected in the range of 50–300 $/MWh (in ascending order).

The peak hours of PJM (8:00 a.m. to 11:00 p.m.) were chosen to be the DR curtailment hours. To mitigate rebound for this simulation, DR was allowed to be shifted to any other time during the day. This DR shift flexibility was chosen for this case study to analyze the capability of the algorithm to shift demand to a time that does not significantly increase utility payments (i.e., LMP). Because the demand is capable of being shifted to any hour of the day (except the curtailment hour), the shift window size for every offer is set to $s_i = 23, \forall i = 1, \ldots, n$ between 1:00 a.m. to 12:00 a.m. With the above conditions, $n = 128$ DR offers were generated for each day. These numbers are chosen for this system, but the DRX clearing method described above works for DR offers generated from any source (e.g., real customer data, state-of-the-art literature DR techniques).

### 5.5.3 GA Parameters

The GA for the case study used a population size $p = 250$. Each chromosome contained $n = 128$ genes (i.e., the number of DR offers). At each generation, the index for the parent chromosomes are chosen using the linear bias function in Equation (5.10) using a bias parameter $l = 1.5$, which means that the best solution according to its fitness has a 50% greater chance of being selected than the median solution. For each child chromosome, each gene has a probability of mutation $p_m = 0.1$. The GA continues until stopping criteria of 10,000 generations is reached, or 300 consecutive generations without improvement in the best fitness value.

## 5.6 Simulation Results and Discussion

To determine the impact of a DRX minimizing utility payments in a day-ahead market, we analyzed the weighted-load average system marginal energy price, and the utility payments before and after DR. Although our proposed DRX model was described with a trigger in Section 5.3, no technique was incorporated any technique for triggering the DRX in this chapter, two dates were pre-selected as the candidates for DR-as-a-service. The simulations were conducted in MATLAB R2017a on an Intel Core i7 3.6 GHz computer with 16 GB of RAM. The GA took 42:06 and 53:11 minutes to converge for 23 January and 27 January, respectively. Out of 128 offers, 66 were selected and scheduled for 23 January, and 71 were selected for 27 January. A detailed discussion of the impact of these DR offers on the bulk-power market are presented in this section.

As discussed in Section 5.5.1, the hourly demand on the augmented test case was derived by scaling down the PJM system hourly demand. This demand serves as the

baseline demand for the two days, and is represented by the solid black line in Figure 5.5. The solid blue line with the right *y*-axis represents the marginal DR cost for each DR hour. Based on the DR offers selected by DRX for the two days, the demand curve is modified for demand curtailment as shown by the red dashed curve, and the demand rebound (shift) is represented by the green dotted curve. The DR offers considered in this study (described in Section 5.5.2) operate only between 8:00 a.m. to 11:00 p.m., which is reflected in Figure 5.5 where the red dashed line only varies between those hours. All demand plots in Figure 5.5 are the aggregated demand of the load buses on the test case. Because the shift hours considered in those offers were flexible throughout the day, the demand rebound is spread across the day and mostly concentrated during off-peak hours.

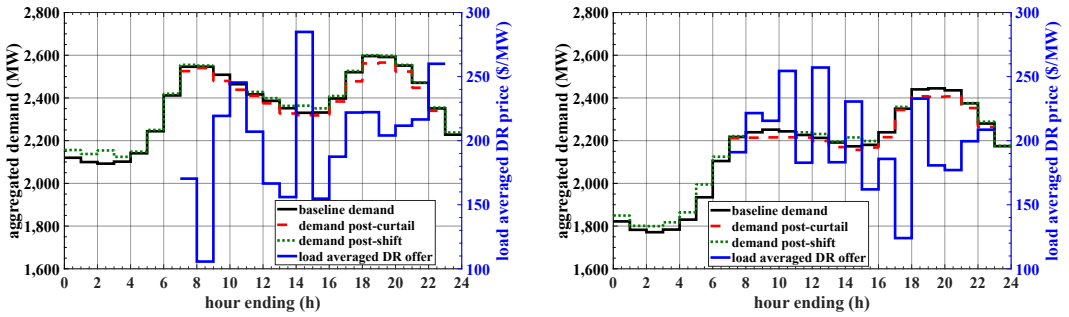



Figure 5.5. The cumulative baseline load for each hour across the network (solid black line), compared to the demand post-curtailment (red dotted line) and demand post-shift (green dotted line) for the day of (**a**) 23 January 2014, and (**b**) 27 January 2014. The solid blue line represents the weighted-load average DR cost for each hour of DR.

For a given time period, DR is offered on multiple load buses. Based on the cleared offers, curtailment and rebound can occur at the same hour, but at different load buses. This can be observed in Figure 5.5, where in most hours curtailment and shift happen simultaneously. The magnitude of curtailment is relatively high during the

peak-hours (6:00 p.m.–9:00 p.m.), and shift is relatively high during the mid-day demand *valley* (1:00 p.m.–3:00 p.m.). The objective of the DRX is to minimize utility payments, which occurs when curtailment offers are selected during those hours when the demand intersects the supply curve at a steep slope, and shifts the demand to those hours when the demand intersects the flat supply curve (as shown in Figure 5.1). In this case study, the peak DR activity occurred on 23 January at hour 6:00 p.m. During this DR period, 42.8 MW was curtailed out of the aggregated system demand of 2520.5 MW, i.e., 1.7% curtailment. For the 27 January simulation, the peak curtailment of 1.6% was deployed during hour 8:00 p.m. The number of generators in a test case is much lower than in the real network. The IEEE 24-bus RTS consists of 32 generators, with modified cost functions as described in the Section 5.5.1. Cumulative generator supply curves similar to Figure 5.1 were developed for the augmented test case for the two simulation dates, as shown in Figure 5.6. Even though the test case has significantly fewer generators than the real system, the supply curve in this simulation is statistically similar to the real PJM system curve of the same day in terms of price band and shape. The blue-dots represent each generator on the test case with the corresponding peak offer price. The daily system load is between the dashed-green and dashed-red lines, representing the daily minimum and peak demand, respectively. A steeper peak and flatter base supply curve represents ideal conditions for DR, as small changes in demand can result in significant decrease of LMP and utility payments. The supply curve of 27 January, shown in Figure 5.6b, is steeper when compared to 23 January, shown in Figure 5.6a, in the range of demand for the respective days. Thus, between the two days, 27 January is expected to perform better for DR.

