6-24-2019

Adding renewables to the grid: Effects of Storage and Stochastic Forecasting

Naga Srujana Goteti
nsg3545@rit.edu

Adding renewables to the grid: Effects of Storage and Stochastic Forecasting

by

Naga Srujana Goteti

A DISSERTATION

Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Sustainability

Department of Sustainability

Golisano Institute for Sustainability

Rochester Institute of Technology

June 24, 2019

Certificate of approval

Golisano Institute for Sustainability
Rochester Institute of Technology
Rochester, New York

Ph.D. DEGREE DISSERTATION

The Ph.D. Degree Dissertation of Naga Srujana Goteti has been examined and approved by the dissertation committee as satisfactory for the dissertation requirement for the Ph.D. degree in Sustainability.

____
Dr. Thomas Trabold, Director of Ph.D. program and Department Chair

____
Dr. Clark Hochgraf, Dissertation External Chairperson

____
Dr. Eric Williams, Co-Advisor & Committee Chairperson

____
Dr. Eric Hittinger, Co-Advisor, Committee Member

____
Dr. Gabrielle Gaustad, Committee Member

Date __________________________

Abstract

The electricity sector contributes to a quarter of global greenhouse emissions, and managing its evolution is a critical sustainability challenge. The context for the development and operation of electricity grids has dramatically changed in recent years. Wind and solar power have become much less expensive. Lower costs combined with increased policy action to address carbon emissions is leading to substantial shares of electricity generated by intermittent renewables. Maintaining a stable electricity supply with intermittency is a critical challenge; storage and natural gas are possible solutions. While policymakers promote storage as green grid technology, low-cost natural gas from hydrofracturing extraction raises the economic hurdle for storage.

Researchers have developed complicated energy system models to help plan grids in the face of the above trends. The research in this dissertation introduces new modeling features that affect the economic and environmental outcomes of the adoption of renewable and storage technologies. First, prior models that explore the future build-out of electricity grids are nearly always deterministic, i.e., they assume that decision-makers have perfect information. Here a stochastic optimization grid expansion model is developed that presumes that expected future fluctuations, e.g. in fuel prices, influence build-out decisions. This stochastic model thus includes uncertainty and risk as core elements: Grid build-out depends on the distribution of system costs. A genetic algorithm with Monte-Carlo simulation is used for co-optimization using two objective functions: "risk-neutral," which optimizes to minimize average system cost and "risk-averse," which optimizes to minimize average of the top 5% of costs (also called 95% Conditional Value at Risk (CVaR)). This model is tested for the US Midwest regional grid. The results show that the risk-averse scenario does not increase mean system costs but adds significantly more wind. These results corroborate prior work showing that electricity system costs can be surprisingly inelastic to renewable adoption and further introduces quantification of how increased renewables lowers cost risk.

Second, the economic and environmental performance of storage is complicated by how its introduction affects the operation of both renewable and fossil plants. In this dissertation, a

model is developed that accounts for how storage operation would affect prices on the grid and in turn, the operational schedule that yields optimal revenue. Results from modeling the US Midwest region shows that this treatment of storage as a "price maker" affects results. The model indicates that storage increases carbon emissions when it enables a high emissions generator, such as a coal plant, to substitute for a cleaner plant, such as natural gas. In this case, low cost; efficient natural gas generation is relatively better than coal to realize emissions reductions with storage under economic arbitrage until renewables dominate the grid mix.

Third, the operational strategies of energy storage alter the generation and profits of the other electricity generation systems. The operational effects of storage on the change in generation is investigated for all the eGRID subregions across the US based on actual historical electricity prices and the generation mix for the year 2016. Results show that storage increases the coal generation and affects the natural gas generation in the west – except in California and the Midwest regions of the US; and increases the generation of the natural gas in the eastern US regions. California, upstate New York and New England regions show an exception with an increase in natural gas generation and decrease in coal generation. The model also investigates the operational effects of storage on the profits of other generating units in California, Midwest and New York regions. Profits of other generating units are significantly affected when large capacities of storage operate as price-makers. Coal has a small increase in profits by 2% and all the other fuels continue to see a decline in profits in New York and the Midwest regions. The decrease in profits of the other generating units is because of the offset/retirements of the peaker natural gas plants that set the electricity prices. On the other hand, in California, the profits for renewables increase from the increase in electricity clearing prices set by the natural gas combined cycle plants to meet the additional demand from the storage charging.

Acknowledgements

First and foremost, I am indebted to my Ph.D. advisors Dr. Eric Hittinger and Dr. Eric Williams. Thank you for your continuous support, valuable guidance, and constant encouragement all throughout my Ph.D. journey. On the academic level, you taught me the fundamentals of conducting scientific research and effective communication of the results. Outside the academic world, you helped me with the travel and networking during the conferences, secure fellowship to work at National Renewable Energy Laboratory and participate in other outreach activities. Thanks, Dr. Erics, as I fondly address you both for believing in me and supporting me all throughout my time here. I will look up to you and emulate your research principles as I continue to grow.

My sincere appreciation to Dr. Gabrielle Gaustad for readily agreeing to be part of my committee. Your class on advanced optimization techniques is a backbone to my research work. Thank you for always being there, when I wanted to ask you for career advice, and recommendations. Your valuable insights, constructive feedback, and positivity are treasured.

Thank you to my husband Prathmesh Savargaonkar for pushing me to pursue Ph.D., when I was not sure if this was the path I wanted to take. It has been a roller coaster journey for both of us. I am so proud of your positive attitude towards life, and without you, I wouldn't have reached this point in my life. Thank you for everything and can't wait to start a new journey with you after my Ph.D.

I am eternally grateful to my late grandfather Mr. Goteti Veereshwar Rao who always wanted me to take up an advanced degree in science. Thank you to my family, in-laws, and my second family in the United States for your kind words and emotional support. I wish to especially thank my father Goteti V. Prasad, my mom Goteti S. Sharada, my sisters Manisha and Manjari for always being there during my high and low days. I wish to also thank my father-in-law Dr. Vijay Savargaonkar, mother-in-law Dr. Ujwala Savargaonkar, and brothers-in-law Mayuresh and Kalyan for your encouragement and positivity. I immensely cherish the visits of my in-laws all the way from India and the food my mother-in-law cooked for me to boost my morale and help me focus on my work.

Dr. Clark Hochgraf, I thank you for readily agreeing to chair my dissertation defense and ensuring that the process is conducted fairly and smoothly.

I acknowledge my funding sources for my research during Ph.D and for my internship at National Renewable Energy Laboratory: The National Science Foundation and the Golisano Institute for Sustainability (GIS). I also would like to express my gratitude to Dr. Thomas Trabold, and Dr. Callie Babbit for their pep talks, career advice, and inspiring me to think as a 'sustainabologist'. A massive shout out to GIS staff- Lisa Dammeyer and Donna Podeszek for all the administrative help and the follow-ups. Lisa Dammeyer, thank you so much for your patience and lending ears to my occasional banters.

Thanks to my internship opportunity at NREL, I acquired immense technical skills much needed in the energy modeling world from my mentor- Dr. Brady Stoll, and manager- Daniel Steinberg.

Finally, all this would not be possible without the support from my friends in GIS and at NREL. A special thanks to Wendy Harmon and Vernon Harmon, William Armington, Ashok Sekar, Sherwyn Millette, Aleena Humayun, Haleh Moghaddasi, Vineet Nair, Neha Chengappa, Matthew Irish, Sarah Awara, Elisabeth Mcclure, Parangat Bhaskar, and many many more.

You will all occupy a special place in my heart for everything you have done, my gratitude knows no bounds.

Table of Contents

Table of Figures

Table of Tables

Table of Appendix Figures

Table of Appendix Tables

GLOSSARY OF TERMS

Deregulation	Wholesale markets for trading electricity generation.
Discharge	Releasing energy/electricity into the system from storage.
eGRID	The Emissions & Generation Resource Integrated Database by EPA.
ISO	Independent System Operators, coordinates and monitors the electricity grid such that supply meets the demand.
Marginal generator	Plant used to meet the last unit of demand.
MW	Unit of power output, e.g. nameplate capacities of the power plants.
MWh	Unit of energy output.
NGCC	Natural Gas Combined Cycle.
NGCT	Natural Gas Combustion Turbine.
Peaker plant	Plants used during the peak demand periods.
Plant Retirement	Plants not in active operation.
Price Maker	System's operation affects the market prices.
Price Taker	System exogenous to the market prices and its operation does not affect the prices.
Ramping	Rate at which the power plant's output changes.
Storage Charge	State of using energy for storing.

Chapter 1: Introduction

1.1 Background

Managing the evolution of the electricity grid is a critical sustainability challenge. As of 2017, electricity contributed to the 28% of total greenhouse gas emissions in the U.S. [1] and to a quarter of global GHG emissions [2]. Electricity production is also economically significant, generating $380 billion in revenue in 2016 in the U.S. [3] and is a backbone of many other industries. Not only is electricity a major carrier of energy, its role in the energy system continues to expand. Electric vehicles (EV) could be on a trajectory to enable electricity to replace much of the demand for liquid fuels, prompting expansion and transformation of the electrical grid.

Like any other infrastructure, the electric grid is long-lived. An average age of the coal power plant, accounting for 30% of annual generation in the U.S. is 40 years [4]. Long life exacerbates lock-in effects: Capital investments, once made, last for decades. Sunk investments crowd out the potential to adopt new technology and indeed, some elements of the grid, such as photovoltaic panels, batteries and wind energy, are undergoing rapid technological improvement.

Deregulated electricity markets

In addition to the long-life of the electricity infrastructure, the deregulation of the electricity industry has shifted the capital availability and risk preferences of generation companies. As of 2018, eighteen states in the U.S. participate in the deregulated electricity markets [5]. Traditionally, in a vertical integrated structure/regulated markets, utilities controlled the transmission, distribution, and generation of the electricity to the consumers. The utilities conformed to the regulations set by the governments and were assured a guaranteed return on the investments. This enabled them to participate in an almost risk-free environment and procure finance for capital-intensive power plants [6]. Whereas, in the de-regulated markets, electricity is a tradeable commodity, creating competition. Competition lowers the prices and generators should be able to produce cheaper electricity while changing the output to meet the real-time demand. For example, in a wholesale electricity market, generators bid a price at which they can supply a specific number of megawatt-hours of electricity. Independent System Operators (ISO) like Midcontinent ISO in the Midwest region, or New York ISO in the NYISO, clear the market by selecting the generators with lowest bid till the supply meets the demand. The price of the last resource to offer such that the demand meets the supply decides the wholesale price of the power. In such a market with volatile electricity prices, often coal power plants cannot match with the almost zero marginal prices of the renewables and cannot change output like gas power plants in response to demand.

In recent years, a combination of shale revolution and reduction in the wind/solar costs changed the dynamics of electricity markets, pushing flexible natural gas generation and renewables to the forefront over the usage of coal- with higher emission rates. As of 2015, natural gas generation surpassed coal and contributed to 33% of the total generation, followed by coal at 32%, nuclear at 19%, non-hydro renewables at 8%, and hydro at 6% [7]. Also, realizing the economic and environmental benefits of the renewables, states are revising and introducing Renewable Portfolio Standards (RPS) to achieve a certain percentage of the renewable mix in the overall electricity generation. As of today, twenty-nine states in the U.S. have RPS, and eight states have set some form of renewable energy goals [8].

One of the significant limitations of the wind and solar is that they are variable and do not have a firm (constant) output generation. Natural gas complements the variable generation from the renewables and is easy to turn on and off (ramping), based on the availability of renewable energy. Thus, as the renewables grow in the system, natural gas could become even more reliable as a cheap source of energy during the intermittent hours when the renewable resources are unavailable. Also, the clearing prices set by the gas in the current de-regulated markets drives the revenues of the renewables, further bolstering a stable symbiotic relationship between the gas and the renewables [9].

Energy Information Administration (EIA) projects that the gas will be the leading source of electricity supply and the CO_2 emissions will remain unchanged until at least 2050 [10]. The current average age of natural gas power plants is 25 years, and further investments in natural gas will lock-in the infrastructure for a substantial period of time into the future.

Energy storage for renewables

On the other hand, energy storage is gaining traction, too, notably amongst policymakers to support the variable renewable generation. It stores energy when the wind blows and the sun shines and discharges into the grid when there is a demand. Storage can be a chemical battery, flywheel, or pumped hydro systems [11]. While the cost of energy storage, especially Lithium Ion batteries are expensive than a traditional gas plant, there has been a steep drop in the recent years from $800/MWh in 2013 to $200/MWh in 2018 [12]. It is a 75% drop in the costs and can soon compete with the natural gas whose Levelized Cost of Energy (LCoE) is between $30-60/MWh as of 2018 [10]. Since the storage overcomes the limitation of producing and consuming electricity in the real-time, policymakers are increasingly keen on promoting utility-scale and distributed storage systems enabling the integration of the renewables. As of 2018, some of the major states that passed energy storage target mandates are California [13], Massachusetts [14], Maryland [15], New York [16], and New Jersey [17].

Current challenges

Both natural gas and storage are catalysts for increasing renewable penetration, or are they? Annual Energy Outlook by EIA shows that gas prices are historically low as of 2019 and could continue to remain the same at $3/MMBtu till 2050, in a high oil and gas resources scenario. The same publication also shows that the gas will continue to rule the electricity generation mix by contributing to 40% of the generation supply by 2050 [10]. The big challenge is once these capital investments are made based on the signal of low gas prices, the plant could keep running for a very long time, even with cheaper renewables in the system. Most of the utilities [18] and even government agencies like EIA [19] run financial models, or forecast models for certain future scenarios. These methods undermine the risks of higher costs and future uncertainties, and the capital investments once made are sunk. In a world of volatile fuel prices, changing demand, cheaper renewables and deregulated markets: shifting the capital availability and risks, it is imperative to understand the uncertainties before modeling future grid infrastructure on certain sets of inputs. Especially if the results from these models include carbon-intensive technologies like natural gas, the bridge to renewables cannot be crossed for any time soon.

The challenge with the storage systems is that it is promoted by policy makers to solve the intermittency issues of renewable energy. But, storage in the US is rarely used to prevent curtailment of renewable energy. 88% of the total storage capacity in the US operates for profit maximization in an arbitrage scenario [11]. Deregulated grids feature generators (and consequently storage) as profit-maximizing agents. A profit-maximizing bulk energy storage system charges during low price/low demand periods and discharges during high price/high demand periods, regardless of the type of generation being used. The effect that this economic dispatch of storage has on grid emissions depends upon generation mix, dispatch order, demand, and storage round-trip efficiency [20]. Thus, the storage may or may not benefit the renewables, depending on the grids it operates, and a comprehensive analysis must be performed for different grid types before incentivizing them.

1.2 Problem Statement, research questions and novel contributions

Current energy models do not consider the future weather and fuel price uncertainties, risk preferences in the deregulated energy markets, and the current economic operation of the storage systems. The results from these models could lead to lock-in investments in natural gas for years to come or significant investments in bulk storage. While both the technologies complement the variable nature of the renewables, does the assumption of promoting storage, or results from deterministic grid expansion models help in the growth of the renewables? The research questions develop to answer these challenges are:

(1) **Does stochastic forecasting of the future grid for different risk preferences of the market enable more renewables over cheaper natural gas?**

In chapter 2, this dissertation introduces a novel model for the future build-out of grid infrastructure, accounting for future uncertainties, risk preferences in the de-regulated markets, and evolving technologies. The build-out decisions are made by minimizing total cost of service constrained by the need to meet current and future load. Uncertainty in fuel prices and load demand is treated through Monte Carlo analysis and the grid build-outs are analyzed for different risk preferences of the market. This analysis is performed in the Midwest Region of the US, which has the largest concentration of coal in the generation mix [4]. Application of the model on a grid heavily dependent on the fossil fuels yields insights into trends towards a more sustainable electrical grid.

