Data-Driven Inverse Optimization with Applications in Electricity Markets

by

Ali Rafieepouralavialavijeh

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Management Sciences

Waterloo, Ontario, Canada, 2022

Author's Declaration

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

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

Abstract

Due to the increasing penetration of renewable resources and demand response instruments in the electricity markets, generation planning models have become more complex and require detailed information on the inherent structure of the system, including generator and demand parameters. Demand should be met by cost-effective, adaptable, and efficient power plants, to ensure that it is met even in the worst-case scenarios, such as an unanticipated peak or the failure of a critical generating unit. On the other hand, there is a need to consider short-term details in the Planning problems to address the needed system flexibility due to sudden changes in demand and renewables' generation. Such short-term details increase the size of the models and their related computations. As a result, there is a trade-off between the complexity of the computation and the level of short-term operational details, which should be considered.

Accessing electricity infrastructure data in North America is often difficult due to the lack of open data standards and the proprietary nature of much of the data. The regulations and policies surrounding the data also vary significantly from province to province, making it difficult to access the data uniformly. Additionally, privacy and security considerations can limit access even further. Despite these limitations, there are indirect methods such as inverse optimization(IO) to derive the market parameters using publicly available data; examples of these parameters include generator costs of generation, their capabilities, etc. The discovery of unobservable information via IO could aid energy models to account for operational details without increasing the complexity of their problem. Furthermore, this information can inform policymakers on potential interventions to improve the efficiency of the electricity market.

In this research, a MIP model is developed to incorporate capital and operational costs associated with long-term planning problems. The operational costs of each technology are assumed to be approximated by a series of step-wise functions so that model outcomes, such as generation output, are as close as possible to real-world electricity market generation. The proposed method employs a two-stage algorithmic framework using data-driven inverse optimization and regression. In the first stage, a series of constraints are generated based on relationships between cost and electricity prices. In the second stage, these constraints on costs are added to a problem that finds and reconciles the parameters of the cost functions. To evaluate the performances of the proposed IO-based method, it was applied to a DC-OPF model using the IEEE 24-bus system, which helped eliminate power flow constraints. This approach was then applied to a long-term planning model using Ontario's electricity market data. The results indicate that the proposed approach could find a close solution to the conventional models. In the long-term planning model, the IO-based approach showed more moderate investment policies, while the traditional methods tend to over or under-invest.

Acknowledgements

I would like to express my sincere gratitude to my supervisors Dr. Michael Pavlin, and Dr. Mehrdad Pirnia, for the invaluable guidance and support throughout my thesis process. I am especially thankful for their patience, encouragement, and inspiring critiques. I would also like to thank my family and friends for their love, support, and encouragement. Finally, I thank everyone who helped me directly or indirectly in completing this project.

Dedication

This thesis is dedicated to my mother,

whose unconditional love and unwavering support have been the cornerstones of my life. She has been my guiding light and my source of strength throughout my life. I am deeply grateful for her love, caring, and guidance. I will forever be thankful for her presence in my life.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ATB Annual Technology Baseline

EDP Economic Dispatching Problem

EV Electric Vehicle

GEP Generation Expanssion Planning

GW Gigawatt

IEA International Energy Agency

IESO Independent Electricity System Operator

IO Inverse Optimization

LMP Locational Marginal Prices

MAE Meaning Average Error

MIP Mixed Integer Programming

MW Megawatt

NREL National Renewable Energy Laboratory

OEB Ontario Energy Board

OPF Optimal Power Flow

OPG Ontario Power Generation

SSE Sum of Squared Errors

UC Unit Commitment

Nomenclature

Indices and Sets

$d \in D$ Set of days selected in the clustering analysis as representative days

$g \in \Omega_r$ Set of generators of connected to bus r

$g \in G$ Set of generators

$g \in G_i$ Set of generators of technology i

$h \in H$ Set of hours

$i \in I$ Set of generating technology types

$i \in I_d$ Set of dispatchable technology types

$i \in I_r$ Set of renewable technology types

$j \in J$ Set of Energy storage (Battery) types

$r, s \in \Psi$ Set of buses

$s \in E_r$ Set of buses connected to bus s

$t \in T$ Set of years in the planning horizon

Deterministic Parameters

$\gamma_{i,g}$ Nominal capacity of a cluster generator g of technology type i (MW)

λ_j A fixed proportion of storage capacity that may be charged or released

ρ_t^c Penalty for curtailment of generation from renewable resources during the fiscal year t (\$/MWh)

ρ_t^{rn} Penalty for not satisfying renewable energy planned strategies target during the year t (\$/MWh)

$C_{i,t}^f$ Fixed operation cost of generating technology type i (\$/MW)

$C_{i,t}^v$ Variable operations and maintenance cost of generating technology type i (\$/MWh)

C_i^e Life span extension cost for generator type i

D_b Depreciation rate of batteries

D_g Depreciation rate of generators

F_i^s The maximum percentage of any generator's nominal capacity that can add to spinning reserves

$L_{t,d,h}$ Load in hour h of representative day d of year t (MW)

L_t^{max} Maximum load in year t (MW)

Q_i^v A fraction of available capacity of technology i that can be reliably calculated toward the projected reserve need

R_t^{min} The minimum reserve margin of the power system for year t (fraction of the maximum load, L_t^{max})

$Rd_{i,g}^{max}$ Maximum ramp-down rate for power generator g of technology type i

RN^{min} The renewable generation quota target

$Ru_{i,g}^{max}$ Maximum ramp-up rate for power generator g of technology type i

S^{min} Minimum spinning operating reserve (MW)

W_d Weights of the representative days d

W_g Weights of the generators g

Decision Variables

δ_t Gap from renewable energy targets during year t (MWh)

$b^{ch}_{j,t,d,h}$ Binary variable associated with whether battery j is in charging condition, in year t, in representative day d, and in hour h or not

$b^{dis}_{j,t,d,h}$ Binary variable associated with whether battery j is in discharging condition, in year t, in Representative day d, and in hour h or not

$cu_{t,d,h}$ Curtailment for extra generation during hour h of representative day d of year t (MW)

$e^{G}_{i,t}$ Total expansion of generation for generating type i during year t (MW)

$e^{B}_{j,t}$ Total expansion of battery storage for storage unit type j in area r in year t (MWh)

$m_{i,g,t,d,h}$ Binary variable associated with whether generator g of technology type i starts up during hour h of representative day d in year t or not

$n_{i,g,t,d,h}$ Binary variable associated with whether generator g of technology type i is shutting down during hour h of representative day d in year t or not

$p_{i,g,t,d,h}$ Electricity generation of generator g of technology type i during hour h of representative day d of year t (MW)

$p^{ch}_{j,t,d,h}$ The energy charged into the storage unit j during hour h of representative day d of year t (MW)

$p^{dis}_{j,t,d,h}$ The energy discharged into the storage unit j during hour h of representative day d of year t in MW

$q^{h}_{i,g,t,d,h}$ Spinning reserve capacity of power generator g of technology type i throughout hours s of representative day d of the year t (MW)

$q^{Q}_{i,g,t,d,h}$ Quick-start capacity reserve of power generator g of technology type i during hours s of representative day d of year t (MW)

$r^{G}_{i,g,t}$ Total generation for generator type i in year t that has retired (MW)

$r^{B}_{j,t}$ Total battery storage that has been retired for storage unit type j in year t (MWh)

$sc_{j,t,d,h}$ State of charge of the storage unit j throughout hours of a sample day of the year (MWh)

$u_{i,g,t,d,h}$ Binary variable associated with whether generator g of technology type i is in working condition during hour h of representative day d in year t or not

Chapter 1

Introduction

The world's population is growing, and the trend is anticipated to be ascending in the future. Since there is a strong correlation between population growth and electricity usage, it can be concluded that electricity demand will also rise. However, due to the scarcity of fossil fuels in the long run, as the world population grows and the available fossil fuel resources gradually deplete, these resources cannot be reliable as our primary source of electricity generation. Furthermore, it has been established that the consumption of fossil fuels is the direct source of carbon emissions [4]. Since the beginning of the industrial revolution, a 40 percent growth in the emissions of CO2 has occurred and this is primarily due to the consumption of fossil fuels. Despite the uncertainty about the scale of the effects of greenhouse gas emissions, developing non-fossil fuel-based systems is still a reasonable decision.

Based on reports of the International Energy Agency (IEA), [3], there are estimates that by 2050, more than half of the world's power generation may be from renewable resources. Besides, there is currently a trend toward using renewable energy resources more frequently worldwide. For instance, 29 percent of the energy generated worldwide in 2021 came from renewable resources. Increased use of renewable energy in the power mix will undoubtedly increase uncertainty and unpredictability. The main origin of this uncertainty is the variable nature of these renewable resources and atmospheric and climatic conditions. The amount of power generated may be accurately predictable when utilizing conventional non-renewable energy sources like coal, natural gas, and oil. Unless there are unanticipated equipment problems, electricity market operators typically had total control over how much power they generate each hour when renewable resources had little share in the generation mix. The main driver behind this total control over power generation is the ability of operators to alter typical generator outputs in response to power demand and constraints placed on the generators.

Adding renewable energy resources to a power system will significantly impact both short-term and long-term planning in the markets. Short-term planning involves deciding which generators are On and how much they are generating. At the same time, long-term planning models not only include these decisions but also involve decisions regarding expanding the existing capacities. The primary reason for the impact of renewables on these models is the intermittent nature of renewable energy generation. One of the main downsides of this intermittency for renewable resources is their reliability because operators have little control over when the wind blows or the sun shines. As a result, power system problems had become more complex than when renewables did not have much role, and operators have begun to plan generation differently.

In short-term operational problems, such as unit commitment (UC), electricity market operators generally deal with decisions such as selecting the generators that need to be On, and when they should be On. In this problem, decisions are generally taken hourly. One of the main features affecting these decisions is ramping capability of generators. A generator's ramping capacity is the rate at which it can raise or decrease its output. Market operators consider ramping limitations to limit which generators can ramp up or down to meet changing demand. Renewable energy sources' variability and unpredictability require having these ramping capabilities to reconcile the unpredictable rise and fall of electricity generation.

In the long-term planning problem, in addition to the decisions that need to be taken in a short-term planning problem, decisions such as investment in different generators or transmission lines need to be incorporated. Short-term planning decisions are usually incorporated in long-term planning models to make investment decisions that respect the fluctuations and dynamics of generation and load on different days of the year and in different scenarios. Generally, considering all of these fluctuations and scenarios in a long-term planning model makes the problem computationally intractable. Nevertheless, these fluctuations are necessary for today's energy markets considering the more significant role of renewable resources, while historically, they were omitted and left for short-term planning problems. Hence, there is a trade-off between computational tractability and how closely the dynamics of load and generation are expected to be met.

This thesis attempts to demonstrate the importance of including short-term planning features as specified before with low granularity in long-term expansion planning problems. In addition, a novel inverse optimization algorithmic framework is proposed to compensate for the short-term details that are not considered. These details are either not directly available as a source of information, or omitted to have a more computationally tractable problem.

This thesis is organized as follows: Chapter 2 provides background on long-term planning, electricity markets, and inverse optimization, a technique that finds unknown parameters in an operations research problem, Chapter 3 provides a two-step algorithmic framework to derive the unknown parameters that approximate cost function of different generators, Chapter 4 discusses two long-term planning models using fixed operational costs, and an IO approximated cost functions, and Finally, Chapter 5 provides a summary and some potential areas for further study.

1.1 Motivation

1.1.1 Long-term planning

Based on the annual reports published by IESO, power consumption has many drivers, including but not limited to the economy, population, technology, energy price, and consumer behavior, [1]. Even though many factors affect power consumption and as a result system-wide load, grid-level demand in Ontario's energy market has been mostly flat during the past five years (2016-2021). Ontario electricity market forecasts predict ongoing growth, notably in the residential and commercial sectors, and continued agricultural expansion. Hence, the province is entering a period of increased demand and decarbonization measures and economic recovery are the two main key factors leading the way for new technologies and sources of energy.

In the 2020s, Ontario's nuclear fleet's available capacity will undergo a significant transformation. Resource requirements will rise due to Pickering NGS's retirement and additional renovations at Darlington NGS and Bruce NGS that might result in protracted outages. It is anticipated that by the end of 2023 a total of 9.6 GW of capacity in the nuclear sector would need refurbishment. Refurbishments of some of these generators including Darlington and Bruce are anticipated to finish in 2026 and 2033. Furthermore, several IESO and Ontario Electricity Financial Corporation commitments and generating contracts will expire in the coming decade. Many contracts have already reached the end of their terms, and expirations are expected to skyrocket as by the end of 2030. Figure 1.1 illustrates the trend of these expirations and how they affect the total installed capacity if no investments happen.

Resource adequacy, which specifies the system's supply and demand balance, is integral

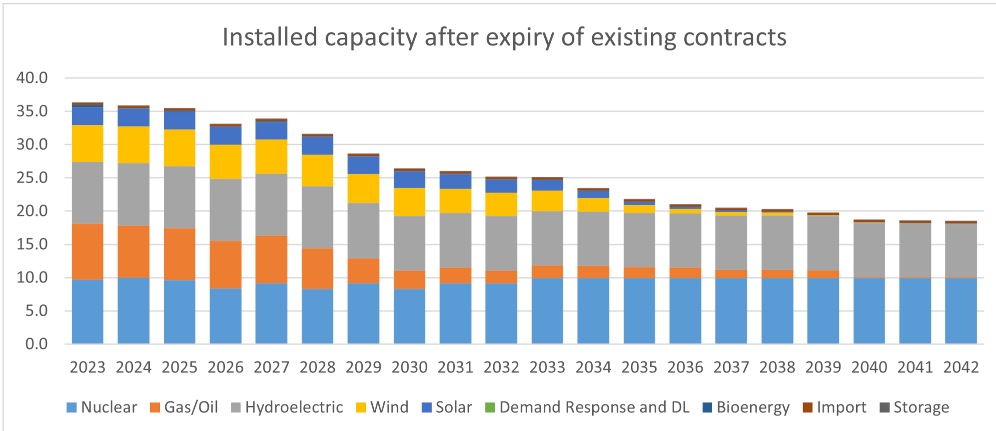


Figure 1.1: Total installed capacity in GW without reacquisition of expired contracts

to power system dependability. For example, severe weather and generator failures might cause demand to outpace supply for an extended period. A sound system can reduce these hazards. The IESO analyses several hazards when determining resource sufficiency. Because of weather circumstances, actual demand may be higher or lower than projected. Due to planned maintenance or equipment breakdown, resources may be unavailable in real-time. Variable generators, such as wind and solar, may deliver relatively low levels of adequate capacity due to their reliance on climatic circumstances. Finally, big projects, such as ongoing nuclear refurbishments, may have delays in resuming service and a greater outage rate once they do.

In order to illustrate the need for future investments the continuous availability of the existing resources would be assumed. Figure 1.2 shows the energy adequacy outlook in Ontario with this assumption. As can be seen, existing generation units are sufficient to supply the majority of demand through the mid-2030s. An energy shortage arises at the end of the planned horizon, mainly owing to rising demand. In the absence of continuous availability of existing supplies beyond contract expiration, an energy deficit occurs in 2026 and accelerates beginning in 2029.

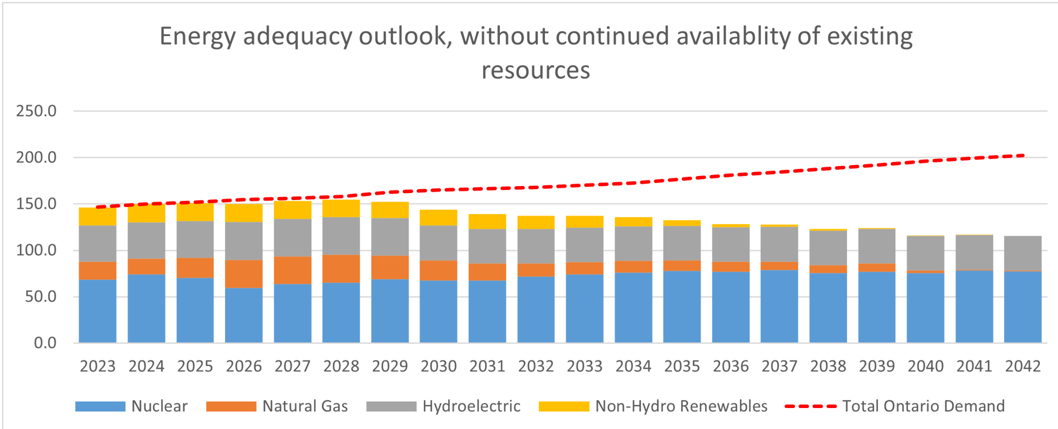


Figure 1.2: Energy adequacy outlook without continued availability of existing resources (driven based on the annual reports published by IESO)

All in all, given the ongoing trends in the Ontario energy market, long-term capacity expansion planning is of utmost importance.

1.1.2 IO-approximated cost functions

As more renewable variability and price-responsive agents are added to the electricity grid, more accurate modeling of their behaviors and interactions becomes a necessity. This necessity roots in the significant variability of these agents in short-term operations of the electricity market. Neglecting the growing participation of these agents would mean investment needs in the future would not incorporate the flexible operations of renewables, gas-fired units, or storage units such as batteries.

While employing estimated cost of operations in the market can be a good source in many studies, they cannot provide operational outcomes close to real-world operations. Market outcomes in each hour are dependent on the bids and offers the market operator receives from market participants and the technical limitations of the grid. From system to system, only parts of these details are available to the general public, making the modeling of these systems prone to errors. As a result, inverse optimization, an area of operations research is generally employed to approximate these invisible or hidden details in the market. Inverse optimization assumes the market outcomes are optimal or close to optimality and tries to reconcile and approximate model parameters given these assumptions.

A series of non-decreasing linear convex cost functions for each of the market participants allows for modeling and reproducing the market outcomes using the limited knowledge of the market. This would allow more reliable investment planning on the market participants with high variability in their generation or operations.

1.2 Literature review

1.2.1 Expansion planning

Models for generation expansion planning include two layers of decision-making: investment and operational layers. While investment decisions are often made on an annual or decennial basis, operational short-term decisions are made more frequently (for example, hourly). As a result, a typical long-term investment planning model would incorporate more operational decisions than investment decisions. Adding more operational decisions leads to increased computational complexity and as a result a difficult mathematical problem. Several papers in the literature have proposed various solutions to this problem.

In the literature, most researchers apply a series of techniques to reduce the size of their problem because of memory and time limits. However, there are other benefits associated with reducing the size of a problem. The authors of [30] emphasize one advantage of this approach: running multiple planning models quickly and doing sensitivity analysis and model debugging. While picking representative days/hours/time slices can substantially influence model outputs, the benefits of lower computational expenditure exceed the penalties associated with less granular model inputs [55].

One of the early works in this area, [17], attracted attention to the impacts of non-dispatchable technologies, such as Solar, and Wind, on the system load profile. The authors propose a stochastic approach to modifying the annual load in the system and selecting several operating conditions from it. In [63], a deterministic approach is devised to expansion planning that considers the unpredictability in demand and resource profiles by using representative time slices. Each season is represented by one day, and each day by 4 time slices. In addition, they take into account a super-peak time slice in the summer for a total of 17 time slices. In [51], every year is considered as 4 seasons, and each season is represented by three days with hourly granularity. The authors of [9] recommended two approaches: the first was to utilize wind and load duration curves, and the second was to use k-means clustering to model geographical correlations. Because of the examination of representative hours within each day, the work of this study violates the chronological sequence of hours and hence intertemporal correlations and operational constraints (for example, ramp-up and ramp-down constraints). The authors of [66] offered a completely different technique, using states and transition tables to represent operational details in a generation expansion planning model utilizing principles from control theory. The authors state this method could offer a more realistic picture of system outcomes such as power pricing and overall cost. The authors of [56] created a measure for selecting representative days in expansion planning exercises in storage systems, and they demonstrated how to employ this measure. In [7], the authors worked on transmission-expansion planning and developed a novel technique to determine representative operating conditions. In contrast to earlier works that concentrate on more frequent situations, the authors of this research emphasize essential network conditions. The authors of [54] proposed the term "Snapshot selection" to reduce the size of the planning problem as well as capture the temporal variability. The authors of [42] employed two algorithms, k-means, and hierarchical clustering methods, to find representative days instead of operating hours or time slices. With this approach, they could model the intertemporal operating constraints and capture correlations in the data.

This research shares the same goals with the literature in that it tries to reduce the size of the long-term planning problem and increase its computational traceability while considering realistic operational, short-term, constraints. As a result of this, investment decisions are made based on all operational dynamics of the system and therefore, are more reliable. The realistic operational constraints in a problem set could cause computational intractability due to the growing number of decision variables and constraints. This thesis proposes a two-stage algorithmic framework that considers the econometric behavior of the generators in the electricity market, without considering all of these realistic operational constraints to alleviate the interactivity issues.

1.2.2 Inverse optimization

Inverse optimization is a branch of operations research that aims to impute unknown parameters in standard optimization problems. These parameters can be found in either the objective function or the constraints, or both. In the literature, there are numerous examples of applications of inverse optimization. This method is employed in different problem contexts including linear problems [6], conic problems[34], convex problems [40], integer problems [61], multi-objective problems [18], and variational inequality problems [11].

Inverse optimization has found applications in a wide range of disciplines, in addition to distinct problem formulations. In particular, it has been applied to the area of power systems in two sets of models, Market-clearing models, and Agent demand models [21].

The modern-day electricity grid is facing an increasing number of price-responsive consumers. Price-responsive consumers can control their expenditures in response to wholesale electricity prices. Accurately predicting the demand for these consumers helps retailers, and market operators make better planning decisions. Inverse optimization is employed to estimate the utility functions of consumer demand models, which can eventually be used to predict demands. Authors of [60] proposed an inverse optimization methodology that uses bi-level programming to infer the market bid parameters of a pool of price-responsive consumers. In [59] a novel IO approach is proposed to estimate a total load of a pool of price-responsive buildings in the short term. Authors of [43] employed IO to estimate the demand response characteristics of price-responsive consumers. In [27] inverse optimization is employed to propose a two-step estimation procedure and model the behavior of a pool of electric Vehicles (EVs).

In most North American electricity markets, a market operator is normally in charge of arranging output and power flow. In the meantime, a market operator should also consider different operational restrictions as well as demand bids for future dates. The problem of arranging the outputs of generators and power flow is called a market-clearing problem. This market-clearing problem is often described as a linear optimization problem that is solved on a daily basis at scheduled intervals. Moreover, the final allocations (solutions, or to be more specific generation levels), prices (shadow prices of the supply-demand constraints), and bid functions are released to ensure market transparency. The authors of [13] employed inverse optimization models that use publicly accessible information as input to infer parameters that were not previously observable and represent institutional restrictions. Being aware of these unobservable characteristics in an energy market, for example, can lead to a better knowledge of transmission capacity, which can lead to information regarding bidding tactics or investment decisions. In [58] an IO-based method is proposed to estimate rival marginal offer prices for a strategic producer in day-ahead market with network constraints. The authors of [69] employed an IO approach in the context of generation expansion planning to design an incentive policy for the promotion of renewable resources.

This thesis proposes a two-step algorithmic framework to estimate the cost functions of each of the generators in an electricity market. This market has unknown features such as transmission lines and flow information. Inverse optimization, as a result, is used to find a series of cost functions that best satisfy the known constraints and adjust to the market outcomes even though some information was unknown. These cost functions have a non-decreasing convex linear structure to find a linear relationship between a set of explanatory variables. This proposed algorithmic framework has been implemented for two sample test cases, namely IEEE 24 bus test system, and the Ontario energy market to show its effectiveness in following the short-term dynamics of an electricity market with unknown parameters. Finally, the calculated cost functions are employed in a long-term planning model to evaluate their effectiveness in long-term problems and incentivizing units such as renewables, and batteries.

Chapter 2

Background

In this chapter, the background of this thesis is discussed in more detail. Different electricity markets are discussed in the first section, and then the capacity expansion planning problem is provided. Finally, inverse optimization methods are discussed and some theoretical backgrounds are provided.

2.1 Electricity markets

An electricity market is a system that enables the exchange of electricity-related goods and services [64]. More than a century of evolution in the electric power industry has changed the economics of electricity markets significantly due to advances in technology, dynamics of supply and demand, and policies. In the early 21st century heavily controlled traditional electricity markets were replaced with several competitive markets for electricity generation, transmission, distribution, and retailing [29]. The transition from traditional market approaches to competitive ones involved a transformation of electricity from a public service into a tradable good [67]. As of the 2020s, there are still large areas of the United States and Canada that are operating under a traditional energy market.

The day-ahead market is a financial market in which various entities, including people and businesses, can sell and buy electricity at a contractually binding price for the following day. Because the prices are fixed at the start of the day, sellers and buyers are hedged against changes and volatility during the day.

In a real-time electricity market, market participants are allowed to buy or sell electricity during an operating day. This market balances the differences between day-ahead commitments and the actual real-time demand and generation of electricity. Further, it generates a different second financial settlement. This settlement would determine locational marginal prices (LMP), which could be used to charge participants in demand or production that deviate from the day-ahead commitments.