For brevity, only the 27 January results are presented in detail in Figure 5.5b, where the loads are always curtailed when the CBL is above 2200 MW. By observing the corresponding supply curve in Figure 5.6b, it can be confirmed that the curve is steeper starting at 2250 MW. The supply curve has a flat offer price between the minimum demand (1770 MW) and 2050 MW, and, as a result, most of the loads are shifted to hours when the demand falls below 2000 MW. These observations show that DR-as-a-service is ideal for curtailing when the supply curve is steep, and shifting to time periods when the supply curve is relatively flat.

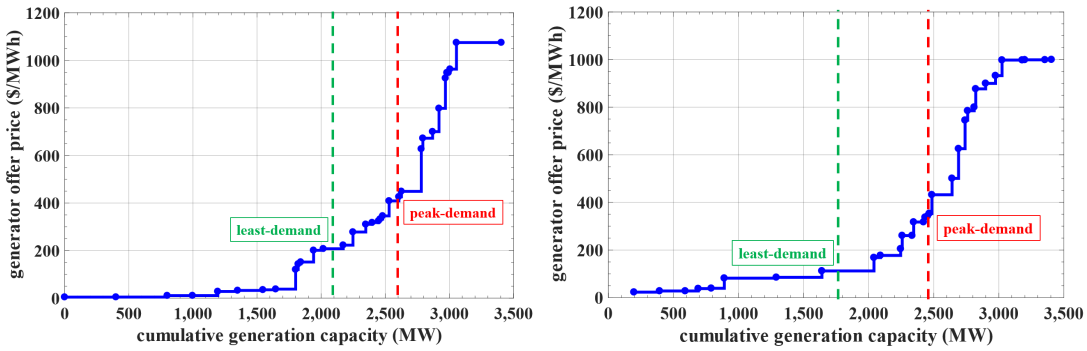



Figure 5.6. Generator supply curve for the test case (solid blue line), with each generator represented by a dot for the day of (**a**) 23 January 2014 and (**b**) 27 January 2014. The dashed green line represents the minimum daily demand, and the dashed red line represents the peak daily demand for each respective day.

In Figure 5.7, the solid blue curve is the PJM system marginal price, and the red dashed line is the simulated marginal price based on the scaled demand of PJM. An exact reproduction of PJM prices is not possible, as the test case network structure and number of generation units is significantly different than the real system. As discussed in Section 5.5.1, the augmented test case of the IEEE 24-bus system to represent the PJM network marginal prices. With the available data resources, the augmented test case better

emulates the real PJM marginal prices than the default fuel-based cost curves (shown in detail in [54]), observed for the marginal prices in Figure 5.7b during the evening peak hours that match that of the real system.

In the test case, the marginal price at the peak demand (∼450 $/MWh at 2600 MW) from the supply curve in Figure 5.6a matches the maximum marginal price (∼450 $/MWh at 7:00 p.m.) in the pre-DR case in Figure 5.7a for 23 January. However, in the 27 January case, the pre-DR marginal price at the peak hour was ∼450 $/MWh at 8:00 p.m. in Figure 5.7b, where the marginal offer at this peak demand was lower at ∼350 $/MWh for 2450 MW. This is evidence of network congestion, indicating the cheaper generator (∼350 $/MWh) could not be dispatched due to network constraints, and instead the next available generator offer of ∼450 $/MWh was cleared. This also explains the large decrease in LMP post-DR between hours 6:00 p.m. and 9:00 p.m. for 27 January.

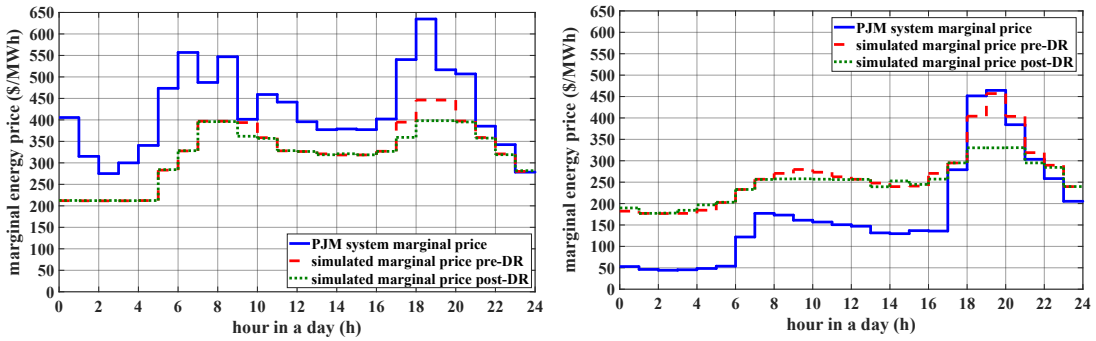



Figure 5.7. Actual marginal energy price for the PJM interconnection (solid blue line) compared to the pre-DR marginal price (dashed red line) and the post-DR price (dotted green line) for (**a**) 23 January 2014, and (**b**) 27 January 2014.

Even though the entire curtailed demand (328 MW on 23 January and 340 MW on 27 January) was shifted to another hour, there is no significant increase in the marginal

price during any off-peak hour. This holds true because the algorithm was designed to select only those DR offers that will reduce utility payments during curtailment, while simultaneously not significantly increasing the payments during rebound hours. Because the DRX allows DR offers as a service rather than resource, the system marginal price is only set by the generators that are part of the supply curve. The DR marginal price, shown in Figure 5.7, only serves as the price for the DR service, but does not set the marginal price of energy in the bulk-power market.

The objective function of the DRX in this study was to minimize utility payments, which depends on the aggregated demand and LMP of the load bus. From the discussion above, it was found that the LMP is reduced during peak-hours, and does not significantly increase during off-peak hours. This directly reflects utility payment reductions, shown in Figure 5.8. The red dashed lines represent the hourly payments before DR, and the green dotted line represents the payments after DR. It can be observed in Figure 5.8b that between hours 6:00 p.m. and 9:00 p.m., the utility payments reduced to the same point. This is likely because pre-DR, the marginal generator was creating an LMP for a small quantity of demand (e.g., the $\sim$450 \$/MWh generator discussed above). Just like in the marginal energy plot, there is no significant increase in utility payments after the DR and demand shift.