(2) **Can bulk energy storage compliment renewables and enable environmental benefits to the grid?**

In chapter 3, storage model is built to evaluate the carbon implications of storage on the current electric grid and grids, and with expanded wind and solar energy. In the next few years, many states around US are keen to implement policy incentives/mandates for encouraging large storage systems on the electricity grids, hoping that it will support renewable growth and decrease emissions. Thus, chapter 3 explores how does storage affect grid emissions with differing amounts of added renewables, and differing natural gas prices? For this case, the New York ISO (NYISO) and Midcontinent ISO (MISO) regions are modeled. The choice of these two case studies allows us to contrast between grids not dependent and dependent on coal. The two grids studied here are representative of many systems in the US and around the world.

(3) **Can bulk energy storage enable profit benefits to the renewables?**

In Chapter 4, the model examines if adding storage benefits the profits and generation of the renewable energy power plant operators, especially as compared to cheaper natural gas in the system. This study is conducted in two parts. In the first part, the study investigates how storage affects the likely generation from other fuels based on the actual electricity prices and the fuel mix across 22 eGRID regions of the U.S. In this, case the storage does not affect the market prices. Though it captures the dynamics of the actual operation, it does not show the change in profit for other generators as it is a marginally small capacity. To answer the change in price/profit of other generators, a simulated dispatch model with larger storage capacities from 3GW-10GW is used in the second part of the chapter. In this case, storage distorts the prevailing prices. Therefore, in the first part the model captures the dynamics of the actual operation, and in the second part, it captures the changes in the profit when storage is no more a marginal operator. For this case, the New York ISO, Midcontinent ISO (MISO) and California ISO (CAISO) regions are modeled. These choices allow us to compare across a spectrum of low, medium, and high renewable energy penetration into the grid.

Table 1. Overview of knowledge gaps and the contribution of the dissertation

Knowledge Gaps	Contribution
Impacts of future uncertainty and risk preferences in the deregulated markets on the integration of renewables, Effects of adding storage on the integration of renewables	Analytical models to the decision makers on the environmental, and economical implications of including assessment of storage operation and future uncertainties.

Chapter 2: An alternative structure for integration of uncertainty and risk aversion into capacity expansion models

Abstract

Current capacity expansion models forecast the grid by deterministic optimization to minimize total system cost. Uncertainty analysis is ex-post, via running scenarios varying input parameters. A capacity expansion model with a built-in uncertainty would enable Monte Carlo analysis and consideration of alternative objective functions, e.g. accounting for cost risk. This paper introduces a proof-of-concept stochastic model that includes uncertainty and risk as core elements. Grid build-out now depends on a distribution of system costs; a genetic algorithm is used for co-optimization. Two objective functions are considered: "risk-neutral", which optimizes to minimize average system cost and "risk-averse", which optimizes to minimize average of the top 5% of costs (also called 95% Conditional Value at Risk (CVaR)). This study implements the model for the U.S. Midwest region, accounting for distributions in future electricity demand and fuel prices. Curiously, the risk-averse scenario does not increase mean system cost but adds significantly more wind (~ 20GW) and solar capacity (~15 GW) by 2050 compared to the risk-neutral objective. These results corroborate prior work showing that electricity system costs can be surprisingly inelastic to renewable adoption and adds quantification of how increased renewables lowers cost risk.

2.1 Introduction and Literature Review

Electricity generation is a major contributor to climate change, accounting for 25% of global carbon emissions [2]. Electricity is also economically significant, e.g. generating $380 billion in revenue in 2016 in the U.S. and affecting the profitability of many industrial sectors [3]. Efforts are underway around the globe to decarbonize the grid. Federal, state and provincial governments grant tax credits to select technologies, set targets for renewable energy adoption (Renewable Portfolio Standards (RPS)[21]) and implement carbon taxes [22].

The development of decarbonization policies and future energy trends is informed by a variety of energy system models with differing scope and spatial scales. One example at the global scale is the World Energy Model (WEM) by International Energy Agency [23]. It provides medium-long term energy projections on energy consumption, energy transformation, and energy supply levels across the globe, published annually in the World Energy Outlook. Future uncertainties are treated through three different scenario analyses- new policies scenario, current policies scenario and the sustainable development scenario. Another energy system model at the global scale is the World Energy Projection System (WEPS) Model by Energy Information Administration (EIA)[24]. Similar to WEM, they also model the demand, transformation, and supply at the region scale across sixteen different regions globally. The future scenarios in this model are developed based on the economic growth trajectories in the different regions. Modeling efforts in the U.S. are mostly centered at government agencies and national laboratories. The National Energy Modelling System (NEMS), coordinated by EIA, is a complex multi-module model that simulates the entire U.S. energy system, including electricity infrastructure, used as the basis for the Annual Energy Outlook published annually. The NEMS is an economy-wide model that also includes the future supply of natural gas, coal, and oil whose data is used by the other popular models for the future price and load trajectories. Another model is The Regional Energy Deployment System Model (ReEDS) by National Renewable Energy Laboratory (NREL), which is a capacity expansion model extensively focusing on the future scenarios of the U.S. power system, taking into account the technology innovation impacts, and policy scenarios [25]. The capacity expansion of the power system in the above models is optimized such that the generators and infrastructure are built to minimize the total system cost to meet future demand. The usual capacity expansion model is deterministic: The output is a single set of generators and infrastructure that minimizes discounted system costs when faced with a set of deterministic inputs [23, 24, 26].

What is the purpose of grid capacity expansion models? One might first think these models are intended to give reasonable forecasts of the future grid. A capacity expansion model may turn out to be good forecast, but I argue this is not the main purpose for which they are constructed. If energy system models were intended to good forecasts, this would imply that, as with other forecasts such as weather, model construction would involve retrospective analysis to determine what model and data most accurately "forecast" the past. This is not done in any systematic way in building energy models. There is a small literature that notes dramatic differences between retrospective forecasts and actual evolution of energy systems [27], but this is not part of a larger modeling effort to incorporate model structure into retrospective evaluation of forecasting.

The real purpose of a capacity expansion model is to construct a future in which society sensibly builds out the future grid. "Sensibly build" usually is defined as minimal total system cost, though often with constraints built in for other societal objectives such as renewable energy adoption targets or a carbon tax. It may be that the actors involved, and their decision-making processes result in a grid reasonably reflecting minimum total cost to society. In spite of this, energy system modelers can argue that there is value in knowing what choices are good for society aside from issues such as market failures.

Uncertainty, however, significantly complicates what it means for society to sensibly build an electricity grid. The usual deterministic optimization used in capacity expansion assumes decisions are made assuming perfect information, i.e. parameter values are fixed and there is one grid-build out that minimizes system costs. Many parameters driving electricity system profitability are, however, highly uncertain, including future fuel prices, technology prices, and policies. A choice that minimizes system costs for baseline values of input parameters may in fact, incur high risk of cost increases. Society knows that drivers of energy systems are uncertain and rightly ought to be concerned about lowering risk. To give an example from the sphere of personal decision making, many consumers prefer new cars over used ones because they are sensibly concerned with the costs if something goes wrong, i.e. the car breaking down.

Decision-making that accounts for uncertainty is mathematically formulated as stochastic optimization. Stochastic optimization treats decision makers as possessing knowledge of uncertainty. The goal is to find a grid build-out that which optimizes an expected (but not certain) outcome. Following the idea that sensibly building the energy system should account for risk, capacity expansion models should thus use stochastic optimization.

The need for stochastic capacity expansion models has been recognized by prior analysts. The Switch energy system model developed by Fripp [28] uses stochastic linear optimization to minimize the total system cost. Stochastic linear optimization recasts a fundamentally non-linear problem (a decision as a function of a distribution) as a linear problem by using discrete values for alternative parameter values and creating new optimization variables according to the probabilities of alternate parameter values. Switch takes into account the uncertainty in the renewable power supply and future demand. The model co-optimizes the investment and operational decisions under different scenarios of weather conditions. These weather conditions are generated through sample demand patterns, wind, solar and hydro availability chosen from a sample historical date.

The Tools for Energy Model Optimization and Analysis (TEMOA) model [29] also uses similar stochastic linear programming [30–33] for optimization with uncertainty. In this method, uncertain future outcomes are encoded as possible scenarios in an event tree with an assigned likelihood of occurrence. For example: in the sample region of their study using TEMOA, each combination of high, medium, and low growth rates of coal, oil, and gas were divided into 9 branches in an event tree and were assigned an equal probability of 1/9 to each branch. These branches with their assigned probabilities were solved through linear optimization [29, 32]

Stochastic linear optimization has two major limitations in its application to capacity expansion models. The first is the 'curse of dimensionality' [34], i.e. the need to define additional optimization variables for each new uncertain variable leads to an event tree increasing exponentially with the number of uncertain parameters. As a result, the current models limit the number of scenario branches/uncertainties from a tree to eight or have to use high-performance computers [35]. The second limitation is lack of flexibility in exploring different ways society might account for risk. For example, society might choose to aim for minimal mean system cost (from a distribution) or show aversion to high cost scenarios. While these different optimization objectives can in principle be treated with stochastic optimization, it is not a natural framework to do so.

This study develops a stochastic capacity expansion model that does not suffer from the curse of dimensionality (i.e. more uncertain parameters can be treated) and is flexible in allowing different optimization objectives that manage risk. To achieve this, this study pursues a stochastic non-linear approach, note this is distinct from prior efforts using stochastic linear optimization. Given that there is no general analytic solution to non-linear optimization, its use implies numerical approach as opposed to true optimization. There are many possible numerical techniques for nonlinear optimization, this model uses a genetic algorithm because it is suitable for global optimization problems with complex fitness landscapes [36, 37].To summarize our approach, given an initial test build-out model computes a distribution of system costs based on uncertain input parameters. I select an optimization objective, here the two explored are distribution mean and Conditional Value at Risk' (CVaR) 95%, i.e. the average of top 5% of the distribution of system costs. From the initial trial, the genetic algorithm generates a sequence of build-outs designed to converge towards the optimization objective. The final result of a run is the build-out with smallest value of the optimization objective. Given the expected variability in numerical approaches, I follow the common practice of running the algorithm 10 times. The resulting distribution of grid outcomes reflects the uncertainty in numerical optimization.

For this case study, this model covers the Midwest region of the U.S., accounting for distributions in two uncertain variables: electricity demand and natural gas prices. Uncertainty in electricity demand is treated by using historical hourly fluctuations in daily to model future variability. The uncertainty in natural gas prices is treated with a mean reversion model, with historical data informing the mean and fluctuations. Other driving variables such as technology cost are deterministic. Build-out decisions are made by minimizing the total cost of service over a 20-year horizon, constrained by the need to meet current and future load.

While there are uncertain variables other than demand and natural gas prices, I treat only these two with the intent of demonstrating the plausibility of stochastic nonlinear optimization in capacity expansion models. Note that using these two distributions is already beyond the computational capacity of stochastic linear optimization to handle. The weak scaling of Monte Carlo based numerical optimization makes it computational feasible to treat additional variables. In our understanding, this is the first model to do stochastic non-linear optimization of grid expansion.

The three critical advantages of the approaches used in our model are: 1) the genetic algorithm allows co-optimization over several criteria, such as minimizing cost and meeting RPS or emissions requirements, 2) it permits examination of the dynamics of the decision-making process rather than simply choosing an optimal grid mix at a fixed point in the future, and 3) uncertainty is integrated into the core of the model, allowing exploration of issues such as the effect that uncertainty over fuel prices and load has on choice of new generation. These advantages are possible due to the use of a general search algorithm, and a model design that accounts for both uncertainty in many elements and the non-myopic nature of the model enables to see effects of decisions in earlier periods have on decisions in later periods. Also, this approach does not suffer from the curse of dimensionality for the input distributions, and any number of uncertainties can be considered over the input parameters. In the case study presented in this paper, this study limits the input uncertainties to fuel prices and load, but our model is capable of including other uncertainties such as capital costs, policy constraints, learning rates, emission rates, RPS, etc.

2.2 Method

I use a simplified electricity dispatch model to estimate the total costs of electricity service over a 30-year time horizon (2020-2050), given decisions regarding the construction of new electricity generation in each period. A genetic algorithm optimization is used to search for the generation build-out plans that minimize discounted expected total costs of meeting electricity load [37]. Monte Carlo simulation in the core of the model integrates uncertainty in inputs, such as fuel price, and load.

To further explore the risk aversion as a scenario by itself, the model optimizes for not only minimizing the total cost, but also by minimizing the Conditional Value at Risk (CVaR) in a risk averse scenario. CVAR measures the worst-case costs/tail risk in distribution by taking an average of the extreme tail costs after a chosen cut-off. This technique is commonly used in portfolio optimization of stocks. The cut-off used in this study is 95% which means an average of the worst 5% of the costs gives the CVAR value of the distribution at 95%.

I consider the coal-heavy grid mix of the Midcontinent Independent System Operator (MISO) of the Midwest region in the United States over a 30-year horizon from 2020-2050.

2.3 Modelling Framework

The model has several levels, which interact, as shown in Fig. 1. At the core is a simplified dispatch model that determines whether a set of generation technologies can meet the load and estimates the variable cost of doing so. The long-term assessment model calculates the discounted expected electricity system costs over a 30-year planning horizon from the fixed costs, capital costs of the new power plants, and the variable costs from the dispatch model. Monte Carlo simulations are performed over the long-term assessment model for a distribution of fuel prices and demand inputs. Given the computational limitations, it is unreasonable to run 30 years of hourly electricity simulation from 2020-2050, considering that the Monte Carlo simulation will perform simulations over the distribution of the inputs. Therefore, the model operates at 5-year intervals over the 30- year horizon. The output from the Monte-Carlo is a distribution of costs. The distributions are aggregated (like mean) to a single value, based on the risk preference of the markets and given to the decision model. At the highest level, the decision model -genetic algorithm determines the best set of generation technologies to build from 2020-2050. Using this approach allows modification of variables between periods, integrating changes such as experience curves through learning rates and their effect on prices for generating technologies.

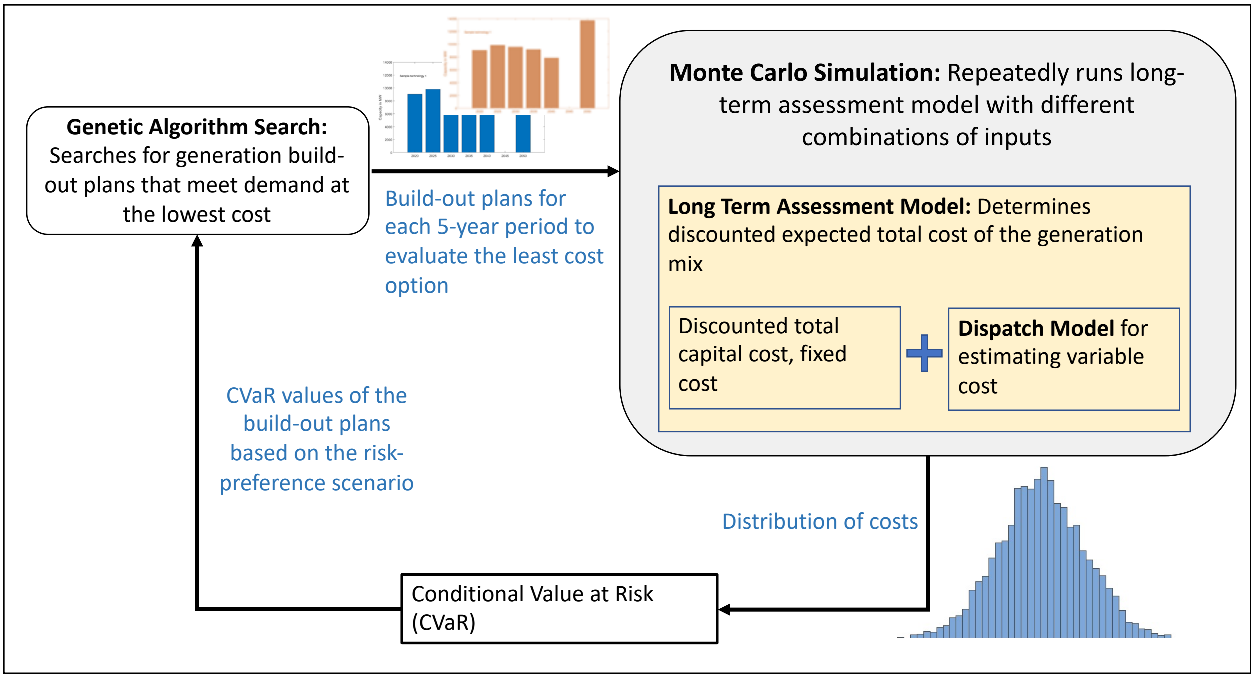

Fig. 1 Framework of methodology for determining the optimized grid build out plan under uncertainty of inputs from 2020-2050 across the Midcontinent Independent System Operation (MISO) region.
The core of the system is the Monte-Carlo simulations of the Long-Term Assessment (LTA) Model, which calculates the distribution of the expected total cost of electricity for different stochastic inputs. Based on a risk preference scenario, Conditional Value at Risk estimates the single point output from the distribution of outputs from the LTA model. Genetic algorithm search is used to identify generation build-out plans that minimize the total system costs while meeting the future uncertain demand.