2.1.1 Ontario Energy Market

The energy market in Ontario is structured as a competitive market that includes various companies and stakeholders. Energy companies in Ontario are regulated by the Ontario Energy Board (OEB) and must abide by the rules and regulations set out by this authority. Market players include electricity generators, wholesalers, retailers, electric transmission and distribution systems operators, and independent power producers.

Power producers are the primary source of energy; they are responsible for the creation of electricity through their energy resources. In Ontario, there are three primary energy sources: nuclear power, hydroelectric power, and natural gas. These power producers sell the electricity they produce directly to large commercial and industrial customers, thereby setting the spot price.

Wholesalers purchase electricity from power producers and sell it to electricity retailers, who in turn sell it to residential and commercial customers. Reciprocally, retail customers buy electricity from retailers and submit payments to the Ontario Energy Board. In order to provide reliable service to customers, the system must always be stable and secure at a high level.

Electricity generators, wholesalers, retailers, electric transmission and distribution systems operators, and independent power producers are all participants in the Ontario energy market. However, the main players are Ontario Power Generation (OPG), Hydro One, and Independent Electricity System Operator (IESO). OPG is the largest energy provider in Ontario, responsible for generating approximately half of Ontario's electricity needs. Hydro One is responsible for operating and maintaining the province's transmission and distribution systems, as well as managing hydroelectric assets. IESO is responsible for managing the province's power grid, monitoring electricity supply and demand in real time, and ensuring that the market remains stable and secure.

The structure of the energy market in Ontario provides a competitive and reliable energy delivery system. This structure ensures that energy prices remain competitive, and energy is available to the public in an efficient and economic manner.

2.2 Capacity expansion planning

The generation expansion planning (GEP) problem is an optimization problem that ensures enough generation capacity is installed within the electric grid to satisfy the electricity demand. Due to the aging of the current generation units and the increase in demand over time [24], this problem is becoming a vital topic for energy system planning, which necessitates careful examination of the required capacity by either refurbishment or the construction of new ones. The planning horizon for such problems is 10 years to 20 years [23], [45].

Satisfying demand continuously, economically, and efficiently is a complex task that requires the inclusion of the worst-case scenarios, such as when an unanticipated load occurs, or a key market participant fails to dispatch electricity or is unavailable due to refurbishment.

There are two methods for solving the generation expansion planning problems: a) a market framework, [38], [39] and b) a centralized approach [14], [47]. The former suggests that profit-driven market players choose their expansion strategies to maximize their predicted earnings by recouping their investment expenses and selling their electricity production in the market. In contrast, the latter suggests that a central planner chooses the generation growth strategy that is the most economical for the system.

The central planners determine a strategy for expanding the generation infrastructure that would ideally result in an optimal system operation to effectively satisfy demand even in worst-case scenarios. Various objective functions, such as maximizing social welfare, or minimizing generation costs, might be considered for this purpose. This thesis focuses on social welfare maximization in which the capital cost of building new generation capacities is considered.

The aim of GEP is to determine the type and size of facilities to be built in the electricity system in the coming years. The optimal type and size of the generating units are calculated based on the current energy system's structure, future demand, and modifications to the system's design. The decisions are impacted by the retirement of existing generators and the costs of investing in and producing the suggested generating units.

The system topology must be represented in the GEP problem [37], [38]. As the penetration of renewable energy generating units grows, the transmission network becomes increasingly overloaded [12], forcing the installation of additional generating units to relieve system congestion.

Aside from determining the best size and location for new generation capacities, it is also critical to establish the best time to develop them. The generation investment planning could be done statically [22], [65], or dynamically [10], [14]. In the static approach, all investment decisions are made at a single point in time. This reformulation provides a rather simple problem. On the other hand, in the dynamic approach, we must make investment decisions at different points. The dynamic approach provides a more accurate outcome at the expense of model complexity or intractability.

2.2.1 Model Overview

The planning problems are commonly designed to minimize the cost of construction and operations or maximize social welfare in which the construction decision variables are represented by binary variables and generation dispatch instructions are given by continuous variables. This would create a MIP/MINLP problem while the system constraints should be satisfied. The following sections describe the objective function and constraints of such models.

2.2.2 Objective function

In the following equation, a cost minimization objective function is considered.

$$\begin{aligned}\min\ Z = \sum_{t\in T}\Big(&\sum_{i\in I, g\in G_i} K^g_{i,t} e^G_{i,g,t} + \sum_{j\in J} K^b_{j,t} e^B_{j,t} + \sum_{d\in D, h\in H, i\in I, g\in G_i} W_d C^v_{i,t} p_{i,g,t,d,h} + \\ &\sum_{i\in I, g\in G_i} C^f_{i,t} \kappa^G_{i,g,t} + \sum_{j\in J} C^f_{j,t} \kappa^B_{j,t} + \rho^{rn}_t \delta_t + \sum_{g\in G_i, d\in D} \rho^g_t cu_{t,d,h}\Big)\end{aligned} \tag{2.1}$$

The objective function stated in 2.1 has several components. The first term is the total generator capital costs, $\sum_{i\in I, g\in G_i, t} K^g_{i,t} e^G_{i,g,t}$. In this component, $K_{i,t}$ stands for inflated investments for technology i in year t throughout the planning horizon, and $e^G_{i,g,t}$ stands for the expansion for technology i, representative generator g in year t and is greater than or equal to zero. As can be seen, the capital cost of each MW investment is multiplied by the number of MW of investment. The term $\sum_{j\in J, t\in T} K^b_{j,t} e^B_{j,t}$ is the total storage capital costs of storage j, where $K^b_{j,t}$ stands for the capital costs associated with storage unit j in year t, and $e^B_{j,t}$ is a decision variable associated with the expansion of storage unit j. This variable is greater than or equal to zero. The term $\sum_{d\in D, h\in H, i\in I, g\in G_i} W_d C^v_{i,t} p_{i,g,t,d,h}$ is variable generator costs, which are costs associated with each MW of electricity generation, where W_d stands for the weight of the cluster that day d belongs to, $C^v_{i,t}$ is the variable cost of technology i in year t, and $p_{i,g,t,d,h}$ is the generation of representative generator g with a greater than or equal to zero sign. The term $\sum_{i\in I,g\in G_i,t} C^f_{i,t}\kappa^G_{i,g,t}$ is associated with the fixed generator cost, where $C^f_{i,t}$ is fixed cost of technology i at year t, and $\kappa^G_{i,g,t}$ is a decision variable associated with the capacity at year t with a greater than or equal sign. The term $\sum_{j\in J,t\in T} C^f_{j,t}\kappa^B_{j,t}$ is the total fixed storage costs, where the total cost is a multiplication of a fixed cost $C^f_{j,t}$ and the storage capacity variable $\kappa^B_{j,t}$ in year t. This variable has a greater than or equal to zero sign. The term $\sum_{t\in T}\rho^{rn}_t\delta_t$ is the penalty ρ^{rn}_t for not reaching planned goals about the market share of renewable resources, which is dependent on the deficiency δ_t from renewable goals. This δ_t variable can take positive values. The last term, $\sum_{g\in G_i,t\in T,d\in D,h\in H}\rho^g_t cu_{t,d,h}$, is curtailment costs, in which the penalty (ρ^g_t) is associated with each MW of curtailment variable $cu_{t,d,h}$ of renewable resources' generation. This variable is greater than or equal to zero.

2.2.3 Model constraints

This section provides a more detailed description of the constraints included in long-term planning. We would first provide the operational constraints, such as load balance, ramp-up, and ramp-down, and then focus our attention on the constraints of long-term planning, such as renewable generation penetration policies, retirement, refurbishment, and investments.

Operational constraints

The operational constraints refer to those which are reflecting the operations of the system in each hour as follows:

Load balance constraint: Equations 2.2 demonstrate the balance between total demand and supply by all energy storage and generators, in which the generations of representative generators $p_{i,g,t,d,h}$ and discharge of energy storages $p^{dis}_{j,t,d,h}$ which is multiplied by their efficiency μ_j is equal to the load $L_{t,d,h}$, charging of energy storages $p^{ch}_{j,t,d,h}$ multiplied by charging efficiency $1/\mu_j$, and renewable curtailments $cu_{t,d,h}$. The dual of this constraint will provide us with the market price.

$$\sum_{i\in I}\sum_{g\in G_i} p_{i,g,t,d,h} + \sum_{j\in J}\mu_j p^{dis}_{j,t,d,h} = L_{t,d,h} + \sum_{j\in J}\frac{1}{\mu_j}p^{ch}_{j,t,d,h} + cu_{t,d,h}, \qquad \forall t\in T, d\in D, h\in H \tag{2.2}$$

Ramp up constraint: Equations 2.3 impose a bound, $Ru^{max}_{i,g}$, on the amount a particular generator can increase (ramp up) its production from one hour to another. This constraint only applies to hours after 2 AM as transversality constraints are the equations that focus on the relation between hour 24 and hour 1.

$$p_{i,g,t,d,h} - p_{i,g,t,d,h-1} \leq Ru^{max}_{i,g}, \qquad \forall I\in I_d, g\in G_i, t\in T, d\in D, h\in H \tag{2.3}$$

Ramp down constraint: Equations 2.4 impose a bound, $Rd^{max}_{i,g}$, on the amount a particular generator can decrease (ramp down) its production from one hour to another. This constraint only applies to hours after 2 AM as transversality constraints are the equations that focus on the relation between hour 24 and hour 1.

$$p_{i,g,t,d,h-1} - p_{i,g,t,d,h} \leq Rd^{max}_{i,g}, \qquad \forall I\in I_d, g\in G_i, t\in T, d\in D, h\in H \tag{2.4}$$

Transversality constraints: Equations 2.5, and 2.6 are to ensure that the status and production levels of a particular generator follow $Rd^{max}_{i,g}$, $Ru^{max}_{i,g}$, from one day to another.

$$p_{i,g,t,d,1} - p_{i,g,t,d,24} \leq Ru^{max}_{i,g}, \qquad \forall I\in I_d, g\in G_i, t\in T, d\in D \tag{2.5}$$

$$p_{i,g,t,d,24} - p_{i,g,t,d,1} \leq Rd^{max}_{i,g}, \qquad \forall I\in I_d, g\in G_i, t\in T, d\in D \tag{2.6}$$

Generator capacity constraint: Equations 2.7 ensure the amount of production of a particular generator, $p_{i,g,t,d,h}$, in addition to the capacity variable that is reserved, $q^s_{i,g,t,d,h}$, would always be less than or equal to the existing capacity and the capacity that was invested in year t, $\kappa^G_{i,g,t}$. Variable $q^s_{i,g,t,d,h}$ is greater than or equal to zero.

$$p_{i,g,t,d,h} + q^s_{i,g,t,d,h} \leq \kappa^G_{i,g,t}, \qquad \forall i \in I_d, t \in T, d \in D, h \in H \tag{2.7}$$

Minimum total spinning reserve constraint: Equations 2.8 ensure that the total available reserved capacity ($q^s_{i,g,t,d,h}$) across all representative generators of different technology types is more than a ratio (S^{min}) of the load ($L_{t,d,h}$) at year t, day d, and hour h.

$$\sum_{i \in I^{dis}} \sum_{g \in G_i} q^s_{i,g,t,d,h} \geq S^{min} L_{t,d,h}, \qquad \forall t \in T, d \in D, h \in H \tag{2.8}$$

Minimum spinning reserve capacity constraint: Equations 2.9 ensure that the reserved capacity ($q^s_{i,g,t,d,h}$) of a representative generator g of a particular technology i should be greater than or equal to a generator's nominal capacity $\kappa_{i,g,t}$ multiplied by the maximum ratio (F^s_i) that can add to spinning reserves.

$$q^s_{i,g,t,d,h} \geq F^s_i \kappa_{i,g,t}, \qquad \forall i \in I_d, t \in T, d \in D, h \in H \tag{2.9}$$

Minimum reserve margin constraint: Equations 2.10 ensure the minimum amount of renewable energy generated in year t (fraction of maximum annual load, R^{min}_t) should be less than or equal to the total available capacity, $\kappa^G_{i,g,t}$, that can be reliably calculated toward the projected reserve need. Variable $\kappa^G_{i,g,t}$, takes positive values greater than or equal to zero.

$$(1 + R^{min}_t) L^{max}_t \leq \sum_{i \in I_r} Q^v_i \kappa^G_{i,g,t} + \sum_{i \in I_d} \kappa^G_{i,g,t} \qquad \forall t \in T \tag{2.10}$$

State of charge constraint: Equations 2.11 ensure that the state of charge of storage unit j at hour h, $sc_{j,t,d,h}$, is dependent on the state of charge of the previous hour $h-1$ in addition to the amount of charging, $p^{ch}_{j,t,d,h}$, and discharging $p^{dis}_{j,t,d,h}$, that happened at the current hour h. Variables $sc_{j,t,d,h}$, $p^{ch}_{j,t,d,h}$, $p^{dis}_{j,t,d,h}$ take values greater than or equal to zero.

$$sc_{j,t,d,h} = sc_{j,t,d,h-1} + \frac{1}{u_j} p^{ch}_{j,t,d,h} - u_j p^{dis}_{j,t,d,h}, \qquad \forall j \in J, t \in T, d \in D, h \in H \tag{2.11}$$

Charging rate constraint: Equations 2.12 ensure that if charging happens in one of the storage units, it should be less than or equal to a maximum, $b^{ch}_{j,t,d,h}\lambda_j\kappa^B_{j,t}$. This maximum is calculated by multiplying the current existing capacity, $\kappa^B_{j,t}$, and a fraction of the capacity that can be charged or discharged, λ_j. Variable $b^{ch}_{j,t,d,h}$ takes a binary value.

$$p^{ch}_{j,t,d,h} \leq b^{ch}_{j,t,d,h}\lambda_j\kappa^B_{j,t}, \qquad \forall j \in J, t \in T, d \in D, h \in H \tag{2.12}$$

Discharging rate constraint: Equations 2.13 ensure that if discharging, $p^{dis}_{j,t,d,h}$, happens in one of the storage units, it should be less than or equal to a maximum, $b^{dis}_{j,t,d,h}\lambda_j\kappa^B_{j,t}$. This maximum is calculated by multiplying the current existing capacity and a fraction of the capacity that can be charged or discharged. Variable $b^{dis}_{j,t,d,h}$ takes a binary value.

$$p^{dis}_{j,t,d,h} \leq b^{dis}_{j,t,d,h}\lambda_j\kappa^B_{j,t}, \qquad \forall j \in J, t \in T, d \in D, h \in H \tag{2.13}$$

Constant state of charge in start and end of the day constraint: Equations 2.14 ensure that the state of charge of each battery at the end of the day, $sc_{j,t,d,24}$, should be equal to that of the start of the day, $sc_{j,t,d,1}$.

$$sc_{j,t,d,1} = sc_{j,t,d,24}, \qquad \forall j \in J, t \in T, d \in D \tag{2.14}$$

Battery capacity constraint: Equations 2.15 and 2.16 ensure that the state of charge of a storage unit would stay within a specific boundary with a maximum and minimum.

$$sc_{j,t,d,h} \leq \kappa^B_{j,t}, \qquad \forall j \in J, t \in T, d \in D, h \in H \tag{2.15}$$

$$sc_{j,t,d,h} \geq \kappa^{B,min}_{j,t}, \qquad \forall j \in J, t \in T, d \in D, h \in H \tag{2.16}$$

No simultaneous charging and discharging constraint: Equations 2.17 ensure that charging and discharging do not happen simultaneously. In these equations, $b^{ch}_{j,t,d,h}$ and $b^{dis}_{j,t,d,h}$ are binary variables that stand for whether charging or discharging happened for storage unit j, at year t, day d, and hour h.

$$b^{ch}_{j,t,d,h} + b^{dis}_{j,t,d,h} \leq 1, \qquad \forall j \in J, t \in T, d \in D, h \in H \tag{2.17}$$

Planning constraints

Renewable quota target constraint: Equations 2.18 ensure that the objectives and policies to meet the minimum renewable deficiencies, $RN^{min} \sum_{d\in D} \sum_{h\in H} L_{t,d,h}$, are satisfied or otherwise δ_t, the gap from renewable energy targets during the year t, is penalized in the objective function.

$$RN^{min} \sum_{d\in D} \sum_{h\in H} L_{t,d,h} \leq \delta_t + \sum_{d\in D} \sum_{h\in H} W_d (\sum_{i\in I_r} p_{i,g,t,d,h} - cu_{t,d,h}), \qquad \forall t \in T \tag{2.18}$$

Generator active capacity constraint: Equations 2.19 ensure that the capacity of generator g of technology types i at each year $t+1$ during the planning horizon is equal to the capacity of the previous year t, investment $e^G_{i,g,t+1}$, and depreciation that happened (introduced by multiplying D_i a fraction of available capacity after depreciation, by the capacity). The previous year's capacity is subject to depreciation, which means a small fraction of the capacity is depreciated each year.

$$\kappa^G_{i,g,t+1} = D_i \kappa^G_{i,g,t} + e^G_{i,g,t+1} \qquad \forall i \in I, g \in G_i, t \in T \tag{2.19}$$

Battery active capacity constraint: Equations 2.20 ensure that the capacity of storage unit j in year $t+1$ is a summation of capacity in the previous year t, investments, and depreciation.

$$\kappa^B_{j,t+1} = D_j \kappa^B_{j,t} + e^B_{j,t+1}, \qquad \forall j \in J, t \in T \tag{2.20}$$

2.2.4 Transforming the model to MIP

Given the instances that make this model MINLP, the Big-M method is used to transform these non-linear terms.

The term $b^{ch}_{j,t,d,h}\kappa^{B}_{j,t}$ is replaced by the variable $\eta^{ch}_{j,t,d,h}$ which takes values greater than or equal to zero. As such these new constraints are added to our problem:

$$\eta^{ch}_{j,t,d,h} \leq \kappa^{B}_{j,t}, \qquad \forall j \in J, t \in T, d \in D, h \in H \tag{2.21a}$$

$$\eta^{ch}_{j,t,d,h} \geq \kappa^{B}_{j,t} - M(1 - b^{ch}_{j,t,d,h}), \qquad \forall j \in J, t \in T, d \in D, h \in H \tag{2.21b}$$

$$\eta^{ch}_{j,t,d,h} \leq M b^{ch}_{j,t,d,h}, \qquad \forall j \in J, t \in T, d \in D, h \in H \tag{2.21c}$$

$$\eta^{ch}_{j,t,d,h} \geq 0, \qquad \forall j \in J, t \in T, d \in D, h \in H \tag{2.21d}$$

Similarly, the term $b^{dis}_{j,t,d,h}\kappa^{B}_{j,t}$ is replaced by the variable $\eta^{dis}_{j,t,d,h}$ which takes values greater than or equal to zero. As such these new constraints are added to our problem:

$$\eta^{dis}_{j,t,d,h} \leq \kappa^{B}_{j,t}, \qquad \forall j \in J, t \in T, d \in D, h \in H \tag{2.22a}$$

$$\eta^{dis}_{j,t,d,h} \geq \kappa^{B}_{j,t} - M(1 - b^{dis}_{j,t,d,h}), \qquad \forall j \in J, t \in T, d \in D, h \in H \tag{2.22b}$$

$$\eta^{dis}_{j,t,d,h} \leq M b^{dis}_{j,t,d,h}, \qquad \forall j \in J, t \in T, d \in D, h \in H \tag{2.22c}$$

$$\eta^{dis}_{j,t,d,h} \geq 0, \qquad \forall j \in J, t \in T, d \in D, h \in H \tag{2.22d}$$

2.3 Inverse optimization

Inverse optimization is a subfield of operations research. In contrast to the common stream of most optimization problems, where the goal is to compute optimal decisions given a set of specified constraints and objective functions, in the inverse optimization context, the optimal solutions to the problem are given, and it is expected to estimate specific parameters either in the objective function or in the constraints. This issue has gotten much interest in recent years, both in terms of mathematics and applications. This section will give some background information on this area of operations research.

In recent years, there has been a surge of interest in both the mathematics and applications of inverse optimization. In most applications, we face an observable data set of decisions and must develop an optimization model that best reproduces these observations. These application fields cover a wide range of themes, including but not limited to power systems, healthcare, and transportation. Decision variables might include electricity consumption patterns in power systems, and utility functions can be approximated. In healthcare systems, decision variables might be medical treatments, and to be estimated could be clinical treatment objectives that best fit with observed decisions. In transportation systems, decision variables might represent routing decisions, and route-choice preferences could be approximated.

The forward (traditional) problem and the required attributes of the inverse model are two of the most important aspects influencing the tractability of an inverse optimization problem. Because each application field has unique assumptions and details, several inverse models and solution methodologies exist. All models, however, may be thought of as a mixture of three dimensions.

The forward problem's structure is the first dimension. A problem's structure might be linear, convex, integer (discreet optimization), or sequential (for example, Markov decision processes). The second dimension is the type of parameters to be approximated. These parameters might be in the objective function, the set of constraints, or both, which significantly impacts the tractability of the inverse optimization problem. The third and last dimension is the inverse optimization model's expectations, whether it should approximate the parameters such that it fully optimizes the observed decisions or whether an imperfect measure of fitness is also desired. These three dimensions, in order, can also be considered as three steps in a sequence that must be considered for building an inverse optimization problem.

Classical and data-driven inverse optimization are labels associated with two sets of problems discussed in the third dimension. In early literature on inverse optimization, classical inverse optimization approaches, mainly used to reformulate the inverse model, dominated. These models are applied to problems that necessitate optimality of the decisions. On the other hand, data-driven models focus on scenarios when we have noisy data or imperfections in our decision variables. For that, other levels of complexity are added to the problem on top of the classical methodologies.

2.3.1 Fundamentals of Inverse Optimization

Formulating an inverse optimization problem starts by defining a parametric forward optimization problem that incorporates the decision-making process of one or more agents. Next, given a set of observed decisions, the inverse problem tries to find the parameters of this problem such that the decisions are exact exactly optimal or close to optimality. Here, we start by formulating these two problems.

The forward problem

The parametric forward problem can include parameters in both the objective function and the set of constraints. Equation 2.23 below identifies the general formulation for a constrained optimization problem. The term g is the vector of cost parameters associated with each decision variable x, A is the matrix of constraint parameters associated with inequality constraints, and likewise, G is the matrix of constraint parameters associated with the equality constraints.

$$\begin{aligned} \min \quad & c^T x & (2.23a) \\ \text{s.t.} \quad & Ax \leq b & (2.23b) \\ & Gx = h & (2.23c) \\ & x \geq 0 & (2.23d) \end{aligned}$$

The inverse problem

When looking at a data collection of decisions, these decisions might be the consequence of several variations on a particular problem. Consider an agent that solves the decision-making process in a variety of settings. Given this data set, an inverse optimization method computes a parameter vector that optimizes the aggregate fit of the linked forward models to the decision set. An estimated parameter set is considered a perfect fit for the decision set if the decisions that are the outcome of the rebuilt model are in the decision set. 2.24 is a general formulation of an inverse problem based on the forward problem formulated before. In this model, x^0 represents the optimal solutions to the primal problem, c^0 represents the prior knowledge about cost c and L is a loss function that minimizes the distance between c and c^0.

$$\begin{aligned} \min \quad & L(c, c^0) & (2.24a) \\ \text{s.t.} \quad & A\pi \leq c & (2.24b) \\ & \pi(Ax^0 - b) = 0 & (2.24c) \end{aligned}$$

$$\lambda(Gx^0) = 0 \tag{2.24d}$$
$$\pi \in \mathcal{R}^+ \tag{2.24e}$$

This problem includes the set of optimality conditions such that the set of decisions x^0 becomes optimal, given the loss function L that is considered.

Choosing a Structure for an Inverse Model

Inverse optimization, as previously indicated, offers a wide range of applications. As a result, the chosen structure is heavily influenced by the application area—most applications in the literature deal with design or estimation problems.

Inverse optimization is used in design applications to create systems that provide the desired outputs or decisions, [46], [57], [5]. The decision sets of the agents are considered to be known in these sets of problems. Hence the emphasis is mainly on classical inverse optimization approaches.

In the estimation applications of inverse optimization, which are the subject of this study, model parameters are calculated using decision inputs reflecting observed agent behavior, [50], [11], [41]. In this case, the data-driven inverse optimization can generate parameters no matter whether some of the decision sets are infeasible, noisy, inaccurate, or even not optimal. This feature of data-driven inverse optimization enables it to cope with massive decision data sets obtained from noisy environments, irregular behavior, or unexpected contextual characteristics.

2.3.2 Classical inverse optimization

This section will examine inverse optimization methods for forward problem structures. For each of these problem structures, we demonstrate that we can formulate and solve the forward problem by considering the optimality conditions associated with our forward problem.

The literature's most common objective function in the context of inverse optimization is stated in the equation below, which is a general formulation for an inverse optimization problem. g is the feasible set, and C^{inv} is the inverse feasible set.

$min_g\{\|c - c^*\| \ \mid c \in C^{inv}, c \in c\}$

The following characterization of the inverse feasible set is provided in early classical inverse optimization problems such as [6].

