ADVANCES IN THE OPTIMIZATION OF ENERGY SYSTEMS AND MACHINE LEARNING HYPERPARAMETERS

A Dissertation

by

WILLIAM WEIKANG TSO

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Efstratios N. Pistikopoulos
Committee Members,	Mahmoud El-Halwagi
	M.M. Faruque Hasan
	Bruce A. McCarl
Head of Department,	Arul Jayaraman

May 2020

Major Subject: Chemical Engineering

ABSTRACT

Intensifying public concern about climate change risks has accelerated the push for more tangible action in the transition toward low-carbon or carbon-neutral energy. Concurrently, the energy industry is also undergoing a digital transformation with the explosion in available data and computational power. To address these challenges, systematic decision-making strategies are necessary to analyze the vast array of technology options and information sources while navigating this energy transition. In this work, mathematical optimization is utilized to answer some of the outstanding issues around designing cleaner processes from resources such as natural gas and renewables, operating the logistics of these energy systems, and statistical modeling from data.

First, exploiting natural gas to produce lower emission liquid transportation fuels is investigated through an optimization-based process synthesis. This extends previous studies by incorporating chemical looping as an alternative syngas production method for the first time. Second, a similar process synthesis approach is implemented for the optimal design of a novel biomass-based process that coproduces ammonia and methanol, improving their production flexibility and profit margins.

Next, operational difficulties with solar and wind energies due to their temporal intermittency and uneven geographical distribution are tackled with a supply chain optimization model and a clustering decomposition algorithm. The former describes power generation through energy carriers (hydrogen-rich chemicals) connecting resource-dense rural areas to resource-deficient urban centers. Results show the potential of energy carriers for long-term storage. The latter is developed to identify the appropriate number of representative time periods for approximating an optimization problem with time series data, instead of using a full time horizon. This algorithm is applied to the simultaneous design and scheduling of a renewable power system with battery storage.

Finally, building machine learning models from data is commonly performed through k-fold cross-validation. From recasting this as a bilevel optimization, the exact solution to hyperparameter optimization is obtainable through parametric programming for machine learning models that are LP/QP. This extends previous results in statistics to a broader class of machine learning models.

DEDICATION

To my wife-to-be, Melissa, for her endless patience, loving support, and keeping me grounded

To my parents, Christine and Kevin, for instilling in me wisdom and inspiration to reach my goals

To my brother, Benjamin, for being my soundboard and having my back through ups & downs

To my cousins, Stephanie and Brandon, for their fortitude and perseverance through adversity

To the rest of my family, for everything you have provided me with

ACKNOWLEDGMENTS

My professional and personal trajectories in life have been greatly influenced by the late Professor Christodoulos Floudas, whom I owe an immense amount of gratitude. Thank you Professor Floudas for introducing me to optimization and process systems engineering in undergraduate. I will always remember your process design course and setting aside time in your busy schedule to talk to me about graduate school. The redirection of my career interests from biological engineering to energy systems optimization was sparked by your course and our conversations.

Your passing on August 14, 2016 was a sudden shock to my plans for graduate school. It has been a difficult road to recovery for all of us, but I believe there is always a silver lining in the midst of tragedy. I met my future wife, Melissa, while traveling to your memorial service at Princeton University in May 2017. Even in your absence, you amazingly found a way to impact my life for the better. I truly believe that where I am today and the path set forth before me now are because of your early encouragement and strong belief in me. You continue to be an inspiration in all of my endeavors. May you rest in peace. Much heartfelt appreciation to the Floudas family: Mrs. Fotini, Ismini, and Stefanos Baratsas for keeping his students on your minds, motivating us to finish our Ph.D. degrees, and encouraging us in our chosen careers.

I express my sincerest gratefulness to my advisor Professor Efstratios Pistikopoulos for allowing me to join his research group and providing me the opportunity to finish my Ph.D. studies upon the passing of Professor Floudas. I admire your resolve and determination to preserve the legacy of Professor Floudas and his research group, despite the lack of resources at the time. Thank you for your guidance and support in my research pursuits, your enthusiasm through the difficult times, an unmatched energy level for any professional or personal conversation, and shielding me away from the politics of graduate school. Many thanks to my other committee members: Professors Mahmoud El-Halwagi, M.M. Faruque Hasan, and Bruce McCarl for their astute suggestions and precious time spent on this dissertation. I would also like to thank Professor Phanourios Tamamis for his warm passion and care in helping me when I was his TA for the undergraduate thermody-

namics course. I appreciate Professor Arul Jayaraman and Ashley Henley for all of their assistance in answering my questions and addressing my concerns during the graduate studies.

I express my deepest gratitude to Dr. Joseph Powell from Royal Dutch Shell for securing the financial funds to support my dissertation and offering valuable insight at every step. Thank you for giving me career advice and helping me decide where to start my career, even though ultimately it was not the desired outcome you had wanted. Many thanks to the Shell New Energies colleagues: Dr. Clara Heuberger and Dr. Mark Klokkenburg for their inputs on my energy systems optimization work. My two summer internships at Emerson Automation Solutions and ExxonMobil Research & Engineering also afforded me with great insight on relevant problems that industry are tackling with optimization and analytics. I am grateful for my time spent at these two companies and cherish the learning experiences. These helped to sharpen my research acumen and shape my career aspirations. Thank you to Dr. Dirk Thiele, Eric Schulz, and Caroline Cisneros from Emerson Automation Solutions for creating an opportunity for me to do a process control & data science project. Thank you to Dr. Ravishankar Sethuraman, Dr. Ryan Gwaltney, Dr. Jose Gomez, Dr. Ramasamy Selvaraj, Dr. Manjiri Moharir, Dr. Yunfei Chu, Dr. Nathan Yee, Dr. Peter Hanratty, and John Righi from ExxonMobil Research & Engineering for providing me the freedom to work on two projects related to real-time optimization and process constraint analysis.

At Texas A&M, I am honored to have cross paths with many members of the Floudas and Pistikopoulos research groups. I have enjoyed all of our conversations and camaraderie throughout the past five years. I am thankful for the senior Floudas lab members: Dr. Chris Kieslich, Dr. Fani Boukouvala, Dr. Yannis Guzman, Dr. Alexander Niziolek, Dr. Onur Onel, Dr. Logan Matthews, and Dr. Melis Onel for immersing me into research and graduate school life. I especially acknowledge Drs. Onur Onel and Alexander Niziolek for training me on my initial research projects and giving me career advice. I also sincerely thank Dr. Logan Matthews for presenting a spiritual avenue outside of the lab and inviting me to join his church & Bible study groups. Thank you Joseph Costandy and Utkarsh Shah for being a part of my early days in the Floudas lab as well.

From the Pistikopoulos lab, I am particularly grateful for Justin Katz, Baris Burnak, Burcu

Beykal, and Doga Demirhan. We entered graduate school at the same time, have gone through all the struggles together, and shared many laughs along the way. I would like to especially thank Doga Demirhan, my research and discussion partner on all of our collaborated energy systems optimization studies. I also have enjoyed my time with other group members: Dr. Styliani Avraamidou, Dr. Gerald Ogumerem, Dr. Nikolaos Diangelakis, Dr. Ioana Nascu, Dr. Maria Papathanasiou, Dr. Richard Oberdieck, Dr. Rajib Mukherjee, Dr. Hari Ganesh, Cory Allen, Yuhe Tian, Iosif Pappas, Christopher Gorden, Denis Su-Feher, Dustin Kenefake, and Rahul Kakodkar. I sincerely appreciate Yuhe Tian and Iosif Pappas, my deskmates, for their support and engaging conversations during the last few months as I finish up my dissertation. I would also like to acknowledge former visiting scholars: Zhihao Chen, Yaling Nie, Marcello Di Martino, Mario Villanueva, and Wenjie Xu and master's students: Haneol Song and Seungyeon Lee for their time spent in the lab. Thanks to my classmates Yixi Chen, Azhar Ali, Jingyao Wang, and Emre Demirel for sympathizing with their own graduate school experiences and giving me perspective.

To my chemical & biological engineering (CBE) classmates from Princeton, especially Caitlin Wood, Lisa Kojima, Helen Yao, Ming-Ming Tran, Sunny Niu, Elliot Horlick, Joshua Taliaferro, and Ismael Catovic, thank you for getting me through the challenges of our undergraduate courses. I never told you guys this before, but I would never have made it through Princeton or applied to graduate school without your friendship. Many thanks to Professors Robert Prud'homme and Mark Brynildsen for writing my letters of recommendation and their continued encouragement throughout the years. Thanks also to Thoa Bui for writing a recommendation letter for me. The summer internship at MedImmune before my senior year made me realize that I wanted to pursue graduate research. To my Bishop Eustace Prep friends, Matthew Thompson, Eric Carpizo, and Jayanth Vatson, thank you for getting me through the growing pains of high school.

I have been blessed by God who has given me abundant opportunities and unwavering guidance. To my Bible study members: David and B.J. Fox, Felix and Teegee Thompson, George and Carol Parry, Steve and LeAnn Strong, Ray and Cindy Montoya, Ron and Sylvia Toole, Chip and Diane Jennings, Murphy and Kathy Smith, I present my wholehearted appreciation for providing

me a comforting space and respite away from graduate school. You guys remind me to always keep faithful and look toward God for hope in this broken world. I especially thank David Fox and Steve Strong for continuously mentoring me and the many thought-provoking conversations shared over meals together. I would also like to thank Matt Purcell for his early counseling in the Sunday School class at Grace Bible Church - Southwood Campus. I am forever grateful that I discovered God and the Gospel during my time in College Station. I am thankful to have met each one of you and everyone else from church whom I have failed to mention. My sincerest condolences to James and Maria Childs for the lost of their son. Rest in peace David Childs. Thank you for trusting me to be part of your son's life, even though it was only for a short while.

I cannot express my gratitude enough for my family and close friends who have witnessed my journey through graduate school. I have been blessed to have Melissa by my side for the last three years of my Ph.D. journey. Her bubbly spirit and persistent love has given me the strength to push through the finish line. I know I have tested your patience. I love you so much for bearing with me through the tough and happy times. My parents, Christine and Kevin, always stressed higher education to me from the time I entered kindergarten and sacrificed to give me every possible learning opportunity. I am proud to say that 22 years later my formal education trek is finally coming to a conclusion. This one is for you Mom and Dad. My brother, Benjamin, is the carefree and goofy character in the family, but truly knows how to deliver important viewpoints when they are needed. I always value your being there for me and providing a soundboard to bounce ideas off of. I cannot wait to see all the successes that lie ahead of you. To the rest of my immediate family members: Stephanie, Brandon, Aunt Judy, Uncle Tony, Austin, Grandma & Grandpa Wong, Aunt Mindy, Uncle Kit, Alina, Desmond, Aunt Terri, Uncle Johnny, Clayton, Isabelle, and Grandma & Grandpa Tso, and Minji, I am so grateful for all your concerns and well wishes from the Northeast. To my new family members: Suree, Ted, Sayuri, and Wilson, I truly appreciate you guys making me feel at home in Texas and welcoming me in. To my groomsmen: Johnathan Yao, Marcus Lee, Joseph Laseter, Adam Fisch, Erick Chen, Sunny Patel, and Jamie Gillespie, thanks so much for being my close friends and always providing moral support from afar.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professors Efstratios Pistikopoulos, Mahmoud El-Halwagi, and M.M. Faruque Hasan of the Department of Chemical Engineering and Professor Bruce McCarl of the Department of Agricultural Economics.

The ammonia process models utilized in Section 3 were jointly developed with Doga Demirhan. The design and scheduling model implemented in Section 5 was also created in collaboration with Doga Demirhan. Initial work on the energy carrier supply chain optimization presented in Section 4 was conducted by Haneol Song and Seungyeon Lee, two former master's students in the Pistikopoulos group, and supervised by Doga Demirhan and the author. Haneol Song is currently a researcher at Hyundai Motor Company and Seungyeon Lee is a Ph.D. student at the University of Delaware. Suggestions on the clustering algorithm in Section 5 were offered by Dr. Joseph Powell and Dr. Clara Heuberger from Royal Dutch Shell. Baris Burnak assisted in the initial conception of the work presented in Section 6. Works related to the ongoing collaboration between the Pistikopoulos group and Royal Dutch Shell are located in Sections 3, 4, and 5.

All other work conducted for the dissertation was completed by the author independently.

Funding Sources

My graduate studies was funded by financial support from Royal Dutch Shell, Texas A&M Energy Institute, National Science Foundation (NSF CBET-1158849), and a Texas A&M Graduate Assistant Teaching (GAT) fellowship.

NOMENCLATURE

A	Ammonia
AHC	Agglomerative hierarchical clustering
AR	Air reactor
ASU	Air separation unit
ATR	Autothermal reforming
BAM	Biomass-based ammonia and methanol coproduction
BEOP	Break-even oil price
BEP	Break-even price
CAES	Compressed air energy storage
CAPEX	Capital expense
CART	Classification and regression trees
CLC	Chemical looping combustion
CLR	Chemical looping reforming
DNI	Direct normal solar irradiance
ECN	Energy carrier network
FBBT	Feasibility-based bounds tightening
FR	Fuel reactor
GHG	Greenhouse gas
GHGAE	Greenhouse gas avoided from electricity
GHGAM	Greenhouse gas avoided from methanol
GHGAN	Greenhouse gas avoided from ammonia
GHGI	Greenhouse gas emissions index

GHI	Global horizontal solar irradiance
GTL	Natural gas to liquid transportation fuels
H	Hydrogen
H_2 eq.	Hydrogen equivalent
HY-POP	Hyperparameter optimization through parametric programming
I	Integrated
IoT	Internet of Things
LASSO	Least absolute shrinkage and selection operator
LAR	Least angle regression
LB	Lower bound
LCOE	Levelized cost of electricity
LHV	Lower heating value
LGHG	Lifecycle greenhouse gas emissions
LP	Linear programming
LPG	Liquefied petroleum gas
kBPD	Thousand barrels per day
M	Methanol
MILP	Mixed-integer linear programming
MINLP	Mixed-integer nonlinear programming
MIQP	Mixed-integer quadratic programming
MM	Million
MPC	Model predictive control
mp	Multi-parametric
MSE	Mean squared error
MSW	Municipal solid waste

NLP	Nonlinear programming
NYC	New York City
O&M	Operation and maintenance
OBBT	Optimality-based bounds tightening
OLS	Ordinary lease squares regression
OPEX	Operating expense
p	Parametric
POP	Parametric optimization
PSE	Process systems engineering
PSH	Pumped-storage hydroelectricity
PV	Solar photovoltaic
QP	Quadratic programming
RMSE	Root mean squared error
RTN	Resource task network
SMR	Steam methane reforming
STS	Shale gas-to-syngas
SVM	Support vector machine
UB	Upper bound
WCSS	Within cluster sum of squares
W_D	Ward's distance

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION

1.1 The Energy Future: Business as Usual?

The world population is projected to surpass 10 billion people by 2055, approximately 30% greater than that in 2017 [1]. At the same time, the global middle class will undergo an unprecedented expansion [2], increasing from around 3 billion in 2015 to about 5.4 billion by 2030, driven largely by burgeoning economies in Asia. With more people, especially those who will adopt a higher standard of living, global energy consumption is expected to greatly appreciate in the coming decades, even when accounting for efficiency gains to offset the population and prosperity increases. Major energy producers like ExxonMobil and Royal Dutch Shell expect the annual global energy demand to reach around 700-800 exajoules by 2040 and 2050, respectively [3, 4]. This is roughly a 25% increase from the energy consumption amount in 2018 [5].

Moreover, the types of energy consumed will also change. Energy growth will not be uniformly experienced throughout all areas. Oil demand is expected to marginally grow (15% increase through 2040) [6] to support rising transportation volume from commerce and leisure, but this is significantly counteracted by vehicles with much greater fuel economy. On the other hand, electricity is projected to increase its share of the total energy consumption from 19% to a quarter or more by 2040, reflecting an absolute growth of approximately 60% from its current output [5]. This is due to the push toward electrification in vehicle transportation and new residential demand as nearly 1 billion people gain access to electricity. Commodity chemicals is another fast growing sector with an expected increase of about 40% from 2016 to 2040 [3] as individuals with rising living standards, especially those in China and India, utilize more petrochemical products such as plastics and consumer goods in their daily lives.

To supply this energy, today's production systems are primarily driven by fossil fuels such as oil, coal, and natural gas, comprising about 85% of the world's energy consumption [7]. Dependence on fossil fuels for affordable and reliable energy has powered the astronomical ascension

of developed countries like the United States through several industrial revolutions and created enormous capital wealth. However, this remarkable progress over the last few hundred years and improvement in the quality of human life have come at a dire environmental cost. In 2018, CO_2 emissions reached a record 37.1 billion tons [8], which is unfortunately not surprising since emission levels have been increasing almost every year since the end of World War II [9]. In the coming years, as emerging markets like China and India gear up to develop into advanced economies and demand a greater energy share, continuing business as usual will almost certainly exacerbate environmental damage and accelerate the risks of irreversible climate change.

Therefore, adjustments to current energy production systems are needed to adapt to the new approaching energy landscape. Continuing to utilize fossil fuels at the same rate moving forward would be an obsolete choice in terms of meeting increasingly diverse energy requirements in a more eco-friendly and sustainable manner.

One option is to update and retrofit existing fossil fuel technologies to be more efficient and cleaner. Carbon capture and process control automation are such examples. Because the capital cost of constructing new production systems are often very expensive and risky, this is a safe option, but also a low-hanging fruit that does not directly address the underlying issue. Another option is to develop and implement new fossil fuel technologies that are vastly superior to their predecessors in terms of efficiency and carbon intensity. Examples include the on-purpose production of ethylene and propylene from natural gas liquids as opposed to crude oil and chemical looping with inherent carbon sequestration for power production. While these systems are costly to build, they better address the problem of reducing CO_2 emissions, but are still not sustainable solutions. The final option is to completely revamp and rebuild energy infrastructures to run on renewable energies such as biomass, hydropower, solar, and wind. An example is using solar or wind powered electrolysis to produce hydrogen for vehicle transportation. This is certainly the most costly and risky option, but also the only one that can provide a sustainable energy future.

Which of these paths should be taken? There is not a simple or single solution. Different government leaders, corporate businessmen, influential economists, and outspoken scientists all

have contrasting opinions and interests in the right direction moving forward. Ideally speaking, producing all of the world's energy solely from renewable resources right now would be highly desirable, but this is not practical. Fossil fuel energy systems have evolved and matured over hundreds of years to meet the world's energy demands. It would be unreasonable to expect renewable energy systems to do the same in a much shorter amount of time. Even though there are different thoughts about what to do when and how to transform the energy network, almost everyone agrees that continuing business as usual is not an option. Fossil fuels or renewable energies alone cannot support the world's energy needs. There is no silver bullet. Near-term future energy systems will likely resemble a combination of all options, but the transition must begin sooner rather than later if permanent climate change is to be avoided and decarbonization is to be realized.

1.2 An Energy Transition

Fossil fuels maintain an advantage over renewables because they are cheaper, more robustly available, and denser in energy. Oil, coal, and natural gas exist as physical entities that result from millions of years of geological formation. They are accessible energy forms and using already advanced technologies can be drilled or mined for with relative ease. Whereas, solar irradiance and wind speed are intangible assets that are dispersed and need to be purposely concentrated in order to use. This process of collecting and densifying solar and wind into usable energy forms is the reason why their maturing technologies have been prohibitively expensive until very recently. Unlike solar or wind, biomass is a physical material that is directly usable and can be renewable from CO_2 photosynthesis. However, biomass has not been widely utilized because of concerns regarding the resources required to harvest it, potential land competition with food agriculture, feedstock variability & impurities, and carbon intensity associated with land preparation [10, 11].

An energy transition period is ultimately necessary because renewable technologies are not ready to thoroughly supplant their fossil fuel predecessors just yet. Besides greater capital cost, the performance efficiency of these technologies need to improve even further to compete with fossil fuels. For example, the lower heating value (LHV) of gasoline is around 32 MJ/L, while the LHV of compressed hydrogen gas at 700 bar is about 5.6 MJ/L [12]. Renewable energies

like hydrogen need to compensate for what they lack in energy density with superior performance efficiency in their production and end-usage. In terms of production, about 50 kWh of electricity is required to produce 1 kg of hydrogen from electrolysis [13, 14], and roughly 5 kWh is needed to refine 1 gallon of gasoline [15–17]. On an energy basis, 1 kg of hydrogen ($\sim$120 MJ/kg) has the same energy content as 1 gallon of gasoline ($\sim$120 MJ/gal). Therefore, for the same equivalent energy output, hydrogen requires ten times more electricity input than gasoline. On the other hand, in terms of end-usage, a hydrogen fuel cell is approximately twice as efficient as a gasoline engine in converting the energy to mobility [18]. As such, it appears that hydrogen production efficiency is a major bottleneck that needs to be overcome for hydrogen fuel cars to become a bigger strategy for decarbonizing transportation.

In addition to affordability, energy density, and efficiency, a final complicating issue with utilizing renewables is their intermittent availability. For example, solar irradiance and wind speeds vary throughout the day and can fluctuate seasonally. Likewise, they are unevenly distributed geographically, and sometimes even stranded away from large urban centers. As a result, the temporal and spatial variabilities of solar and wind are asynchronous with consumer energy demands, especially for electricity purposes. Therefore, power from solar and wind is non-dispatchable, meaning that their output cannot be adjusted on-demand by operators. This severely limits their ability to satisfy demand load and introduces curtailment & overgeneration issues. Technologies like batteries, pumped-storage hydroelectricity (PSH), and compressed air energy storage (CAES) [19] attempt to solve these shortcomings by storing renewable energy during periods of excess supply to be utilized during periods of excess demand. Currently, there is about 4.67 TWh of energy storage worldwide for electricity, and PSH is by far the most dominant (96% of total capacity) [20]. Both PSH and CAES are mature technologies with average levelized costs of electricity (LCOE) values of \$0.12/kWh and \$0.16/kWh, respectively [21]. Batteries are still advancing, and current costs are around \$0.30/kWh [21], but expected to decrease as the technology matures and economies of scale are realized [22]. As the amount of renewables entering the energy mix doubles, global energy storage capacity is expected to double as well [20]. It is very clear that energy storage will

play an important facilitating role in the energy transition as more renewables are integrated.

Despite these shortcomings, renewables are already transitioning into the energy picture. First, power from renewables reached 25% of the total global electricity production in 2017 [23], with solar and wind energy representing the fastest growing portion and comprising approximately 5% of the total production [3]. By 2040, renewables are expected to reach over 40% penetration [5], half of which will be made up by solar and wind [3], into global electricity production. In 2019, 64% of the planned new power capacity installation in the United States came from solar and wind alone [24]. Dramatic cost reductions in solar photovoltaic (PV) panels and wind turbines are responsible for the recent surge in renewable electricity. From 2009 to 2017, PV modules became 81% cheaper and turbine prices dropped by almost half [25]. This has caused LCOE values for solar and wind ($0.05/kWh and $0.04/kWh, respectively) to become extremely competitive with those from fossil fuels [26]. However, overgeneration and curtailment are outstanding challenges.

Ethanol production from biomass for use as an advanced biofuel is another area where renewables have made an impact. Global production of ethanol reached over 27 million gallons in 2017, with the United States and Brazil accounting for almost 85% of this amount [27]. Due to government mandates for blending ethanol with gasoline, ethanol was 10% of the United States vehicle fuel consumption [28] and made up 16.4% of the energy consumed for Brazilian transportation [29] in 2017. Corn bushels are the main feedstock for ethanol in the United States, while sugarcane is mostly used in Brazil. Depending on crop economics and other variable operating costs, optimistic prices for ethanol production may be about $1.20/gal from corn [30] and $0.80/gal from sugarcane [31]. On an equivalent energy basis, these values correspond to roughly $50-75/bbl of crude oil, which is competitive with historical prices. Biofuels consumption is projected to grow modestly (2.5% per year [3]) to satisfy about 10% of all ground transportation globally in 2040 [5].

The solar & wind power and cellulosic ethanol examples highlight the promising potential of integrating renewables toward satisfying the growing energy demand and alleviating environmental impact from CO_2. The technical development progress and experience gained so far in these areas position renewables to be well-aligned with the energy trends toward increased power and

chemicals production. In fact, future energy systems need not to be so restrictive in their product output and can instead exploit the synergies that exist in a coordinated production of electricity and chemicals. For example, in lieu of storing energy in batteries, PSH, or CAES, renewable energy can be used to power the production of chemicals. In this way, the energy is chemically stored, instead of electrochemically or physically, in the bonds of the product. This setup not only allows for a flexible production but also utilization because the product can later be consumed as is or converted back to electricity depending on the consumer demand. Renewables are more appropriate for this dual-purposed application because the round-trip process of producing electricity this way from fossil fuels is nonsensical. Likewise, the round-trip efficiency to electricity is too low for fossil fuels to be effective given their carbon intensity. Efficiencies matter less for solar and wind energies because their resource availability is essentially unlimited, free, and carbon-neutral.

Finally, the energy transition is undoubtedly influenced by the shale gas revolution, technological advances in hydraulic fracturing and horizontal drilling for extracting fossil resources from shale formations, that has endowed the United States with an unexpected windfall of natural gas and tight oil. The United States now has the ninth-most proven oil reserves (50 thousand million barrels) and the fifth-largest proven natural gas reserves (308.5 trillion cubic feet) in the world [7]. Compared to values from the 1990s, when worries about peak oil arose during the Gulf War, the United States has expanded its oil reserves by nearly 70%, and natural gas reserves have nearly doubled [7]. The United States is also now the largest crude oil and natural gas producer in the world [32]. This new abundance of cheap natural gas is especially attractive for electricity and chemicals production to convert those processes previously ran on crude oil or coal to natural gas, which is notably cleaner due to less CO_2 emissions (50% less than coal, 20-30% less than oil [33]). Therefore, natural gas is going to play a vital role in the energy transition. Moreover, fossil fuels in general are not disappearing any time soon and are critical energy resources during the transition period, especially for processes that are not easily amended to utilizing renewables like heavy industry. In their optimistic projections, Royal Dutch Shell expects fossil fuels to cover about 15-25% of the global energy consumption in 2100 [4]. Hybrid energy systems that utilize both natural

gas and renewables are likely in the transition period to bridge the gap between the former and latter. While the length and ultimate path of an energy transition are unknown, the indefinite usage of finite resources like fossil fossils is not a sustainable solution.

1.3 A New Digital Era for Energy

As the world prepares itself to confront an energy transition, a concurrent digital transformation is already underway. This digitalization promises to revolutionize the way businesses are monetized through exploiting a vast trove of available data and increased computational power to inform better decision-making [34]. In the energy sector, the advent of this digital era has materialized as having smart sensors installed on production sites that transmit process information to the cloud. Other remote devices such as a personal cellphone or supercomputer are then able to access this data and perform necessary manipulations or computations. Facilitating automated decision-making work flows, this connected network of shared devices is called the Internet of Things (IoT). In process industries, the data from these sensors is utilized in models for real-time optimization, advanced control, predictive maintenance, demand forecasting, production scheduling, and fault detection & diagnosis applications [35–37].

As a real industrial example, ExxonMobil has an ongoing program to collect all operating data from its refineries and chemical plants. This project is called the "ExxonMobil Manufacturing Support Data Lake" [38]. The projected velocity of the incoming data to their high-performing computing environment is up to 1 billion bits per minute. With this "big" data, the company aims to discover new directions for its production sites to run more efficiently and with fewer CO_2 emissions through applying advanced analytics.

However, exploiting "big" data is not restricted to just improving existing operations with IoT, but also very relevant for planning future energy systems. For example, meteorological stations positioned around the United States continuously gather measurements on solar irradiance and wind speed at minute resolutions. This time series data is valuable for designing renewable power systems with storage to ensure unit capacities are properly sized to meet demand load and handle volatilities within a day or across several years. Geological measurements from oil and shale fields

can also intelligently inform drillers how to plan their rig operations across an investment horizon to maximize throughput. Nonetheless, this abundance of data may become a double-edge sword*. On one hand, having more data gives greater possible insight into a phenomena that was poorly understood before or lacks a first-principles explanation. On the other hand, an over-dependence on data for enlightenment is prone to misleading conclusions when bad information is not clearly distinguishable from good data or contaminates the dataset.

Therefore, a core basis of this digital era needs to focus on the development of *appropriate* models that reliably make decisions or draw conclusions from the *right* data. There is an endless number of model candidates to choose from within machine learning, artificial intelligence, and statistics. It is often difficult to make an optimal choice. For each individual energy application, a different model may be more advantageous than another. Unfortunately, this is not an one size fits all approach, and there is little theory to *a priori* predict which model will work best. Tried-and-tested model selection techniques and metrics are utilized to separate out the better performing models, but are not guaranteed to work well. In addition, dimensionality reduction and clustering methods are employed to sift out the important aspects of the data and filter out any noise.

Using "big" data to unlock solutions for existing energy assets is a near-term goal with potentially huge upside for operational gains. Nevertheless, the long-term challenge for a digitalized energy sector is figuring out the right modeling approaches for "big" data to help bring about the energy transition goals. Without appropriate models, the data is meaningless. Correctly analyzing the data will likely help accelerate the energy transition.

1.4 Process Systems Engineering's Role in Addressing the Energy Challenges

The problem with no silver bullet for the energy transition is that there are numerous alternatives to consider without any obvious best path or set of paths forward. Additionally, empirically testing different choices and observing their effects on the global energy infrastructure is not practically possible. And even if this was a feasible option, the economic investment, human capital

*There are also cybersecurity issues with hosting large data centers and cloud computing, but these are not the subject of this dissertation.

cost, and potential repercussions of ill-informed decisions would be too risky to do so. This is the challenge decision-makers involved in the energy transition face. Without a systematic and quantitative methodology (i.e. an *in silico* testing protocol) to navigate all possible options, using an *ad hoc* trial and error approach toward implementing future energy systems is not an effective strategy. There are too many energy system permutations to consider and too many requirements for them to meet for anyone to simply stumble onto the right choices. Moreover, there is not enough time to try everything out. Therefore, mathematical modeling & optimization are needed to inform decision-makers of the more promising courses of action to take toward maneuvering the energy transition and screen out unlikely answers. Similar optimization approaches are also necessary to select the correct data-driven model from a list of other candidate ones.

Process systems engineering (PSE) is a research field within chemical engineering that evolved during the 1960s to reform the heuristic and rule of thumb methods that existed for designing and operating industrial chemical processes. This is not to say that these previous methods were unsuccessful. For example, the industrial process for ammonia production was remarkably developed at BASF in the 1910s without the use of modern computing [39]. Its success, especially as a highly pressurized process for its time, set an important precedent for future chemical engineers to reference in their designs and operations. Instead, the motivation for PSE's development was the desire to unify the heuristics and rule of thumbs and translate them into mathematical language. With a developed mathematical representation of the chemical process, the goal is to methodically identify opportunities to improve the process design and/or production output during operation. Therefore, the theoretical foundation of PSE has been rooted in the mathematical optimization of chemical processes for optimal process design and operation. Examples of PSE's success in optimizing and upgrading the process industry are well-documented [40–44].

Given its strengths, PSE is aptly positioned to tackle the energy transition. First, chemicals production is a sector very familiar to all chemical engineers. The expected growth in chemicals production is likely best addressed by those who are specifically trained for this area. Second, electricity production is also a closely related field since most power will still be generated from

the combustion of fuels to produce thermal heat to drive turbines. These power cycles are typical topics covered in thermodynamics courses in chemical engineering. Third, and most importantly, the systematic consideration of all energy system possibilities and requirements is well-addressed through mathematical optimization, the exact area that PSE specializes in. Likewise, model selection and clustering methods in deriving appropriate machine learning models from data are also optimization problems in disguise. The challenges with energy transition and digitalization are the same problems PSE has aimed to model and address in chemical engineering for many years. No matter the energy system of interest, PSE methodologies can provide optimal answers to guide decision-makers and data modelers without having to exhaustively enumerate all options.

The decisions made during the energy transition are important because they will influence the direction of energy development for the next several decades. For example, many of today's energy systems are products of a bygone era when modern computing and digital technologies were in their infancy. The average age for existing United States petrochemical and power plants is about 30 years [45, 46], while refineries average around 40 years old [47]. Older energy infrastructure is overdue for an update, and new capacity installations are required to keep up with the increased demand and supplant retired plants. The conventional strategy of simply retrofitting older facilities is not sufficient or efficient enough to satisfy the projected energy consumption. Therefore, more capital investment into rebuilding and revamping the energy infrastructure is needed. Moreover, making renewables a priority now ensures that the requisite technology components and expertise are in place for years ahead when we will require more from them. PSE can play a key role in shaping the energy future, through providing non-intuitive insights into how to modernize existing plants and which new projects to pursue, such that the most gain is made out of each investment.

1.5 Dissertation Objectives

Thereby, this dissertation aims to answer some of the outstanding questions regarding the energy transition and digitalization. In particular, the main focus areas here are twofold. First, the optimal design and operation of future energy systems in the transition are unknown. Second, the incorporation of data into the investigation and decision-making process for these systems and the

optimal selection of models from data are underdeveloped. To address these two main objectives, a mathematical optimization and systems approach from PSE is taken to provide insights.

For the first objective, detailed optimization models are built and solve for the following energy systems: (1) natural gas to liquid transportation fuels, (2) coproduction of ammonia and methanol from biomass, and (3) energy carriers supply chain network. These systems represent production processes from natural gas and renewables that may become relevant and important during the energy transition. For the former part of the second objective, a clustering algorithm is created to reduce the complexity of solving a design and scheduling optimization problem with time series data. This algorithm is applied to optimizing a renewable power system with storage to assess its LCOE sensitivity with respect to the clustering. For the latter part of the second objective, a novel bilevel optimization and parametric programming method for optimal model selection is developed to give the exact solution without any approximation.

In next sections, a brief introduction and overview to each topic are given before going into more details in later sections. My contribution to each study is highlighted.

1.5.1 Natural Gas to Liquid Transportation Fuels (GTL)

Although vehicle electrification and fuel economy are improving, the demand for liquid transportation fuels is expected to increase in the near-term future due to growth in personal mobility and commercial services. Therefore, fossil fuels are still needed in the transportation sector to provide these dense fuels that renewables are not yet suited to produce. At the same time, relying on crude oil increases a nation's dependence on other countries for supply, and most major oil producing countries have unstable political environments that are difficult to predict. Developing an alternative to crude oil for transportation, especially for oil poor nations, is important for diversifying a nation's energy portfolio [48] and strengthening national energy security [49].

An attractive feedstock is natural gas because it is abundantly available [7] and inexpensive [50] due to the shale gas revolution. Assuming the current consumption rate, the United States also has enough natural gas to last almost a century [51]. For synthesizing liquid fuels, natural gas is appealing because it has a high hydrogen to carbon ratio, and thus favors a greater yield of

higher hydrocarbons. Several academic [52–54] and industrial [55–57] parties have focused on developing GTL processes that convert natural gas to a syngas intermediate before upgrading to liquid products. Large companies such as Royal Dutch Shell [58] and Sasol Limited [59] have commercial GTL plants operating in Malaysia, Qatar, and South Africa, respectively.

My contribution here to investigate chemical looping as an alternative method for producing syngas from natural gas, as opposed to the traditional steam reforming process, within a process synthesis & global optimization framework. This is first time chemical looping has been studied for use in a GTL process within such a framework, and it is shown to dramatically decrease the cost of production.

1.5.2 Coproduction of Ammonia and Methanol from Biomass (BAM)

Ammonia and methanol are two essential chemicals for sustaining life and the lifestyle many people enjoy. Over 85% of ammonia is utilized as synthetic fertilizer, a form of fixated nitrogen, for agriculture. It is estimated that the world's population would be about half the size it is today without artificial fertilizers to promote crop growth because biological nitrogen fixation would be too slow to feed everyone [39]. Methanol is an important feedstock for producing numerous commodity chemicals such as olefins, aromatics, formaldehyde, and acetic acid that are utilized to make plastics, fibers, solvents, cosmetics, pharmaceuticals, building materials, and other commonly used household products. Because they are vital precursors to other chemicals, demand for ammonia and methanol is expected to increase faster than demand for products from the overall chemicals sector.

At the same time, the productions of ammonia and methanol are both energy and carbon intensive. About 2.9 tons of CO_2 is emitted per ton of ammonia production, and this value is approximately 0.9 tons for methanol. Using the most advanced technologies, both utilize roughly 30 GJ of energy to produce one ton of product. In their synthesis, both require hydrogen which is derived from the same downstream processes. Therefore, there is an opportunity to integrate their production in an attempt to share costs for similar units and decrease the overall cost of production. Likewise, to reduce their carbon footprint in a similar spirit as ethanol does in the transportation

sector, biomass-based production of ammonia and methanol is a promising avenue .

My contribution here is to investigate the coproduction of ammonia and methanol from biomass for the first time using a process synthesis & global optimization strategy. While most industrial production is based on natural gas or coal, biomass is shown to be competitive, especially under carbon-constrained scenarios.

1.5.3 Energy Carrier Supply Chain Network (ECN)

While PSH and CAES are mature utility-scale technologies, they are geographically restrictive in their suitable construction sites. On the other hand, batteries have smaller storage capacities and are more fit for distributed applications. Additionally, at current costs, the scale-up of batteries is prohibitively expensive, and they require rare earth metals that may become inaccessible in the future. Present battery technologies are also not suitable for long-term energy storage. As an alternative, energy can be stored in chemical compounds through renewable energy powered water electrolysis to produce hydrogen, which can be further converted to ammonia and methanol. Storing energy in chemicals is promising because their production is well-studied, easily scalable to large volumes, and benefits from economies of scale. Moreover, chemicals have higher energy density and are geographically flexible in terms of where they can be produced and consumed.

In this way, chemicals act as energy carriers, storing and transporting renewable energy from regions of excess supply to resource-deficient demand areas. Energy carriers are then converted back to electricity on-demand. However, there are trade-offs between which chemicals to choose. Hydrogen is cheaper to produce, but has lower energy density and is more expensive to transport than ammonia or methanol. On the other hand, ammonia and methanol are more expensive to manufacture, but are more energy dense and have cheaper transportation costs. Hydrogen and ammonia are carbon-neutral, while methanol is not. Over the years, there have many separate discussions about a hydrogen, ammonia, or methanol economy [60–62]. Delineations like these are limiting and do not consider the potential interactions that exist among the three.

My contribution here is to investigate the integrated supply chains of hydrogen, ammonia, and methanol for the first time in the context of transporting renewable energy. A case study on Texas

highlights that a supply chain network of energy carriers is competitive with other storage options, especially for long-term applications.

1.5.4 Clustering in the Optimization of Renewable Power Systems

Intermittent solar and wind availabilities pose design and operational challenges for renewable power systems because they are asynchronous with consumer demand. For renewable energies to increase their penetration into the power grid and avoid curtailment & over-generation issues, such as those currently observed in California [63,64], greater investment in energy storage technologies is necessary. However, supply and demand uncertainties pose significant investment risks [65] for renewable power systems with energy storage, which are large capital intensive projects. Likewise, without considering the future operations of a renewable power system simultaneously with its design during the investment planning phase, undersizing or oversizing power and storage capacities become costly miscalculations [66, 67].

To align this supply-demand mismatch, optimization-based design and scheduling models have been developed to minimize the capital and operational costs associated with power production and energy storage. However, hourly time discretization and large time horizons used to describe short- and long-term solar and wind dynamics, demand fluctuations, & price changes significantly increase the computational burden of solving these models. Time aggregation [68] or temporal clustering [69] have been applied to reduce the complexity of these models. The main concept is to decrease the number of unique hours that are modeled through finding patterns or clusters of similar weather, load, and price behavior in the time series data.

My contribution here is the development of a decomposition algorithm based on agglomerative hierarchical clustering (AHC) that separates decision decisions from operational ones, while alleviating the computational burden of solving the design and scheduling optimization. The algorithm is applied to investigate a renewable power system with battery storage in New York City and its LCOE sensitivity to clustering.

1.5.5 Hyperparameter Optimization Through Parametric Programming (HY-POP)

In a digital era of energy, the shift toward machine learning models to aid decision-makers better understand their current operations and forecast the future is accelerating. Given all the modeling choices, optimal model selection is vital to realizing this digitalized economy. Fitting a machine learning model often requires presetting parameter values (hyperparameters) that control how an algorithm learns from the data. Selecting an optimal model that minimizes error and generalizes well to unseen data becomes a problem of tuning or optimizing these hyperparameters.

Typical hyperparameter optimization strategies involve discretizing the parameter space and implementing an iterative search procedure to approximate the optimal hyperparameter and model selection through k-fold cross-validation. Bayesian optimization is another common method to tune hyperparameters. This approach still involves iteratively exploring the hyperparameter space. A probabilistic model approximates the optimal hyperparameter by assigning probability values to its location and selecting the one with the highest probability. These existing methods can suffer from selecting suboptimal models because of the inexact parameter space discretization.

My contribution here is the development of an exact solution to the hyperparameter optimization problem for machine learning algorithms that are formulated as linear or quadratic programming (LP/QP) models. The optimal model selection is guaranteed and obtained through bilevel optimization and parametric programming without any approximation. The method is then applied on example data for verification.

2. ENHANCING GTL PROCESSES THROUGH CHEMICAL LOOPING FOR SYNGAS PRODUCTION*

2.1 Motivation

Interest in GTL processes is growing due to the promise of transforming natural gas, a suddenly abundant and inexpensive feedstock due to the shale gas revolution, into valuable liquid products like gasoline, diesel, and kerosene. Adoption of a GTL process in rich gas but oil poor countries is one application. Another one is to monetize stranded natural gas that otherwise would be left unused. However, the economics must be competitive with crude oil prices. GTL processes first convert natural gas to a syngas intermediate before upgrading it to liquid products. Syngas production remains the most expensive and energy intensive part of any GTL process. Chemical looping is an alternative to traditional reforming methods for producing syngas that may lower the economic and energy cost of the GTL process.

2.2 Background

A previous work proposed a comprehensive process superstructure of natural gas conversion to liquid transportation fuels [54]. A top-down view of a GTL refinery is shown in Figure 2.1. This work incorporates two chemical looping technologies, one based on a NiO and one based on Fe_2O_3, into the previous superstructure as process alternatives for syngas generation to evaluate their performance against traditional reforming. The GTL refinery consists of the following sections: (i) natural gas conversion, (ii) syngas cleaning, (iii) hydrocarbon production and upgrading, (iv) light gas recycle, (v) hydrogen/oxygen production, and (vi) wastewater treatment. These sections, their units, and modeling equations are described in the paper's supplementary material [70]. Emphasis is placed here on modeling chemical looping, the new component in this work. Some background on chemical looping is provided before moving onto the mathematical modeling of

*Reprinted from "Enhancing natural gas-to-liquids (GTL) processes through chemical looping for syngas production: Process synthesis and global optimization" by Tso, W.W. and Niziolek, A.M. and Onel, O. and Demirhan, C.D. and Floudas, C.A. and Pistikopoulos, E.N., Computers & Chemical Engineering, 2018, Vol. 113, pp 222-239, with permission from Elsevier and Copyright Clearance Center.

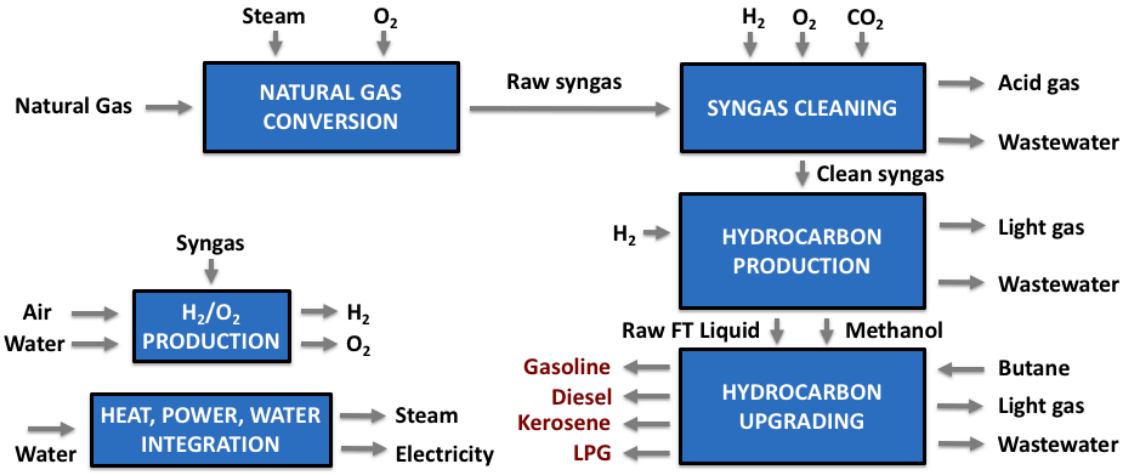


Figure 2.1: An overview of the GTL refinery.

the reactors and global optimization to solve the GTL process synthesis problem.

Chemical looping has been predominantly studied as an advanced combustion system for electric power generation with *in situ* CO_2 capture [71–74]. The overall combustion is carried out in two interconnected reactors, an air reactor (AR) and a fuel reactor (FR), that separates the reaction into oxidation and reduction parts (Figure 2.2). Circulating metal oxides between the two reactors serve as oxygen carriers. In the FR, metal oxides react with a carbonaceous fuel, producing CO_2 & H_2O and reducing the metal oxides to a lower oxidation state. In the AR, reduced metal oxides are fully oxidized with air and then recycled back to the FR. By this loop construction, metal oxides provide oxygen to the fuel without needing additional air separation. As a result, chemical looping combustion (CLC) avoids directly contacting air with the fuel and facilitates CO_2 capture since CO_2 & N_2 are in separate streams.

Chemical looping is adjusted for syngas production by either decreasing the air-to-fuel ratio or the metal oxides-to-fuel molar ratio to favor syngas conversion instead of combustion [75–78]. From lowering the amount of available oxygen, partial oxidation of the fuel occurs [79]. Energy efficiency and exergy is greater for chemical looping compared to autothermal reforming (ATR)

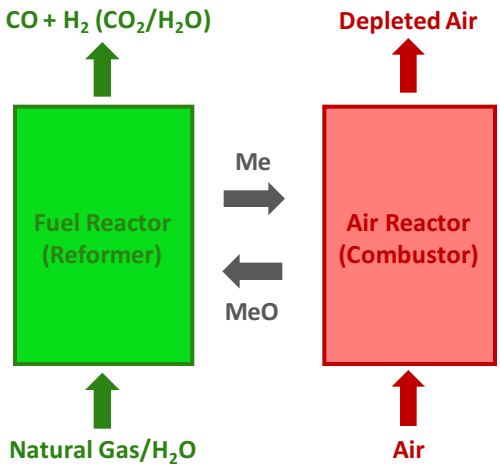


Figure 2.2: A simplified block diagram of the chemical looping process with natural gas as the carbonaceous fuel.

and steam reforming (SMR), the traditional methods for syngas production [80, 81]. In reforming, excess O_2 and/or H_2O are co-fed with natural gas to improve gas conversion, prevent coking, and sustain catalyst activity. This creates extra CO_2 and H_2O in the effluent, reducing syngas purity and process efficiency. Additional conditional units become necessary to adjust the syngas composition before downstream processing. Chemical looping produces a higher concentration of CO and H_2, without requiring excess co-fed reactants [82, 83]. This reduces the capacities of downstream process units and offers potential capital cost savings. Chemical looping is also operable at lower pressures than reforming, which facilitates greater natural gas conversion due to Le Châtelier's principle. Lastly, chemical looping systems have smaller land impact [74]. For example, autothermal reforming requires pure oxygen from a costly air separation unit (ASU), which adds another unit and more cost to the process. Chemical looping has inherent oxygen separation by design.