2.3.1 Inputs

The inputs to the model can be broadly categorized into stochastic and deterministic inputs. The stochastic inputs to the model are distribution of expected natural gas prices, and distribution of expected electricity demand as a function of season and hour-of-day. The deterministic inputs are expected capital costs, discount rate, technology learning rates, and hourly variations of wind/solar as summarized in Table 2.

The model is capable of incorporating other uncertainties such as expected future subsidies, distribution of capital costs, RPS constraints, etc. but is not considered in the current study for MISO.

Table 2. Summary of the inputs, and data sources used in the stochastic model.

Deterministic Inputs	**Source**	**Note**
Capital Cost of new power plants	EIA [38]	Shown in Table 3
Discount Rate	EIA [38], NREL [39]	5%
Fixed and operating cost of new power plants	EIA [38]	Shown in Table 3
Fixed and operating cost of existing plants	eGRID [40]	Appendix A
Heat Rate of new technologies	EIA [38]	Shown in Table 3
Heat Rate of existing plants	eGRID [40]	Appendix A
Learning Rates (single factor)	IEA [23], EIA [24],Rubin et. al [41]	Shown in Table 3
Existing Power Plant Fleet	eGRID [40]	Appendix A
Fuel Prices	EIA [7]	Shown in Fig. 2
Carbon Emissions of existing plants	eGRID [40]	Appendix A
Wind and Solar variability (hourly capacity factors)	NREL [42, 43]	Appendix A
Stochastic Inputs	**Source**	
Natural Gas Prices	Simulated	Shown in Fig. 5
Demand	Simulated	Shown in Fig. 4

2.3.1.1 Deterministic Inputs

This section covers about the deterministic inputs used in the model and summarized in Table 2. The model uses the existing portfolio of generation in the studied area, including the age, efficiencies, emissions, capacities of each plant, to define the starting point for future portfolios from eGRID database [40]. Sample data from the eGRID database is shown in Appendix A.

New Power Plant's characteristics

New generation technologies and their corresponding capital cost, fixed cost, and efficiencies are considered based on the EIA's estimates used for modeling the NEMS' electricity market module [44], shown in Table 3. Overnight capital costs are considered, excluding the financing/interests during the construction and development of the power plants. The discount rate in the model is assumed to be 5% [19, 39].

Table 3. Cost and efficiency characteristics of new generation technologies considered in the study based on EIA's estimates [13].
All the costs are expressed in 2016$. *The capital costs are the overnight costs that exclude interest during the construction and development. ** The capital costs are for the year 2015 and the future capital costs are estimated accounting for the technological progress through learning rates.

Technology	Capital Cost*, ** (2016$/kW)	Fixed Cost (2016$/kW)	Operating Cost (2016$/MWh)	Heat rate (Btu/kWh)	Fuel	Learning Rate (%)
Coal with Carbon Sequestration (CCS)	4983	69.23	7.02	1070	Coal	8.3
Combined Cycle (CC) Natural Gas	962	10.80	3.47	6450	Gas	14
Combustion Turbine	875	17.3	3.47	9900	Gas	15
Biomass	3757	110.00	5.46	13500	Biomass	11
Wind	1622	47	0	0	Wind	12
Solar	1812	22	0	0	Solar	23
Nuclear	5822	99.17	2.27	232930	Uranium	2

Capital costs (accounting for technological progress)

Rapid growth, competition, and technology improvements lead to a significant cost reduction over the time. These cost reductions are generally determined through learning rates. The learning rates *('LR')* assumed in this study are based on the mean learning rates observed from the literature review by Rubin et.al. in their study [41], given in Table 3.

Learning coefficient α determines the capital cost of the technologies *('CC')* based on the initial cost *('CC_o')*, initial capacity of the technology *('P_o')*, and the current cumulative capacity after the new additions *('P')* (Eq. 6). Coefficient α is determined from the learning rate of the technologies, which specifies the cost reduction rate, as the technology capacity is doubled *('LR')* [45] (**Error! Reference source not found.**). The total installed capacity of the technology are determined based on the global level projections from the EIA data [24] and the future capacities in MISO determined by the model.

$$LR = (1 - 2)^{\alpha} \qquad \text{Eq. 1}$$

$$\alpha = \frac{\ln(1 - LR)}{\ln(2)} \qquad \text{Eq. 2}$$

$$CC_t = CC_{t-1}\left(\frac{P_t}{P_{t-1}}\right)^{\alpha} \quad \text{Eq. 3}$$

Where, Subscript t – given year t

LR- Learning Rate

α - Learning Rate coefficient

P – Cumulative capacity (Initial capacity + new capacity additions)

P_0- Initial Capacity

CC – Capital cost

Fuel Prices: Fuel prices of coal, uranium for nuclear power, and oil prices are taken from EIA database [7], shown in Fig. 2, all the units expressed in \$/MMBtu for an easy comparison.

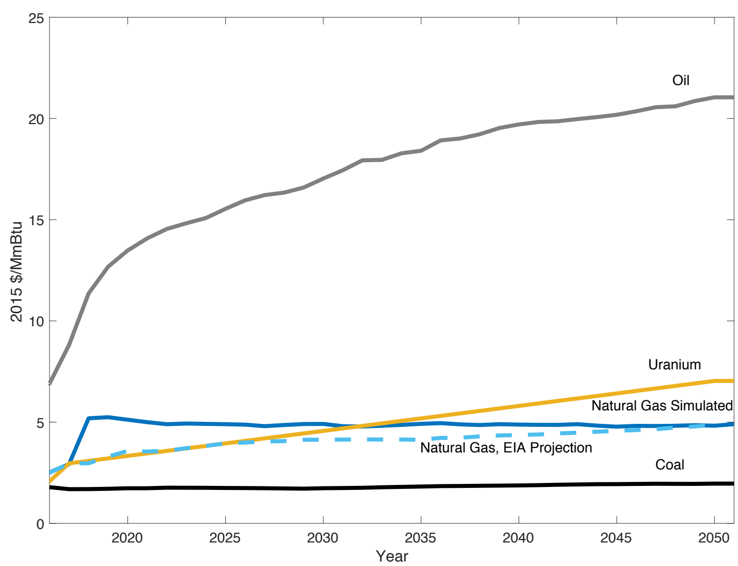

Fig. 2 Fuel prices of coal, natural gas, uranium, and oil considered for the deterministic scenario. The blue dotted line indicates the EIA projection of the natural gas prices. For this study, natural gas prices for the deterministic scenario are estimated from the mean of the stochastic scenario.

Wind/Solar Variability

The hourly generation profiles of solar and wind energy across various locations in MISO are estimated according to the Wind Integration National Database (WIND) toolkit [42] and Eastern Solar Integration Data [43].

The WIND Toolkit provides data related to wind energy production for over 126,000 current and potential locations across the United States for 7 years from 2007–2013 [42]. This dataset consists of meteorological data, 5-min resolution of wind power production, and capacity factors. I consider thirty potential locations in the Midwest region and the corresponding hourly wind output/MW. The average wind energy output (kWh/hour) for a 1kW system across these locations is used to generate the hourly variations of incremental wind capacities considered in the study. Similarly, the Eastern Solar Integration dataset by NREL consist of 5-minute solar power and hourly day-ahead forecasts for approximately 6,000 simulated PV plants. 30 potential sites from 15 states in the Midwest region are considered and a similar procedure to wind energy output is used to generate solar energy output/hour. Annual capacity factors of most of the potential wind power sites in MISO are greater than 40% and most of the solar power sites are greater than 16%. More details on the hourly variation of solar/wind energy output/hour and potential locations considered are provided in the Appendix B section.

2.3.1.2 Stochastic Inputs

In order to model uncertainty, the model will require distributions of possible future values for each input whose uncertainty is considered. In our case, distributions of fuel prices, and load are inputs to the model along with the other deterministic inputs.

Distribution of Natural gas prices:

Volatility in natural gas prices generally exhibit mean reversion and seasonality [46]. Mean reversion is the tendency of natural gas prices to revert to a long-term equilibrium value after fluctuations due to extreme weather, supply, or demand surges. Seasonality is the cyclic variations over the seasons because of the cyclic changes in demand [46]. In the current model, seasonality of the fuel prices is not considered but only the annual variations using Ornstein-Uhlenbeck (OU) mean-reversion process [47]. Historical variations of the Henry Hub natural gas spot prices since 1986 are used to estimate the future uncertainties.

OU process is a variation of the Markov process, i.e. the future value is independent of the past but depends upon the present value [9]. OU process using the stochastic differential equation given in the Eq. 4 [47] is used to determine the discrete natural gas prices for a given period.

Natural logarithm of natural gas prices p_t is used in the equation to avoid negative stochastic prices. Mean reversion rate (α) determines the attraction or repulsive speed from a long-term mean value (μ) of the historical natural gas prices. The volatility (σ) is the 'noise' in the system based on historical standard deviations of monthly natural gas prices from 1986-2017, obtained from EIA [48]. The annual prices input to the model are an average of monthly prices for a given year. All the historical prices are adjusted to 2016-dollar value.

$$\Delta x_t = \underbrace{\alpha(\mu - x_t)\Delta t}_{\text{drift}} + \underbrace{\sigma dZ_t}_{\text{Brownian motion}} \text{, where } dZ_t \sim N\left(0, \sqrt{\Delta_t}\right) \qquad \text{Eq. 4}$$

Where, x_t – $\ln(p_t)$, logarithm of price for a given time t

α - Mean reversion rate

μ- Mean of the log of historical natural gas prices

σ- Volatility

N – Random normal distribution

Reversion rate (α), mean (μ) and, volatility (σ) are calibrated by dividing the Eq. 5 with Δt and by determining the coefficients of $\frac{\Delta x_t}{\Delta t}$ based on historical data. Calibration was performed using 'polyfit' function in Matlab. The values obtained from the coefficients are α- 0.5022, μ- 1.5030, and σ- 0.3963. The solution for the Stochastic Differential Equation in Eq. 4 is as given in the Eq. 5 which is used to generate random time-series of natural gas prices from 2020-2050. A sample of the simulated time series is as shown in Fig. *3*.

$$x_{t+1} = x_t e^{-\alpha\delta} + \mu\left(1 - e^{-\alpha\delta}\right) + \sigma\sqrt{\frac{1-e^{-2\alpha\delta}}{2\alpha}} N(0,1) \qquad \text{Eq. 5}$$

Where, δ- time difference between t and t+1 which in our case is 1 year.

N – Random generator between 0-1 from a normal distribution.

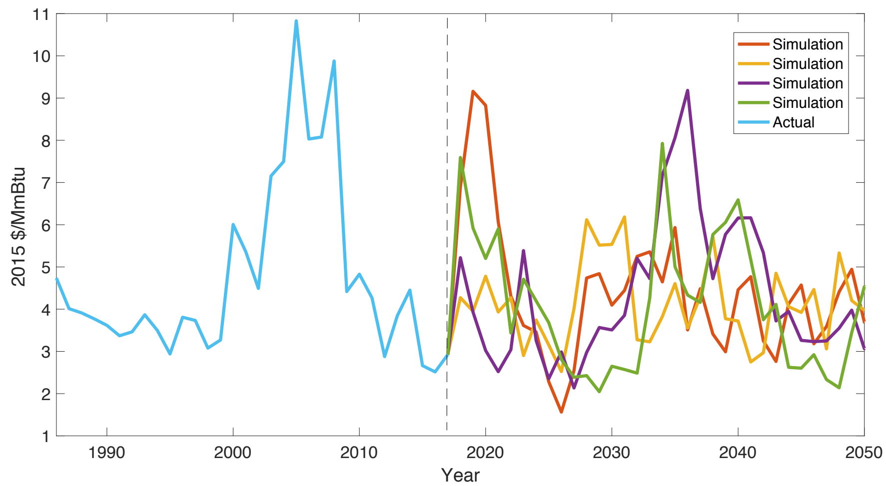

Fig. 3 Natural gas henry hub spot prices and simulated price scenarios from 2018-2050. Ornstein-Uhlenbeck (OU) mean-reversion process is used to create stochastic natural gas prices as an input to the Long-Term Assessment Model.

Input Demand: Similar to natural gas, demand also exhibits two distinct characteristics –long-term demand growth and the seasonality.

Long term demand growth:

The demand growth is estimated using a simple Brownian motion equation as shown in Eq. 6. The *'a'* coefficient of the deterministic part in (Eq. 6) is calibrated for a growth rate of 1% every year based on MISO forecast [49]. The σ is the volatility calculated based on the standard deviation of the change in the historical data which is 415.6 MWh from 2007-2017.

$$L_{t+1} = \underbrace{L_t + a\Delta t}_{\text{Deterministic}} + \underbrace{\sigma dZ_t}_{\text{Diffusion}} \text{, where } dZ_t \sim N\left(0, \sqrt{\Delta_t}\right) \qquad \text{Eq. 6}$$

Where, L_t- Average load for a given year t

a – Linear coefficient of first order linear equation

σ- Volatility

Seasonality and hourly variations:

Seasonality and the hourly variation of the demand are based on the historical load patterns observed in MISO [49]. Percentage change in the load over 8760 hours in a year with respect to the mean load for a given historical year is estimated. These percentage changes provide the information on how the hourly load historically varied with respect to the mean demand in a given year. From Eq. 6, a random mean demand value for a year is estimated. Then a historical sample year *'s'* is chosen with 8760 hourly values of percentage change with respect to the mean demand in that sample year. The new hourly variations are estimated by multiplying the historical variation *('V')* with the random mean demand *('L')* (Eq. 7).

$$l_{h,t,s} = L_t * \frac{V_{h,s}}{100} \qquad \text{Eq. 7}$$

Where, L_t- Average load for a given year t

s- historical sample year

h – hour

V- percentage change with respect to mean demand for a sample year s

l – hourly load value

Stochastic distributions

Distribution of the annual load is constructed from an average load increase of 1%/year and the volatility/noise of the future load at 415.6 MWh (Fig. 4). The hourly load in each Monte Carlo run constructed from the historical variations and the random point of an average load in the distribution for a given year captures both the annual load growth and the seasonality of the load changes (section 2.3).