$C^{inv} = \{c | f(x^*, g) \leq f(x, g), \forall x \in X(c)\}$

In this equation, for x^* to be optimal, the value associated with objective function $f(x^*, c)$ should be necessarily less than or equal to all feasible solutions. Using this equation to find the parameters requires us to iterate through all feasible solutions. This is the main reason why early research in inverse optimization focused on problems with a finite set of feasible solutions. Examples of these forward models include shortest path problems [15], spanning tree models [68], and minimum cost flow models [32]. Since practically enumerating all of the technically feasible solutions for the forward problem is intractable, in future research, new alternative characterizations are developed for the feasible inverse set.

Estimating the objective function

One of the earliest frameworks for inverse optimization and, in particular, for estimating the objective function was proposed by [6]. This framework estimates the objective for a general bounded linear forward problem. In this work, the authors demonstrate that the inverse problem associated with the forward problem is linear. They characterize the inverse feasible set C^{inv} by considering the complementary slackness constraints that are one of the principal conditions of the optimality of a solution. On the other hand, other researchers use strong duality; examples of these works include [20], [62], and [28].

These works emphasize that multiple equivalent formulations for an inverse optimization problem are associated with a forward problem. For example, the property 1 shows the formulation of an inverse problem based on the complementary slackness.

Property 1 (Complementary Slackness) *x^* is an optimal solution for the forward problem, and π an optimal solution for the dual problem, if and only if these conditions hold: 1) dual feasibility constraints $A^T\pi = c$, $\pi \geq 0$ and 2) complementary slackness* $(Ax^* - b)^T\pi = 0$

As a result of this property, the final inverse problem that should be solved to uncover the unknown parameters is stated in equations 2.25.

$$\min_{\pi,c} \quad \|c - c^*\|_n \tag{2.25a}$$

$$\text{s.t.} \quad A^T\pi = c \tag{2.25b}$$

$$(Ax^* - b)^T\pi = 0 \tag{2.25c}$$
$$c \in C \tag{2.25d}$$
$$\pi \geq 0 \tag{2.25e}$$

In this problem, n represents the norm that is used. If the norm is 1 or ∞, the inverse problem 2.25 will become a linear program.

An equivalent formulation for the inverse problem is the application of strong duality. Property 2 states the conditions that are needed to be met to construct an inverse problem based on strong duality.

Property 2 (Strong Duality) *x^* is an optimal solution for the forward problem, and π an optimal solution for the dual problem, if and only if these conditions hold: 1) dual feasibility constraints $A^T\pi = c$, $\pi \geq 0$ and 2) strong duality $c^Tx^* = \pi^Tb$.*

As a result of this property, the final inverse problem that should be solved to uncover the unknown parameters is stated in the equations 2.26.

$$\min_{\pi,c} \quad \|c - c^*\|_n \tag{2.26a}$$
$$\text{s.t.} \quad A^T\pi = c \tag{2.26b}$$
$$c^Tx^* = \pi^Tb \tag{2.26c}$$
$$c \in C \tag{2.26d}$$
$$\pi \geq 0 \tag{2.26e}$$

Estimating parameters in the constraints

Parameters could be on either the left-hand-side of a constraint, A, or the right-hand-side, b. Authors of [19] consider a forward problem where the right-hand-side of each constraint b is known, and the left-hand-side A should be estimated through inverse optimization. They explained that for a solution x^* to be optimal, it should at least satisfy one equality constraint. As a result, the equation below demonstrates the inverse feasible region:

$X^{inv} = \{A|\exists\pi \geq 0 : A^T\pi = c, c^Tx^* = \pi^Tb\}$

According to [19], we can get an optimal solution to the inverse problem by altering the adjacent facet until $x*$ satisfies the appropriate constraint. Given a set of constraints such as $Ax^* \geq b$, we will try to find the closest facet $\{x|A_jx \geq b_j$ and then perturb it until the corresponding constraint is satisfied.

2.3.3 Data-driven inverse optimization

As mentioned at the beginning of this chapter, the emphasis in a data-driven inverse optimization environment is on recovering model parameters while considering solutions that may be infeasible, noisy, erroneous, or even not optimum. This allows us to work with enormous volumes of data. In this part, we will offer a quick introduction to several loss functions and solution techniques in the context of inverse optimization. Most inverse optimization literature focuses on convex forward models with observable input parameters like $u \in \mathcal{U}$. We can define the convex forward model stated below for a given input $\hat{u}_i$.

$$min_x \quad \{f(x, \hat{u}_i, c) | g(x, \hat{u}_i, c) \leq 0\} \tag{2.27}$$

In practice, because there are visible differences between instances, these u_Is capture them. In the context of a data-driven problem, we witness a data set of decisions, and optimum solution sets $\mathcal{D} := (\hat{x}_i, \mathcal{X}_i^{opt}(c))_{i=1}^N$ that are drawn independently and identically from a probability distribution $\mathcal{P}$. To represent the empirical distribution associated with the data set, we may use $\mathcal{P}_N := \sum_{i=1}^N \delta_{(\hat{x}_i, \mathcal{X}_i^{opt}(c))}$. A unit probability mass is represented by δ in this equation. If we assume $l()$ to represent the loss function, we can write a data-driven inverse optimization problem as follows:

$$Z(l, \mathcal{P}_N) := min_{c \in C} \quad \frac{1}{N} \sum_{i=1}^{N} l(\hat{x}_i, \mathcal{X}_i^{opt}(c)) \tag{2.28}$$

The equation mentioned in 2.28 might incorporate another term beyond the loss function associated with breaches in the inverse-feasibility. This term can refer to any application-specific objective.

Distance from the optimal solution set

A straightforward method for estimating the error of a forward problem concerning an optimal solution such as $\hat{x}$ is to calculate its distance for $\hat{x}$. We can define the minimum distance loss function as follows:

$$L(\hat{x}, \mathcal{X}^{opt}(\theta)) := min_{x \in \mathcal{X}^{opt}} ||x - \hat{x}||_2 \tag{2.29}$$

This equation calculates the 2-norm distance of x from the optimal set. The Inverse Distance problem is the data-driven inverse optimization problem Z with the loss above.

Consider the convex forward models described in the model 2.27 parameterized by $\hat{u}_i$ in [8]. The authors show that 2.27 is a statistically consistent estimator of g for these models.

Violating the KKT conditions

In [40], an inverse optimization problem method is considered in which KKT optimality conditions could be violated. The authors proposed a series of loss functions to incorporate the degrees of deviation to optimality. For each sample (x^k, p^k), we denote r^k_{ineq} the residuals associated with the violation of the inequality constraints in the forward problem, r^k_{eq} the residuals associated with the violation of equality constraints in the forward problem, r^k_{comp} the residuals associated with the violation of complementary slackness constraints, and r^k_{stat} as residuals associated with the violation of stationary constraints. For each sample, the primal residuals are fixed and do not depend on the parameters we are estimating in the objective. As a result, the problem below is the final model to be solved.

$$\min \sum_k L(r^k_{comp}, r^k_{stat}) \tag{2.30}$$

$$\text{s.t.} \;\; \pi \geq 0 \quad \forall k \tag{2.31}$$

Chapter 3

Inverse optimization for DC-OPF

This chapter presents two DC-OPF models and then determines their unknown parameters using a novel optimization method. Long-term planning models frequently employ these operational models to incorporate more short-term operational details into investment decisions. A correct estimation of these unknown parameters can provide many advantages in modeling different market agents and reducing the computational complexity of expansion planning models that include operational constraints.

As the long-term model in the background section specifies, each planning model has a set of parameters. Some of these parameters can be estimated using publicly available data sources from the electricity market operators or academic studies. In general, most parameters are specific to the type and location of a generator. For example, different technologies have different technical features (parameters), and if they are renewable generators their location comes into the scene. These features in addition to details of grid topology are proprietary information and therefore, they are difficult to obtain, and the provided estimations need to generate the desired outcomes.

In this chapter, these unknown features are approximated into a cost function, the parameters of which are to be estimated. Estimating the cost function of each market agent would reduce the need to consider the operational details of their behavior and activity. As a result of these cost approximations, the size of the problem is significantly reduced, and future problems become computationally tractable. Hence, an inverse optimization method is employed to find an approximation for these parameters. This section discusses the suggested method in more detail.

3.1 Operational Model

In this section, two operational models are presented. The first is a DC optimal power flow (OPF) problem. The OPF problem is frequently used in practice to determine day-ahead and real-time allocations and prices in energy markets. While in practice, transmission constraints are non-linear in nature, this model features linearized transmission constraints. Despite the frequent use of OPF problems, details of transmission constraints are difficult to obtain. As a result, a second approach is presented: the approximate optimal power flow with extended costs (OPF-EC). The OPF-EC incorporates the transmission costs implied by each generator's power output as part of their generation cost.

3.1.1 DC Optimal Power Flow problem

The OPF problem is the benchmark allocation problem in this thesis and determines energy production at a set of generators over a single operating day. This model features a set of buses connected by transmission lines. Each bus has a load and a set of co-located generators. Each generator is indexed by g, is located at a particular bus r, and can produce $p_{i,g,d,h}$ MWh of electricity. Each generator faces some technical limits, including an upper bound for its generation $\kappa_{i,g}$, ramping limits $Ru^{max}_{i,g,d}$ and $Rd^{max}_{i,g,d}$, and bounds on the difference in generation at starting and ending hours of the day. Each transmission line between buses r and s, namely rs, has a maximum capacity of P^{max}_{rs} and the voltage angles $\theta_{r,d,h}$. The objective function minimizes the generation costs. These generation costs are a function of a set of parameters β, $p_{i,g,d,h}$, and the load $L_{r,d,h}$. The parameters β are estimated in the following sections using inverse optimization.

Model 3.1 is the OPF problem. Equations 3.1b are supply-demand balance constraints; these constraints ensure that the generation units that are connected to a bus in addition to the transmitted electricity from other buses would satisfy the load at that specific bus. Equations 3.1c represent the ramping-up constraints, and 3.1d represent the ramping-down constraints; these constraints ensure that the generations units do not exceed their technical capabilities to increase or to decrease their generation levels. Equations 3.1e and Equations 3.1f show the transversality constraints; these constraints ensure that the generation units would be within a boundary when a day starts and ends (hour 1 and hour 24). This is intuitive, as generation from the end of one day cannot abruptly change in the first hour of the next day. Equations 3.1g and Equations 3.1j represent a lower-bound and an upper-bound on the capacity of each of the generation units. Finally, Equations 3.1h demonstrate transmission constraints, and Equations 3.1i represent reference bus constraints that ensure the direction of electricity flow satisfies the demand in different buses.

$$
\begin{aligned}
\min_d \;\; & Z = \sum_{(i,g)\in\Omega_r, r\in\Psi} \Big(\sum_{h\in H} p_{i,g,d,h} C_i^v(\theta, p_{i,g,d,h}, L_{r,d,h})\Big) && && \text{(3.1a)}\\
\text{s.t.} \;\; & \sum_{(i,g)\in\Omega_r} p_{i,g,d,h} + \sum_{s\in E_r} b_{rs}.(\theta_{r,d,h} - \theta_{s,d,h}) - L_{r,d,h} = 0, && \sigma_{1,r,d,h}, \forall r \in \Psi, h \in H && \text{(3.1b)}\\
& - p_{i,g,d,h} + p_{i,g,d,h-1} + Ru_{i,g}^{max} \geq 0, && \sigma_{2,i,g,d,h}, \forall (i,g) \in \Omega_r, h \neq 1 && \text{(3.1c)}\\
& - p_{i,g,d,h-1} + p_{i,g,d,h} + Rd_{i,g}^{max} \geq 0, && \sigma_{3,i,g,d,h}, \forall (i,g) \in \Omega_r, h \neq 1 && \text{(3.1d)}\\
& - p_{i,g,d,24} + p_{i,g,d,1} + Ru_{i,g}^{max} \geq 0, && \sigma_{4,i,g,d,h}, \forall (i,g) \in \Omega_r && \text{(3.1e)}\\
& - p_{i,g,d,1} + p_{i,g,d,24} + Rd_{i,g}^{max} \geq 0, && \sigma_{5,i,g,d,h}, \forall (i,g) \in \Omega_r && \text{(3.1f)}\\
& - p_{i,g,d,h} + \kappa_{i,g} \geq 0, && \sigma_{6,i,g,d,h}, \forall (i,g) \in \Omega_r, h \in H && \text{(3.1g)}\\
& - b_{rs}.(\theta_{r,d,h} - \theta_{s,d,h}) + P_{rs}^{max} \geq 0, && \sigma_{8,r,s,h}, \forall r \in \Psi, s \in E_r, r \neq s, h \in H && \text{(3.1h)}\\
& \theta_{REF,h} = 0, && \sigma_9, \forall h \in H && \text{(3.1i)}\\
& p_{i,g,d,h} \geq 0, && \sigma_{7,i,g,d,h}, \forall (i,g) \in \Omega_r, h \in H && \text{(3.1j)}
\end{aligned}
$$

3.1.2 Approximate Optimal Power Flow with Extended Costs (OPF-EC)

The Approximate Optimal power flow with Extended Costs (OPF-EC) is a model used throughout this study due to the limitations in the data sources employed in the case studies. This model is built on the original OPF problem but lacks transmission line constraints. As a result, this model can be interpreted as a single bus system where all the market participants and loads are connected to one bus. All other features are similar to the OPF problem: electricity generation by each unit is bound to an upper-bound $\kappa_{i,g}$, and generation difference from one hour to the next hour should be within a bound $Ru_{i,g,d}^{max}$ and $Rd_{i,g,d}^{max}$, and finally the generation of a unit at the end of the day and the start of the day should respect a boundary.

Model 3.2 represents the OPF-EC problem. In this model, Equations 3.2b represent supply-demand balance constraints, Equation 3.2d represent ramp-up constraints, Equation 3.2e represent ram-down constraints, Equation 3.2f and Equation 3.2g represent transversality constraints, Equation 3.2h represent capacity constraints for power generators, and Equation 3.2i represent the positivity of generation as a set of constraints.

$$\min_{p_{i,g,d,h}} \quad Z(p;d) = \sum_{i \in I, g \in G_i} w_{i,g}\Big(\sum_{h} p_{i,g,d,h} C^{v}_{i,g,d,h}(\beta, p_{i,g,d,h}, L_{d,h})\Big) \tag{3.2a}$$

$$\text{s.t.} \sum_{i \in I, g \in G_i} w_{i,g} p_{i,g,d,h} - L_{d,h} = 0, \sigma_{1,d,h},\ \forall h \in H \tag{3.2b}$$

$$- p_{i,g,d,h} + p_{i,g,d,h-1} + Ru^{max}_{i,g,d} \geq 0, \sigma_{2,i,g,d,h},\ \forall i \in I, h \neq 1, g \in G_i \tag{3.2c}$$

$$- p_{i,g,d,h} + p_{i,g,d,h-1} + Ru^{max}_{i,g,d} \geq 0, \sigma_{2,i,g,d,h},\ \forall i \in I, h \neq 1, g \in G_i \tag{3.2d}$$

$$- p_{i,g,d,h-1} + p_{i,g,d,h} + Rd^{max}_{i,g,d} \geq 0, \sigma_{3,i,g,d,h},\ \forall i \in I, h \neq 1, g \in G_i \tag{3.2e}$$

$$- p_{i,g,d,24} + p_{i,g,d,1} + Ru^{max}_{i,g,d} \geq 0,\ \sigma_{4,i,g,d,h},\ \forall i \in I, g \in G_i \tag{3.2f}$$

$$- p_{i,g,d,1} + p_{i,g,d,24} + Rd^{max}_{i,g,d} \geq 0,\ \sigma_{5,i,g,d,h},\ \forall i \in I, g \in G_i \tag{3.2g}$$

$$- p_{i,g,d,h} + \kappa_{i,g} \geq 0,\ \sigma_{6,i,g,d,h},\ \forall i \in I_d, h \in H, g \in G_i \tag{3.2h}$$

$$p_{i,g,d,h} \geq 0,\ \sigma_{7,i,g,d,h},\ \forall i \in I, h \in H, g \in G_i \tag{3.2i}$$

3.2 Inverse optimization

Model 3.2 which was introduced in the previous section is employed to estimate the parameters, β, of a set of ascending cost functions, $C^{v}_{i,g}$. The objective is to ensure that the model outcomes such as $p_{i,g,d,h}$, as well as locational market price $\sigma_{1,d,h}$ would be optimal or close to optimality. In inverse optimization, this model is called a forward problem.

Since the goal is to estimate an approximation for the cost function of each generator, there is a need to reformulate the OPF_EC problem as an inverse optimization problem. This means there is a need to build another model considering the optimality conditions of the OPF_EC operational model. Equation 3.3 is the Lagrangian function associated with this model, from which most optimality conditions are driven. These constraints are dual feasibility, complementarity, and primal feasibility constraints.

$$
\begin{aligned}
L = &\sum_{i\in I, g\in G_i} w_{i,g}(\sum_{h\in H} p_{i,g,d,h} C^v_{i,g,d,h}) - \sigma_{1,d,h}(\sum_{i\in I, g\in G_i} w_{i,g} p_{i,g,d,h} - L_{d,h}) - \sum_{i\in I, h\neq 1, g\in G_i} (\sigma_{2,i,g,d,h}(\\
&- p_{i,g,d,h} + p_{i,g,d,h-1} + Ru^{max}_{i,g,d})) - \sum_{i\in I, h\neq 1, g\in G_i} \sigma_{3,i,g,d,h}(-p_{i,g,d,h-1} + p_{i,g,d,h} + Rd^{max}_{i,g,d}) \\
&- \sum_{i\in I, g\in G_i} (\sigma_{4,i,g,d}(-p_{i,g,d,24} + p_{i,g,d,1} + Ru^{max}_{i,g,d})) - \sum_{i\in I, g\in G_i} (\sigma_{5,i,g,d}(-p_{i,g,d,1} + p_{i,g,d,24} \\
&+ Rd^{max}_{i,g,d})) - \sum_{i\in I, h\in H, g\in G_i} (\sigma_{6,i,g,d,h}(-p_{i,g,d,h} + \kappa_{i,g})) - \sum_{i\in I, h\in H, g\in G_i} (\sigma_{7,i,g,d,h}(p_{i,g,d,h}))
\end{aligned} \tag{3.3}
$$

Equations 3.4a-d show the dual problem associated with the operational model. This model is built to derive dual feasibility constraints, one of the three set of equations that should be considered to make an inverse optimization problem.

$$
\begin{aligned}
\max_\sigma \; Z = &\sum_{i\in I, g\in G_i, h\in H} (-\sigma_{1,d,h} L_{d,h} + \sigma_{2,i,g,d,h} Ru^{max}_{i,g,d} + \sigma_{3,i,g,d,h} Rd^{max}_{i,g,d} + \sigma_{4,i,g,d} Ru^{max}_{i,g,d} \\
&+ \sigma_{5,i,g,d} Rd^{max}_{i,g,d} + \sigma_{6,i,g,d,h}\kappa_{i,g})
\end{aligned} \tag{3.4a}
$$

$$
\begin{aligned}
\text{s.t.} \;\; & w_{i,g} C^v_{i,g,d,h} - w_{i,g}\sigma_{1,d,h} - \sigma_{2,i,g,d,h+1} + \sigma_{3,i,g,d,h+1} - \sigma_{4,i,g,d} + \sigma_{5,i,g,d} + \sigma_{6,i,g,d,h} \\
& - \sigma_{7,i,g,d,h} = 0, \forall i \in I, h = 1, g \in G_i
\end{aligned} \tag{3.4b}
$$

$$
\begin{aligned}
& w_{i,g} C^v_{i,g,d,h} - w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h} - \sigma_{2,i,g,d,h+1} - \sigma_{3,i,g,d,h} + \sigma_{3,i,g,d,h+1} \\
& + \sigma_{6,i,g,d,h} - \sigma_{7,i,g,d,h} = 0, \forall i \in I, h \neq 1, 24, g \in G_i
\end{aligned} \tag{3.4c}
$$

$$
\begin{aligned}
& w_{i,g} C^v_{i,g,d,h} - w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h} - \sigma_{3,i,g,d,h} + \sigma_{4,i,g,d} - \sigma_{5,i,g,d} + \sigma_{6,i,g,d,h} \\
& - \sigma_{7,i,g,d,h} = 0, \forall i \in I, h = 24, g \in G_i
\end{aligned} \tag{3.4d}
$$

Equations B.2 below demonstrate complementarity constraints associated with the operational model. Complementarity constraints are one of the three constraints in an inverse optimization problem.

$$
\sigma_{1,d,h}(\sum_{i\in I, g\in G_i} w_{i,g} p_{i,g,d,h} - L_{d,h}) = 0, \qquad \forall h \in H \tag{3.5a}
$$

$$\sigma_{2,i,g,d,h}(-p_{i,g,d,h} + p_{i,g,d,h-1} + Ru^{max}_{i,g,d}) = 0, \qquad \forall i \in I, h \neq 1, g \in G_i \tag{3.5b}$$
$$\sigma_{3,i,g,d,h}(-p_{i,g,d,h-1} + p_{i,g,d,h} + Rd^{max}_{i,g,d}) = 0, \qquad \forall i \in I, h \neq 1, g \in G_i \tag{3.5c}$$
$$\sigma_{4,i,g,d}(-p_{i,g,d,24} + p_{i,g,d,1} + Ru^{max}_{i,g,d}) = 0, \qquad \forall i \in I, g \in G_i \tag{3.5d}$$
$$\sigma_{5,i,g,d}(-p_{i,g,d,1} + p_{i,g,d,24} + Rd^{max}_{i,g,d}) = 0, \qquad \forall i \in I, g \in G_i \tag{3.5e}$$
$$\sigma_{6,i,g,d,h}(-p_{i,g,d,h} + \kappa_{i,g}) = 0, \qquad \forall i \in I_d, h \in H, g \in G_i \tag{3.5f}$$
$$\sigma_{7,i,g,d,h}(p_{i,g,d,h}) = 0, \qquad \forall i \in I, h \in H, g \in G_i \tag{3.5g}$$

The third set of constraints to be considered is primal feasibility constraints. Given our problem's assumption that production levels in the market are optimal, these constraints are satisfied and, therefore, trivial.

Since for the IO model $p_{i,g,d,h}$s have known parameters, all the values within the parentheses in the constraint set B.2 are known, and as a result, based on whether the value is zero, greater than zero, or less than zero these constraints can be split into two categories:

- If the constraints in the primal problem are binding (which means the value within the parentheses is zero), then the dual variable would be free in its value
- If the constraints in the primal problem are non-binding (the value within the parentheses is non-zero), then the dual variable would be zero.

As a result of this step, a set of simplified constraints are generated on how different generators behave in the scenarios of each observation.

Equations 3.6a, 3.6b, and 3.6c are showing how the cost function f is calculated in each of the cases ($h = 1$, $h = 24$, and $h \neq 1, 24$).

$$C^v_{i,g,d,h}(p_{i,g,d,h}, L_{d,h}, \beta) = (w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h+1} - \sigma_{3,i,g,d,h+1} + \sigma_{4,i,g,d} - \sigma_{5,i,g,d} - \sigma_{6,i,g,d,h} + \sigma_{7,i,g,d,h})/w_{i,g}, \forall i \in I, h = 1, g \in G_i \tag{3.6a}$$

$$C^v_{i,g,d,h}(p_{i,g,d,h}, L_{d,h}, \beta) = (w_{i,g}\sigma_{1,d,h} - \sigma_{2,i,g,d,h} + \sigma_{2,i,g,d,h+1} + \sigma_{3,i,g,d,h} - \sigma_{3,i,g,d,h+1} - \sigma_{6,i,g,d,h} + \sigma_{7,i,g,d,h})/w_{i,g}, \forall i \in I, h \neq 1, 24, g \in G_i \tag{3.6b}$$

$$C^v_{i,g,d,h}(pf_{i,g,d,h}, L_{d,h}, \beta) = (w_{i,g}\sigma_{1,d,h} - \sigma_{2,i,g,d,h} + \sigma_{3,i,g,d,h} - \sigma_{4,i,g,d} + \sigma_{5,i,g,d} - \sigma_{6,i,g,d,h} + \sigma_{7,i,g,d,h})/w_{i,g}, \forall i \in I, h = 24, g \in G_i \tag{3.6c}$$

Below a set of scenarios are mentioned. Each of these scenarios for the status of the dual variables helps us understand the IO model in more detail. With these cases in mind, a combination of cases could happen for our $C^v_{i,g,d,h}$. These combinations of cases are represented in the equations A.1, A.2, and A.3 in Appendix.

$$\text{Production capacity constraints} = \begin{cases} \text{Binding UB,} & \sigma_{6,i,g,d,h} \geq 0, \sigma_{7,i,g,d,h} = 0 \quad (3.7a) \\ \text{Binding LB,} & \sigma_{6,i,g,d,h} = 0, \sigma_{7,i,g,d,h} \geq 0 \quad (3.7b) \\ \text{Not binding,} & \sigma_{6,i,g,d,h} = 0, \sigma_{7,i,g,d,h} = 0 \quad (3.7c) \end{cases}$$

$$\text{Ramping constraints} = \begin{cases} \text{Binding UB,} & \sigma_{2,i,g,d,h} \geq 0, \sigma_{3,i,g,d,h} = 0 \quad (3.8a) \\ \text{Binding LB,} & \sigma_{3,i,g,d,h} \geq 0, \sigma_{2,i,g,d,h} = 0 \quad (3.8b) \\ \text{Not binding,} & \sigma_{2,i,g,d,h} = 0, \sigma_{3,i,g,d,h} = 0 \quad (3.8c) \end{cases}$$

$$\text{Transversality constraints} = \begin{cases} \text{Binding UB,} & \sigma_{4,i,g,d,h} \geq 0, \sigma_{5,i,g,d,h} = 0, \quad \text{if } h = 1, 24 \quad (3.9a) \\ \text{Binding LB,} & \sigma_{5,i,g,d,h} \geq 0, \sigma_{4,i,g,d,h} = 0, \quad \text{if } h = 1, 24 \quad (3.9b) \\ \text{Not binding,} & \sigma_{4,i,g,d,h} = 0, \sigma_{5,i,g,d,h} = 0, \quad \text{if } h = 1, 24 \quad (3.9c) \end{cases}$$

Using the formula stated above, some economic insights can be extracted from the operations of different market participants and how they affect the cost attribute associated with them. For example, a time slot can be understood wherein a particular generator is not producing at its maximum, and it does not hit its ramping limits, nor is it in starting or ending hours of the day. In this particular scenario, using equations mentioned before it can be realized that $C^v_{i,g}(p_{i,g,d,h}, L_{d,h}, \beta) = \sigma_{1,d,h}$. Using this idea, a constrain and reconcile approach is proposed that uses the economic rationality of market participants to find constraints over the costs being approximated. In addition, a set of assumptions should be considered for approximating the costs. Among these assumptions is that the cost functions should be non-decreasing, and these estimations should not be time-dependent, which means the parameters of the cost function should not be indexed by time, but the model inputs could be indexed by the time given the dynamic nature of the market.