Table 5.1 presents a comparison of the utility payments and generator revenue before and after DR for both days. The surplus described in this table is the *market surplus*, which is the difference between the utility payments and generator revenue. Market surplus is an indication of congestion in the network. DR-as-a-service resulted in higher benefits in terms of utility payment reduction on 27 January than 23 January

because of the difference in generator supply curves of these two days (discussed earlier in this section). The peak demand on 27 January intersects on a much steeper region of the supply curve, where the small changes in demand resulted in larger reduction of LMP when compared to the supply curve of 23 January. The largest factor in utility payment reduction is the decrease in LMP, which is highly dependent on the generator supply curve and network congestion. The utility saved 6.2% in payments by spending only 0.5% of the initial payment for 27 January, and 2.1% savings by spending 0.4% of the initial payment for 23 January. The results for 23 January show savings in utility payments, but the market surplus of the system has increased when compared to the pre-DR case.

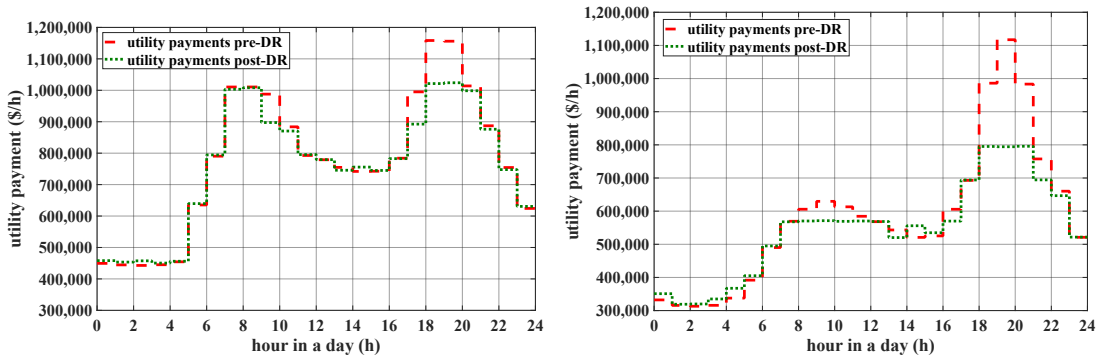



Figure 5.8. Comparison of utility payments pre-DR (red dashed line), and post-DR (green dotted line) for (**a**) 23 January 2014, and (**b**) 27 January 2014.

Table 5.1. Comparison of payments and revenue for pre-DR and post-DR conditions (in Million $).

| Date | Case | Utility Payments | Generator Revenue | Surplus | DR Operation Cost | Payment Benefit |
|---|---|---|---|---|---|---|
| 23 January | pre-DR | 18.73 | 18.33 | 0.40 | N.A | N.A |
| | post-DR | 18.3 | 17.9 | 0.43 | 0.068 | 0.39 |
| 27 January | pre-DR | 13.98 | 13.05 | 0.92 | N.A | N.A |
| | post-DR | 13.12 | 12.27 | 0.84 | 0.069 | 0.86 |

## 5.7 Conclusions

An extended review of DR in electricity markets in terms of advantages, challenges, and opportunities was presented. In this work, a market model is designed for the DRX, to integrate DR-as-a-service into existing energy markets. The DRX model minimizes utility payments using DR services, while simultaneously providing an opportunity for DR service providers to offer the incremental cost of the service. A multi-period market clearing algorithm to select DR offers was proposed, where DR offers have both curtailment and shift information. The multi-period nature of the algorithm ensures that the selected DR offers do not adversely affect utility payments when the demand rebounds post-DR. The algorithm was implemented using a GA on an augmented IEEE 24-bus RTS test case that statistically represents the PJM day-ahead energy market. The results of the case study of the proposed DRX market clearing method determined important factors that influence the ability of DRX to reduce utility payments.

From the two simulations on the augmented IEEE RTS-79 test case, significant reduction in utility payments was obtained for both days. The utility payments reduced by 6.2% for 27 January, creating a benefit of $864,199 by spending $69,955 on the DR service, and $393,066 of savings by spending a similar amount ($68,092) for the DR service for 27 January. The peak DR deployment for 23 and 27 January was 1.7% and 1.6% of the aggregated demand of the system, respectively. By comparing the results of the two days, it is clear that the benefit (utility payments) from DR-as-a-service depends heavily on the generator supply curve, demand, and the state of congestion in the network. It is important that the demand during the load curtailment meets the supply curve at a steep slope region, where small changes in demand can result in significant savings for utilities and customers. This factor is why 27 January shows greater savings when compared to 23 January, even though the quantity of DR curtailed is similar. Additionally, the demand rebound (shift) should be planned such that it is scheduled during hours where the supply curve has a relatively flat slope so the price during that time does not significantly increase.

Even though the quantity of DR was chosen statistically to represent the real installed DR capacity of a market, there is still a need to design methods to determine DR offer blocks based on real customer load models, and offer prices based on customer willingness models. Along with the realistic DR offer blocks and costs, the shift window should be developed with realistic window sizes according to customer behavior.

# CHAPTER 6 Conclusions and Future Work

## 6.1 Conclusions

A three-layer augmentation-based technique has been explored to develop test case generator data to represent the cost and emission performance of a real deregulated power system. This augmentation based technique utilizes the existing state-of-art test cases which contain synthetic bus, line, and generator information that statistically represent a real power system. The first layer of augmented data proposed in Chapter 3 changes the cost functions of the test case generators so that they produce dynamically changing electricity cost of an electricity market over a time-series simulation. The data required to develop the cost functions come from the electricity market generator offer data. This offer data is masked of the generators' identity. All the similar offers were grouped using statistical pattern recognition to utilize them on a small test case. The proposed technique was tested on eight different test cases from six buses to 2000 buses and produced three years of time-series simulation. This time-series of energy price was compared to the real-market energy price. The proposed technique was found to accurately represent the electricity market with the goodness of fit close 65%.

The second layer of augmented data proposed in Chapter 4.4.1 uses the hourly fuel data of a real electricity market to develop synthetic generator data with an energy mix to represent the real system. The capacity factor of the test case generators and a capacity factor of a fuel type was used as the attribute to augment the market energy mix on the test case. Nineteen fuel-turbine types and their capacity factors have been analyzed to develop the augmented test case data. To the best of my knowledge, this is the highest number of

fuel types on a publicly available test case. Three test cases were augmented using the proposed method and one-year time-series simulation results were used to compare the energy mix with the real system. The energy mix on the test case was found to represent the real system closely. The maximum error in the energy mix of any fossil-fueled source was evaluated close to 2.3% (coal) when compared to the 46% (natural gas) on the state-of-art test case.