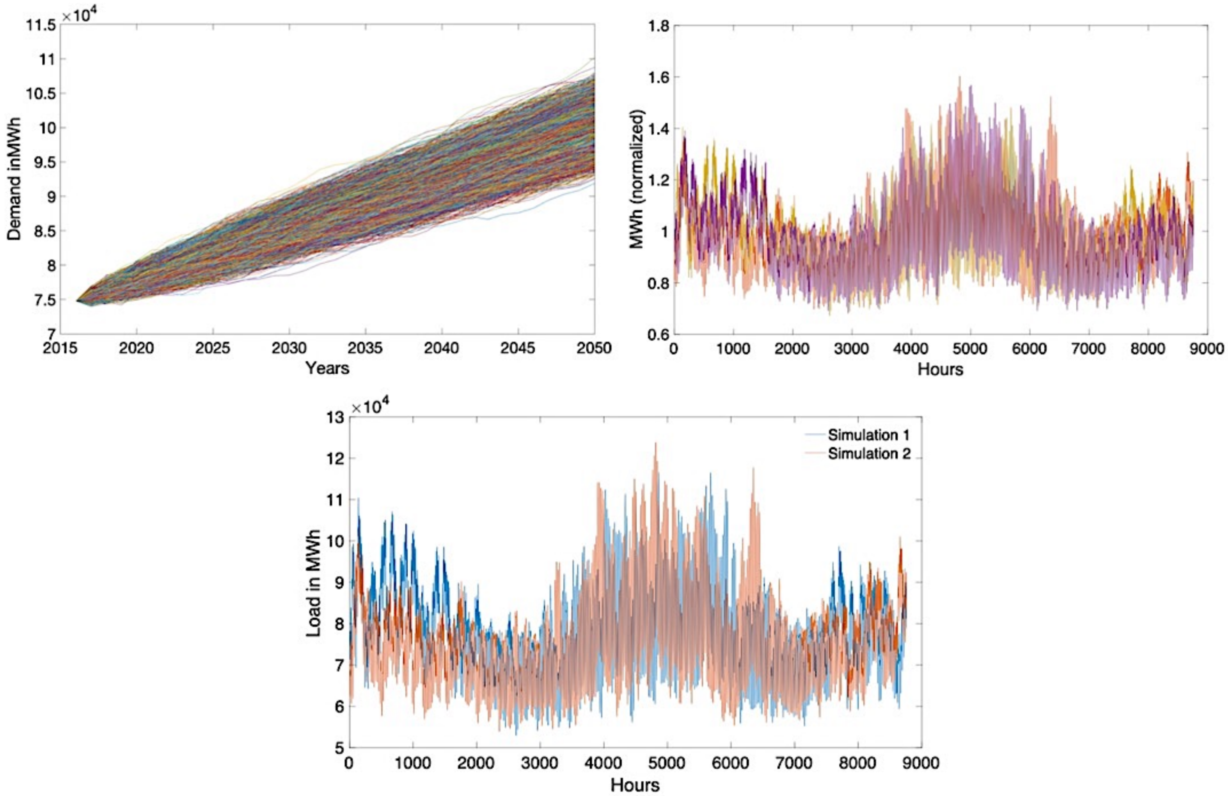

Fig. 4 Simulated hourly load patterns used for Monte Carlo runs in the model.
Top-left figure shows the distribution of average annual load growth from 2015-2050. Each color indicates the stochastic load forecasts from 2020-2050. Top-right figure shows the samples of historical normalized hourly load patterns seen in MISO since 2013. Each color represents the sample of historically observed hourly load patterns in MISO. The bottom most figure shows two random samples of hourly load patterns in the year 2020. They are created from multiplying a random sample point in the distribution from the year 2020 in the top-left figure with a random normalized hourly load pattern for an year in the top-right figure. Similar patterns were created at 5-year steps from 2020-2050.

This study estimates the distribution of the natural gas prices from the historical variations using OU mean-reversion process. Because the natural gas prices cannot go negative, they are skewed towards the positive side of the mean of the distribution (Fig. 5).

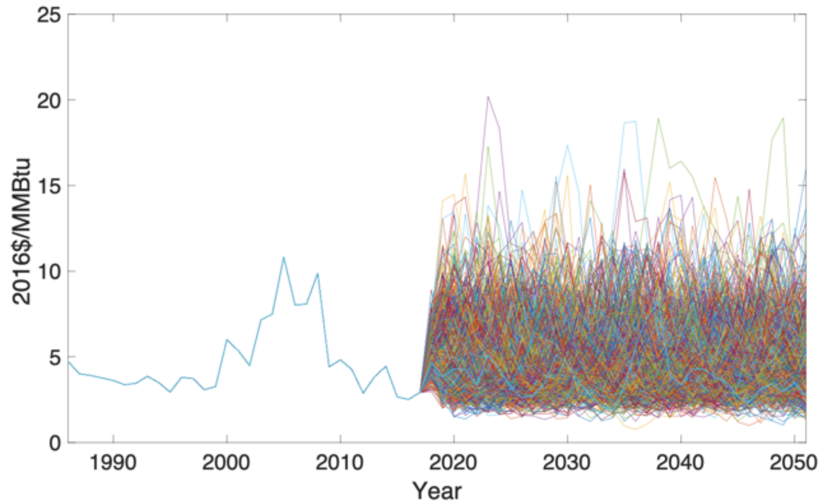

Fig. 5 Distribution of natural gas prices from 2020-2050 used for Monte Carlo runs in the model. Each color indicates the possible price forecast from 2020-2050.

2.3.2 Dispatch Model for variable costs

The lowest level of the proposed optimization model is a dispatch routine that uses simplified rules to determine the variable cost of electricity generation over a year [25]. Due to the need to run many scenarios, both for the Monte Carlo simulation and the genetic algorithm search, there are limitations on computational time. Therefore, the dispatch model is limited to choosing the generation in each hour based on the marginal cost of operation for 8760 hours in a year.

More sophisticated electricity system elements, such as transmission constraints, ramp limits, startup time and spinning reserves, are not included. However, with a high computational ability, modular nature of the model allows a replacement with sophisticated dispatch models, without major changes to the modeling framework.

The principle of the model is to sequentially add plants to the generation mix in order of the marginal cost *('MC')* until the demand is met. The output generation *('e')* of each power plant in a given hour is the capacity of the power plants used to meet the demand. The total variable cost *('VC')* of the electricity generation is the marginal cost *('MC')* incurred by the power plants to produce electricity energy *'e'* for every hour in a year, as shown in Eq. 9.

The current fleet of power plants for electricity generation are taken from EPA's eGRID database [40], and the new generation fleet is added based on the inputs from the decision model and the plant characteristics from EIA data (Table 2) [38]. The marginal cost (assumed as bid price) of operation for each power plant is calculated based on the heat rate [40], and the subsequent fuel costs as given in Eq. 8.

$$\mathrm{MC}_{i,t,p}\left(\frac{\$}{\mathrm{MWh}}\right) = \mathrm{HR_p} * \frac{\mathrm{Price_{i,t,f}}}{1000} + \mathrm{O\&M_p} \qquad \text{Eq. 8}$$

$$\mathrm{VC_{i,t}} = \sum_{p,h,t} MC_{i,t,p} * e_{p,t,i,h} \qquad \text{Eq. 9}$$

Where, Subscript t – hours in a given year

Subscript i- i^{th} Monte-Carlo run

Subscript p – Power Plant

e – Energy output in hour t (MWh/h)

MC- marginal cost of operation of a power plant ($/MWh),

HR- hear rate (Btu/kWh)

Price- average spot price of fuel ($/MMBtu)

O&M – Operations and maintenance cost of the power plant ($/MWh)

VC – Total Variable cost

This study does not model imports of electricity from regions outside of MISO and penalize the model with a high cost of $ 5,000/MWh, when the demand is not met.

2.3.3 Long Term Assessment Model

The long-term assessment model calculates a distribution of discounted expected total system costs (Eq. 10) for meeting load over a 30-year horizon.

$$\text{Total system cost (\$)} = \text{Capital cost} + \text{Fixed cost} + \text{Variable cost} \qquad \text{Eq. 10}$$

$$\mathrm{TC_{i,t}} = CC_t + FC_t + VC_{i,t}$$

Where, Subscript t – Given year

Subscript i- i^{th} Monte-Carlo run

CC – Capital cost ($)

FC – Fixed cost ($)

VC – Total Variable cost ($)

The model operates in 5-year intervals, reducing the model calculations over the 30- year horizon to seven periods. When combined with data on capital and operating costs, the discounted expected total cost of electricity service over the 30-year horizon is calculated. The start year is 2015 and the costs of other years are extrapolated based on the costs estimated at 5-year intervals.

Cashflow

I assume a discount rate (r) of 5% and calculate costs in the 2016-dollar value as shown in Eq. 11 for a given future value (fv) in the year t.

$$C_i = \sum_{t=2015}^{2030} \frac{TC_{i,t}}{(1+r)^{(t-Y)}} \qquad \text{Eq. 11}$$

Where, Subscript i- i^{th} Monte-Carlo run

C- Discounted present value of the cost

TC – Total cost

r-Discount rate, 5%

t-for a given year

Y- reference year, 2016

2.3.4 Monte Carlo Simulation

To include the effects of uncertainty in the Long-Term Assessment (LTA) Model, I use Monte-Carlo simulations by running the Long-Term Assessment model iteratively for random combinations of natural gas prices and load, and the output is a distribution of discounted total system costs.

A total of 20 Monte-Carlo runs are performed for each iteration. Because the process is computationally intensive, the LTA model is not subjected to larger number of Monte Carlo runs. However, the genetic algorithm is an iterative process, it identifies an optimized build-out plan, each time subjected to a distribution of the inputs, and identifies a build-out plan that consistently has the lowest discounted cost of the electricity service. The number of iterations of the genetic algorithm are around 800 before the model converges to a solution.

Deterministic scenario is run using the mean natural gas prices and demand growth and the number of Monte Carlo runs is set to 1 in the model.

2.3.5 Conditional Value at Risk (CVaR)

The output from Monte Carlo simulation is a distribution of output costs for various build-out plans. The cost for optimization to the genetic algorithm in the decision model is calculated based on the CVaR scenario. Risk preference of the model is set at this phase. For a risk neutral scenario, CVaR is 0% and thus, mean of the distribution is fed into the decision model, and for a risk-averse scenario, CVaR is 95% and thus, mean of the worst 5% values are fed into the decision model for optimization (Eq. 12).

$$\mathrm{DC_{CVaR}}\ (\$) = \frac{\sum_{\left(i=\frac{CVaR}{100}*n\right)}^{n} \mathrm{C_i}}{\left(1-\frac{CVaR}{100}\right)*n}\ \ after\ arranging\ C_i in\ ascending\ order \qquad \text{Eq. 12}$$

Where, C- Total discounted present value of the cost

Subscript i- i^{th} Monte-Carlo run

Subscript n – total Monte-Carlo runs

2.3.6 Decision Model -Genetic algorithm for optimization

The decision model uses the genetic algorithm optimization to minimize the cost output from the CvaR, based on the risk preference.

Genetic Algorithm Search: Genetic algorithm is good at rapidly identifying a set of reasonably fit solutions using a heuristic optimization algorithm derived from natural selection process. The genetic algorithm iteratively modifies a population of individual solutions, in our case, generation build-out plans. After every step, random individuals/generation build out plans are selected from the current population as parents to produce children/new build-out plans for the next step/generation. Fitter individuals/lower cost grid build-out plans have a higher probability of getting chosen as parents. The children are created using different crossover techniques, and over time, the population 'evolves' towards an optimal solution [50].

The algorithm broadly works on four essential rules: 1) evaluation, 2) selections rule, 3) crossover rule, and 4) mutation. Evaluation rule applies to calculating the cost of the build-out plans after each iteration in the genetic algorithm. Selection rule applies to assigning probabilities and choosing individuals/generation build-out plans for creating children/new-grid build-out plans. Probability scores to each individual are assigned based on their fitness/expected total cost. Lower the cost (output from CVaR), higher the probability of becoming a parent for the next generation. Crossover rule applies to the process of creating children/new build-out plans from the chosen parents. These children replace weaker individuals with a high cost. In order to avoid local optimization, mutation rule is applied to create 'genetic diversity' in the pool. Based on a user-defined mutation probability, a random bit in the child chromosome/binary form of the new generation build-out is altered before re-converting the children to decimal forms. In our case, the probability of a mutation is set to 3%. The model uses global optimization toolbox in MATLAB R2017b version for implementing the genetic algorithm. Further information on the genetic algorithm from this toolbox can be found on their website [51].

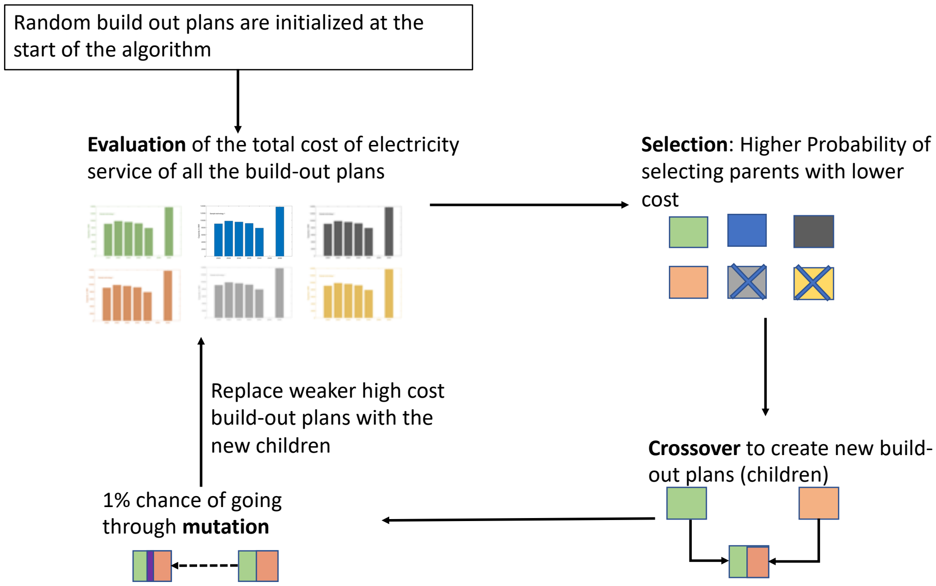

Fig. 6 Brief illustration of flow of steps in the genetic algorithm.
The four principles of genetic algorithm are 1) evaluation of discounted total cost of the build-out plans based on a risk-preference scenario, 2) selection of parents for creating next generation of build-out plans, 3) crossover of parents to create children, and 4) random mutation to avoid local optimization. These steps are repeated till the change in total cost of electricity service remains constant within the set tolerance levels.

2.3.7 Objective function and summary of equations:

The total cost *('TC')* for a given year in a given Monte-Carlo run *('i')* is estimated from the marginal cost of operation from the dispatch model, fixed cost, and capital cost of new power plants (Eq. 15). Maximum electricity generated by a power plant in an hour *'T'* does not exceed the name plate capacity *('P')* of the power plant, shown in equation (Eq. 18). Name plate capacity also includes an additional large capacity of 1000 GW at a high penalty cost of $5,000/MWh, in case the new power plants from the genetic algorithm fail to meet the total demand *('L')*. This is to ensure that the total generation always meets the demand which varies for each Monte-Carlo run (Eq. 16). All the total costs *('TC')* for a given year *'t'* are adjusted to the reference year 2016 dollar value to evaluate the total net present value of the cost *('C')* (Eq. 14). In the final step, Eq. 13 shows the objective function of minimizing total discounted electricity service cost *('DC')* depending on the CVaR scenario.