3.2.1 Constrain and reconcile approach

The constrain and reconcile approach is a two-step algorithmic framework. In the first step, constrain step, based on the existing econometric knowledge set, a set of constraints is generated. These constraints limit the scope of the parameters of cost functions that need to be estimated. In the second step, the generated constraints in the previous step are included in another optimization problem. This optimization problem is called reconcile step and its objective is to find cost functions within the scope of their limits that best satisfy the requirements of the problem. These requirements may be violated due to the approximations and assumptions so far; as such, the objective is to minimize these violations. This algorithmic framework has a principal assumption underlying it. That is, the extended cost function approximated using inverse optimization is consistent concerning the parameters in β. In addition, it is assumed that the system cost for each generator increases in a manner that is affine concerning total renewable generation and total load in the system.

Step 1 - Constrain:

Given the econometric knowledge set, in the first step of this proposed approach, the attention is focused on finding three categories of constraints on compatible supply cost functions. These constraints separate each of the time slots of the data set based on whether the cost associated with that specific generator has been below, equal, or higher than the price in the market. For example, if a generator at a particular time is generating at its maximum capacity and ramping constraints are not active, it means that it must have been economically sound for it to behave like that, as a result of which, it can be concluded that the cost of production at this level must have made sense, which is equivalent to saying the cost was at or below the market price.

Problem 3.10 shows a formulation of a problem called the constrain step. The main goal is to find a set of feasible costs where their absolute distance from the market price is as far as possible. This is the reason why the sense of the objective in this problem is maximization. The estimated costs in this step are only used for the purpose of distinguishing between the three sets of constraints. These costs are indexed by generator and period and used to identify the constraints. For example, if the feasible cost is below the market price a constraint ensures this is added to a set of constraints.

Constraints 3.11 identify the measure of distance. Equations 3.12, 3.13, and 3.14 are the set of dual constraints associated with the operational problem. Equations 3.15 to 3.21 are complementarity constraints associated with the equality, and inequality constraints in the operational problem, respectively. The three sets of constraints derived in this step will be the foundation of our second step, reconcile. Our second stage algorithm tries to minimize violations of these three categories of constraints.

$$\max \ Z = |\epsilon_{i,g,d,h}| \tag{3.10}$$

s.t.

$$\epsilon_{i,g,d,h} = C^v_{i,g,d,h}(p_{i,g,d,h}, L_{d,h}, \beta) - \sigma_{1,d,h} \qquad \forall i \in I, g \in G_i, d \in D, h \in H \tag{3.11}$$

$$\begin{aligned} &w_{i,g}C^v_{i,g,d,h} - w_{i,g}\sigma_{1,d,h} - \sigma_{2,i,g,d,h+1} + \sigma_{3,i,g,d,h+1} - \sigma_{4,i,g,d} + \sigma_{5,i,g,d} + \sigma_{6,i,g,d,h} \\ &- \sigma_{7,i,g,d,h} = 0, \forall i \in I, h = 1, g \in G_i \end{aligned} \tag{3.12}$$

$$\begin{aligned} &w_{i,g}C^v_{i,g,d,h} - w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h} - \sigma_{2,i,g,d,h+1} - \sigma_{3,i,g,d,h} + \sigma_{3,i,g,d,h+1} \\ &+ \sigma_{6,i,g,d,h} - \sigma_{7,i,g,d,h} = 0, \forall i \in I, h \neq 1, 24, g \in G_i \end{aligned} \tag{3.13}$$

$$\begin{aligned} &w_{i,g}C^v_{i,g,d,h} - w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h} - \sigma_{3,i,g,d,h} + \sigma_{4,i,g,d} - \sigma_{5,i,g,d} + \sigma_{6,i,g,d,h} \\ &- \sigma_{7,i,g,d,h} = 0, \forall i \in I, h = 24, g \in G_i \end{aligned} \tag{3.14}$$

$$\sigma_{1,d,h}\Big(\sum_{i \in I, g \in G_i} w_{i,g}p_{i,g,d,h} - L_{d,h}\Big) = 0, \ \forall h \in H \tag{3.15}$$

$$\sigma_{2,i,g,d,h}(-p_{i,g,d,h} + p_{i,g,d,h-1} + Ru^{max}_{i,g,d}) = 0, \ \forall i \in I, h \neq 1, g \in G_i \tag{3.16}$$

$$\sigma_{3,i,g,d,h}(-p_{i,g,d,h-1} + p_{i,g,d,h} + Rd^{max}_{i,g,d}) = 0, \ \forall i \in I, h \neq 1, g \in G_i \tag{3.17}$$

$$\sigma_{4,i,g,d}(-p_{i,g,d,24} + p_{i,g,d,1} + Ru^{max}_{i,g,d}) = 0, \ \forall i \in I, g \in G_i \tag{3.18}$$

$$\sigma_{5,i,g,d}(-p_{i,g,d,1} + p_{i,g,d,24} + Rd^{max}_{i,g,d}) = 0, \ \forall i \in I, g \in G_i \tag{3.19}$$

$$\sigma_{6,i,g,d,h}(-p_{i,g,d,h} + \kappa_{i,g}) = 0, \ \forall i \in I_d, g \in G_i \tag{3.20}$$

$$\sigma_{7,i,g,d,h}(p_{i,g,d,h}) = 0, \ \forall i \in I, h \in H, g \in G_i \tag{3.21}$$

Figure 3.1 demonstrates the three most common econometric scenarios for a generator when ramping constraints are not active. If a generator in a particular hour h is generating electricity at its maximum limits, it can be concluded that its generation cost must have been below the market price of electricity; otherwise, generation would not be economical. In this scenario, the constraint associated with time-slot h would be a member of $\mathcal{C}_{le}$. Likewise, suppose a generator generates at a level between its lower and upper bound limits in the time slot h. In that case, its generation cost must have been equal to the market price of electricity. Otherwise, this generator should have generated more electricity closer to its upper bound if the cost was below. If the cost was higher than the market price of electricity, it should have generated closer to its lower bound limits. In this case, the constraint associated with time-slot h would be a member of $\mathcal{C}_{eq}$. Finally, if a generator is not generating at time-slot h, or in other words, is Off, it must not have been economical for it to generate electricity; in other words, its cost of generation must be above the market price of electricity. In this case, the constraint associated with time-slot h would be a member of $\mathcal{C}_{ge}$.

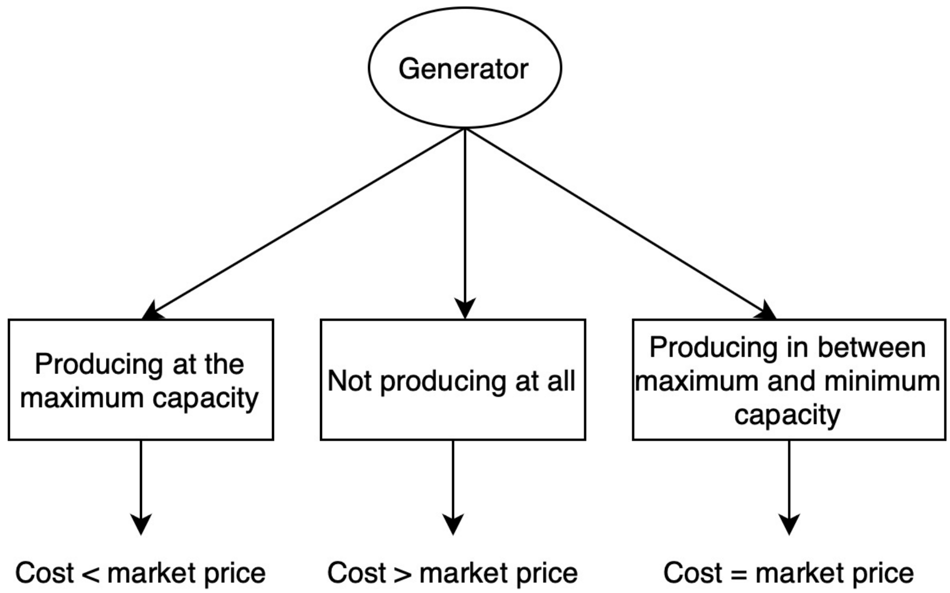


Figure 3.1: The three scenarios in the behavior of a generator in the market when ramping constraints are not active

Step 2 - Reconcile:

The reconciliation step derives cost function parameters consistent with the constraints ($\mathcal{C}_{eq}$, $\mathcal{C}_{gr}$, $\mathcal{C}_{le}$) determined in the constrain step. The algorithmic framework allows for many possibilities regarding the parametric form of the generator cost functions and the parameter estimation. These parameters may include a range of factors that impact the cost to generate and deliver energy from the target generator to the target destination. In the empirical work discussed in 4, these parameters include a constant β_1, the coefficient of productions β_2, the coefficient of the load β_3, and the coefficient of total renewable generator productions β_4. It is assumed that generator costs are a linear function of these variables and coefficients. Problem 3.22 shows a formulation solving this particular problem. This simple linear regression algorithm minimizes the squared magnitude of constraint violations.

In Equation 3.23, on the left-hand side, generator costs are calculated in day d and hour h. On the right-hand side, there are a market price and two variables $\alpha_{i,g,d,h}$ and $\delta_{i,g,d,h}$ which penalize cases when it is hard to precisely model a cost function that respects the set of constraints from the constrain step of the algorithm. The variable $\alpha_{i,g,d,h}$ takes a value when there is a violation of the three sets of constraints discussed before. However, the variable $\delta_{i,g,d,h}$ takes a value when there is no violation of any of the constraints, in other words, it only takes a value for the sake of maintaining the balance of left-hand-side and right-hand-side this constraint.

$$\min_{\beta,\alpha,\delta} \ Z = \sum_{i,g,d,h} \alpha_{i,g,d,h}^2 \tag{3.22}$$

s.t.

$$\begin{aligned} &\beta_1^{i,g} + \beta_2^{i,g} p_{i,g,d,h} + \beta_3^{i,g} L_{d,h} + \beta_4^{i,g} \sum_{i \in I_r} p_{i,g,d,h} = \sigma_{d,h} + \alpha_{i,g,d,h} \\ &+ \delta_{i,g,d,h}, \ \ \forall i \in I, g \in G_i, d \in D, h \in H \end{aligned} \tag{3.23}$$

$$\delta_{i,g,d,h} = 0, \quad \forall i \in I, g \in G_i, (d,s) \in \mathcal{C}_{i,g}^{eq} \tag{3.24}$$

$$\delta_{i,g,d,h} \geq 0, \quad \forall i \in I, g \in G_i, (d,s) \in \mathcal{C}_{i,g}^{gr} \tag{3.25}$$

$$\delta_{i,g,d,h} \leq 0, \quad \forall i \in I, g \in G_i, (d,s) \in \mathcal{C}_{i,g}^{le} \tag{3.26}$$

For instance, if a time slot has a constraint in $\mathcal{C}_{\rceil\mathrm{II}}$ then if the supply cost 3.24 is not able to be equal to $\sigma_{d,h}$ then $\delta_{i,g,d,h}$ takes value. Depending on whether the left-hand side of the 3.23 is greater than or lesser than $\sigma_{d,h}$ on the right-hand side, the value $\delta_{i,g,d,h}$ takes range from negative to positive values.

3.3 Application of Inverse Optimization to Derive Costs

Inverse optimization is employed to determine the parameters of the OPF-EC and approximate the generation and system costs. Figure 3.2 shows the procedure of how IO is applied and validated in the context of an independent energy system operator (IESO).

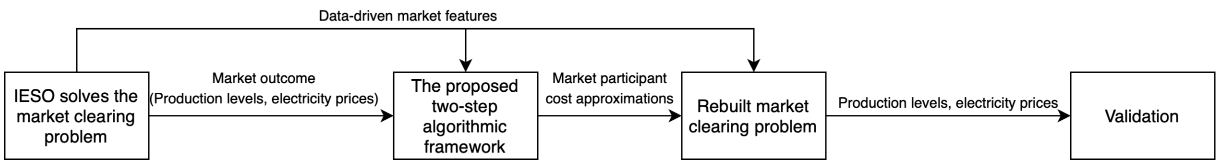


Figure 3.2: Summary of how the constrain-and-reconcile algorithm is used

In this diagram, the market operator, such as IESO, solves a mathematical problem to determine the market outcomes. These outcomes are published to the public and are accessible. Using these outcomes, a set of market features are derived. These market features include ramping rates and hourly demand. These features are parameters of the OPF-EC. The proposed two-step algorithmic framework is built on top of these two sources of data and employs inverse optimization to rebuild the model that the market operator solves such that the market outcomes are close to real-world outcomes.

3.3.1 IEEE test Case Results

In order to check the validity of the proposed two-stage constrain and reconcile approach, it is employed for a modified version of an IEEE 24-BUS test system. Figure 3.3 shows a schematic view of this test system. The IEEE Reliability Test System single-area version [31] serves as the foundation and is modified for use in the energy market and power system operating studies. The most important reason why this test system is selected for our study is that most test systems are only a snapshot of a power system and therefore lack operational details such as ramping constraints, reserve constraints, and any other constraint that incorporates operational details, the updated version of IEEE 24-BUS test system provided in [48] include most details that are required for a power system operation studies. For limiting the computational difficulty, and making it easier to understand the correlation of the most important factors (for example, productions) with the cost function approximations, certain details are omitted. These details include reserve costs and constraints, minimum up-time, and downtime constraints.

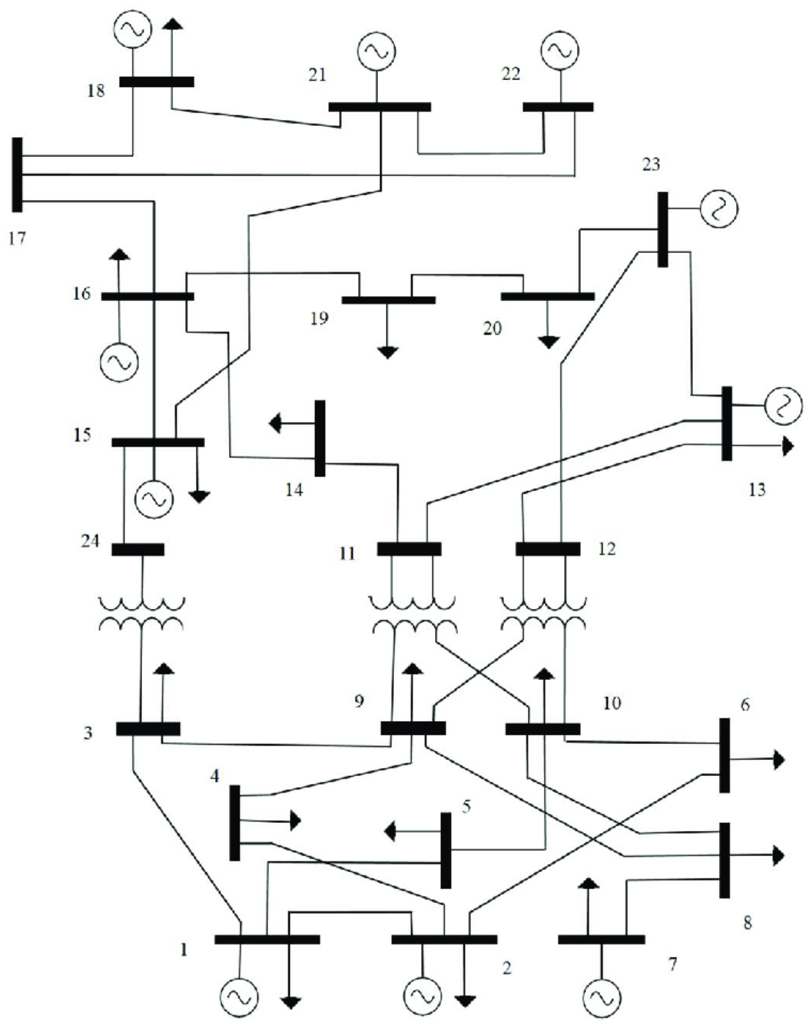


Figure 3.3: IEEE 24-bus test system, taken from [48]

Before applying the proposed methods to the IEEE 24 BUS Test system, it is worth mentioning how we have used this system. Since this test system comes with 24 buses, as such we will have different locational marginal prices for each bus. In order to make our IEEE 24 bus system resemble the level of information, and technical considerations that are visible in the larger electricity market considered, here we propose a method. This approach is also an inverse optimization method that finds uniform prices that are good enough to satisfy the optimality conditions of a market clearing problem. This IO approach will help extrapolate and relate the results of this system with future larger systems.

For the sake of simplicity, some of the features of the original model have been simplified and as such model, a simplified IEEE 24-bus-test-system is used. This simplified model is solved to generate market outcomes that are needed for further steps. These market outcomes are production levels $p_{c,d,h}$, and locational marginal prices $\sigma_{1,i,d,h}$. We assume that these data are known to all market participants as it is in many energy markets.

IO model to recover a uniform price for the whole market

As is the case for many energy markets, there are some components of data that are deemed confidential and as a result not published. An example of such data is transmission limits and power transmissions. If the real-world model is reduced such that there would only be one bus in the system, and as a result, all the loads and generators are connected to that one bus, electricity prices would not be locational. In a real-world energy market such as Ontario, there are some side payments to recover for the costs of some participants. These side payments are necessary because many operational details have not been accounted for when clearing the market. In addition, there is no information about the extra payments market operators make, but it can be assumed that these payments are a result of not considering operational details.

Appendix B contains a market clearing that is suitable for all locations. This problem tries to find and reconcile a market price that is reflective of the market realities such as transmission line congestion. Finding such a market price might violate some of the requirements or assumptions of the problem such as optimality of the generation levels and dispatching.

The uniform market price is essential for the analysis of this section as the models to recover unobservable information in the market rely on the limited information available in the market. In addition, this part discusses how omitting some of this locational information would affect the ability to recover them, or reproduce market outcomes.

Load

The IEEE test case in this section is a snapshot of a 24-bus test system. As a result, the study of various scenarios requires having different demand levels and trajectories. Here, load profiles are generated by scaling the original IEEE 24-BUS system loads using data from the Ontario energy market. Table 3.1 shows the node locations as well as the load at each node as a percentage of the total system demand. Based on this table and representative days in a particular year in the Ontario energy market, a set of load profiles have been produced that match the capabilities of the generators in the IEEE 24 bus system. Here, it is assumed that percentages of system load remain fixed for all of the generated demand scenarios.

Table 3.1: Node location and distribution of the total system demand [48]

Load #	Node	% of system load	Load #	Node	% of system load
1	1	3.8	10	10	6.8
2	2	3.4	11	13	9.3
3	3	6.3	12	14	6.8
4	4	2.6	13	15	11.1
5	5	2.5	14	16	3.5
6	6	4.8	15	18	11.7
7	7	4.4	16	19	6.4
8	8	6	17	20	4.5
9	9	6.1			

Table 3.2: IEEE 24-BUS system: Technical data of generating units

Unit #	Node	P_i^{max} (MW)	P_i^{min} (MW)	R_i^U	R_i^D
1	1	152	30.4	120	120
2	2	152	30.4	120	120
3	7	350	75	350	350
4	13	591	206.85	240	240
5	15	60	12	60	60
6	15	155	54.25	155	155
7	16	155	54.25	155	155
8	18	400	100	280	280
9	21	400	100	280	280
10	22	300	300	300	300
11	23	310	108.5	180	180
12	23	350	140	240	240

The modifications that have been considered for the updated IEEE 24-BUS test system data set are provided here. Since one of the common ways to reduce the complexity of capacity expansion planning models is to consider representative days, here, some load profiles are generated for each of these representative days in the context of the IEEE 24-BUS test system. The Ontario energy market is considered as a sample and its representative days are found. Using the load profiles of a sample market such as the Ontario energy market, the load distributions across all of the buses can be scaled to make sure there are different demand scenarios.

The process of finding representative days in the Ontario energy market and in the context of load profile generation includes an application of k-means clustering on the demand data. Figure 3.4 demonstrates what some of the load profiles look like. As can be seen, the load profiles from Ontario have been scaled based on the maximum load that is reported by [48]. In addition, a careful look at the figure provided shows how the clustering algorithm has identified similar load profiles as representatives.

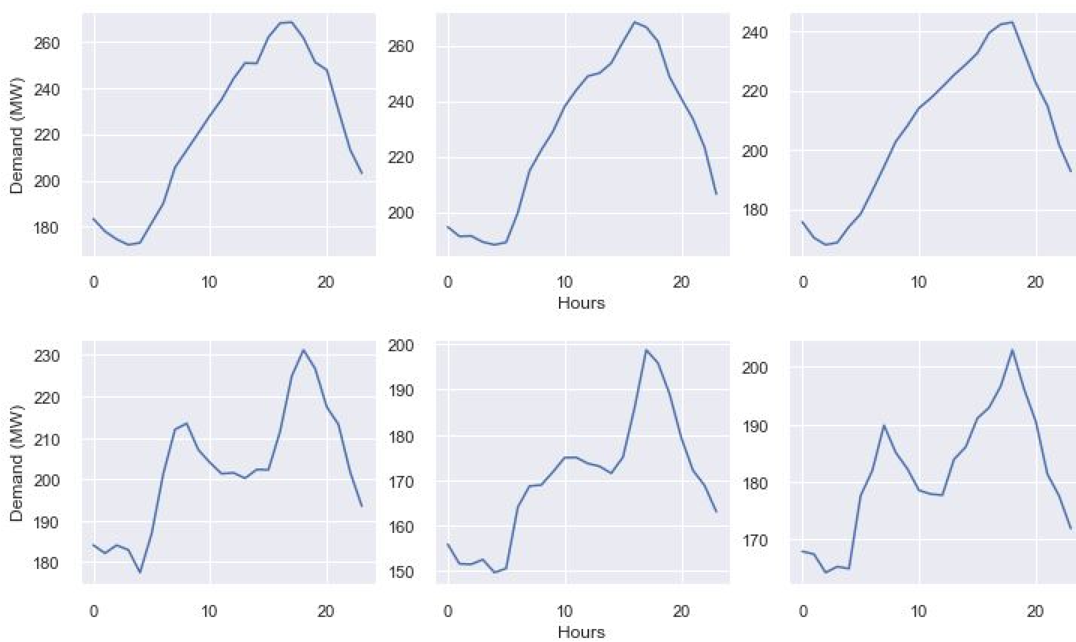


Figure 3.4: Scaled total demand for three daily profiles with a maximum demand of 300 MWs

Actual production levels

Figures 3.5 demonstrate the utilization of each of the 12 generators in an IEEE 24-BUS test system from three representative generators. The utilization starts from 0, with no use of the generator capacity, to 1, using the capacity at its maximum level. As can be seen, generators 6 to 12 are mostly dispatched and produced at their maximums, however, generators 1 to 5 are mostly idle or marginal. These states of the time-slots would form the constraints on the costs estimated for each generator.