From a thermodynamics analysis, the modified Ellingham diagram shows that CeO_2, NiO, and Fe_2O_3 are more favorable metal oxides for syngas production from natural gas [74]. Studies have shown that severe carbon deposition and slow kinetics hindered the practical usage of CeO_2 [84, 85]. Results have been more promising for NiO and Fe_2O_3 with several independent

research groups [75–78, 86] demonstrating experimental and pilot-scale operations. de Diego et al. [83] and Ortiz et al. [87] reported that NiO/Al_2O_3 in chemical looping reforming (CLR) could achieve methane conversions to syngas at near thermodynamic equilibrium under different operating temperatures, pressures, and oxygen carrier-to-fuel molar ratios. Luo et al. [82] developed a shale gas-to-syngas (STS) process using Fe_2O_3/TiO_2 that operates close to the thermodynamic equilibrium for various process conditions. Deshpande et al. [88] also verified feasibility for higher pressurized STS systems.

2.3 Mathematical Modeling of Chemical Looping Reactors

While there are very detailed models for chemical looping reactor designs, such as computational fluid dynamics with reaction kinetics information [89,90], they are impractical to implement within a process synthesis & global optimization framework. This is because these models have a large number of variables and nonlinear equations that require several hours to simulate and the process synthesis model itself is a computationally expensive mixed-integer nonlinear programming (MINLP) model that is solved many times within a branch-and-bound algorithm. Instead, simplified input-output models of chemical looping reactors are developed here to closely approximate the underlying phenomena with as little complexity as appropriate.

2.3.1 Fuel Reactor Model

In accordance with the experimental results, thermal equilibrium is assumed for the FR model. The inlet streams to the FR are CH_4, H_2O, and an oxidized metal oxide (MeO_x) coupled with a support material. The incoming metal oxides are fully oxidized from the AR, and the oxygen-containing support materials (SuO) are inert. The outlet vapor effluent is comprised of H_2, H_2O, CO, CO_2, and CH_4. The leaving solids, reduced metal oxide (MeO_y) and pure metal (Me), and their support materials are recycled back to the AR. Carbon formation is assumed to be negligible by constraining the inlet $\frac{[O]}{[C]} \geq 1$.

Let n represent the number of solid products leaving the FR. NiO is fully reduced to Ni or stays unreacted, while Fe_2O_3 is reduced to either Fe, FeO, or Fe_3O_4. The support materials for NiO and

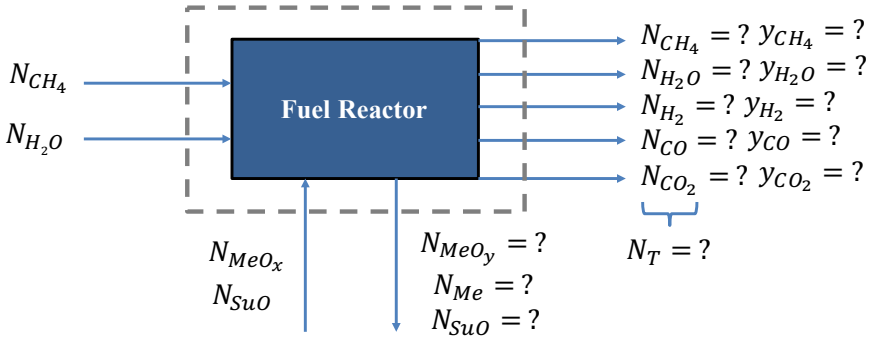


Figure 2.3: The known and unknown streams entering and leaving the fuel reactor.

Fe_2O_3 are Al_2O_3 and TiO_2, respectively. As such, n = 3 for NiO and n = 4 for Fe_2O_3. There are n+11 unknown variables: n+5 molar species flows (N_s), 5 vapor species mole fractions (y_s), and 1 total molar vapor flow (N_T). To fully specify the system, n+11 equations are needed. The described scenario for the FR model is shown in Figure 2.3. After writing down atom balances for C, H, O, Me, and Su, mole fraction equations, total mole balance, and two thermodynamic equilibrium equations for methane reforming (Eq. 2.1) and water-gas shift (Eq. 2.2), there are n-2 degrees of freedom left. The CLR process has one degree left, while for the STS process has two.

$$CH_4 + H_2O \rightleftharpoons CO + 3H_2 \tag{2.1}$$

$$CO + H_2O \rightleftharpoons CO_2 + H_2 \tag{2.2}$$

To close these last degrees of freedom, the conversion of MeO_x into MeO_y and Me is specified. For the CLR process, NiO conversion to Ni is calculated using a parameter estimation; and for the STS process, Fe_2O_3 conversion to Fe, FeO, or Fe_3O_4 is determined using disjunctive programming.

2.3.1.1 Parameter Estimation for NiO Conversion

To determine the NiO conversion, the vapor effluent compositions in Figure 6 from de Diego et al. [83] are compared to the theoretical output from a parameter estimation model described below.

Indices. The following indices are used:

a: Atom index

s: Species index

t: Experimental dataset

Sets. We define the set of all atoms, A, as follows:

$a \in A = \{\text{C, H, O, Ni, Al}\}$

The set of all species, S, is defined as:

$s \in S = \{\text{CH}_4, \text{H}_2, \text{H}_2\text{O}, \text{CO}, \text{CO}_2, \text{NiO}, \text{Ni}, \text{Al}_2\text{O}_3\}$

For convenience, the set of vapor species, S_v, is given as:

$S_v = \{\text{CH}_4, \text{H}_2, \text{H}_2\text{O}, \text{CO}, \text{CO}_2\}$

The set I is the total number of experimental datasets:

$t \in I = \{1, 2, 3,...,21\}$

Parameters. The known model parameters are:

$E_{s,a}$: number of atom(s) a in species s

$N_{s,t}^{in}$: inlet molar flow of species s in experimental dataset t

$y_{s,t}^{out}$: effluent mole fraction of vapor species $s \in S_v$ in experimental dataset t

$K_{eq,SR}$: steam reforming equilibrium constant

$K_{eq,WGS}$: water-gas shift equilibrium constant

P_f: pressure of the CLR fuel reactor

Variables. The model has the following variables:

$N_{s,t}$: predicted outlet molar flow of species s for experimental dataset t

$N_{T,t}$: predicted effluent total molar flow of vapor species $s \in S_v$ for experimental dataset t

$y_{s,t}$: predicted effluent mole fraction of vapor species $s \in S_v$ for experimental dataset t

ED_t: Euclidean distance (error metric) for experimental dataset t

Avg_{err}: averaged Euclidean distance across all experimental datasets

X_{NiO}: fitting variable for NiO conversion to Ni

Constraints. Equations describing the model are shown below:

Atom balances:

$$\sum_{s \in S} N_{s,t}^{in} E_{s,a} = \sum_{s \in S} N_{s,t} E_{s,a} \qquad \forall a \in A \ , \ \forall t \in I \tag{2.3}$$

Total mole balance for vapor species:

$$\sum_{s \in S_v} N_{s,t} = N_{T,t} \qquad \forall t \in I \tag{2.4}$$

Vapor mole fractions:

$$N_{s,t} = y_{s,t} N_{T,t} \qquad \forall s \in S_v \ , \ \forall t \in I \tag{2.5}$$

Steam reforming:

$$K_{eq,SR} = \frac{y_{CO,t}\, y_{H_2,t}^3\, P_f^2}{y_{CH_4,t}\, y_{H_2O,t}} \qquad \forall t \in I \tag{2.6}$$

Water-gas shift:

$$K_{eq,WGS} = \frac{y_{CO_2,t}\, y_{H_2,t}}{y_{CO,t}\, y_{H_2O,t}} \qquad \forall t \in I \tag{2.7}$$

Fit for NiO conversion:

$$N_{Ni,t} = X_{NiO} N_{NiO,t}^{in} \qquad \forall t \in I \tag{2.8}$$

Error between experimental and predicted vapor mole fractions:

$$ED_t = \sqrt{\sum_{s \in S_v} (y_{s,t} - y_{s,t}^{out})^2} \qquad \forall t \in I \tag{2.9}$$

Objective function. The objective of the parameter estimation model is to minimize the overall averaged error between the experimental and model predicted values (Eq. 2.10).

$$Avg_{err} = \frac{\sum_{t \in I} ED_t}{|I|} \tag{2.10}$$

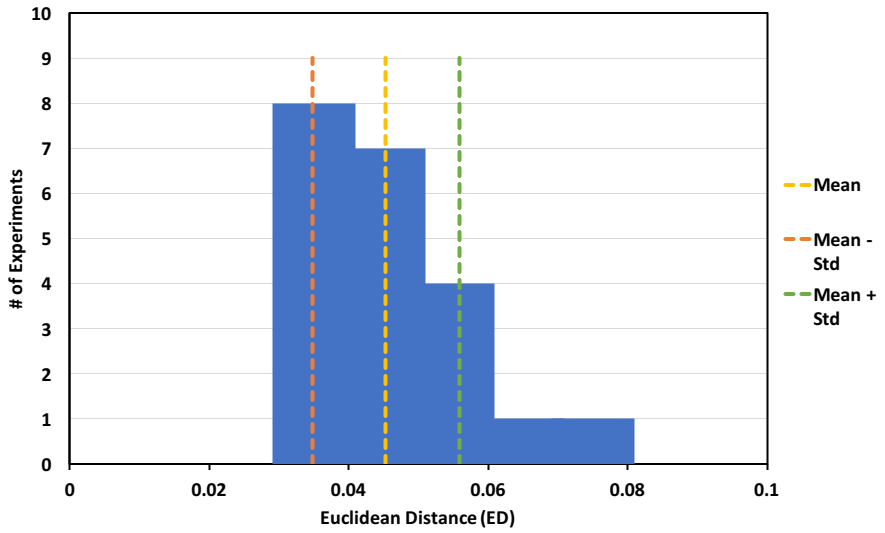


Figure 2.4: Error distribution for the parameter estimation of NiO conversion in the CLR fuel reactor. Experimental data gathered from de Diego et al. [83]

Thereby, the final model formulation becomes:

$$\begin{aligned} &\min \ Avg_{err} \\ &\text{s.t. Eqs. 2.3-2.9} \end{aligned} \tag{2.11}$$

The parameter estimation is essentially a linear regression model with nonlinear constraints, which formulates as a nonlinear programming (NLP) problem. The model is solved to global optimality using ANTIGONE in GAMS. The optimal X_{NiO} is calculate as 0.923 with an overall averaged error of 4.53% and standard deviation of 1.06%. A histogram of the individual errors is shown in Figure 2.4 with the mean and one standard deviation above and below the mean. More than 85% of the datasets have an error less than 5.5%, and sixteen have an error that is within one standard deviation of the mean. By specifying the value for X_{NiO}, the last degree for the CLR FR is satisfied.

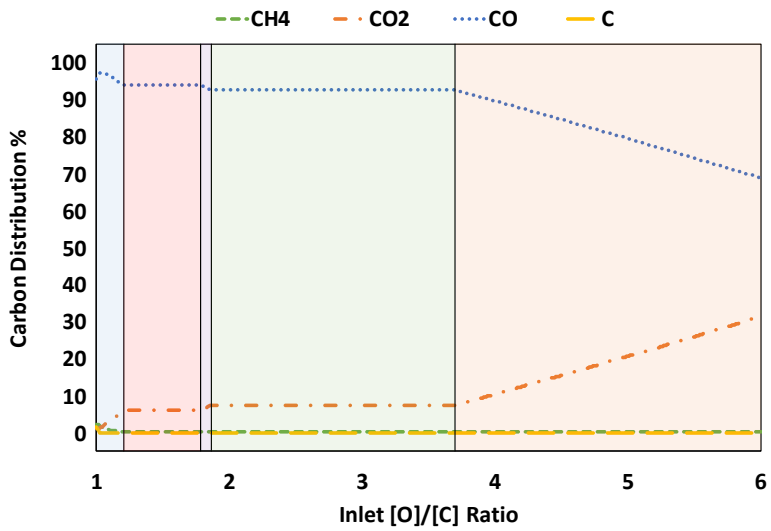


Figure 2.5: Distribution % of carbon products across different inlet $\frac{[O]}{[C]}$ ratios for Fe_2O_3/TiO_2 oxidizing CH_4 at 900 °C and 1 bar. Data is from a RGIBBS simulation in Aspen Plus and is in agreement with Figure 3 reported in Luo et al. [82]

2.3.1.2 Disjunctive Programming for Fe_2O_3 Conversion

A similar parameter estimation analysis is not applicable for Fe_2O_3 because there is a very limited set of data points provided by Luo et al. [82]. Instead, we use the RGIBBS module in Aspen Plus V8.8 to simulate the STS FR at thermodynamic equilibrium and generate input-output values to model after. The incoming solids to the STS FR are assumed to have a Fe:Ti ratio of 2. Depending on the inlet $\frac{[O]}{[C]}$ ratio, Fe_2O_3/TiO_2 reduces to different amounts of its reduction products. There are five distinct zones of $\frac{[O]}{[C]}$ ratios corresponding to different product compositions of carbon species (Figure 2.5) and reduced metal oxides (Figure 2.6). Luo et al. [82] and Deshpande et. al. [88] showed that Fe/TiO_2, $FeTiO_3$, and Fe_2TiO_4 are the reduced solid species that exist at equilibrium. These oxides are composite analogues to Fe, FeO, and Fe_3O_4 with slightly different thermodynamic behavior. Two degrees of freedom still remain for the STS reactor because the composite iron oxides can be related to the original iron oxides through the following mathematical

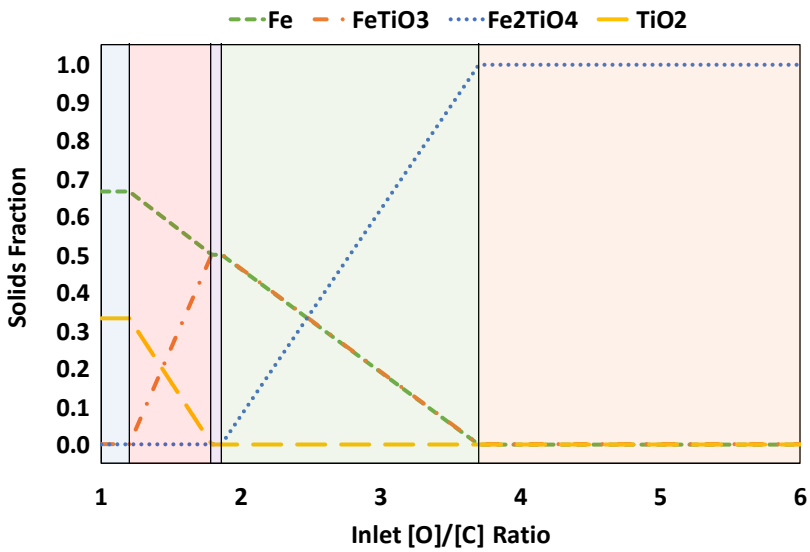


Figure 2.6: Fraction of solid products across different inlet $\frac{[O]}{[C]}$ ratios for Fe_2O_3/TiO_2 oxidizing CH_4 at 900 °C and 1 bar. Data is from a RGIBBS simulation in Aspen Plus.

relationships (these are not chemical reactions):

$$FeO + TiO_2 = FeTiO_3 \tag{2.12}$$

$$Fe + Fe_3O_4 + 2TiO_2 = 2Fe_2TiO_4 \tag{2.13}$$

The disjunctive program models the linear conversion behavior within each zone instead of capturing the overall nonlinearity in a single unified expression. The discretization of these zones by their $\frac{[O]}{[C]}$ values are used to formulate the disjunctive constraints, which determine the set of active and inactive equations describing the iron oxide conversion in each zone. The Big-M formulation, where an appropriately large M value relaxes constraints that are inactive for a given zone, is utilized. The disjunctive constraints for each zone share the following characteristics:

$$N^{out}_{solid} \leq a_{i,solid} N^{in}_{s} + b_{i,solid} + (1 - \lambda_i) M_{1,i,solid} \quad \forall i = 1, 2, \ldots, 5\ ,\ s = Fe_2O_3 \text{ or } TiO_2, \tag{2.14}$$

$$solid = Fe, FeTiO_3, Fe_2TiO_4, \text{ or } TiO_2$$

$$N_{solid}^{out} \geq a_{i,solid} N_s^{in} + b_{i,solid} + (\lambda_i - 1) M_{2,i,solid} \quad \forall i = 1, 2, \ldots, 5\ ,\ s = \text{Fe}_2\text{O}_3 \text{ or } \text{TiO}_2, \text{ (2.15)}$$

$$solid = \text{Fe, } \text{FeTiO}_3, \text{ } \text{Fe}_2\text{TiO}_4, \text{ or } \text{TiO}_2$$

$$\lambda_i = \begin{cases} 1, & \text{if } LB_i \leq \left(\frac{[O]}{[C]}\right)_{in} \leq UB_i \\ 0, & \text{otherwise} \end{cases}, \qquad 1 \leq i \leq 5 \tag{2.16}$$

$$\sum_{i=1}^{5} \lambda_i = 1 \tag{2.17}$$

where N_s^{in} is the inlet molar flow of solid s, N_{solid}^{out} is the outlet molar of the product $solid$, $a_{i,solid}$ and $b_{i,solid}$ are specific constants defined according to the zone and product, $M_{1,i,solid}$ and $M_{2,i,solid}$ are sufficiently large values relaxing the constraints to be redundant when λ_i = 0, and λ_i are binary variables describing the disjunction logic according to the lower bound (LB_i) and upper bound (UB_i) of each $\frac{[O]}{[C]}$ zone.

Because there are two degrees of freedom remaining, the disjunctive constraints are written for only two of the four solid products (Fe, FeTiO_3, Fe_2TiO_4, TiO_2). Therefore, Eqs. 2.14 and 2.15 are each written twice for a given zone i, once for each selected product $solid$. The disjunctive constraints are based upon the inlet amounts of Fe_2O_3, TiO_2, or both depending on which solid products are selected for the disjunction. When λ_i = 1, $N_{solid}^{out} = a_{i,solid} N_s^{in} + b_{i,solid}$; when λ_i = 0, the bounds on N_{solid}^{out} are relaxed. $M_{1,i,solid}$ and $M_{2,i,solid}$ are defined uniquely for each product $solid$ and are not necessarily the same value. Explicit formulations for Eqs. 2.14 and 2.15 are found in the journal publication [70].

In addition, the $\frac{[O]}{[C]}$ ratio is a nonlinear term that discretizes the STS operation into zones using λ_i (Eq. 2.16) with only one zone allowed to be active (Eq. 2.17). Eq. 2.16 is rewritten as:

$$\sum_{i=1}^{5} LB_i N_{[C]}^{in} \lambda_i \leq N_{[O]}^{in} \leq \sum_{i=1}^{5} UB_i N_{[C]}^{in} \lambda_i \tag{2.18}$$

where $N_{[C]}^{in}$ is the inlet molar flow of carbon species and $N_{[O]}^{in}$ is the inlet molar flow of active oxygen species. Because of Eq. 2.17, only one LB_i and UB_i is active in Eq. 2.18, constraining

$N^{in}_{[O]}$ to be between $LB_i N^{in}_{[C]}$ and $UB_i N^{in}_{[C]}$. In this formulation, $N^{in}_{[C]}\lambda_i$ is a bilinear term comprised of a continuous variable and binary variable. Eq. 2.18 is reformulated into the following linear constraints:

$$\sum_{i=1}^{5} LB_i H_i \leq N^{in}_{[O]} \leq \sum_{i=1}^{5} UB_i H_i \tag{2.19}$$

$$N^{in}_{[C]} - N^{in,up}_{[C]}(1-\lambda_i) \leq H_i \leq N^{in}_{[C]} - N^{in,lo}_{[C]}(1-\lambda_i) \qquad \forall i = 1,2,\ldots,5 \tag{2.20}$$

$$N^{in,lo}_{[C]}\lambda_i \leq H_i \leq N^{in,up}_{[C]}\lambda_i \qquad \forall i = 1,2,\ldots,5 \tag{2.21}$$

When λ_i = 1, H_i is constrained to equal $N^{in}_{[C]}$ with redundant lower and upper bounds ($N^{in,lo}_{[C]}$ and $N^{in,up}_{[C]}$). Likewise, when λ_i = 0, H_i is constrained to equal zero with superfluous bounds from Eq. 2.20. The disjunctive programming model for Fe_2O_3 conversion is fully described using Eqs. 2.14-2.15, 2.17 and 2.19-2.21.

The predicted values are compared to the experimental data described by Luo et al. [82] in Table 2.1, and the error is calculated using Eq. 2.9. The error for sub-pilot condition # 1 is very small (1.08%), but the other two cases have slightly greater errors. Due to the limited data, the quality of the overall averaged error (4.70%) is unknown. More experimental datasets are needed to compare with. However, the disjunctive programming model is sufficient enough to describe the STS FR for subsequent use in the process synthesis. Both Luo et al. [82] and Deshpande et al. [88] concluded that since Fe_2O_3/TiO_2 has very fast kinetics, their experimental results were consistent enough with the thermodynamic calculation to assume equilibrium in process simulations.

2.3.2 Air Reactor Model

The metal oxides (MeO_x) leaving the AR for the CLR and STS processes are assumed to be fully oxidized, since this is what is experimentally observed [82, 83, 87, 88]. Me, MeO_y, SuO, and air (primarily comprised of N_2 and O_2 with trace amounts of CO_2 and Ar) are inlet streams to the AR. Air oxidizes Me and MeO_y to MeO_x, which is recycled back to the FR, and the oxygen-depleted air is vented back to the environment.

A degrees of freedom analysis is also used to build an input-output model for the AR (Fig-

Table 2.1: Comparison between disjunctive model for Fe_2O_3/TiO_2 conversion and experimental results from Luo et al. [82] Experimental vapor compositions were not reported in their work, but are derived here using the given information.

Data	Bench Scale	Model Eq.	Sub-Pilot # 1	Model Eq.	Sub-Pilot #2	Model Eq.
Temp. (°C)	960		975		900	
Pressure (atm)	1		1		1	
Inlet $\frac{[O]}{[C]}$	2.5		2.2		2.8	
CH_4 conversion (%)	>99	99.93	>99.9	99.95	>99.9	99.76
H_2 : CO	1.85	1.94	1.97	1.93	1.85	1.96
CO : CO_2	9.9	13.3	11.8	13.6	9.1	12.5
Syngas selectivity (%)	86.3	91.0	91.3	90.9	85.6	91.5
Carbon deposition (mol C per mol CH_4)	<0.005	0	0.015	0	0.03	0
y_{CH_4}	0.003	~ 0	~ 0	~ 0	~ 0	0.001
y_{CO_2}	0.030	0.023	0.026	0.023	0.033	0.025
y_{CO}	0.302	0.310	0.307	0.310	0.300	0.308
y_{H_2O}	0.106	0.067	0.061	0.068	0.111	0.061
y_{H_2}	0.558	0.600	0.605	0.599	0.555	0.605
ED	0.0584		0.0108		0.0718	

ure 2.7). Though there are n solid reactants entering the AR, there are only two solid products leaving due to full oxidation. All incoming nickel solids are oxidized to NiO, and all incoming iron solids are oxidized to Fe_2O_3. Support materials are inert. In the AR, there are 6 unknown molar species flows. To fully specify the system, 6 equations are needed. Conveniently, there are 6 equations from the atom balances for C, H, O, Ar, Me, and Su. Therefore, under the full oxidation assumption, the AR is fully specified and no additional equations are required.

2.4 Chemical Looping Configurations in a GTL Process

Chemical looping is an alternative to reforming options for converting natural gas to syngas, which is upgraded to liquid fuels in the GTL process. Detailed flowsheets of the CLR and STS process are shown in Figures 2.8 & 2.9. Besides the shared commonalities in the FR and AR setup, the CLR and STS process differ from each other as shown in Table 2.2. Pressure variations are included for the chemical looping processes because of the natural gas expansion and syngas compression immediately before and after the FR. Operating the chemical looping FRs at atmospheric pressure for liquid fuels synthesis, the setup that has been suggested by much of the literature, may

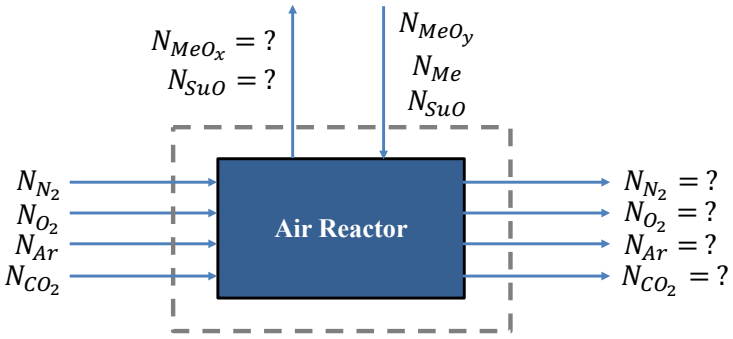


Figure 2.7: The known and unknown streams entering and leaving the air reactor.

Table 2.2: The operating conditions considered for the CLR and STS processes.

Operation	CLR	STS
Oxygen carrier	NiO/Al_2O_3	Fe_2O_3/TiO_2
FR Temps. (°C)	800 | 850 | 900	900 | 950 | 1000
FR Pressures (bar)	1 | 5 | 10	1 | 5 | 10
AR Temp. (°C)	950	1200
AR Pressure (bar)	1 | 5 | 10	1 | 5 | 10
Max inlet H_2O : C ratio to FR	0.5	0.2
Max inlet MeO_x : C ratio to FR	3.5	6

not be optimal.

Expanding the inlet natural gas all the way down to 1 bar (or atmospheric pressure) and later compressing the outlet syngas back up to the higher pressures required for downstream process units could be an ineffective use of energy. This sequence is referred to as a low pressure operation. An intermediate (5 bar) or high pressure (10 bar) operation for the chemical looping process might be more advantageous. However, higher pressures also have an adverse effect on the methane reforming equilibrium (Eq. 2.1) due to Le Châtelier's principle. Therefore, there are trade-offs among these different operational modes that need to be investigated.

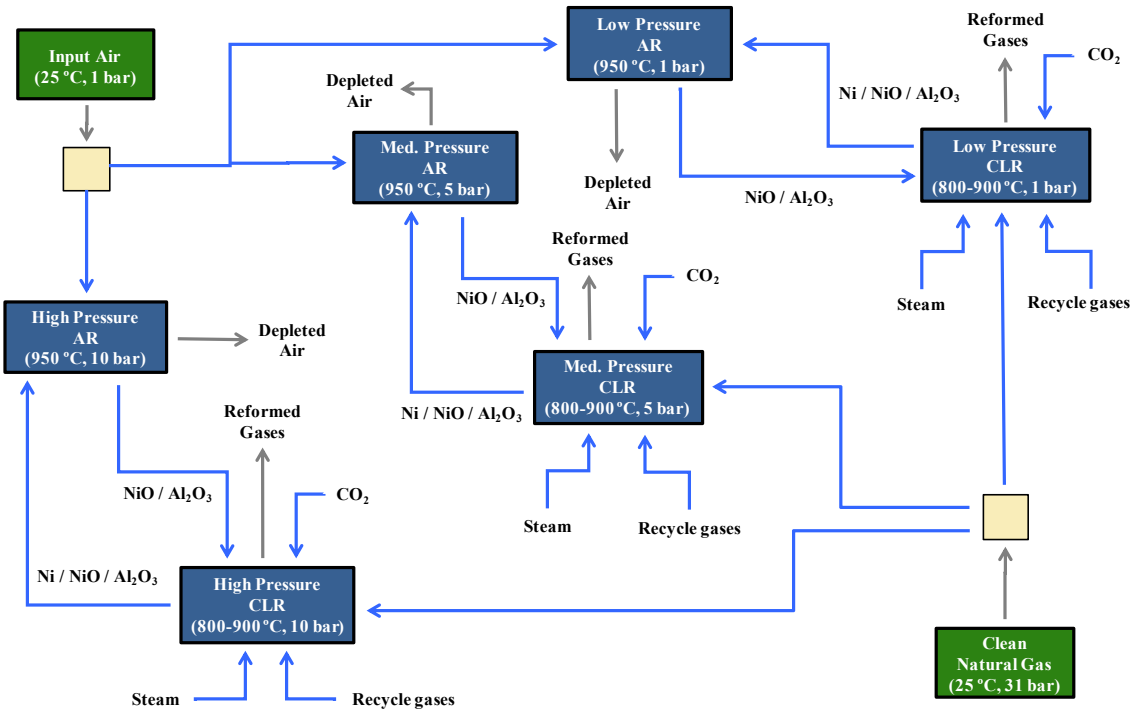


Figure 2.8: Chemical looping reforming (NiO) process.

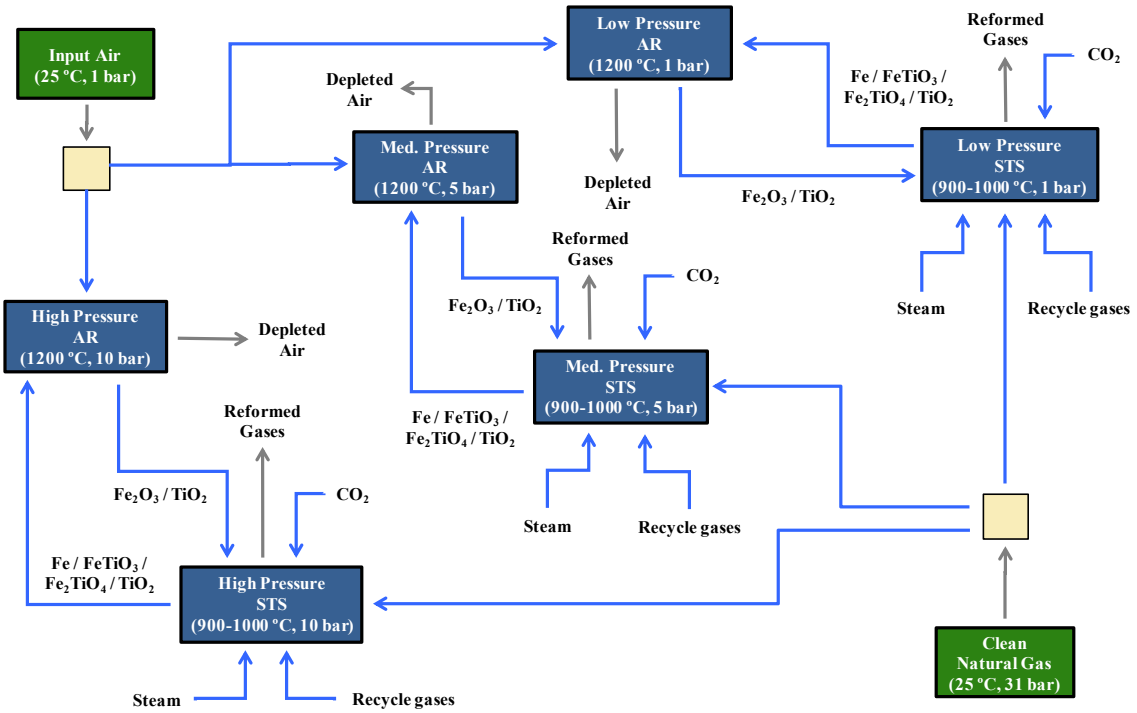


Figure 2.9: Shale gas-to-syngas (Fe_2O_3) process.

2.5 GTL Case Studies

To study the effect of chemical looping for natural gas conversion to liquid fuels, the CLR and STS processes are incorporated as alternatives into a previously developed GTL process superstructure [54]. Information about cost estimations for the process units is found in the published paper [70]. Detailed model equations for other units besides chemical looping are also found in the publication [70]. The objective function of the GTL process (Eq. 2.22) is to minimize the overall cost of liquid fuels production.

$$\min \sum_{f \in Feed} Cost_f + Cost^{El} + Cost^{Seq} + \sum_{u \in U_{Inv}} Cost_u^U - Sales_{LPG} \tag{2.22}$$

Natural gas and fresh water are the considered feedstock costs ($Cost_f$). Electricity cost ($Cost^{El}$) is negative if excess electricity is produced as a byproduct and positive elsewise. A cost ($Cost^{Seq}$) may be paid to sequester CO_2 to meet greenhouse gas (GHG) emission constraints. By summing over the individual unit costs ($Cost_u^U$), the total investment cost of the GTL process is obtained. Finally, the sales of LPG ($Sales_{LPG}$) as a byproduct are also included. Each of these terms is normalized with respect to the total energy of liquid fuels production ($/GJ).

2.5.1 Global Optimization Strategy

The complete mathematical model describing the GTL process synthesis is a large-scale nonconvex MINLP with >30,000 constraints >25,000 continuous variables, and >50 binary variables. Since commercial solvers struggle to find a feasible solution for a problem of this size, a customized deterministic branch-and-bound algorithm (Figure 2.10) is implemented to solve the MINLP to global optimality. The overall problem is decomposed into a lower bound problem and an upper bound problem. The algorithm's aim is to close the gap between the solutions to these two subproblems and converge toward the global optimum.

Tight linear underestimators are used to relax nonlinear expressions (piecewise linearization for concave terms and McCormick envelopes for bilinear & higher terms), forming a mixed-integer

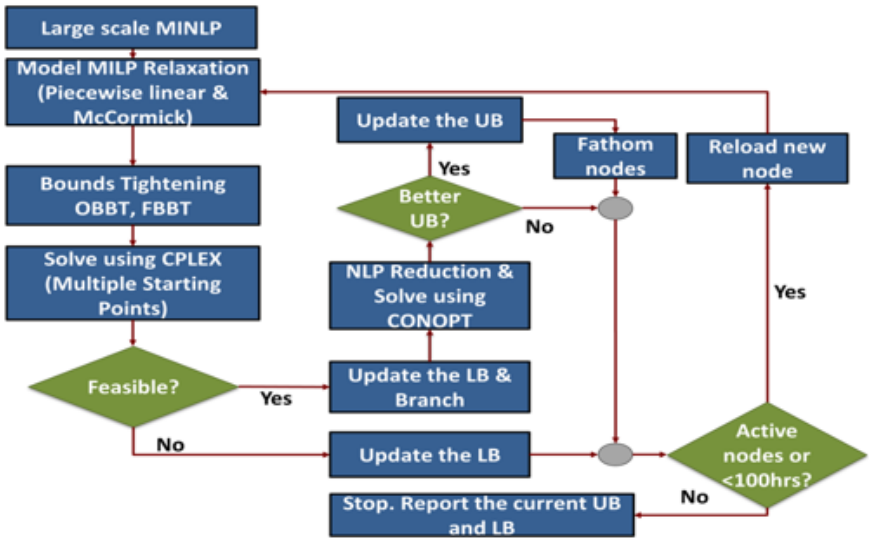


Figure 2.10: Branch-and-bound algorithm to solve MINLP.

linear program (MILP) as the lower bound problem. The MILP is solved using CPLEX, and this solution is a lower bound on the original MINLP. From the lower bound solution, a pool of initial starting points is also generated. A nonlinear program (NLP) is created as the upper bound problem from fixing binary variables at these starting points. The NLP is solved using CONOPT, and this solution is an upper bound on the original MINLP.

Optimality-based bounds tightening (OBBT) and feasibility-based bounds tightening (FBBT) routines are also performed to tighten the bounds on continuous variables. In constructing the branch-and-bound tree, the heuristic is to branch on the continuous variable involved in a nonlinear term with the largest relaxation error. When iterating through the tree, the upper bound to the MINLP is updated if the NLP solution at a given node is lower than the current best. Nodes with a lower bound solution that is ϵ-greater than the upper bound are fathomed. The algorithm is run until the time limit has expired or all the nodes in the tree have been explored, whatever occurs first. The final solution to the MINLP is the best upper bound that is found. For a more detailed discussion about global optimization theory and algorithms, several textbooks are referred [91,92].

2.5.2 Description of Case Studies

Sixteen case studies are investigated to compare chemical looping technologies against reforming for natural gas conversion and their effects on overall GTL process economics. Four different plant sizes (1, 10, 50, and 100 thousand barrels per day) are studied, and four set of case studies are performed at each plant size: (i) ATR is used for natural gas conversion, (ii) SMR is used for natural gas conversion, (iii) CLR is used for natural gas conversion, and (iv) STS is used for natural gas conversion. The case studies are represented as $[T][SC]$, where T is the technology restriction (A: ATR, S: SMR, N: CLR, and F: STS) and SC is the plant scale in thousand barrels per day (kBPD). For example, A10 represents a 10 kBPD GTL plant that uses ATR for natural gas conversion.

The liquid fuels output is constrained to match the U.S. demand (66.6% gasoline, 21.5% diesel, and 11.9% kerosene by volume [93]). The GTL process is also constrained not to produce more GHG emissions than typical petroleum refining based upon a well-to-wheel lifecycle analysis from the GREET model [94]. If electricity is a required input, greenhouse (GHG) emissions from electricity are added to the GTL process emissions; if electricity is a produced output, GHG emissions are discounted from the total. Only the key economic results are discussed here. For further detailed results, please refer to the published paper [70].

2.5.3 Overall Cost Breakdown

A lower overall production cost corresponds to a more profitable plant. Break-even oil prices (BEOP), measured in dollars per barrel, translate the production cost into an estimate of when the GTL process becomes competitive with petroleum refining. The overall cost breakdown of each case study is shown in Table 2.3. Across all the plant sizes, ATR is the best performing reforming process and CLR is the worst performing chemical looping process. In the worst case, CLR offers about 25% cost savings over ATR. In the best case, STS offers about 30-40% cost savings over SMR. By switching to chemical looping from reforming, cost savings of 25-40% are observed. This suggests that the GTL process economics could be greatly enhanced from utilizing chemical

Table 2.3: Overall cost ($/GJ) of the GTL refinery with lower bound and optimality gap reported for each case study.

Case Study	A1	S1	N1	F1	A10	S10	N10	F10
Natural Gas	7.93	8.11	7.35	7.91	8.15	8.11	8.97	9.27
Water	0.02	0.06	0.01	0.01	0.02	0.03	0.02	0.02
Investment	14.45	14.44	10.42	10.30	6.15	6.93	5.26	5.12
CO_2 TS&M	0.00	0.02	0.00	0.00	0.00	0.01	0.00	0.00
O&M	3.81	3.81	2.75	2.72	1.62	1.83	1.39	1.35
Electricity	-0.84	-0.44	-0.57	-1.93	-1.14	-0.48	-3.22	-3.98
LPG	-0.84	-0.84	-0.84	-0.84	-0.84	-0.84	-0.84	-0.84
Total ($/GJ)	24.53	25.16	19.11	18.17	13.97	15.58	11.57	10.93
Lower Bound ($/GJ)	24.25	23.85	18.22	17.20	13.30	14.43	11.17	10.52
Optimality Gap (%)	1.12	5.47	4.90	5.64	5.05	7.96	3.60	3.86
BEOP ($/bbl)	122.22	125.71	92.41	87.25	64.11	72.98	50.91	47.37
Case Study	A50	S50	N50	F50	A100	S100	N100	F100
Natural Gas	8.15	8.11	8.97	9.27	8.17	8.11	8.97	9.27
Water	0.01	0.03	0.01	0.01	0.01	0.03	0.01	0.01
Investment	4.11	4.84	3.34	3.33	3.75	4.35	3.02	2.92
CO_2 TS&M	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00
O&M	1.09	1.28	0.88	0.88	0.99	1.15	0.80	0.77
Electricity	-1.13	-0.48	-3.22	-3.98	-1.17	-0.48	-3.22	-3.98
LPG	-0.84	-0.84	-0.84	-0.84	-0.84	-0.84	-0.84	-0.84
Total ($/GJ)	11.39	12.94	9.15	8.67	10.90	12.32	8.74	8.16
Lower Bound ($/GJ)	11.22	11.87	9.13	8.32	10.22	11.90	8.44	8.09
Optimality Gap (%)	1.49	9.01	0.19	4.15	6.59	3.54	3.60	0.87
BEOP ($/bbl)	49.88	58.46	37.55	34.91	47.20	55.04	35.32	32.10

looping. It is interesting to note that the BEOP values of the N10 & A50 case studies and F10 & A100 case studies are very similar, suggesting that the GTL refinery with chemical looping is cost competitive with 5-10x larger GTL refineries with reforming. This is significant because an economy-of-scale effect is experienced in using chemical looping without having to increase the plant size.

Overall, lowered investment costs and increased electricity sales outweigh higher natural gas costs in making chemical looping more profitable than reforming. Unit sizes are smaller for chemical looping because there is less material that flows through them due to having no light gas recycle and greater per pass natural gas conversion. The higher syngas product from chemical looping also reduces the size of downstream units, decreasing their investment costs as well. Because of the lack of a light gas recycle, chemical looping generally intakes more natural gas than reforming. This causes the natural gas cost contribution to be higher for chemical looping. However, this is compensated through using the light gases that are not recycled for fuel combustion to generate process heat and electricity instead. By having more available fuel gas, the electricity production in all the chemical looping cases is greater. The expanded electricity production ultimately gives CLR & STS a major cost advantage over ATR & SMR.

2.5.4 Investment Cost Breakdown

Investment costs broken down for each section in the GTL process are displayed in Table 2.4. Oxygen generation cost for the ATR is included in the H_2/O_2 production section and not in the syngas generation section. This is because the ASU may provide oxygen to other units that are not associated with syngas generation.

In the ATR cases, $>$99% of the H_2/O_2 production cost is contributed by the ASU, and the produced oxygen is all sent to the ATR. Adding the ASU cost to the ATR syngas generation cost, it is clear that the chemical looping processes are less expensive than reforming after considering all process units associated with syngas generation (1 kPBD: 52-57% less than ATR, 21-30% less than SMR; 10 kBPD: 37-47% less than ATR, 33-44% less than SMR; 50 kBPD: 42-49% less than ATR, 47-53% less than SMR; 100 kBPD: 42-50% less than ATR, 47-54% less than SMR).

Table 2.4: Investment costs ($MM) of the major sections of the GTL refinery in each case study.

Case Study	A1	S1	N1	F1	A10	S10	N10	F10
Syngas Generation	20.48	41.20	32.41	28.80	108.96	247.04	164.86	138.87
Syngas Cleanup	0.01	1.86	0.00	0.01	0.05	7.16	0.02	0.03
HC Production	16.86	18.40	12.26	12.65	70.58	81.69	60.69	58.77
HC Upgrading	55.32	58.91	55.31	54.76	247.17	262.49	245.27	244.65
H_2/O_2 Production	46.69	0.21	0.00	0.20	152.71	0.93	0.91	4.61
H&P Integration	22.04	37.96	16.27	17.48	106.39	166.68	109.61	116.59
Wastewater Treatment	8.94	11.76	6.58	7.55	39.35	50.82	38.30	39.56
Total ($MM)	170.34	170.29	122.83	121.45	725.21	816.80	619.66	603.08
Total ($/BPD)	170341	170289	122827	121454	72521	81680	61966	60308
Case Study	A50	S50	N50	F50	A100	S100	N100	F100
Syngas Generation	463.79	1051.43	559.59	492.24	862.76	1962.06	1042.45	899.53
Syngas Cleanup	0.16	20.03	0.08	0.08	1.89	31.90	0.12	0.12
HC Production	285.91	335.45	225.60	220.54	527.19	625.89	421.11	410.05
HC Upgrading	770.94	826.15	764.63	761.92	1364.59	1483.01	1351.10	1345.42
H_2/O_2 Production	502.71	2.66	2.61	13.16	939.04	4.17	4.10	20.67
H&P Integration	288.59	479.54	308.82	360.33	545.87	805.84	573.57	594.55
Wastewater Treatment	111.86	139.86	109.34	112.38	174.75	216.89	171.01	175.82
Total ($MM)	2423.95	2855.11	1970.67	1960.65	4416.10	5129.77	3563.46	3446.16
Total ($/BPD)	48479	57102	39413	39213	44161	51298	35635	34462

Downstream processing units in the hydrocarbon production and upgrading sections also have reduced investment costs due to higher syngas purity coming from chemical looping.

Generating syngas from CLR is slightly more expensive than STS. NiO has one active oxygen atom, while Fe_2O_3 has three active oxygen atoms. In other words, around 9% of the total mass of NiO/Al_2O_3 is active oxygen, while about 20% of the total mass of Fe_2O_3/TiO_2 is active oxygen. More NiO/Al_2O_3 needs to be circulated between the FR and AR to achieve the same catalytic outcome as a given amount of Fe_2O_3/TiO_2, increasing the reactor sizes in CLR and leading to the higher costs in syngas generation.

Overall, by decreasing the capital expenses associated with syngas generation (including the O_2 production cost for ATR) and hydrocarbon production and upgrading, the two most costly sections in the GTL refinery, CLR and STS lead to improved GTL economics.

2.6 Conclusion

The prospective benefits of chemical looping for syngas production in GTL processes are shown through a process synthesis and global optimization approach. By switching to chemical looping from conventional reforming for natural gas conversion, cost savings of approximately 25-40% are potentially achievable. Even though chemical looping utilizes more natural gas, it is ultimately more profitable due to greater electricity sales and reduced investment costs. This is highlighted by the economic performance of small-scale chemical looping processes rivaling those of conventional reforming at 5-10x greater size. Integrating chemical looping into GTL processes is an enormous opportunity to enhance their capabilities.

3. BIOMASS-BASED COPRODUCTION OF AMMONIA AND METHANOL*

3.1 Motivation

Interest in ammonia and methanol coproduction initiated from companies wanting to hedge their product output with fluctuating selling prices. The first operating coproduction plant in the United States started in 1994 in Westward, OK when Haldor-Topsøe was commissioned to retrofit an ammonia facility with methanol capability to take advantage of the widening price margin between the two products [95]. Similar market conditions will likely exist sometime in the future because their supply and consumption are not always in sync. Therefore, the flexibility afforded by an intensified process is highly advantageous. Likewise, the sharing of common process units and utilities reduces the capital investment required compared to separate operations. As in the ethanol case, biomass is an attractive feedstock for reducing the carbon intensity of ammonia and methanol production in meeting their growing demand.

3.2 Background

While there are several patents and company brochures describing coproduction [96, 97], there is a dearth of scholarly work concretely quantifying the benefits [98]. Moreover, a biomass-based ammonia and methanol process has never been investigated before. Previous works proposed separate process superstructures for ammonia and methanol production [99–101]. This work integrates these two superstructures together, forming a MINLP process synthesis problem, and solves for the optimal coproduction using global optimization. A plant overview of ammonia and methanol coproduction from biomass is shown in Figure 3.1. The process contains four main components. First, biomass is converted to syngas through gasification. The syngas is then cleaned and purified to the inlet specifications required for ammonia and methanol synthesis. Finally, synthesis loops with compression, reaction, and separation steps produce the final product. Some background on

*Part of this section is reprinted from "Toward Optimal Synthesis of Renewable Ammonia and Methanol Processes (RAMP)" by Tso, W.W. and Demirhan, C.D. and Powell, J.B. and Pistikopoulos, E.N., Computer Aided Chemical Engineering, 2018, Vol. 44, pp 1705-1710, with permission from Elsevier and Copyright Clearance Center.