Objective function, minimize:

$$\mathrm{DC_{CVaR}}\ (\$) = \frac{\sum_{\left(i=\frac{CVaR}{100}*n\right)}^{n} \mathrm{C_i}}{\left(1-\frac{CVaR}{100}\right)*n}\ after\ arranging\ C_i in\ ascending\ order \qquad \text{Eq. 13}$$

Subject to:

$$C_i = \sum_{t=2015}^{2030} \frac{\mathrm{TC_{i,t}}}{(1+\mathrm{r})^{(\mathrm{t-Y})}} \qquad \text{Eq. 14}$$

$$TC_{i,t} = \sum_{\mathrm{p,h=1}}^{8760}(\mathrm{MC}_{i,t,h,p} * \mathrm{e}_{i,t,h,p}) + \sum_{p}(FC_{i,t,p} * P_p + CC_{i,t,p_{new}} * P_{p_{new}}) \qquad \text{Eq. 15}$$

$$\sum_{\mathrm{T,p}} \mathrm{e_{i,t,h,p}} \geq \mathrm{L_{i,t}} \qquad \text{Eq. 16}$$

$$\mathrm{e_{t,T,p}} > 0 \qquad \text{Eq. 17}$$

$$\mathrm{e_{t,Tp}} \leq P_t \qquad \text{Eq. 18}$$

$$1 \leq \mathrm{T} \leq 8760 \qquad \text{Eq. 19}$$

Where, C- Discounted present value of the cost,

DC – Discounted cost of electricity service for a given CVaR scenario

TC – Total cost of electricity service

i- i^{th} Monte-Carlo run

n – total Monte-Carlo runs

r-Discount rate, 5%

t-for a given year

Y- reference year, 2016

h- hours in a given year

Subscript p - Power plant,

p_{new}- new power plants (MW),

MC - marginal cost of operation of power plant ($/MWh)

e – electricity generated by power plant in a given hour (MWh)

FC- Fixed cost ($/MW)

CC- Capital cost ($/MW)

P – name plate capacity of the power plants (MW)

L_t - load (MWh).

Retirement

The retirement of the power plants is case-specific and depends on a number of factors such as wholesale electricity prices, inefficiency and high costs of operation, and environmental regulations [52]. In 2017, most of the retirement decisions in MISO were from uneconomic power plant units [53].

In the current study, economy of the power plants is based on their cost of operation for every 5-year period. The model endogenously retires the power plants by allowing the optimization to randomly choose positive or negative capacity additions. Positive additions denote new generation technologies, and 'negative' additions denote retirement of the power plants for a specific fuel type. For the negative capacities, power plants with a high annual cost of operation per unit nameplate capacity for a given fuel type are assumed to be uneconomical to operate and are retired until the retired capacities equal the negative capacities by the genetic algorithm. The total cost of electricity service is then calculated for the resultant build-out.

2.3.8 Reporting output distributions of cost and emissions

The output cost and emissions distributions of the resultant build-outs for different risk preferences are presented in the section 2.4 below. These distributions are plotted to understand the probability distributions of NPV for deterministic, risk-averse, and risk neutral scenarios.

These distributions are calculated by running the resultant build-outs for different risk preferences through LTA model and Monte Carlo model. A fixed sample distribution of natural gas prices and demand is assumed for all the build-outs and when run through 1000 Monte Carlo simulations provides a distribution of output costs and emissions. A fixed sample of 1000 different natural gas prices and demand growth patterns is used to ensure a fair comparison between scenarios.

The total annual CO_{2eq}. emissions (in million metric tonnes) for each Monte Carlo run are calculated based on the hourly dispatch of plants as shown in equations (19-20). The plant-level emission rates are in metric tonnes/MWh, taken from the eGRID database [40] and the emission characteristics of the new power plants are based on the EIA data [38]. Total CO_{2eq}. emissions are comprised of all greenhouse gas emissions measured on a common scale based on their Global Warming Potential (GWP) relative to CO_2 [54].

The total CO_{2eq} emissions in a given hour for a given operation schedule of generator plants is given by Eq. 46:

$$em_i = \sum_{p,t} m_{p,i} * e_{p,t,i} \quad t = 1,2 \dots ,8760 \qquad \text{Eq. 20}$$

Where, Subscript t – hours in a given year

Subscript i- i^{th} Monte-Carlo run

Subscript p – Power Plant

e – Energy output in hour t (MWh/h)

m – CO_2 eq. emissions of plant p per MWh (in metric tonnes/MWh)

em – total emissions (metric tonnes)

2.4 Results

I run an alternative structure for capacity expansion modeling in MISO with uncertainty considered for inputs load and natural gas prices to show comparisons between Deterministic scenario and stochastic scenario; and to compare between risk-neutral and risk-averse scenarios when the inputs are stochastic. The risk-neutral scenario optimizes for the mean of the distribution, i.e., CVaR at 0, and risk-averse scenario optimizes for CVaR at 95%.

2.4.1 Risk-Neutral Scenario

The base-case scenario is the risk-neutral scenario that optimizes for the mean of the distribution. Output results for MISO show a dominant mix of wind and natural gas in the total capacity by 2050. Though natural gas and coal dominate more than 50% of the total capacity, wind constitutes a significant 36% of the rest of the capacity (Fig. 7). For a risk-neutral scenario, the capacity mix by 2050 is 41% natural gas (87 GW), 32% wind (70 GW), 25% coal (53 GW), 2% solar (3 GW), 2% hydro (3 GW), 1% biomass (2 GW) and solar, 1% nuclear (2 GW) and less than 1% Oil (0.2 GW) (Fig. 7). The energy mix differs depending upon the natural gas prices and demand, which are considered stochastic for this scenario. For the average natural gas prices in the distribution and average demand, coal dominates the output generation mix in 2050 at

41%, followed by wind (34%), natural gas (20%), nuclear (3%), biomass (2%), hydro (2%), solar (<1%), and oil (<1%) (Fig. 7).

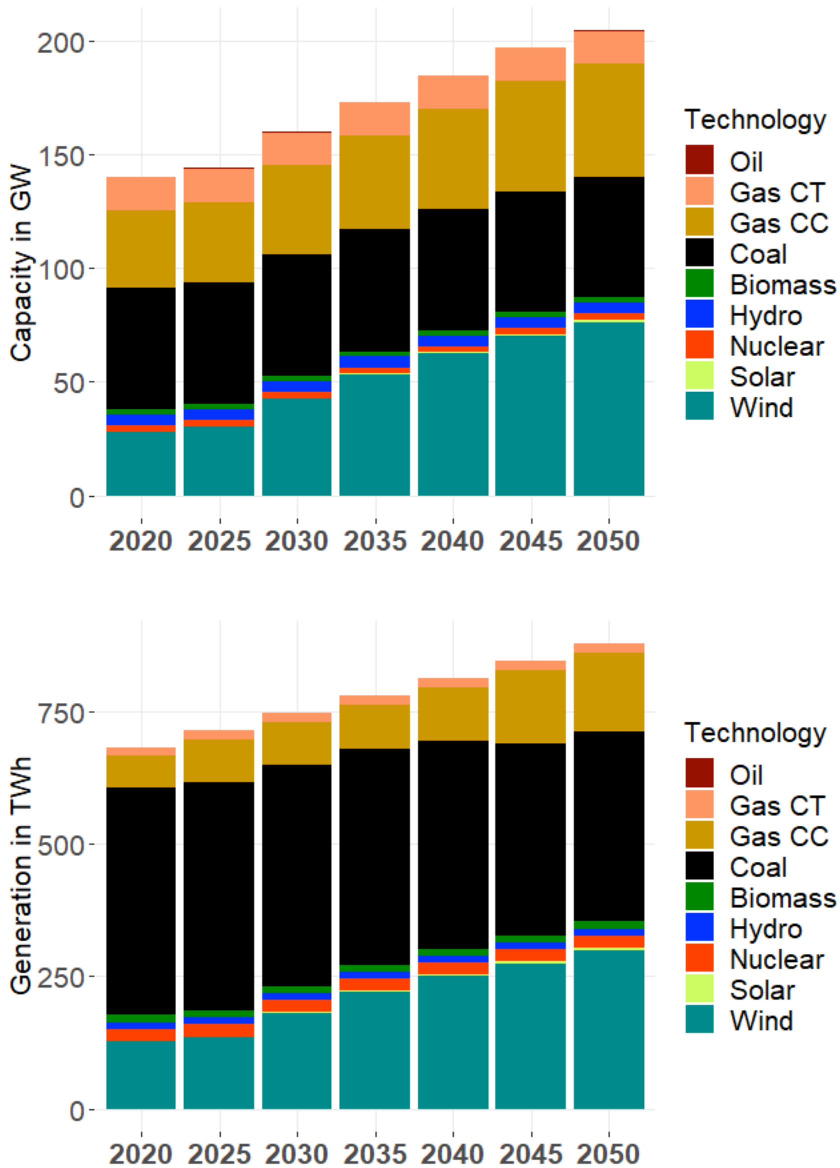

Fig. 7 Top-figure represents the Capacity mix from 2020-2050 in a risk-neutral scenario, bottom-figure represents the generation mix from 2020-2050 in a risk-neutral scenario, for mean natural gas prices and demand in the overall input distribution .
X- axis represents the year and, in the top, y-axis represents the capacity in GW, and in the bottom-figure, y-axis represents the generation in TWh. Colors of the bars indicate the technology type.

2.4.2 Risk-Neutral Scenario and Deterministic Scenario Cost Distributions

When a deterministic scenario is run without any uncertainties in the input parameters or Monte-Carlo simulations, results show lower capacity additions as compared to the risk-neutral scenario (Fig. 9). The risk-neutral scenario adds ~5GW additional wind capacity, ~20 GW more Gas CC, but ~8GW lower solar capacity by 2050.

Probability distributions of NPV/discounted total cost of electricity service from the resultant build-outs of risk-neutral scenario and deterministic scenario are shown in Fig. 8. Results from this comparison show that the mean NPV of the probability distributions of the deterministic scenario is slightly higher than the risk-neutral scenario by $5 billion. The mean NPV of the output distribution of the risk-neutral scenario is ~480 billion dollars, and the probability of extreme/tail cost are lower than the deterministic scenario (Fig. 8). The deterministic scenario optimizes for the average demand and natural gas prices but does not include the lower or higher price/demand shocks in the system. The probability of high total system cost for a risk-neutral scenario is lower than the deterministic scenario, as the stochasticity in the model optimizes for the expected distribution of load and, natural gas prices. Therefore, between both the scenarios, the stochasticity in the inputs for the risk-neutral scenario optimizes better for output probability distribution of total cost of electricity service than the deterministic scenario.

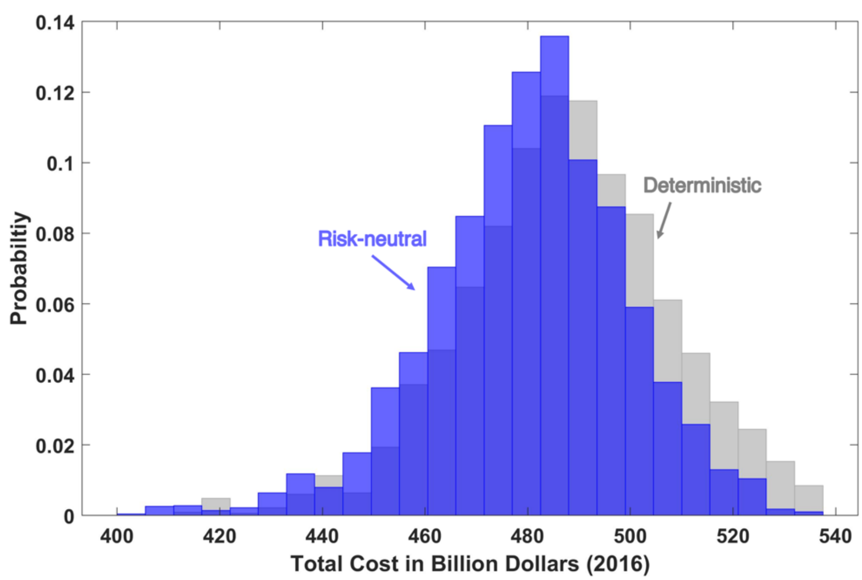

Fig. 8 Probability distribution of the discounted total cost of electricity service for risk-neutral and deterministic scenarios from 2020-2050, generated when the resultant build-out plans are run through a sample of 1000 random natural gas prices and demand.

X-axis represents the total discounted system cost of electricity and the y-axis represents the probability. Colors represent the scenarios. Risk-neutral scenario is optimized for mean of the distribution and deterministic scenario is optimized without any uncertaintites in the inputs.

2.4.3 Risk-Averse Scenario, Risk-Neutral Scenario and Deterministic Scenario Comparisons

In the previous sections, results so far compare the deterministic and stochastic scenarios. In this section, within the stochastic scenario, results between risk-averse and risk-neutral scenarios are compared for different risk preferences. The risk-averse scenario optimizes for the CVaR at 95% cut-off and risk-neutral scenario optimizes for the CVaR at 0% (which is for the mean of the output distribution).

Resultant build-out plans are plotted using boxplots as the inherent nature of the non-linear optimization does not provide a single unique value. Boxplots of risk-averse scenario indicate higher capacity additions of wind and solar, and lower capacities of natural gas by 2050 as compared to the risk-neutral scenario (Fig. 9). The risk-averse optimizes for the probability of high costs and the lower additions of natural gas capacity reduce the likelihood of high NPV from price uncertainties. The risk-averse scenario adds ~ 20GW of more wind capacity, and ~15 GW of more solar capacity and ~5GW of lower Natural gas combined cycle capacity by 2050 as compared to the risk-neutral scenario (Fig. 9).

When all the scenarios are compared, deterministic scenario has the lowest capacity additions as it does not optimize for the high/low demand and gas price scenarios. Risk-neutral scenario relatively adds more wind and combined cycle gas plants for meeting the likely high demand in the future. In the risk-averse scenario, more renewables are added into the system to meet the future uncertain high demands while avoiding the high gas prices.

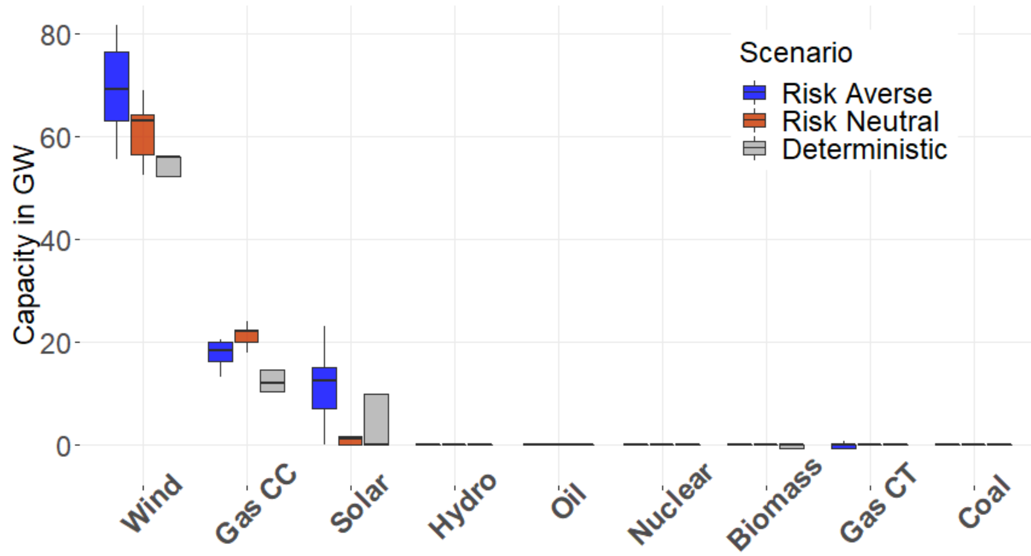

Fig. 9 Boxplot of cumulative additions of different generation technologies by 2050, comparing the deterministic scenario and different risk preferences under the stochastic scenario. The risk preferences considered in the study are risk-averse scenario and risk-neutral scenario. Risk-averse scenario is optimized for CVaR at 95% and risk-neutral scenario is optimized for the mean of the output NPV distribution.
X-axis represents the generation technologies and y-axis represents the capacity additions in GW. Colors represent different scenarios. The bars represent the variations in build-out plans as a resultant of using genetic algorithm search.