The third augmented layer uses the EIA-923 national generator data to develop the heat curves of each of the fuel-generator type. Over 30 fuel-turbine types were analyzed to obtain the heat curves. These heat curves are assigned to the test case generators over the fuel type so that the emissions can be evaluated. To the best of my knowledge, this is the highest number of fuel-turbine types of heat curves available on a publicly available test case. The augmented test case are provided with emission factors along with the emission control factor to estimate the harmful GHG and AP emissions accurately. The GHG ($CO_2$) emission estimated from this technique was 93% accurate to the real system whereas state of the art could not represent the emissions.

With the available market-based test cases, the factors affecting energy demand response was explored. The generator supply curve was found to be one of the most influential factors that determine the revenue for the DR in the energy market. An aggregated DR model is presented in Chapter 5 that presents the advantages of having a dedicated market to trade DR as-a-service. The proposed technique shows the potential revenue on a day that the real market did not call for a DR event.

The contribution of this dissertation will provide the power systems research community with one of the essential test case data that will result in accurate estimation of

costs and emission from the power system. These test case data will allow the researchers to explore opportunities to improve revenue and reduce emissions to create a sustainable power system.

## 6.2 Limitations

Even though this dissertation provides the research community a set of simulation resources, there are limitations to what extent this work can be used. This work was intended to emulate the deregulated energy market operation, which includes unit commitment and economic dispatch. With all the available realistic data from the electricity markets, only economic dispatch was implemented in this dissertation. This work is aimed for researchers who have limited access to electricity market data and want to conduct economic analysis based on steady state operation.

The wholesale electricity markets have multiple sectors such as capacity markets, energy, and ancillary service market. Most of the transactions in the ancillary service market depend on the dynamics of the system. Regulation market is one such service which compensates generators for providing frequency regulation during network disturbances such as loss of generation, sudden change in load. The test cases used in this work are not equipped with the data required to conduct dynamic analysis. Voltage regulation is another ancillary service that the test cases used in this work are not capable of simulating.

No part of this work is developed to predict the electricity market prices. All the results presented are based on the recent past data, and is intended to analyse the cost and emission changes with new energy technologies. No part of this work is intended to be

used to predict the future prices or operation of any electricity market and used for any market malpractices. All publicly available data from this work is intended for research purpose and not for any commercial purpose. No real transactions for any electricity market is to be made from the results from this work.

## 6.3 Future Work

This work presented a data-driven technique to augment test cases. There is a possibility of developing a complete test case by analyzing the enormous open-source data from various organizations. These test cases can be utilized to perform complex emission constrained optimal power flows. The DRX can be extended to trade carbon credits or renewable portfolio credits, using this realistic test cases. The potential value of DR service can be extended not only to benefit the entities by the economy and physical security, but also improve the sustainability of the existing power system.

A fully deregulated system simulated can be realized in which the test case will be capable of simulating all the major markets and replicate the costs. The additional information that needs to be added to the test case for performing all the tasks of a day-ahead market would be (a) generator operational limits, (b) unit commitment information. The ancillary service market for reserve can also be implemented provided the data for regulation and reserve can be attributed to generators in test cases. The cost of a forward capacity market can be added to the entire installed capacity of the test case as a scaled cost from the real market. The electricity markets publish the forward capacity market annual cost per MW-year.

# REFERENCES


[1] *U.S. Energy Information Administration (EIA), Electricity Data Browser*, Available Online: https://www.eia.gov/electricity/data/browser/.

[2] G. Eliasson, "Deregulation, innovative entry and structural diversity as a source of stable and rapid economic growth," *Journal of Evolutionary Economics*, vol. 1, no. 1, pp. 49–63, Mar. 1991.

[3] U.S. Federal Energy Regulatory Commission, *Order No. 888, Promoting Wholesale Competition Through Open Access Non-discriminatory Transmission Services by Public Utilities; Recovery of Stranded Costs by Public Utilities and Transmitting Utilities*, accessed 23 March 2019.

[4] Federal Energy Regulatory Commission, *FERC order 2000*, accessed 23 March 2019. [Online]. Available: https://www.ferc.gov/legal/maj-ord-reg/land-docs/RM99-2A.pdf.

[5] U.S. Environmental Protection Agency, *Sources of greenhouse gas emissions*, Available Online: https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions.

[6] U.S. Energy Information Administration, "Annual Energy Outlook 2017, with projections to 2050," *US Department of Energy*, pp. 1–127, Jan. 2017, Available Online: https://www.eia.gov/outlooks/archive/aeo17/.

[7] U.S. Department of Energy, "Benefits of Demand Response in Electricity Markets and Recommendations for Achieving them. A Report to the United States Congress Pursuant to section 1252 of the Energy Policy Act of 2005," in *U.S. DoE, Washington D.C.*, Available online: http://eetd.lbl.gov/ea/EMP/reports/congress-1252d.pdf (accessed on 10 Oct. 2018), 2006.

[8] H. Allcott, "Rethinking real-time electricity pricing," *Resource and Energy Economics*, vol. 33, no. 4, pp. 820 –842, 2011, Special section: Sustainable Resource Use and Economic Dynamics, ISSN: 0928-7655.

[9] H. T. Haider, O. H. See, and W. Elmenreich, "A review of residential demand response of smart grid," *Renewable and Sustainable Energy Reviews*, vol. 59, pp. 166 –178, 2016, ISSN: 1364-0321.

[10] D. T. Nguyen and L. B. Le, "Risk-constrained profit maximization for microgrid aggregators with demand response," *IEEE Transactions on Smart Grid*, vol. 6, no. 1, pp. 135–146, Jan. 2015.

[11] S. Pineda and A. Bock, "Renewable-based generation expansion under a green certificate market," *Renewable Energy*, vol. 91, pp. 53 –63, Jun. 2016.

[12] L. Gkatzikis, I. Koutsopoulos, and T. Salonidis, "The role of aggregators in smart grid demand response markets," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1247–1257, July 2013.

[13] T. M. Hansen, E. K. P. Chong, S. Suryanarayanan, A. A. Maciejewski, and H. J. Siegel, "A partially observable markov decision process approach to residential home energy management," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 1271–1281, Mar. 2018.

[14] IEEE RTS Task Force of APM Subcommittee, "IEEE Reliability Test System," *IEEE PAS*, vol. 98, no. 6, pp. 2047–2054, Dec. 1979.

[15] A. B. Birchfield, T. Xu, K. M. Gegner, K. S. Shetye, and T. J. Overbye, "Grid structural characteristics as validation criteria for synthetic networks," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3258–3265, Jul. 2017.