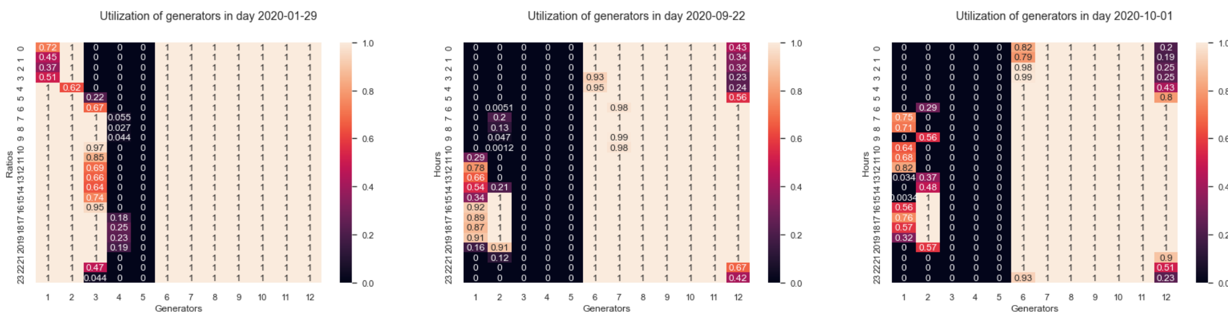


Figure 3.5: Percentage utilization of the maximum capacity of each of the 12 generators in 3 different representative days

Locational marginal prices

In order to reflect the operational dynamics of our test case, figures 4.3 are provided. These prices in the market at each hour and bus are locational marginal prices driven directly by the EDP (Economic dispatching problem) dual variables associated with the load-balance constraints at each bus. As seen during the early morning and late nights, the prices across the network fall because of the overall decrease in total demand and rise during the day as the demand increases. In some time slots, we see spikes and falls in prices, mainly because of the congestion in transmission lines.

Figure 3.6: IEEE distribution of prices in 3 different sample representative days

IO two-step algorithmic framework results

This section employs the proposed two-step algorithmic framework in the IEEE 24-Bus test system. The result of the first step is three sets of constraints. These constraints are used in the second step to find cost function parameters such that the output of the cost function would respect them. The second-step model places constraints on how closely the predicted cost resembles the actual market price as more equality constraints are added to the equation. More inequality constraints also lead to anticipated costs deviating from the base market price and rising or falling below it. Figure 3.7 represents the number of constraints associated with each of the discussed sets for each generator. As can be seen, generators 6 to 12 have a majority of lesser than constraints, while in contrast, generators 3, 4, and 5 have a majority of Greater than constraints. Given this distribution of constraints, it can be expected that a generator 8, at most hours, would be producing at its maximum. Likewise, it can be expected that generator 5 would be Off and inactive almost all the time.

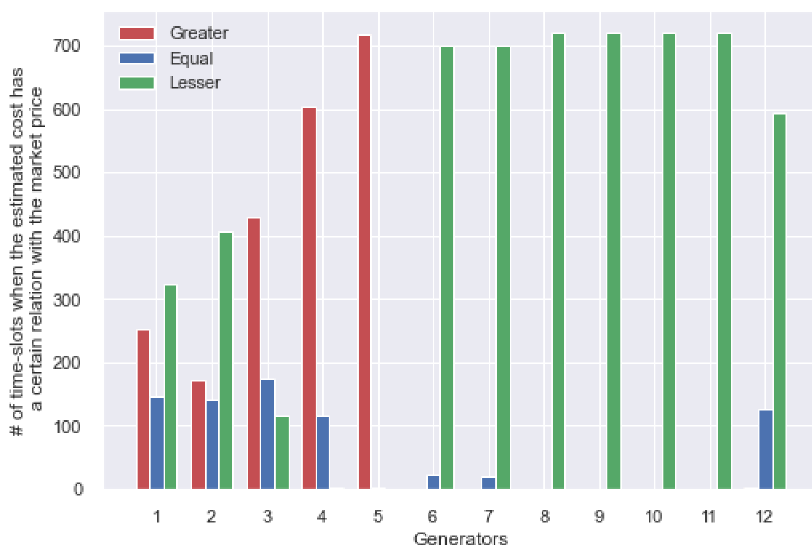


Figure 3.7: Distribution of time-slots in terms of their relationship with the market price

Comparison of market outcomes

Figures 3.8 demonstrate a comparison of the market outcomes of the IEEE model and its simplified rebuilt model with estimated costs. These example results are estimated using in-sample data. A rebuilt model is a market-clearing problem based on observable information and assumptions. This model is used for validation to see whether the results of this assumed market problem would match the real-world market outputs. Here, While there are visible differences in times, we can find an exact match between the results in most examples. This shows the inverse optimization approach recovered the most fundamental basics of operational behaviors. However, this does not mean that all confidential information is retrieved, as there are also some mismatches. Model bias and unobservable market details, such as transmission conditions and contracts, could be the main reasons for a mismatch. For example, for generator 11, although it is expected that the market outcome of the rebuilt model matches the reality, there are some apparent differences. The reason is mismatching in other generators cause over-generation and under-generation, and as a result, this generator should compensate for these generations, which means even more mismatches.

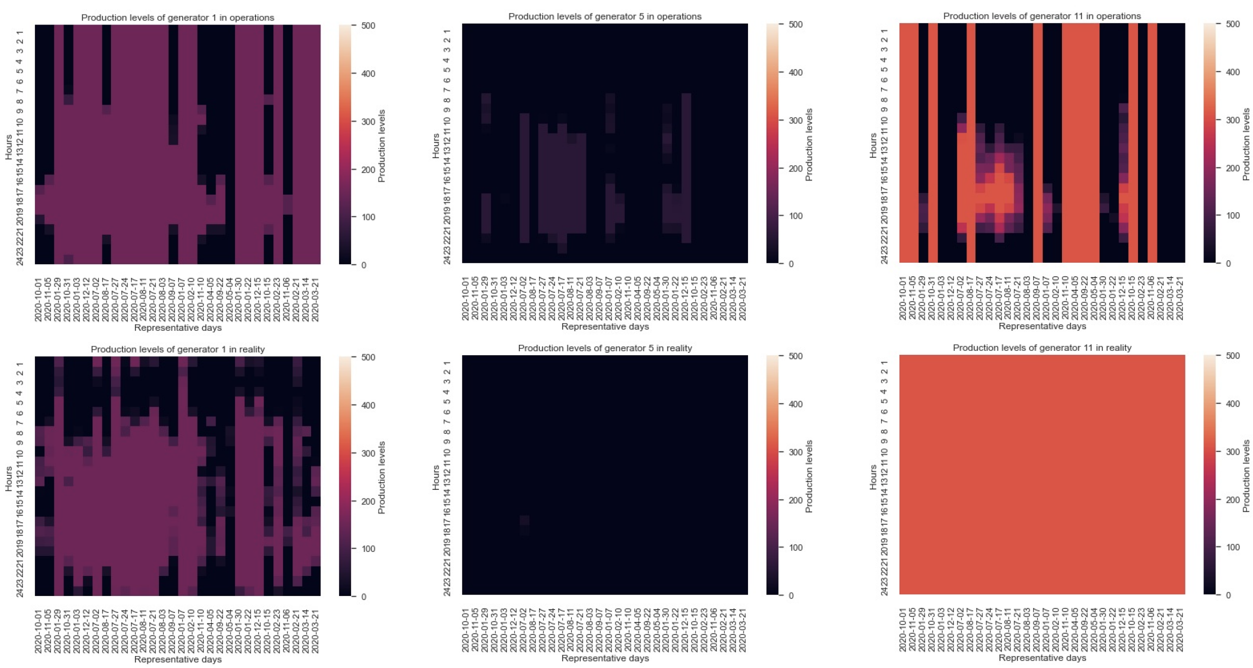


Figure 3.8: Comparison between rebuilt model and actual generation levels in different generators

Evaluation

In this section, numerical metrics are identified to assess the two-step IO approach's effectiveness in simulating market participants' operational behavior. Here, the accumulated absolute hourly error is used as the metric. The proportion of the error about the original production is measured using an MAE percentage, calculated to show a comparative meaning to comprehend the dimensions of each of the MAE values.

Table 3.3 presents an hourly calculated measure of fit between the results of the rebuilt model using IO-approximated cost functions and real-world data. Most generators with a mix of the three categories of constraints have an error rate ranging from 16 percent to 36 percent. The generators 6 to 12, with a majority of lesser than constraints, have an error range of 0 to as high as 66 percent of the initial capacity. In particular, generators 10 and 11 have the highest error rate among all of the generators. When generators have many equality constraints, we have more precise information regarding the cost function, otherwise just a couple of bounds.

Table 3.3: Error metrics for inverse optimization application to IEEE 24-bus test system

Generator	Accumulated absolute hourly error	MAE	MAE percentage
1	25427.505	35.315	23.2%
2	17507.627	24.316	15.9%
3	90688.078	125.955	35.9%
4	134152.959	186.323	31.5%
5	6623.491	9.199	15.3%
6	369.098	0.512	0.3%
7	3317.463	4.607	2.9%
8	0	0	0%
9	0	0	0%
10	142061.555	197.307	65.7%
11	124886.333	173.453	55.9%
12	6431.083	8.932	2.5%

3.3.2 Application of IO Algorithm to IESO

This section employs the proposed two-stage IO-based algorithm for a case study of the Ontario energy market, IESO. The IESO is a wholesale market serving the province of

Ontario and publishes its market outcomes regularly. These market outcomes include the generation levels of each generator and market prices. In addition to market outcomes, some technical features of the market participants can be indirectly calculated using the published market outcomes. These technical features include ramping capabilities and the minimum and maximum capacity of the generators. This section presents these market features and data for a subset of the generators. The two-stage algorithm is applied to this problem. Finally, a market-clearing problem is rebuilt using the cost approximations in the previous step, and the results (generation levels) are compared with real-world data.

IESO market characteristics & data

IESO publishes hourly data on the generation levels for each of the 112 generators spread across Ontario. In addition, the uniform market price is available on the same granularity level, hourly. The market outcomes are essential to the proposed two-stage algorithmic framework in that it not only gives the problem the optimal or close to optimal generation levels but also gives an indirect perspective of the technical capabilities. For example, considering the generation levels of a particular generator throughout its lifespan, maximum and minimum levels of generation can be estimated. Likewise, considering the changes in the generation levels of these generators would allow the calculation of an upper bound and a lower bound on the ramping abilities of these generators. Here, a snapshot of the data is provided. Instead of representing all days, a small subset of days and generators are selected for representation purposes.

Capabilities of the generators in the Ontario energy market include the maximum generation levels, and the maximum and minimum ramp-up and ramp-down rates. These features of our model are determined using publicly available data.

Representative generator	Capability	Cluster weight
PICKERINGB-G8	523	3
BRUCEB-G5	873	7
PICKERINGA-G1	520	1
PICKERINGB-G6	531	4
BRUCEA-G4	839	1
DARLINGTON-G2	913	1
BRUCEB-G6	810	1

Table 3.4: Capability and weight of each of the representative generators of nuclear type

Representative generator	Capability	Clustering weight
HALTONHILLS-LT.G3	304	7
DPNTMTLND	49	34
GREENFIELD ENERGY CENTRE-G4	492	1
STCLAIRCGS	585	1
NAPANEE-G3	447	3
GREENFIELD ENERGY CENTRE-G1	195	5
TASARNIA	409	1

Table 3.5: Capability and weight of each of the representative generators of gas type

Representative generator	Capability	Clustering weight
DESJOACHIMS	404	2
REDROCK	37	47
SAUNDERS	1002	1
CHENAUX	128	5
BECK2	1438	1

Table 3.6: Capability and weight of each of the representative generators of hydro type

Tables 3.4, 3.5, and 3.6, provide a subset of nuclear, gas, and hydro generators selected using a clustering method, namely k-means, and their capability.

Likewise, ramping rates stated in the tables 3.7, 3.8, and 3.9 are driven by the market outcomes in the Ontario energy market. These rates are calculated by measuring each generator's maximum increase or decrease in generation levels.

Nuclear representative generators	Maximum ramp-up rate	Maximum ramp-down rate
BRUCEA-G4	135	151
BRUCEB-G5	127	154
BRUCEB-G6	3	5
DARLINGTON-G2	10	9
PICKERINGA-G1	13	8
PICKERINGB-G6	12	4
PICKERINGB-G8	34	10

Table 3.7: Ramping rates for each representative generator selected from the nuclear generators.

Gas representative generators	Maximum ramp-up rate	Maximum ramp-down rate
DPNTMTLND	32	36
GREENFIELD ENERGY CENTRE-G1	87	119
GREENFIELD ENERGY CENTRE-G4	165	196
HALTONHILLS-LT.G3	91	154
NAPANEE-G3	159	252
STCLAIRCGS	351	315
TASARNIA	142	140

Table 3.8: Ramping rates for each representative generator selected from the gas generators.

Hydro representative generators	Maximum ramp-up rate	Maximum ramp-down rate
BECK2	295	358
CHENAUX	56	91
DESJOACHIMS	143	145
REDROCK	19	19
SAUNDERS	216	208

Table 3.9: Ramping rates for each representative generator selected from the hydro generators.

Application of the proposed two-stage algorithmic framework

Here, the proposed two-stage IO-based algorithmic framework is employed for the Ontario energy market case study. In the first stage of the algorithm, three categories of economic constraints are derived. Figure 3.9 demonstrates the distribution of each of these three categories in a subset of 112 generators existing in the Ontario energy market. As can be seen, most generators have a majority of constraints belonging to the equality set. This is mainly because limits on the generation of these generators are calculated in a data-driven approach. As a result, the higher limits could be higher than the limits in reality.

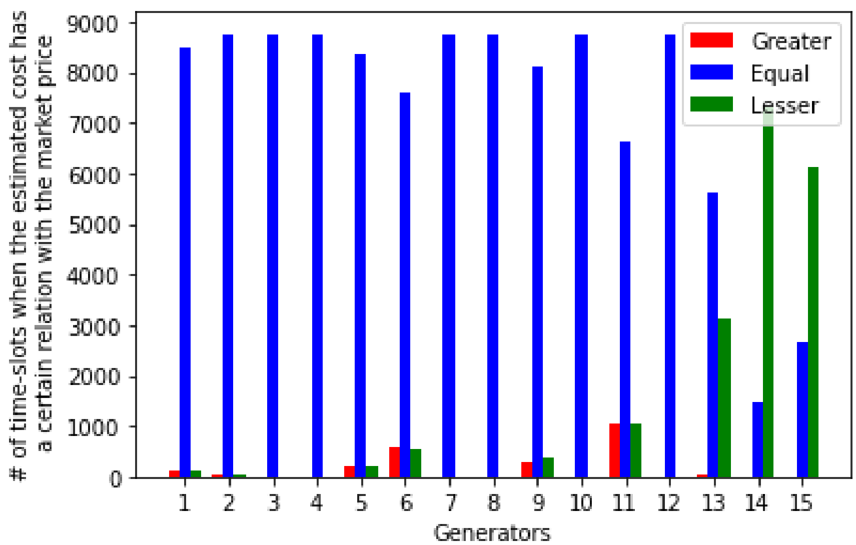


Figure 3.9: Distribution of time-slots for each of the generators

One of the primary reasons a novel two-stage inverse optimization algorithm is proposed is to ensure that when operational models are run based on market characteristics, the market results match, correlate, and eventually roughly align with real-world market behavior. Figures 3.10, 3.11, and 3.12 demonstrate three different days of generation levels comparison between the IO model generation and actual generation of gas-fired units, hydro units, and biofuel units. These example results are estimated using in-sample data. As can be seen, there is a close match on most days between Gas and Hydro generators. However, biofuel generators do not match the generations in reality. These mismatches in the biofuel sector are negligible compared to the extent of generation in Gas and Hydro and, as a result, can be omitted. Nevertheless, model bias and unobservable market details such as transmission constraints and contracts could be the main reasons if a mismatch occurs.

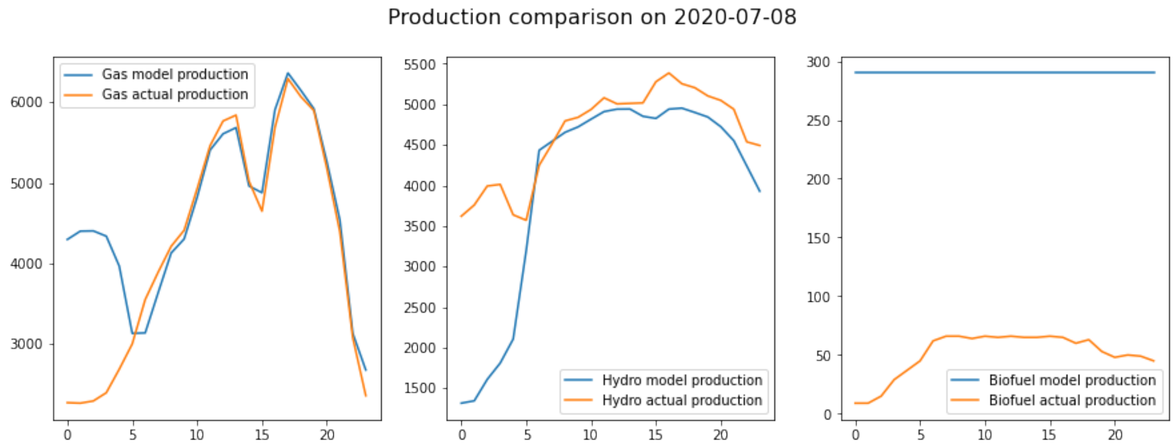


Figure 3.10: Production comparison between IO generated market outcome and real-world Ontario energy market outcomes on 2020-07-08

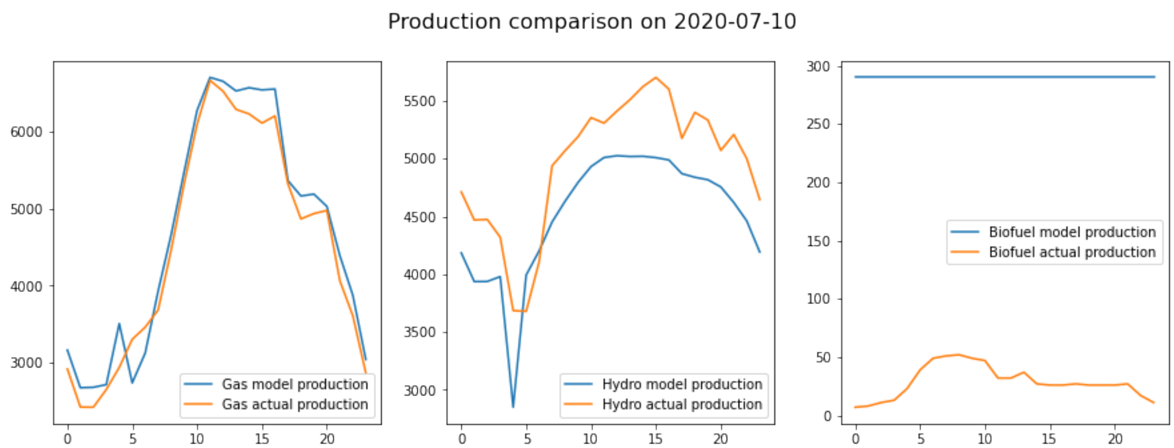


Figure 3.11: Production comparison between IO generated market outcome and real-world Ontario energy market outcomes on 2020-07-10

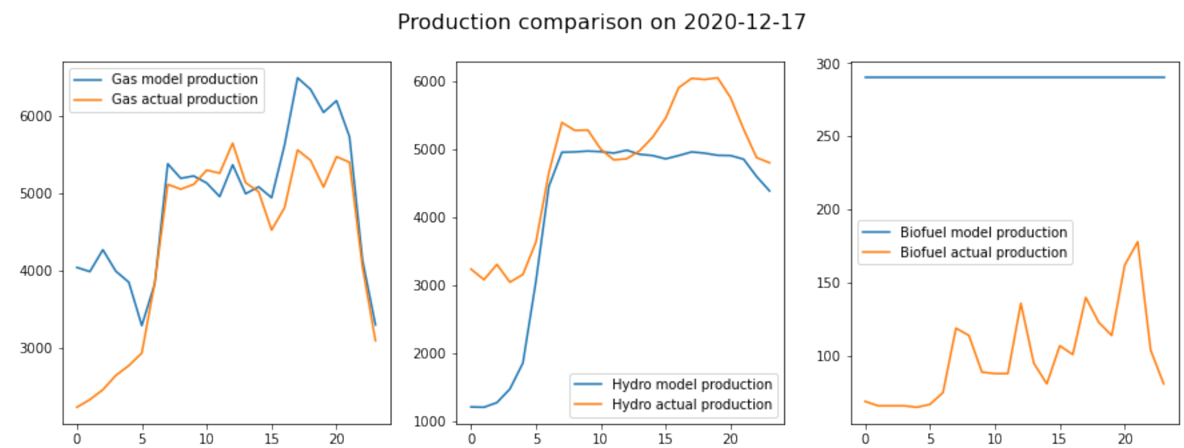


Figure 3.12: Production comparison between IO generated market outcome and real-world Ontario energy market outcomes on 2020-12-17

Evaluation

The three categories of Gas, Hydro, and Biofuel were the focus of the IO method employed here because renewable resources like wind, solar, and nuclear energy are not as price responsive as other generating technologies. The proportion of the error about the original production is measured using an MAE percentage, calculated to show a comparative meaning to comprehend the dimensions of each of the MAE values.

For gas, the MAE percentage is 24.27 percent of the original capacity, while the MAE for hydro generators is higher and 32.17 percent. This may be for a number of reasons, for example, hydro generators have other non-price considerations such as not flooding small villages that could affect the results and are not modeled in this study. Likewise, biofuel generators generally have stable levels of generation which mean lower levels in comparison to their capacity.

Table 3.10: Error metrics for inverse optimization application to Ontario energy market

Technology	Total absolute error	MAE	MAE percentage
Gas	16154269.03	1844.09	24.27%
Hydro	15860452.65	1810.55	32.17%
Biofuel	322007.28	36.75	12.67%

Chapter 4

Inverse Optimization for Long-term Planning Models

As shown in Chapter 2, long-term planning models are used to prescribe the optimal generation mix to satisfy demand while minimizing cost or maximizing social welfare, and meeting the technical and regulatory constraints. These long-term planning models could be used by system operators such as IESO. There are many different cost components including but not limited to operational, and investment costs that are involved in such models. Since the estimation of cost parameters in these models is a complex task, in this chapter an IO-based method is proposed to both simplify the models and reach plausible outcomes. This problem can be used by independent energy system operators to reduce the original size of their problems and propose an optimal generation mix.

This chapter is outlined as follows: first, the developed methodology is described including the modifications to the long-term planning models based on the IO approach provided in Chapter 2, Chapter 3, and methods for model reduction, and then numerical examples using data from Ontario is provided. The IO-based model is further validated by comparing results with a similar model with the given cost parameters from other sources.

4.1 Methodology

The provided methodology in this section, which is to estimate the cost parameters using inverse optimization is based on Chapter 2 where the formula of long-term planning is provided and Chapter 3 where the approximations of IO-based model is provided.

In addition, to reduce the computational complexity of the long-term planning model, we employ a clustering method, namely k-means, to find representative days d and generators g. These model reductions, help have a more computationally tractable model and pave the way for more scenario analysis in the future.

4.1.1 Applications of inverse optimization to long-term planning models

The long-term planning model is presented in Chapter 2, equations 2.1 to 2.22d. In order to estimate the cost parameters $C_{i,t}$, in this section, a set of non-decreasing linear convex cost functions are provided. As a result, the proposed model is not different from the model introduced in Chapter 2 except for the cost of operations, where costs are replaced by functions in the form of $f(\beta, p_{i,g,t,d,h}, L_{t,d,h}) = \beta_0^{i,g} + p_{i,g,t,d,h}\beta_1^{i,g} + \sum_{i \in I_r} p_{i,g,t,d,h}\beta_2^{i,g} + L_{t,d,h}\beta_3^{i,g}$. These functions are estimated using a two-step algorithmic framework and discussed in more details in Chapter 3. In the first step, three categories of daily time slots are derived based on econometric concepts and inverse optimization. In the second step, a model based on regression analysis is provided to estimate cost parameters that group daily time-slots into the three categories derived in the previous step. These functions and their parameters β are estimated for each generator g of each technology i. The summary of these parameters is provided in Table 4.1.

Table 4.1: Summary of parameter estimations using a novel inverse optimization algorithmic framework

Technology	# generators	β_0	β_1	β_2	β_3
Gas	52	-37617.6 (3075.6)	555.01 (1170.909)	3.904 (0.456)	9.362 (0.381)
Hydro	56	-53358.1 (16151.24)	845.08 (998.401)	4.349 (0.667)	9.481 (0.952)
Biofuel	4	-14.09 (0.263798)	0.169 (0.142)	0.0037 (0.00033)	-1E-08 (2.51E-15)

4.1.2 Methods for model reduction

Long-term planning models are computationally costly due to a large number of variables and parameters defined over a long planning horizon. For example, when presenting demand in each hour, through a 20-year planning horizon, there is a need for $24 \times 365 \times 20$ demand parameters. Therefore, a wide variety of methods to reduce the size of the planning problems while ensuring the reliability and accuracy of the power systems are maintained.

One common approach is to select the representative days or timeslices representing an electricity grid's behavior, instead of including all of the details over a long time. These representative time periods can be obtained using various methods including clustering algorithms in which similar entities are grouped together.

While every algorithm has its own strengths and weaknesses, the k-means algorithm is among the most popular clustering methods which are widely employed in the literature. We assume that in this section, different items we need to group are provided as vectors.