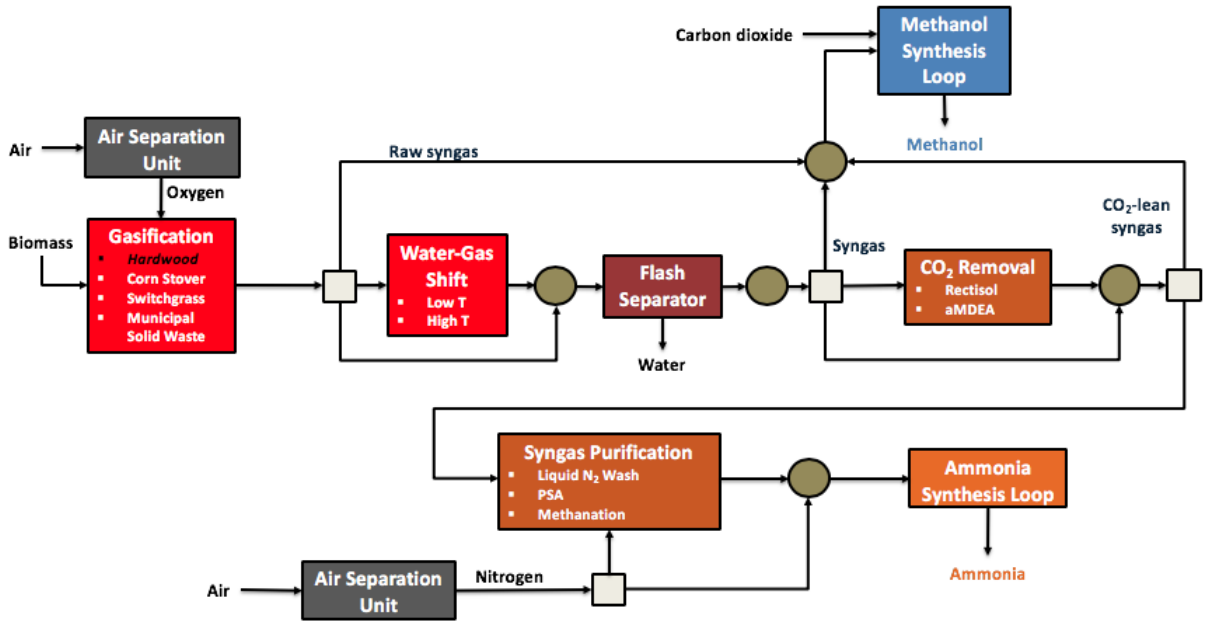


Figure 3.1: An overview of the BAM plant.

integration options between ammonia and methanol is provided before moving onto mathematical modeling of the synthesis reactors and results from the global optimization solution.

The overall reactions for ammonia (Eq. 3.1) and methanol synthesis (Eq. 3.2) are shown below. Hydrogen is a key ingredient in both reactions and is obtained from syngas production. Upstream units for producing hydrogen are the main components shared by ammonia and methanol. Because syngas production is typically the most expensive section, a joint utilization mitigates this cost. Given thermodynamic and kinetic limitations, not all the hydrogen sent to the synthesis reactors will be consumed in a single pass. Therefore, one obvious way to integrate ammonia and methanol production is to exchange hydrogen-rich off gases from the reactors with each other. Inherent recycling of off gases to the same reactor is already performed to increase yield and improve

efficiency. A combined recycling scheme can enhance productivity even further.

$$N_2 + 3H_2 \rightarrow 2NH_3 \tag{3.1}$$

$$CO + 2H_2 \rightarrow CH_3OH \tag{3.2}$$

Another integration design is to use methanol synthesis as a carbon oxide purifying step for ammonia synthesis. Syngas is compromised of mostly H_2, CO, and CO_2, but these carbon oxides contaminate and poison the catalyst promoting ammonia synthesis. As such, syngas has to be completely removed of CO and CO_2 before being sent to the ammonia synthesis reactor. This is typically carried out in two steps. First, CO_2 is separated out in bulk quantity, and then the remaining carbon oxides are hydrogenated to methane, consuming some of the valuable hydrogen, or washed out with liquid nitrogen. Instead of treating the carbon oxides as waste contaminants, they could be utilized for methanol production. Methanol synthesis is broken down to two reaction steps (Eqs. 3.3 & 3.4), where CO_2 is the actual main reactant. Therefore, producing methanol monetizes the value of CO_2, an otherwise wasted stream from ammonia synthesis, and reduces the raw material needed as input.

$$CO_2 + 3H_2 \rightarrow CH_3OH + H_2O \tag{3.3}$$

$$CO + H_2O \rightarrow CO_2 + H_2 \tag{3.4}$$

Figure 3.2 summarizes these two main integration methods. The sharing of utilities such as heat, electricity, and cooling water is another way of cooperative operation to reduce costs. The optimization of utility use is performed through a simultaneous heat, power, and water integration with the process synthesis. This methodology followed is similar to the one described in previous process synthesis works [99–101].

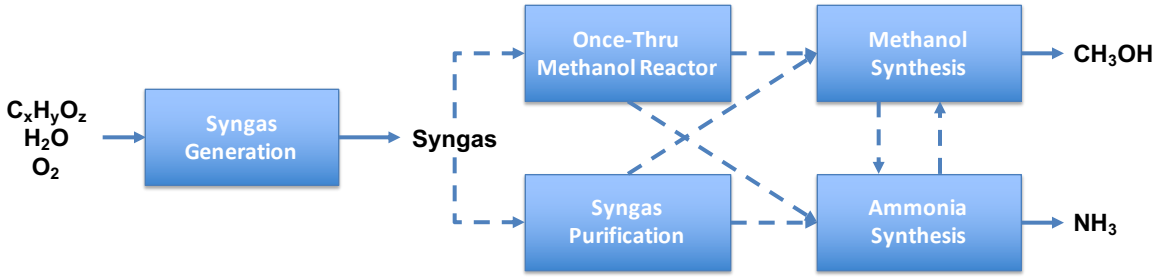


Figure 3.2: Integration options between ammonia and methanol production.

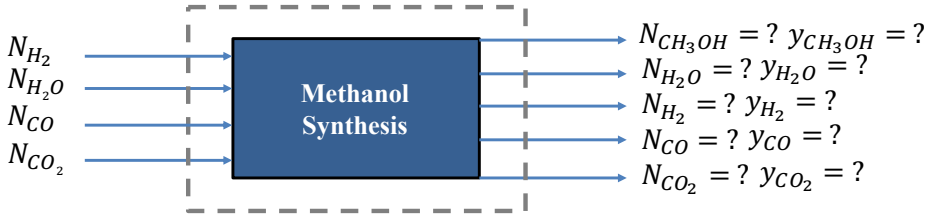


Figure 3.3: Methanol synthesis reactor.

3.3 Mathematical Modeling of Synthesis Reactors

Simplified models of ammonia and methanol synthesis are assumed to be adequate enough to describe their reactors. Like previously described in Section 2.3, more complex models would be too complicated and ill-suited for the MINLP process synthesis and global optimization.

3.3.1 Methanol Synthesis Model

An equilibrium-based model, assuming Eqs. 3.3 & 3.4 are at equilibrium, is developed for methanol synthesis. In the reactor (Figure 3.3) , there are five unknown molar species flows and five unknown mole fractions. After writing down atom balances for C, H, and O, five mole fraction equations, and the two thermodynamic equilibrium equations (Eqs. 3.5 & 3.6), there are zero degrees of freedom left. Therefore, under an equilibrium assumption, methanol synthesis is fully specified. Industrial methanol synthesis reactors do often reach equilibrium, so this is a sufficient approximation.

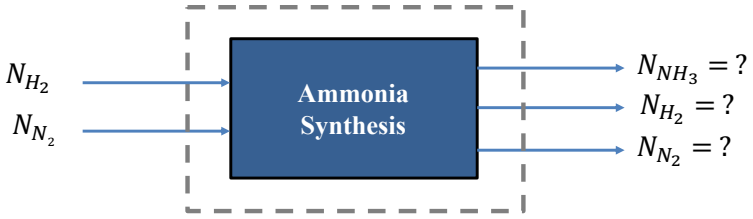


Figure 3.4: Ammonia synthesis reactor.

$$K_{eq,MS} = \frac{y_{H_2O}\, y_{CH_3OH}}{y_{H_2}^3\, y_{CO_2}\, P^2} \tag{3.5}$$

$$K_{eq,WGS} = \frac{y_{CO_2}\, y_{H_2}}{y_{CO}\, y_{H_2O}} \tag{3.6}$$

3.3.2 Ammonia Synthesis Model

An equilibrium model for the ammonia synthesis reactor is not assumed because the reaction does not reach the equilibrium value due to slow kinetics. Instead, the reactor is modeled through a data-driven approach using a dataset of 25 industrial (Imperial Chemical Industries, Kellogg Brown & Root, Haldor Topsøe, Uhde GmbH, Casale) and experimental values. There are three unknown molar species flows coming out of the reactor (Figure 3.4). Inerts and ammonia, due to recycling, are also inlets to the reactor, but they are not shown for clarity. Two degrees of freedom are specified from atom balances for N and H. The final degree of freedom is closed by determining the conversion of either N_2 or H_2, depending on which is the limiting reactant, to NH_3. For simplicity, the conversion X_r is assumed to be linearly dependent on inlet composition, temperature, and pressure (Eq. 3.7).

$$X_r(T, P, y_i) = A \cdot T + B \cdot P + C \cdot y_{H_2} + D \cdot y_{N_2} + E \cdot y_{NH_3} + F \cdot y_{Inert} + G \tag{3.7}$$

Fitted parameters A-G are determined from the data-driven approach. The dataset is separated into training, validation, and testing subsets. 5-fold cross-validation with the training and validation sets is performed. Then, the fitted parameters are used to predict the testing data not included in

Table 3.1: Fitted parameters from 5-fold cross-validation for each model (different training and validation sets). Error is reported for each model.

Terms	Model 1	Model 2	Model 3	Model 4	Model 5
A	-0.888	-0.547	-0.680	-0.718	-0.659
B	0.099	0.086	0.094	0.093	0.093
C	0.367	2.141	8.678	8.550	5.838
D	0.229	0.389	1.012	1.053	0.764
E	0.025	0.146	0.857	0.856	0.554
F	0.099	0.292	0.933	0.920	0.646
G	-0.361	-2.222	-8.654	-8.630	-5.892
Train RMSE	0.127	0.143	0.151	0.144	0.158
Validation RMSE	0.228	0.186	0.204	0.180	0.101

Table 3.2: Model predictions of testing data using the fitted parameters from cross-validation. Error is reported for each prediction.

Testing Set	Model 1	Model 2	Model 3	Model 4	Model 5	Actual
Test 1	0.327	0.329	0.332	0.324	0.327	0.338
Test 2	0.303	0.299	0.300	0.297	0.301	0.305
Test 3	0.315	0.316	0.284	0.329	0.318	0.331
Test 4	0.213	0.209	0.209	0.211	0.209	0.224
Test 5	0.209	0.238	0.220	0.214	0.220	0.214
Test RMSE	0.061	0.089	0.130	0.056	0.062	-

training and validation. These results are shown below in Tables 3.1 & 3.2.

The fitted parameters from all models are observed to closely approximate the actual conversion values. Parameters from Model 4 are used to model ammonia synthesis in the process optimization because it has the smallest average absolute error of 0.007 and the smallest root mean squared error (RMSE). It is important to note that the reported RMSE figures are measures of the *spread* of error between the actual value and predicted value. They do not quantify the *absolute* error. Finally, with the calculated X_r, the production of ammonia is determined through stoichiometric relationships. For example, if N_2 is the limiting reactant, the following equations are used:

$$N_2^{out} = N_2^{in}(1 - X_r) \tag{3.8}$$

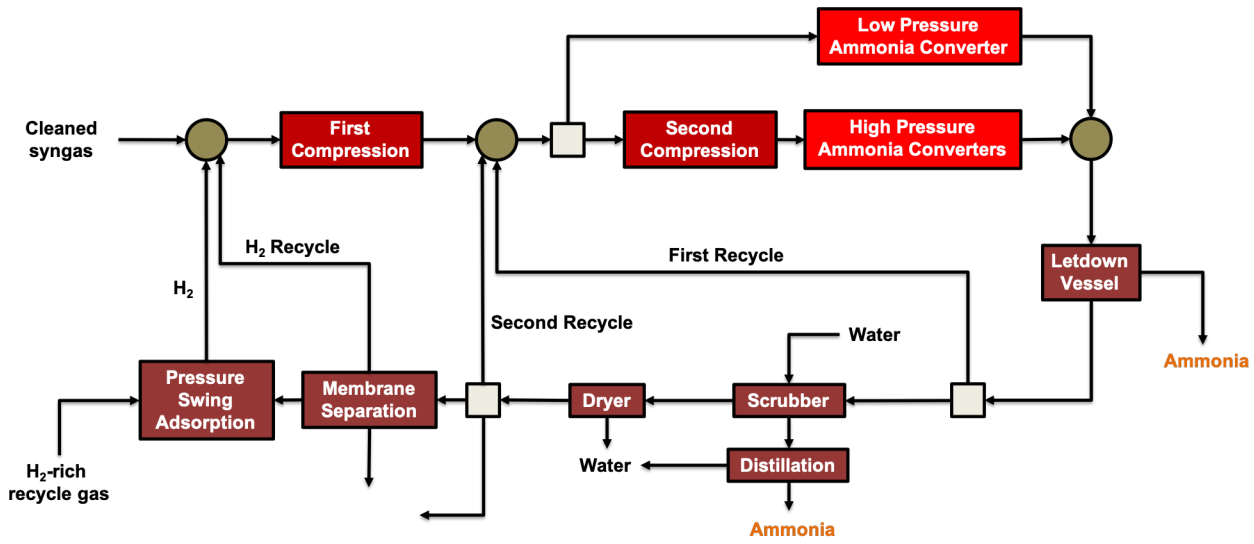


Figure 3.5: Ammonia synthesis loop.

$$H_2^{out} = H_2^{in} - 3(N_2^{in} - N_2^{out}) \tag{3.9}$$

$$NH_3^{out} = NH_3^{in} + 2(N_2^{in} - N_2^{out}) \tag{3.10}$$

3.4 Synthesis Loop Configurations for Coproduction

Detailed flowsheets of the ammonia and methanol synthesis loops are shown in Figures 3.5 & 3.6. Syngas containing carbon oxide species can be sent to an once-through reactor to produce methanol and purify the syngas stream for ammonia synthesis. Flash vessels and distillation processes separate out the ammonia and methanol as final products. Each synthesis loop has an inherent recycle where unreacted syngas is sent back to the reactor to increase product yield. Likewise, the loops are connected through the collective recycling of unreacted gases. The published paper explains in further detail the process descriptions for each section of the BAM process [102].

3.5 BAM Case Studies

Ammonia and methanol coproduction is studied for the first time within a process synthesis & global optimization approach. Moreover, biomass is utilized as a feedstock to reduce carbon

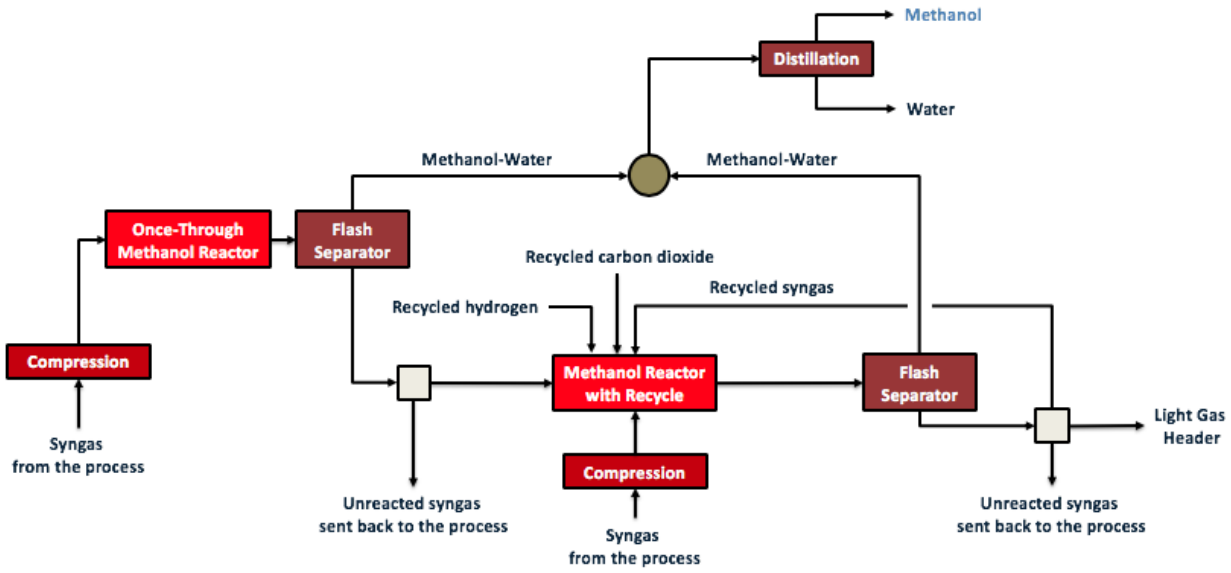


Figure 3.6: Methanol synthesis loop.

intensity. Information about cost estimations for the process units is found in previous publications [99–101]. Detailed model equations for other units are also found in these publications [99–101]. The objective function of the BAM process (Eq. 3.11) is to minimize the overall cost of the coproduction. To weigh the production of ammonia and methanol on an equal basis, each cost contribution is levelized with respect to the total hydrogen content in the product (\$/kg H_2 eq.). Ammonia has 17.6 wt.% of hydrogen, and methanol has 12.6 wt.% of hydrogen.

$$\min \sum_{f \in Feed} Cost^{f} + Cost^{El} + Cost^{Seq} + \sum_{u \in U_{Inv}} Cost^{u} \tag{3.11}$$

The cost contribution terms have the same meaning as those in Section 2.5, except the feedstock costs here include biomass and not natural gas.

3.5.1 Global Optimization Strategy

The complete mathematical model describing the BAM process synthesis is a large-scale non-convex MINLP with >25,000 constraints >20,000 continuous variables, and >50 binary variables.

A similar global optimization strategy as that described in Section 2.5.1 is adopted.

3.5.2 Description of Case Studies

Two sets of case studies are explored to quantify the cost savings from integration and the reduction in CO_2 emissions from using biomass. The first case study compares producing 250 kg H_2 eq./day of ammonia and methanol separately or together. As an example of grasping the weight basis of hydrogen content, a 250 kg H_2 capacity corresponds to about 1400 MT/day for ammonia and 2000 MT/day for methanol. Figure 3.7 outlines this case study. The second case study examines the effect of different product ratios of ammonia and methanol. The product output is 500 kg H_2 eq./day, and product ratios explored are 1:0, 1:3, 1:1, 3:1, and 0:1. Figure 3.9 outlines the second case study.

The case studies are labeled as $[\%A]/[\%M] - [C]$, where $\%A$ is percentage of the product output by hydrogen content that is ammonia, $\%M$ is percentage of the product output by hydrogen content that is methanol, and C is the plant scale in kg H_2 eq./day. For example, 75A/25M-500 represents a 500 kg H_2 eq./day BAM plant that produces 75% ammonia and 25% methanol. All case studies assume hardwood (forest residue) as the biomass feedstock at a purchase price of \$70/MT. The BAM process is also constrained not to produce more than 50% of the GHG emissions typically emitted from producing ammonia and methanol from natural gas (1.7 tons of CO_2 per ton of NH_3 and 0.8 tons of CO_2 per ton of CH_3OH). Only economic and GHG emissions results are shown in the following sections. For additional results on the optimal topology, refer to the publication [102].

3.5.3 Case Study I: Single vs. Coproduction

3.5.3.1 Overall Cost Breakdown

A break-even price (BEP), measured in \$/kg or \$/MT, estimates the minimum ammonia or methanol selling price for the BAM plant to turn a profit. The hydrogen BEP is converted to ammonia and methanol BEPs through their hydrogen weight percentages. Lower BEP values represent plants with more promising profit margins. The overall cost breakdown by contribution

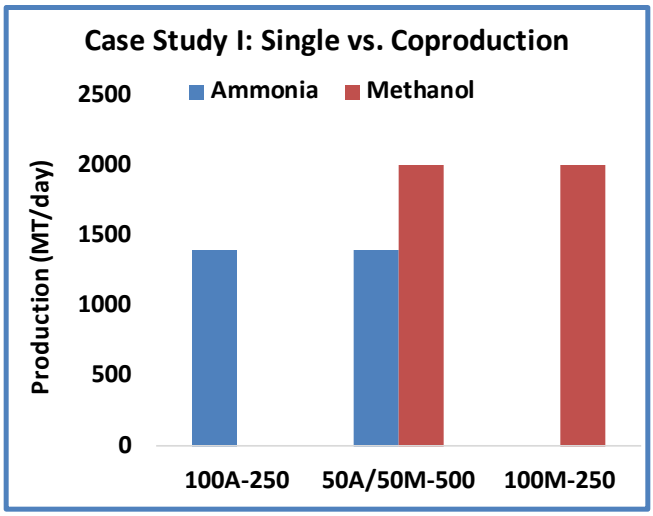


Figure 3.7: Description of first case study.

is shown in Table 3.3.

For the separate productions, methanol is about 25% cheaper to produce than ammonia, which is expected because there are fewer unit operations required in methanol production, lowering its capital investment cost. Methanol has more electricity production because it is less energy intensive to produce, and therefore more fuel gas is available for power. All other costs are similar between the two. From 2000-2009, ammonia prices averaged about $500/MT [103] and the current methanol price is $432/MT [104]. The calculated ammonia and methanol BEPs are very competitive with recent historical prices, suggesting the economic viability of biomass-based ammonia and methanol production.

From integrating the production, the BEP value moves somewhere between that of ammonia and methanol. The ammonia BEP decreases by approximately 20%, while the methanol BEP increases by about 8%. For every $1 increase in the methanol BEP, there is a $4.63 decrease in the ammonia BEP. Therefore, for a small sacrifice in methanol's selling price, the economics of ammonia have greatly improved. This is reflected in the cost contributions, where the coproduction plant utilizes less biomass and has an investment cost comparable to the methanol plant.

Table 3.3: Overall cost of the BAM plant in Case Study I.

Cost Contribution ($/kg of H_2 eq.)	100A-250	100M-250	50A/50M-500
Hardwood Biomass	1.013	1.036	0.996
Water	0.009	0.006	0.007
Investment	1.093	0.840	0.855
O&M	0.288	0.222	0.226
Electricity	-0.343	-0.584	-0.439
BEP ($/kg H_2 eq.)	2.060	1.520	1.644
BEP ($/MT NH_3)	366	-	292
BEP ($/MT CH_3OH)	-	191	207

3.5.3.2 Investment Cost Breakdown

Investment costs broken down for each section in the BAM plant are displayed in Table 3.4. The syngas generation section costs around the same for separate ammonia and methanol produc-

Table 3.4: Investment costs ($MM) of the major sections of the BAM plant in Case Study I.

Investment Costs ($MM)	100A-250	100M-250	50A/50M-500
Syngas Generation	255.68	232.18	417.37
Syngas Cleanup	69.88	87.62	162.36
Ammonia Synthesis Loop	130.52	0.00	130.52
Methanol Synthesis Loop	0.00	33.05	31.06
H&P Integration	91.10	71.51	126.68
Wastewater Treatment	37.69	25.48	46.98
Total	584.87	449.83	914.98

tion. About 40% of ammonia's total cost is from syngas generation, while this is approximately 50% for methanol. The largest investment difference comes from the synthesis loops, where ammonia is about four times more expensive. Costs of the same sections of ammonia and methanol

production are added together to compare with the coproduction result. Figure 3.8 highlights the total investment cost reduction from coproduction and the percentage of each section's contribution to this reduction. The overall investment decreases by about 12% for coproduction. Syngas generation is responsible for more than half of the savings, while the sharing of heat and power utilities reduces investment cost by nearly 30%.

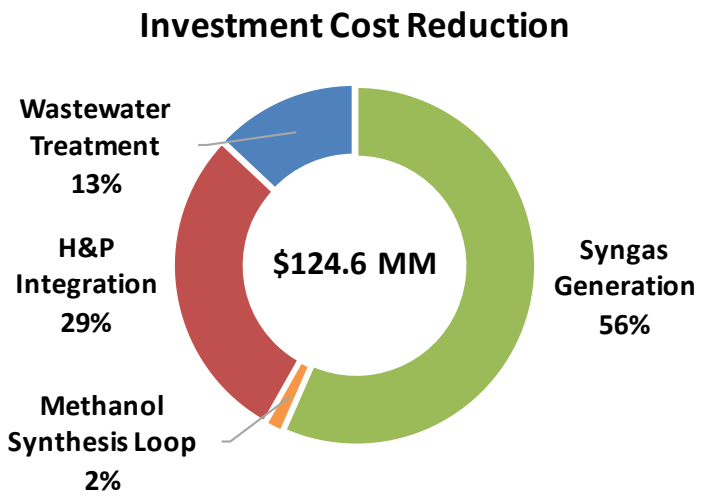


Figure 3.8: Investment cost reduction from coproduction.

3.5.3.3 Greenhouse Gas Emissions

The GHG emissions are determined through considering emissions accrued during: (i) feedstock acquisition and transportation [94], (ii) delivery and end usage of chemicals, and (iii) venting or sequestration of CO_2. Three metrics, the GHG emissions avoided from ammonia (GHGAN), methanol (GHGAM), and electricity (GHGAE) production, are defined to compare the GHG emissions from the BAM plant to fossil fuel based production. GHGAN and GHGAM are reported values for the carbon intensity of ammonia and methanol production from natural gas (1.7 kg CO_2/kg NH_3 and 0.8 kg CO_2/kg CH_3OH). GHGAE is calculated using a typical natural gas based power

plant emissions level (101.3 kg $CO_{2eq.}$/GJ). The sum of GHGAN, GHGAM, and GHGAE represents the total GHG emissions avoided by a BAM plant. An emissions index (GHGI), defined as the ratio of GHG to the previous sum, indicates the relative GHG emissions of the BAM plant. The BAM plant is constrained to emit no more than 50% of that a typical fossil fuel process (Eq. 3.12).

$$GHGI = \frac{GHG}{GHGAN + GHGAM + GHGAE} \leq 0.5 \tag{3.12}$$

Table 3.5 highlights the GHG emissions from the single and coproduction. All biomass-based processes have a GHGI significantly lower than 0.5. The BAM plant is over 80% less carbon intensive than typical ammonia and methanol production from fossil fuels, while still remaining economically competitive. Satisfying some of the growing ammonia and methanol demand with biomass to mitigate CO_2 emissions is a feasible alternative.

Table 3.5: GHG emissions of the BAM plant in Case Study I.

GHG Emissions (kg CO_2 eq./s)	100A-250	100M-250	50A/50M-500
Biomass	-68.25	-69.81	-134.25
Vented CO_2	77.07	47.34	120.14
LGHG	8.82	9.11	17.48
GHGAE	5.17	8.80	13.24
GHGAN	27.61	0.00	27.61
GHGAM	0.00	49.29	49.29
GHGI	0.27	0.16	0.19

3.5.4 Case Study II: Different Product Ratios

3.5.4.1 Overall Cost Breakdown

The overall cost breakdown is shown in Table 3.6 for BAM plants from high ammonia percentage to low. The cost contribution of biomass is lowest for the 75A/25M plant, suggesting that

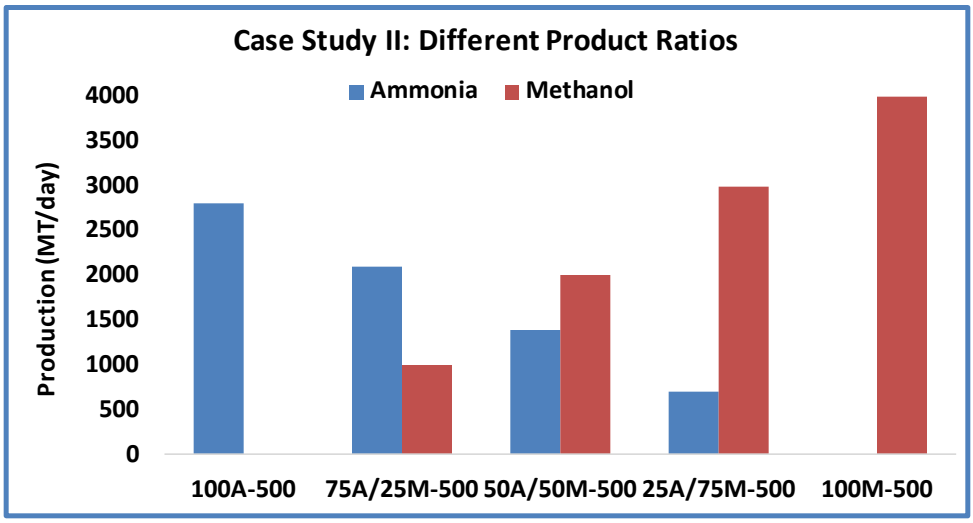


Figure 3.9: Description of second case study.

coproduction plants producing more ammonia utilize less biomass. Methanol contains carbon in its structure and likely needs more biomass than ammonia, which does not have carbon. Investment costs for coproduction plants are more expensive than methanol, but cheaper than ammonia. With increasing methanol percentage, the BEP values of ammonia and methanol decrease, again reiterating that methanol production is less costly than ammonia. However, all the BEP values are competitive with historical ammonia and methanol prices. The optimal output of the BAM plant will depend on actual market conditions and needs to be adjusted accordingly to meet demand.

3.5.4.2 Investment Cost Breakdown

Figure 3.10 shows the investment cost breakdown for the BAM plants when varying the product ratios. Syngas generation dominates the majority of the investment cost (44-54%). This percentage increases as more methanol is produced, but the absolute value stays fairly constant. The ammonia synthesis loop is more costly than methanol synthesis loop. Coproduction leads to dramatically reduced costs for the former at the expense of increasing the latter by a lesser amount. Investment cost decreases with more methanol production. One takeaway is that adding methanol production capability to an existing ammonia plant is a cheaper option than adding ammonia production capa-

Table 3.6: Overall cost of the BAM plant in Case Study II.

Cost Contribution ($/kg of H_2 eq.)	100A	75A/25M	50A/50M	25A/75M	100M
Hardwood Biomass	1.021	0.924	0.996	1.002	1.04
Water	0.009	0.007	0.007	0.006	0.006
Investment	0.948	0.873	0.855	0.788	0.713
O&M	0.250	0.230	0.226	0.208	0.188
Electricity	-0.405	-0.262	-0.439	-0.497	-0.58
BEP ($/kg H_2 eq.)	1.824	1.772	1.644	1.507	1.369
BEP ($/MT NH_3)	324	315	292	267	-
BEP ($/MT CH_3OH)	-	223	207	190	172

bility to an existing methanol plant. Therefore, in terms of retrofitting existing facilities, it is more advantageous to add methanol into the ammonia picture than vice versa. Moreover, constructing a new coproduction plant is marginally more expensive than building a solely ammonia plant, but the former is much more flexible in adjusting to market demand.

3.5.4.3 Greenhouse Gas Emissions

The GHG emissions for each BAM plant in Table 3.7. As expected, methanol production is less carbon intensive than ammonia production because the former stores some carbon in the chemical structure of methanol. All BAM plants show a significant reduction in GHG emissions compared to fossil fuels and remain economically competitive. Coproduction allows the carbon intensity of ammonia to decrease (about 25% lower with 50% methanol).

3.5.4.4 Sensitivity Analysis

Purchase price of biomass is a key parameter because it has the greatest cost contribution to the BEP value. In Table 3.6, biomass contributes between 52-76% to the final cost. Considering a low price of $40/ton and a high price of $100/ton [105], Figure 3.11 displays the BEP values for the various product ratios. Even with higher biomass costs, the BAM plants remain profitable with respect to historical prices for ammonia and methanol. Therefore, BAM plants are not only flexible

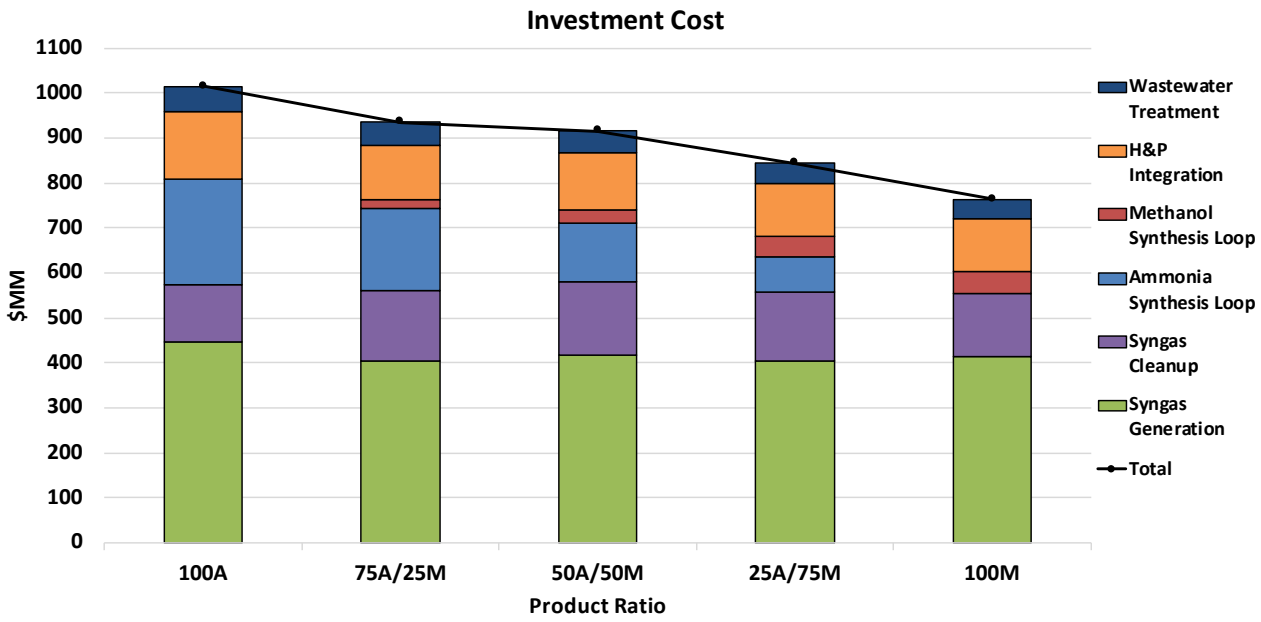


Figure 3.10: Investment costs ($MM) of the major sections of the BAM plant in Case Study II.

in navigating market conditions and less carbon intensive, but also highly resilient to changing biomass economics.

3.6 Conclusion

Promising results for BAM plants are shown through a process synthesis and global optimization approach. By switching from fossil fuels to biomass, an estimated 80% reduction in GHG emissions is possible, without any additional CO_2 capture units, is achievable. Through integrating ammonia and methanol production, the investment cost decreases over 10% compared to separate plants. Ammonia BEP decreases faster than the increase in methanol BEP for coproduction of all product ratios, which improves the plant economics and flexibility. Ultimately, there is enormous potential in producing ammonia and methanol from biomass, and this is a possible extension of biomass utilization beyond ethanol.

Table 3.7: GHG emissions of the BAM plant in Case Study II.

GHG Emissions (kg CO_2 eq./s)	100A	75A/25M	50A/50M	25A/75M	100M
Biomass	-137.66	-124.52	-134.25	-135.02	-140.21
Vented CO_2	155.45	124.91	120.14	105.13	95.26
LGHG	17.79	16.17	17.48	17.49	18.21
GHGAE	12.21	7.89	13.24	14.98	17.47
GHGAN	55.21	41.42	27.61	13.80	0.00
GHGAM	0.00	24.63	49.29	73.94	98.58
GHGI	0.26	0.22	0.19	0.17	0.16

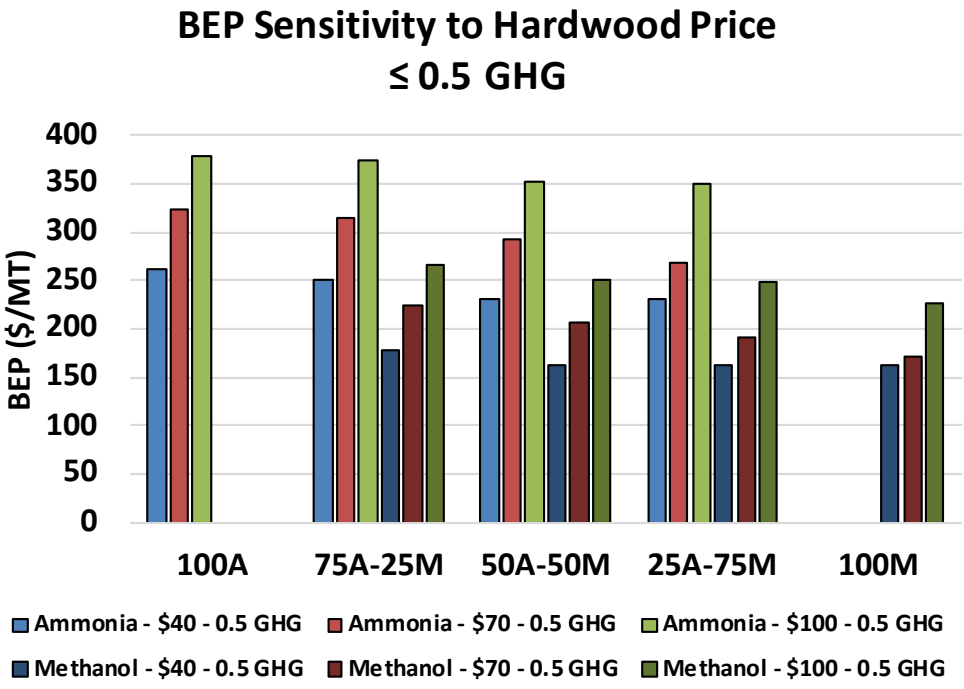


Figure 3.11: Biomass price sensitivity of the BAM plant in Case Study II.

4. SUPPLY CHAIN OPTIMIZATION OF AN ENERGY CARRIER NETWORK*

4.1 Motivation

Increased penetration of renewable energy will require energy storage solutions to align power generation with consumer demand. Energy storage in chemicals have advantages over PSH, CAES, and batteries as mentioned in Section 1.5.3, giving them flexibility as energy carriers, but which chemical to choose is not an obvious decision. Moreover, the additional complexity of chemical storage (more intermediary steps) is capital intensive and efficient strategies are needed for it to be competitive with other storage options. Process synthesis approaches mentioned in Sections 2 & 3 can determine the optimal process to produce the energy carriers, but give no guideline of how to design an infrastructural network connecting supplier to consumer. Instead, a supply chain optimization model can elucidate the optimal network topology.

4.2 Background

Hydrogen, ammonia, and methanol are compared as energy carriers in Table 4.1. While hydrogen has the most attractive mass energy density, volumetric capacity is a key design criterion for energy storage applications. In this respect, gaseous and liquid hydrogen suffer. As described in Section 3.2, ammonia and methanol are synthesized from hydrogen. They have respective volumetric energy densities that are about twice that of liquid hydrogen and four times greater than hydrogen gas. In fact, ammonia and methanol have even greater H_2 volume density than hydrogen itself. Ammonia and methanol also possess favorable storage properties compared to the extreme cryogenic or pressurized conditions required for hydrogen. However, these benefits for ammonia and methanol come at higher production costs, which need to be weighed against the extra storage costs for hydrogen.

Previous works have studied hydrogen, ammonia, and methanol as separate economies [60–

*Reprinted from "Energy Carrier Supply Chain Optimization: A Texas Case Study" by Tso, W.W. and Demirhan, C.D. and and Lee, S. and Song, H. and Powell, J.B. and Pistikopoulos, E.N., Computer Aided Chemical Engineering, 2019, Vol. 47, pp 1-6, with permission from Elsevier and Copyright Clearance Center.

Table 4.1: Comparison of hydrogen, ammonia, and methanol as energy carriers.

Attribute	H_2 (Liquid)	H_2 (Gas)	NH_3	CH_3OH
Mass Energy Density (kWh/kg)	≈ 33.3	≈ 33.3	≈ 5.1	≈ 6.4
Volumetric Energy Density (kWh/L)	≈ 2.5	≈ 1.0	≈ 4.3	≈ 4.6
H_2 Weight %	100	100	17.8	12.6
H_2 Volume Density (kg/m^3)	71.2	24-40	105	99.8
Storage Temperature (°C)	-253	20	a. -33.3 b. 20	20
Storage Pressure (atm)	1	350-700	a. 1 b. 10-20	1

62]. This division is limiting and does not investigate the potential interactions among the three. A more holistic perspective could exploit the energy carriers together in a hybrid economy. In addition, quantitative studies analyzing energy carriers and their infrastructural design are few in number and have primarily focused on hydrogen [106–108]. Toward addressing these gaps and designing an optimal network of energy carriers, a MILP model is developed that considers the trade-offs and coordinates logistical decisions within the supply chain. The MILP model is then applied to investigate energy carrier scenarios in Texas.

Utilizing energy carriers in Texas is an interesting application because nearly half of its residents live in the five largest metropolitan areas, but high solar irradiance and wind speeds are concentrated away from this urban population (Figure 4.1). In terms of currently installed renewable capacity, Texas leads the United States in wind power ($\sim$22 GW) [109] and is fifth in solar power ($\sim$2.9 GW) [110]. In 2017, solar and wind energies generated about 18% of the electricity in Texas [109]. Therefore, Texas is well-suited as a case study for the ECN. In the next section, the MILP is described before results are shown.

4.3 Energy Carrier Supply Chain

Inputs to the energy supply chain model are shown in Figure 4.2. Annual average global horizontal solar irradiance (GHI) and wind speeds are obtained from NREL NSRDB [111] and

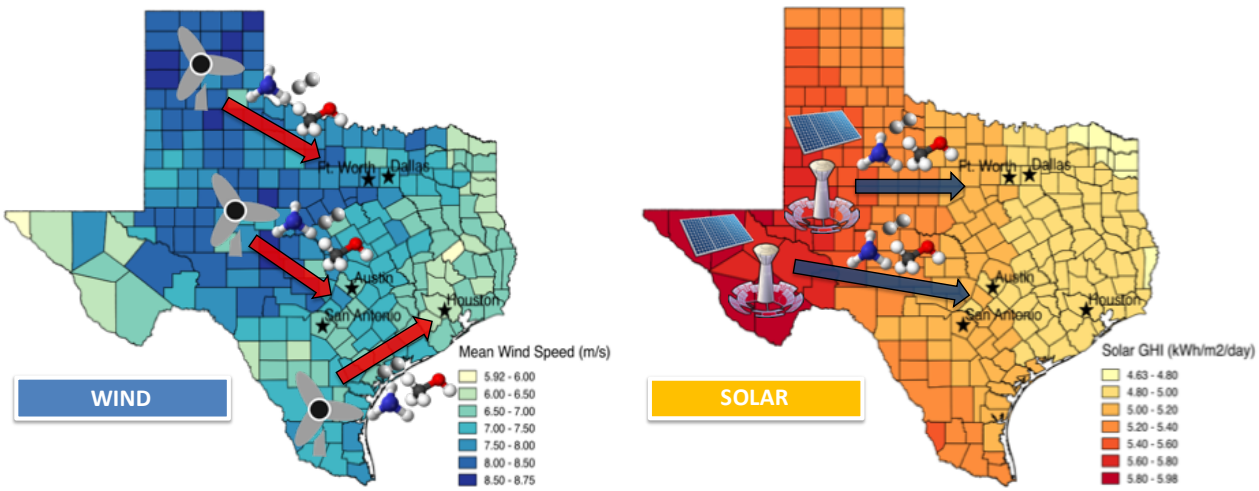


Figure 4.1: Energy carrier overview for Texas.

WIND Toolkit [112], respectively. Land available for solar PV and wind turbine development are derived from a NREL study [113]. Water supplies are taken from USGS [114], and CO_2 point sources are determined from NETL ATLAS. Candidate facilities for energy carrier production are located at the centroids of each county, and the demand locations are Houston, Dallas, Ft. Worth, San Antonio, and Austin. Truck and rail are transportation options from a facility location to a demand city. CO_2 and water are transported through pipeline from resource areas to facility location. Optimal parameters for energy carrier production facilities are determined from a process synthesis approach similar to that in Sections 2 & 3, but only considering air, water, CO_2, solar, and wind as process inputs. Fuel cell and gas turbines convert energy carriers back to power at the demand location site.

Binary variables in the MILP model are assigned to model the selection of facility location, type, & size and conversion technologies for each energy carrier. Continuous variables account for material flows in the network and power capacities. Constraints are specified for flow balances, resource availability, land restrictions, facility requirements, and electricity demand. The overall objective function is to minimize the LCOE of the supply chain network, Full detailed equations of the MILP model are described below. The set, parameter, and variable definitions are found in

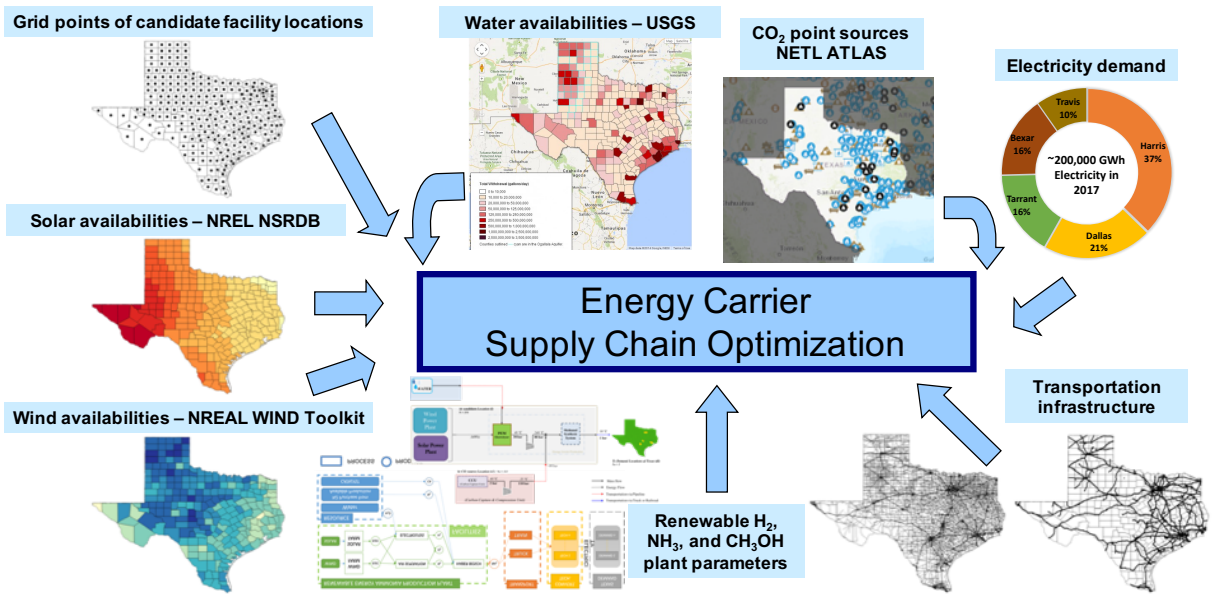


Figure 4.2: Energy carrier supply chain model overview.

the Appendix 7.3.4. Parameter values are also given in Appendix 7.3.4.