[16] J. A. Cullen and E. T. Mansur, "Inferring carbon abatement costs in electricity markets: A revealed preference approach using the shale revolution," *American Economic Journal: Economic Policy*, vol. 9, no. 3, pp. 106–133, Aug. 2017.

[17] H. Lund, "Renewable energy strategies for sustainable development," *Energy*, vol. 32, no. 6, pp. 912 –919, 2007.

[18] H. Lund, S. Werner, R. Wiltshire, S. Svendsen, J. E. Thorsen, F. Hvelplund, and B. V. Mathiesen, "4th Generation District Heating (4GDH): Integrating smart thermal grids into future sustainable energy systems," *Energy*, vol. 68, Apr. 2014, 11 pp. ISSN: 0360-5442.

[19] I. Simonsen, "Volatility of power markets," *Physica A: Statistical Mechanics and its Applications*, vol. 355, no. 1, pp. 10 –20, 2005, Market Dynamics and Quantitative Economics, ISSN: 0378-4371.

[20] M. Parvania, M. Fotuhi-Firuzabad, and M. Shahidehpour, "ISO's optimal strategies for scheduling the hourly demand response in day-ahead markets," *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 2636–2645, 2014.

[21] C Fong, S Haddad, and D Patton, "The IEEE Reliability Test System - 1996," *IEEE Transactions on Power Systems*, vol. 14, no. 3, Aug. 1999.

[22] S. S. Reddy, "Multi-Objective Based Congestion Management Using Generation Rescheduling and Load Shedding," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 852–863, 2017.

[23] C.-l. Su and D. Kirschen, "Quantifying the Effect of Demand Response on Electricity Markets," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1199–1207, Aug. 2009.

[24] E. Dehnavi and H. Abdi, "Determining optimal buses for implementing demand response as an effective congestion management method," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1537–1544, Mar. 2017.

[25] D. T. Nguyen, M. Negnevitsky, and M. de Groot, "Market-based demand response scheduling in a deregulated environment," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 1948–1956, Dec. 2013.

[26] Y. Hayashi et al., "Versatile modeling platform for cooperative energy management systems in smart cities," *Proceedings of the IEEE*, vol. 106, no. 4, pp. 594–612, Apr. 2018.

[27] Y. Wang, C. Wang, C. Miller, S. McElmurry, S. Miller, and M. Rogers, "Locational marginal emissions: Analysis of pollutant emission reduction through spatial management of load distribution," *Applied Energy*, vol. 119, pp. 141 –150, 2014.

[28] V. Raviprabakaran and R. C. Subramanian, "Enhanced ant colony optimization to solve the optimal power flow with ecological emission," *International Journal of System Assurance Engineering and Management*, vol. 9, no. 1, pp. 58–65, Feb. 2018.

[29] M. Parvania, M. Fotuhi-Firuzabad, and M. Shahidehpour, "Assessing impact of demand response in emission-constrained environments," in *2011 IEEE Power and Energy Society General Meeting*, Detroit, MI, July 2011, 6 pp.

[30] C. Kang, T. Zhou, Q. Chen, J. Wang, Y. Sun, Q. Xia, and H. Yan, "Carbon emission flow from generation to demand: A network-based model," *IEEE Transactions on Smart Grid*, vol. 6, no. 5, pp. 2386–2394, Sep. 2015.

[31] Z. Zhao, W. C. Lee, Y. Shin, and K. Song, "An optimal power scheduling method for demand response in home energy management system," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1391–1400, Sep. 2013.

[32] C. Su and D. Kirschen, "Quantifying the effect of demand response on electricity markets," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1199–1207, Aug. 2009.

[33] Q. Hu, F. Li, X. Fang, and L. Bai, "A framework of residential demand aggregation with financial incentives," *IEEE Transactions on Smart Grid*, vol. 9, no. 1, pp. 497–505, Jan. 2018.

[34] Z. Chen, L. Wu, and Y. Fu, "Real-time price-based demand response management for residential appliances via stochastic optimization and robust optimization," *IEEE Transactions on Smart Grid*, vol. 3, no. 4, pp. 1822–1831, Dec. 2012.

[35] T. M. Hansen, R. Roche, S. Suryanarayanan, A. A. Maciejewski, and H. J. Siegel, "Heuristic optimization for an aggregator-based resource allocation in the smart grid," *IEEE Transactions on Smart Grid*, vol. 6, no. 4, pp. 1785–1794, July 2015.

[36] U.S. Federal Energy Regulatory Commission, *Order No. 745, Demand Response Compensation in Organized Wholesale Energy Markets*. Washington D.C.: FERC, Mar. 2011, Availble online: https://www.ferc.gov/EventCalendar/Files/20110315105757-RM10-17-000.pdf (accessed on 10 Oct. 2018).

[37] O. Ma, N. Alkadi, P. Cappers, and et. al., "Demand response for ancillary services," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 1988–1995, Dec. 2013.

[38] D. T. Nguyen, M. Negnevitsky, and M. de Groot, "Pool-based demand response exchangeconcept and modeling," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1677–1685, Aug. 2011.

[39] G. H. Brundtland, "Our common future," The World Commission on Environment and Development, Tech. Rep., Mar. 1987.

[40] United Nations General Assembly, "World summit outcome," Sep. 2005.

[41] S. Sharma and et. al., "Metrics-based assessment of sustainability in demand response," in *IEEE 15th International Conference on Smart City*, Bangkok, Thailand, Dec. 2017, 8 pp.

[42] Q. Zhou and J. W. Bialek, "Approximate model of european interconnected system as a benchmark system to study effects of cross-border trades," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 782–788, May 2005.

[43] SEI, "The European Fossil-fueled Power Station Database Used in the SEI CASM Model," Stockholm Environment Institute, York, England, Tech. Rep., 1996.

[44] T. Xu, A. B. Birchfield, K. M. Gegner, K. S. Shetye, and T. J. Overbye, "Application of large-scale synthetic power system models for energy economic studies," in *2017 Hawaii International Conference on System Science (HICSS)*, Waikolowa, HI, Jan 2017, 7 pp.

[45] Yongjun Ren and F. D. Galiana, "Pay-as-bid versus marginal pricing – Part I: Strategic generator offers," *IEEE Transactions on Power Systems*, vol. 19, no. 4, pp. 1771–1776, Nov. 2004.