In the k-means clustering algorithm, each item in a data set is assigned a cluster with a centroid that is closer to that particular point than all other cluster centroids [44]. One of the most significant advantages of the k-means algorithm over other algorithms is its speed. On the other hand, one of the downsides of the k-means clustering algorithm is that it requires that the data would be in similarly sized hyper-spherical clusters [35]. In this study, the most commonly used algorithm k-means approach [36], is employed by providing two entries as inputs, a set of points or vectors, and a number that specifies the number of clusters, k. In the second stage, the algorithm is initialized by setting the iteration counter to zero and then randomly assigning each point to one of two starting clusters. In general, the clustering algorithm's final result is heavily reliant on the initial cluster assignment.

Using these assignments, the next step is to iteratively find new assignments by recalculating the centroid of each of the clusters. After the centroids are found, the closest centroid to each point is the primary determinant of which cluster that point will belong to. In this step, even though many distance metrics could be employed, the Euclidean distance $d(x, x^{'})$ would be used as in 4.1. This iterative procedure is continued as long as reassignments are not necessary between two rounds.

One common method to find the k, number of clusters, in the k-means algorithm is the elbow method. This method visualizes the cluster solutions by plotting the sum of squared errors (SSE) or inertia for different values of k and determining the "elbow" or "inflection point". The inflection point is the point after which further increase in k produces a negligible decrease in SSE. This method helps to choose the right value of k, which is the number of clusters that best suits the data. Due to its simplicity, the elbow method is the most commonly used method for finding the optimal value of k for k-means clustering.

$$d(x, x^{'}) = \sqrt{\sum_{j=1}^{n}(x_j - x_j^{'})} \tag{4.1}$$

Representative day selection

There are two requirements when selecting representative days. First, the employed method should be able to generate representative days quickly, and second, they should respect all of the crucial correlations in the data. In order to respect related correlations, the electricity demand is clustered and renewable availability is represented in the capacity-expansion model.

When choosing representative days, there are two criteria to consider. First, the technique we use should be able to create these representative days quickly, and second, they should also respect all of the data's critical relationships. To account for these relationships, renewable availability and demand are clustered in a group. First, vectors of demand are defined in day d, V_d^D:

$$V_d^D = (V_{d,1}^D, V_{d,2}^D, V_{d,3}^D, ..., V_{d,24}^D), \quad \forall d = 1, 2, 3, ...365 \tag{4.2}$$

Similarly, vectors of wind generation in day d, V_d^W are defined as follows:

$$V_d^W = (V_{d,1}^W, V_{d,2}^W, V_{d,3}^W, ..., V_{d,24}^W), \quad \forall d = 1, 2, 3, ...365, \tag{4.3}$$

And finally, the vectors of solar generation, V_d^S are defined as follows:

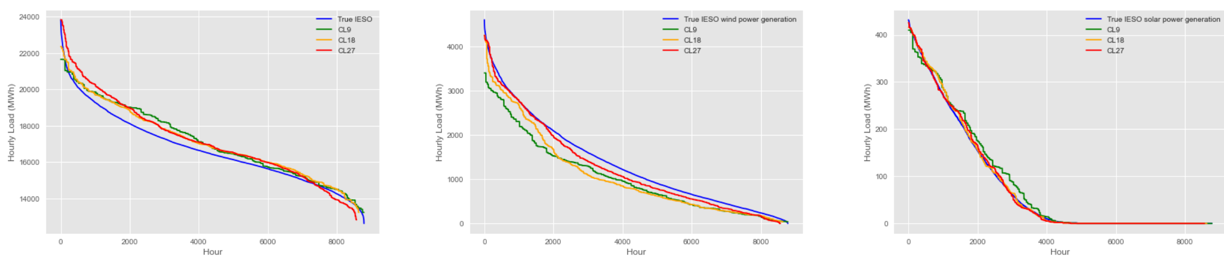
$$V_d^S = (V_{d,1}^S, V_{d,2}^S, V_{d,3}^S, ..., V_{d,24}^S), \quad \forall d = 1, 2, 3, ...365, \tag{4.4}$$

These vectors are ordered by the hour of the day in order not to lose the temporal sequence of observations. Vectors V_d are also defined as the vector of day d short-term operational details data:

$$V_d = (V_d^D, V_d^W, V_d^S), \forall d = 1, 2, 3, ...365, \tag{4.5}$$

Finally, V_d has all the information related to the operational conditions on which the clustered days are calculated. As a result, clustering based on the collection of points $V_1, ..., V_{365}$ would result in a set of days ensuring to capture of a wide variety of operating conditions. To respect the intraday serial correlation between the operational details, the temporal sequence of the operational data should be preserved. In the proposed approach, the chosen representative days would completely respect these correlations.

Figure 4.1a illustrates the load duration curve of clustered and unclustered data. Here, a $k = 10$ is considered for the purpose of clustering. The number of clusters, k, is calculated through an elbow method. Now, a different number of samples could be selected from each of these clusters. The figure shows mismatches for each number of representative days that are selected. Likewise, figures 4.1b and 4.1c show the wind and solar duration curve. Significant matches in solar-duration curves, especially when the number of representative days is close to one month can be noticed. In general, overestimations in total load in Ontario, underestimations for wind generators, and suitable matches in solar duration-curve are observed.



(a) Total load (b) Wind (c) Solar

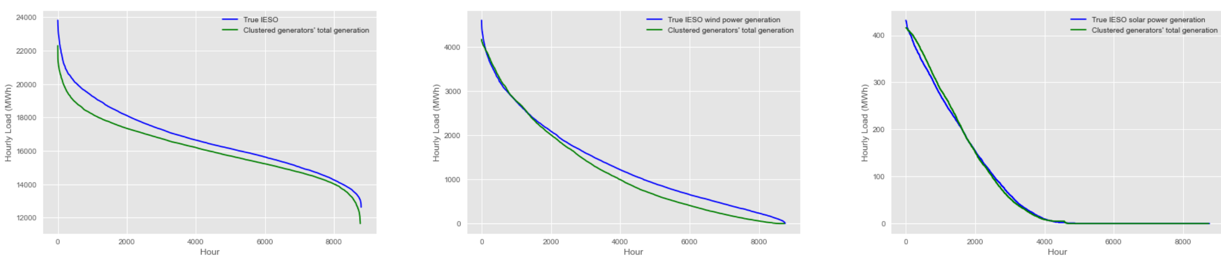
Figure 4.1: Load curve comparison of different numbers of representative days

Representative generator selection

Generators participating in an electricity market could have various capacities and technical details. They are in different locations, and each has specific geographical features. Each province's area, such as Ontario, has different levels of availability for wind, solar, and hydro, making some of the generators more valuable to the grid. Refraining from considering these operational details when investing would mean our investment decisions did not consider geographical and technical constraints that exist in reality. While there is a trade-off here regarding the number of generators selected to represent technology, considering all of them would also be trivial because some generators are technically and geographically similar. On the other hand, considering more representative generators for a particular technology would mean many more constraints are needed to be considered and, as a result, a more complex cost function to be estimated with many more break points. This would result in a function that considers more operational scenarios and, as a result, is less prone to noise and outliers.

k-means algorithm is employed for the purpose of this section. The k is estimated using an elbow method. This algorithm is explained in previous sections and the results are also presented when the operational models in 3 are studied.

To show the effectiveness of this clustering approach to model the specifics of the market load, here, load duration curves for renewable resources and total load are provided. Figure 4.2a provides a comparison between clustered generators' total generation and actual IESO data set. As can be seen, the clustered set of generators tends to slightly underestimate loads. Further, figures 4.2b and 4.2c demonstrate the effectiveness of clustered generators for respectively solar and wind resources. As can be seen, in contrast to the total load, wind and solar generation have been closely matched by clustered generators.



(a) Total load (b) Wind (c) Solar

Figure 4.2: Comparison of load duration curve in Ontario generation and generation built based on representative generators

4.2 Numerical results

This section provides the results of a generation expansion planning model for the Ontario energy market by first using fixed-cost parameters affecting investment decisions over time and then providing the results of the planning model with estimated cost functions obtained in Section 4.1.1.

4.2.1 Model parameters and inputs

A long-term expansion planning model is covered in section 2.2. This model featured a large number of decision variables that would be obtained by the solution of the optimization problem, as well as parameters and constants that must be included. An algorithmic framework for recovering estimated cost functions is covered in Chapter 3. On the other hand, other model parameters can be derived from yearly reports and a literature study or analysis of market outcomes.

This section discusses model inputs such as technical parameters, investment costs, capacity factors, and long-term parameters and assumptions.

Generator technical parameters

Most operating technical parameters, such as maximum capability, ramp-up, and ramp-down rates, are already provided in the previous chapter. Tables 3.4, 3.5, 3.6 represent the capability of each of the representative generators, and 3.7, 3.8, and 3.9 represent the ramp-up and ramp-down rate capabilities.

Operating costs

Fixed operating costs shown in Table 4.2 provide the data that are used for showing the significance of IO-approximated cost functions.

Table 4.2: Variables cost of generation for each technology (\$/MWh)

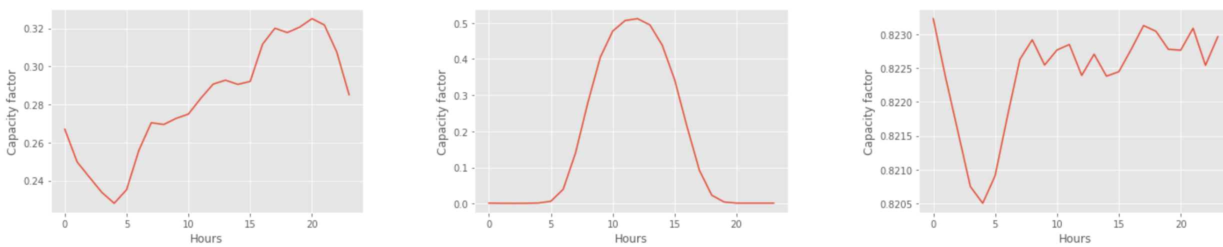
	Nuclear	Gas	Hydro	Biofuel	Solar	Wind
Operating cost	10.3	27.77	4.13	30.07	0	0

Capacity factors

The capacity factor measures a power generation unit's total usage. It is computed by dividing a power plant's yearly generation by the product of the nameplate capacity and the number of hours. In other words, the capacity factor evaluates a power plant's output in relation to its most significant output potential under ideal conditions. Because most power plants typically operate below their maximum capacity, a capacity factor may be considered a measure of how many hours the power plant worked in a period of time and at what percentage of its overall capability.

Figures 4.3 demonstrate the average capacity factor of three types of technologies in the Ontario energy market. As shown in the capacity factor of renewable generators such as solar and wind, they vary significantly throughout the day due to the availability of such resources, while nuclear is constantly available during the day.

These capacity factors play a significant role in the long-term planning model discussed before. They model the output of price-takers in the market using a number between zero and one. This number is then multiplied by the available capacity of that generator or technology to form the final generation.



(a) Wind (b) Solar (c) Nuclear

Figure 4.3: Average hourly capacity factor of Wind power plants in Ontario energy market

Long-term planning parameters and assumptions

Load: As predicted by IESO, Ontario's total load is expected to increase by 1.7 percent per year until 2042 [2]. As a result, Figure 4.4 shows how the total demand in Ontario would increase over the span of the next twenty years from 2021 to 2041.

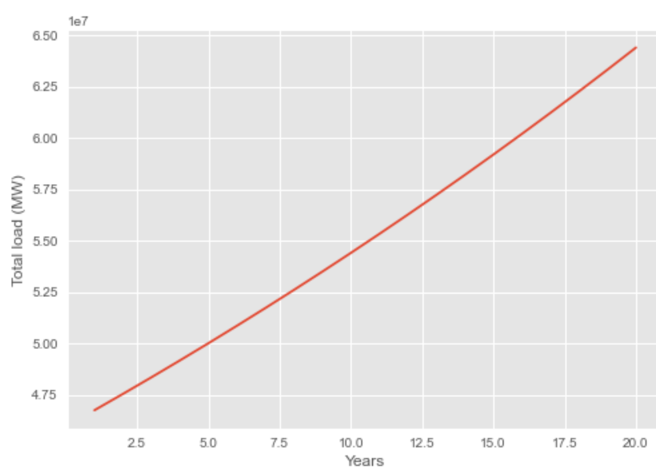


Figure 4.4: Total annual demand over the span of planning horizon

Expected lifetime of generation units: Each generator based on its type has an expected lifetime which plays an important role in the long-term planning models as it identifies how depreciation factors affect the available generation mix. Table 4.3 presents these parameters.

Table 4.3: The expected lifetime of generators of different technology types

Technology	Expected lifetime
Nuclear	40
Hydro	25
Gas	50
Biofuel	40
Solar	30
Wind	25

Capital costs: Investment costs on different technologies play a significant role in determining particular investments. These costs could change over time due to the advent of newer technologies or an update to their efficiency. On the other hand, operational costs affect the daily operations of generations within an electricity market. They could affect the investment costs if the operating expenses add up and justify the investment in other units.

Table 4.4: Capital investment costs of different market participants (1000$) driven from Electricity Annual Technology Baseline (ATB) Data published by NREL

Technology	**Capital investment cost per year per MW in the planning horizon**																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Nuclear	7334	7279	7257	7198	7044	7044	7007	6972	6934	6893	6846	6807	6755	6699	6653	6606	6556	6508	6463	6420
Gas	1049	1044	1042	1030	1025	1025	1021	1012	1008	1003	999	996	991	985	981	977	971	967	964	959
Hydro	4141	4141	4141	4141	4141	4141	4200	4259	4319	4378	4438	4497	4556	4616	4675	4735	4735	4735	4735	4735
Biofuel	4332.67	4319	4306	4278	4275	4275.1	4253	4232	4209	4185	4156	4133	4101	4068	4040	4012	3981	3953	3925	3900
Wind	1392	1348	1303	1259	1215	1171	1127	1083	1038	994	950	941	931	922	912	903	893	884	874	865
Solar	1377	1366	1256	1196	1136	1076	1016	956	896	836	776	769	762	755	748	741	734	728	721	714
Battery	1851	1736	1620	1512	1409	1310	1261	1218	1180	1145	1112	1098	1084	1070	1057	1043	1029	1015	1001	987

4.2.2 Fixed operational cost approach

A fixed cost approach for identifying costs in the objective function of our long-term planning problem, is a benchmark for representing the significance of our IO method. A long-term planning model with fixed operational costs uses a data similar to what is provided in Table 4.2. These costs are inflated throughout the planning horizon and remain fixed for all generators of a specific technology. That means two generators A and B of a technology type are equally dispatchable at a particular timeslot. The tie-breaker is only their generation and ramp-up or ramp-down capability. In addition, a fixed operating cost would mean the costs are not responsive to changes in the generation mix or the advent of new cheaper technologies and demand volatility. Further, these costs are generally offered competitively by market generators. In a fixed-cost approach, such competitive behavior is not expected.

Operational outcomes

Operational outcomes are an essential aspect of planning models, and significant literature has been developed on methods that try to incorporate them. This part discusses the day-to-day operational outcomes during the planning horizon. A reliable investment model would include market dynamics to propose wiser investment decisions, especially when more uncertainty is being introduced daily.

Figures 4.5, 4.6, and 4.7 demonstrate the generation mix in a day in the first year, after 10 years, and after 20 years. As can be seen, the extent of renewables and, as a result, the generation of gas-fired units increased over time. This is reasonable, as more participation in renewables means more uncertainly added to the grid and a more need for generators that can accommodate it. Nevertheless, comparing these results per technology with generation levels in the real world gives better comparison metrics. Figures 4.8, 4.9, and 4.10 provide a more detailed comparison of generation levels in the first year. As can be seen, there are significant gaps and mismatches between the generation levels in the real world and the model, resulting in a need for methods to narrow down these gaps.

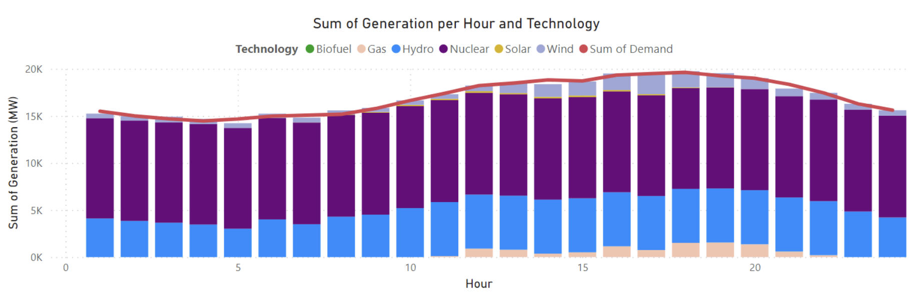


Figure 4.5: Generation by the hour and technology in a sample day in the first year of the planning horizon

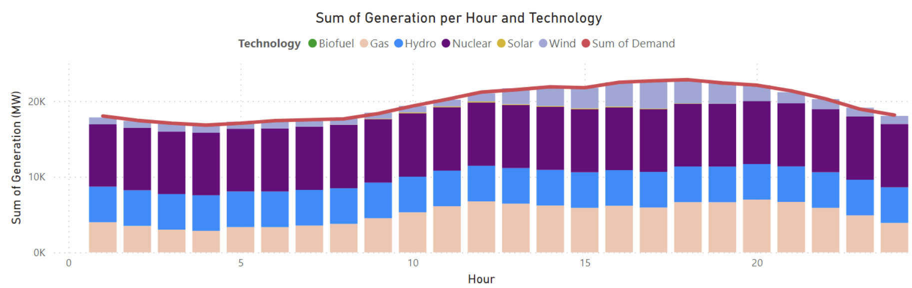


Figure 4.6: Generation by the hour and technology in a sample day in the tenth year of the planning horizon

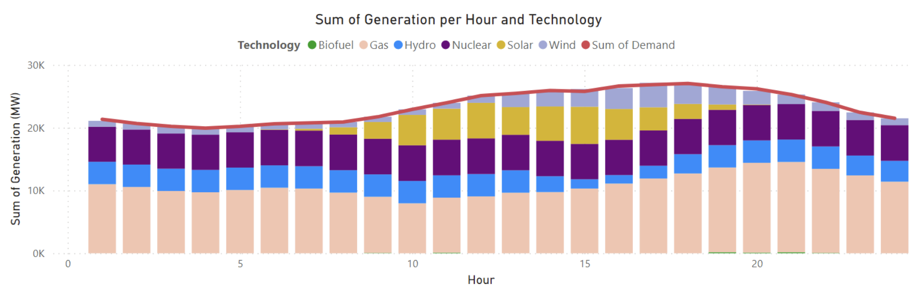


Figure 4.7: Generation by the hour and technology in a sample day in the last year of the planning horizon

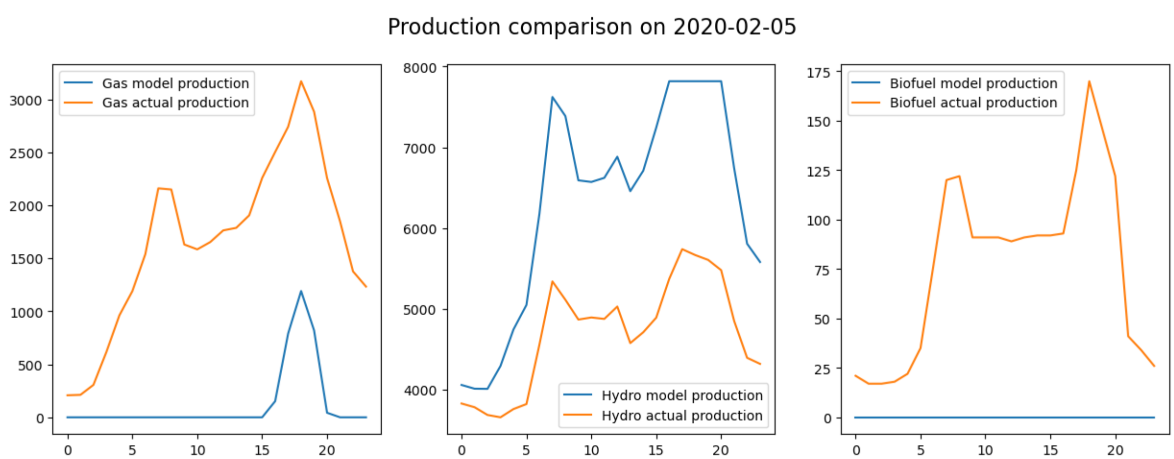


Figure 4.8: Comparison of each generator's aggregated generation on day 2020-02-05

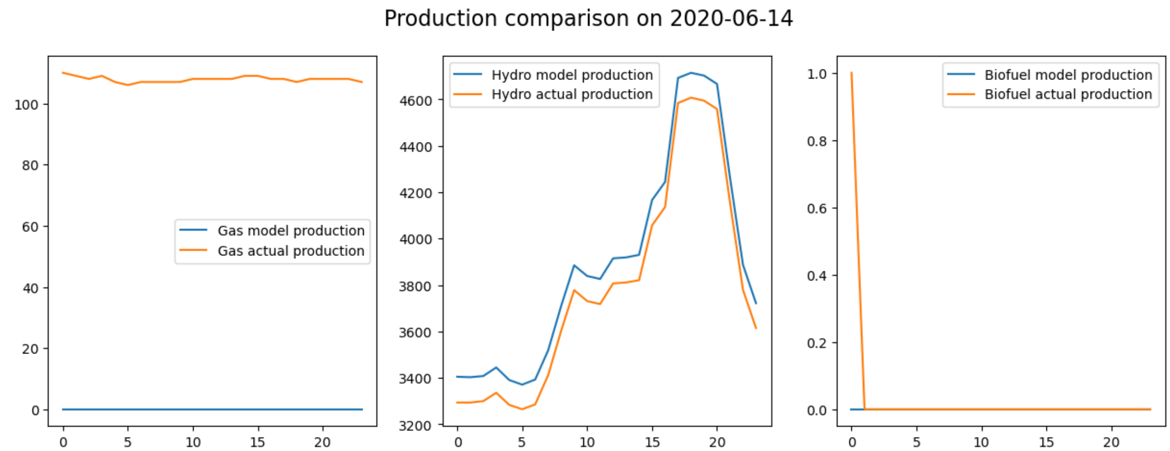


Figure 4.9: Comparison of each generator's aggregated generation on day 2020-06-14

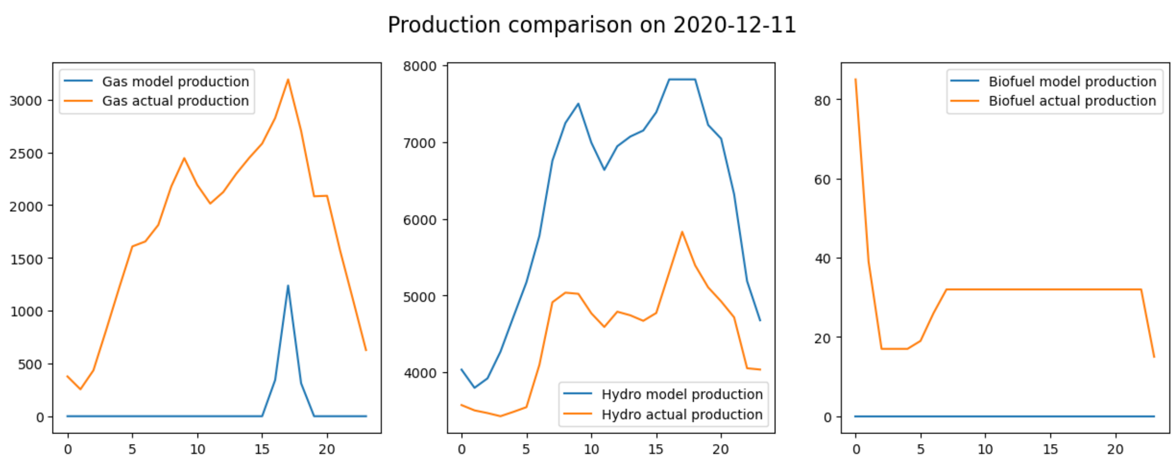


Figure 4.10: Comparison of each generator's aggregated generation on day 2020-12-11

Investments

Figures 4.11 and 4.12 provide a summary of the investments throughout the planning horizon and aggregated investments on each technology. One of the most prevalent points regarding these investments is that they are comparatively higher than the long-term planning model with extended operational costs. This over-investment or under-investment issue is because some scenarios are neglected or deemed less necessary.