4.3.1 MILP Mathematical Model

4.3.1.1 Logistics Constraints

Eq. 4.1 restricts the existence of at most one energy carrier production facility at each candidate location, and Eq. 4.2 restricts the usage of at most one conversion technology for each energy carrier at each demand location.

$$\sum_{(t,q)} y_{l,t,q} \leq 1 \qquad \forall l \in L^F \tag{4.1}$$

$$\sum_{k} y^D_{p,d,k} \leq 1 \qquad \forall p \in P,\ \forall d \in L^D \tag{4.2}$$

4.3.1.2 Facility Location Flows

Eqs. 4.3 to 4.5 deal with network flows around the candidate facility locations. Eq. 4.3 specifies the feedstocks utilized at a facility location ($Feed_{f,l}$) to be the feedstock requirement of

a candidate production facility ($FR_{f,t,q}$) calculated by an earlier process synthesis. Eq. 4.4 limits the total feedstock flow to facility locations from a resource area to be less than the feedstock availability at the given area ($FA_{f,s}$). Eq. 4.5 sets the total feedstock flow to a facility location to be equal to the amount of feedstock required there.

$$\sum_{(t,q)} y_{l,t,q} FR_{f,t,q} = Feed_{f,l} \qquad \forall f \in F,\ \forall l \in L^F \tag{4.3}$$

$$\sum_{(l,m)} x_{f,s,l,m} \leq FA_{f,s} \qquad \forall f \in F,\ \forall s \in S \tag{4.4}$$

$$\sum_{(s,m)} x_{f,s,l,m} = Feed_{f,l} \qquad \forall f \in F,\ \forall l \in L^F \tag{4.5}$$

4.3.1.3 Utility Requirements

$$\sum_{(t,q)} y_{l,t,q} ER_{t,q} = Elec_l \qquad \forall l \in L^F \tag{4.6}$$

$$PV_l + Turb_l = Elec_l \qquad \forall l \in L^F \tag{4.7}$$

$$PV_l \leq \frac{3.6 \times 10^6}{86400} PV_{CF} PV_{Eff} GHI_l SL_l \qquad \forall l \in L^F \tag{4.8}$$

$$Turb_l \leq \frac{T_{CF}(1 - T_{Eff}) WP_l WL_l}{T_{Land}} \qquad \forall l \in L^F \tag{4.9}$$

$$WP_l = \frac{10^{-6} C_p \pi \rho D^2 v_l^3}{8 \times 10^6} \qquad \forall l \in L^F \tag{4.10}$$

Eqs. 4.6 to 4.10 are related to the electricity requirements of the facilities and restrictions on the electricity producible from the renewable farms. Eq. 4.6 specifies the electricity needed at a facility location ($Elec_l$) to be the electricity requirement of a candidate production facility ($ER_{t,q}$), and Eq. 4.7 divides this electricity generation between solar PV (PV_l) and wind turbines ($Turb_l$). Eqs. 4.8 and 4.9 restrain the amount of electricity producible from solar PV and wind turbines, respectively, by taking into consideration their capacity factors (CF), conversion efficiencies (Eff), resource availabilities (GHI_l & v_l), and land usages (SL_l & WL_l). Eq. 4.10 relates wind speed (v_l) to wind

power (WP_l).

4.3.1.4 Demand Location Flows

Eqs. 4.11 to 4.13 are concerned with network flows around the demand locations. Eq. 4.11 specifies the total product flow from a facility location to demand locations is equal to the production output of the candidate facility ($PR_{p,t,q}$). Eq. 4.12 represents the total product flow of an energy carrier to each demand location ($Prod_{p,d}$). Conversion from energy carrier back to power ($Cap_{p,d,k}$) is calculated using each carrier's lower heating value (LHV_p) and conversion efficiencies ($Eff_{p,k}$) in Eq. 4.13.

$$\sum_{(t,q)} y_{l,t,q} PR_{p,t,q} = \sum_{(d,m)} z_{p,l,d,m} \qquad \forall p \in P,\ \forall l \in L^F \tag{4.11}$$

$$\sum_{(l,m)} z_{p,l,d,m} = Prod_{p,d} \qquad \forall p \in P,\ \forall d \in L^D \tag{4.12}$$

$$\frac{y^D_{p,d,k} Prod_{p,d} LHV_p Eff_{p,k}}{86400} = Cap_{p,d,k} \qquad \forall p \in P,\ \forall d \in L^D,\ \forall k \in K \tag{4.13}$$

$$\frac{h_{p,d,k} LHV_p Eff_{p,k}}{86400} = Cap_{p,d,k} \qquad \forall p \in P,\ \forall d \in L^D,\ \forall k \in K \tag{4.14}$$

$$Prod_{p,d} - M(1 - y^D_{p,d,k}) \leq h_{p,d,k} \qquad \forall p \in P,\ \forall d \in L^D,\ \forall k \in K \tag{4.15}$$

$$h_{p,d,k} \leq Prod_{p,d} \qquad \forall p \in P,\ \forall d \in L^D,\ \forall k \in K \tag{4.16}$$

$$h_{p,d,k} \leq M y^D_{p,d,k} \qquad \forall p \in P,\ \forall d \in L^D,\ \forall k \in K \tag{4.17}$$

Notice that Eq. 4.13 contains a bilinear term ($y^D_{p,d,k} Prod_{p,d}$) between a binary variable and a continuous variable. Through substituting for this nonlinear term with a positive variable $h_{p,d,k}$, the exact replacement of Eq. 4.13 is performed with the linear Eqs. 4.14 to 4.17, where M represents a value large enough to relax the constraint.

4.3.1.5 Total Power Demand

The total power demand output from the entire supply chain network ($Total$) is set in Eq. 4.18, assuming an amount of storage time ($Store_{Time}$) passes between when the energy carrier is produced at the facility locations and consumed at the demand sites. Eq. 4.19 limits the capacity contributions ($Demand_d$) of a demand site's conversion technologies toward satisfying the overall power demand.

$$\sum_{(p,d,k)} Cap_{p,d,k} \times 86400 Store_{Time} = Total \tag{4.18}$$

$$Cap_{p,d,k} \leq \frac{y^D_{p,d,k} Demand_d}{86400 Store_{Time}} \qquad \forall p \in P,\ \forall d \in L^D,\ \forall k \in K \tag{4.19}$$

4.3.1.6 Cost Contributions

Finally, Eqs. 4.20 to 4.29 explain the cost contributions from production facilities, renewable farms, conversion technologies, storage, feedstock purchases, transportation, land leases, operation & maintenance, and oxygen sales.

$$Cost_{Fac} = Store_{Time} \sum_{(p,l,t,q)} y_{t,l,q} IC_{t,q} PR_{p,t,q} HC_p \tag{4.20}$$

$$Cost_{Renew} = Store_{Time} \frac{CCR}{OPT} \sum_{l} \frac{PV_l}{PV_{CF}} Cost_{PV} + \frac{Turb_l}{T_{CF}} Cost_{Turb} \tag{4.21}$$

$$Cost_{Conv} = Store_{Time} \frac{CCR}{OPT} \sum_{(p,d,k)} \frac{Cost_k}{CF_k} Cap_{p,d,k} \tag{4.22}$$

$$Cost_{Store} = Store_{Time} \sum_{(p,d)} Prod_{p,d} HC_p Cost_p \tag{4.23}$$

$$Cost_{FP} = Store_{Time} \sum_{(f,p,t,l,q)} y_{t,l,q} FC_{f,t,q} PR_{p,t,q} HC_p \tag{4.24}$$

$$Cost_{FT} = Store_{Time} \sum_{(f,s,l,m)} x_{f,s,l,m} Cost_{f,s,l,m} \tag{4.25}$$

$$Cost_{PT} = Store_{Time} \sum_{(p,l,d,m)} z_{p,l,d,m} Cost_{p,l,d,m} \tag{4.26}$$

$$Cost_{Land} = \sum_{l} LF_l \left(\frac{PV_l PV_{Lease}}{PV_{CF}} + \frac{Turb_l T_{Lease}}{T_{CF}} \right) \tag{4.27}$$

$$Cost_{O\&M} = Store_{Time} \left(\sum_{(p,l,t,q)} y_{t,l,q} OM_{t,q} PR_{p,t,q} HC_p + \frac{OMR}{365} (Cost_{Renew} + Cost_{Conv}) \right) \tag{4.28}$$

$$Sales_{Oxy} = Store_{Time} \sum_{(p,l,t,q)} y_{t,l,q} Oxy_{t,q} PR_{p,t,q} HC_p \tag{4.29}$$

4.3.1.7 Objective Function

The overall aim is to minimize the LCOE of the energy carrier supply chain. Each of the cost contributions is levelized with respect to the total electricity output from the network ($/MWh). The final MILP model consists of Eq. 4.30 as the objective function and all the equations above, except for Eq. 4.13, as constraints.

$$\begin{aligned} \min \quad & (Cost_{Fac} + Cost_{Renew} + Cost_{Conv} + Cost_{Store} \\ & + Cost_{FP} + Cost_{FT} + Cost_{PT} \\ & + Cost_{Land} + Cost_{O\&M} - Sales_{Oxy})/Total \end{aligned} \tag{4.30}$$

4.4 Texas Case Studies

Eight case studies are chosen to show the effect of storage time and the integration of energy carriers. The first set of case studies deal with separate hydrogen (H), ammonia (A), and methanol (M) networks, assuming 1-month and 3-month storage times to examine their effect. The second set of case studies consider integrated (I) supply chains. H-1 is an example of how case studies are labeled, and this represents the 1-month hydrogen network. In the first seven case studies, energy carriers fulfill 100% of the electricity demand at the five cities. The final study looks at a simpler scenario, which location's demand to satisfy to achieve 10% replacement of the total demand, assuming 50% reduction in renewable energy investment cost. The average monthly demand is

assumed to be 16320 GWh. The total demand is assumed to be the monthly demand multiplied by the storage time. The MILP model is solved using CPLEX in GAMS.

4.4.1 Separated Economies

4.4.1.1 1-Month Storage Time

From Figures 4.3 to 4.5, the H-1 network has a much lower LCOE than that of A-1 and M-1. This demonstrates that the additional production costs for ammonia and methanol production outweigh hydrogen's increased storage costs for an assumed 1-month storage time. Taking a closer look, the storage costs are 35% of the LCOE for hydrogen ($0.17/kWh), while it is only 4.2% for ammonia ($0.03/kWh) and <1% for methanol ($0.004/kWh). However, the investment cost contributions (including the production facility, renewable farm, and conversion technologies) for ammonia and methanol total $0.49/kWh and $0.29/kWh, respectively, which is greater than the $0.21/kWh of hydrogen. Likewise, captured CO_2 costs is the most significant factor in the LCOE for methanol, contributing $0.24/kWh by itself.

Hydrogen facilities are concentrated in northern and central Texas, where wind is abundant, while more ammonia and methanol facilities are needed and expanded elsewhere because extra energy is required for their production. The H-1 network requires 64.7 GW of renewable energy (95% wind, 5% solar), while the A-1 network consumes 144.4 GW (70% wind, 30% solar) and the M-1 network utilizes 80.4 GW (91% wind, 9% solar).

4.4.1.2 3-Month Storage Time

When the storage time is extended to 3 months, the LCOE starts to slightly favor ammonia over hydrogen and methanol. A similar LCOE breakdown is observed for the A-3 and M-3 networks (Figure 4.7 & 4.8) compared to their 1-month cases; however, the hydrogen picture (Figure 4.6) is vastly different. Storage cost now covers nearly 2/3 of the hydrogen LCOE or $0.50/kWh. Therefore, energy storage in hydrogen appears more suitable for shorter time periods, and it is more advantageous to utilize other chemicals like ammonia if a longer storage period is expected. In both 1-month and 3-month cases, transportation costs, land lease, and conversion technologies

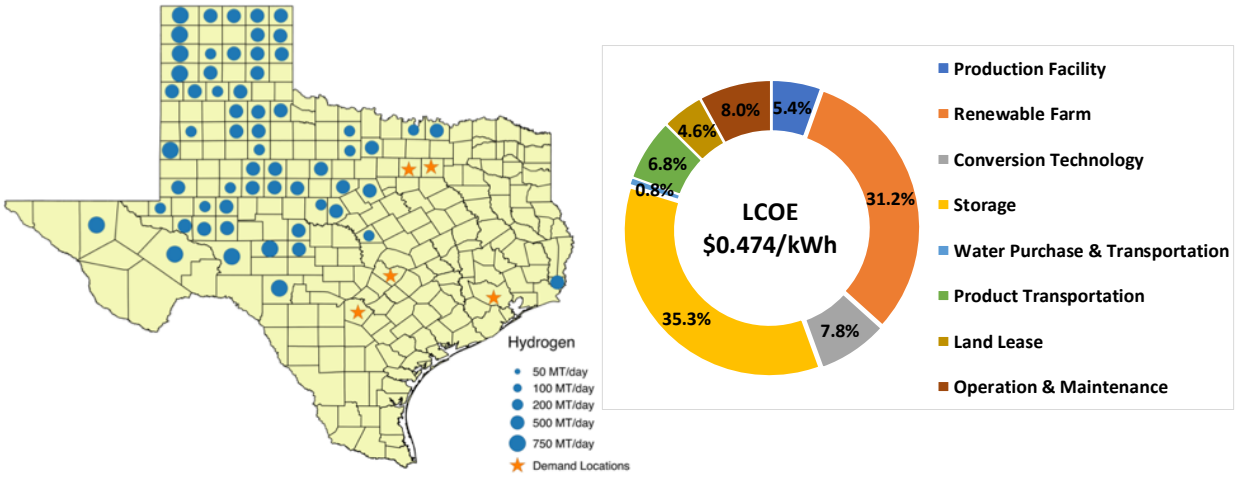


Figure 4.3: H-1 network.

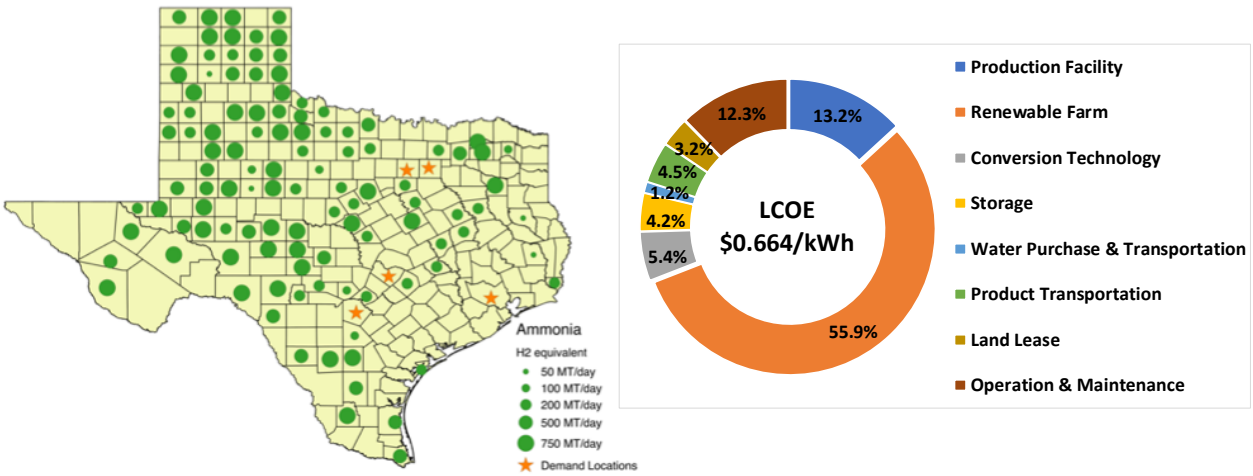


Figure 4.4: A-1 network.

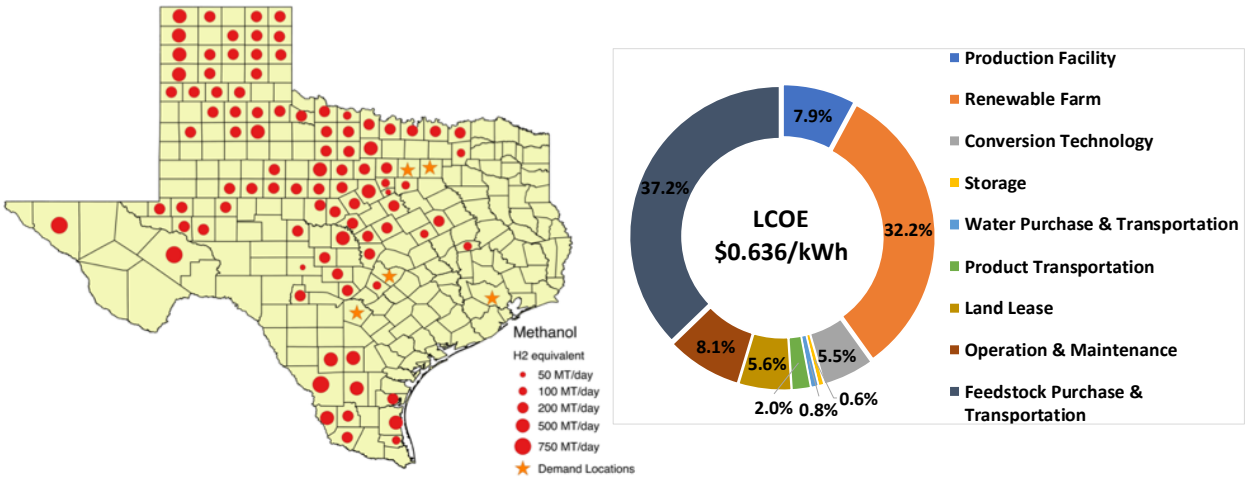


Figure 4.5: M-1 network.

are much smaller portions of the overall LCOE, indicating that they are not limiting factors toward realizing an energy carrier supply chain. The main costs are capital investment associated with energy carrier production, storage in the case of hydrogen, and feedstock in the case of methanol.

Although not explicitly shown in Figure 4.8, methanol facilities have moved closer to CO_2 point sources to mitigate feedstock costs. The hydrogen and ammonia network looks similar to before. The H-1 network requires 62.8 GW of renewable energy (99% wind, 1% solar), while the A-1 network consumes 148.4 GW (60% wind, 40% solar) and the M-1 network utilizes 102 GW (35% wind, 65% solar). Due to increased energy requirements for ammonia and methanol production comprising 3 month storage, their solar energy utilization is now a major factor.

4.4.2 Integrated Economies

4.4.2.1 3-Month Storage Time

The integration of hydrogen, ammonia, and methanol networks is examined for an assumed 3-month storage period. In Figure 4.9, about a 25% reduction in the LCOE is observed (about $0.20/kWh) from the best value obtained for ammonia in Section 4.4.1.2. All three energy carriers are utilized with ammonia holding the largest share (55.5%), followed by methanol (44.3%), and hydrogen (0.2%). The percentage of demand each energy carrier satisfies at each city is shown in

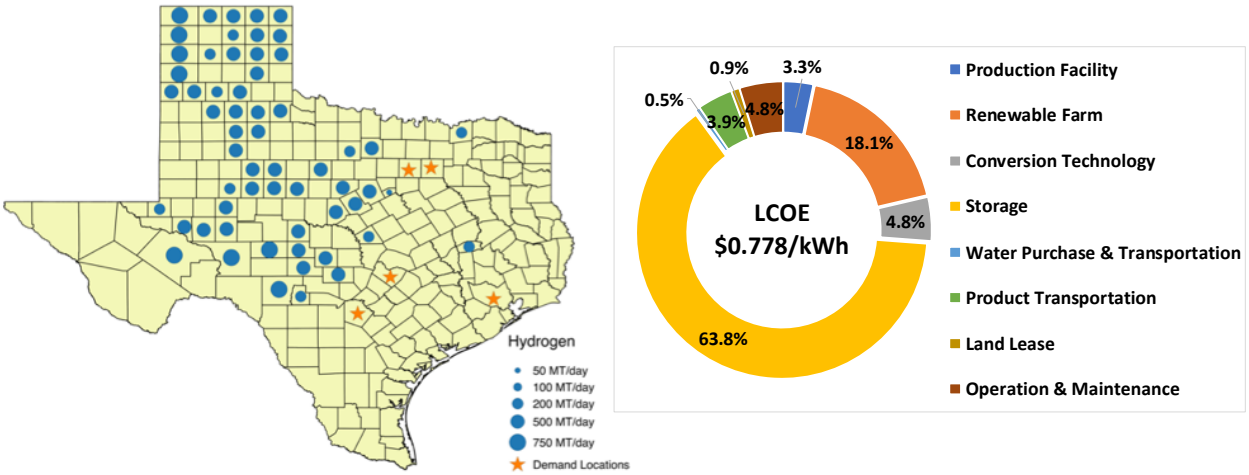


Figure 4.6: H-3 network.

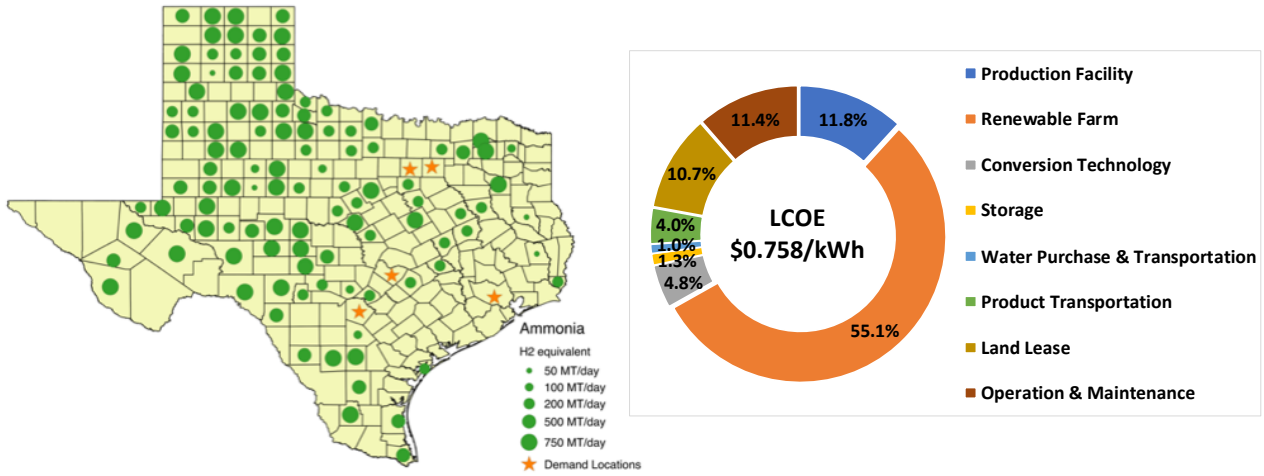


Figure 4.7: A-3 network.

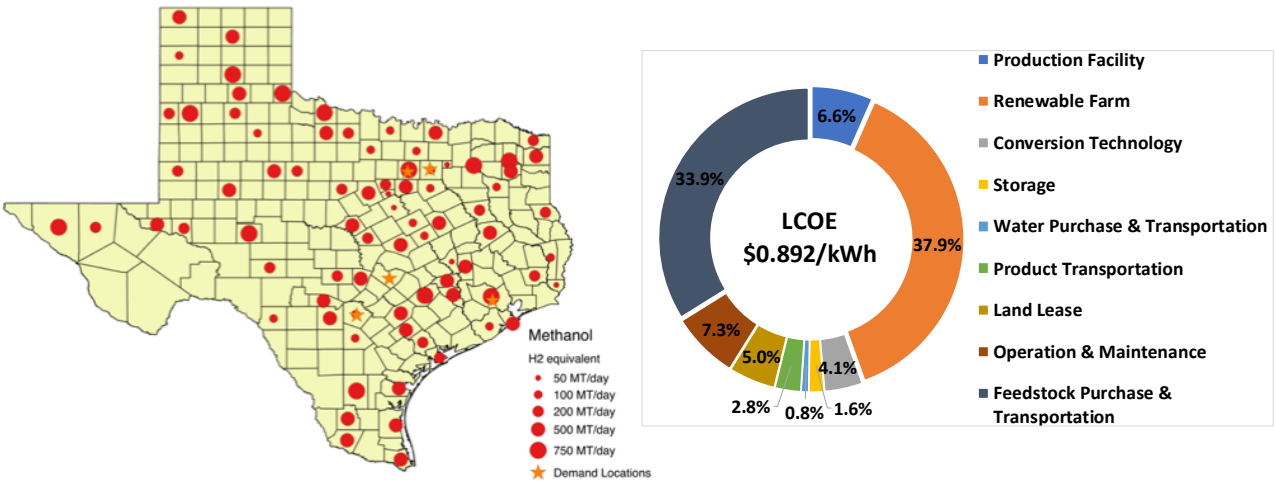


Figure 4.8: M-3 network.

Table 4.2: Energy carrier composition at each demand city.

City	H_2	NH_3	CH_3OH
Houston	2.2%	97.8%	0%
Dallas	0%	33.6%	66.4%
Ft. Worth	0.4%	67.4%	32.2%
San Antonio	0%	62.7%	37.3%
Austin	0%	40.8%	59.2%

Table 4.2. Ammonia and methanol mitigate the hefty storage costs of hydrogen, while hydrogen and methanol alleviate the extra energy burden and capital cost of producing ammonia. Investment costs total \$0.32/kWh, while CO_2 capture costs add \$0.14/kWh and storage contribute \$0.02/kWh. Facilities are concentrated mostly in northern and central Texas, reflecting the high utilization of wind resources. The integrated network requires 97.2 GW of renewable energy (92% wind, 8% solar).

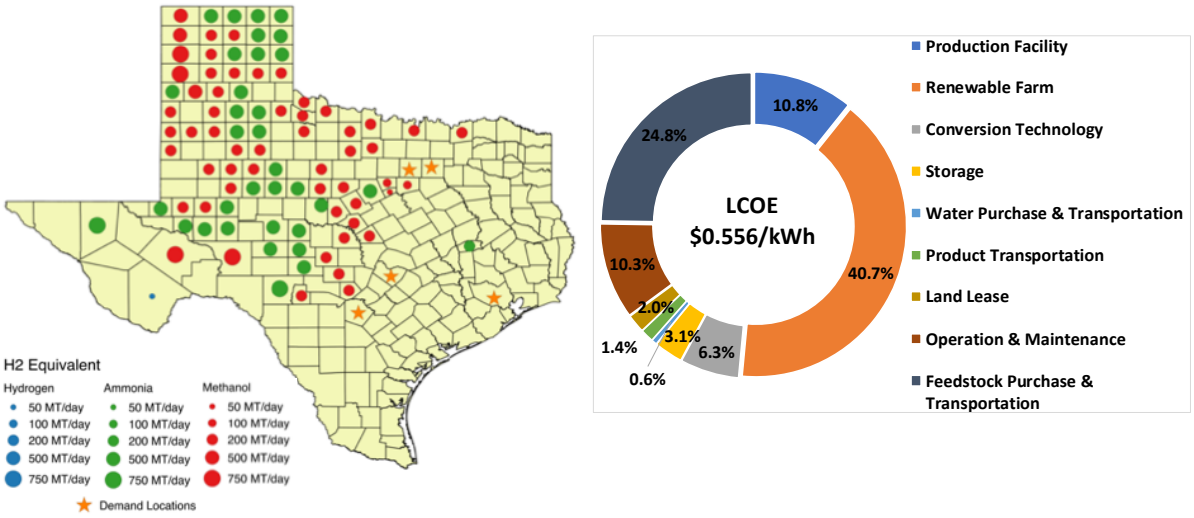


Figure 4.9: I-3 network.

4.4.2.2 10% Demand and 50% Decrease in Renewable Investment Cost

For this case study, it is assumed that 10% of the total demand is met, renewable farm costs have decreased by half, and the storage period is 1 month. Figure 4.10 highlights the results. For these assumptions, the LCOE becomes very competitive with those reported for batteries, but those values usually do not consider the construction of renewable farms as part of their cost evaluation. If the renewable farm cost is subtracted from the network's LCOE, this value becomes $0.24/kWh, which is extremely competitive with PSH, CAES, and batteries. This modified LCOE would be the additional cost for building an energy carrier network with facilities located next to currently operating solar and wind farms. Also comparing this value to the hydrogen LCOE in Section 4.4.1.1 with the renewable farm cost removed, the integration of networks leads to a $0.09/kWh decrease or 27% reduction.

Only ammonia (74%) and hydrogen (26%) are utilized to satisfy electricity requirements in Austin, Dallas, and San Antonio. All of the ammonia is sent to San Antonio, and hydrogen is divided between the other two. Investment costs contribute $0.20/kWh to the LCOE or 57%, and storage costs add $0.06/kWh. Rail is chosen for the long-distance transportation of hydrogen to

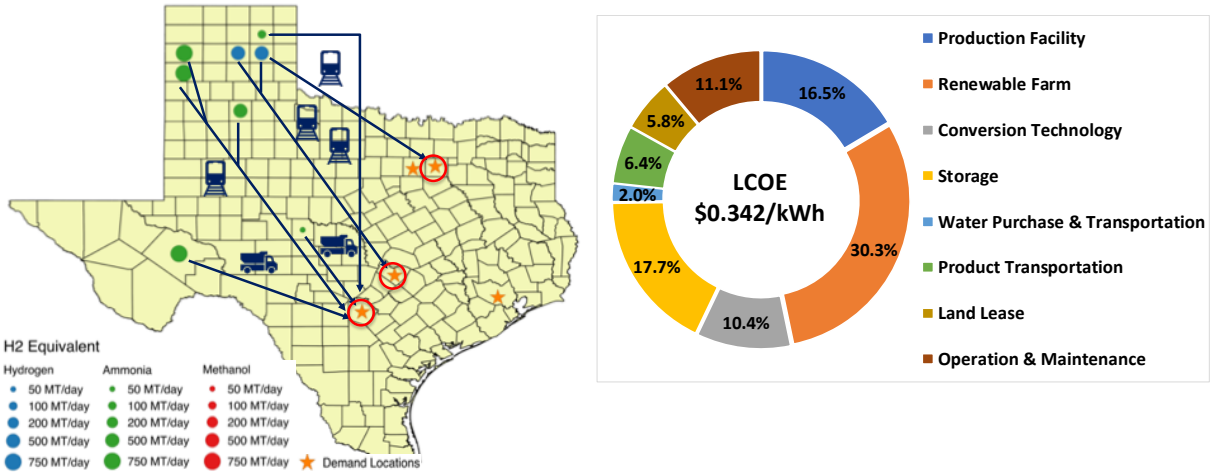


Figure 4.10: I-1 network.

Austin and Dallas, while rail and truck are utilized to ship ammonia to San Antonio depending on the distance. The facilities are spread out between northern, western, and central Texas, exploiting mostly the wind speeds that are concentrated in these areas. The renewable energy required is 10.6 GW (98% wind, 2% solar).

4.5 Remarks on MILP Model

There are two shortcoming of this MILP, which are addressed in Demirhan et al. [115] and Section 5. Time is not explicitly modeled, and storage time is assumed as a parameter. By extending the model to a multi-period form, storage time and inventory levels of each energy carrier can be explicitly accounted for as variables to optimize. Similarly, while geographical availability of solar and wind is captured, their temporal intermittency is not. An annual average is assumed for solving the MILP in one instance. This assumption is made because the goal of this work was to achieve a basic understanding of the economic feasibility of designing an ECN to meet power demand. Incorporating temporal intermittency into the model is important for achieving feasible operation.

4.6 Conclusion

Energy carriers are shown to be competitive with other storage alternatives, especially if geographical flexibility is desired. LCOE values for energy carriers are calculated to be as low as \$0.24/kWh, if solar PV and wind turbines are already constructed. For short storage periods, the extra production costs of ammonia and methanol outweigh hydrogen's storage cost. As the storage time increases, hydrogen becomes less desirable and ammonia grows more attractive. CO_2 capture costs cause methanol to be less preferred than ammonia. Integration of energy carriers leads to about 25% cost savings compared to separate networks, highlighting the importance of not being shortsighted in choosing which chemical to choose. Ultimately, energy carriers are a very promising option to increase the penetration of renewable energy.

5. CLUSTERING DECOMPOSITION ALGORITHM FOR OPTIMIZING RENEWABLE POWER SYSTEMS WITH STORAGE

5.1 Motivation

To tackle curtailment, overgeneration, and capacity sizing issues of renewable power systems with storage, integrated design and scheduling models in the form of mixed-integer optimization problems [116, 117] are useful for minimizing capital and operational costs. Such methods are also called capacity expansion [118] and unit commitment [119] models, where investments into new power and storage capacities (expansion) are concurrently decided with how the units are operated to meet the demand (commitment). The results from these models give the decision-maker important information about the optimal sizing of solar photovoltaic (PV) modules, wind turbines, and energy storage options such as batteries. In addition, key operational choices like optimal battery charging and discharging are elucidated.

The disadvantage is that the models grow very large from the hourly time discretization and long time horizons that are needed to describe weather, load, and price dynamics. This makes them very computationally demanding to solve. As the discretized granularity and time horizon increases, the models become even more complicated. Fig. 5.1 shows the similarity between solar and wind availabilities for two independent days in College Station, TX. Instead of optimizing over the entire time horizon, it is assumed that optimizing over an aggregated number of these clusters offers a good enough approximate solution to the problem [120]. A systematic procedure is needed to address how good of an assumption this is.

5.2 Background

Several authors have employed clustering in mixed-integer optimization models for renewable energy systems. Gabrielli et al. [121] used k-means clustering to optimize multi-energy systems with seasonal storage through developing designs off of clustered time periods and operating them for the full time horizon. Lara et al. [122] utilized k-means clustering to group together gener-

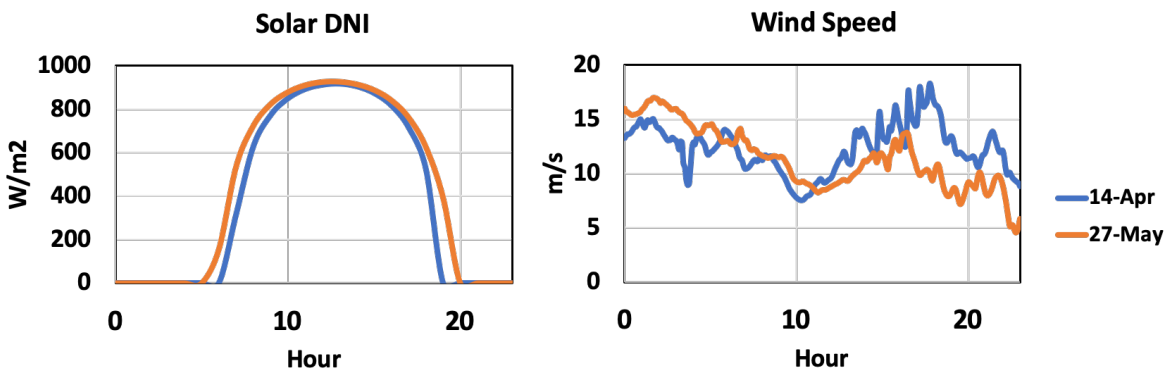


Figure 5.1: Similar solar and wind profiles for April 14 and May 27 in College Station, TX

ators in a multi-scale electric power infrastructure planning model. Heuberger et al. [123] employed k-means clustering in analyzing the system value of energy storage. Peng et al. [124] optimized a concentrated solar power design across scenarios identified through k-means clustering. Pineda and Morales [125] compared different representative hour, day, and week profiles from agglomerative hierarchical clustering (AHC) for capacity expansion planning. Teichgraeber and Brandt [126] used k-shape clustering to optimally schedule a battery storage operation. Tejada-Arango et al. [127] compared different short- and long-term battery storage models through k-medoids clustering. Domínguez et al. [128] selected typical days using k-medoids clustering to optimize the configuration and operation of a residential heat & power system.

In most studies, an underlying assumption is that when the clusters, which are calculated through minimizing the within cluster variance during data processing, are incorporated into an optimization problem, the resulting optimal objective value will also have the least error difference from the true optimal solved using the full time horizon. This is not true; the optimal time aggregation based on clustering error does not necessarily give the best approximation to the true optimal solution. Teichgraeber and Brandt [126] and Bahl et al. [129] have also noticed this error in optimal objective values between the aggregated and full time horizons.

Likewise, while k-means clustering and its variants are common techniques for time aggregation, they do not preserve the time chronology of the input data. Information from non-sequential

time points are grouped together into the same cluster. Because time chronology is not kept, clever techniques are then applied to reconstruct the right timeline [120, 121, 130], increasing the model complexity. Instead, an AHC approach as utilized by Pineda and Morales [125] maintains time chronology in its cluster assignments through a connectivity matrix, enforcing the consecutive ordering of clusters. Having a correctly sequenced time aggregation is important for energy storage studies because inventory levels carry over between adjacent time periods.

This work attempts to address the aforementioned issues. A decomposition algorithm that selects the proper number of representative periods to aggregate a time horizon based on objective value error is developed from elements of previous works [121, 129]. The main aim of this algorithm is to quantify how adequate of an approximation is a specific aggregated time horizon, with respect to the objective value. The algorithmic concept is irrespective of clustering method, but AHC is chosen over k-means clustering because it is a more appropriate choice for energy storage [125]. The decomposition algorithm is then applied to a renewable power system with battery storage in New York City (NYC) to demonstrate how sensitive the levelized cost of electricity (LCOE) is to the number of clusters in time aggregation.

5.3 Decomposition Algorithm Through AHC

An overview of AHC is first provided, and the decomposition algorithm is described next. In general, a similar approach is applicable to any optimization problem with time discretization.

5.3.1 Agglomerative Hierarchical Clustering (AHC)

Let *X* be a matrix of various time series data, where each column X_j is a different feature (i.e. solar irradiance, wind speed, load demand, electricity price) and each row represents a particular time period (i.e. hour, day, week). A data point, x_i, is then all feature values that occur during the ith time period. The AHC method is a bottom-up approach, where the data points start off in their own cluster and are successively merged together to create new clusters as one moves up the hierarchy. A linkage criterion (similarity metric between clusters) determines how clusters are combined together.

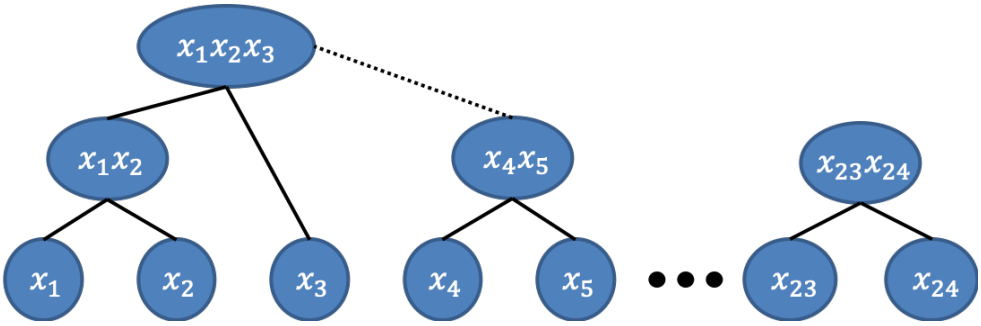


Figure 5.2: An agglomerative hierarchical clustering example with 24 data points

As an example, in Fig. 5.2 , each circle represents a separate data point. At the bottom, each data point is its own cluster. Moving up the hierarchy, the clusters are joined in a step-wise fashion using a linkage criterion, where only neighboring clusters are candidates for merging. Solid lines show clusters that have joined together, and the dashed line represents the next two cluster candidates to be merged if the AHC method is continued. The final clusters in this example are (x_1, x_2, x_3), (x_4, x_5), (x_{23}, x_{24}), and each of the remaining data points as its own individual cluster. Notice how x_1 and x_2 are linked together before x_3 is allowed to join them later. Also, (x_4, x_5) is not permitted to merge with (x_1, x_2) until x_3 offers a bridges between the two clusters. Through these connectivity restrictions, AHC preserves time chronology. The number of clusters decreases as one moves up the hierarchy.

For each of the final clusters, the mean of its data points, the medoid of its data points, or a single data point nearest to the mean/medoid is chosen to represent the entire cluster. This selected data value becomes the representative time period. The size of a cluster is the weight factor applied to the representative period to capture the same total eclipsed time as all members in the cluster. For an example, if x_2 is selected to represent (x_1, x_2, x_3), it receives a weight factor of 3. In the aggregated time horizon, x_2 is cycled 3 times in lieu of x_1, x_2, and x_3 in the full time horizon. This cycling with a weight factor is performed for each cluster such that all the represented periods connected together comprise the aggregated time horizon (Fig. 5.3). In this way, the complexity of an optimization problem with time series data is decreased.

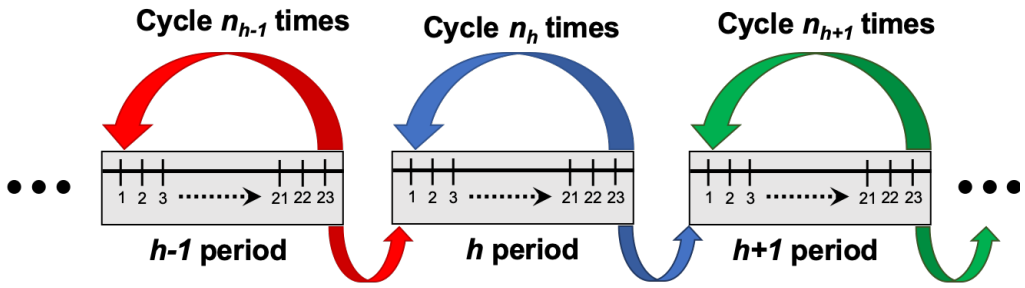


Figure 5.3: Each representative period h is cycled by a weight factor n_h to form the aggregated time horizon that simulates the full time horizon

A commonly used linkage criterion is Ward's distance (W_D), the difference in the within cluster variance (WCSS) before and after joining two clusters together (Eq. 5.1). WCSS is defined as the squared sum of deviations from the cluster mean of all members in the cluster.

$$W_D(C_j, C_k) = \sum_{x_i \in C_{jk}} (x_i - \mu_{jk})^2 - \sum_{x_i \in C_j} (x_i - \mu_j)^2 - \sum_{x_i \in C_k} (x_i - \mu_k)^2 \tag{5.1}$$

W_D between two clusters C_j and C_k is calculated by comparing the WCSS of having their respective data points separately in C_j and C_k versus having them together in one cluster C_{jk}. μ_j, μ_k, and μ_{jk} are the respective cluster means for C_j, C_k, and C_{jk}. W_D is always non-negative for any pair of clusters because some information is lost when agglomerating two clusters together, and this increases the WCSS. Therefore, at each step, AHC merges together the two candidate clusters with the smallest W_D, minimizing the increase in total WCSS for all clusters.

Fig. 5.4 shows an example scree plot that visualizes the total WCSS as a function of the number of clusters. At the right end, having 365 clusters means that every day is its own cluster, leading to a total WCSS value of 0. At the left end, having 1 cluster means that all days are in the same cluster, leading to WCSS being equal to the total variance in the data set. As a result, the more clusters there are, the lower the WCSS becomes. However, there is a trade-off between increasing the cluster complexity and the amount of new information added (measured as the decrease in WCSS). One way to select the optimal number of clusters is to implement a cut-off threshold ϵ

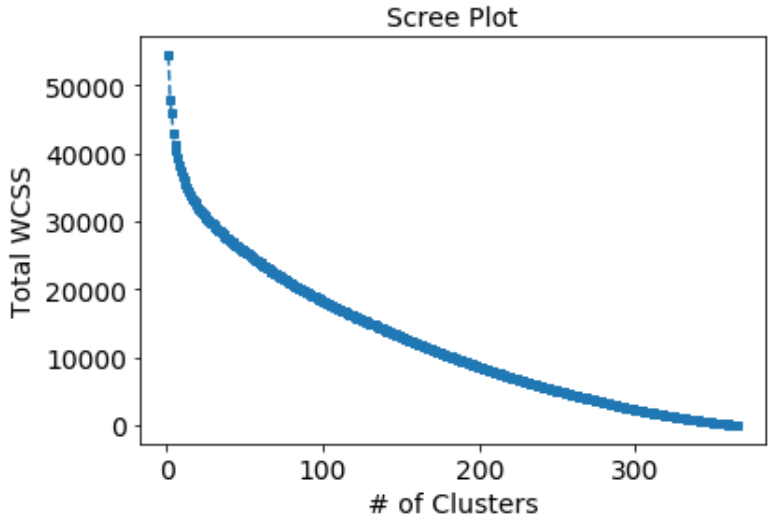


Figure 5.4: An example scree plot of the total WCSS as a function of the number of clusters

measuring the percentage of slope decrease in WCSS (Eq. 5.2).

$$\Delta WCCS\% = \frac{WCSS_n - WCCS_{n-1}}{WCSS_{n-1}} \leq \epsilon \tag{5.2}$$

The cut-off threshold aims to approximate the point at which the curve in Fig. 5.4 begins to bend, indicating that the decrease in variance is slowing with respect to the number of clusters added. Algorithm 1 summarizes the steps to the AHC method.

Algorithm 1 AHC for finding representative periods

1: Set each data point as its own cluster
2: Calculate the cluster mean
3: Compute W_D between each cluster pair
4: Agglomerate two clusters with the smallest W_D
5: Repeat Steps 2-4 until desired number of clusters N is reached
6: Determine representative periods as the cluster mean, medoid, or the cluster member closest to mean/medoid
7: Cluster size is the weight factor of each representative period in the aggregated time horizon

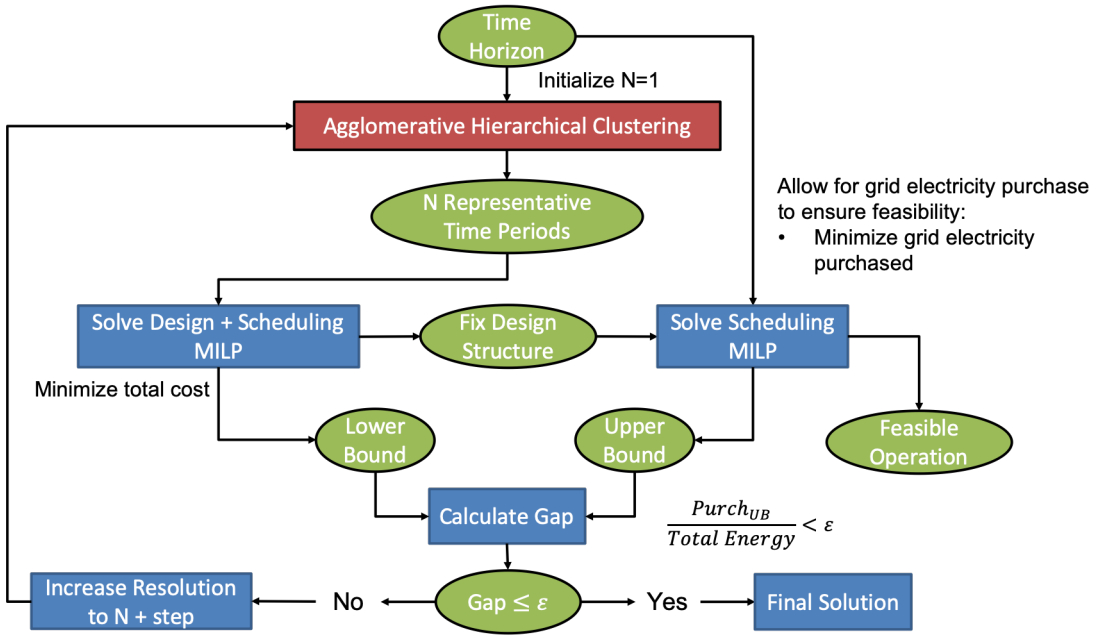


Figure 5.5: Decomposition algorithm through AHC

5.3.2 Decomposition Algorithm

In mixed-integer optimization models for simultaneous design and scheduling, there are separate binary variables associated with capacity expansion and unit commitment. The design binaries decide how to size the capacity of an unit (i.e. solar PV, wind turbine, battery) if it is selected, and the scheduling binaries express which time periods the unit operates in. The motivation for a decomposition algorithm is to decouple the design and scheduling decisions and solve for their optimum separately. Fig. 5.5 depicts an overview of the decomposition algorithm through AHC for the design and scheduling of a renewable energy system. Gabrielli et al. [121] and Bahl et al. [129] also discussed similar concepts.

At the initialization step, the full time horizon and data is grouped into a single cluster through AHC. Using this aggregated time horizon, the design and scheduling model is solved, minimizing the total capital and operational cost of the renewable energy system. From this optimal solution, the binaries and capacity variables associated with the design are then fixed. These together with the full time horizon are inputted into a second version of the same design and scheduling model,

but with a new objective of minimizing the purchase of backup electricity to ensure feasible operation. An auxiliary variable is introduced to capture this because the design based on the aggregated time horizon may not be feasible for operating on the full time horizon. Note the purchase of backup electricity is not allowed for the first model.