[46] Y. Ren and F. D. Galiana, "Pay-as-bid versus marginal pricing – Part II: Market behavior under strategic generator offers," *IEEE Transactions on Power Systems*, vol. 19, no. 4, pp. 1777–1783, Nov. 2004.

[47] F. Wen and A. K. David, "Optimal bidding strategies and modeling of imperfect information among competitive generators," *IEEE Transactions on Power Systems*, vol. 16, no. 1, pp. 15–21, Feb. 2001.

[48] R. J. L. Fan J. Wang and Y. Guan, "Min-max regret bidding strategy for thermal generator considering price uncertainty," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2169–2179, Sep. 2014.

[49] N. Zhang, "Generators' bidding behavior in the NYISO day-ahead wholesale electricity market," *Energy Economics*, vol. 31, no. 6, pp. 897–913, Nov. 2009.

[50] I. Pea, C. B. Martinez-Anido, and B. Hodge, "An extended ieee 118-bus test system with high renewable penetration," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 281–289, Jan. 2018.

[51] D. Krishnamurthy, W. Li, and L. Tesfatsion, "An 8-zone test system based on iso new england data: Development and application," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 234–246, Jan. 2016.

[52] E. Krall, M. Higgins, and R. P. O'Neill, "RTO Unit Commitment Test System," Federal Energy and Regulatory Commission (FERC), Washington D.C., Tech. Rep., July 2, 2012, Available Online: https://www.ferc.gov/industries/electric/indus-act/market-planning/rto-commit-test.asp (accessed on: 03/23/2019).

[53] *U.S. Energy Information and Administration (EIA), Electricity: Form EIA-411*, Available Online: https://www.eia.gov/electricity/data/eia411/ (accessed on: 03/23/2019).

[54] V. Durvasulu and T. M. Hansen, "Market-based generator cost functions for power system test cases," *IET Cyber-Physical Systems: Theory Applications*, vol. 3, no. 4, pp. 194–205, Dec. 2018.

[55] H. Wang, C. E. Murillo-Sanchez, R. D. Zimmerman, and R. J. Thomas, "On computational issues of market-based optimal power flow," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1185–1193, Aug. 2007, ISSN: 0885-8950. DOI: 10.1109/TPWRS.2007.901301.

[56] C. L. Su and D. Kirschen, "Quantifying the effect of demand response on electricity markets," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1199–1207, Aug. 2009.

[57] R. S. Fang and A. K. David, "Transmission congestion management in an electricity market," *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 877–883, Aug. 1999.

[58] *Power Systems Test Case Archive*, accessed 23 March 2019. [Online]. Available: https://www2.ee.washington.edu/research/pstca/.

[59] *PJM Data Miner 2, Generation by Fuel Type*, Available Online: http://dataminer2.pjm.com/list (accessed on: 03/23/2019).

[60] Z. Zhao and L. Wu, "Impacts of high penetration wind generation and demand response on LMPs in day-ahead market," *IEEE Transactions on Smart Grid*, vol. 5, no. 1, pp. 220–229, Jan. 2014, ISSN: 1949-3053.

[61] H. W. Dommel and W. F. Tinney, "Optimal power flow solutions," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-87, no. 10, pp. 1866–1876, Oct. 1968, ISSN: 0018-9510. DOI: 10.1109/TPAS.1968.292150.

[62] K. Y. Lee and Y. M. Park and J. L. Ortiz, "A united approach to optimal real and reactive power dispatch," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-104, no. 5, pp. 1147–1153, May 1985.

[63] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "MATPOWER steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, June, 2011.

[64] The Market Monitoring Unit, PJM, "State of the market report for PJM 2014," Monitoring Analytics LLC, Tech. Rep., 2015, Available online: http://www.monitoringanalytics.com/reports/PJM_State_of_the_Market/2014/2014-som-pjm-volume2.pdf (accessed on 10 Oct. 2018), 510 pp.

[65] U.S. Energy Information Administration (EIA), *Electric Power Monthly, Net Generation by Energy source*, accessed 10 December 2017. [Online]. Available: https://www.eia.gov/electricity/monthly/.

[66] U.S. Energy Information Administration (EIA), *U.S. Natural Gas Electric Power Price*, accessed 20 Jan 2018. [Online]. Available: https://www.eia.gov/dnav/ng/hist/n3045us3m.htm.

[67] PJM Interconnection, *Daily Energy Market Offer Data*. [Online]. Available: http://www.pjm.com/markets-and-operations/energy/real-time/historical-bid-data/unit-bid.aspx.

[68] PJM, Cost Development Subcommittee, *PJM Manual 15: Cost Development Guidelines*, accessed 20 Jan 2018. [Online]. Available: http://www.pjm.com/~/media/documents/manuals/m15.ashx.

[69] V. Durvasulu, T. M. Hansen, and R. Tonkoski, "Classification of generators participating in the bulk-power market," in *2017 IEEE International Conference on Industrial Technology ICIT*, Mar. 2017, 5 pp.

[70] B. Bezerra, L. A. Barroso, and M. V. Pereira, "Bidding strategies with fuel supply uncertainty in auctions of long-term energy call options," *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 653–660, May 2011.

[71] P. K. Nag, "Power Plant Engineering," in *Introduction: Economics of power generation*, 3rd ed. New Delhi: Tata McGraw-Hill Publication Company Limited, 2008, ch. 1, p. 7.

[72] R. Billinton, S. Kumar, N. Chowdhury, K. Chu, K. Debnath, L. Goel, E. Khan, P. Kos, G. Nourbakhsh, and J. Oteng-Adjei, "A reliability test system for educational purposes-basic data," *IEEE Transactions on Power Systems*, vol. 4, no. 3, pp. 1238–1244, Aug. 1989.

[73] ISO-New England, *ISO-NE Pricing Reports: Day-Ahead Energy Offer Data*, accessed Jan. 20, 2018. [Online]. Available: https://www.iso-ne.com/isoexpress/web/reports/pricing/-/tree/day-ahead-energy-offer-data.

[74] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4–37, Jan. 2000.

[75] T. Caliński and J. Harabasz, "A dendrite method for cluster analysis," *Communications in Statistics*, vol. 3, no. 1, pp. 1–27, 1974.

[76] C. Fraley and A. E. Raftery, "Model-based clustering, discriminant analysis, and density estimation," *Journal of the American Statistical Association*, vol. 97, no. 458, pp. 611–631, 2002.

[77] M. E. Tipping and C. M. Bishop, "Probabilistic principal component analysis," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 61, no. 3, pp. 611–622, 1999.