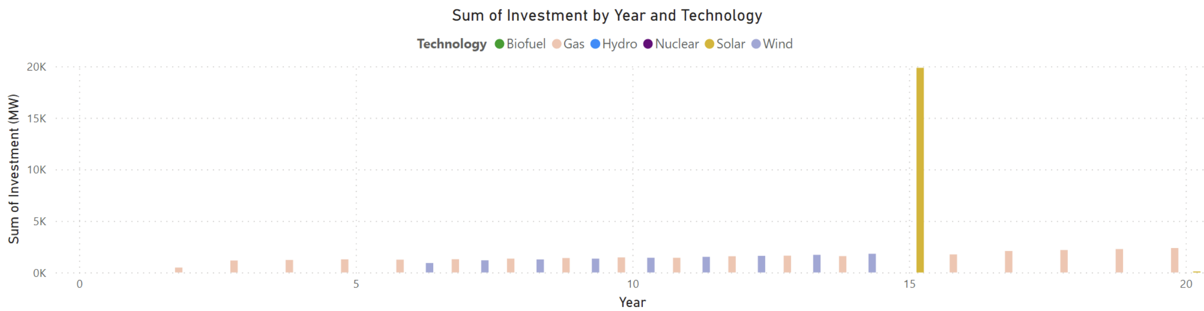


Figure 4.11: Investments in different technologies over the span of the planning horizon using a fixed-cost approach

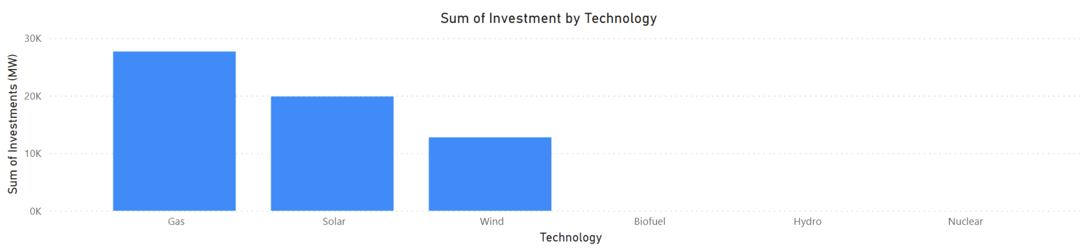


Figure 4.12: Aggregated investments in each technology

Generation mix

Figure 4.13 shows how the capacity of each of the technologies changes over the span of the planning horizon. Renewables tend to have more share in the generation mix as there are policy constraints in the long-term planning that penalize not meeting them. As a result of more renewable contributions to the generation mix, more flexible generators are needed which can be seen as well. Nevertheless, the maximum capacities throughout the planning horizon are not comparable with the long-term model with an IO approximated cost function as there are a lot more investments needed. This is one of the main reasons long-term planning models in the literature tend to consider more short-term operational details.

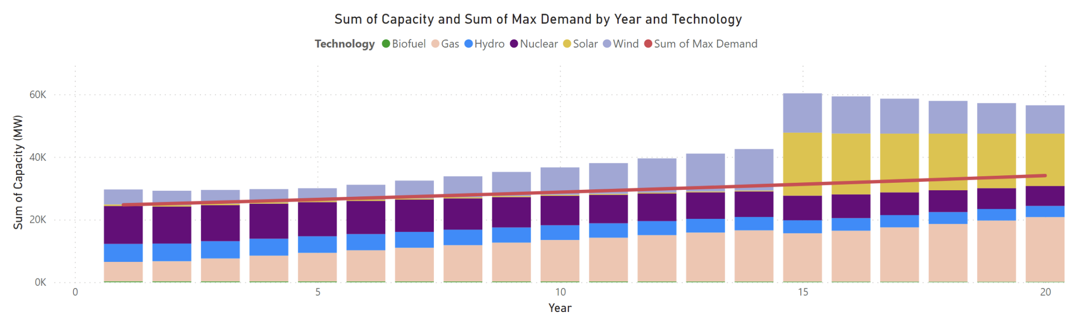


Figure 4.13: Capacity of different technologies in the planning horizon using fixed-cost approach

4.2.3 IO approximated operating cost approach

A long-term planning model with extended operational costs uses the data provided in Table 4.1. These costs are calculated by factoring in different inputs such as total load in the system, total renewable generation, and generation of the generator itself. These cost functions approximate the precedence of dispatching for all of the generators. All other features and parameters are the same as the fixed operational cost approach.

Operational outcomes

In order to check the reliability and performance of each of the market generators, here the aggregated operational results of the Ontario energy market are provided over the span of different years in the planning horizon. As the generation mix changes, the participation of each of the generators also changes. This is shown using figures that illustrate operations in the first year, after 10 years, and finally after 20 years. These figures are provided in 4.14, 4.15, and 4.16 respectively.

In the first year of operations, a large share of the nuclear technology and gas sector is noticed in the generation mix which stands first and second respectively. This generation mix changes as generation units depreciate and capacity expansion occurs in the market. After ten years, the gas sector takes a more significant percentage of the generation mix. This aligns with real-world scenarios that are foreseen for the Ontario energy market. After twenty years, and as more nuclear technology retires, a larger share of wind power plants and a significant increase in the participation of renewable in total are noticeable. This participation, in part, is due to an increase in the available capacity of flexible generators such as gas power plants.

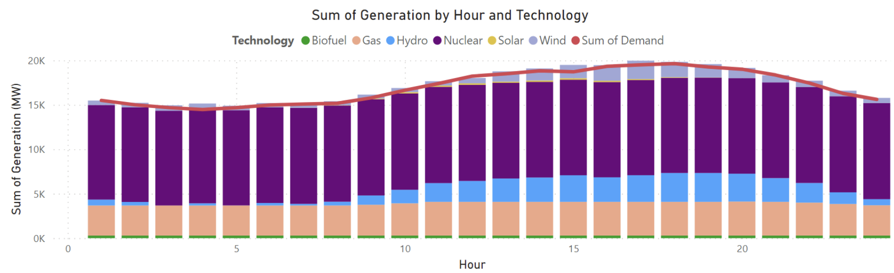


Figure 4.14: Generation by the hour and technology in a sample day in the first year of the planning horizon using an IO-based cost approximation approach

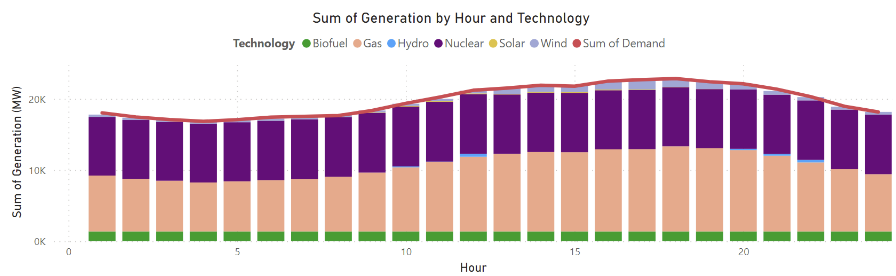


Figure 4.15: Generation by the hour and technology in a sample day in the tenth year of the planning horizon using an IO-based cost approximation approach

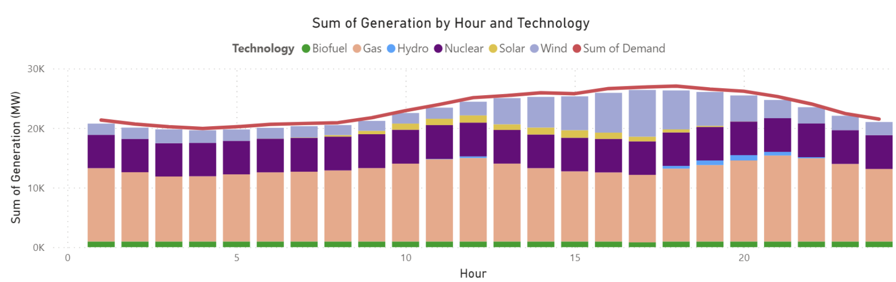


Figure 4.16: Generation by the hour and technology in a sample day in the last year of the planning horizon using an IO-based cost approximation approach

Investments

Figures 4.17, and 4.18 respectively provide a detailed distribution of investments throughout the planning horizon and an aggregated sum of these investments in each technology.

As can be seen, the IO estimated extended costs have helped the model have more investment diversity. In addition, having a distribution of representative generators would mean capacity factors of each of these generators, their technical capabilities, and their extended operational costs are considered in the investment decisions. Generators are generally built in geographical areas with specific features in favor of that technology. As a result, investment decisions neglecting these details would fail to consider these feasibility requirements.

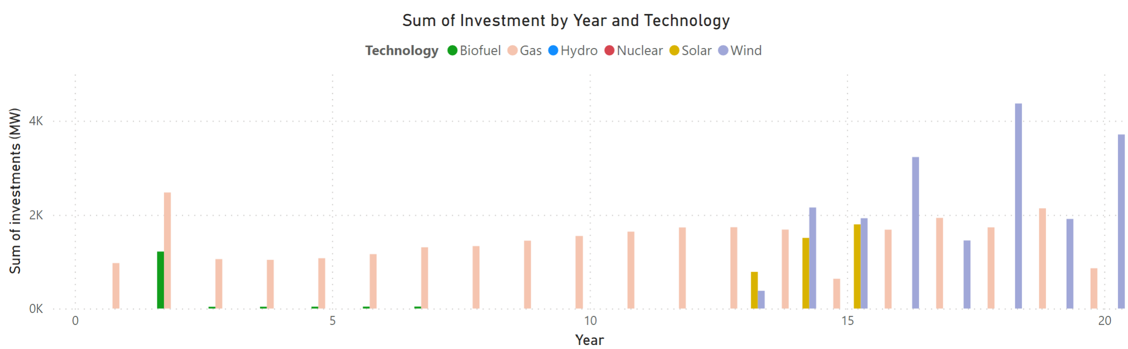


Figure 4.17: Investments in different technologies over the span of the planning horizon using an IO-based cost approximation approach

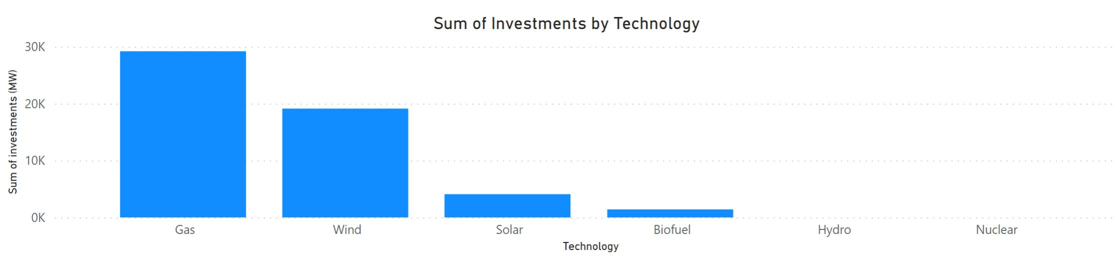


Figure 4.18: Aggregated investments in each technology

Generation mix

Generation mix changes over time as older technologies retire, new policies about the participation of different technologies are introduced, and demand increases over time. Figure 4.19 provides a brief overview of changes in the capacity of different technologies throughout the planning horizon. We see steady increases in the availability of gas-fired generators from the early years until the end of the planning horizon, as well as a steady decrease in the capacity of nuclear power plants due to their depreciation. Spikes in investment in renewables start mid-way in the planning horizon around 2030, with a higher priority given to wind power generators.

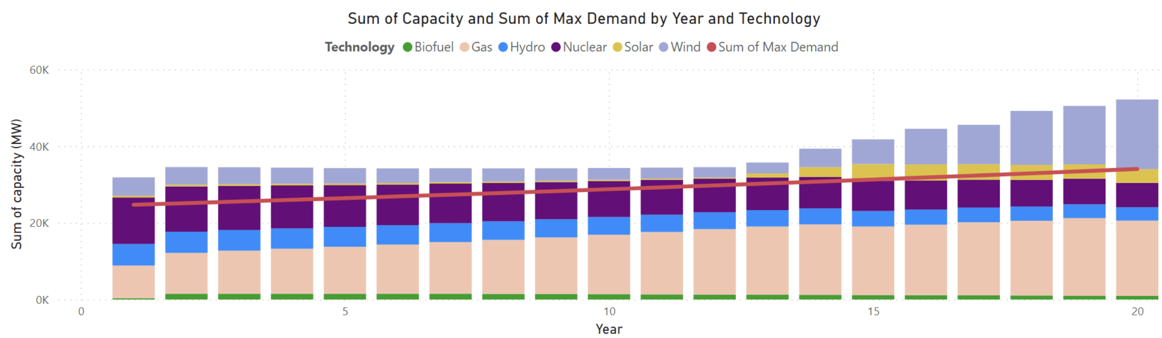


Figure 4.19: Capacity of each of the technologies in the planning horizon

4.2.4 Summary of the results

The two-stage algorithmic framework provided in this study, can be generalized to any of other markets with market clearing prices and partial information available. The results of the fixed-cost and IO-based approaches are provided in the previous sections. Besides the operational differences that these models have, and examples of these differences that are provided in different parts of this thesis, the investment scenario and the final generation mix of the two models have significant differences too. The first and most crucial difference is the total final capacity between the two models. The fixed-cost approach tends to over-invest in different technologies to meet the same operational and long-term constraints and policies. This is justifiable as the fixed-cost approach cannot follow the generation levels usually happening in the real world, thus requiring it to invest in more capacities as it tends to underestimate or overestimate the role of different generators. The other difference between the two approaches is the difference in the portfolio of investment. While the IO-based approach tends to invest more in Solar generators, the fixed-cost approach tends to invest more in Wind generators. Investments in the IO-based approach tend to be spread over time, while a spike in investment can be noticed in the fixed-cost approach.

4.3 Challenges and limitations

Like any other model, the model given in this study has inherent issues and limits. It does not include all generators from all technology categories but clusters them to discover a minimum collection. The model does not consider demand-side load increases, such as intelligent thermostats or electric vehicles. Representative days are selected using one of the clustering methods, and this does not mean that all scenarios including the extreme cases are also considered. Most long-term expansion planning model results significantly change based on the parameters that are included in the analysis, as such these parameters could be altered for further scenario analysis, and as a result, they are a major limitation.

The proposed inverse optimization two-step algorithmic framework also comes with certain limitations. First of all, these cost estimations are approximate and reconciliation of different ideal parameters for different time slices. There are overestimations and underestimations in generation levels of different generators in certain time slices because of these approximations. Any mismatch would cause a domino effect for other mismatches as the generation level of each of the generators complements each other. In addition, cost functions are reliant on the generation mix available in the market, and as a result, cost functions should be periodically estimated to have a better alignment with updated market realities.

Chapter 5

Conclusions

5.1 Summary

Electricity markets throughout the world are undergoing major generation mix transitions as new, cheaper technology is deployed, older generating units retire, and new carbon-emission-reduction measures are implemented. Long-term planning is required to extend or repair existing generation technologies or introduce new ones. As the market becomes more unpredictable due to rapid changes in customer behavior and the arrival of new inexpensive technologies, long-term growth planning models must consider operational instability within the market in their calculations. Prior research has demonstrated that incorporating this detailed information has benefits, including reduced over- and under-investment and the capacity to undertake trials and perform sensitivity analysis.

This research introduces a novel inverse optimization algorithmic framework for the first time to model these operational details without explicitly introducing computational complexity to the original long-term planning model. We compared our results with a fixed-cost benchmark which is commonly used in the literature.

5.2 Contributions

This thesis proposes a two-stage algorithmic framework based on inverse optimization and linear regression to approximate the cost functions of different generators in the market. This approach is considered a data-driven inverse optimization approach because unlike the classical inverse optimization approaches which impose the optimality of the solutions, the optimality conditions in data-driven approaches can be violated due the suboptimality of solutions or their noisiness. This method is applied to long-term expansion planning models for the first time, and the results have been analyzed. Its applicability in the operational models is studied with two use cases, namely IEEE 24 Bus test system and the Ontario energy market. Finally, the estimated non-decreasing convex cost functions are employed in the long-term planning models to analyze the generation mix in the planning horizon of the problem.

5.3 Future Work

The operational models used in this study primarily included linear terms. This ensured a more tractable inverse-optimization approach and long-term planning model. Studying operational models with integer and non-linear terms would be a topic for future studies.

The approximated cost functions estimated in this thesis only capture some of the operational features of the generators. The estimations of market outcomes in our validation step clearly show that errors in fitting the behavior of one generator could result in cascading errors in other generator behaviors. As such, the study of scenarios when inverse optimization can model the behavior of generators closely and when it cannot, would be a topic for future research. We have used only three of the most intuitive scenarios in this thesis for our research. However, these scenarios could be further expanded, and the results verified. On the other hand, the cost estimations are considered linear; in future research, this could be considered non-linear, and more data sources could be included in the process.

Another future direction could be estimating these cost approximations every year in the planning horizon based on the changes in the generation mix. The structure of the cost function we estimate in this thesis is highly dependent on the generators' capabilities and the total electricity demand. Hence, having a series of these functions every year in the planning horizon could provide an updated version of how generators respond to the prices in the market and, as a result, more reliable operational details.

Finally, the estimated costs can be driven using other IO methods in the literature and the results could be compared with our proposed approach.

References

[1] Annual Planning Outlook | Ontario's electricity system needs: 2023-2042. https://www.ieso.ca/en/Sector-Participants/Planning-and-Forecasting/Annual-Planning-Outlook. [Online; accessed 2022-11-21].

[2] The Future of Electricity Demand in Ontario. https://www.ieso.ca/en/Powering-Tomorrow/2021/The-Future-of-Electricity-Demand-in-Ontario. [Online; accessed 2022-11-29].

[3] World energy outlook 2021.

[4] Global Greenhouse Gas Emissions. https://www.epa.gov/climate-change, aug 12 2013. [Online; accessed 2022-11-16].

[5] Sezin Afşar, Luce Brotcorne, Patrice Marcotte, and Gilles Savard. Revenue optimization in energy networks involving self-scheduled demand and a smart grid. *Computers & Operations Research*, 134:105366, 2021.

[6] Ravindra K Ahuja and James B Orlin. Inverse optimization. *Operations Research*, 49(5):771–783, 2001.

[7] Ricardo Alvarez, Albert Moser, and Claudia Andrea Rahmann. Novel methodology for selecting representative operating points for the tnep. *IEEE Transactions on Power Systems*, 32(3):2234–2242, 2016.

[8] Anil Aswani, Zuo-Jun Shen, and Auyon Siddiq. Inverse optimization with noisy data. *Operations Research*, 66(3):870–892, 2018.

[9] L Baringo and AJ Conejo. Correlated wind-power production and electric load scenarios for investment decisions. *Applied energy*, 101:475–482, 2013.

[10] Luis Baringo and Antonio J Conejo. Risk-constrained multi-stage wind power investment. *IEEE Transactions on Power Systems*, 28(1):401–411, 2012.

[11] Dimitris Bertsimas, Vishal Gupta, and Ioannis Ch Paschalidis. Data-driven estimation in equilibrium using inverse optimization. *Mathematical Programming*, 153(2):595–633, 2015.

[12] Lori Bird, Jaquelin Cochran, and Xi Wang. Wind and solar energy curtailment: Experience and practices in the united states. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2014.

[13] John R Birge, Ali Hortaçsu, and J Michael Pavlin. Inverse optimization for the recovery of market structure from market outcomes: An application to the miso electricity market. *Operations Research*, 65(4):837–855, 2017.

[14] Audun Botterud, Marija D Ilic, and Ivar Wangensteen. Optimal investments in power generation under centralized and decentralized decision making. *IEEE Transactions on Power Systems*, 20(1):254–263, 2005.

[15] Didier Burton and Ph L Toint. On an instance of the inverse shortest paths problem. *Mathematical programming*, 53(1):45–61, 1992.

[16] Michael L. Bynum, Gabriel A. Hackebeil, William E. Hart, Carl D. Laird, Bethany L. Nicholson, John D. Siirola, Jean-Paul Watson, and David L. Woodruff. *Pyomo–optimization modeling in python*, volume 67. Springer Science & Business Media, third edition, 2021.

[17] Michael C. Caramanis, Richard D. Tabors, Kumar S. Nochur, and Fred C. Schweppe. The introduction of nondiispatchable technologies a decision variables in long-term generation expansion models. *IEEE Transactions on Power Apparatus and Systems*, PAS-101(8):2658–2667, 1982.

[18] Timothy CY Chan, Tim Craig, Taewoo Lee, and Michael B Sharpe. Generalized inverse multiobjective optimization with application to cancer therapy. *Operations Research*, 62(3):680–695, 2014.

[19] Timothy CY Chan and Neal Kaw. Inverse optimization for the recovery of constraint parameters. *European Journal of Operational Research*, 282(2):415–427, 2020.

[20] Timothy CY Chan, Taewoo Lee, and Daria Terekhov. Inverse optimization: Closed-form solutions, geometry, and goodness of fit. *Management Science*, 65(3):1115–1135, 2019.

[21] Timothy CY Chan, Rafid Mahmood, and Ian Yihang Zhu. Inverse optimization: Theory and applications. *arXiv preprint arXiv:2109.03920*, 2021.

[22] Angela S Chuang, Felix Wu, and Pravin Varaiya. A game-theoretic model for generation expansion planning: problem formulation and numerical comparisons. *IEEE transactions on power systems*, 16(4):885–891, 2001.

[23] Wesley Cole, Bethany Frew, Trieu Mai, Yinong Sun, John Bistline, Geoffrey Blanford, David Young, Cara Marcy, Chris Namovicz, Risa Edelman, et al. Variable renewable energy in long-term planning models: a multi-model perspective. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2017.

[24] Antonio J Conejo, Luis Baringo, S Jalal Kazempour, and Afzal S Siddiqui. Investment in electricity generation and transmission. *Cham Zug, Switzerland: Springer International Publishing*, 119, 2016.

[25] Antonio J Conejo, Miguel Carrión, Juan M Morales, et al. *Decision making under uncertainty in electricity markets*, volume 1. Springer, 2010.

[26] Anna Creti and Natalia Fabra. Capacity markets for electricity. 2003.

[27] Ricardo Fernández-Blanco, Juan Miguel Morales, Salvador Pineda, and Álvaro Porras. Inverse optimization with kernel regression: Application to the power forecasting and bidding of a fleet of electric vehicles. *Computers & Operations Research*, 134:105405, 2021.

[28] Kimia Ghobadi and Houra Mahmoudzadeh. Inferring linear feasible regions using inverse optimization. *European Journal of Operational Research*, 290(3):829–843, 2021.

[29] Jean-Michel Glachant, Paul L Joskow, and Michael G Pollitt. Introduction to the handbook on electricity markets. In *Handbook on Electricity Markets*, pages 1–11. Edward Elgar Publishing, 2021.

[30] Richard Green, Iain Staffell, and Nicholas Vasilakos. Divide and conquer k-means clustering of demand data allows rapid and accurate simulations of the british electricity system. *IEEE Transactions on Engineering Management*, 61(2):251–260, 2014.

[31] Cliff Grigg, Peter Wong, Paul Albrecht, Ron Allan, Murty Bhavaraju, Roy Billinton, Quan Chen, Clement Fong, Suheil Haddad, Sastry Kuruganty, et al. The ieee reliability test system-1996. a report prepared by the reliability test system task force of the application of probability methods subcommittee. *IEEE Transactions on power systems*, 14(3):1010–1020, 1999.

[32] Çiğdem Güler and Horst W Hamacher. Capacity inverse minimum cost flow problem. *Journal of Combinatorial Optimization*, 19(1):43–59, 2010.

[33] William E Hart, Jean-Paul Watson, and David L Woodruff. Pyomo: modeling and solving mathematical programs in python. *Mathematical Programming Computation*, 3(3):219–260, 2011.

[34] Garud Iyengar and Wanmo Kang. Inverse conic programming with applications. *Operations Research Letters*, 33(3):319–330, 2005.

[35] Anil K Jain, M Narasimha Murty, and Patrick J Flynn. Data clustering: a review. *ACM computing surveys (CSUR)*, 31(3):264–323, 1999.

[36] Richard A Johnson, Dean W Wichern, et al. Applied multivariate statistical analysis. *New Jersey*, 405, 1992.

[37] Pinar Kaymaz, Jorge Valenzuela, and Chan S Park. Transmission congestion and competition on power generation expansion. *IEEE Transactions on Power Systems*, 22(1):156–163, 2007.

[38] S Jalal Kazempour and Antonio J Conejo. Strategic generation investment under uncertainty via benders decomposition. *IEEE Transactions on Power Systems*, 27(1):424–432, 2011.

[39] S Jalal Kazempour, Antonio J Conejo, and Carlos Ruiz. Strategic generation investment using a complementarity approach. *IEEE transactions on power systems*, 26(2):940–948, 2010.

[40] Arezou Keshavarz, Yang Wang, and Stephen Boyd. Imputing a convex objective function. In *2011 IEEE international symposium on intelligent control*, pages 613–619. IEEE, 2011.

[41] Sheng Liu, Auyon Siddiq, and Jingwei Zhang. Planning bike lanes with data: Ridership, congestion, and path selection. *Congestion, and Path Selection (March 12, 2022)*, 2022.

[42] Yixian Liu, Ramteen Sioshansi, and Antonio J Conejo. Hierarchical clustering to find representative operating periods for capacity-expansion modeling. *IEEE Transactions on Power Systems*, 33(3):3029–3039, 2017.

[43] Tianguang Lu, Zhaoyu Wang, Jianhui Wang, Qian Ai, and Chong Wang. A data-driven stackelberg market strategy for demand response-enabled distribution systems. *IEEE Transactions on Smart Grid*, 10(3):2345–2357, 2018.

[44] J MacQueen. Classification and analysis of multivariate observations. In *5th Berkeley Symp. Math. Statist. Probability*, pages 281–297, 1967.

[45] Trieu Mai, John Bistline, Yinong Sun, Wesley Cole, Cara Marcy, Chris Namovicz, and David Young. The role of input assumptions and model structures in projections of variable renewable energy: A multi-model perspective of the us electricity system. *Energy Economics*, 76:313–324, 2018.