The optimal solution for the first model becomes a lower bound (LB) to the overall problem, and the optimal solution to the second model is an upper bound (UB). The gap between LB and UB is calculated as the percentage of total energy provided by a backup source in the UB solution ($\frac{Purch_{UB}}{TotalEnergy}$). If this gap is within an acceptable threshold ϵ, the decomposition algorithm terminates. If not, the algorithm continues by increasing the resolution of the aggregated time horizon (a step up in the number of clusters). When the algorithm exits, the optimal solution to overall system are the design variables from LB and the operating variables from the UB.

The decomposition algorithm through AHC for an optimization-based design and scheduling model of a renewable energy system is summarized in Algorithm 2. In general, a similar approach is possible with any optimization problem that involves time discretization. In the following section, Algorithm 2 is applied to investigating a MILP design and scheduling model for a renewable power system with battery storage in NYC.

Algorithm 2 Decomposition algorithm through AHC

1: Determine the N clusters using Algorithm 1; $N = 1$ if this is initial step
2: Solve the design and scheduling optimization, minimizing total cost
3: Fix the design-related variables from Step 2 and incorporate back the full time horizon
4: Solve the design and scheduling optimization, minimizing auxiliary variable for backup grid electricity cost
5: Calculate gap between the solutions in Step 2 and Step 4
6: Terminate if gap $\leq \epsilon$; otherwise, increase N by a step size and continue back to Step 1

5.4 Renewable Power System with Battery Storage in NYC

The setup for a renewable power system with battery storage in NYC is described below.

5.4.1 Problem Definition

Given:

- time series data of local solar & wind availabilities, demand loads, and electricity prices
- land, capital, and operating costs for solar PV, wind turbine, battery, and DC-AC inverter
- process performance coefficients for the above units

Solve for:

- minimal levelized cost to meet electricity demand (LCOE)
- optimal capacity sizing for all units
- optimal operating schedule for all units

5.4.2 Process Description & Model

A resource-task network (RTN) portrayal for the renewable power system with battery storage in NYC is shown in Fig. 5.6. Resources are represented by circles, and the process tasks are shown with rectangles. Solar radiation, wind, battery charge, AC power, and DC power are resources. Process tasks are the units (solar PV, wind turbine, DC-AC inverter, and battery) utilized to transform one resource into another. Only the battery charge is a storable resource, and this is highlighted by its darker color. AC power is the demand load that needs to be satisfied. When allowed, grid purchase of electricity is fed directly to AC power.

Demirhan et al. [115] describes a general MILP model for a RTN representing the simultaneous design and scheduling of a renewable power system with storage. In their study, the usage of dense energy carriers storing and transporting renewable energy from Amarillo, TX to NYC is investigated. Similar equations are adapted here for a design and scheduling model for the local production of electricity in NYC from solar PV and wind turbine with battery storage. The complete MILP model and its information are included in the Appendix A.2.

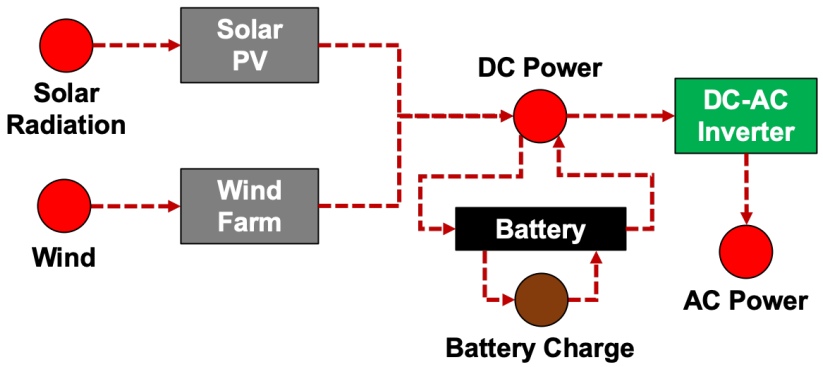


Figure 5.6: RTN representation of a renewable power system with battery storage in NYC

5.4.3 Data Input

Solar DNI and wind speed data are retrieved from NREL NSRDB [111] and Wind Prospector [112], respectively. Demand load and grid electricity prices are taken from the NYISO [131]. Available land for solar and wind development is assumed to be located in Suffolk County, NY. From a GIS analysis [113], these values are 839 km^2 and 164 km^2, respectively. All other assumptions remain the same as in Demirhan et al. [115].

The X matrix for the system depicted in Fig. 5.6 comprises daily solar, wind, demand load, and electricity price data. X is a 365×96 matrix, where the rows are days and the columns are feature values for a particular hour during the day. The AHC method (Algorithm 1) is implemented in Python using packages from the scikit-learn library. Setting a threshold of $\epsilon = 0.01$ in Eq. 5.2, the scree plot of ΔWCSS% versus number of clusters in Fig. 5.7 is produced. From a clustering error perspective, it appears that 15 clusters (representative days) are good enough to describe the variance in X.

5.5 Case Studies

While the 15 representative days selected in Section 5.4.3 minimized ΔWCSS%, they might not necessarily comprise the best time aggregation to approximate the true optimal LCOE from an objective error perspective. Moreover, how sensitive the optimal solution is to the number of clusters is unknown, and what is the least number of representative days needed to adequately

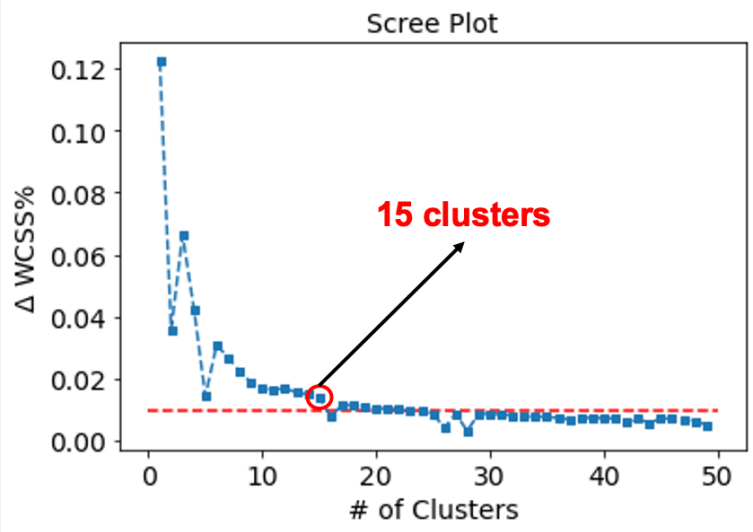
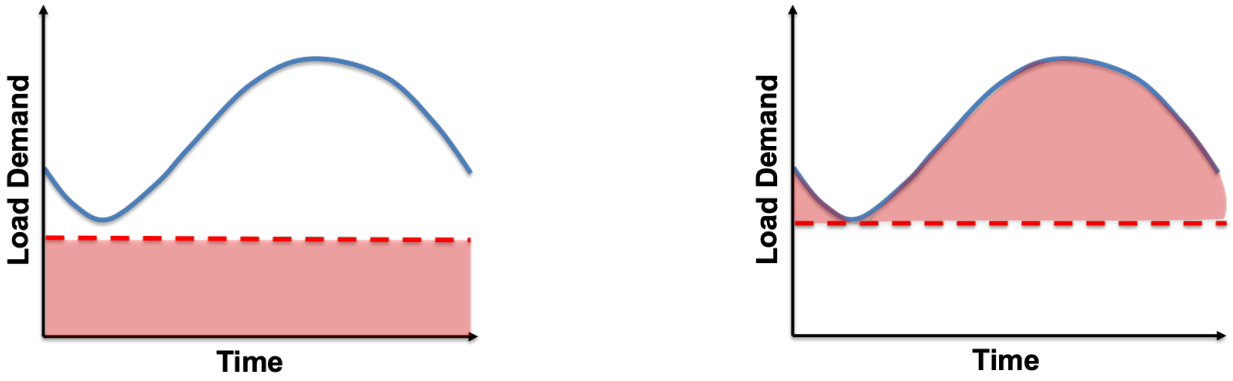


Figure 5.7: Scree plot of the ΔWCSS% as a function of the number of clusters for the NYC time series data

approximate the true optimal LCOE is unanswered.

To address these issues, the 15 representative days are used as a baseline to compare the resulting solutions from implementing Algorithm 2. The different objective functions in the decomposition algorithm (Fig. 5.5) for the LB and UB are found in Appendix A.2. The two types of case studies vary in terms of the demand profile that is met in NYC. Base load (Fig. 5.8a) refers to a flat profile that constitutes 5% of the annual power consumption in NYC. Peak shaving (Fig. 5.8b) tracks the demand fluctuations above a minimum load, and 25% of this total variation is met. The total power demand for both demand profiles is about 2,700 GWh/yr.

The case studies are labeled as [DP]-[RE]-[C], where DP is the demand profile followed (BL: base load, PS: peak shaving), RE is the renewable energy resources utilized (S: solar, SW: solar & wind), and C is the number of representative days in the aggregated time horizon. For example, BL-S-10 corresponds to a solar PV supplying the base load in NYC, modeled using 10 representative days. Battery storage is present in all case studies. In following subsections, the 15 representative days from Section 5.4.3 and Algorithm 2 are utilized to investigate the optimal LCOE for the renewable power systems with battery storage for base load and peak-shaving.



(a) Base load

(b) Peak-shaving

Figure 5.8: Blue line is the an example of the daily demand load in NYC; demand profiles to meet with renewable power are shown in red (not drawn to scale)

Unless otherwise stated, cluster sizes (N=1, N=5 to N=20 by 5, N=30 to N=60 by 10, N=80 to N=200 by 20, N=225 to N=325 by 25, & N=365) and $\epsilon \approx 5\%$ are assumed for the decomposition algorithm. All MILP models are coded in GAMS and solved using CPLEX.

5.5.1 Base Load

BL-S case study results are shown in Table 5.1. BL-S-5 is the first solution that terminates Algorithm 2. BL-S-15 is the baseline solution with clusters determined from ΔWCSS%, and BL-S-40 is the first solution that leads to $\epsilon = 0$ (100% of power is generated without backup). Subsequent tables follow this logic as well.

The first important observation is that BL-S-5 and BL-S-15 both calculate similar LCOE and ϵ values. This suggests that 5 representative days captures enough of the input data variance to give as good of an approximation to the optimal LCOE as 15 representative days does. In terms of ϵ, there is no extra information gained from adding an additional 10 clusters to the aggregated time horizon. This reiterates previous results that minimizing clustering error is not the same as approximating the objective value better.

The second important observation is that the optimal LCOE with $\epsilon = 0$ for BL-S-40 is much higher compared to the other two case studies. This dramatic difference is largely due to the

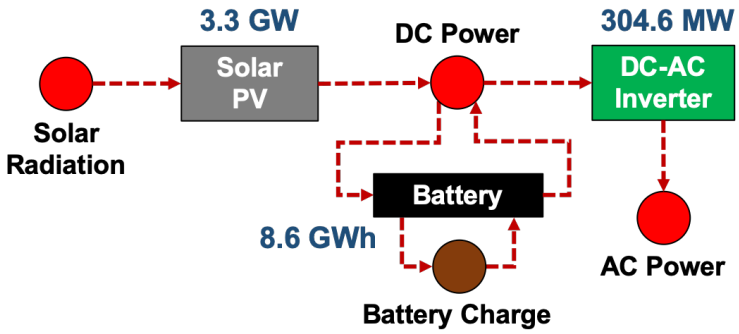


Figure 5.9: RTN for BL-S-5

Table 5.1: BL-S case study results

Case ID#	LB LCOE ($/kWh)	UB LCOE ($/kWh)	CAPEX ($MM)	OPEX ($MM/yr)	Total Cost ($MM/yr)	ϵ (% Backup)
BL-S-5	0.221	0.223	7057	30.27	594.86	4.93
BL-S-15	0.220	0.222	7032	30.29	592.86	4.96
BL-S-40	0.704	0.704	23136	26.99	1877.92	0

increased capital investment cost involved in battery oversizing to account for bad weather days and completely meet the demand. In BL-S-40, the capacities for solar PV and battery are 2.7 GW and 57.9 GWh, respectively. This is compared to a 3.3 GW solar PV and 8.6 GWh battery in BL-S-5 and BL-S-15. Fig. 5.9 shows the network representation of BL-S-5.

Therefore, it becomes economically unfavorable to enforce complete demand satisfaction ($\epsilon = 0$) from solar PV and battery for 5% base load in NYC. ϵ is viewed as an auxiliary variable representing how much demand is negotiable to be met by a backup source instead of the main renewable power system. Here, with about 5% of the demand coming from backup, the LCOE for BL-S-5 is about 30% that of BL-S-40. While the backup source considered in this study is grid power, Demirhan et al. [115] noted that this source could be an energy carrier such as hydrogen coming from another geographical location with dense renewable resources.

Another possible source of extra energy is utilizing the local wind in addition to solar. BL-SW case study results are shown in Table 5.2. The optimal LCOE for BL-SW-15 is about half that

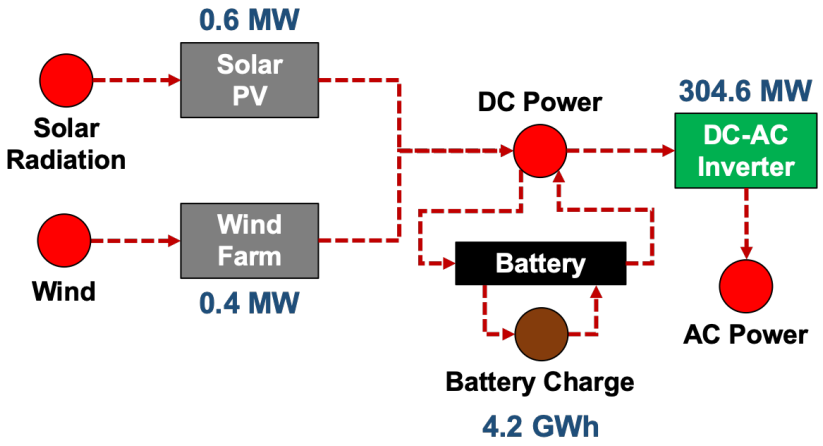


Figure 5.10: RTN for BL-SW-10

of BL-S-15. Wind complements solar in making up for when it is not available. The additional power from wind (0.4 GW) reduces the capacity sizes of the solar PV and battery to 0.6 GW and 4.2 GWh, respectively. This decreases the capital investment cost of BL-SW-15 by over 50% compared to BL-S-15, with a modest increase in operational cost because wind turbines are more expensive to operate. The reduction in solar PV and battery sizes reiterates the earlier point that having an additional energy source can avoid oversizing to handle bad weather days. Fig. 5.10 shows the network representation of BL-SW-10.

For the BL-SW cases, there is also little information gained from using 15 representative days over 10 days in the aggregated time horizon. BL-SW-10 has nearly the same results as BL-SW-15. This again highlights that maximizing the amount of variance covered from the input data is not equivalent to better approximating the objective value. Even though wind energy is now utilized, a similar escalation in the LCOE value is observed when $\epsilon = 0$ for BL-SW-225. In this case, the solar PV, wind turbine, and battery capacities are 1.2 GW, 0.3 GW, and 34.4 GWh, respectively.

Capital costs associated with oversizing again becomes an issue in enforcing every bit of the demand to be satisfied by the renewable power system. BL-SW-10 has a LCOE value that is about 25% that of BL-SW-225. Therefore, even with added wind power, some backup energy source is still required to completely meet the demand with renewable energies in an economic manner.

Table 5.2: BL-SW case study results

Case ID#	LB LCOE ($/kWh)	UB LCOE ($/kWh)	CAPEX ($MM)	OPEX ($MM/yr)	Total Cost ($MM/yr)	ϵ (% Backup)
BL-SW-10	0.106	0.109	3195	34.21	289.86	4.13
BL-SW-15	0.107	0.110	3237	34.22	293.19	4.08
BL-SW-225	0.429	0.430	13971	31.97	1149.63	0

This reinforces the motivation and support for energy carriers that are produced with renewable energies in a different geographical location [132].

Finally, it appears that many more clusters are required in the BL-SW case studies to capture the wind variability and drive ϵ to 0 than in the BL-S case studies. In general, needing more representative days in the aggregated time horizon for ϵ to approach 0 is a common trend observed in the BL-S and BL-SW case studies. As more representative days are included in the aggregated time horizon, the complexity to solve the MILP model approaches the original full time horizon.

5.5.2 Peak Shaving

Peak-shaving results for PS-S and PS-SW case studies are shown in Tables 5.3 and 5.4. Figs. 5.11 and 5.12 display the network descriptions of PS-S-10 and PS-SW-10, respectively. Similar conclusions as those in the base load studies are also observed for these peak-shaving results.

Table 5.3: PS-S case study results

Case ID#	LB LCOE ($/kWh)	UB LCOE ($/kWh)	CAPEX ($MM)	OPEX ($MM/yr)	Total Cost ($MM/yr)	ϵ (% Backup)
PS-S-10	0.188	0.214	7035	31.38	594.17	5.28
PS-S-15	0.205	0.213	7030	31.38	593.79	5.29
PS-S-30	0.849	0.850	32079	24.15	2590.47	0

First, a fewer number of representative days than 15 is needed to adequately capture the LCOE objective value. In both PS-S and PS-SW cases, 10 representative days give similar results as their

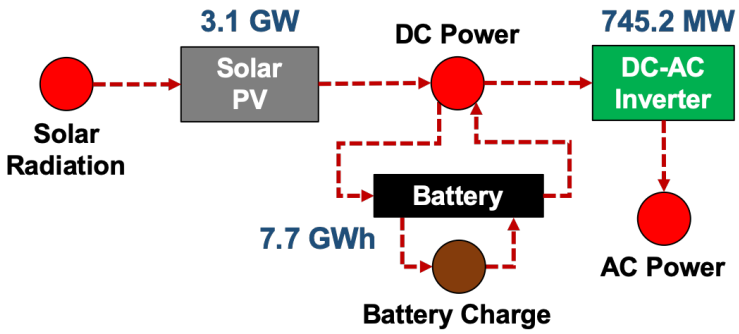


Figure 5.11: RTN for PS-S-10

Table 5.4: PS-SW case study results

Case ID#	LB LCOE ($/kWh)	UB LCOE ($/kWh)	CAPEX ($MM)	OPEX ($MM/yr)	Total Cost ($MM/yr)	ϵ (% Backup)
PS-SW-10	0.102	0.117	3646	35.48	327.17	5.03
PS-SW-15	0.111	0.117	3645	35.47	327.11	5.04
PS-SW-225	0.325	0.325	10951	30.12	906.17	0

15 day counterparts. This shows further empirical proof that minimizing clustering error does not necessarily lead to a better approximation of the the objective value. Second, for the demand to be completely satisfied by a renewable power system ($\epsilon = 0$), several more representative days comprise the aggregated time horizon, and the LCOE values grow prohibitively expensive. Again, this suggests the necessity for some sort of energy backup to avoid oversizing the battery capacity. The battery sizes in PS-S-30 and PS-SW-225 are 72.9 GWh and 22.9 GWh, respectively. Moreover, solving the MILP model with an aggregated time horizon comprising a very large number of representative days is almost as difficult as solving the original model with the full time horizon. Third, adding in wind energy into the mix makes the power economics more attractive than using solar on its own, but does not solve the backup issue.

There is one interesting observation from the peak-shaving studies. Despite outputting the same amount of energy (2,700 GWh/yr) during base load or peak-shaving operation, the renewable power system has a slightly lower LCOE for peak-shaving than base load. For solar-only cases

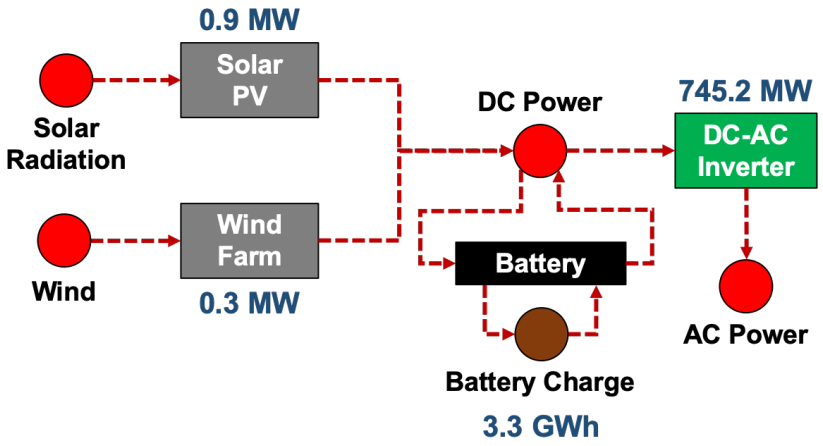


Figure 5.12: RTN for PS-SW-10

(BL-S-5 vs. PS-S-10), there is about a \$0.03/kWh difference, and for the added-in wind cases (BL-SW-10 vs. PS-SW-10), there is <0.01/kWh. This suggests that the demand profile has a bigger effect on solar-only power systems. By itself, solar appears cheaper for tracking the demand load peaks and struggles a bit more to satisfy base load operations. Though, once wind is added into the picture, this effect of different demand profiles is mitigated.

5.5.3 Time Aggregation Effect

The effect of time aggregation on ϵ and the optimal LCOE is depicted in Fig. 5.13 for the base load and peak-shaving studies. In general, there is a decreasing trend for ϵ and an increasing trend for LCOE as the number of clusters in the aggregated time horizon grows.

It is clear that the 15 clusters from minimizing ΔWCSS% do not give the best approximation of the optimal LCOE with least error difference. The solar only case studies closed the ϵ gap to 0 with a reasonable number of clusters, while the BL-SW and PS-SW studies struggled to do so. This is due to the increased variability in the wind data compared to solar data. Finding representative hours instead of days [125] may be more advantageous for studying wind power systems.

Moreover, depending on the decision-maker's specified tolerance on ϵ, 15 clusters may be an overestimation or underestimation on the suitable number of clusters to approximate the optimal LCOE. Therefore, selecting an appropriate number of periods in the time aggregation should be

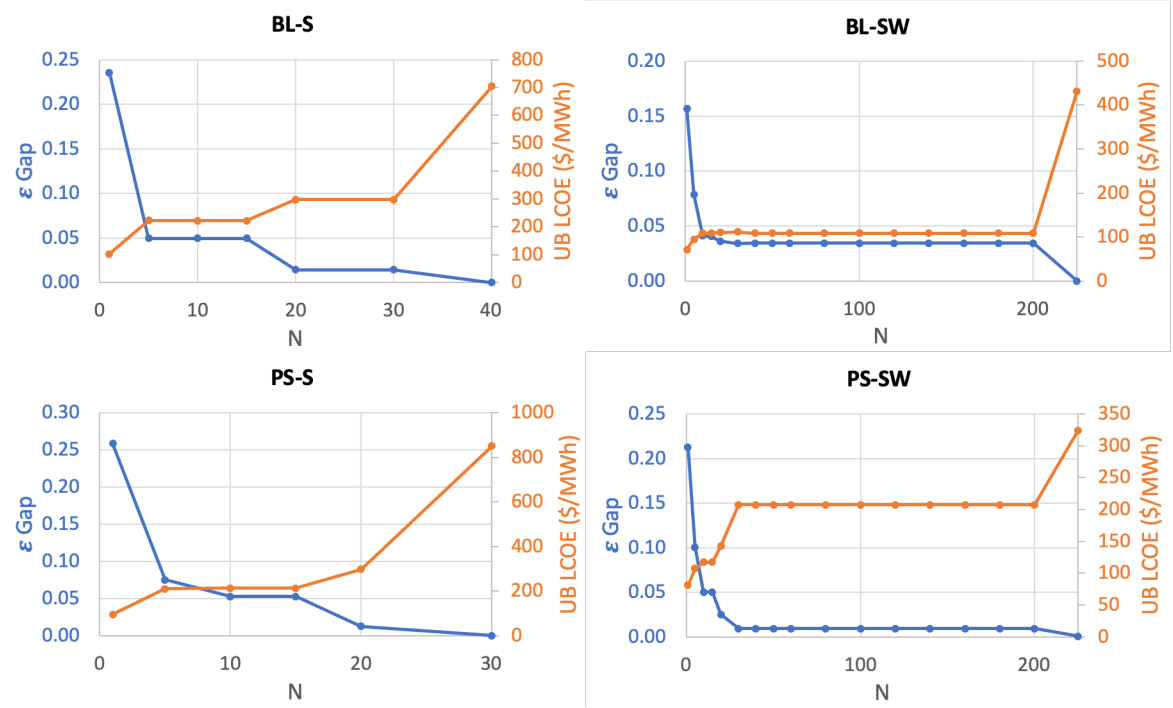


Figure 5.13: ϵ and LCOE vs. cluster size N in time aggregation

evaluated in the objective function space during the optimization solution and not through comparing statistical error metrics like WCSS during data processing. The ability to assess clustering performance in the objective value domain is the major benefit from a decomposition algorithm.

5.5.4 Breakdown Summary

Capital cost breakdowns for select base load and peak-shaving case studies (BL-S-5, BL-SW-10, PS-S-10, PS-SW-10) are shown in Fig. 5.14. Increased costs from oversizing solar PV and battery in BL-S-5 and PS-S-10 are clearly reduced from the introduction of wind turbines into the mix in BL-SW-10 and PS-SW-10. Wind has a complementary effect in making up for the poor solar days. There is no substantial difference in capital investment cost for a renewable power system following base load or peak-shaving demand profiles, assuming some backup is available.

Operational cost breakdowns for these select base load and peak-shaving studies are shown in Fig. 5.15. It is observed that introducing wind turbines into the system slightly increases the overall operational cost. Moreover, the majority of the cost is associated with renewable power

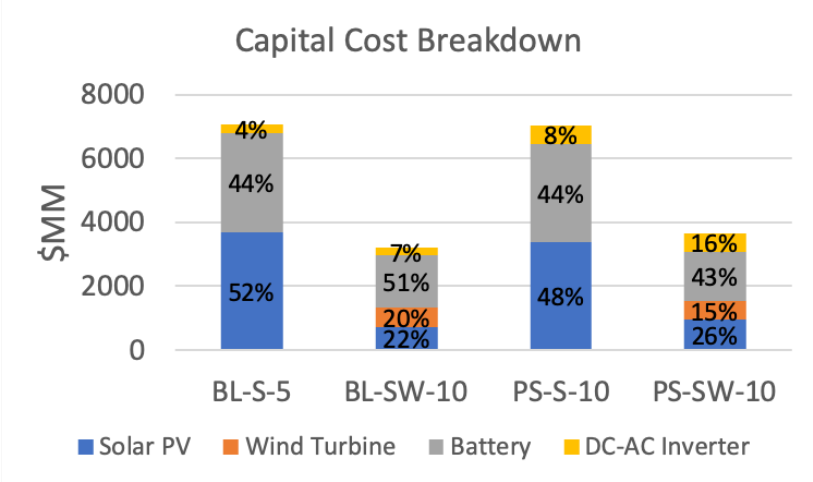


Figure 5.14: Capital cost breakdown for select case studies

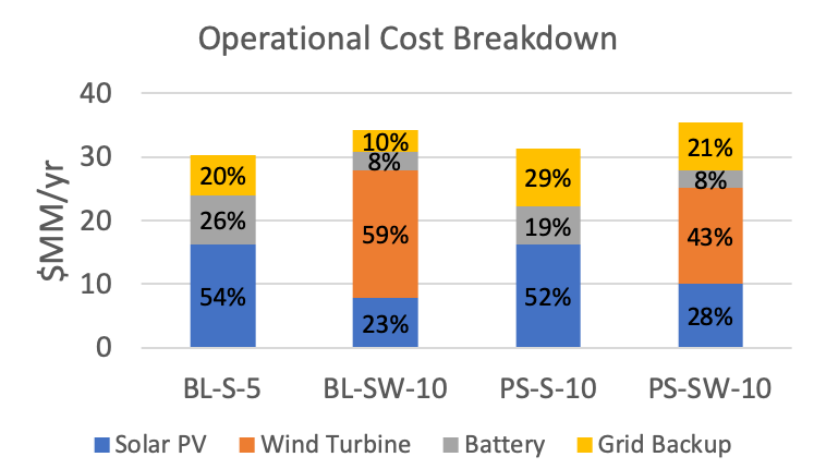


Figure 5.15: Operational cost breakdown for select case studies

generation. Solar PV comprise about half of the operational cost in BL-S-5 and PS-S-10, while solar PV and wind turbine comprise about three-fourths of the operational cost in BL-SW-10 and PS-SW-10. Grid backup is never more than 30% of the total operating cost, and batteries are at most about a quarter of the total operating cost.

The share of power demand satisfied through solar, wind, battery, or grid backup for these select base load and peak-shaving studies is displayed in Fig. 5.16. The inclusion of wind turbines shifts some load burden away from the battery. This is the main driver for the decreased investment

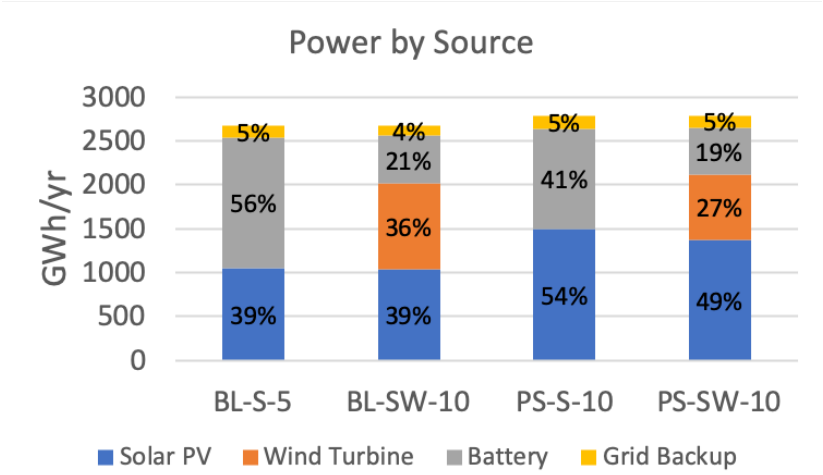


Figure 5.16: Power demand share for select case studies

costs observed in BL-SW-10 and PS-SW-10 compared to BL-S-5 and PS-S-10. Power from solar is about 40% in the base load studies and about 50% in the peak-shaving studies, suggesting that it is may be more appropriate for the latter application.

5.6 Conclusion

Time aggregation through clustering is one method to solve capacity expansion and unit commitment optimization models for renewable power systems with storage. In this work, a decomposition algorithm through agglomerative hierarchical clustering (AHC) is developed and applied to investigating a solar PV and wind turbine system with battery storage in NYC. The renewable power system is described using a previously developed MILP design and scheduling model.

Results showed that selecting clusters in the aggregated time horizon based off of minimizing ΔWCSS% overestimated the number of representative periods necessary to capture the LCOE within a reasonable tolerance. Nevertheless, a general trend of increasing the number of clusters and decreasing ϵ is observed. If strict demand satisfaction ($\epsilon = 0$) is enforced, LCOE values for the renewable power system grow unreasonably large as the number of representative days greatly increase, approaching the full time horizon. Therefore, the decomposition algorithm (with a ϵ threshold) quantifies for the decision-maker how much energy backup is required to maintain the

LCOE within a reasonable value and keeps the optimization problem manageable to be solved.

This motivates the design and scheduling of renewable power systems meeting a flexible demand, allowing some violation within a tolerance level ($\epsilon \leq tol.$), instead of a fixed demand profile to avoid oversizing issues. Further investigation is also needed to compare and quantify different clustering strategies within the decomposition algorithm. Finally, the renewable power system in NYC is a relatively small problem. Additional refinements to the decomposition algorithm may be necessary as the problem size grows larger.

6. PARAMETRIC PROGRAMMING FOR HYPERPARAMETER OPTIMIZATION OF MACHINE LEARNING MODELS

6.1 Motivation

Let *X* be a $i \times j$ data matrix, and *Y* be a $i \times 1$ response vector, where *i* is the sample size and *j* is the number of predictors. In supervised learning (regression or classification) problems, it is assumed that there exists a function f that maps the relationship between a set of input predictors $X = (X_1, X_2, \ldots, X_j)$, where X_j is the jth column vector of *X*, and output responses *Y*.

$$Y = f(X) + \epsilon \tag{6.1}$$

f represents the learnable information that *X* provides about *Y*, while ϵ is a random error term containing information that is unmeasured or unavailable in the data for the learning process. Because f is not exactly known, machine learning algorithms are needed to estimate f and predict *Y*.

$$\hat{Y} = \hat{f}(X) \tag{6.2}$$

$\hat{f}$ represents the estimate for f, and $\hat{Y}$ is the resulting prediction. ϵ is not included in the prediction because it averages out to be zero.

The accuracy of this estimation and prediction is the squared error between *Y* and $\hat{Y}$. Assuming $\hat{f}$ is fixed, the expected error of a single predicted point $\hat{y}_0$ from one observation x_0, a given row vector of *X*, is decomposed into reducible and irreducible quantities [133].

$$\begin{aligned} \mathbf{E}[y_0 - \hat{y}_0]^2 &= \mathbf{E}[y_0 - \hat{f}(x_0)]^2 \\ &= [f(x_0) + \epsilon - \hat{f}(x_0)]^2 \\ &= \underbrace{[f(x_0) - \hat{f}(x_0)]^2}_{\text{reducible}} + \underbrace{Var(\epsilon)}_{\text{irreducible}} \end{aligned} \tag{6.3}$$

Because $\hat{f}$ is not a perfect estimate for *f*, this inaccuracy introduces some error. This error is reducible because it is possible to improve the fit of $\hat{f}$ by using a better performing algorithm. Even if $\hat{f}$ were to exactly match *f*, the prediction of Y still has some error associated with it due to ϵ. This error is irreducible because a model cannot account for information that is not contained in the data while the algorithm is learning. This provides an upper bound [134] on the accuracy of any estimated $\hat{f}$.

As such, the goal of any machine learning method is to minimize the reducible error in order to maximize the accuracy of $\hat{f}$ to be closer to its upper bound. The reducible error is made up of two components: bias & variance [133].

$$\underbrace{[f(x_0) - \hat{f}(x_0)]^2}_{\text{reducible}} = \underbrace{Var[\hat{f}(x_0)]}_{\text{variance}} + \underbrace{Bias[\hat{f}(x_0)]^2}_{\text{bias}} \tag{6.4}$$

Bias is the error introduced by approximating a real-world phenomena, which is often complex, with a simpler model that has less fidelity. Variance measures the sensitivity of $\hat{f}$ to the training data set and how much its fit would change if estimated using different data. In general, as a model becomes more complicated or flexible (including more parameters to estimate), bias decreases and variance increases [133, 134]. A flexible $\hat{f}$ fits closer to given training data set (lowering bias), but is more sensitive to training data variability (raising variance). The reverse of this bias & variance trade-off for $\hat{f}$ is also generally true for a simpler model. Fig. 6.1 illustrates the bias & variance trade-off with respect to model complexity.

Therefore, selecting an optimal $\hat{f}$ from a set of various candidate ones, ranging in complexity from simple linear to highly nonlinear, involves balancing the bias & variance trade-off. The ideal machine learning method is one that simultaneously achieves low bias and low variance [135]. Possessing these characteristics gives the learned model a higher probability of generalizing well to new data unseen during model training and predict more accurately [134]. Too much variance leads to $\hat{f}$ overfitting the data, and too much bias leads to $\hat{f}$ underfitting the data (Fig. 6.2).

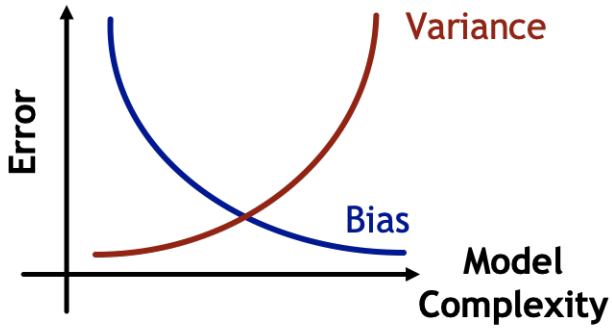


Figure 6.1: More complex models generally have lower bias and higher variance, while less flexible models generally have higher bias and lower variance

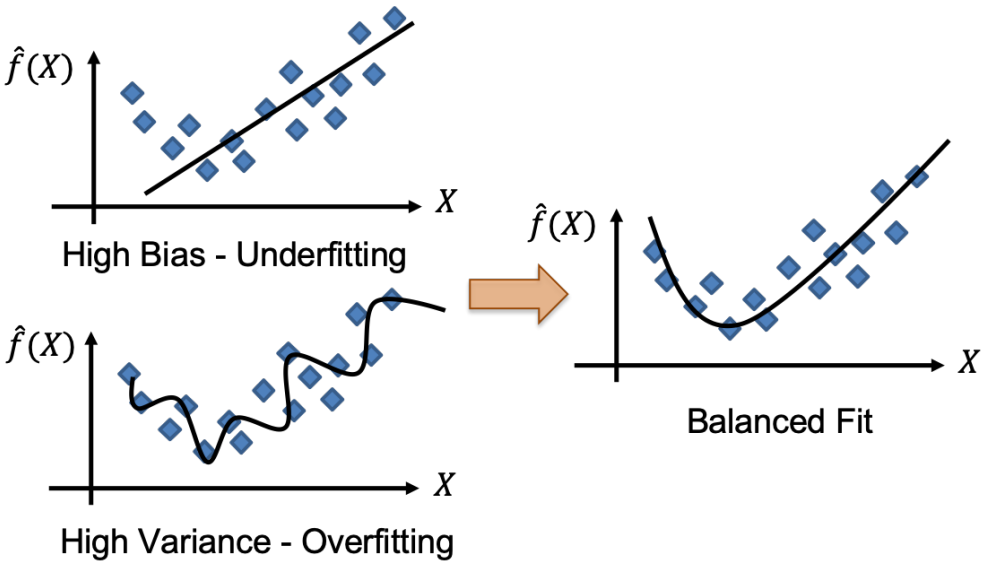


Figure 6.2: Finding a good $\hat{f}$ fit requires balancing bias and variance

6.2 Background

The most utilized approach for finding an optimal $\hat{f}$ is to incorporate an additional regularization term in a machine learning algorithm's loss function formulation [136]. Typically, the basic loss function for a supervised learning problem is the minimization of the mean squared error (MSE) between *Y* and $\hat{Y}$, where *N* is the sample size.

$$\min \quad \frac{1}{N}\|Y - \hat{Y}\|_2^2 = \frac{1}{N}\|Y - \hat{f}(X)\|_2^2 \tag{6.5}$$

A regularization term is an exogenous penalty parameter (hyperparameter) whose value is set prior to training the model. This hyperparameter λ controls the importance and weight of the regularization term, which affects the resulting optimization solution of a machine learning algorithm. A common regularization term [136] is λ penalizing the *q*-norm of w, the model weights of $\hat{f}$ (an example is $\hat{f}(X) = Xw$), raised to the power *p*.

$$\min_{w} \quad \frac{1}{N}\|Y - \hat{f}(X)\|_2^2 + \lambda\|w\|_q^p \tag{6.6}$$

In general, machine learning algorithms may have multiple hyperparameters that are prespecified [137, 138], such as a λ_j for each w_j.

In Eq. 6.6, the aim of regularization is to control the complexity of $\hat{f}$ that is fitted. As the value of λ varies from 0 to ∞, the resulting estimated $\hat{f}$ will have different reducible error realizations. Likewise, the bias and variance values for each $\hat{f}$ are different. By including a regularization term, the reducible errors for several $\hat{f}$ candidates are comparable and are an implicit function of the hyperparameter. Finding an optimal $\hat{f}$ with low bias and low variance amounts to correctly tuning λ. However, what is the best value for λ is not known *a priori*. Therefore, the selection of an optimal machine learning model is really a hyperparameter optimization problem (Fig. 6.3).

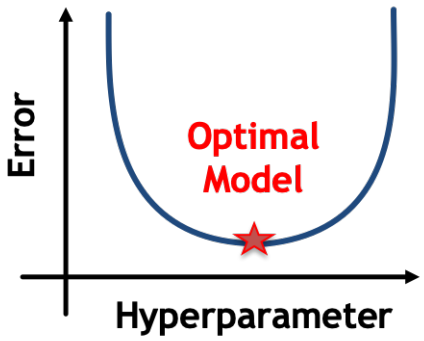


Figure 6.3: Optimal model selection amounts to determining the optimal hyperparameter

6.2.1 Hyperparameter Optimization

Common strategies for hyperparameter optimization [139–142] involve dividing the parameter space into D evenly or randomly discretized points and performing an iterative optimization procedure through k-fold cross-validation (Fig. 6.4). First, the data is split into K subsets. Within each subset, the data is further separated into training and testing sets. Next, for each discretized λ value, a separate optimization problem for the machine learning model (Eq. 6.6) is constructed on the training data in each fold and solved to estimate $\hat{f}$. The validation error $\frac{1}{N}\|Y - \hat{Y}\|_2^2$ is then computed using the estimated $\hat{f}$ and testing data in the same fold to predict Y. Finally, after iterating through all the λ values, the validation errors for each λ across all folds are averaged together. The optimal $\hat{f}$ is the one with the λ value that gives the smallest mean validation error.

K-fold cross-validation with grid [143] or random search [144] for hyperparameter optimization is a generalizable way to approximate optimal model selection. The advantages of k-fold cross-validation are that, in most cases, it captures the actual test error as an implicit function of λ well-enough and calculates a $\hat{f}$ that balances bias & variance [133, 134]. The disadvantages are that sometimes KxD optimization problems may become a computational burden to solve and possibly lead to inexact solutions that are troublesome due to poor discretization.

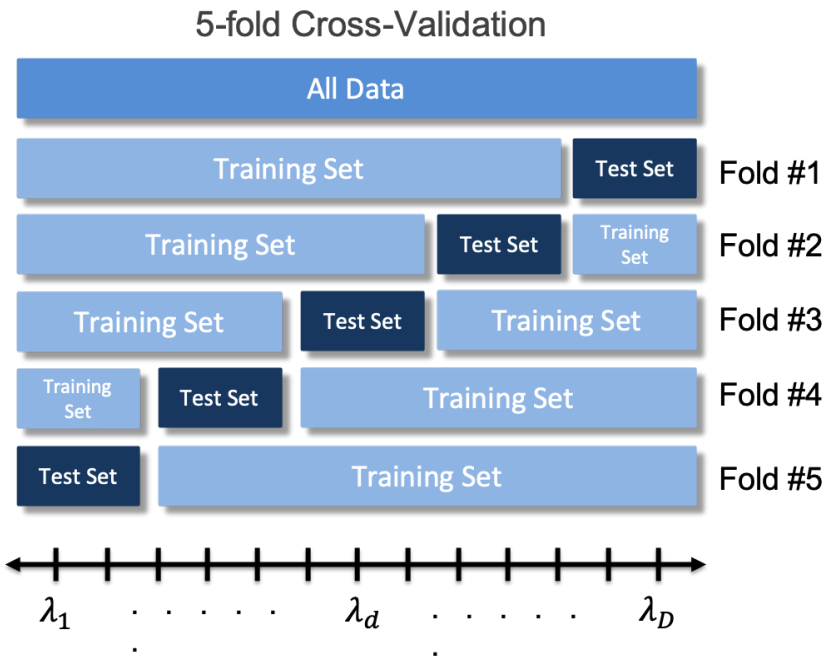


Figure 6.4: An overview of k-fold cross-validation for hyperparameter optimization

6.3 A Parametric Programming Perspective

Instead, the developments in this work translate the k-fold cross-validation for hyperparameter optimization into a bilevel optimization problem that is exactly solvable through parametric programming without any approximation or probabilistic modeling. This applies for hyperparameter optimization problems that have a machine learning algorithm that is formulated as a linear or quadratic programming (LP/QP) model.

Parametric programming is an optimization strategy, popularized by explicit model predictive control (MPC) [145], that determines the optimal solution as a function of a varying parameter θ, without exhaustively traversing the entire parameter space. The general form is seen in Eq. 6.7. The objective (loss) function F, inequality constraints g, and equality constraints h are all functions of the decision variables u and parameter θ. The optimal solution comprises a set of finite areas (critical regions in Fig. 6.5), where a particular solution is valid for a given realization of θ, along with explicit expressions relating the decision variables u to the θ parameter ($u = A\theta + b$). Using

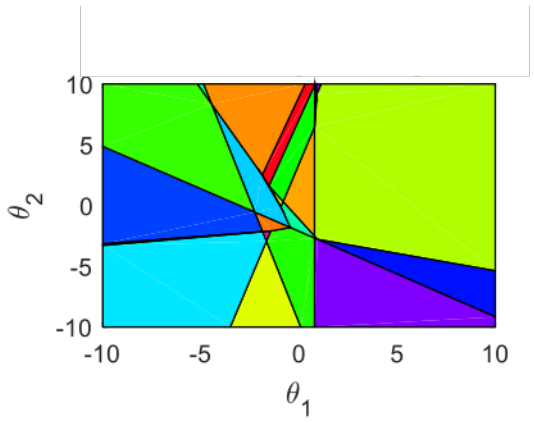


Figure 6.5: An example of critical regions comprising the optimal solution to a parametric programming problem (Eq. 6.7)

this, the objective (loss) function is also solely expressed as a function $J(\theta)$.

$$\begin{aligned} J(\theta) = \quad & \min_{u} F(u, \theta) \\ & \text{s.t. } g(u, \theta) \leq 0 \\ & \quad h(u, \theta) = 0 \end{aligned} \tag{6.7}$$

In the classic explicit MPC problem [145], decision variables u are the control inputs and parameters θ are the system states. For the hyperparameter optimization setup, decision variables u are the model weights w to $\hat{f}$ and parameters θ are the hyperparameters λ. In this work, $\hat{f}$ is assumed to be linear with respect to w.

$$\hat{f}(X, w) = \phi(X)w \tag{6.8}$$

w is a $j \times 1$ vector of model weights and ϕ is a vector of basis functions transforming each predictor of X such that $\phi(X) = [\phi_1(X_1), \phi_2(X_2), \ldots, \phi_j(X_j)]$. $\hat{f}$ is a function of w because the model weights are learned from training the machine learning model.

By viewing Eq. 6.6 as an optimization problem in the form of Eq. 6.7, a new understanding of the hyperparameter optimization problem is reached. From this parametric programming perspec-

tive, a different affine expression $w = A_r\lambda + b_r$ governs each critical region *r* of λ and the learned model is redefined as $\hat{f}(X, \lambda) = \phi(X)A_r\lambda + \phi(X)b_r$. Having $\hat{f}$ as an explicit function of λ has important ramifications for the bilevel optimization approach to *k*-fold cross-validation discussed later in Section 6.4.

The algorithmic procedure (graph, geometrical, combinatorial) for computing the critical regions and affine expressions for Eq. 6.7 depends on the optimization problem structure (linear, nonlinear, convex, differentiable) and the nature of the variables & parameters (continuous or binary). In general, parametric programming is also extendable to the case of multiple varying parameters (multi-parametric programming). The reader is referred to several review papers and books for further discussion on multi-parametric programming theory [146–150] and its applications [151–156].

6.4 Bilevel Optimization of k-fold Cross-Validation

Within each *k*th fold of cross-validation for hyperparameter optimization, there are two different objectives. On the training set level, the goal is to minimize the training error in Eq. 6.6, where *Y* is the output response from the training data. On the testing set level, the validation error $\frac{1}{N}\|Y - \hat{Y}\|_2^2$ is evaluated using the learned $\hat{f}$ from model training to predict the output response *Y* from the testing data. After model training and recording the validation errors across all λ values for every fold, the goal is to select the optimal $\hat{f}$ that minimizes the mean validation error across all folds. Overall, *k*-fold cross-validation seeks λ such that when the optimal training is solved for each training set, the validation error over the test errors is also minimized.