[78] K. He and G. Meeden, "Selecting the number of bins in a histogram: A decision theoretic approach," *Journal of Statistical Planning and Inference*, vol. 61, no. 1, pp. 49 –59, 1997.

[79] G. F. Reid and L. Hasdorff, "Economic dispatch using quadratic programming," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-92, no. 6, pp. 2015–2023, Nov. 1973.

[80] T. Athay, R. Podmore, and S. Virmani, "A practical method for the direct analysis of transient stability," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-98, no. 2, pp. 573–584, Mar. 1979.

[81] Texas A and M University, *Electric grid test case repository*, accessed 20 Jan 2018. [Online]. Available: https://electricgrids.engr.tamu.edu/electric-grid-test-cases/activsg500/.

[82] "State of the Market Report for PJM, 2016," Monitoring Analytics, LLC, Tech. Rep., Available Online: http://www.monitoringanalytics.com/reports/PJM_State_of_the_Market/2016.shtml (accessed on: 03/23/2019).

[83] United Nations Organization, "The role of gas in sustainable energy," in *United Nations Economic Commission for Europe*, Geneva, 2015, 8 pp.

[84] *U.S. Energy Information and Administration (EIA), Today in Energy: Natural gas-fired power plants are being added and used more in PJM Interconnection*, Available Online: https://www.eia.gov/todayinenergy/detail.php?id=37293 (accessed on: 03/23/2019), Washington, D.C.

[85] *U.S. Energy Information and Administration (EIA), Electricity: Form EIA-923*, Available Online: https://www.eia.gov/electricity/data/eia923/ (accessed on: 03/23/2019).

[86] *U.S. Environmental Protection Agency (EPA), Emission Factors for Greenhouse Gas Inventories*, Available Online: https://www.epa.gov/sites/production/files/2018-03/documents/emission-factors_mar_2018_0.pdf (accessed on: 03/23/2019).

[87] *U.S. Environmental Protection Agency (EPA), AP-42: Compilation of Air Emissions Factors*, Available Online: https://www.epa.gov/air-emissions-factors-and-quantification/ap-42-compilation-air-emissions-factors (accessed on: 03/23/2019).

[88] *ISO New England "Operations Report: Dispatch fuel Mix"*, Available Online: https://www.iso-ne.com/isoexpress/web/reports/operations/-/tree/gen-fuel-mix (accessed on: 03/23/2019).

[89] *Pjm member list*, Available Online: https://www.pjm.com/about-pjm/member-services/member-list.aspx (accessed on: 03/23/2019).

[90] V. Durvasulu and T. M. Hansen, "Market-based generator cost functions for power system test cases," *IET Cyber-Physical Systems: Theory Applications*, vol. 3, no. 4, pp. 194–205, Dec. 2018.

[91] *PJM, Environmental Information Services: Public Repots*, Available Online: https://www.pjm-eis.com/reports-and-events/public-reports.aspx (accessed on: 03/23/2019).

[92] U.S. Federal Energy Regulatory Commission, *Order No. 888, Promoting Wholesale Competition Through Open Access Non-discriminatory Transmission Services by Public Utilities*. Washington D.C.: FERC, 1996, Available online: http://www.ferc.gov/legal/maj-ord-reg/land-docs/rm95-8-00v.txt (accessed on 10 Oct. 2018).

[93] M. Albadi and E. El-Saadany, "A summary of demand response in electricity markets," *Electric Power Systems Research*, vol. 78, no. 11, pp. 1989 –1996, Nov. 2008.

[94] J. Aghaei and M.-I. Alizadeh, "Demand response in smart electricity grids equipped with renewable energy sources: A review," *Renewable and Sustainable Energy Reviews*, vol. 18, pp. 64 –72, Feb. 2013.

[95] P. Siano, "Demand response and smart gridsa survey," *Renewable and Sustainable Energy Reviews*, vol. 30, pp. 461 –478, Feb. 2014.

[96] F. Wang, H. Xu, T. Xu, K. Li, M. Shafie-khah, and J. P. Catalo, "The values of market-based demand response on improving power system reliability under extreme circumstances," *Applied Energy*, vol. 193, pp. 220 –231, 2017, ISSN: 0306-2619.

[97] S. S. Reddy, "Multi-objective based congestion management using generation rescheduling and load shedding," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 852–863, Mar. 2017.

[98] F. Stern, M. Shober, M. Tanner, and D. Violette, "Greenhouse gas reductions from demand response: Impacts in three U.S. markets," in *2016 International Energy Policies and Programmes Evaluation Conference*, Sep. 2016, 15 pp.

[99] P. Cappers, C. Goldman, and D. Kathan, "Demand response in u.s. electricity markets: Empirical evidence," *Energy*, vol. 35, no. 4, pp. 1526 –1535, 2010.

[100] Federal Energy Regulatory Commission Staff Team, *Assessment of Demand Response and Advanced Metering*. Washington D.C.: FERC, Dec. 2017, Available online: https://www.ferc.gov/legal/staff-reports/2017/DR-AM-Report2017.pdf (accessed on 10 Oct. 2018).

[101] P. Ponnaganti, J. R. Pillai, and B. Bak-Jensen, "Opportunities and challenges of demand response in active distribution networks," *Wiley Interdisciplinary Reviews: Energy and Environment*, vol. 7, no. 1, 16 pp. Feb. 2018.

[102] H. Zhong, L. Xie, and Q. Xia, "Coupon incentive-based demand response: Theory and case study," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1266–1276, May 2013.

[103] P. Cappers, L. Hans, and R. Scheer, *American Recovery and Reinvestment Act of 2009: Interim Report on Customer Acceptance, Retention, and Response to Time-Based Rates from the Consumer Behavior Studies*. Berkely, CA: Energy Analysis and Environmental Impacts Division Lawrence Berkeley National Laboratory, 2015, Available online: https://emp.lbl.gov/sites/all/files/lbnl1830290.pdf (accessed on 10 Oct. 2018).

[104] P. Palensky and D. Dietrich, "Demand side management: Demand response, intelligent energy systems, and smart loads," *IEEE Transactions on Industrial Informatics*, vol. 7, no. 3, pp. 381–388, Aug. 2011.

[105] U.S. Federal Energy Regulatory Commission, *Order No. 719, Wholesale Competition in Regions with Organized Electric Markets*. Washington D.C.: FERC, 2008, Available online: https://www.ferc.gov/whats-new/comm-meet/2008/101608/E-1.pdf (accessed on 10 Oct. 2018).