[46] Patrice Marcotte, Anne Mercier, Gilles Savard, and Vedat Verter. Toll policies for mitigating hazardous materials transport risk. *Transportation science*, 43(2):228–243, 2009.

[47] Frederic H Murphy and Yves Smeers. Generation capacity expansion in imperfectly competitive restructured electricity markets. *Operations research*, 53(4):646–661, 2005.

[48] Christos Ordoudis, Pierre Pinson, Juan M Morales, and Marco Zugno. An updated version of the ieee rts 24-bus system for electricity market and power system operation studies. *Technical University of Denmark*, 13, 2016.

[49] Bryan S Palmintier and Mort D Webster. Heterogeneous unit clustering for efficient operational flexibility modeling. *IEEE Transactions on Power Systems*, 29(3):1089–1098, 2013.

[50] Michael Patriksson. *The traffic assignment problem: models and methods*. Courier Dover Publications, 2015.

[51] André Pina, Carlos Silva, and Paulo Ferrão. Modeling hourly electricity dynamics for policy making in long-term scenarios. *Energy Policy*, 39(9):4692–4702, 2011.

[52] André Pina, Carlos A Silva, and Paulo Ferrão. High-resolution modeling framework for planning electricity systems with high penetration of renewables. *Applied Energy*, 112:215–223, 2013.

[53] Mehrdad Pirnia and J David Fuller. A mixed integer nonlinear program for electric generation expansion with energy and capacity pricing. In *IEEE PES General Meeting*, pages 1–7. IEEE, 2010.

[54] Quentin Ploussard, Luis Olmos, and Andrés Ramos. An operational state aggregation technique for transmission expansion planning based on line benefits. *IEEE Transactions on Power Systems*, 32(4):2744–2755, 2016.

[55] K Poncelet, H Höschle, E Delarue, and W D'haeseleer. Selecting representative days for investment planning models. *KU Leuven*, 2015.

[56] Kris Poncelet, Hanspeter Höschle, Erik Delarue, Ana Virag, and William D'haeseleer. Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems. *IEEE Transactions on Power Systems*, 32(3):1936–1948, 2017.

[57] Himanshu Rathore and Suresh Kumar Jakhar. Differential carbon tax policy in aviation: One stone that kills two birds? *Journal of Cleaner Production*, 296:126479, 2021.

[58] Carlos Ruiz, Antonio J Conejo, and Dimitris J Bertsimas. Revealing rival marginal offer prices via inverse optimization. *IEEE Transactions on Power Systems*, 28(3):3056–3064, 2013.

[59] Javier Saez-Gallego and Juan M Morales. Short-term forecasting of price-responsive loads using inverse optimization. *IEEE Transactions on Smart Grid*, 9(5):4805–4814, 2017.

[60] Javier Saez-Gallego, Juan M Morales, Marco Zugno, and Henrik Madsen. A data-driven bidding model for a cluster of price-responsive consumers of electricity. *IEEE Transactions on Power Systems*, 31(6):5001–5011, 2016.

[61] Andrew J Schaefer. Inverse integer programming. *Optimization Letters*, 3(4):483–489, 2009.

[62] Zahed Shahmoradi and Taewoo Lee. Quantile inverse optimization: Improving stability in inverse linear programming. *Operations Research*, 70(4):2538–2562, 2022.

[63] Walter Short, Patrick Sullivan, Trieu Mai, Matthew Mowers, Caroline Uriarte, Nate Blair, Donna Heimiller, and Andrew Martinez. Regional energy deployment system (reeds). Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States), 2011.

[64] Georgios Tsaousoglou, Juan S Giraldo, and Nikolaos G Paterakis. Market mechanisms for local electricity markets: A review of models, solution concepts and algorithmic techniques. *Renewable and Sustainable Energy Reviews*, 156:111890, 2022.

[65] Jianhui Wang, Mohammad Shahidehpour, Zuyi Li, and Audun Botterud. Strategic generation capacity expansion planning with incomplete information. *IEEE Transactions on Power Systems*, 24(2):1002–1010, 2009.

[66] Sonja Wogrin, Pablo Dueñas, Andrés Delgadillo, and Javier Reneses. A new approach to model load levels in electric power systems with high renewable penetration. *IEEE Transactions on Power Systems*, 29(5):2210–2218, 2014.

[67] Hellmut Wollmann and Gérard Marcou. *The provision of public services in Europe: Between state, local government and market*. Edward Elgar Publishing, 2010.

[68] Jianzhong Zhang, Zhenhong Liu, and Zhongfan Ma. On the inverse problem of minimum spanning tree with partition constraints. *Mathematical methods of operations research*, 44(2):171–187, 1996.

[69] Ying Zhou, Lizhi Wang, and James D McCalley. Effective incentives design for renewable energy generation expansion planning: An inverse optimization approach. In *IEEE PES General Meeting*, pages 1–7. IEEE, 2010.

[70] Ying Zhou, Lizhi Wang, and James D McCalley. Designing effective and efficient incentive policies for renewable energy in generation expansion planning. *Applied Energy*, 88(6):2201–2209, 2011.

APPENDICES

Appendix A

Generator behavior scenario trees

$$
\text{if } h \neq 1, 24 \text{ then }
\begin{cases}
\text{Case 3.7a} \begin{cases}
\text{Case 3.8a} \begin{cases}
\text{Case 3.8b:} & f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{2,i,g,d,h} - \sigma_{3,i,g,d,h+1} - \sigma_{6,i,g,d,h})/w_{i,g} \\
\text{Case 3.8c:} & f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{2,i,g,d,h} - \sigma_{6,i,g,d,h})/w_{i,g}
\end{cases} \\
\text{Case 3.8c} \begin{cases}
\text{Case 3.8b:} & f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{3,i,g,d,h+1} - \sigma_{6,i,g,d,h})/w_{i,g} \\
\text{Case 3.8c:} & f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{6,i,g,d,h})/w_{i,g}
\end{cases}
\end{cases} \\
\text{Case 3.7b} \begin{cases}
\text{Case 3.8b} \begin{cases}
\text{Case 3.8a:} & f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h+1} + \sigma_{3,i,g,d,h} - \sigma_{7,i,g,d,h})/w_{i,g} \\
\text{Case 3.8c:} & f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{3,i,g,d,h} - \sigma_{7,i,g,d,h})/w_{i,g}
\end{cases} \\
\text{Case 3.8c} \begin{cases}
\text{Case 3.8a:} & f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h+1} + \sigma_{7,i,g,d,h})/w_{i,g} \\
\text{Case 3.8c:} & f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{7,i,g,d,h})/w_{i,g}
\end{cases}
\end{cases} \\
\text{Case 3.7c} \begin{cases}
\text{Case 3.8a} \begin{cases}
\text{Case 3.8a:} & f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{2,i,g,d,h} + \sigma_{2,i,g,d,h+1})/w_{i,g} \\
\text{Case 3.8b:} & f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{2,i,g,d,h} - \sigma_{3,i,g,d,h+1})/w_{i,g} \\
\text{Case 3.8c:} & f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{2,i,g,d,h})/w_{i,g}
\end{cases} \\
\text{Case 3.8b} \begin{cases}
\text{Case 3.8a:} & f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h+1} + \sigma_{3,i,g,d,h})/w_{i,g} \\
\text{Case 3.8b:} & f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{3,i,g,d,h} - \sigma_{3,i,g,d,h+1})/w_{i,g} \\
\text{Case 3.8c:} & f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{3,i,g,d,h})/w_{i,g}
\end{cases} \\
\text{Case 3.8c} \begin{cases}
\text{Case 3.8a:} & f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h+1})/w_{i,g} \\
\text{Case 3.8b:} & f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{3,i,g,d,h+1})/w_{i,g} \\
\text{Case 3.8c:} & f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h})/w_{i,g}
\end{cases}
\end{cases}
\end{cases}
\tag{A.1}
$$

In the cases A.1, an example case is when Capacity constraints are not active (3.7c), in this case ramping constraints, could be active or inactive, let's assume it is in active (3.8c), now, the next hour could have any of the two situations already mentioned, either the ramping constraint is active for it or not. If it is active (3.8a) then it could be claimed that $f_v(i,g,h) \geq \sigma_{1,d,h}$.

$$
\text{if } h = 1 \text{ then } \begin{cases}
\text{Case 3.7a} \begin{cases}
\text{Case 3.9a} \begin{cases}
\text{Case 3.8a: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h+1} + \sigma_{4,i,g,d} - \sigma_{6,i,g,d,h})/w_{i,g} \\
\text{Case 3.8b: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{3,i,g,d,h+1} + \sigma_{4,i,g,d} - \sigma_{6,i,g,d,h})/w_{i,g} \\
\text{Case 3.8c: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{4,i,g,d} - \sigma_{6,i,g,d,h})/w_{i,g}
\end{cases} \\
\text{Case 3.9b} \begin{cases}
\text{Case 3.8a: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h+1} - \sigma_{5,i,g,d} - \sigma_{6,i,g,d,h})/w_{i,g} \\
\text{Case 3.8b: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{3,i,g,d,h+1} - \sigma_{5,i,g,d} - \sigma_{6,i,g,d,h})/w_{i,g} \\
\text{Case 3.8c: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{5,i,g,d} - \sigma_{6,i,g,d,h})/w_{i,g}
\end{cases} \\
\text{Case 3.9c} \begin{cases}
\text{Case 3.8a: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h+1} - \sigma_{6,i,g,d,h})/w_{i,g} \\
\text{Case 3.8b: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{3,i,g,d,h+1} - \sigma_{6,i,g,d,h})/w_{i,g} \\
\text{Case 3.8c: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{6,i,g,d,h})/w_{i,g}
\end{cases}
\end{cases} \\
\text{Case 3.7b} \begin{cases}
\text{Case 3.9a} \begin{cases}
\text{Case 3.8a: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h+1} + \sigma_{4,i,g,d} + \sigma_{7,i,g,d,h})/w_{i,g} \\
\text{Case 3.8b: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{3,i,g,d,h+1} + \sigma_{4,i,g,d} + \sigma_{7,i,g,d,h})/w_{i,g} \\
\text{Case 3.8c: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{4,i,g,d} + \sigma_{7,i,g,d,h})/w_{i,g}
\end{cases} \\
\text{Case 3.9b} \begin{cases}
\text{Case 3.8a: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h+1} - \sigma_{5,i,g,d} + \sigma_{7,i,g,d,h})/w_{i,g} \\
\text{Case 3.8b: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{3,i,g,d,h+1} - \sigma_{5,i,g,d} + \sigma_{7,i,g,d,h})/w_{i,g} \\
\text{Case 3.8c: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{5,i,g,d} + \sigma_{7,i,g,d,h})/w_{i,g}
\end{cases} \\
\text{Case 3.9c} \begin{cases}
\text{Case 3.8a: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h+1} + \sigma_{7,i,g,d,h})/w_{i,g} \\
\text{Case 3.8b: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{3,i,g,d,h+1} + \sigma_{7,i,g,d,h})/w_{i,g} \\
\text{Case 3.8c: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{7,i,g,d,h})/w_{i,g}
\end{cases}
\end{cases} \\
\text{Case 3.7c} \begin{cases}
\text{Case 3.9a} \begin{cases}
\text{Case 3.8a: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h+1} + \sigma_{4,i,g,d})/w_{i,g} \\
\text{Case 3.8b: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{3,i,g,d,h+1} + \sigma_{4,i,g,d})/w_{i,g} \\
\text{Case 3.8c: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{4,i,g,d})/w_{i,g}
\end{cases} \\
\text{Case 3.9b} \begin{cases}
\text{Case 3.8a: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h+1} - \sigma_{5,i,g,d})/w_{i,g} \\
\text{Case 3.8b: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{3,i,g,d,h+1} - \sigma_{5,i,g,d})/w_{i,g} \\
\text{Case 3.8c: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{5,i,g,d})/w_{i,g}
\end{cases} \\
\text{Case 3.9c} \begin{cases}
\text{Case 3.8a: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h+1})/w_{i,g} \\
\text{Case 3.8b: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{3,i,g,d,h+1})/w_{i,g} \\
\text{Case 3.8c: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h})/w_{i,g}
\end{cases}
\end{cases}
\end{cases} \tag{A.2}
$$

$$
\text{if } h = 24 \text{ then }
\begin{cases}
\text{Case 3.7a}
\begin{cases}
\text{Case 3.9a}
\begin{cases}
\text{Case 3.8a: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{2,i,g,d,h} - \sigma_{4,i,g,d} - \sigma_{6,i,g,d,h})/w_{i,g} \\
\text{Case 3.8b: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{3,i,g,d,h} - \sigma_{4,i,g,d} - \sigma_{6,i,g,d,h})/w_{i,g} \\
\text{Case 3.8c: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{4,i,g,d} - \sigma_{6,i,g,d,h})/w_{i,g}
\end{cases} \\
\text{Case 3.9b}
\begin{cases}
\text{Case 3.8a: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{2,i,g,d,h} + \sigma_{5,i,g,d} - \sigma_{6,i,g,d,h})/w_{i,g} \\
\text{Case 3.8b: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{3,i,g,d,h} + \sigma_{5,i,g,d} - \sigma_{6,i,g,d,h})/w_{i,g} \\
\text{Case 3.8c: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{5,i,g,d} - \sigma_{6,i,g,d,h})/w_{i,g}
\end{cases} \\
\text{Case 3.9c}
\begin{cases}
\text{Case 3.8a: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{2,i,g,d,h} - \sigma_{6,i,g,d,h})/w_{i,g} \\
\text{Case 3.8b: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{3,i,g,d,h} - \sigma_{6,i,g,d,h})/w_{i,g} \\
\text{Case 3.8c: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{6,i,g,d,h})/w_{i,g}
\end{cases}
\end{cases} \\
\text{Case 3.7b}
\begin{cases}
\text{Case 3.9a}
\begin{cases}
\text{Case 3.8a: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{2,i,g,d,h} - \sigma_{4,i,g,d} + \sigma_{7,i,g,d,h})/w_{i,g} \\
\text{Case 3.8b: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{3,i,g,d,h} - \sigma_{4,i,g,d} + \sigma_{7,i,g,d,h})/w_{i,g} \\
\text{Case 3.8c: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{4,i,g,d} + \sigma_{7,i,g,d,h})/w_{i,g}
\end{cases} \\
\text{Case 3.9b}
\begin{cases}
\text{Case 3.8a: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{2,i,g,d,h} + \sigma_{5,i,g,d} + \sigma_{7,i,g,d,h})/w_{i,g} \\
\text{Case 3.8b: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{3,i,g,d,h} + \sigma_{5,i,g,d} + \sigma_{7,i,g,d,h})/w_{i,g} \\
\text{Case 3.8c: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{5,i,g,d} + \sigma_{7,i,g,d,h})/w_{i,g}
\end{cases} \\
\text{Case 3.9c}
\begin{cases}
\text{Case 3.8a: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{2,i,g,d,h} + \sigma_{7,i,g,d,h})/w_{i,g} \\
\text{Case 3.8b: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{3,i,g,d,h} + \sigma_{7,i,g,d,h})/w_{i,g} \\
\text{Case 3.8c: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{7,i,g,d,h})/w_{i,g}
\end{cases}
\end{cases} \\
\text{Case 3.7c}
\begin{cases}
\text{Case 3.9a}
\begin{cases}
\text{Case 3.8a: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h+1} + \sigma_{4,i,g,d})/w_{i,g} \\
\text{Case 3.8b: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{3,i,g,d,h+1} + \sigma_{4,i,g,d})/w_{i,g} \\
\text{Case 3.8c: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{4,i,g,d})/w_{i,g}
\end{cases} \\
\text{Case 3.9b}
\begin{cases}
\text{Case 3.8a: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h+1} - \sigma_{5,i,g,d})/w_{i,g} \\
\text{Case 3.8b: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{3,i,g,d,h+1} - \sigma_{5,i,g,d})/w_{i,g} \\
\text{Case 3.8c: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{5,i,g,d})/w_{i,g}
\end{cases} \\
\text{Case 3.9c}
\begin{cases}
\text{Case 3.8a: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} + \sigma_{2,i,g,d,h+1})/w_{i,g} \\
\text{Case 3.8b: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h} - \sigma_{3,i,g,d,h+1})/w_{i,g} \\
\text{Case 3.8c: } f_v(i,g,h) = (-w_{i,g}\sigma_{1,d,h})/w_{i,g}
\end{cases}
\end{cases}
\end{cases}
\tag{A.3}
$$

Appendix B

IO model to recover a uniform price for the whole market

As is the case for many energy markets, there are some components of data that are deemed confidential and as a result not published. An example of such data is transmission limits and power transmissions. If we reduce the real-world model such that we would have only one bus and as a result, all the loads and generators are connected to that one bus, we would end up having electricity prices that are not locational. In a real-world energy market such as Ontario, we know that there are some side payments to recover for the costs of some participants. These side payments are necessary because many operational details have not been accounted for when clearing the market. In addition, we have no information about the extra payments market operators make, but can assume that these payments are a result of not considering operational details.

This section contains a market clearing that is suitable for all locations. This procedure breaches some of the initial limits; however, we recognize that these violations are similar to the side payments that a market operator must pay while simplifying its own operations.

Lagrangian function:

$$L = \sum_{r\in\Psi,g\in\Omega_r}(\sum_{h\in H} p_{i,g,d,h}C_g^v) - \sum_{r\in\Psi,d\in D,h\in H}\sigma_{1,d,h}(\sum_{c\in\Omega_r} p_{i,g,d,h} + \sum_{s\in E_r} b_{rs}.(\theta_{r,h}-\theta_{s,h}) - L_{r,d,h}) +$$
$$\sum_{i\in I,g\in\Omega_r,d\in D,h\neq 1}\sigma_{2,i,g,d,h}(p_{i,g,d,h}-p_{i,g,d,h-1}-Ru_g^{max}) + \sum_{i\in I,g\in\Omega_r,d\in D,h\neq 1}\sigma_{3,i,g,d,h}(p_{i,g,d,h-1}$$
$$-p_{i,g,d,h}-Rd_g^{max}) + \sum_{r\in\Psi,g\in\Omega_r,d\in D,h\in H}\sigma_{4,i,g,d}(p_{i,g,d,24}-p_{i,g,d,1}-Ru_g^{max}) +$$
$$\sum_{r\in\Psi,g\in\Omega_r,d\in D,h\in H}\sigma_{5,i,g,d}(p_{i,g,d,1}-p_{i,g,d,24}-Rd_g^{max}) + \sum_{r\in\Psi,g\in\Omega_r,d\in D,h\in H}\sigma_{6,i,g,d,h}(p_{i,g,d,h}-\kappa_g)$$
$$+\sum_{r\in\Psi,g\in\Omega_r,d\in D,h\in H}\sigma_{7,i,g,d,h}(-p_{i,g,d,h}) + \sum_{r\in\Psi,s\in\Psi,h\in H}\sigma_{8,r,s,h}(b_{rs}.(\theta_{r,h}-\theta_{s,h})-P_{rs}^{max})$$
$$+\sigma_9\theta_{REF,h}$$

Dual constraints

$$\begin{aligned}&C_g^v - \sigma_{1,d,h} - \sigma_{2,i,g,d,h+1} + \sigma_{3,i,g,d,h+1} - \sigma_{4,i,g,d} + \sigma_{5,i,g,d} + \sigma_{6,i,g,d,h} - \sigma_{7,i,g,d,h} = 0,\\ &\forall r \in \Psi, h = 1\end{aligned} \tag{B.1a}$$

$$\begin{aligned}&C_g^v - \sigma_{1,d,h} + \sigma_{2,i,g,d,h} - \sigma_{2,i,g,d,h+1} - \sigma_{3,i,g,d,h} + \sigma_{3,i,g,d,h+1} + \sigma_{6,i,g,d,h} - \sigma_{7,i,g,d,h} = 0,\\ &\forall r \in \Psi, h \neq 1, 24\end{aligned} \tag{B.1b}$$

$$\begin{aligned}&C_g^v - \sigma_{1,d,h} + \sigma_{2,i,g,d,h} - \sigma_{3,i,g,d,h} + \sigma_{4,i,g,d} - \sigma_{5,i,g,d} + \sigma_{6,i,g,d,h} - \sigma_{7,i,g,d,h} = 0,\\ &\forall r \in \Psi, h = 24\end{aligned} \tag{B.1c}$$

$$\begin{aligned}&-\sigma_{1,d,h}\sum_{s\in E_r} b_{rs} + \sum_{\forall k|r\in E_k}\sigma_{1,d,h}b_{kr} + \sigma_{8,r,s,h}\sum_{s\in E_r} b_{rs} - \sum_{\forall k|r\in E_k}\sigma_{8,k,r,h}b_{kr} = 0,\\ &\forall r \in \Psi, h \in H\end{aligned} \tag{B.1d}$$

$$\begin{aligned}&-\sigma_{1,d,h}\sum_{s\in E_r} b_{rs} + \sum_{\forall k|r\in E_k}\sigma_{1,d,h}b_{kr} + \sigma_{8,r,s,h}\sum_{s\in E_r} b_{rs} - \sum_{\forall k|r\in E_k}\sigma_{8,k,r,h}b_{kr} + \sigma_9 = 0,\\ &r = REF, \forall h \in H\end{aligned} \tag{B.1e}$$

Complimentary slackness constraints

$$\sigma_{1,d,h}(\sum_{c\in\Omega_r} p_{i,g,d,h} + \sum_{s\in E_r} b_{rs}.(\theta_{r,t}-\theta_{s,t}) - L_{r,d,h}) = 0, \qquad \forall r \in \Psi, h \in H \tag{B.2a}$$

$$\sigma_{2,i,g,d,h}(p_{i,g,d,h} - p_{i,g,d,h-1} - Ru_{i,g}) = 0, \qquad \forall r \in \Psi, h \neq 1 \qquad \text{(B.2b)}$$
$$\sigma_{3,i,g,d,h}(p_{i,g,d,h-1} - p_{i,g,d,h} - Rd_{i,g}) = 0, \qquad \forall i \in I, g \in G_i, h \neq 1 \qquad \text{(B.2c)}$$
$$\sigma_{4,i,g,d}(p_{i,g,d,24} - p_{i,g,d,1} - Ru_{i,g}) = 0, \qquad \forall i \in I, g \in G_i \qquad \text{(B.2d)}$$
$$\sigma_{5,i,g,d}(p_{i,g,d,1} - p_{i,g,d,24} - Rd_{i,g}) = 0, \qquad \forall i \in I, g \in G_i \qquad \text{(B.2e)}$$
$$\sigma_{6,i,g,d,h}(p_{i,g,d,h} - \kappa_g) = 0, \qquad \forall i \in I, g \in G_i, h \in H \qquad \text{(B.2f)}$$
$$\sigma_{7,i,g,d,h}(-p_{i,g,d,h}) = 0, \qquad \forall i \in I, g \in G, h \in H \qquad \text{(B.2g)}$$
$$\sigma_{8,r,s,h}(b_{rs}.(\theta_{r,h} - \theta_{s,h}) - P_{rs}^{max}) = 0, \qquad \forall r \in \Psi, s \in E_r, h \in H \qquad \text{(B.2h)}$$
$$\sigma_9 \theta_{REF,h} = 0, \qquad \forall h \in H \qquad \text{(B.2i)}$$

Model B.3 below is the final model solution which provides us with a uniform market price. Since some of the constraints (B.1a) - (B.1e), and (B.2a) - (B.2g) could be violated penalty variables ϵ are introduced and put in the right-hand-side of each of these constraints. Finally, we try to minimize this amount of penalty by including that in the objective function.

$$\begin{aligned}
\min_{\sigma,\epsilon} \quad & \sum_{r\in\Psi,d\in D,h\in H} |\sigma_{1,d,h} - \sigma_{1,r,d,h}| + \sum_{r\in\Psi,h=1} |\epsilon_{3a,r,h}| + \sum_{r\in\Psi,h\neq 1,24} |\epsilon_{3b,r,h}| + \sum_{r\in\Psi,h=24} |\epsilon_{3c,r,h}| \\
& + \sum_{r\in\Psi,h\in H} (|\epsilon_{3d,r,h}| + |\epsilon_{3e,r,h}|) + \sum_{r\in\Psi,h\in H} |\epsilon_{4a,r,h} + \sum_{r\in\Psi,h\neq 1} |\epsilon_{4b,r,h}| + \sum_{i\in I,g\in G_i,h\neq 1} |\epsilon_{4c,i,g,h}| \\
& + \sum_{g\in G_i} (|\epsilon_{4d,g}| + |\epsilon_{4e,g}|) + \sum_{i\in I,g\in G_i,h\in H} (|\epsilon_{4f,i,g,h}| + |\epsilon_{4g,i,g,h}|) + \sum_{r\in\Psi,s,h} |\epsilon_{4h,r,s,h}| + \\
& \sum_{h\in H} |\epsilon_{4i,h}|
\end{aligned} \qquad \text{(B.3a)}$$

$$\text{s.t.} \quad \text{(B.1a) - (B.1e)} \qquad \text{(B.3b)}$$
$$\text{(B.2a) - (B.2g)} \qquad \text{(B.3c)}$$

After this step which results in finding $\sigma_{1,d,h}$, we can consider the assumptions of our problem and consider a reduced model without any knowledge of the transmission constraints.