In this setup, the dual objectives are captured using a bilevel optimization [157, 158] formulation (Eq. 6.9). Note this is an example formulation, and modifications may be necessary depending on the particular machine learning algorithm, as seen later in Section 6.6. However, the general concepts described in this section remain valid. In the inner level, the objective is to minimize each kth fold's training error with a regularization penalty, the decision variables are the model weights w_k, and the parameter is λ. In the outer level, the objective is to minimize the mean squared

validation error across $|K|$ folds, the decision variable is λ, and the parameters are w_k.

$$\begin{aligned} \min_{\lambda} \quad & \frac{1}{|K|}\sum_{k=1}^{|K|}\frac{1}{N_k^{tst}}\|y_k^{tst}-\hat{y}_k\|_2^2 \\ \text{s.t.} \quad & \min_{w_k} \quad \frac{1}{N_k^{trn}}\|y_k^{trn}-\hat{f}_k\|_2^2+\lambda\|\hat{w}_k\|_q^p \qquad \forall k\in K \end{aligned} \tag{6.9}$$

K is the set of all data folds. For each kth fold, N_k^{trn} is the training set size, N_k^{tst} is testing set size, y_k^{trn} is a $N_k^{trn}\times 1$ vector of output responses in the training set, y_k^{tst} is a $N_k^{tst}\times 1$ vector of output responses in the testing set, $\hat{f}_k$ is the trained machine learning model of form $\hat{f}_k(X,w_k)=\phi(X)w_k$, and $\hat{y}_k$ is a $N_k^{tst}\times 1$ vector of predicted responses from $\hat{f}_k$ using the testing set. Again, $\phi(X)$ is vector of basis functions transforming the columns of *X* and w_k is a $j\times 1$ vector. Note that $\hat{f}_k$ is estimated using X from the training set, $\hat{f}_k(X_k^{trn},w_k)=\phi(X_k^{trn})w_k$, and then X from the testing set is inputted with w_k fixed to predict the output response, $\hat{y}_k=\hat{f}_k(X_k^{tst},w_k)=\phi(X_k^{tst})w_k$.

This observation of *k*-fold cross-validation as a bilevel optimization problem has also been noted by earlier works [159–163]. Some of these authors [159, 160] attempted to solve the bilevel optimization by replacing the inner level optimization problem with its Karush-Kuhn-Tucker (KKT) conditions. The KKT conditions are Lagrangian and complementarity constraints, reducing the bilevel optimization into a single level constrained mixed-integer optimization problem after reformulation. However, nonlinear terms containing Lagrange multipliers and decision variables arise with the complementarity constraints from reformulating the bilevel optimization using the KKT approach. This renders the single level optimization to be a MINLP, a very difficult problem to solve to global optimality.

If the original machine learning algorithm with regularization penalty in the inner level of Eq. 6.9 is well-posed as a LP ($p=1$ & $q=1$) or QP ($p=2$ & $q=2$), parametric programming is another viable strategy for reformulating the bilevel optimization into a single level optimization that is a mixed-integer quadratic (MIQP) problem. Although both methods will give the same optimal λ, the advantage of parametric programming is that it preserves useful information about the optimal solution profile that the KKT approach does not. The KKT approach only provides a

single optimal solution, and a MINLP is more difficult to solve than a MIQP. Through parametric programming, the model weights are derived as explicit affine functions of the hyperparameter ($w_k = A_{kr}\lambda + b_{kr}$) for each kth fold and critical region *r* of λ. As such, $\hat{f}_k$ is expressible as a function of λ, and the machine learning model becomes $\hat{f}_k(X, \lambda) = \phi(X)A_{kr}\lambda + \phi(X)b_{kr}$. Note that exactly one critical region is active (one corresponding pair of A_{kr} and b_{kr} coefficients are nonzero) for each kth fold. This is because λ is the single decision variable in the outer optimization in Eq. 6.9.

This new form of $\hat{f}$ is important because the training error (inner level objective in Eq. 6.9) and the mean validation error (outer level objective in Eq. 6.9) are both now explicit functions of just λ. This is easily seen through substituting the affine function of $w_k(\lambda)$ into the inner and outer objectives. For the training error, $\hat{f}_k(X_k^{trn}, w_k) = \hat{f}_k(X_k^{trn}, \lambda)$, and for the validation error, $\hat{y}_k = \hat{f}_k(X_k^{tst}, w_k) = \hat{f}_k(X_k^{tst}, \lambda)$. The affine function of $w_k(\lambda)$, capturing the optimal solution in the inner optimization, passes information between the two levels of Eq. 6.9 to reduce the bilevel optimization into a single level.

By utilizing parametric programming, the implicit function of error versus λ discussed in Section 6.2 is no longer unknown and now has a closed-form expression. This is a huge advantage for parametric programming over using KKT conditions for hyperparameter optimization because, in addition to the optimal λ, the complete training & validation error versus λ profiles are given. Having these profiles makes it easier to understand the trained machine learning model and visualize the prediction results.

The model formulation of the HY-POP approach for the hyperparameter optimization example in Eq. 6.9 is formally described below. Bilevel optimization through parametric programming (B-POP) [164] for other applications have also been demonstrated in previous works [165–169]. Fig. 6.6 shows an overview of the HY-POP strategy for hyperparameter optimization. Multi-parametric quadratic programming models (mpQP) is the general form of machine learning models that are allowed for the inner level optimization problem.

The first step is to replace the inner level optimization in Eq. 6.9 with constraints that define the

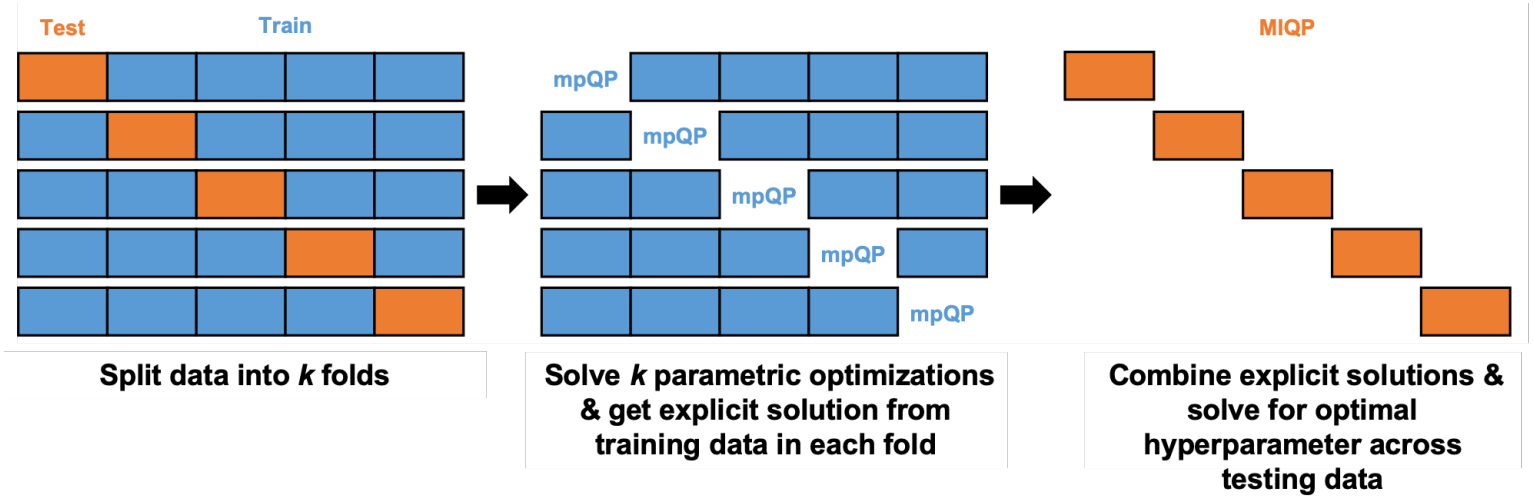


Figure 6.6: HY-POP gives the explicit solutions for the training optimizations through parametric programming and passes along this information to solve for the optimal hyperparameter in a single optimization

optimal solution profile from the critical regions. Since there are $|K|$ folds, there are $|K|$ training optimization problems to replace with parametric programming. The critical regions for each kth training optimization are separately calculated, and then they are all combined together into the appropriate constraint set. These constraints control the affine expressions of $w_k(\lambda)$ and restrict only one critical region to be active for each fold to represent the optimal training. This is conveyed through introducing Big-M constraints (Eq. 6.10), critical region bound constraints (Eq. 6.11), and a SOS1 constraint (Eq. 6.12).

$$\begin{aligned} w_k &\leq A_{kr}\lambda + b_{kr} + M(1 - y_{kr}^{CR}) \qquad \forall k \in K, \forall r \in R^k \\ w_k &\geq A_{kr}\lambda + b_{kr} + M(y_{kr}^{CR} - 1) \qquad \forall k \in K, \forall r \in R^k \end{aligned} \tag{6.10}$$

M is an appropriately large-enough constant value. For each kth fold, R^k is the set of all critical regions that comprise the optimal training solution, w_k is a $j \times 1$ vector of model weights to the trained machine learning model, A_{kr} and b_{kr} are $j \times 1$ coefficient vectors in the affine expression for w_k from a critical region *r*, and y_{kr}^{CR} are binary variables fixing/relaxing w_k for active/inactive critical regions. The Big-M constraints determine which critical region *r* in each fold *k* defines w_k for the machine learning model $\hat{f}_k$.

$$\sum_{r=1}^{|R^k|} LB_{kr}^{CR} y_{kr}^{CR} \leq \lambda \leq \sum_{r=1}^{|R^k|} UB_{kr}^{CR} y_{kr}^{CR} \qquad \forall k \in K \tag{6.11}$$

$$\sum_{r=1}^{|R^k|} y_{kr}^{CR} = 1 \qquad \forall k \in K \tag{6.12}$$

To ensure that only one critical region is active for each fold, Eq. 6.12 enforces this discrete decision. The lower (LB_{kr}^{CR}) and upper (UB_{kr}^{CR}) bounds to λ in a critical region *r* from fold *k* define the range of values for λ such that a particular affine expression for w_k and an resulting optimal solution apply. To enforce that the same λ value is utilized across all folds, when selecting an active critical region in each fold to represent the optimal training, Eq. 6.11 defines that the lower and upper bounds of λ from these $|K|$ critical regions must overlap each other. Together,

Eqs. 6.10, 6.11, & 6.12 replace the inner optimization in Eq. 6.9.

The second step is to substitute $\hat{y}_k = \phi(X_k^{tst})w_k$ into the objective (loss) function of the outer level in Eq. 6.9. Since the inner level decision variable w_k is a function of the outer level decision variable λ from parametric programming, the bilevel optimization is converted into a single level optimization. Eq. 6.13 constitutes the HY-POP reformulation of the hyperparameter optimization example in Eq. 6.9, assuming the machine learning algorithm is a LP ($p = 1$ & $q = 1$) or QP ($p = 2$ & $q = 2$) and $\hat{f}_k$ is linear with respect to its model weights w_k. Because the outer validation error objective is in MSE form and binary variables y_{kr}^{CR} are added for the critical regions, this single level optimization is also a MIQP.

$$
\begin{aligned}
\min_{\lambda, w_k, y_{kr}^{CR}} \quad & \frac{1}{|K|}\sum_{k=1}^{|K|}\frac{1}{N_k^{tst}}\|y_k^{tst} - \phi(X_k^{tst})w_k\|_2^2 \\
\text{s.t.} \quad & w_k \le A_{kr}\lambda + b_{kr} + M(1 - y_{kr}^{CR}) \qquad \forall k \in K, \forall r \in R^k \\
& w_k \ge A_{kr}\lambda + b_{kr} + M(y_{kr}^{CR} - 1) \qquad \forall k \in K, \forall r \in R^k \\
& \sum_{r=1}^{|R^k|} LB_{kr}^{CR} y_{kr}^{CR} \le \lambda \le \sum_{r=1}^{|R^k|} UB_{kr}^{CR} y_{kr}^{CR} \qquad \forall k \in K \\
& \sum_{r=1}^{|R^k|} y_{kr}^{CR} = 1 \qquad \forall k \in K
\end{aligned}
\tag{6.13}
$$

Depending on the actual machine learning algorithm, the example formulations of Eqs. 6.9 and 6.13 may need some modifications. One instance of this is the LP L_1-norm SVM in Section 6.6. However, the general concept shown here of using parametric programming to connect the two levels of a bilevel optimization problem (assuming LP or QP) through affine expressions relating model weights to the hyperparameter remains valid. In Section 6.6, similar steps, as performed for Eqs. 6.9 and 6.13, are taken to formulate the k-fold cross-validation for LP L_1-norm SVM hyperparameter optimization through a HY-POP approach. Nevertheless, the example formulations of Eqs. 6.9 and 6.13 are useful for LASSO regression in the next section.

6.5 LASSO Regression

LASSO [170] is a popular regression technique that performs model selection through regularization. It introduces a L_1-norm penalty on β, a vector of regression coefficients (model weights), to the ordinary least squares (OLS) model. This attempts to improve the regression fit by reducing the variance observed in the OLS estimation for β and better balance the bias & trade-off. LASSO regression is one useful method to build sparse surrogate models for data-driven optimization [171, 172]. The LASSO regression form is shown in Eq. 6.14.

$$\min_{\beta} \quad \frac{1}{2N}\|Y - \phi(X)\beta\|_2^2 + \lambda\|\hat{\beta}\|_1 \tag{6.14}$$

Eq. 6.14 is a parametric programming problem in the form of Eq. 6.7 and fits the problem structure shown in the inner level of Eq. 6.9. Because the L_1-norm regularization term is nonlinear, $\|\beta\|_1 = \sum_j |\beta_j|$, Eq. 6.14 is first reformulated before it is solved through parametric programming. It is observed that β_j is piecewise linear. After substituting $\alpha_j = |\beta_j|$ and adding two constraints to describe the piecewise behavior, the parametric quadratic programming (pQP) model for LASSO regression is the following.

$$\begin{aligned} \min_{\beta,\alpha} \quad & \frac{1}{2N}\|Y - \phi(X)\beta\|_2^2 + \lambda\sum_j \alpha_j \\ \text{s.t.} \quad & \alpha_j \geq \beta_j \qquad \forall j \in J \\ & \alpha_j \geq -\beta_j \qquad \forall j \in J \end{aligned} \tag{6.15}$$

To optimize λ in Eq. 6.15 and select an optimal LASSO model, the bilevel optimization depiction of k-fold cross-validation from Eq. 6.9 is implemented. Eq. 6.15 is the inner optimization problem, and the outer level objective still is to minimize the validation MSE across all folds. The critical regions representing the optimal solution profile to Eq. 6.15 for each kth fold are the similar to those in Eqs. 6.10, 6.11, & 6.12, with the only difference being the parametric programming solution now accounts for the two added constraints for the reformulation of $|\beta_j|$. The final HY-

POP formulation (MIQP) of the LASSO hyperparameter optimization is in Eq. 6.16. β_k is a $j \times 1$ vector of regression coefficients.

$$\begin{aligned}
\min_{\lambda,\beta_k,y_{kr}^{CR}} \quad & \frac{1}{|K|}\sum_{k=1}^{|K|}\frac{1}{N_k^{tst}}\|y_k^{tst}-\phi(X_k^{tst})\beta_k\|_2^2 \\
\text{s.t.} \quad & \beta_k \le A_{kr}\lambda + b_{kr} + M(1-y_{kr}^{CR}) \qquad \forall k\in K, \forall r \in R^k \\
& \beta_k \ge A_{kr}\lambda + b_{kr} + M(y_{kr}^{CR}-1) \qquad \forall k\in K, \forall r \in R^k \\
& \sum_{r=1}^{|R^k|} LB_{kr}^{CR} y_{kr}^{CR} \le \lambda \le \sum_{r=1}^{|R^k|} UB_{kr}^{CR} y_{kr}^{CR} \qquad \forall k \in K \\
& \sum_{r=1}^{|R^k|} y_{kr}^{CR} = 1 \qquad \forall k \in K
\end{aligned} \tag{6.16}$$

Eq. 6.16 is a new form for the hyperparameter optimization of LASSO regression through k-fold cross-validation. Next, this HY-POP formulation is validated on an ammonia reactor data example and against a coordinate descent algorithm (with grid search) from the glmnet package in R.

6.5.1 Ammonia Reactor Data Example

A dataset of 29 samples are collected from different sources on the performance of an industrial ammonia synthesis reactor [99, 102]. The exact data values are included in the Appendix B.3. These values are normalized and centered before training. Reactor temperature T & pressure P, inlet molar concentration of hydrogen x_{H_2}, nitrogen x_{N_2}, ammonia x_{NH_3} & inert species x_{Inert}, and the molar ratio between hydrogen & nitrogen $\frac{x_{H_2}}{x_{N_2}}$ are 7 predictors for the reactor conversion y_X. It is assumed that the predictors are linear, $\phi(X) = X$. Therefore, the proposed LASSO model that is trained has the following form in Eq. 6.17. An intercept term is not included because the data is centered, and therefore, the intercept has been zeroed out.

$$y_X = \beta_1 T + \beta_2 P + \beta_3 x_{H_2} + \beta_4 x_{N_2} + \beta_5 x_{NH_3} + \beta_6 x_{Inert} + \beta_7 \frac{x_{H_2}}{x_{N_2}} \tag{6.17}$$

While this is a small dataset, the goal is not a comprehensive computation study, but to validate

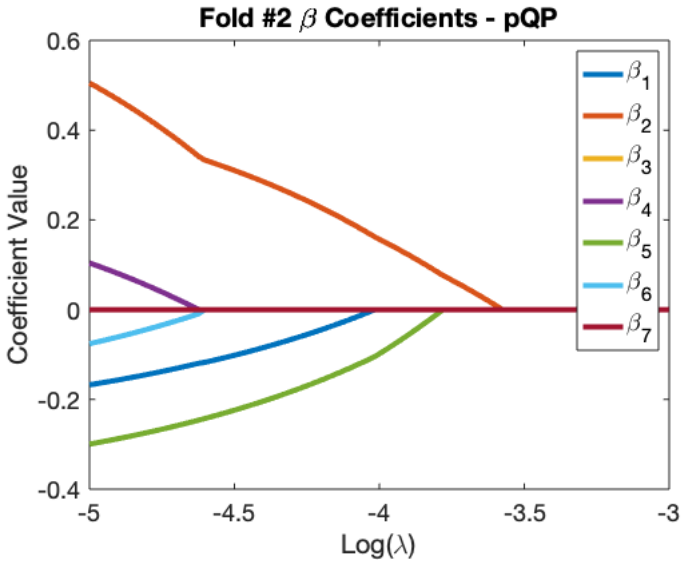


Figure 6.7: LASSO regularization path for the ammonia reactor data in fold #2 from the training pQP (Eq. 6.15) solved using POP

that the HY-POP approach correctly identifies the optimal λ & β, comparing to an established coordinate descent algorithm. The data is randomly divided into 5 folds for cross-validation. The fold identification of the data points is also provided in the Appendix B.3. Each training optimization problem (Eq. 6.15) is formulated in MATLAB 2019b. An in-house developed and state-of-the-art software, the Parametric Optimization (POP) toolbox [173], is then used to solve for the critical regions, using the built-in QP solver from MATLAB and the geometrical algorithm. An example result of the critical regions for fold #2 is shown in Fig. 6.7.

The piecewise linear relationship between β & λ is referred to as the LASSO regularization path [136]. We expected this behavior from the affine expressions $\beta(\lambda)$ provided by the parametric programming solution to Eq. 6.15. Each line segment piece represents a critical region, where a unique $\beta(\lambda)$ function is valid for the values of λ. Having the β change in a piecewise linear fashion has also been previously observed in least angle regression (LAR) [174] and coordinate descent [175], the first efficient algorithms developed to solve LASSO regression. While these methods were specifically tailored for LASSO regression [174,175], parametric programming is the general theory of solving problems in the form presented in Eq. 6.7. Therefore, both LAR and coordinate

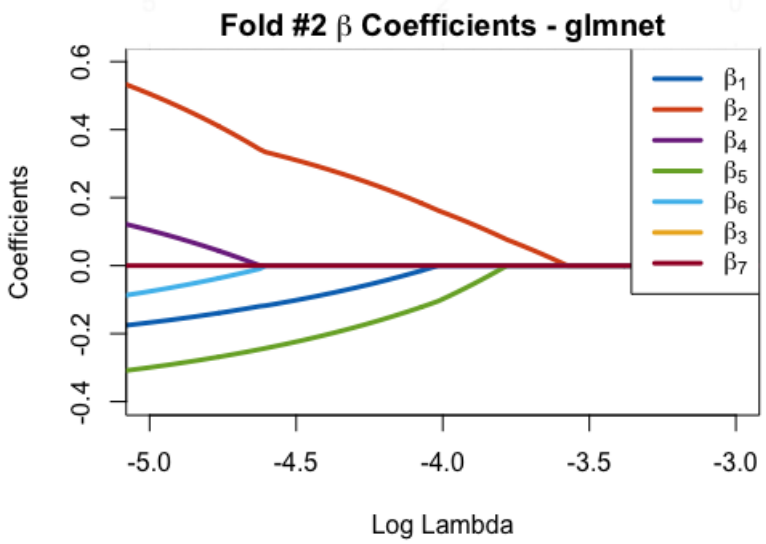


Figure 6.8: LASSO regularization path for the ammonia reactor data in fold #2 solved using a coordinate descent algorithm from glmnet in R

descent algorithms can actually be viewed as specialized parametric programming approaches. In Fig. 6.8, it observed that the coordinate descent algorithm (with 10^3 evenly discretized points for $\lambda \in [10^{-3}, 1]$) gives exactly same regularization path as Fig. 6.7. This verifies that the critical regions for each fold exactly represent the optimal solution profile for the pQP (Eq. 6.15).

With $\beta(\lambda)$ given from parametric programming, calculating the training and validation errors for each fold, the objectives in Eqs. 6.15 & 6.16, respectively, are simple function evaluations. Likewise, these errors are also piecewise functions with respect to λ, but they are not linear due to the squaring of the error term. Fig. 6.9 is an example of this nonlinear piecewise behavior for the testing error in fold #2. With error as a function of λ, finding the optimal λ is an easy calculation, pinpointing the minimum of these validation error profiles aggregated across all folds.

To find this minimum mean validation error, after computing the critical regions, the MIQP for hyperparameter optimization (Eq. 6.16) is formulated in MATLAB 2019b and solved using CPLEX. The MIQP solution is compared to result given from using cv.glmnet, the cross-validation function included in glmnet, with a grid of 10^3 evenly discretized points for $\lambda \in [10^{-3}, 1]$. Fig.

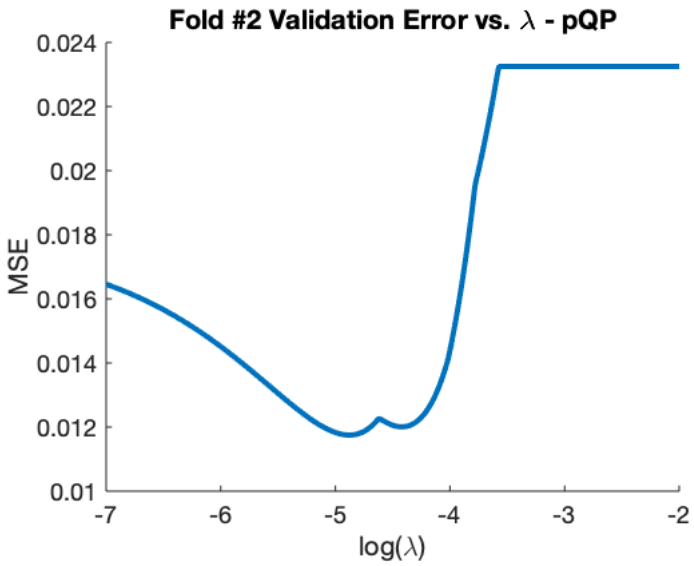
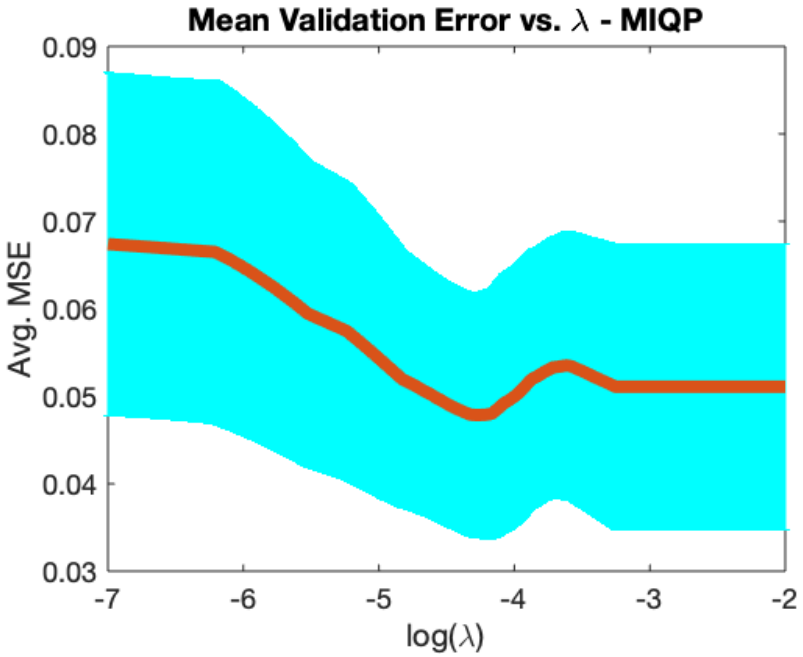


Figure 6.9: Validation error for the ammonia reactor data in fold #2 from the training pQP (Eq. 6.15) solved using POP

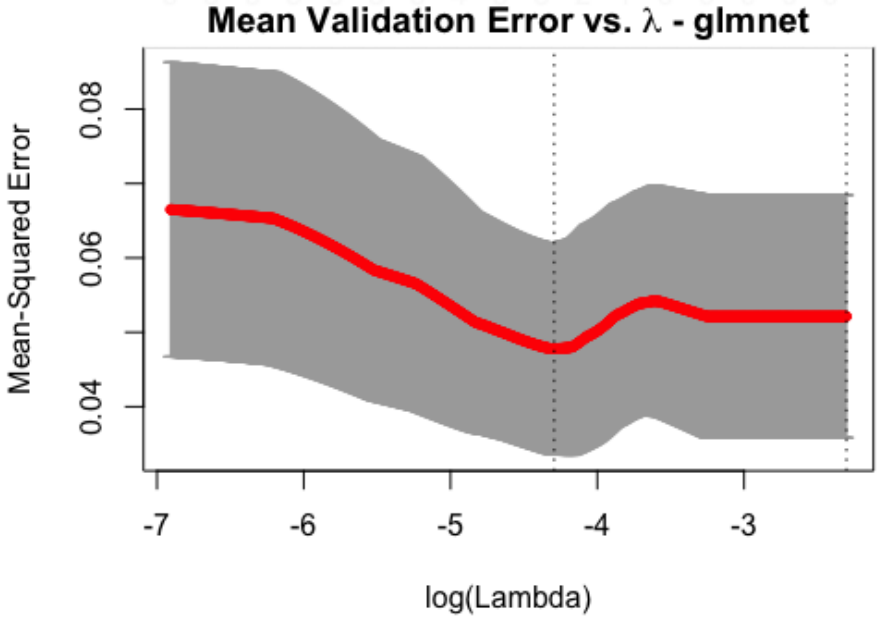
6.10 highlights that the resulting mean validation error profiles calculated from these two methods. Shaded blue and gray areas represent one standard error above and below the mean validation error.

The validation error profiles appear exactly the same, confirming that the HY-POP approach leads to the same solution as the established coordinate descent algorithm in glmnet. Table 6.1 depicts some computational results. The HY-POP approach calculates an optimal λ of 0.140, while cv.glmnet calculates an optimal λ of 0.1365. This slight difference in value is attributed to numerical sensitivity in the algorithmic computations and how the λ space is the discretized. Not controlling for programming environment, the coordinate descent algorithm appears faster that the HY-POP approach. This is expected because the former is a tailored algorithm with warm starts for solving the pQP of LASSO regression, while the former uses a generic algorithm applicable to any pQP with no specializations for LASSO regression. Moreover, the MATLAB code has not been optimized for speed.

However, depending on the size of X, the number of folds, and the granularity of how λ is discretized, these computational conclusions may change. There may be corner cases where a much finer discretization of λ is needed for a good approximation. These are subjects for further study.



(a)



(b)

Figure 6.10: Mean validation error for the hyperparameter optimization of LASSO regression on the ammonia reactor data through 5-fold cross-validation. (a) Error profile solved using the MIQP (Eq. 6.16). (b) Error profile solved using cv.glmnet

.

Table 6.1: CPU times for 5-fold cross-validation on ammonia reactor data

Method	Optimal MSE	Optimal λ	CPU Time[1](s)
pQP + MIQP	0.0479	0.0140	1.44 ± 0.06
cv.glmnet[2]	0.0478	0.0137	0.19 ± 0.01

[1] Averaged over 10 runs
[2] Grid of 10^3 points

Albeit for a small example, it is promising that the generic POP solver is only a order of magnitude slower. For the HY-POP approach here, about 84% of the CPU time is spent determining the critical regions for the folds. Time spent constructing the model is included in the CPU time.

6.6 LP L_1-Norm Support Vector Machine (SVM)

SVM [176] is a common classification technique that identifies a maximal margin hyperplane separating different classes of labeled responses and utilizes it as a decision boundary. It has been applied for fault detection and process monitoring in chemical engineering [177, 178]. The standard general form is the C-parameterized SVM, where a hyperparameter *C* penalizes the slack variables ε_i controlling how much margin violation to tolerate for each misclassified observation i.

$$\begin{aligned} \min_{w,b,\varepsilon} \quad & \frac{1}{m}\|w\|_q^p + C\sum_i \varepsilon_i^d \\ & y_i(\phi(x_i)w + b) \geq 1 - \varepsilon_i \qquad \forall i \in I \\ & \varepsilon_i \geq 0 \qquad \forall i \in I \end{aligned} \tag{6.18}$$

I is the set of all training observations, w is a $j \times 1$ vector of model weights, b is a constant bias, and m is a constant value. For the ith observation, x_i is a $1 \times j$ vector of predictors, y_i is the given class label, and ε_i is the slack variable. Binary classification is assumed, where the y_i is either $+1$ or -1.

The power d to which ε_i is raised in the objective (loss) function dictates whether Eq. 6.18 is the L_1-norm ($d = 1$) or the L_2-norm ($d = 2$) SVM. When $p = 1$ & $q = 1$, Eq. 6.18 is the LP SVM, and when $p = 2$ & $q = 2$, it is the QP SVM. Typically, $m = 1$ for the LP and $m = 2$ for the

QP. Since LASSO regression in previous section is an example of the HY-POP approach applied to a QP, in this section, LP L_1-norm SVM [179] is used as an example of the HY-POP approach implemented on a LP. The LP L_1-norm SVM [179] is shown below.

$$\begin{aligned} \min_{w,b,\varepsilon} \quad & \sum_j |w_j| + C \sum_i \varepsilon_i \\ & y_i(\phi(x_i)w + b) \geq 1 - \varepsilon_i \qquad \forall i \in I \\ & \varepsilon_i \geq 0 \qquad \forall i \in I \end{aligned} \tag{6.19}$$

Eq. 6.19 is a parametric programming problem in the form of Eq. 6.7, but does not have a squared error loss as LASSO regression (Eq. 6.14) did. Instead, the objective is to minimize the sum of margin violations, while selecting which predictors are more important to construct the hyperplane. The C hyperparameter is associated with the slack variables ε_i accounting for this margin violation sum, instead of the model weights w for LASSO regression. An absolute value reformulation of $\sum_j = |w_j|$ is performed for Eq. 6.19, where $|w_j| = p_j + q_j$ and $w_j = p_j - q_j$. The parametric linear programming (pLP) for LP L_1-norm SVM is the following.

$$\begin{aligned} \min_{p,q,b,\varepsilon} \quad & \sum_j p_j + q_j + C \sum_i \varepsilon_i \\ & y_i(\phi(x_i)p - \phi(x_i)q + b) \geq 1 - \varepsilon_i \qquad \forall i \in I \\ & \varepsilon_i \geq 0 \qquad \forall i \in I \\ & p_j \geq 0 \qquad \forall j \in J \\ & q_j \geq 0 \qquad \forall j \in J \end{aligned} \tag{6.20}$$

Eq. 6.20 is the inner training optimization to the bilevel problem to optimize C. For each kth fold, the critical regions describing the optimal solution profile to Eq. 6.20 are in the same form as those in Eqs. 6.10, 6.11, & 6.12, taking into account the reformulation of $|w_j|$. The only difference here are the two sets of Big-M constraints (Eq. 6.10) to account for w and b. The critical regions

for Eq. 6.20 for all folds are defined by the following constraints, where w is evaluated from $p-q$.

$$
\begin{aligned}
w_k &\leq A_{kr}C + b_{kr} + M(1 - y_{kr}^{CR}) \qquad \forall k \in K, \forall r \in R^k \\
w_k &\geq A_{kr}C + b_{kr} + M(y_{kr}^{CR} - 1) \qquad \forall k \in K, \forall r \in R^k \\
b_k &\leq G_{kr}C + h_{kr} + M(1 - y_{kr}^{CR}) \qquad \forall k \in K, \forall r \in R^k \\
b_k &\geq G_{kr}C + h_{kr} + M(y_{kr}^{CR} - 1) \qquad \forall k \in K, \forall r \in R^k \\
\sum_{r=1}^{|R^k|} LB_{kr}^{CR} y_{kr}^{CR} &\leq C \leq \sum_{r=1}^{|R^k|} UB_{kr}^{CR} y_{kr}^{CR} \qquad \forall k \in K \\
\sum_{r=1}^{|R^k|} y_{kr}^{CR} &= 1 \qquad \forall k \in K
\end{aligned}
\tag{6.21}
$$

For critical region r in the kth fold, the affine expressions for the model weights are $w_k = A_{kr} + b_{kr}$ and the constant bias is $b_k = G_{kr} + h_{kr}$. These define the decision boundary $w_k \phi(x_{ik}) + b_k$ that classifies each observation into either $+1$ or -1 class, depending on the sign of its evaluated value (positive is $+1$ and negative is -1).

The outer level objective in the bilevel optimization is to minimize the misclassification rate across all folds. For the kth fold with N_k^{tst} sample points in the testing set, the misclassification rate is defined below.

$$
\frac{1}{N_k^{tst}} \sum_{i=1}^{N_k^{tst}} F(y_{ik}, \hat{y}_{ik}) \tag{6.22}
$$

Here, $\hat{y}_{ik}$ is the predicted class label for the ith observation in the testing set for the kth fold, and y_{ik} is the given class label. F is an indicator function that equals 1 if $y_{ik} \neq \hat{y}_{ik}$ and 0 if $y_{ik} = \hat{y}_{ik}$. If $F(y_{ik}, \hat{y}_{ik}) = 0$, then the ith observation is correctly classified. Otherwise, it is misclassified.

To capture this discrete decision and identify a function form for F, $\hat{y}_{ik}$ is redefined as a binary variable equal to 1 for a prediction belonging to the $+1$ class and equal to 0 for a prediction belonging to the -1 class. During testing validation in the outer level, the $+1$ and -1 class labels y_{ik} are also redefined as 1 and 0, respectively, to accommodate the binary variable $\hat{y}_{ik}$. However, during training (Eq. 6.20) and to compute the critical regions (Eq. 6.21), the class labels y_{ik} remain

$+1$ and -1. In this way, the misclassification rate is rewritten as a quadratic loss function.

$$\frac{1}{N_k^{tst}} \sum_{i=1}^{N_k^{tst}} (y_{ik} - \hat{y}_{ik})^2 \tag{6.23}$$

$(y_{ik} - \hat{y}_{ik})^2$ equals 1 for a misclassified observation, when $y_{ik} = 1$ & $\hat{y}_{ik} = 0$ or $y_{ik} = 0$ & $\hat{y}_{ik} = 1$. $(y_{ik} - \hat{y}_{ik})^2$ equals 0 for a correctly classified observation, when $y_{ik} = 1$ & $\hat{y}_{ik} = 1$ or $y_{ik} = 0$ & $\hat{y}_{ik} = 0$. The reason for a redefinition of the class labels during testing validation is clear from this form of Eq. 6.23.

To connect the critical regions in the inner level describing the trained decision boundaries $(w_k \phi(x_{ik}) + b_k)$ to the misclassification rate in the outer level, two additional Big-M constraints are needed to handle the redefinition of class labels made between the training and testing validation.

$$\begin{aligned} w_k \phi(x_{ik}) + b_k &\geq M(\hat{y}_{ik} - 1) \qquad \forall k \in K, \forall i \in I^k \\ w_k \phi(x_{ik}) + b_k &\leq M \hat{y}_{ik} \qquad \forall k \in K, \forall i \in I^k \end{aligned} \tag{6.24}$$

I^k is the set of all observations in the testing set in the kth fold. x_{ik} are the predictors for the ith observation in the kth fold. When the decision boundary is positive ($+1$ class), $\hat{y}_{ik} = 1$. When the decision boundary is negative (-1 class), $\hat{y}_{ik} = 0$. Therefore, Eq. 6.24 captures the binary decision of predicting a class label and connects the learned w_k & b_k from the critical regions (Eq. 6.21) to the misclassification rate in Eq. 6.23.

The final HY-POP formulation (MIQP) of LP L_1-norm SVM hyperparameter optimization comprises of Eqs. 6.21, 6.23, and 6.24 and is shown below. Eq. 6.25 represents a new construction for the hyperparameter optimization of LP L_1-norm SVM through k-fold cross-validation. This HY-POP formulation is next demonstrated on a breast cancer data example.

$$
\begin{aligned}
\min_{C,w_k,b_k,\hat{y}_{ik},y_{kr}^{CR}} \quad & \frac{1}{|K|}\sum_{k=1}^{|K|}\frac{1}{N_k^{tst}}\sum_{i=1}^{N_k^{tst}}(y_{ik}-\hat{y}_{ik})^2 \\
\text{s.t.} \quad & w_k\phi(x_{ik})+b_k \geq M(\hat{y}_{ik}-1) \qquad \forall k\in K, \forall i \in I^k \\
& w_k\phi(x_{ik})+b_k \leq M\hat{y}_{ik} \qquad \forall k\in K, \forall i \in I^k \\
& w_k \leq A_{kr}C+b_{kr}+M(1-y_{kr}^{CR}) \qquad \forall k\in K, \forall r \in R^k \\
& w_k \geq A_{kr}C+b_{kr}+M(y_{kr}^{CR}-1) \qquad \forall k\in K, \forall r \in R^k \\
& b_k \leq G_{kr}C+h_{kr}+M(1-y_{kr}^{CR}) \qquad \forall k\in K, \forall r \in R^k \\
& b_k \geq G_{kr}C+h_{kr}+M(y_{kr}^{CR}-1) \qquad \forall k\in K, \forall r \in R^k \\
& \sum_{r=1}^{|R^k|} LB_{kr}^{CR}y_{kr}^{CR} \leq C \leq \sum_{r=1}^{|R^k|} UB_{kr}^{CR}y_{kr}^{CR} \qquad \forall k\in K \\
& \sum_{r=1}^{|R^k|} y_{kr}^{CR}=1 \qquad \forall k\in K
\end{aligned}
\tag{6.25}
$$

6.6.1 Breast Cancer Data Example

A dataset of 116 samples with 9 predictors for breast cancer [180] is downloaded from the UCI Machine Learning Repository. The data values are normalized and randomly divided into 5 folds. The fold identification of the data points is provided in the Appendix B.3. Healthy patients are labeled $+1$, and cancer patients are labeled -1. The predictors are age, BMI, and levels of glucose, insulin, HOMA, leptin, adiponectin, resistin, & MCP-1. It is assumed that the predictors are linear, $\phi(X)=X$. Therefore, the proposed LP L_1-norm SVM that is trained is the following.

$$
\begin{aligned}
y_{cancer} = {} & w_1x_{age}+w_2x_{BMI}+w_3x_{glu}+w_4x_{insu}+w_5x_{HOMA} \\
& +w_6x_{lep}+w_7x_{adi}+w_8x_{res}+w_9x_{MCP1}+b
\end{aligned}
\tag{6.26}
$$

Each training LP L_1-norm SVM (Eq. 6.20) for each fold is formulated in MATLAB 2019b. The critical regions are solved using the POP toolbox [173] with the CPLEX LP solver and the geometrical algorithm. An example result of the critical regions for fold #3 is shown in Fig. 6.11.

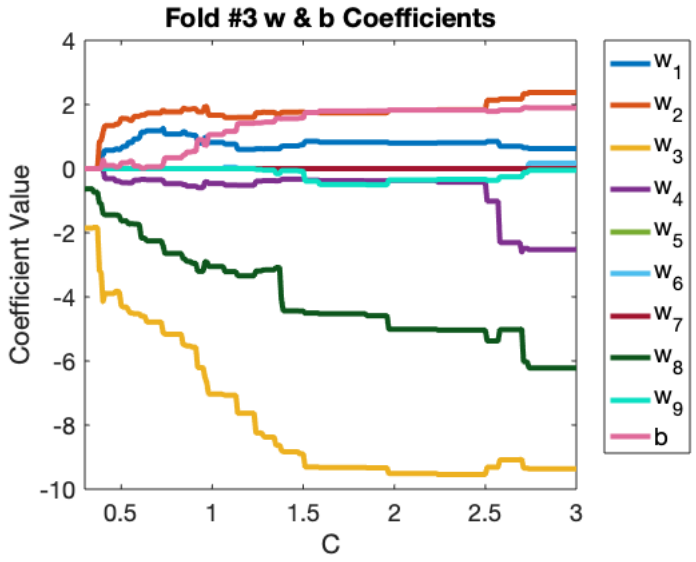
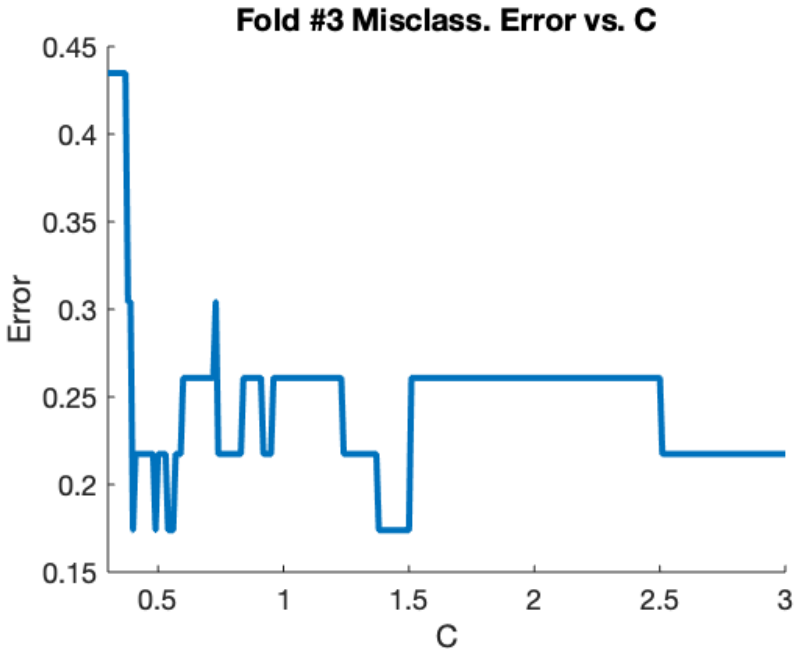


Figure 6.11: SVM regularization path for the breast cancer data in fold #3 from the training pLP (Eq. 6.20) solved using POP

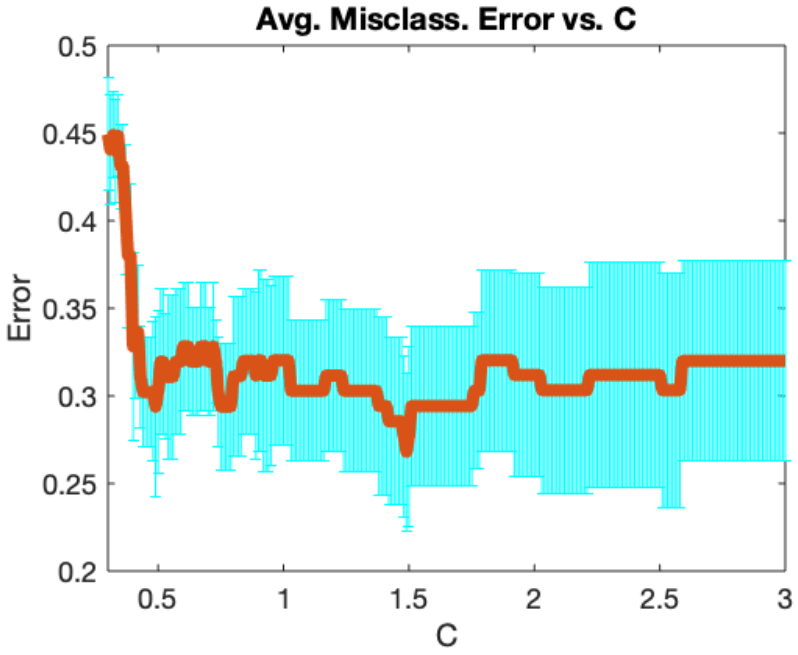
Like the LASSO regression, a piecewise relationship between the model weights w & b with the hyperparameter C is also observed here. This piecewise relationship is referred to as the SVM regularization path [136]. We expected this behavior from the affine expressions, $w(C)$ & $b(C)$, determined from the parametric programming solution to Eq. 6.20. Earlier extensions of LAR and coordinate descent to investigating SVM also discovered the SVM regularization path [181, 182], noticing similarities between the algorithms used to solve SVM and LASSO regression. In general, this is expected because all these solution methods fall under the general theory of parametric programming (Eq. 6.7).

Unlike the LASSO regression, there are many more individual line segments (critical regions) that are stitched together to represent the optimal solution profile in Fig. 6.11. This creates a very non-smooth behavior in $w(C)$ & $b(C)$ and the resulting misclassification error profile (Fig. 6.12). The non-smoothness is expected because the objective (loss) function in Eq. 6.25 is quadratic with respect to a binary variable $\hat{y}_{ik}$ and not a continuous variable. The shaded blue area in Fig. 6.12b represent one standard error above and below the average misclassification rate.

Because of this non-smooth behavior in the misclassification error profile, characteristic of



(a)



(b)

Figure 6.12: (a) Misclassification error in fold #3 (b) Misclassification error averaged across all the folds

.

Table 6.2: CPU times for 5-fold cross-validation on breast cancer data

Method	Optimal Error	Optimal C	CPU Time[1](s)
pLP + MIQP	0.2681	1.4869	50.49 ± 0.61

[1] Averaged over 10 runs

classification problems in machine learning, it is inherently more difficult to find an accurate approximation to an optimal C from a discretized grid search. The optimal C is more sensitive to the granularity of the discretization. This is one advantage of having the optimal solution in explicit form through a HY-POP approach. For the breast cancer example, the optimal C value is 1.4869 with an average misclassification error of 0.2681. Table 6.2 shows the computational results. Even though Eq. 6.20 is a pLP, a longer CPU time for this breast cancer example is observed than for the LASSO regression because there are many more critical regions to compute in this instance. Eq. 6.25 has more constraints in its problem than Eq. 6.16, which causes more active set explorations for the parametric programming algorithm to solve the former. About 87% of the CPU time is spent solving for the critical regions in the folds.

6.7 Conclusion

The novelty in this work is constructing hyperparameter optimization through k-fold cross-validation as a bilevel optimization problem that is solvable as a single level optimization through parametric programming. We refer to this as the bilevel & parametric optimization approach to hyperparameter optimization (HY-POP).

This parametric programming perspective ties together previous studies that first recognized the regularization paths of LASSO regression and SVM as piecewise linear functions and extends these results to optimize hyperparameters in k-fold cross-validation. The HY-POP approach to hyperparameter optimization is demonstrated on ammonia reactor data, a QP example, and breast cancer data, a LP example.