[106] J. Kang and J.-H. Lee, "Data-driven optimization of incentive-based demand response system with uncertain responses of customers," *Energies*, vol. 10, no. 10, Oct. 2017 17 pp.

[107] A. Asadinejad and K. Tomsovic, "Optimal use of incentive and price based demand response to reduce costs and price volatility," *Electric Power Systems Research*, vol. 144, pp. 215 –223, Mar. 2017.

[108] A. Khodaei, M. Shahidehpour, and S. Bahramirad, "Scuc with hourly demand response considering intertemporal load characteristics," *IEEE Transactions on Smart Grid*, vol. 2, no. 3, pp. 564–571, Sep. 2011.

[109] H. Wu, M. Shahidehpour, A. Alabdulwahab, and A. Abusorrah, "Demand response exchange in the stochastic day-ahead scheduling with variable renewable generation," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 2, pp. 516–525, Apr. 2015.

[110] M. Shafie-khah, E. Heydarian-Forushani, and et.al., "Strategic offering for a price-maker wind power producer in oligopoly markets considering demand response exchange," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 6, pp. 1542–1553, Dec. 2015.

[111] H. T. Nguyen, L. B. Le, and Z. Wang, "A bidding strategy for virtual power plants with the intraday demand response exchange market using the stochastic programming," *IEEE Transactions on Industry Applications*, vol. 54, no. 4, pp. 3044–3055, July 2018.

[112] D. T. Nguyen, M. Negnevitsky, and M. de Groot, "Walrasian market clearing for demand response exchange," *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 535–544, Feb. 2012.

[113] S. K. Reddy, L. K. Panwar, B. K. Panigrahi, and R. Kumar, "Computational intelligence for demand response exchange considering temporal characteristics of load profile via adaptive fuzzy inference system," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 3, pp. 235–245, June 2018.

[114] P. B. Luh, W. E. Blankson, and et. al., "Payment cost minimization auction for deregulated electricity markets using surrogate optimization," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 568–578, May 2006.

[115] The Market Monitoring Unit, PJM, "State of the market report for PJM 2016," Monitoring Analytics LLC, Tech. Rep., 2017, Available online: http://www.monitoringanalytics.com/reports/PJM_State_of_the_Market/2016/2016-som-pjm-volume2.pdf (accessed on 10 Oct. 2018), 510 pp.

[116] H. Yin and N. Powers, "Do state renewable portfolio standards promote in-state renewable generation?" *Energy Policy*, vol. 38, no. 2, pp. 1140 –1149, Feb. 2010, ISSN: 0301-4215.

[117] N. O'Connell, P. Pinson, H. Madsen, and M. O'Malley, "Benefits and challenges of electrical demand response: A critical review," *Renewable and Sustainable Energy Reviews*, vol. 39, pp. 686 –699, 2014, ISSN: 1364-0321.

[118] California ISO, *What the duck curve tells us about managing a green grid*, Available online: https://www.caiso.com/Documents/FlexibleResourcesHelpRenewables_FastFacts.pdf (accessed on 10 Oct. 2018).

[119] Z. Zhao and L. Wu, "Impacts of high penetration wind generation and demand response on lmps in day-ahead market," *IEEE Transactions on Smart Grid*, vol. 5, no. 1, pp. 220–229, Jan. 2014.

[120] F. Rahimi and A. Ipakchi, "Demand response as a market resource under the smart grid paradigm," *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 82–88, June 2010.

[121] L. Wu, "Impact of price-based demand response on market clearing and locational marginal prices," *IET Generation, Transmission and Distribution*, vol. 7, no. 10, pp. 1087–1095, Oct. 2013.

[122] Z. Liu, A. Wierman, Y. Chen, B. Razon, and N. Chen, "Data center demand response: Avoiding the coincident peak via workload shifting and local generation," *Performance Evaluation*, vol. 70, no. 10, pp. 770 –791, 2013.

[123] A. Safdarian, M. Fotuhi-Firuzabad, and M. Lehtonen, "Integration of price-based demand response in discos' short-term decision model," *IEEE Transactions on Smart Grid*, vol. 5, no. 5, pp. 2235–2245, Sep. 2014.

[124] V. Durvasulu, H. Syahril, and T. M. Hansen, "A genetic algorithm approach for clearing aggregator offers in a demand response exchange," in *2017 IEEE Power Energy Society General Meeting*, July 2017, 5pp.

[125] T. K. Wijaya, M. Vasirani, and K. Aberer, "When bias matters: An economic assessment of demand response baselines for residential customers," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1755–1763, July 2014.

[126] J. Lee, S. Yoo, J. Kim, D. Song, and H. Jeong, "Improvements to the customer baseline load (cbl) using standard energy consumption considering energy efficiency and demand response," *Energy*, vol. 144, pp. 1052 –1063, 2018, ISSN: 0360-5442.

[127] M. R. Sarker, M. A. Ortega-Vazquez, and D. S. Kirschen, "Optimal coordination and scheduling of demand response via monetary incentives," *IEEE Transactions on Smart Grid*, vol. 6, no. 3, pp. 1341–1352, May 2015.

[128] A. Asadinejad, A. Rahimpour, K. Tomsovic, H. Qi, and C. fei Chen, "Evaluation of residential customer elasticity for incentive based demand response programs," *Electric Power Systems Research*, vol. 158, pp. 26 –36, 2018.

[129] V. Durvasulu and T. M. Hansen, “Classifying day-ahead electricity markets using pattern recognition for demand response,” in *North American Power Symposium (NAPS)*, Sep. 2016, 6 pp.

[130] PJM, Cost Development Subcommittee, *PJM Manual 15: Cost Development Guidelines*, Available online: http://www.pjm.com/~/media/documents/manuals/m15.ashx (accessed on 10 Oct. 2018), PJM Interconnection.

[131] Whitley, L Darrell, “The genitor algorithm and selection pressure: Why rank-based allocation of reproductive trials is best,” in *ICGA*, vol. 89, 1989, pp. 116–123.

[132] J. S. Vardakas, N. Zorba, and C. V. Verikoukis, “A survey on demand response programs in smart grids: Pricing methods and optimization algorithms,” *IEEE Communications Surveys Tutorials*, vol. 17, no. 1, pp. 152–178, 2015.

[133] D. Neves and C. A. Silva, “Optimal electricity dispatch on isolated mini-grids using a demand response strategy for thermal storage backup with genetic algorithms,” *Energy*, vol. 82, pp. 436–445, 2015.