When we consider the output probability distributions of NPV of resultant build-plans for the different risk preferences in the stochastic scenario, the mean NPV of the risk-averse scenario is slightly higher than the risk neutral scenario by $1 billion. Mean NPV of the risk-neutral scenario is ~$480 billion. T sample test shows that the difference in means of output probability distributions of both the scenarios are statistically insignificant. However, the extreme/worst tail NPVs in both the distributions after a cut-off of 95% were statistically different and the risk-averse scenario’s CVaR value of $514 billion is lower than risk-neutral scenario’s CVaR value of $518 billion (Fig. 10).

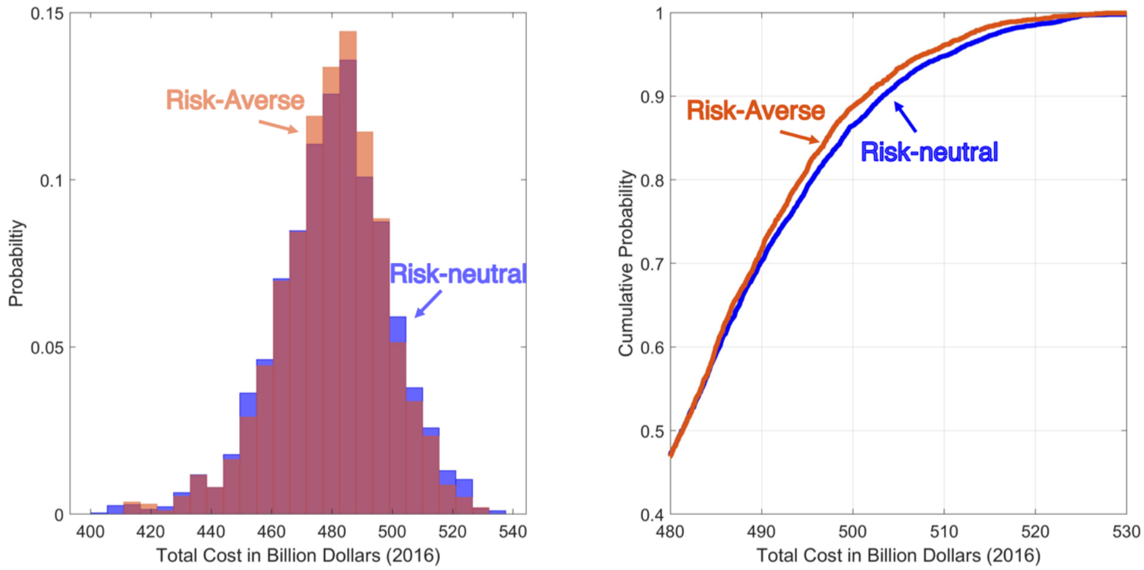

Fig. 10 Probability distribution (left) and cumulative distributions (right) of the discounted total cost of the electricity service from 2020-2050 for risk-neutral and risk-averse scenarios.
X-axis represents the total discounted system cost and the y-axis represents the probability. Colors represent the scenarios. Risk neutral scenario is optimized for mean of the distribution and risk averse scenario is optimized for CVaR at 95%.

2.4.4 Comparison of Emissions

The distribution of output CO_2 emissions are estimated from subjecting the results to a sample of 1000 different natural gas prices and demand. The emissions are lowest for the risk averse-scenario at an average of 23 Metric tons/MWh, second for the risk-neutral scenario at an average of 25 Metric tons/MWh, and highest for the deterministic scenario at 28 Metric tons/MWh (Fig. 11). Higher additions of renewables in the risk averse scenario enables least emissions of all the scenarios.

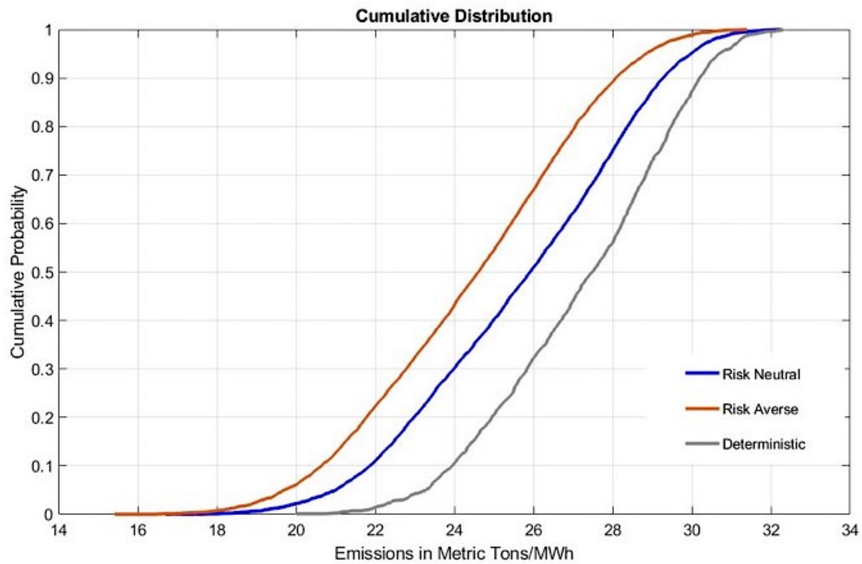

Fig. 11 Cumulative distributions of the output emissions for deterministic, risk-neutral and risk-averse scenario. X-axis represents the total discounted system cost and the Y-axis represents the probability. Colors represent the scenarios. Risk-neutral scenario is optimized for mean of the distribution and risk-averse scenario is optimized for CVaR at 95%.

2.4.5 Summary of Results

Summary of the results comparing the deterministic, risk-neutral, and risk-averse scenarios is shown in Table 4.

Table 4. Comparison of all the scenarios and summary of the results. The scenarios include deterministic scenario, risk-neutral scenario, and risk-averse scenario.

Scenario	Mean discounted total system cost (2016 billion$)	Conditional Value at Risk at 95% (2016 billion$)	Mean of output distribution of emissions Metric Tons/MWh	Mean % contribution of wind and solar to the total grid capacity by 2050
Deterministic	485	530	28	30%
Risk-neutral	480	518	25	34%
Risk-averse	481	514	23	40%

2.5 Contribution to the literature and discussion

In conclusion, this study develops the first stochastic capacity expansion model that does not suffer from the curse of dimensionality (i.e., more uncertain parameters can be treated) and is flexible in allowing different optimization objectives that manage risk, which are the major limitations in the stochastic linear approach used in the current stochastic models [28, 29, 55].

Rather than discretizing the non-linear problem into limited discrete scenarios, a stochastic non-linear approach is used in this study to allow the flexibility of any number of uncertain parameters and different risk objectives, distinct from widely used stochastic linear optimization approach in the prior studies [28, 29, 55].

Current energy system models construct a future grid for the society across a long-time horizon of two decades or more, often assuming a static social change, institutional change, technological change, and innovations [56]. However, longer the time frame, more uncertain are the static underlying assumptions, widely changing the outcomes. A choice that minimizes system costs for static baseline values of input parameters may, in fact, incur a high risk of cost increases. Given the complexity of these models, the validation is difficult, but these models should be able to 'sensibly' provide an outcome for the current planners and avoid future lock-in effects of investments. Therefore, stochastic optimization is critical to optimize for an expected, but not a specific outcome. Following the idea that sensibly building the energy system should account for risk and uncertainty, capacity expansion models should thus use stochastic optimization. Not accounting for uncertainty could be one of the reasons why most of the deterministic models underpredict renewable adoption [27] or overpredict future demand [56].

A key finding from this study is that comparing the distribution of costs allows us to see that the risk-averse scenario has almost the same mean as the risk-neutral scenario but has more renewables. Also, minimizing for deterministic inputs does not necessarily produce optimized results when subjected to uncertainty Overall, risk-averse scenario has the least emissions of all the scenarios, while deterministic scenario results in the highest emissions.

The results show that electricity system costs are inelastic to renewable adoption and the optimization space for the total cost of electricity service is like a 'flat bowl,' i.e., a small increase in net system costs packs more renewables, in our case shown through a small degree of risk aversion. Most of the current models do not adequately explore the optimization space but conclude fewer renewables at a fixed minimization point. Also, risk aversion attitude in the electricity markets could have led to more renewables than the predictions in the recent past.

Further research with a broader range of uncertainties in policies, other costs, deployment of energy storage, etc. could help policy makers frame and address the pertinent environmental issues from the electricity grid. Our model integrated with uncertainty can also be extended to optimize for a minimum cost of grid buildout meeting RPS standards and with several technology subsidies. While RPS specifies target years for minimum adoption levels for renewables, public subsidies and utility costs depend on the trajectory through which targets are met. Uncertainty integrated into the core of the model will allow exploration of issues such as the effect of subsidies/changing policies over the choice of new generations and address the probability of sustainability challenges that could be addressed.

Chapter 3: How much wind and solar are needed to realize emissions benefits from storage?

Abstract

Environmental outcomes from energy storage depend on its usage patterns, the existing generation fleet, and fossil fuel prices. This work models the deployment of large, non-marginal quantities of energy storage and wind and solar power to determine their combined effects on grid system emissions. Two different grid environments are analyzed: a coal-heavy grid (Midcontinent ISO) and non-coal grid (New York ISO). An iterative dispatch model is used that operates storage to maximize income, considering that this operation can influence wholesale energy prices. With current low natural gas prices ($2.6 per MMBtu), adding storage slightly reduces carbon emissions in New York, while increasing them in the Midcontinent ISO (MISO). Storage increases carbon emissions when it enables a high emissions generator, such as a coal plant, to substitute for a cleaner plant, such as natural gas. The study estimates that adding storage operated to maximize revenue in the MISO region will not be carbon neutral until wind or solar power reach around 18% of the generation capacity. Different operation patterns for storage could realize higher carbon reductions. For example, a carbon price on emissions from generators would shift operation to make energy storage carbon neutral even with current wind and solar capacities. Sensitivity analysis shows that a higher natural gas price ($5 per MMBtu) yields much higher storage-induced carbon emissions in both NYISO and MISO and storage in MISO will not be carbon neutral unless 35% of total generation capacity is from wind/solar. This illustrates that low cost; efficient natural gas generation is important to realize emissions reductions with storage under economic arbitrage.

3.1 Introduction and Literature Review

As of 2015, emissions from electricity generation in the United States contribute 27% of total US energy-related greenhouse gas emissions [57]. Renewable electricity technologies are a widely-discussed solution to reduce electricity system emissions of all kinds. However, given the intermittent nature of renewable technologies (wind and solar), large-scale integration is challenging [58],[59]. Energy storage is a potential solution to the intermittency of renewables. However, the discourse on storage often presumes it to be inherently neutral or beneficial with regards to greenhouse gas emissions.

Policymakers in some jurisdictions have been promoting storage, e.g. through tax incentives or regulatory mandates, partly on the assumption that storage is an obvious or necessary complement to renewables. In 2017, Maryland passed a bill to provide tax credits for up to 30% of the cost of residential and commercial storage systems [15], becoming the first US state to provide exclusive tax credits for energy storage systems. In 2013, the California Public Utilities Commission required the state's three largest utilities to add 1.3 GW of energy storage through 2020 [13], arguing that storage "...stores [energy] when consumption is low and puts it back onto the grid when needed at peak demand times …it is beginning to revolutionize the electric system by enabling increased renewables integration, increasing grid optimization, and reducing greenhouse gas emissions" [60].

Electricity grids are complex techno-economic systems, and it is important to explore whether the above assumptions about storage are correct. While storage systems certainly can solve the intermittency issues of renewable energy, storage in the US is rarely used to prevent curtailment of renewable energy. 88% of the total storage capacity in the US operates for profit maximization in an arbitrage scenario [11]. Deregulated grids feature generators (and consequently storage) as profit-maximizing agents. A profit-maximizing bulk energy storage system charges during low price/low demand periods and discharges during high price/high demand periods, regardless of the type of generation being used. The effect that this economic dispatch of storage has on grid emissions depends upon generation mix, dispatch order, demand, and storage round-trip efficiency [20].

Comprehensive evaluation of the environmental outcomes from the deployment of energy storage is only recently being explored. Reviewing prior studies on the operation of economically arbitraged storage, Lin et al model emissions changes due to storage under different grid configurations in IEEE 9- and 30-bus systems using a dispatch model [61]. Their results indicate that net emissions from additional storage are likely to increase when non-flexible, high-emission systems provide base load and flexible, low-emission systems meet peak load. Similarly, Hittinger and Azevedo calculated emissions from new storage using a Marginal Emission Factor approach. They conclude that, subject to the location and operation of storage, net CO_2 emissions consistently rise with addition of storage to the grid, varying between 100-400 kg/MWh (of delivered electricity)[62]. That work was performed for 20 eGRID sub-regions of the United States and modeled small storage systems (20 MW) as price-takers. However, large storage systems will substantially alter demand patterns, prices, and dispatched generation; and marginal system emissions will change as more storage or renewable generation is added [63]. Prior work has established that, depending on the grid mix and how it is operated, storage can have positive or negative effects on carbon emissions. But an important and unresolved question is how emissions due to storage change as intermittent renewables (wind and solar) are added to a grid and as natural gas prices vary. More renewables increase the likelihood that storage is used to substitute fossil generation with excess wind or solar energy. But when natural gas is expensive relative to coal, storage tends to provide more peak power using energy from coal plants.

This study models the economic dispatch of price-making energy storage on two electricity systems while adding increasing quantities of wind and solar generation. The hypothesis is that, while storage will initially increase CO_2 emissions (or break even), emissions induced by storage will decrease as wind and solar are added and eventually become negative. Large storage systems often have a noticeable effect on electricity prices and should be modelled as price-makers [64]. Not accounting for this dynamic can lead to incorrect assessment of operation, revenue, and emissions [65]. In this work, I consider storage as a price-maker, and build an iterative dispatch model to investigate the effect of bulk energy storage additions. I apply our model to two electricity systems - the New York and Midcontinent ISO regions - and investigate the system emissions as wind/solar capacity is added, with the goal of better understanding how large quantities of new renewables and storage will interact to affect emissions.

The New York ISO (NYISO) and Midcontinent ISO (MISO) regions are modeled. There are plans in both to add large quantities of new wind or solar and, potentially, bulk energy storage [66],[67]. The mix of current generation resources is very different in the two regions. In NYISO, the power plants' capacity mix is 47% natural gas, 18% oil, 13% nuclear, 11% hydro, 6% coal, 4% solar & wind, and 2% biomass. In contrast, the MISO power plants' capacity mix is 41% coal, 29% natural gas, 10% nuclear, 10% solar & wind, 3% oil, 2% biomass and 2% other fuels[1][40]. The choice of these two case studies thus allows us to contrast between grids dependent and not dependent on coal. The NYISO and MISO grids are first modeled with the current generator mix and fuel prices, then modified with increased capacities of wind and/or solar power, keeping the capacity of other generators constant. This is not a forecast of an expected future grid mix, rather an exploration of relationship between new storage and renewables: How does storage affect grid emissions with differing amounts of added renewables, and differing natural gas prices? The result informs qualitative trends, in particular identification of a transition when storage decreases rather than increases emissions on a coal heavy grid. The two grids that I study are representative of many systems in the US and around the world. While there are no numerical results for other grids, the existence of an emissions transition with increased renewables for both suggests a general result: there is some level of renewable adoption for which storage is ensured to deliver emissions benefits, but this level can vary considerably based on the existing generation fleet and fuel prices.

These results are contingent on operating storage under economic arbitrage, i.e. maximizing income. Different operational modes could lead to different emissions outcomes. I explore an economic mechanism to shift storage operation towards emission reductions: a carbon tax on emission from generators. Also, currently low natural gas prices may not hold in the long term - an increase would affect the dispatch order of generation and, in turn, the $CO_{2eq.}$ emissions from storage. To address this, I calculate results for both a low (current during 2015-2016) and high natural gas price scenario.