Advantages of recognizing hyperparameter optimization as a parametric programming problem are threefold. First, the HY-POP approach to hyperparameter optimization is applicable to

any general machine learning algorithm that is a LP/QP model. In fact, mixed-integer linear or quadratic (MILP/MIQP) models can also be used because parametric programming theory exists for these problem types. Second, when there multiple hyperparameters in a machine learning model (a common occurrence), there is theory to solve these problems exactly through multi-parametric programming. Third, and most importantly, no discretization of the hyperparameter space is required for HY-POP.

Finally, the aims of this work are to lay the introductory foundation and present an unified view to hyperparameter optimization of machine learning models from a parametric programming perspective. The multi-parametric programming, mixed-integer, and computational aspects of a HY-POP approach to hyperparameter optimization are subjects of further investigation.

7. CONCLUSION AND FUTURE WORK

7.1 Conclusion

In this dissertation, imminent challenges society faces in the energy transition and digitalization are discussed. While there are many outstanding issues, the focus here is to present a few ideas that address some of the challenges. For the energy transition, the optimal design and operation of select natural gas & renewable energy systems for producing fuels, chemicals, and power are investigated. For a digital energy sector, an algorithm for incorporating time series data into optimizing a renewable power system with storage and a novel method for the optimal selection of machine learning models are developed.

In all these studies, mathematical optimization and a systems approach from PSE show their utility in elucidating insights into problems that have a plethora of different options. This ultimately is invaluable for decision-makers to make well-informed choices on what future energy systems to pursue and which models to monetize data with. Some insights into future energy systems are the following. In Section 2, a MINLP process synthesis highlighted the potential benefits of chemical looping in reducing the syngas production costs over traditional autothermal or steam reforming in GTL plants. Small-scale chemical looping processes are equivalent in performance to conventional reforming ones that are 5-10 times larger. In Section 3, a MINLP process synthesis showcased the synergies in integrating ammonia and methanol production from biomass together to decrease CO_2 emissions and costs. Ammonia BEP declines faster than the increase in methanol BEP, suggesting a potential to hedge production based on varying economic situations. The MILP supply chain model in Section 4 described the long-term energy storage values of hydrogen, ammonia, & methanol and how storage times affect the trade-off between choices. Hydrogen is more suitable at shorter times, and ammonia becomes attractive as the storage time increases.

Some advances on utilizing data for energy system analysis and model-building are the following. A clustering decomposition algorithm for a design and scheduling MILP model in Section 5

concluded that renewable power systems with storage are cost-competitive with power production from fossil fuels, and LCOE values remain reasonable if demand satisfaction is flexible. Battery storage is necessary to match renewable supply with consumer demand, but may not be sufficient. Backup options like energy carriers (hydrogen-rich chemicals) may be required to provide the last portions of demand for a completely renewable power system to avoid costly oversizing capacity issues. Section 6 introduced the exact solution for optimal model selection of machine learning LP/QPs through bilevel optimization and parametric programming. This extends available explicit solutions beyond those of LASSO- & SVM-type models and avoids suboptimal data fitting. As more data is utilized to influence operational decisions in energy systems, optimal machine learning models become even more important.

Overall, mathematical optimization is a common theme throughout this dissertation. Options for which energy systems to design and models to build from data will only continue to multiply over time in an increasingly digital and globalized world. Mathematical optimization and other fundamental PSE strategies such as process synthesis, scheduling, supply chain optimization, and data-driven modeling will become even more essential tools than they are today. This dissertation has demonstrated the power of these optimization-based strategies for providing some insights into understanding the complex issues surrounding the design and operation of energy systems. As shown, they are especially attractive strategies for studying chemicals and electricity production, the two areas with rapid growth in the coming decades. Optimization is also a crucial component for advancing machine learning. Utilizing bilevel optimzation and parametric programming theory, the exact solutions of LP/QP machine learning models is possible.

7.2 Key Contributions

The key contributions of the dissertation are summarized below.

1. Chemical looping has predominantly been studied as a post-combustion CO_2 capture system. There is little previous work into the techno-economic analysis and optimization of chemical looping as a syngas production method. The MINLP process synthesis of a GTL system with

chemical looping is a first effort at understanding this.

2. Coproduction of ammonia and methanol has been hypothesized before, but there is very little work on its techno-economics and optimization. The MINLP process synthesis of a biomass-based ammonia and methanol coproduction system is an initial attempt at investigating this.

3. A MILP supply chain optimization model is developed to better understand the trade-offs and synergies between hydrogen, ammonia, and methanol for chemical energy storage. This is first effort at understanding the hydrogen, ammonia, and methanol "economies" together.

4. An improved decomposition algorithm based on agglomerative hierarchical clustering is developed for unit commitment and capacity expansion optimization problems with time series data. Evaluating the accuracy of an aggregated time horizon in approximating the true optimal solution is possible with this algorithm.

5. A novel method for the hyperparameter optimization of machine learning models is developed through bilevel optimization and parametric programming theory. This extends previous results from the statistics community in algorithms for regularization path solutions.

7.3 Future Work

There are many challenges in the energy transition and digitalization that are not coverable in a single dissertation. A non-exhaustive list of future research directions is presented below. Similar MILP/MINLP modeling approaches such as the ones utilized in this dissertation are appropriate for investigating any one of these directions.

7.3.1 Process Synthesis for a Circular Economy

Sections 2 and 3 discussed a MINLP process synthesis model for producing fuels and chemicals from natural gas and biomass. These proposed processes represent the traditional way of viewing production systems. A feedstock is first transformed into a product. After product consumption, the waste is disposed of and fresh feedstock is required to manufacture a new product all over again. Motivated by sustainability and climate change, there is a growing interest expressed by

industrial companies in shifting to a different paradigm where the waste is recycled to manufacture more product. This has been termed the "circular economy" where reusable products are favored over single-use products [183, 184].

The concept of recycling waste streams to increase product yield is not new for chemical engineering. This is a common setup seen in many equilibrium-constrained reactors. However, there has been very little work done on investigating entire energy systems where the overall exiting waste is utilized as feedstock to manufacture new product [185]. For example, the world currently creates about amount 2 billion metric tons of municipal solid waste (MSW) per year, and this is expected to increase 70% by 2050 as the world population grows [186]. Previous process synthesis works have investigated MSW to fuels, olefins, and aromatics [187, 188] processes. This concept could be extended to include other products such as plastics, textiles, cosmetics, and fertilizers/urea. Other circular economy energy systems could include CO_2 utilization and wastewater purification. Again, the possible options for a circular economy are numerous, and PSE tools are suitable to tackle the challenge.

7.3.2 Strategic Planning of Renewable Energy Systems

In Sections 4 and 5, the optimization models for these renewable energy systems were all formulated as here-and-now decisions, assuming that these systems were already built and operating today. This is a myopic assumption because capital costs and process efficiencies are fixed to current day values, whereas there may be a learning curve over time associated with them. For example, as economies of scale and manufacturing know-how are developed, there will be a deceleration in capital costs and an increase in process efficiencies. Moreover, the growing penetration of renewable energies into the world energy mix and installation of renewable technologies are multi-year phenomena and not something that will happen over night.

Therefore, a more realistic portrayal would be a strategic planning of renewable energy systems over a long time horizon, where investment and operational decisions are both time-dependent and made across different points in the horizon. Elia et al. [189] describes an optimization model for the strategic planning of GTL systems and solves it using a rolling horizon strategy. A similar

approach can be adopted for analyzing the roll-out of renewable energy systems between different geographical locations with varying resource densities of solar and wind. Optimal renewable energy systems calculated from either a process synthesis (Sections 2 and 3) or design & scheduling model (Section 5) can be used as candidate facilities to be built in the strategic planning optimization problem.

7.3.3 Data Variability in the Optimization of Renewable Energy Systems

The time series data of solar DNI, wind speed, demand load, and electricity prices used for the optimization of a renewable power system with storage in Section 5 is taken from a single year. An assumption was made that this data adequately captures enough of the input variability for designing and scheduling the power system. However, there are yearly fluctuations that are not considered here, and a renewable energy system designed off of a single year's data may become suboptimal or infeasible when operating in another year. Moreover, the lifetime of any renewable energy system spans multiple decades. The system design needs to take into account the potential weather changes in the future.

Therefore, additional modeling techniques are required to quantitatively assess the impact of data variability in the optimization of renewable energy systems. Uncertainties in the realization of solar DNI, wind speed, demand load, and electricity prices can be addressed with either robust optimization [190–192] or stochastic programming [193, 194]. The integration of a clustering algorithm like the one presented in Section 5 with robust optimization or stochastic programming has not yet been investigated. Results from these modeling approaches will quantify the sensitivity of optimal energy system designs to varying weather and economic situations.

7.3.4 HY-POP for Selecting Optimal MILP/MIQP Machine Learning Models

LASSO regression and L_1-norm SVM classification in Section 6 are examples of QP and LP machine learning models, respectively. There exists multi-parametric programming theory to exactly solve MILP/MIQP optimization problems as well. Moreover, several recent works have shown that nonparametric machine learning algorithms such as neural networks with rectified lin-

ear unit (ReLU) activation functions [195] and classification & regression (CART) trees [196, 197] are solvable as mixed-integer optimization problems. However, the hyperparameter optimization of neural networks and CART trees remains an outstanding challenge because they involve tuning multiple hyperparameters. Likewise, feature selection [198] to improve machine learning models such as SVM can also be casted as a mixed-integer optimization problem.

Therefore, a similar bilevel optimization and parametric programming (HY-POP) method as the one derived in Section 6 is applicable to help tackle the hyperparameter optimization of neural networks and CART trees or the feature selection problem in other machine learning models. Results for the former would be extremely useful because neural networks and CART trees have been empirically demonstrated to be the most reliable and accurate off-the-shelf machine learning models for predicting from data, if their hyperparameters are properly tuned. Moreover, sparsity in machine learning models from feature selection is highly desirable and multi-parametric programming can lend itself to help develop a greater understanding of these sparse models.

REFERENCES

[1] United Nations DESA/Population Divison, "World Population Prospects," 2017. https://population.un.org/wpp/Download/Standard/Population/, Last accessed on 2019-02-07.

[2] H. Kharas, "The unprecedented expansion of the global middle class: An update," 2017. https://think-asia.org/handle/11540/7251, Last accessed on 2019-02-07.

[3] ExxonMobil, "2018 Outlook for Energy: A View to 2040." https://corporate.exxonmobil.com/-/media/Global/Files/outlook-for-energy/2018-Outlook-for-Energy.pdf, Last accessed on 2019-02-07.

[4] Royal Dutch Shell, "Energy Transition Report," 2018. https://www.shell.com/energy-and-innovation/the-energy-future/shell-energy-transition-report.html, Last accessed on 2019-02-07.

[5] International Energy Agency, "World Energy Outlook," 2018. https://www.iea.org/weo2018/, Last accessed on 2019-02-07.

[6] Organization of the Petroleum Exporting Countries, "World Oil Outlook," 2018. https://woo.opec.org/, Last accessed on 2019-02-10.

[7] BP plc, "Statistical Review of World Energy," 2018. https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html, Last accessed on 2019-02-09.

[8] Global Carbon Project, "Carbon Budget," 2018. https://www.globalcarbonproject.org/carbonbudget/index.html, Last accessed on 2019-02-09.

[9] Our World in Data, "CO_2 and other Greenhouse Gas Emissions," 2018. https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions,

Last accessed on 2019-02-09.

[10] V. Smil, *Energy Transitions: History, Requirements, Prospects*. ABC-CLIO, 2010.

[11] V. Smil, *Energy Transitions: Global and National Perspectives*. ABC-CLIO, 2016.

[12] K. T. Møller, T. R. Jensen, E. Akiba, and H.-w. Li, "Hydrogen-a sustainable energy carrier," *Progress in Natural Science: Materials International*, vol. 27, no. 1, pp. 34–40, 2017.

[13] Nel ASA, "Atmospheric Alkaline Electrolyser," 2018. https://nelhydrogen.com/product/atmospheric-alkaline-electrolyser-a-series/, Last accessed on 2019-02-11.

[14] Proton OnSite, "M Series," 2018. https://www.protononsite.com/products-proton-site/m-series, Last accessed on 2019-02-11.

[15] N. R. Council *et al.*, *Hidden costs of energy: unpriced consequences of energy production and use*. National Academies Press, 2010.

[16] S. C. Davis, S. W. Diegel, and R. G. Boundy, *Transportation Energy Data Book - Edition 28*. "Oak Ridge National Laboratory", 2009. https://info.ornl.gov/sites/publications/files/Pub20096.pdf, Last accessed on 2019-02-11.

[17] M. Wang, *Estimation of Energy Efficiencies of U.S. Petroleum Refineries*. "ArgonneNational Laboratory", 2008. https://greet.es.anl.gov/files/hl9mw9i7, Last accessed on 2019-02-11.

[18] U.S. Department of Energy, "Hydrogen Fuel Cells Factsheet," 2006. https://www.californiahydrogen.org/wp-content/uploads/files/doe_fuelcell_factsheet.pdf, Last accessed on 2019-02-11.

[19] International Energy Agency, "Technology Roadmap: Energy Storage," 2014. https://www.iea.org/publications/freepublications/publication/TechnologyRoadmapEnergystorage.pdf, Last accessed on 2019-02-11.

[20] International Renewable Energy Agency, "Electricity Storage and Renewables: Costs and Markets to 2030," 2017. https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2017/Oct/IRENA_Electricity_Storage_Costs_2017.pdf, Last accessed on 2019-02-11.

[21] B. Zakeri and S. Syri, "Electrical energy storage systems: A comparative life cycle cost analysis," *Renewable and Sustainable Energy Reviews*, vol. 42, pp. 569–596, 2015.

[22] U.S. Energy Information Administration, "U.S. Battery Storage Market Trends," 2018. https://www.eia.gov/analysis/studies/electricity/batterystorage/pdf/battery_storage.pdf, Last accessed on 2019-02-11.

[23] Enerdata, "Global Energy Statistical Yearbook," 2018. https://yearbook.enerdata.net/renewables/renewable-in-electricity-production-share.html, Last accessed on 2019-02-10.

[24] U.S. Energy Information Administration, "New electric generating capacity in 2019 will come from renewables and natural gas," 2018. https://www.eia.gov/todayinenergy/detail.php?id=37952, Last accessed on 2019-02-10.

[25] International Renewable Energy Agency, "Renewable Power Generation Costs in 2017," 2018. https://www.irena.org/publications/2018/Jan/Renewable-power-generation-costs-in-2017, Last accessed on 2019-02-11.

[26] U.S. Energy Information Administration, "Levelized Cost and Levelized Avoided Cost of New Generation Resources in the Annual Energy Outlook," 2018. https://www.eia.gov/outlooks/aeo/pdf/electricity_generation.pdf, Last accessed on 2019-02-11.

[27] Renewable Fuels Association, "Markets & Statistics," 2018. https://ethanolrfa.org/resources/industry/statistics/, Last accessed on 2019-02-11.

[28] University of Michigan: Center for Sustainable Systems, "Biofuels Factsheet," 2018. http://css.umich.edu/factsheets/biofuels-factsheet/, Last accessed on 2019-02-11.

[29] Empresa de Pesquisa Energética, "Balanço Energético Nacional," 2018. http://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-303/topico-397/Relat%C3%B3rio%20S%C3%ADntese%202018-ab%202017vff.pdf, Last accessed on 2019-02-11.

[30] Iowa State University: Center for Agricultural and Rural Development, "Historical Ethanol Operating Margins," 2018. https://www.card.iastate.edu/research/biorenewables/tools/hist_eth_gm.aspx, Last accessed on 2019-02-11.

[31] United States Department of Agriculture, "The Economic Feasibility of Ethanol Production from Sugar in the United States," 2006. https://www.usda.gov/oce/reports/energy/EthanolSugarFeasibilityReport3.pdf, Last accessed on 2019-02-11.

[32] U.S. Energy Information Administration, "United States remains the world?s top producer of petroleum and natural gas hydrocarbons," 2018. https://www.eia.gov/todayinenergy/detail.php?id=36292, Last accessed on 2019-02-12.

[33] International Gas Union, "Natural Gas Is The Cleanest Fossil Fuel," 2018. https://www.igu.org/natural-gas-cleanest-fossil-fuel, Last accessed on 2019-02-12.

[34] V. Venkatasubramanian, "The promise of artificial intelligence in chemical engineering: Is it here, finally," *AIChE Journal*, vol. 65, no. 2, pp. 466–78, 2019.

[35] S. J. Qin, "Process data analytics in the era of big data," *AIChE Journal*, vol. 60, no. 9, pp. 3092–3100, 2014.

[36] L. Chiang, B. Lu, and I. Castillo, "Big data analytics in chemical engineering," *Annual Review of Chemical and Biomolecular Engineering*, vol. 8, pp. 63–85, 2017.

[37] S. J. Qin and L. H. Chiang, "Advances and opportunities in machine learning for process data analytics," *Computers & Chemical Engineering*, vol. 126, pp. 465–473, 2019.

[38] ExxonMobil, "Harnessing data to unlock new energy solutions." `https://energyfactor.exxonmobil.com/science-technology/harnessing-data-to-unlock-new-energy-solutions/`, Last accessed on 2020-02-08.

[39] J. W. Erisman, M. A. Sutton, J. Galloway, Z. Klimont, and W. Winiwarter, "How a century of ammonia synthesis changed the world," *Nature Geoscience*, vol. 1, no. 10, pp. 636–639, 2008.

[40] I. Grossmann, "Enterprise-wide optimization: A new frontier in process systems engineering," *AIChE Journal*, vol. 51, no. 7, pp. 1846–1857, 2005.

[41] I. Cameron, F. Wang, C. Immanuel, and F. Stepanek, "Process systems modelling and applications in granulation: A review," *Chemical Engineering Science*, vol. 60, no. 14, pp. 3723–3750, 2005.

[42] K.-U. Klatt and W. Marquardt, "Perspectives for process systems engineering-personal views from academia and industry," *Computers & Chemical Engineering*, vol. 33, no. 3, pp. 536–550, 2009.

[43] G. Stephanopoulos and G. V. Reklaitis, "Process systems engineering: From solvay to modern bio-and nanotechnology.: A history of development, successes and prospects for the future," *Chemical Engineering Science*, vol. 66, no. 19, pp. 4272–4306, 2011.

[44] I. E. Grossmann and I. Harjunkoski, "Process systems engineering: Academic and industrial perspectives," *Computers & Chemical Engineering*, vol. 126, pp. 474–484, 2019.

[45] U.S. Energy Information Administration, "Preliminary Monthly Electric Generator Inventory," 2018. https://www.eia.gov/electricity/data/eia860M/2, Last accessed on 2019-02-12.

[46] Chemical Engineering Magazine, "Comparing Petrochemical Plant Aging," 2010. https://www.chemengonline.com/comparing-petrochemical-plant-aging/, Last accessed on 2019-02-12.

[47] Canary, "Are the Stresses of Middle Age too Much for Today's Refineries," 2014. http://canaryusa.com/stresses-of-middle-age-for-todays-refineries/, Last accessed on 2019-02-12.

[48] N. A. Owen, O. R. Inderwildi, and D. A. King, "The status of conventional world oil reserves - Hype or cause for concern?," *Energy Policy*, vol. 38, no. 8, pp. 4743–4749, 2010.

[49] F. Verrastro and S. Ladislaw, "Providing energy security in an interdependent world," *Washington Quarterly*, vol. 30, no. 4, pp. 95–104, 2007.

[50] Energy Information Administration, "Henry Hub Natural Gas Spot Price," 2019. http://www.eia.gov/dnav/ng/hist/rngwhhdm.htm/, Last accessed on 2019-02-14.

[51] U.S. Energy Information Administration, "How much natural gas does the United States have, and how long will it last?," 2020. https://www.eia.gov/tools/faqs/faq.php?id=58&t=8, Last accessed on 2020-03-13.

[52] E. F. Sousa-Aguiar, F. B. Noronha, and A. Faro Jr, "The main catalytic challenges in gtl (gas-to-liquids) processes," *Catalysis Science & Technology*, vol. 1, no. 5, pp. 698–713, 2011.

[53] K. Aasberg-Petersen, I. Dybkjær, C. V. Ovesen, N. C. Schjødt, J. Sehested, and S. G. Thomsen, "Natural gas to synthesis gas–catalysts and catalytic processes," *Journal of Natural Gas Science and Engineering*, vol. 3, no. 2, pp. 423–459, 2011.

[54] R. C. Baliban, J. A. Elia, and C. A. Floudas, "Novel natural gas to liquids processes: Process synthesis and global optimization strategies," *AIChE Journal*, vol. 59, no. 2, pp. 505–531, 2013.

[55] T. H. Fleisch, R. A. Sills, and M. D. Briscoe, "A review of global gtl developments," *Journal of Natural Gas Chemistry*, vol. 11, pp. 1–14, 2002.

[56] D. A. Wood, C. Nwaoha, and B. F. Towler, "Gas-to-liquids (gtl): A review of an industry offering several routes for monetizing natural gas," *Journal of Natural Gas Science and Engineering*, vol. 9, pp. 196–208, 2012.

[57] D. J. Wilhelm, D. R. Simbeck, A. D. Karp, and R. L. Dickenson, "Syngas production for gas-to-liquids applications: technologies, issues and outlook," *Fuel Processing Technology*, vol. 71, no. 1, pp. 139–148, 2001.

[58] Royal Dutch Shell, "Gas-to-liquids," 2019. `http://www.shell.com/energy-and-innovation/natural-gas/gas-to-liquids.htm`, Last accessed on 2019-02-14.

[59] Sasol Limited, "Gas-to-liquids overview," 2019. `http://www.sasol.com/innovation/gas-liquids/overview`, Last accessed on 2019-02-14.

[60] K. Mazloomi and C. Gomes, "Hydrogen as an energy carrier: Prospects and challenges," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 5, pp. 3024–3033, 2012.

[61] G. A. Olah, "Beyond oil and gas: the methanol economy," *Angewandte Chemie International Edition*, vol. 44, no. 18, pp. 2636–2639, 2005.

[62] A. Klerke, C. H. Christensen, J. K. Nørskov, and T. Vegge, "Ammonia for hydrogen storage: challenges and opportunities," *Journal of Materials Chemistry*, vol. 18, no. 20, pp. 2304–2310, 2008.

[63] R. Golden and B. Paulos, "Curtailment of renewable energy in california and beyond," *The Electricity Journal*, vol. 28, no. 6, pp. 36–50, 2015.

[64] P. Denholm, M. O'Connell, G. Brinkman, and J. Jorgenson, "Overgeneration from solar energy in california. a field guide to the duck chart," tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States), 2015.

[65] O. Tietjen, M. Pahle, and S. Fuss, "Investment risks in power generation: A comparison of fossil fuel and renewable energy dominated markets," *Energy Economics*, vol. 58, pp. 174–185, 2016.

[66] J. Samimi, E. A. Soleimani, and M. Zabihi, "Optimal sizing of photovoltaic systems in varied climates," *Solar Energy*, vol. 60, no. 2, pp. 97–107, 1997.

[67] T. Ma, H. Yang, L. Lu, and J. Peng, "Optimal design of an autonomous solar–wind-pumped storage power supply system," *Applied Energy*, vol. 160, pp. 728–736, 2015.

[68] L. Kotzur, P. Markewitz, M. Robinius, and D. Stolten, "Impact of different time series aggregation methods on optimal energy system design," *Renewable Energy*, vol. 117, pp. 474–487, 2018.

[69] J. H. Merrick, "On representation of temporal variability in electricity capacity planning models," *Energy Economics*, vol. 59, pp. 261–274, 2016.

[70] W. W. Tso, A. M. Niziolek, O. Onel, C. D. Demirhan, C. A. Floudas, and E. N. Pistikopoulos, "Enhancing natural gas-to-liquids (gtl) processes through chemical looping for syngas production: Process synthesis and global optimization," *Computers & Chemical Engineering*, vol. 113, pp. 222–239, 2018.

[71] A. Lyngfelt, B. Leckner, and T. Mattisson, "A fluidized-bed combustion process with inherent CO_2 separation; application of chemical-looping combustion," *Chemical Engineering Science*, vol. 56, no. 10, pp. 3101–3113, 2001.

[72] M. M. Hossain and H. I. de Lasa, "Chemical-looping combustion (clc) for inherent CO_2 separations – a review," *Chemical Engineering Science*, vol. 63, no. 18, pp. 4433–4451, 2008.

[73] J. Adanez, A. Abad, F. Garcia-Labiano, P. Gayan, and L. F. de Diego, "Progress in chemical-looping combustion and reforming technologies," *Progress in Energy and Combustion Science*, vol. 38, no. 2, pp. 215–282, 2012.

[74] L.-S. Fan, *Chemical Looping Systems for Fossil Energy Conversions*. John Wiley & Sons, 2011.

[75] T. Pröll, J. Bolhàr-Nordenkampf, P. Kolbitsch, and H. Hofbauer, "Syngas and a separate nitrogen/argon stream via chemical looping reforming – a 140kw pilot plant study," *Fuel*, vol. 89, no. 6, pp. 1249–1256, 2010.

[76] L. F. de Diego, M. Ortiz, J. Adánez, F. García-Labiano, A. Abad, and P. Gayán, "Synthesis gas generation by chemical-looping reforming in a batch fluidized bed reactor using ni-based oxygen carriers," *Chemical Engineering Journal*, vol. 144, no. 2, pp. 289–298, 2008.

[77] M. Rydén, A. Lyngfelt, and T. Mattisson, "Synthesis gas generation by chemical-looping reforming in a continuously operating laboratory reactor," *Fuel*, vol. 85, no. 12, pp. 1631–1641, 2006.

[78] M. Rydén, A. Lyngfelt, and T. Mattisson, "Chemical-looping combustion and chemical-looping reforming in a circulating fluidized-bed reactor using ni-based oxygen carriers," *Energy & Fuels*, vol. 22, no. 4, pp. 2585–2597, 2008.

[79] S. Bhavsar and G. Veser, "Chemical looping beyond combustion: production of synthesis gas via chemical looping partial oxidation of methane," *RSC Advances*, vol. 4, no. 88, pp. 47254–47267, 2014.

[80] S. C. Bayham, A. Tong, M. Kathe, and L.-S. Fan, "Chemical looping technology for energy and chemical production," *Wiley Interdisciplinary Reviews: Energy and Environment*, vol. 5, no. 2, pp. 216–241, 2016.

[81] M. V. Kathe, A. Empfield, J. Na, E. Blair, and L.-S. Fan, "Hydrogen production from natural gas using an iron-based chemical looping technology: thermodynamic simulations and process system analysis," *Applied Energy*, vol. 165, pp. 183–201, 2016.

[82] S. Luo, L. Zeng, D. Xu, M. Kathe, E. Chung, N. Deshpande, L. Qin, A. Majumder, T.-L. Hsieh, A. Tong, *et al.*, "Shale gas-to-syngas chemical looping process for stable shale gas conversion to high purity syngas with a H_2: Co ratio of 2:1," *Energy & Environmental Science*, vol. 7, no. 12, pp. 4104–4117, 2014.

[83] L. F. de Diego, M. Ortiz, F. García-Labiano, J. Adánez, A. Abad, and P. Gayán, "Hydrogen production by chemical-looping reforming in a circulating fluidized bed reactor using ni-based oxygen carriers," *Journal of Power Sources*, vol. 192, no. 1, pp. 27–34, 2009.

[84] K. Otsuka, E. Sunada, T. Ushiyama, and I. Yamanaka, "The production of synthesis gas by the redox of cerium oxide," *Studies in Surface Science and Catalysis*, vol. 107, pp. 531–536, 1997.

[85] K. Otsuka, Y. Wang, E. Sunada, and I. Yamanaka, "Direct partial oxidation of methane to synthesis gas by cerium oxide," *Journal of Catalysis*, vol. 175, no. 2, pp. 152–160, 1998.

[86] L. K. Rihko-Struckmann, P. Datta, M. Wenzel, K. Sundmacher, N. V. R. A. Dharanipragada, H. Poelman, V. V. Galvita, and G. B. Marin, "Hydrogen and carbon monoxide production by chemical looping over iron-aluminium oxides," *Energy Technology*, vol. 4, no. 2, pp. 304–313, 2016.

[87] M. Ortiz, L. F. de Diego, A. Abad, F. García-Labiano, P. Gayán, and J. Adánez, "Hydrogen production by auto-thermal chemical-looping reforming in a pressurized fluidized bed reactor using ni-based oxygen carriers," *International Journal of Hydrogen Energy*, vol. 35, no. 1, pp. 151–160, 2010.

[88] N. Deshpande, A. Majumder, L. Qin, and L.-S. Fan, "High-pressure redox behavior of iron-oxide-based oxygen carriers for syngas generation from methane," *Energy & Fuels*, vol. 29, no. 3, pp. 1469–1478, 2015.

[89] I. Iliuta, R. Tahoces, G. S. Patience, S. Rifflart, and F. Luck, "Chemical-looping combustion process: Kinetics and mathematical modeling," *AIChE Journal*, vol. 56, no. 4, pp. 1063–1079, 2010.

[90] R. A. Sánchez and H. A. Jakobsen, "Modeling the chemical looping reforming process operated in a circulating fluidized bed reactor consisting of two bubbling bed units: Model validation," *Industrial & Engineering Chemistry Research*, vol. 53, no. 23, pp. 9616–9630, 2014.

[91] C. A. Floudas, *Deterministic Global Optimization: Theory, Methods and Applications*, vol. 37. Springer Science & Business Media, 2013.

[92] C. A. Floudas, *Nonlinear and Mixed-Integer Optimization: Fundamentals and Applications*. Oxford University Press, 1995.

[93] Energy Information Administration, "Monthly Energy Review – March 2017," 2017. http://www.eia.gov/totalenergy/data/monthly/pdf/mer.pdf, Last accessed on 2017-04-01.

[94] Argonne National Laboratory, "GREET 1.8b, The Greenhouse Gases, Regulated Emisssions, and Energy Use in Transportation (GREET) Model," 2007. Released September 2008.

[95] Oil & Gas Journal, "New Integrated Methanol/NH3 Plant Starts Up In Western Oklahoma," 1994. https://www.ogj.com/articles/print/volume-92/issue-32/in-this-issue/petrochemicals/new-integrated-methanol-nh3-plant-starts-up-in-western-oklahoma.html, Last accessed on 2019-03-18.

[96] Haldor-Topsøe, "Co-production | Methanol," 2019. https://www.topsoe.com/processes/methanol/co-production, Last accessed on 2019-03-18.

[97] E. Filippi, "Process for ammonia and methanol co-production," Dec. 25 2001. US Patent 6,333,014.

[98] M. J. Matzen, M. H. Alhajji, and Y. Demirel, "Technoeconomics and sustainability of renewable methanol and ammonia productions using wind power-based hydrogen," 2015.

[99] C. D. Demirhan, W. W. Tso, J. B. Powell, and E. N. Pistikopoulos, "Sustainable ammonia production through process synthesis and global optimization," *AIChE Journal*, vol. 65, no. 7, p. e16498, 2019.

[100] A. M. Niziolek, O. Onel, and C. A. Floudas, "Production of benzene, toluene, and xylenes from natural gas via methanol: Process synthesis and global optimization," *AIChE Journal*, vol. 62, no. 5, pp. 1531–1556, 2016.

[101] L. R. Matthews, A. M. Niziolek, O. Onel, N. Pinnaduwage, and C. A. Floudas, "Biomass to liquid transportation fuels via biological and thermochemical conversion: process synthesis and global optimization strategies," *Industrial & Engineering Chemistry Research*, vol. 55, no. 12, pp. 3203–3225, 2016.

[102] W. W. Tso, C. D. Demirhan, J. B. Powell, and E. N. Pistikopoulos, "Toward optimal synthesis of renewable ammonia and methanol processes (RAMP)," in *Computer Aided Chemical Engineering*, vol. 44, pp. 1705–1710, Elsevier, 2018.

[103] United States Department of Agriculture - Economic Research Service, "Fertilizer Use and Price," 2019. `https://www.ers.usda.gov/data-products/fertilizer-use-and-price.aspx`, Last accessed on 2019-03-01.

[104] Methanex, "Current Posted Prices," 2019. `https://www.methanex.com/our-business/pricing`, Last accessed on 2019-03-01.

[105] U.S. Department of Energy, "2016 Billion-Ton Report: Advancing Domestic Resources for a Thriving Bioeconomy, Volume 1: Economic Availability of Feedstocks." `https://info.ornl.gov/sites/publications/Files/Pub62368.pdf`, Last accessed on 2019-03-02.

[106] A. Almansoori and N. Shah, "Design and operation of a future hydrogen supply chain: multi-period model," *International Journal of Hydrogen Energy*, vol. 34, no. 19, pp. 7883–7897, 2009.

[107] M. Reuß, T. Grube, M. Robinius, P. Preuster, P. Wasserscheid, and D. Stolten, "Seasonal storage and alternative carriers: A flexible hydrogen supply chain model," *Applied Energy*, vol. 200, pp. 290–302, 2017.

[108] H. Dagdougui, "Models, methods and approaches for the planning and design of the future hydrogen supply chain," *International Journal of Hydrogen Energy*, vol. 37, no. 6, pp. 5318–5327, 2012.

[109] U.S. Energy Information Administration, "Texas - State Energy Profile Overview," 2019. https://www.eia.gov/state/?sid=TX, Last accessed on 2019-03-18.

[110] Solar Energy Industries Association, "Texas Solar," 2019. https://www.seia.org/state-solar-policy/texas-solar, Last accessed on 2019-03-18.

[111] National Renewable Energy Laboratory, "National Solar Radiation Database (NSRDB)," 2019. https://nsrdb.nrel.gov/, Last accessed on 2019-03-18.

[112] National Renewable Energy Laboratory, "Wind Integration National Dataset Toolkit," 2019. https://www.nrel.gov/grid/wind-toolkit.html, Last accessed on 2019-03-18.

[113] A. Lopez, B. Roberts, D. Heimiller, N. Blair, and G. Porro, "Us renewable energy technical potentials: A gis-based analysis," tech. rep., NREL, 2012.

[114] U.S. Geological Survey, "Water Use in the United States," 2019. https://water.usgs.gov/watuse/index.html, Last accessed on 2019-03-18.

[115] C. D. Demirhan, W. W. Tso, J. B. Powell, C. F. Heuberger, and E. N. Pistikopoulos, "A multi-scale energy systems engineering approach for renewable power generation and storage optimization," *Industrial & Engineering Chemistry Research*, 2020.

[116] I. Harjunkoski, C. T. Maravelias, P. Bongers, P. M. Castro, S. Engell, I. E. Grossmann, J. Hooker, C. Méndez, G. Sand, and J. Wassick, "Scope for industrial applications of production scheduling models and solution methods," *Computers & Chemical Engineering*, vol. 62, pp. 161–193, 2014.

[117] P. M. Castro, I. E. Grossmann, and Q. Zhang, "Expanding scope and computational challenges in process scheduling," *Computers & Chemical Engineering*, vol. 114, pp. 14–42, 2018.

[118] N. E. Koltsaklis and A. S. Dagoumas, "State-of-the-art generation expansion planning: A review," *Applied Energy*, vol. 230, pp. 563–589, 2018.

[119] B. Saravanan, S. Das, S. Sikri, and D. Kothari, "A solution to the unit commitment problemâĂŤa review," *Frontiers in Energy*, vol. 7, no. 2, pp. 223–236, 2013.

[120] N. Baumgärtner, B. Bahl, M. Hennen, and A. Bardow, "Rises3: Rigorous synthesis of energy supply and storage systems via time-series relaxation and aggregation," *Computers & Chemical Engineering*, vol. 127, pp. 127–139, 2019.

[121] P. Gabrielli, M. Gazzani, E. Martelli, and M. Mazzotti, "Optimal design of multi-energy systems with seasonal storage," *Applied Energy*, vol. 219, pp. 408–424, 2018.

[122] C. L. Lara, D. S. Mallapragada, D. J. Papageorgiou, A. Venkatesh, and I. E. Grossmann, "Deterministic electric power infrastructure planning: Mixed-integer programming model and nested decomposition algorithm," *European Journal of Operational Research*, vol. 271, no. 3, pp. 1037–1054, 2018.

[123] C. F. Heuberger, I. Staffell, N. Shah, and N. Mac Dowell, "A systems approach to quantifying the value of power generation and energy storage technologies in future electricity networks," *Computers & Chemical Engineering*, vol. 107, pp. 247–256, 2017.

[124] X. Peng, T. W. Root, and C. T. Maravelias, "Optimization-based process synthesis under seasonal and daily variability: Application to concentrating solar power," *AIChE Journal*, vol. 65, no. 7, p. e16458, 2019.

[125] S. Pineda and J. M. Morales, "Chronological time-period clustering for optimal capacity expansion planning with storage," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 7162–7170, 2018.

[126] H. Teichgraeber and A. R. Brandt, "Clustering methods to find representative periods for the optimization of energy systems: An initial framework and comparison," *Applied Energy*, vol. 239, pp. 1283–1293, 2019.

[127] D. A. Tejada-Arango, M. Domeshek, S. Wogrin, and E. Centeno, "Enhanced representative days and system states modeling for energy storage investment analysis," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6534–6544, 2018.

[128] F. Domínguez-Muñoz, J. M. Cejudo-López, A. Carrillo-Andrés, and M. Gallardo-Salazar, "Selection of typical demand days for chp optimization," *Energy and Buildings*, vol. 43, no. 11, pp. 3036–3043, 2011.

[129] B. Bahl, A. Kümpel, H. Seele, M. Lampe, and A. Bardow, "Time-series aggregation for synthesis problems by bounding error in the objective function," *Energy*, vol. 135, pp. 900–912, 2017.

[130] L. Kotzur, P. Markewitz, M. Robinius, and D. Stolten, "Time series aggregation for energy system design: Modeling seasonal storage," *Applied Energy*, vol. 213, pp. 123–135, 2018.

[131] New York Independent System Operator, "Energy Market & Operational Data." `https://www.nyiso.com/energy-market-operational-data`, Last accessed on 2019-01-23.

[132] W. W. Tso, C. D. Demirhan, S. Lee, H. Song, J. B. Powell, and E. N. Pistikopoulos, "Energy carrier supply chain optimization: A texas case study," in *Computer Aided Chemical Engineering*, vol. 47, pp. 1–6, Elsevier, 2019.

[133] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning*, vol. 112. Springer, 2013.

[134] Y. S. Abu-Mostafa, M. Magdon-Ismail, and H.-T. Lin, *Learning From Data*, vol. 4. AML-Book New York, NY, USA, 2012.

[135] Z. T. Wilson and N. V. Sahinidis, "The ALAMO approach to machine learning," *Computers & Chemical Engineering*, vol. 106, pp. 785–795, 2017.

[136] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Science & Business Media, 2009.

[137] Y. Bengio, "Gradient-based optimization of hyperparameters," *Neural Computation*, vol. 12, no. 8, pp. 1889–1900, 2000.

[138] C.-s. Foo, C. B. Do, and A. Y. Ng, "Efficient multiple hyperparameter learning for log-linear models," in *Advances in Neural Information Processing Systems*, pp. 377–384, 2008.

[139] J. S. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for hyper-parameter optimization," in *Advances in Neural Information Processing Systems*, pp. 2546–2554, 2011.

[140] M. Claesen and B. De Moor, "Hyperparameter search in machine learning," *arXiv preprint arXiv:1502.02127*, 2015.

[141] F. Hutter, J. Lücke, and L. Schmidt-Thieme, "Beyond manual tuning of hyperparameters," *KI-Künstliche Intelligenz*, vol. 29, no. 4, pp. 329–337, 2015.

[142] G. Luo, "A review of automatic selection methods for machine learning algorithms and hyper-parameter values," *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 5, no. 1, p. 18, 2016.

[143] R. Liu, E. Liu, J. Yang, M. Li, and F. Wang, "Optimizing the hyper-parameters for svm by combining evolution strategies with a grid search," in *Intelligent Control and Automation*, pp. 712–721, Springer, 2006.

[144] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of Machine Learning Research*, vol. 13, no. Feb, pp. 281–305, 2012.

[145] A. Bemporad, M. Morari, V. Dua, and E. N. Pistikopoulos, "The explicit linear quadratic regulator for constrained systems," *Automatica*, vol. 38, no. 1, pp. 3–20, 2002.

[146] E. Pistikopoulos, "Perspectives in multiparametric programming and explicit model predictive control," *AIChE Journal*, vol. 55, no. 8, pp. 1918–1925, 2009.

[147] E. N. Pistikopoulos, "From multi-parametric programming theory to mpc-on-a-chip multi-scale systems applications," *Computers & Chemical Engineering*, vol. 47, pp. 57–66, 2012.

[148] R. Oberdieck, N. A. Diangelakis, I. Nascu, M. M. Papathanasiou, M. Sun, S. Avraamidou, and E. N. Pistikopoulos, "On multi-parametric programming and its applications in process systems engineering," *Chemical Engineering Research and Design*, vol. 116, pp. 61–82, 2016.

[149] E. N. Pistikopoulos, M. C. Georgiadis, and V. Dua, *Multi-Parametric Programming: Theory, Algorithms, and Applications*, vol. 1. 2011.

[150] E. N. Pistikopoulos, N. A. Diangelakis, and R. Oberdieck, *Multi-parametric Optimization and Control*, vol. 1. 2020.

[151] N. A. Diangelakis, B. Burnak, J. Katz, and E. N. Pistikopoulos, "Process design and control optimization: A simultaneous approach by multi-parametric programming," *AIChE Journal*, vol. 63, no. 11, pp. 4827–4846, 2017.

[152] N. A. Diangelakis, I. S. Pappas, and E. N. Pistikopoulos, "On multiparametric/explicit nmpc for quadratically constrained problems," *IFAC-PapersOnLine*, vol. 51, no. 20, pp. 400–405, 2018.

[153] B. Burnak, N. A. Diangelakis, J. Katz, and E. N. Pistikopoulos, "Integrated process design, scheduling, and control using multiparametric programming," *Computers & Chemical Engineering*, vol. 125, pp. 164–184, 2019.

[154] G. S. Ogumerem and E. N. Pistikopoulos, "Parametric optimization and control toward the design of a smart metal hydride refueling system," *AIChE Journal*, vol. 65, no. 10, p. e16680, 2019.

[155] M. Onel, B. Burnak, and E. N. Pistikopoulos, "Integrated data-driven process monitoring and explicit fault-tolerant multiparametric control," *Industrial & Engineering Chemistry Research*, 2019.

[156] Y. Tian, I. Pappas, B. Burnak, J. Katz, and E. N. Pistikopoulos, "A systematic framework for the synthesis of operable process intensification systems–reactive separation systems," *Computers & Chemical Engineering*, vol. 134, p. 106675, 2020.

[157] B. Colson, P. Marcotte, and G. Savard, "An overview of bilevel optimization," *Annals of Operations Research*, vol. 153, no. 1, pp. 235–256, 2007.

[158] A. Sinha, P. Malo, and K. Deb, "A review on bilevel optimization: From classical to evolutionary approaches and applications," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 2, pp. 276–295, 2017.

[159] K. P. Bennett, J. Hu, X. Ji, G. Kunapuli, and J.-S. Pang, "Model selection via bilevel optimization," in *The 2006 IEEE International Joint Conference on Neural Network Proceedings*, pp. 1922–1929, IEEE, 2006.

[160] T. Klatzer and T. Pock, "Continuous hyper-parameter learning for support vector machines," in *Computer Vision Winter Workshop (CVWW)*, pp. 39–47, 2015.

[161] F. Pedregosa, "Hyperparameter optimization with approximate gradient," *arXiv preprint arXiv:1602.02355*, 2016.

[162] L. Franceschi, P. Frasconi, S. Salzo, R. Grazzi, and M. Pontil, "Bilevel programming for hyperparameter optimization and meta-learning," *arXiv preprint arXiv:1806.04910*, 2018.

[163] M. MacKay, P. Vicol, J. Lorraine, D. Duvenaud, and R. Grosse, "Self-tuning networks: Bilevel optimization of hyperparameters using structured best-response functions," *arXiv preprint arXiv:1903.03088*, 2019.

[164] S. Avraamidou and E. N. Pistikopoulos, "B-pop: Bi-level parametric optimization toolbox," *Computers & Chemical Engineering*, vol. 122, pp. 193–202, 2019.

[165] N. P. Faísca, V. Dua, B. Rustem, P. M. Saraiva, and E. N. Pistikopoulos, "Parametric global optimisation for bilevel programming," *Journal of Global Optimization*, vol. 38, no. 4, pp. 609–623, 2007.

[166] L. F. Domínguez and E. N. Pistikopoulos, "Multiparametric programming based algorithms for pure integer and mixed-integer bilevel programming problems," *Computers & Chemical Engineering*, vol. 34, no. 12, pp. 2097–2106, 2010.

[167] R. Oberdieck, N. A. Diangelakis, S. Avraamidou, and E. N. Pistikopoulos, "On unbounded and binary parameters in multi-parametric programming: applications to mixed-integer bilevel optimization and duality theory," *Journal of Global Optimization*, vol. 69, no. 3, pp. 587–606, 2017.

[168] S. Avraamidou and E. N. Pistikopoulos, "A multiparametric mixed-integer bi-level optimization strategy for supply chain planning under demand uncertainty," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 10178–10183, 2017.

[169] S. Avraamidou and E. N. Pistikopoulos, "A multi-parametric optimization approach for bilevel mixed-integer linear and quadratic programming problems," *Computers & Chemical Engineering*, vol. 125, pp. 98–113, 2019.

[170] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[171] B. Beykal, F. Boukouvala, C. A. Floudas, N. Sorek, H. Zalavadia, and E. Gildin, "Global optimization of grey-box computational systems using surrogate functions and application to highly constrained oil-field operations," *Computers & Chemical Engineering*, vol. 114, pp. 99–110, 2018.

[172] B. Beykal, F. Boukouvala, C. A. Floudas, and E. N. Pistikopoulos, "Optimal design of energy systems using constrained grey-box multi-objective optimization," *Computers & Chemical Engineering*, vol. 116, pp. 488–502, 2018.

[173] R. Oberdieck, N. A. Diangelakis, M. M. Papathanasiou, I. Nascu, and E. N. Pistikopoulos, "Pop–parametric optimization toolbox," *Industrial & Engineering Chemistry Research*, vol. 55, no. 33, pp. 8979–8991, 2016.

[174] B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, *et al.*, "Least angle regression," *The Annals of Statistics*, vol. 32, no. 2, pp. 407–499, 2004.

[175] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *Journal of Statistical Software*, vol. 33, no. 1, pp. 1–22, 2010.

[176] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, pp. 144–152, ACM, 1992.

[177] M. Onel, C. A. Kieslich, Y. A. Guzman, C. A. Floudas, and E. N. Pistikopoulos, "Big data approach to batch process monitoring: Simultaneous fault detection and diagnosis using nonlinear support vector machine-based feature selection," *Computers & Chemical Engineering*, vol. 115, pp. 46–63, 2018.

[178] M. Onel, C. A. Kieslich, and E. N. Pistikopoulos, "A nonlinear support vector machine-based feature selection approach for fault detection and diagnosis: Application to the tennessee eastman process," *AIChE Journal*, vol. 65, no. 3, pp. 992–1005, 2019.

[179] P. S. Bradley and O. L. Mangasarian, "Feature selection via concave minimization and support vector machines.," in *ICML*, vol. 98, pp. 82–90, 1998.

[180] M. Patrício, J. Pereira, J. Crisóstomo, P. Matafome, M. Gomes, R. Seiça, and F. Caramelo, "Using resistin, glucose, age and bmi to predict the presence of breast cancer," *BMC Cancer*, vol. 18, no. 1, p. 29, 2018.