3.2 Methods

I model the emissions from storage operations in NYISO and MISO, treating storage as a price-maker. I estimate the total grid emissions with and without storage. This allows us to estimate the change in emissions when storage is added to the system, which I refer to as the “storage-induced emissions”, estimated as given in Eq. 21:

[1] other fuels include waste heat, unknown, or purchased according to e Grid database.

$$Storage\ induced\ emissions = \qquad \text{Eq. 21}$$

$$(Grid\ emissions\ with\ storage) - (Grid\ emissions\ without\ storage)$$

If the outcome is an increase in total grid emissions, I find the additional renewable generation (wind and solar) required to realize reductions, with and without a carbon tax. Finally, I perform sensitivity analysis under high and currently low (2015-2016) natural gas price scenarios.

3.2.1 Modeling Framework

To estimate the storage-induced emissions (equation (1)), I combine models of different elements of grid operation. An economic dispatch model determines the lowest-cost operation of generating facilities that can reliably meet a given demand within the generators' ramping constraints [68] and simulates the market clearing prices for electricity. These electricity prices are used in an optimization model to determine the schedule for storage operation, considering the effects of large storage on electricity prices. A model accounting for the diversity of plant efficiencies in a region estimates carbon emissions with and without storage. These sub-models are sequentially implemented as illustrated in the flowchart in Fig. 12. Wind and solar are incrementally added to the grid mix and the dispatch/storage/emissions sub-modules are run again in-order to determine the joint effect of large-scale renewables and energy storage. All models are developed using the Matlab software package, version R2016a [69].

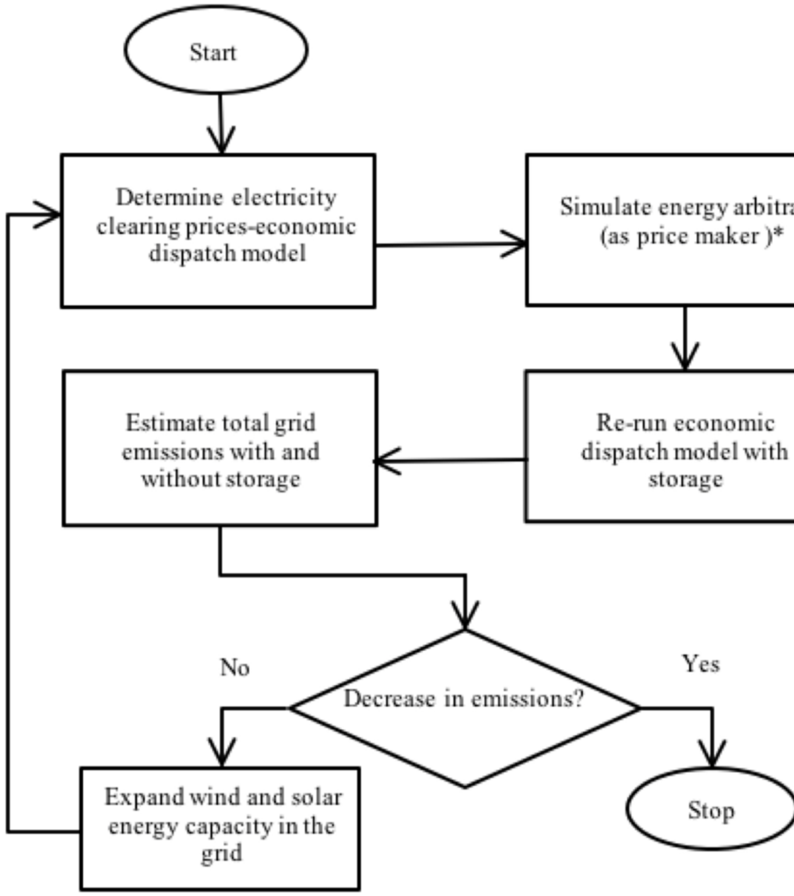

Fig. 12 Flowchart of methodology for evaluating total grid emissions from adding storage and renewable generation. The model produces a "no-storage" time series of prices, simulates storage operation, then calculates the system emissions with and without energy storage. Wind and solar generation are added between simulations until the addition of storage no longer increases system emissions. *For simulating storage operation, I use an iterative dispatch optimization which is shown separately in detail in Fig. 3.

3.2.2 Economic dispatch model and electricity clearing prices:

The economic dispatch model is the first block of our framework (Fig. 12) used to generate electricity clearing prices, which are used as an input to model the operation of the storage. I assume an economic dispatch of generators, where generating facilities place bids based on their marginal costs. After placing bids, ISOs dispatch power plants sequentially from lowest to highest bid, within the ramp rate constraints of each generator, until electricity demand is fully met. This enables determination of market clearing prices. The clearing price is the bid price at which the last unit of electricity is supplied to meet the total demand.

I base the fleet of power plants on data for MISO and NYISO from the EPA eGRID database [40], and calculate each individual power plant's marginal cost of operation based on their respective heat rate, fuel cost, and operations & maintenance (O&M) cost. The dispatch model includes ramping constraints but does not include transmission constraints, assuming new transmission lines will be built in the future to sufficiently accommodate supply expansion.
The reference electricity demand is taken from market data available from NYISO and MISO for 2015 [70, 71]. The fleet of power plants for electricity generation are taken from EPA's eGRID database [40] and the marginal cost (assumed as bid price) of operation for each power plant is calculated based on the power plant's heat rate [40], subsequent fuel costs (Table 5), and variable O&M costs [72].

The Marginal Cost (MC) given in \$/MWh is the summation of the fuel cost incurred per MWh and the variable O&M costs per MWh as shown in Eq. 22. The Heat Rate (HR) for each power plant, expressed in Btu/kWh, is considered (from eGRID data [40]) to estimate the fuel cost incurred to generate one unit of energy in MWh. Variable O&M costs for each power plant are considered based on the generator type and the primary fuel used for the generation of electricity (Table 6). A summary of data sources used in the economic dispatch model are provided in Table 1 of the Supporting Information (SI).

$$MC\ (\$/MWh) = HR * \frac{Price}{1000} + O\&M \qquad \text{Eq. 22}$$

Where, MC- marginal cost of operation of a power plant (\$/MWh),
HR - heat rate (in Btu/kWh),
Price - average spot price of fuel (in \$/MMBtu), and
O&M- variable operations and maintenance cost of the power plant (in \$/MWh).

With an additional carbon tax, the marginal cost of power plants increases based on their emission rates as in equation Eq. 23.

$$MC_{(carbon\ tax)}\ (\$/MWh) = HR * \frac{Price}{1000} + O\&M + \frac{Carbon\ Tax}{1000} * Emission\ Rate \qquad \text{Eq. 23}$$

Where, the carbon tax is expressed in \$/metric tonne, and the emissions rate is expressed in kg/MWh.

Table 5 and Table 6 show fuels costs and variable O&M costs, respectively, used in the modeling.

Table 5. Average fuel costs used for electricity production during the years 2015-2016. Four major types of fuels used for electricity production are considered. The normalized average price of coal includes the different qualities of coal used for electricity production. The original value of crude oil as per the reference is given in $/barrel and converted to MMBtu with the conversion: 1barrel = 5.55MMBtu for crude oil. Constant 2015-$ are used.

Fuel Type	**Cost**	**Units**
Natural Gas	2.6 [73]	$/MMBtu
Coal	2 [73]	$/MMBtu
Uranium	1.4 [74]	$/MWh
Crude Oil	7.99 [75]	$/MMBtu

Table 6. Variable O&M costs of technologies considered in this study [72]. All values are expressed in constant 2015-$. The variable O&M cost of wind and solar power plants is taken as zero.

Technology	**Variable O&M Costs (2015 $/MWh)**
Conventional Hydropower	2.62
Coal power plants with steam turbines	6.96
Combined Cycle power plants (Gas/Oil)	1.96
Conventional Combustion Turbine (Gas/Oil)	3.43
Gas Turbine	3.43
Nuclear	2.26

Using marginal cost as the bid price of power plants, the economic dispatch model is run with an objective of producing electricity at a minimum operating cost using a linear optimization method. Eq. 24Eq. 25 show the objective function without carbon tax scenario and with carbon tax scenario, respectively. Marginal cost of operation (MC) of power plants for these scenarios is calculated as shown in equations (Eq. 22-Eq. 23). The generators run with ramping constraints, shown in Eq. 28Eq. 29 , and the total generation meets the total demand (Eq. 27) in each hour (t). The ramping constraints are expressed in percentage of rated power a generator can ramp up or down in a given hour (% of MW/h). The dispatch model is run for every hour in a year. Ramping constraints for current hour 't' depend upon the electricity generated by the power plant in the previous hour '(t-1)' as shown in equations (Eq. 28Eq. 29). Ramping rates of different types of turbines are shown in Table 7. Maximum electricity generated by a power plant in an hour 't' does not exceed the name plate capacity of the power plant, shown in equation (Eq. 30).

Objective function without carbon tax:

$$minimize\ C_t = \sum_p MC_{tp} * e_{tp} = \sum_p \left(HR * \frac{Price}{1000} + O\&M\right)_{tp} * e_{tp} \quad \text{Eq. 24}$$

Objective function with carbon tax:

$$\text{minimize } \mathrm{C_t} = \sum_{\mathrm{p}} \mathrm{MC_{tp}} * \mathrm{e_{tp}} \quad \text{Eq. 25}$$

$$= \sum_{\mathrm{p}} \left(\mathrm{HR} * \frac{\text{Price}}{1000} + \mathrm{O\&}M + \frac{\text{Carbon Tax}}{1000} * \text{Emission Rate}\right)_{\mathrm{tp}} * \mathrm{e_{tp}}$$

Subject to: Eq. 26

$$\sum_p e_{tp} \geq L_t, \quad \text{Eq. 27}$$

$$e_{tp} \geq e_{(t-1)p} - \frac{RD_p}{100} * P_p, \quad \text{Eq. 28}$$

$$e_{tp} \leq e_{(t-1)p} + \frac{RU_p}{100} * P_p, \quad \text{Eq. 29}$$

$$e_{tp} \leq P_p, \quad \text{Eq. 30}$$

$$e_{tp} > 0 \quad \text{Eq. 31}$$

$$p \leq n \quad \text{Eq. 32}$$

$$t \leq 8760 \quad \text{Eq. 33}$$

Where, Subscript p - Power plant,

Subscript t – Time (in hours),

C_t - cost of electricity generation at hour t (in $),

MC_{tp} - marginal cost of operation of power plant p at hour t ($/MWh)

e_{tp} – electricity generated by power plant at hour t (MWh)

L_t - load demand at t^{th} hour (in MWh), and

n - total number of power plants available for dispatch

RD_p – Ramp down rate of power plant p (% of MW/h)

RU_p – Ramp up rate of power plant p (% of MW/h)

P_p – Nameplate capacity of power plant p (MW)

HR - heat rate (in Btu/kWh),

Price - average spot price of fuel (in $/MMBtu), and

O&M- variable operations and maintenance cost of the power plant (in $/MWh).

Carbon Tax – expressed in $/metric tonne,

Emissions Rate - expressed in kg/MWh.

Table 7. Ramping rates of the electricity generators used in the power plants [39, 76, 77]. The units are percentage change of rated capacity achievable in an hour.

Generator Type	Ramping Rate (% of rated capacity achieved/hour)
Gas Turbine/Combustion Turbine (Natural Gas)	100%
Combined Cycle (Primary Fuel- Natural Gas, Secondary Fuel-Coal)	30%
Steam Turbine (Coal)	15%

Model does not consider the imports of electricity from regions outside of MISO and NYISO. Hourly variations and the resultant power output (e_{tp}) of wind and solar plants for a given location are taken from Eastern Wind Integration dataset [42], and Eastern Solar Integration Dataset [43] respectively.

3.2.3 Energy Storage Model

Using the dispatch model in an iterative storage optimization, I model storage as a revenue-maximizing entity. In other words, I treat storage as an energy arbitrage device used to move bulk energy from low price/demand periods to high price/demand periods. Our treatment of storage applies to operations at utility scale in power networks. Given that a significant percentage (88%) of storage operations in the US are arbitrage-based [11], provision of other grid services from storage, e.g. frequency regulation, is outside of the current scope. Likewise, I do not include Combined Heat and Power (CHP) networks in the modelling effort, as they require more complex coordination between the power and heat portions of the system and are not expected to grow quickly in the US, especially given the competition from cheap natural gas power plants [78]. However, CHP would likely reduce system emissions through much improved efficiency [79]. Further details on the breakdown of all the services provided by storage is provided in the Appendix B section.

Technically, the storage system is described by two parameters: round-trip efficiency and charge rate. Round trip efficiency, set to 80%, is the ratio of energy output from storage against the quantity of energy required to charge it. Charge rate, set at 4 hours in the base-case, reflects how rapidly the storage system can charge and discharge energy, measured here in terms of the minimum time needed for complete charge/discharge. I explore the sensitivity of net emissions on charge rate of storage by varying it from 4-24 hours. A range of charge rate of 4-24 hours is used since 90% of the storage in US is pumped hydro [11] that has charge rates in this range and many emerging bulk storage technologies would have similar charging rates [11]. It is assumed that storage has perfect information of the electricity clearing prices, justified by the fact that most electricity systems forecast electricity prices for the near future up to 48 hours[80]. I model storage operation for capacities ranging between 5% and 20% of the average demand in each of the two systems.

The formulation of storage operation as a price-taker, given perfect information, is a simple maximization problem as shown in Eq. 34.

Objective function:

$$Revenue = max(-\sum_{t=1}^{8760} C_t E_t) \qquad \text{Eq. 34}$$

Subject to:

$$S_0 = \frac{S_{max}}{4} \qquad \text{Eq. 35}$$

$$\forall t, S_t = S_{t-1} + E_t \times \sqrt{\eta} \text{ , } if \ E_t > 0 \qquad \text{Eq. 36}$$

$$\forall t, S_t = S_{t-1} + \frac{E_t}{\sqrt{\eta}} \text{ , } if \ E_t < 0 \qquad \text{Eq. 37}$$

$$\forall t, 0 \leq S_t \leq S_{max} \qquad \text{Eq. 38}$$

$$\forall t, -R \leq \frac{E_t}{hour} \leq +R \qquad \text{Eq. 39}$$

Where, C_t – price of electricity in hour t ($/MWh),

E_t –electricity bought (positive) or sold (negative) by the storage (in MWh),

S_o –initial state of charge of storage (in MWh)

S_t - state of charge in hour t (in MWh)

S_{max}- maximum state of charge of storage (in MWh)

η – Round-trip efficiency of storage

R- Max Charge/discharge rate (in MW)

t- hour in a year (1 to 8,760)

Note that the revenue does not depend on capital cost, as this does not affect optimal operation. In the model, positive E_t indicates energy bought (charging) by the storage, and negative E_t indicates energy sold (discharging). The storage system is assumed to start with a 25% state of charge, given by S_o (in MWh) as shown in equation Eq. 35. S_t (in MWh), the state of charge in each hour, is always less than or equal to the maximum amount of charge attainable by the storage, given by S_{max} in Eq. 35. The round-trip efficiency, η , is equally divided between charge and discharge cycles in Eq. 36Eq. 37 [62]. In any hour, energy in/out (E_t) ranges between the maximum charge/discharge rate, R (in MW) as shown in Eq. 39.