[181] T. Hastie, S. Rosset, R. Tibshirani, and J. Zhu, "The entire regularization path for the support vector machine," *Journal of Machine Learning Research*, vol. 5, no. Oct, pp. 1391–1415, 2004.

[182] J. Zhu, S. Rosset, R. Tibshirani, and T. J. Hastie, "1-norm support vector machines," in *Advances in Neural Information Processing Systems*, pp. 49–56, 2004.

[183] W. R. Stahel, "The circular economy," *Nature*, vol. 531, no. 7595, pp. 435–438, 2016.

[184] M. Geissdoerfer, P. Savaget, N. M. Bocken, and E. J. Hultink, "The circular economy–a new sustainability paradigm?," *Journal of Cleaner Production*, vol. 143, pp. 757–768, 2017.

[185] S. Avraamidou, S. G. Baratsas, Y. Tian, and E. N. Pistikopoulos, "Circular economy-a challenge and an opportunity for process systems engineering," *Computers & Chemical Engineering*, vol. 133, p. 106629, 2020.

[186] Waste Dive, "World bank: Global waste generation could increase 702050." `https://www.wastedive.com/news/world-bank-global-waste-generation-2050/533031/`, Last accessed on 2020-02-11.

[187] A. M. Niziolek, O. Onel, M. F. Hasan, and C. A. Floudas, "Municipal solid waste to liquid transportation fuels–part ii: Process synthesis and global optimization strategies," *Computers & Chemical Engineering*, vol. 74, pp. 184–203, 2015.

[188] A. M. Niziolek, O. Onel, and C. A. Floudas, "Municipal solid waste to liquid transportation fuels, olefins, and aromatics: Process synthesis and deterministic global optimization," *Computers & Chemical Engineering*, vol. 102, pp. 169–187, 2017.

[189] J. A. Elia, J. Li, and C. A. Floudas, "Strategic planning optimization for natural gas to liquid transportation fuel (gtl) systems," *Computers & Chemical Engineering*, vol. 72, pp. 109–125, 2015.

[190] Y. A. Guzman, L. R. Matthews, and C. A. Floudas, "New a priori and a posteriori probabilistic bounds for robust counterpart optimization: I. unknown probability distributions," *Computers & Chemical Engineering*, vol. 84, pp. 568–598, 2016.

[191] Y. A. Guzman, L. R. Matthews, and C. A. Floudas, "New a priori and a posteriori probabilistic bounds for robust counterpart optimization: Ii. a priori bounds for known symmetric and asymmetric probability distributions," *Computers & Chemical Engineering*, vol. 101, pp. 279–311, 2017.

[192] Y. A. Guzman, L. R. Matthews, and C. A. Floudas, "New a priori and a posteriori probabilistic bounds for robust counterpart optimization: Iii. exact and near-exact a posteriori expressions for known probability distributions," *Computers & Chemical Engineering*, vol. 103, pp. 116–143, 2017.

[193] J. Acevedo and E. N. Pistikopoulos, "Stochastic optimization based algorithms for process synthesis under uncertainty," *Computers & Chemical Engineering*, vol. 22, no. 4-5, pp. 647–671, 1998.

[194] Z. Zhou, J. Zhang, P. Liu, Z. Li, M. C. Georgiadis, and E. N. Pistikopoulos, "A two-stage stochastic programming model for the optimal design of distributed energy systems," *Applied Energy*, vol. 103, pp. 135–144, 2013.

[195] M. Fischetti and J. Jo, "Deep neural networks and mixed integer linear optimization," *Constraints*, vol. 23, no. 3, pp. 296–309, 2018.

[196] D. Bertsimas and J. Dunn, "Optimal classification trees," *Machine Learning*, vol. 106, no. 7, pp. 1039–1082, 2017.

[197] I. Gkioulekas and L. G. Papageorgiou, "Optimal regression tree models through mixed integer programming," in *Data Science–Analytics and Applications*, pp. 57–62, Springer, 2019.

[198] S. Maldonado, J. Pérez, R. Weber, and M. Labbé, "Feature selection for support vector machines via mixed integer linear programming," *Information Sciences*, vol. 279, pp. 163–175, 2014.

[199] Texas A&M Real Estate Center, "Rural land." `https://www.recenter.tamu.edu/data/rural-land/`, Last accessed on 2020-02-13.

[200] C. Moné, M. Hand, M. Bolinger, J. Rand, D. Heimiller, and J. Ho, "2015 cost of wind energy review," tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States), 2017.

[201] R. Fu, D. Feldman, R. Margolis, M. Woodhouse, and K. Ardani, "Us solar photovoltaic system cost benchmark: Q1 2017," tech. rep., EERE Publication and Product Library, 2017.

[202] Q. Zhang, M. Martín, and I. E. Grossmann, "Integrated design and operation of renewables-based fuels and power production networks," *Computers & Chemical Engineering*, vol. 122, pp. 80–92, 2019.

[203] A. Pinto, "Ammonia production process," July 21 1987. US Patent 4,681,745.

[204] A. Pinto, "Synthesis process and reactor," Oct. 18 1988. US Patent 4,778,662.

[205] C. P. Singh and D. N. Saraf, "Simulation of ammonia synthesis reactors," *Industrial & Engineering Chemistry Process Design and Development*, vol. 18, no. 3, pp. 364–370, 1979.

[206] S. Lou and R. Lou, "Method of catalytic reaction carried out near the optimal temperature and an apparatus for the method," Apr. 10 2001. US Patent 6,214,296.

[207] S. S. Elnashaie, M. E. Abashar, and A. S. Al-Ubaid, "Simulation and optimization of an industrial ammonia reactor," *Industrial & Engineering Chemistry Research*, vol. 27, no. 11, pp. 2015–2022, 1988.

[208] T. A. Czuppon, "Isothermal ammonia converter," Jan. 9 2001. US Patent 6,171,570.

[209] D. L. Banquy, "Process for the production of ammonia and the corresponding synthesis gas," Oct. 20 1981. US Patent 4,296,085.

[210] U. Zardi, "System for reducing energy consumption improving reactors for heterogeneous catalytic synthesis and relative reactors," July 5 1988. US Patent 4,755,362.

[211] A. Pinto and J. B. Johnson, "Ammonia synthesis process," Sept. 22 1987. US Patent 4,695,442.

[212] M. Azarhoosh, F. Farivar, and H. A. Ebrahim, "Simulation and optimization of a horizontal ammonia synthesis reactor using genetic algorithm," *Rsc Advances*, vol. 4, no. 26, pp. 13419–13429, 2014.

[213] M. Appl, *Ammonia: Principles and Industrial Practice*, vol. 1. 1999.

[214] H. F. Topsoe and E. A. Gam, "Apparatus and process for the synthesis of ammonia," Jan. 1 1980. US Patent 4,181,701.

[215] A. Notman, "Synthesis reactor," Jan. 19 1982. US Patent 4,311,671.

[216] M. Panahandeh, J. Fathikaljahi, and M. Taheri, "Steady-state modeling and simulation of an axial-radial ammonia synthesis reactor," *Chemical Engineering & Technology: Industrial Chemistry-Plant Equipment-Process Engineering-Biotechnology*, vol. 26, no. 6, pp. 666–671, 2003.

[217] J. R. LeBlanc and S. A. Knez, "Ammonia production with enriched air reforming and nitrogen injection into the synthesis loop," Apr. 7 1998. US Patent 5,736,116.

[218] L. D. Gaines, "Ammonia synthesis loop variables investigated by steady-state simulation," *Chemical Engineering Science*, vol. 34, no. 1, pp. 37–50, 1979.

[219] C. Speth, "Process for the preparation of ammonia," Oct. 18 2005. US Patent 6,955,797.

[220] R. B. Strait, "Pseudoisothermal ammonia process," Feb. 22 2011. US Patent 7,892,511.

APPENDIX A

ECN MODEL SUPPLEMENTARY INFORMATION

A.1 Nomenclature Definitions

Indices

f	Feedstock index
p	Product (energy carrier) index
l	Facility location index
t	Capacity index (50, 100, 200, 500, 750 MT H_2 eq./day)
q	Facility type index (H_2, NH_3, CH_3OH)
d	Demand location index
k	Conversion technology index
m	Transportation mode index (rail, truck, pipe)
s	Resource location index

Sets

F	Feedstocks (CO_2, H_2O)
P	Energy carriers (H_2, NH_3, CH_3OH)
L^F	Candidate facility locations (county centers)
L^D	Demand locations (five cities)
S	Resource areas
K	Conversion technologies (fuel cell, gas turbine)

Parameters

$FR_{f,t,q}$	Feedstock f needed for a facility type q of capacity t
$FA_{f,s}$	Availability of feedstock f at area s
$ER_{t,q}$	Electricity required for facility type q of capacity t
PV_{CF}	Capacity factor for solar PV modules
PV_{Eff}	Solar to electricity conversion efficiency
GHI_l	Global horizontal irradiance at location l
SL_l	Available land for solar PV at location l
T_{CF}	Capacity factor for wind turbines
T_{Eff}	Wind to electricity conversion efficiency
WP_l	Wind power at location l
WL_l	Available land for turbines at location l

T_{Land}	Land usage per wind turbine
v_l	Wind speed at location l
C_p	Turbine power coefficient
ρ	Air density
D	Turbine rotator diameter
$PR_{p,t,q}$	Product p from facility q of capacity t
LHV_p	Lower heating value of product p
$Eff_{p,k}$	Power efficiency using conversion technology k for product p
$Store_{time}$	Time period between production and consumption
$Total$	Total power demand from the network
$Demand_d$	Power demand limit at location d
$IC_{t,q}$	Investment cost for facility q of capacity t
HC_p	Hydrogen weight density of product p
CCR	Capital cost discount percentage rate
OPT	Number of days facility is in operation
$Cost_{PV}$	Solar PV capital cost per MW
$Cost_{Turb}$	Wind turbine cost per MW
CF_k	Capacity factor for conversion technology k
$Cost_k$	Capital cost per MW for conversion technology k
$Cost_p$	Product storage cost (\$/kg H_2 eq./day)
$FC_{f,t,q}$	Feedstock f costs for facility type q of capacity t
$Cost_{f,s,l,m}$	Transportation costs for feedstock f from area s to location l using mode m
$Cost_{p,l,d,m}$	Transportation costs for product p from location l to location d using mode m
PV_{Lease}	Land leasing costs for solar PV
T_{Lease}	Land leasing costs for wind turbine
LF_l	Cost factor for land leasing
$OM_{t,q}$	O&M costs for facility type q of capacity t
OMR	Annual O&M cost rate percentage
$Oxy_{t,q}$	Oxygen sales for facility type q of capacity t

Binary Variables

$y_{l,t,q}$	Select facility of type q and capacity t at location l
$y^D_{p,d,k}$	Select conversion technology k for product p at location d

Continuous Variables

$Feed_{f,l}$	Total feedstock f required at location l
$x_{f,s,l,m}$	Feedstock f flow from area s to location l using transportation mode m
$Elec_l$	Total electricity utilized at location l
PV_l	Solar PV electricity at location l
$Turb_l$	Wind turbine electricity at location l

$z_{p,l,d,m}$	Product p flow from location l to location d using transportation mode m
$Prod_{p,d}$	Total product p sent to location d
$Cap_{p,d,k}$	Required capacity of conversion technology k for product p at location d
$Cost_{Fac}$	Total production facilities capital cost
$Cost_{Renew}$	Total renewable farms capital cost
$Cost_{Conv}$	Total conversion technologies capital cost
$Cost_{Store}$	Total storage cost
$Cost_{FP}$	Total feedstock purchases cost
$Cost_{FT}$	Total feedstock transportation cost
$Cost_{PT}$	Total product transportation cost
$Cost_{Land}$	Total land leasing cost
$Cost_{O\&M}$	Total operation & maintenance cost
$Sales_{Oxy}$	Total sales from oxygen

A.2 Parameter Values

In the model, equations for material flows are in kg/day, and power units are in MW or MWh. The parameter values for the model are described below. Solar PV capacity factor (PV_{CF}) is 24%, and solar to electricity efficiency (PV_{Eff}) is 22.5%. Solar PV capital cost ($Cost_{PV}$) is \$985,093/MW, and the leasing cost (PV_{Lease}) is \$6177/MW. Wind turbine capacity factor (T_{CF}) is 40%, and wind to power efficiency loss (T_{Eff}) is 15%. The land utilized per wind turbine (T_{Land}) is 0.69 km^2/MW. C_p is 0.47, ρ is 1.163 kg/m^3, and D is 108 m. Wind turbine capital cost ($Cost_{Turb}$) is \$1,159,000/MW, and leasing cost (T_{Lease}) is \$4000/MW. The capital cost discount rate (CCR) is 15.41%/yr., and the facilities are assumed to operate 330 days/yr. (OPT). O&M rates for the conversion technologies, solar PV, and wind turbines are 4%/yr., 1.8%/yr., 2.7%/yr., respectively.

Transportation costs for feedstocks ($Cost_{f,s,l,m}$) and products ($Cost_{p,l,d,m}$) are calculated off the $distance$ between two locations using the Haversine formula. DFC is the distance fixed cost, DVC is the distance variable cost, and DM is a distance multiplier. DFC for water and CO_2 by pipeline are \$0.0003/kg and \$0.4886/kg, respectively. DFC for water and CO_2 are \$5e-6/kg/mi and \$1.61e-4/kg/mi, respectively. Table A.1 shows the DFC and DVC for the energy carriers. DM accounts for path curvatures and is assumed to be 1.1 for truck & pipeline and 1.05 for rail.

$$TransportCost = DFC + DVC \times distance \times DM \tag{A.1}$$

Table A.1: Transportation Costs

Mode Cost	H_2	NH_3	CH_3OH
Truck DFC (\$/kg)	0.117891	0.007538	0.004290
Truck DVC (\$/kg/mi)	0.002724	0.000224	0.000176
Rail DFC (\$/kg)	0.593454	0.010674	0.014353
Rail DVC (\$/kg/mi)	0.000101	0.000035	0.000037

The storage cost ($Cost_p$), lower heating values (LHV_p), and hydrogen densities (HC_p) are listed in Table A.2. $Store_{Time}$ has unit of days.

Table A.2: Product Information

Product	$Cost_p$ (\$/kg H_2 eq./day)	LHV_p (MJ/kg)	HC_p (%)
H_2	0.106738	119.96	100
NH_3	0.003484	18.60	17.76
CH_3OH	0.003022	20.09	12.58

The capacity factors for the conversion technologies (CF_k) are 55% for a gas turbine and 80% for a fuel cell, respectively. The capital costs ($Cost_k$) are \$1,024,00/MW and \$1,600,000/MW for the turbine and fuel cell, respectively. It is assumed that capacity factor and costs are independent of energy carrier. The conversion efficiencies ($Eff_{p,k}$) are dependent on the product and are shown in Table A.3.

Table A.3: Conversion Technology Efficiencies

$Eff_{p,k}$	H_2	NH_3	CH_3OH
Gas turbine	0.40	0.25	0.40
Fuel Cell	0.60	0.45	0.35

Feedstock availability ($FA_{f,s}$) for water and CO_2 are taken from USGS and NETL ATLAS, respectively. Solar irradiance (GHI_l) in kWh/m^2/day and wind speed (v_l) in m/s are extracted from NREL NSRDB and WIND Toolkit. Land availabilities (km^2) for solar PV (SL_l) and wind farm (WL_l) are derived from a NREL study [113]. The cost factor for land leasing is taken from the Texas A&M Real Estate center database [199]. The average land cost is used a normalizing basis.

Information for optimal hydrogen, ammonia, and methanol plants are shown in Tables A.4 to A.9. These values are determined from an earlier process synthesis similar to the ones presented in Sections 2 and 3.

Table A.4: Hydrogen Plant Input/Output Information

Capacity	FR_{H_2O} (kBPD)	FR_{CO_2} (MT/day)	ER (MW)	PR (MT/day)
50 MT H_2 eq./day	6.86255	0	111.9012	50
100 MT H_2 eq./day	13.72510	0	223.8023	100
200 MT H_2 eq./day	27.45020	0	447.6047	200
500 MT H_2 eq./day	68.62549	0	1119.012	500
750 MT H_2 eq./day	102.9382	0	1678.518	750

Table A.5: Hydrogen Plant Cost Information

Capacity	IC ($/kg H_2 eq.)	OM ($/kg H_2 eq.)	FC_{H_2O} ($/kg H_2 eq.)	FC_{CO_2} ($/kg H_2 eq.)	Oxy_{Sales} ($/kg H_2 eq.)
50 MT H_2 eq./day	0.7301	0.1714	0.0273	0	0.3968
100 MT H_2 eq./day	0.6435	0.1510	0.0273	0	0.3968
200 MT H_2 eq./day	0.5722	0.1343	0.0273	0	0.3968
500 MT H_2 eq./day	0.5014	0.1177	0.0273	0	0.3968
750 MT H_2 eq./day	0.4766	0.1118	0.0273	0	0.3968

Table A.6: Ammonia Plant Input/Output Information

Capacity	FR_{H_2O} (kBPD)	FR_{CO_2} (MT/day)	ER (MW)	PR (MT/day)
50 MT H_2 eq./day	8.75818	0	145.4584	281.6
100 MT H_2 eq./day	17.26955	0	286.7563	563.2
200 MT H_2 eq./day	34.5391	0	573.5127	1126.5
500 MT H_2 eq./day	86.52437	0	1434.426	2816.1
750 MT H_2 eq./day	129.5477	0	2151.228	4224.2

Table A.7: Ammonia Plant Cost Information

Capacity	IC ($/kg H_2 eq.)	OM ($/kg H_2 eq.)	FC_{H_2O} ($/kg H_2 eq.)	FC_{CO_2} ($/kg H_2 eq.)	Oxy_{Sales} ($/kg H_2 eq.)
50 MT H_2 eq./day	1.5565	0.3572	0.0348	0	0.4713
100 MT H_2 eq./day	1.3511	0.3171	0.0343	0	0.4713
200 MT H_2 eq./day	1.1600	0.2723	0.0343	0	0.4713
500 MT H_2 eq./day	1.0024	0.2353	0.0344	0	0.4719
750 MT H_2 eq./day	0.9306	0.2184	0.0343	0	0.4713

Table A.8: Methanol Plant Input/Output Information

Capacity	FR_{H_2O} (kBPD)	FR_{CO_2} (MT/day)	ER (MW)	PR (MT/day)
50 MT H_2 eq./day	12.92597	545.8	174.0552	397.3
100 MT H_2 eq./day	25.78618	1091.6	348.0658	794.7
200 MT H_2 eq./day	51.45858	2183.2	693.8080	1589.4
500 MT H_2 eq./day	113.5816	5458.0	1740.668	3973.5
750 MT H_2 eq./day	192.9694	8187.9	2601.775	5960.2

Table A.9: Methanol Plant Cost Information

Capacity	IC ($/kg H_2 eq.)	OM ($/kg H_2 eq.)	FC_{H_2O} ($/kg H_2 eq.)	FC_{CO_2} ($/kg H_2 eq.)	Oxy_{Sales} ($/kg H_2 eq.)
50 MT H_2 eq./day	1.6391	0.3846	0.0514	0.2406	0.5979
100 MT H_2 eq./day	1.4324	0.3362	0.0512	0.2406	0.5979
200 MT H_2 eq./day	1.2869	0.3020	0.0511	0.2406	0.6261
500 MT H_2 eq./day	1.0802	0.2535	0.0451	0.2406	0.5979
750 MT H_2 eq./day	1.0375	0.2435	0.0511	0.2406	0.6261

APPENDIX B

RENEWABLE POWER SYSTEM WITH STORAGE SUPPLEMENTARY INFORMATION

The information here is extracted from Demirhan et al. [115].

B.1 Nomenclature Definitions

Sets

t	Time discretization in hours
h	Representative periods
a	Location
i	Processes
j	Resources
m	Operating modes
l	Piecewise cost function segments

Subsets

$Time_{h,t}$	Hour t in period h
$Period_h$	Set of periods considered
$Modes_{i,m}$	Operating modes m in process i
$TransModes_{i,m,m'}$	Possible mode transitions from m to m' in process i
$Seq_{i,m,m',m''}$	Predefined sequences of mode transitions for process i
$Demand_j$	Resources j for which demand exists
$PLSegments_{i,l}$	Segments l in piecewise linear approximations for process i

Variables

$TotalCost$	Total annualized cost
$Inv^{excess}_{a,j,h}$	Excess inventory for resource j in period h at location a

Positive Variables

$B_{a,j,h,t}$	Amount of resource j purchased in time t of period h at location a
$Cap^P_{a,i}$	Production capacity for process i at location a based on reference resource
$Cap^S_{a,j}$	Storage capacity for resource j at location a
$Capex_{a,i}$	Overnight capital cost for process i at location a
$Capex^{total}_a$	Total overnight capital expenses at location a

$Inv_{a,j,h,t}$	Inventory level of resource j in time t of period h at location a
$\lambda_{a,i,l}$	Coefficient for segment l in piecewise linear approximation for process i at location a
$Opex_a$	Total annual operation cost at location a
$P_{a,i,h,t}$	Amount of reference resource consumed or produced by process i in time t of period h at location a
$P^m_{a,i,m,h,t}$	Amount of reference resource consumed or produced by process i in mode m in time t of period h at location a
$S_{a,j,h,t}$	Amount of resource j sold in time t of period h at location a

Binary Variables

$w_{a,i,l}$	Equals 1 if the capacity for process i at location a is in the range of line segment l
$x^P_{a,i}$	Equals 1 if process i at location a is built
$x^S_{a,j}$	Equals 1 if storage facility for resource j is built at location a
$y_{a,i,m,h,t}$	Equals 1 if process i at location a operates in mode m in time t of period h
$z_{a,i,m,m',h,t}$	Equals 1 if process i at location a in mode m changes to mode m' in time t of period h

Parameters

$B^{max}_{a,j,h,t}$	Maximum amount of resource j that can be consumed in time t of period h at location a
$BigM$	Big-M parameter
CAP^{P-max}_i	Maximum production capacity for process i
CAP^{S-max}_j	Maximum storage capacity for resource j
$CAP^{segment}_{i,l}$	Capacity of process i at the right end point of segment l
$CAP^{util-min}_{i,m}$	Minimum capacity utilization fraction for mode m of process i
$CAP^{util-max}_{i,m}$	Maximum capacity utilization fraction for mode m of process i
$CAP^{mode-min}_{i,m}$	Minimum production capacity for mode m of process i
$CAP^{mode-max}_{i,m}$	Maximum production capacity for mode m of process i
$CAP^{\Delta rate}_{i,m}$	Maximum rate of change of production for mode m of process i
$CAPEX^{segment}_{i,l}$	Capital cost for process i at the right hand side of segment l
$Cost^{discharge}_{a,j,h,t}$	Cost of discharging resource j in time t of period h at location a
$Cost^{land}_{a,i}$	Land cost for process i at location a
$Cost^{P-fix}_{i,m,h}$	Fixed operational cost for process i operating in mode m of period h
$Cost^{P-var}_{i,m,h}$	Variable operational cost for process i operating in mode m of period h
$Cost^{purchase}_{a,j,h,t}$	Cost of purchasing resource j in time t of period h at location a
$Cost^{S-fix}_j$	Fixed capital cost for storage of resource j
$Cost^{S-var}_j$	Variable capital cost for storage of resource j
$D_{a,j,h,t}$	Demand for resource j in time t of period h at location a
$D^{period}_{a,j,h}$	Demand for resource j for period h at location a
$D^{total}_{a,j}$	Aggregated demand for resource j of the annual operation
$Land^{max}_{a,j}$	Maximum land availability for resource j at location a
$Loss^S_{a,j}$	Fractional loss from storing resource j in period h

n_h	The weight of the representative period h in annual operation
ρ_j	Density parameter to convert standard material flow unit (kg/h) of resource j to a different unit

B.2 Parameter Information

Resource Data

Table B.1: Set of resources and the design and operational parameters for their storage

Resource	Description	CAP_j^{S-max}	$Cost_j^{S-fix}$	$Cost_j^{S-var}$
Solar	Direct normal irradiation (DNI) potential in W/m^2	-	-	-
Wind	Wind power potential in m/s	-	-	-
DC Power	Electrical power	-	-	-
AC Power	Local electrical power demand	-	-	-
Grid Backup	Power provided by the local electricity grid	-	-	-
Battery Charge	Power stored in battery	1e8	0	340,000

Solar and wind resource potentials given in W/m^2 and m/s, respectively are converted to their MW-equivalent resource availabilities using Eqs. B.1 and B.2:

$$B^{max}_{a,Solar,h,t} = (Solar\ DNI)(Land^{max}_{a,Solar}) \tag{B.1}$$

$$B^{max}_{a,Wind,h,t} = \frac{1}{2}\rho_{air}\pi(SA)^2(Wind\ Speed)^3(Land^{max}_{a,Wind})(10^{-6}) \tag{B.2}$$

Where $B^{max}_{a,Solar,h,t}$ and $B^{max}_{a,Solar,h,t}$ are solar and wind resource availability in MW, ρ_{air} is the density of air given in kg/m^3, SA is the swept area by the rotor blades (where rotor blade diameter is 108 m), and $Land^{max}_{a,Solar}$ & $Land^{max}_{a,Wind}$ are the maximum land available for solar and wind energy, respectively. The reference resource for each process is shown in bold in Table B.3.

Process Data

Table B.2: Set of processes

Process	Description
Wind Farm	Turbines that generate *DC Power* from *Wind*
Solar PV	Photovoltaic (PV) cells that produce *DC Power* from *Solar*
DC-AC Inverter	Inverter to convert *DC Power* to *AC Power*
Battery-Store	NaS battery to store *DC Power* as *Battery Charge*
Battery-PCS	NaS battery to produce *DC Power* from *Battery Charge*

Table B.3: Conversion and capacity parameters for the processes

Process	Basis	CAP_i^{P-max}	Inputs	Outputs
Wind Farm [200]	MW	200,000	*Wind* = -2.5	***DC Power = 1***
Solar PV [201]	MW	200,000	*Solar* = -4.44	***DC Power = 1***
DC-AC Inverter	MW	10,000	*DC Power* = -1.053	***AC Power = 1***
Battery-Store [21]	MW	100,000	***DC Power = -1***	*Battery Charge* = 1
Battery-PCS [21]	MW	100,000	*Battery Charge* = -1.176	***DC Power = 1***

Table B.4: Operational cost and rate change parameters for the processes

Process	Basis[1]	$CAP_{i,m}^{\Delta rate}$	$CAP_{i,m}^{util-min}$	$CAP_{i,m}^{util-max}$	$Cost_{i,m,h}^{P-fix}$	$Cost_{i,m,h}^{P-var}$
Wind Farm [200]	MW	1	0	1	0	14.6
Solar PV [201]	MW	1	0	1	0	5.3
DC-AC Inverter	MW	1	0	1	0	0
Battery-Storage [21]	MW	1	0	1	0	2.2
Battery-PCS [21]	MW	1	0	1	0	2.2

Table B.5: Capital investments cost parameters for the processes

		$CAP_{i,l}^{segment}$								$CAPEX_{i,l}^{segment}$ ($ MM)							
	Basis[1]	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
Wind Farm [200]	MW	100	200,000	-	-	-	-	-	-	169	388,000	-	-	-	-	-	-
Solar PV [201]	MW	100	200,000	-	-	-	-	-	-	111	222,000	-	-	-	-	-	-
DC-AC Inverter	MW	1	10,000	-	-	-	-	-	-	0.75	7,500	-	-	-	-	-	-
Battery-Storage [21]	MW	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Battery-PCS [21]	MW	1	100,000	-	-	-	-	-	-	0.65	65,000	-	-	-	-	-	-

B.3 Simultaneous Design and Scheduling MILP Model

Here, we present the modeling constraints for the simultaneous design and operation model. Please see the Section B.1 for the full list of sets, variables, and parameters. The mixed-integer linear programming (MILP) model consists of: (i) network design constraints for production and storage facilities, (ii) operating mode selection with ramp up/down constraints, (iii) resource balance constraints, (iv) continuity constraints, (v) investment and operational cost functions, and (vi) objective function.

Network Design Constraints

Constraints for network design, that allow for processes i and storage vessels for resources j to be selected or not, are shown in Eqs. B.3 and B.4:

$$CAP_{a,i}^{P} \leq CAP_{i}^{P-max} x_{a,i}^{P} \quad \forall a, i \tag{B.3}$$

$$CAP_{a,j}^{S} \leq CAP_{j}^{S-max} x_{a,j}^{S} \quad \forall a, j \tag{B.4}$$

Operating Mode Selection and Ramp Up/Down Constraints

Each process i can have multiple modes m, among which only one can be selected at a time t. Constraints for mode selection are shown in Eqs. B.5 to B.10:

$$\sum_{m \in Modes_{i,m}} y_{a,i,m,h,t} = x_{a,i}^{P} \quad \forall a, i, h \in Period_h, t \in Time_{h,t} \tag{B.5}$$

$$P_{a,i,h,t} = \sum_{m \in Modes_{i,m}} P^m_{a,i,m,h,t} \quad \forall a, i, h \in Period_h, t \in Time_{h,t} \tag{B.6}$$

$$P^m_{a,i,m,h,t} \geq CAP^{P-min}_{i,m} Cap^P_{a,i} \quad \forall i, m \in Modes_{i,m}, h, t \in Time_{h,t} \tag{B.7}$$

$$P^m_{a,i,m,h,t} \leq CAP^{P-max}_{i,m} Cap^P_{a,i} \quad \forall a, i, m \in Modes_{i,m}, h, t \in Time_{h,t} \tag{B.8}$$

$$P^m_{a,i,m,h,t} \geq CAP^{mode-min}_{i,m} y_{a,i,m,h,t} \quad \forall a, i, m \in Modes_{i,m}, h, t \in Time_{h,t} \tag{B.9}$$

$$P^m_{a,i,m,h,t} \leq CAP^{mode-max}_{i,m} y_{a,i,m,h,t} \quad \forall a, i, m \in Modes_{i,m}, h, t \in Time_{h,t} \tag{B.10}$$

Eqs. B.7 and B.8 set the lower and upper bounds of the production range of mode m for process i. During dynamic operation, a processing unit might change its throughput in the same fixed mode or switch to a different operating mode. Below are the equations used to represent mode switch. Eqs. B.11 and B.12 restrict the up-ramping and down-ramping rates of throughput from an unit by using a Big-M parameter to bound the change of $P^m_{i,m,h,t}$ between consecutive time periods $t-1$ and t.

$$P^m_{a,i,m,h,t} - P^m_{a,i,m,h,t-1} \geq -\Delta Rate^{max}_{i,m} - M(2 - y_{a,i,m,h,t} - y_{a,i,m,h,t-1}) \quad \forall a, i, m \in Modes_{i,m}, h, t \in Time_{h,t} \tag{B.11}$$

$$P^m_{a,i,m,h,t} - P^m_{a,i,m,h,t-1} \leq \Delta Rate^{max}_{i,m} + M(2 - y_{a,i,m,h,t} - y_{a,i,m,h,t-1}) \quad \forall a, i, m \in Modes_{i,m}, h, t \in Time_{h,t} \tag{B.12}$$

Eq. B.13 is used to designate mode switch. The binary variable $z_{i,m',m,h,t-1}$ is equal to 1 if process i switches from mode m to m' at time t of period h:

$$\sum_{m' \in Trans_{i,m,m'}} z_{a,i,m',m,h,t-1} - \sum_{m \in Trans_{i,m',m}} z_{a,i,m,m',h,t-1} = y_{a,i,m,h,t} - y_{a,i,m,h,t-1} \quad \forall a, i, m \in Modes_{i,m}, h, t \in Time_{h,t} \tag{B.13}$$

Eqs. B.14 and B.15 put restrictions on the minimum time required to switch between modes and the sequence of changes that needs to be followed if there is a mode switch, respectively.

$$y_{a,i,m',h,t} \geq \sum_{k=1}^{\theta_{i,m,m'}} z_{a,i,m,m',h,t-k} \quad \forall i,(m,m') \in TransModes_{i,m,m'}, a,h,t \in Time_{h,t} \tag{B.14}$$

$$z_{a,i,m,m',h,t-\overline{\theta}_{i,m,m',m''}} = z_{a,i,m',m'',h,t} \quad \forall a,i,(m,m',m'') \in Seq_{i,m,m',m''}, h,t \in Time_{h,t} \tag{B.15}$$

Resource Balance Constraints

The mass balance is written in terms of accounting for the inventory level for each resource in the network. Eq. B.16 is a total mass balance for resource j at any time t in period h.

$$\begin{aligned} Inv_{a,j,h,t} = (1 - Loss_{j,h}) Inv_{a,j,h,t-1} + \sum_{i} \sum_{m \in Modes_{i,m}} Conversion_{i,m,j,t} P^{m}_{a,i,m,h,t} \\ + B_{a,j,h,t} - S_{a,j,h,t} \\ \forall a,j,h,t \in Time_{h,t} \end{aligned} \tag{B.16}$$

$Conversion_{i,m,j,t}$ is a parameter that sets the input-output relationship for each process and is written with respect to the reference resource j for each mode m of process i. The full list of inputs and outputs to processes are given in Section B.2.

Eqs. B.17 and B.18 set upper bounds for the nameplate process and inventory storage capacities, respectively:

$$P_{a,i,h,t} \leq Cap^{P}_{a,i} \quad \forall a,i,h,t \in Time_{h,t} \tag{B.17}$$

$$Inv_{a,j,h,t} \leq Cap^{S}_{a,j} \quad \forall a,j,h,t \in Time_{h,t} \tag{B.18}$$

Supply and demand constraints are written for each resource j in forms of Eqs. B.19 to B.21:

$$B_{a,j,h,t} \leq B^{max}_{a,j,h,t} \quad \forall a,j,h,t \in Time_{h,t} \tag{B.19}$$

$$S_{a,j,h,t} \leq D_{a,j,h,t} \quad \forall a,j \in Demand_j, h, t \in Time_{h,t} \tag{B.20}$$

$$S_{a,j,h,t} = 0 \quad \forall a,j \in NoDischarge_j, h, t \in Time_{h,t} \tag{B.21}$$

Continuity Constraints

Eqs. B.22 to B.28 show the connection between subsequent periods, where n_h is the number of times a representative time period is cycled. The continuity condition represents that inventory needs to be accumulated over the course of a period and carried over to the next one. These constraints are adapted from Zhang et. al. [202].

$$\begin{aligned} y_{a,i,m,h,0} &= y_{a,i,m,h,|Time_{h,t}|} \\ \forall a,i,m &\in Modes_{i,m}, h \end{aligned} \tag{B.22}$$

$$\begin{aligned} z_{a,i,m,m',h,t} &= z_{a,i,m,m',h,t+|Time_{h,t}|} \\ \forall a,i,(m,m') &\in TransModes_{i,m,m'}, h, -\theta_i^{max} - 1 \leq t \leq -1 \end{aligned} \tag{B.23}$$

$$\begin{aligned} y_{a,i,m,h,|Time_{h,t}|} &= y_{a,i,m,h+1,0} \\ \forall a,i,m &\in Modes_{i,m}, h \in Period_h \setminus |Period_h| \end{aligned} \tag{B.24}$$

$$\begin{aligned} z_{a,i,m,m',h,t+|Time_{h,t}|} &= z_{a,i,m,m',h+1,t} \\ \forall a,i,(m,m') &\in TransModes_{i,m,m'}, h \in Period_h \setminus |Period_h|, -\theta_i^{max} - 1 \leq t \leq -1 \end{aligned} \tag{B.25}$$

$$\begin{aligned} Inv_{a,j,h}^{excess} &= Inv_{a,j,h,|Time_{h,t}|} - Inv_{a,j,h,0} \\ &\forall a,j,h \end{aligned} \tag{B.26}$$

$$\begin{aligned} Inv_{a,j,h,0} + n_h Inv_{a,j,h}^{excess} &= Inv_{a,j,h+1,0} \\ \forall a,j,h &\in Period_h \setminus |Period_h| \end{aligned} \tag{B.27}$$

$$Inv_{a,j,|Period_h|,0} + n_{|Period_h|} Inv_{a,j,|Period_h|}^{excess} = Inv_{a,j,1,0} \quad \forall a,j \tag{B.28}$$

Investment and Operational Cost Functions

Capital investment costs are approximated by piecewise linear functions of the processing plant capacities as shown in Eqs. B.29 to B.32:

$$Cap^{P}_{a,i} = \sum_{l \in PL_{i,l}} \left[\lambda_{a,i,j}(CAP^{segment}_{i,l-1} - CAP^{segment}_{i,l}) + CAP^{segment}_{i,l} w_{a,i,l}\right] \quad \forall a, i \tag{B.29}$$

$$Capex_{a,i} = \sum_{l \in PL_{i,l}} \left[\lambda_{a,i,j}(CAPEX^{segment}_{i,l-1} - CAPEX^{segment}_{i,l}) + CAPEX^{segment}_{i,l} w_{a,i,l}\right] \quad \forall a, i \tag{B.30}$$

$$\lambda_{a,i,l} \leq w_{a,i,l} \quad \forall a, (i,l) \in PL_{i,l} \tag{B.31}$$

$$\sum_{l \in PL_{i,l}} w_{a,i,l} = x^{P}_{a,i} \quad \forall i \tag{B.32}$$

Total capital expense is the sum of unit investment costs, storage costs, and land purchase costs as given in Eq. B.33:

$$\begin{aligned} Capex^{total}_{a} = &\sum_{i} \left[Capex_{a,i} + Cost^{land}_{a,i} Cap^{P}_{a,i}\right] \\ &+ \sum_{j} \left[Cost^{S-fixed}_{j} x^{S}_{a,j} + Cost^{S-var}_{j} Cap^{S}_{a,j}\right] \end{aligned} \tag{B.33}$$

Total operational expense as shown in Eq. B.34 is the sum of unit processing costs, resource purchase, and discharge costs:

$$\begin{aligned} Opex_{a} = \sum_{h} \sum_{t \in Time_{h,t}} n_{h} \Bigg[&\sum_{i} \sum_{m \in Modes_{i,m}} (Cost^{P-fixed}_{i,m,h} y_{a,i,m,h,t} + Cost^{P-var}_{i,m,h} P^{m}_{a,i,m,h,t}) \\ &+ \sum_{j} Cost^{Purchase}_{a,j,h,t} B_{a,j,h,t}/\rho_{j} + Cost^{Discharge}_{a,j,h,t} S_{a,j,h,t}/\rho_{j} \Bigg] \end{aligned} \tag{B.34}$$

Objective Function

The model is solved to minimize the annualized production cost of meeting the power demand. This is also called the levelized cost of electricity (LCOE). For this purpose, total capital investment is annualized using 8% annual discount rate. Adding the annualized capital cost and operational cost together gives the objective function value $TotalCost$ as shown in Eq. B.35:

$$TotalCost = \sum_{a} \left[(0.08) Capex_a^{total} + Opex_a \right] \tag{B.35}$$

The resulting optimization problem is:

$$\begin{aligned} \min \quad & TotalCost \\ \text{s.t.} \quad & Eqs. \quad B.3 - B.34 \end{aligned} \tag{B.36}$$

Eq. B.36 is the LB problem in the decomposition algorithm.

For the UB problem, the representative periods are replaced by the original time horizon, and the design variables ($w_{a,i,l}$, $x^P_{a,i}$, $x^S_{a,j}$,$Cap^P_{a,i}$, $Cap^S_{a,j}$, $Capex_{a,i}$ $\lambda_{a,i,l}$) are fixed. The original time horizon is now represented by using only one "representative period" with 8760 hours. Grid backup ($B_{a,GridBackup,h,t} \geq 0$) is allowed for the UB problem to maintain feasibility, and the objective is to minimize the amount of this backup purchased. The resulting optimization problem is:

$$\begin{aligned} \min \quad & \sum_{(a,h,t)} B_{a,GridBackup,h,t} \\ \text{s.t.} \quad & Eqs. \quad B.3 - B.34 \end{aligned} \tag{B.37}$$

Eq. B.37 is the UB problem in the decomposition algorithm.

APPENDIX C

DATASETS UTILIZED FOR HYPERPARAMETER OPTIMIZATION

C.1 Ammonia Reactor Data

Table C.1: Collected dataset of ammonia reactor information from various sources

Fold #	Temp.	Pres.	x_{H_2}	x_{N_2}	x_{NH_3}	x_{Inert}	$\frac{x_{H_2}}{x_{N_2}}$	y_X	Ref.
1	438	93.0	0.620	0.280	0.050	0.050	2.21	0.197	[203]
1	402	85.5	0.605	0.275	0.038	0.082	2.20	0.143	[204]
1	455	229.0	0.670	0.222	0.028	0.081	3.02	0.305	[205]
1	438	283.0	0.650	0.219	0.052	0.079	2.97	0.287	[205]
1	448	240.0	0.658	0.219	0.018	0.105	3.00	0.338	[206]
1	455	229.0	0.670	0.222	0.028	0.080	3.02	0.305	[207]
2	395	85.6	0.660	0.220	0.040	0.080	3.00	0.201	[203]
2	370	199.0	0.653	0.217	0.010	0.120	3.01	0.287	[208]
2	440	210.0	0.621	0.206	0.032	0.141	3.01	0.311	[205]
2	460	272.0	0.658	0.212	0.030	0.100	3.10	0.317	[205]
2	460	177.0	0.682	0.227	0.043	0.048	3.00	0.246	[209]
2	459.5	146.0	0.672	0.224	0.020	0.085	3.00	0.246	[210]
3	432	104.0	0.650	0.216	0.037	0.097	3.01	0.219	[211]
3	450	127.0	0.667	0.230	0.022	0.081	2.90	0.207	[212]
3	469	185.0	0.612	0.235	0.042	0.111	2.60	0.269	[213]
3	472	265.8	0.634	0.211	0.035	0.120	3.00	0.338	[214]
3	419	105.0	0.656	0.222	0.038	0.084	2.96	0.224	[215]
3	445	223.0	0.628	0.209	0.035	0.128	3.00	0.333	[216]
4	401	171.0	0.650	0.240	0.020	0.090	2.71	0.368	[217]
4	473	128.0	0.625	0.215	0.023	0.137	2.91	0.214	[218]
4	480	200.0	0.639	0.213	0.013	0.135	3.00	0.212	[212]
4	443	362.0	0.595	0.188	0.030	0.187	3.16	0.370	[205]
4	460.3	144.0	0.672	0.224	0.020	0.085	3.00	0.210	[210]
4	350	134.0	0.736	0.254	0.000	0.010	2.90	0.430	[219]
5	354	171.0	0.659	0.210	0.018	0.113	3.14	0.313	[217]
5	440	179.0	0.651	0.196	0.032	0.121	3.32	0.295	[205]
5	377	100.0	0.643	0.228	0.015	0.115	2.83	0.278	Haldor-Topsøe
5	453	88.6	0.622	0.309	0.020	0.050	2.01	0.331	[220]
5	455	88.6	0.627	0.309	0.019	0.045	2.03	0.332	[220]

C.2 Breast Cancer Data

Table C.2: Fold identification of breast cancer data [180]

Fold #	Patients
1	3, 10, 13, 22, 24, 42, 46, 54, 60, 65, 66, 68, 71, 72, 73, 74, 79, 83, 86, 91, 92, 99, 110, 114
2	5, 8, 11, 16, 20, 31, 32, 37, 38, 44, 48, 50, 53, 62, 63, 67, 75, 89, 96, 97, 104, 109, 112
3	1, 6, 17, 23, 27, 33, 35, 36, 45, 52, 55, 59, 61, 64, 69, 70, 77, 93, 102, 105, 106, 113, 116
4	2, 4, 7, 9, 14, 15, 18, 21, 25, 26, 28, 47, 56, 76, 78, 80, 82, 85, 88, 95, 101, 111, 115
5	12, 19, 29, 30, 34, 39, 40, 41, 43, 49, 51, 57, 58, 81, 84, 87, 90, 94, 98, 100, 103, 107, 108

APPENDIX D

LIST OF PUBLICATIONS AND PRESENTATIONS

At the time of writing, the publications, conference proceedings, and presentations produced from my graduate studies are listed below.

D.1 Journal Publications

- **Tso, WW**; Niziolek, AM.; Onel, O; Demirhan, CD; Floudas, CA.; Pistikopoulos, EN. Enhancing Natural Gas-to-Liquids (GTL) Processes Through Chemical Looping: Process Synthesis & Global Optimization. *Computers & Chemical Engineering*, 113 (2018): 222-239.

- **Tso, WW**; Demirhan, CD; Floudas, CA; Pistikopoulos, EN. Multi-Scale Energy Systems Engineering for Optimal Natural Gas Conversion. *Catalysis Today*. In press.

- **Tso, WW**; Demirhan, CD; Heuberger, CF; Powell, JB; Pistikopoulos, EN. A Hierarchical Clustering Decomposition Algorithm for Optimizing Renewable Power Systems with Storage. *Applied Energy*. In review.

- **Tso, WW**; Burnak, B; Pistikopoulos, EN. HY-POP: Hyperparameter Optimization of Machine Learning Models Through Parametric Programming. *Computers & Chemical Engineering*. In revision.

- Demirhan, CD; **Tso, WW**; Powell, JB; Pistikopoulos, EN. Sustainable Ammonia Production Through Process Synthesis & Global Optimization. *AIChE Journal*, 65.7 (2019): e16498.

- Demirhan, CD; **Tso, WW**; Ogumerem, GS; Pistikopoulos, EN. Energy Systems Engineering - A Guided Tour. *BMC Chemical Engineering*, 1.1 (2019): 1-19.

- Demirhan, CD; **Tso, WW**; Powell, JB; Heuberger, CF; Pistikopoulos, EN. A Multi-scale Energy Systems Engineering Approach for Renewable Power Generation and Storage Optimization. *Industrial & Engineering Chemistry Research*. Accepted.

- Demirhan, CD.; Boukouvala, F.; Kim, K.; Song, H.; **Tso, WW**; Floudas, CA.; Pistikopoulos, EN. An Integrated Data-Driven Modeling & Global Optimization Approach for Multi-Period Nonlinear Production Planning Problems. *Computers & Chemical Engineering*. In review.

D.2 Conference Proceedings

- **Tso, WW**; Demirhan, CD; Powell, JB.; Pistikopoulos, EN. Toward Optimal Synthesis of Renewable Ammonia and Methanol Production (RAMP). *Computer Aided Chemical Engineering*, 44 (2018): 1705-1710.

- **Tso, WW**; Demirhan, CD; Lee, S; Song, H; Powell, JB; Pistikopoulos, EN. Energy Carrier Supply Chain Optimization: A Texas Case Study. *Computer Aided Chemical Engineering*, 47 (2019): 1-6.

- Ogumerem, GS; **Tso, WW**; Demirhan, CD; Lee, S; Song, H; Pistikopoulos, EN. Toward Optimization of Hydrogen, Ammonia, and Methanol Supply Chains. *IFAC-PapersOnLine*, 52.1 (2019): 844-849.

D.3 Conference Presentations

- Natural Gas to Liquid Transportation Fuels Utilizing Chemical Looping for Syngas Generation: Process Synthesis & Global Optimization
  - TAMU Energy Conference '16, AIChE Meeting '16, TAMU ChESGA Symposium '17

- Advancing the Production of Olefins and Aromatics from Natural Gas via Methanol: Chemical Looping for Syngas Generation
  - TAMU Energy Conference '17, AIChE Meeting '17, TAMU ChESGA Symposium '18

- Toward the Optimal Synthesis of Renewable Ammonia and Methanol Processes (RAMP)
  - PSE Conference '18
- Toward Supply Chain Optimization of Renewable Energy Carriers
  - TAMU Energy Conference '18, AIChE Meeting '18,
    TAMU ChESGA Symposium '19, FOCAPD '19
- A Clustering Decomposition Algorithm for Energy Storage Design & Operation
  - AIChE Meeting '19
- Machine Learning Through a Parametric Programming Lens
  - AIChE Meeting '19