Conventionally, given perfect information, formulation of storage operation as a price taker is a simple profit maximization linear programming problem [62],[80],[81],[82]. However, operation of large amounts of energy storage will influence the market clearing prices and requires a different treatment. I show this in Fig. 13, where I model one week of operation of a 12000 MW storage plant in MISO that takes 4 hours to completely charge without regard for the effect of storage operation on prices. Optimizing storage based on the clearing prices (top) yields an operation schedule (middle) that is subtracted from load and fed back into the dispatch model, determining the actual clearing prices after storage operation (bottom). Due to storage's effect on prices are neglected when calculating storage operation, storage income is 45% lower than expected from the original prices. This illustrates that profit maximization for large storage systems is no longer a simple linear optimization problem and must consider the effect on marginal generation and clearing prices.

Fig. 13 Simulated energy storage operation of 12GW capacity as a price-taker based on clearing prices from an economic dispatch model of Midcontinent ISO (MISO).
The topmost figure shows the simulated clearing prices for a sample week in MISO. The middle figure shows hourly energy storage operation for a plant that ignores its own effect on prices, where positive values indicate charging of storage and negative value indicates discharge. The bottom figure shows clearing prices after the effect of storage on net load has been taken into account. In this scenario, the storage expects to make a revenue of $1.5 million based on the topmost prices but makes only $85,000, a 45% reduction resulting from the non-marginal effect of storage operation on prices.

Modeling of large energy storage as a price-maker is achieved using a self-learning optimization technique. The flowchart of this method is provided in Fig. 14. This is done by considering the moving average of the hourly storage charge/discharge at the end of each iteration until the solution converges. The storage operation converges/remains consistent after about 20 iterations, which I use as the number of iterations for estimating an optimized solution, maximizing revenue given the effect that storage has on prices. The convergence of this process is illustrated in Fig. 15.

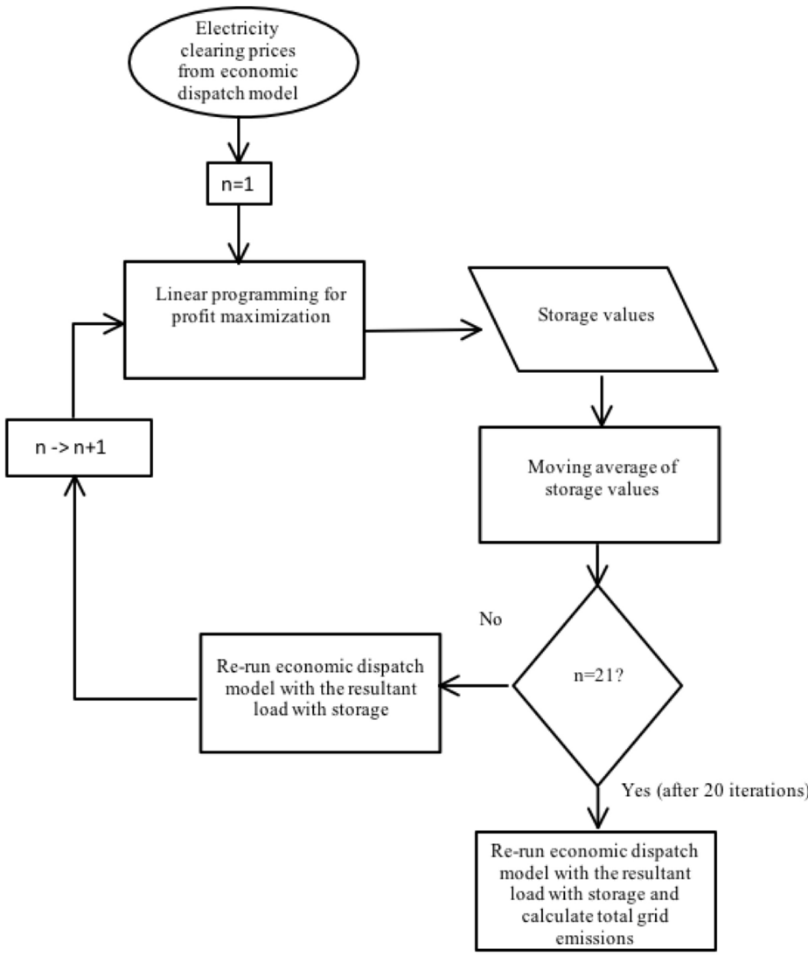

Fig. 14 Flowchart of methodology for modeling energy storage as a price maker.
Market clearing prices are estimated, and the storage operation is generated using linear optimization, maximizing revenue based on the clearing prices. The change in demand pattern due to storage is then taken into account and a new time-series of prices are produced, which are used to re-calculate storage operation. This process is iteratively performed 20 times and the moving average of storage values is taken after end of each iteration. At the end of the cycle, the resultant load from additional storage is taken as the storage operation. Note: I choose 20 as the maximum number of iterations, as the storage operation remains consistent (converges) after this point.

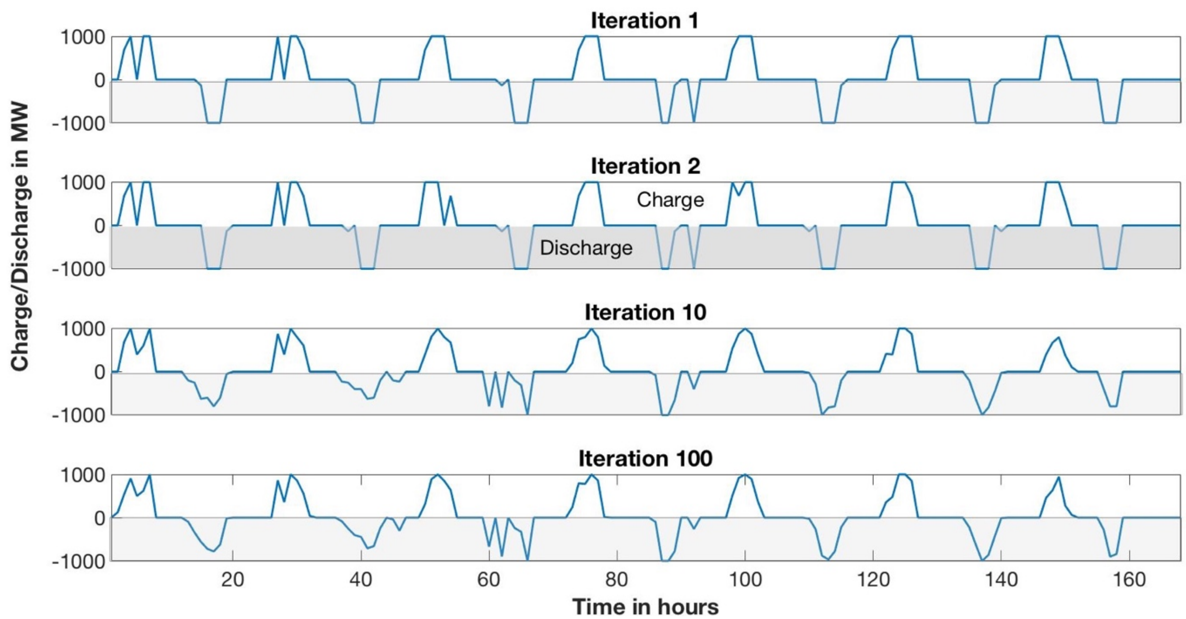

Fig. 15 Output from iterative optimization of storage operation.
Between iterations, the effect of storage operation on prices is considered and the storage adjusts accordingly to ensure that it is maximizing revenue while taking its own effect on prices into account. The storage operation remains consistent (converges) after approximately 20 iterations, which I use as the number of iterations.

3.3 Emissions model

The total annual CO_{2eq}. emissions (in million metric tonnes) from the grid are calculated based on the hourly dispatch of plants as shown in equations (19-20). The plant-level emission rates are in metric tonnes/MWh, taken from the eGRID database [40]. Total CO_{2eq}. emissions are comprised of all greenhouse gas emissions measured on a common scale based on their Global Warming Potential (GWP) relative to CO_2 [54].

The total CO_{2eq} emissions in a given hour for a given operation schedule of generator plants is given by Eq. 40:

$$em_t = \sum_p m_p * E_p * N_{pt}, \quad t = 1,2 \ldots ,8760 \qquad \text{Eq. 40}$$

$$Total\ Emissions = \sum_{t=1}^{8760} em_t \qquad \text{Eq. 41}$$

Where, em_t – total emissions from all operating plants in hour t (metric tonnes)

m_p – emissions of plant p per unit of produced electricity (in metric tonnes/MWh)

E_p – electricity generated by plant p in one hour of operation (MWh)

N_{pt} = 1 when power plant p is operating in hour t, else the value is zero

t- hour in a year (1 to 8,760),

p – index number for power plant

Summing em_t over 8,760 hours in a year gives the annual CO_{2eq} emissions for the grid as shown in Eq. 41.

To identify the point where adding storage becomes carbon neutral for the MISO grid (in NYISO, it is initially emission-reducing), wind and/or solar capacity is incrementally added. These additions are in compounded incremental additions of 10% until the difference between emissions with storage and without storage are zero. The hourly generation profiles of solar and wind energy across various locations in MISO was estimated according to the Wind Integration National Database (WIND) toolkit [42] and Eastern Solar Integration Data [43].
The WIND Toolkit provides data related to wind energy production for over 126,000 current and potential locations across the United States for 7 years from 2007–2013 [42]. This dataset consists of meteorological data, 5-min resolution of wind power production, and capacity factors. I considered 30 potential locations in the Midwest region and the corresponding hourly wind output/MW. The average wind energy output (kWh/hour) for a 1kW system across these locations is used to generate the hourly variations of incremental wind capacities considered in the study. Similarly, the Eastern Solar Integration dataset by NREL consist of 5-minute solar power and hourly day-ahead forecasts for approximately 6,000 simulated PV plants. 30 potential sites from 15 states in the Midwest region are considered and a similar procedure to wind energy output is used to generate solar energy output/hour. Annual capacity factors of most of the potential wind power sites in MISO are greater than 40% and most of the solar power sites are greater than 16%. More details on the hourly variation of solar/wind energy output/hour and potential locations considered are provided in the Appendix B section.

3.4 Results

I first discuss differences between NYISO and MISO grids by showing the hourly mix of generation sources for a typical day for different seasons: winter, spring, and summer. Autumn is not included in the seasons as the demand during this season is similar to that observed in the spring. For illustration, I identified the type of power plants supplying electricity each hour to meet the given load. As seen in Fig. 16, the marginal generators in MISO are a mix of coal and natural gas power plants, though most often coal. Daily coal-based generation is about 30-43% of the total depending upon the season, highest during the summer. Daily natural gas generation is about 25% of the total generation during all the seasons. Summer peaks in MISO are met by coal and natural gas-based plants. As the marginal cost of power from natural gas is close to that of coal in recent years, the dispatch of coal and natural gas power plants are intermixed in the dispatch stack.

In MISO, energy storage tends to increase the emissions if it charges from coal, consequently offsetting cleaner natural gas plants while discharging, which can be seen on a sample summer day in Fig. 17. In the same figure, when there is an increased quantity of wind energy (2-4GWh/75GW capacity), storage tends to charge using efficient combined cycle natural gas power plants, while displacing inefficient natural gas plants while discharging. In the second case, storage does not necessarily discharge only during peak periods because of the large difference in electricity prices when wind energy output is high versus low. However, lower natural gas prices mean that storage is also likely to charge from an efficient combined cycle natural gas plant, even without increased quantity of wind energy.

The grid in NYISO has natural gas plants as the marginal supplier of electricity when the demand peaks during the summer and winter. During spring and autumn seasons, nuclear energy or natural gas is the marginal generator during the off-peak period, and natural gas is the marginal generator during peak period of the day. In Fig.5 below, during the off-peak period, a mix of biomass and natural gas power plants are on the margin on a sample day taken during the spring season. In NYISO, energy storage tends to increase the usage of more efficient natural gas power plants or nuclear power plants (by charging during off-peak periods), consequently offsetting less efficient natural gas plants while discharging.

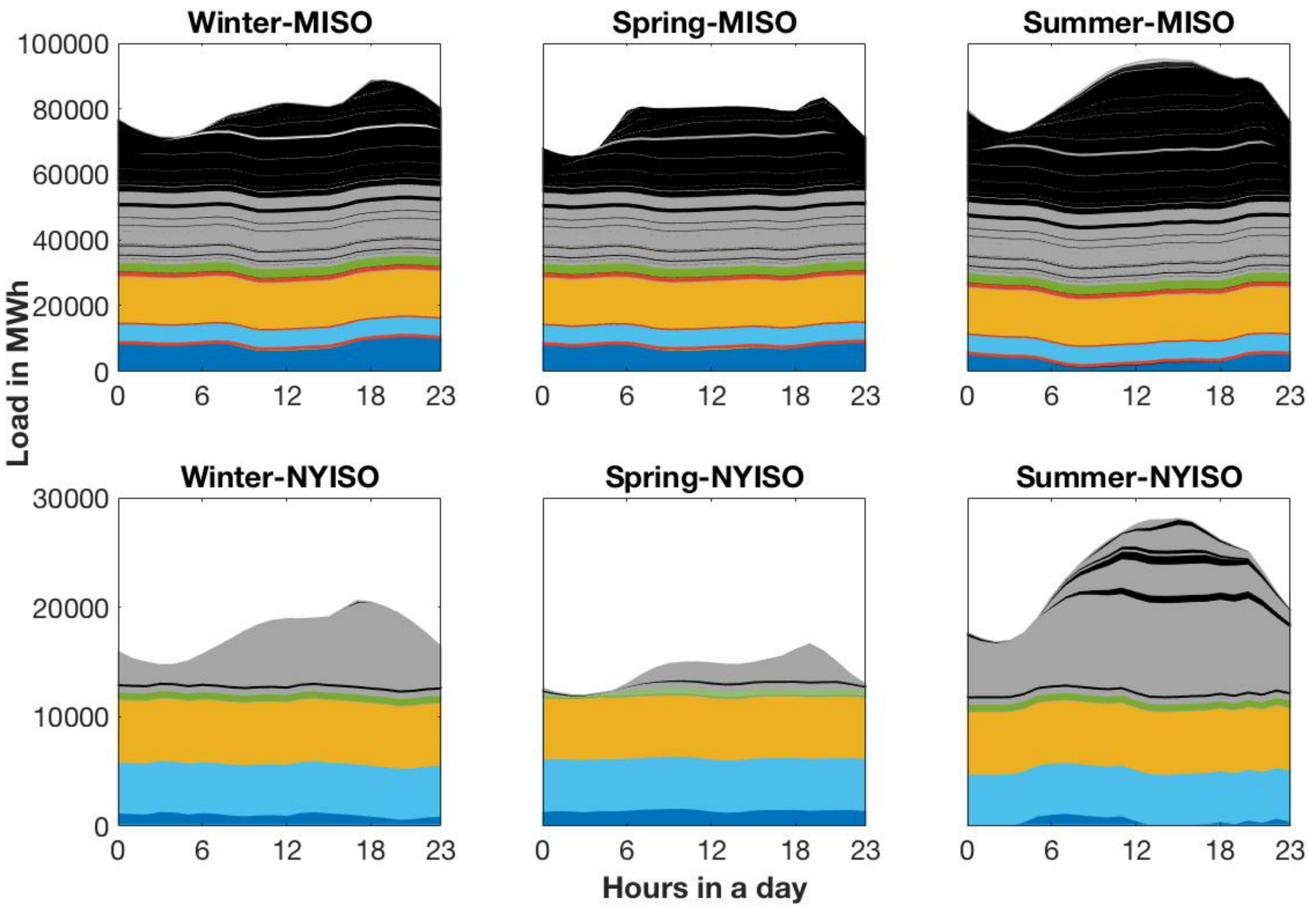

Fig. 16 Grid mix based on hourly dispatch of generators in Midcontinent ISO (MISO, top) and New York ISO (NYISO, bottom) on a sample day in different seasons.
Each color band indicates the typical type of fuel of the dispatched power plants. A generator is said to be "on the margin" if that type of generator is last to be dispatched (ie, at the top of the stack during a given hour). In the MISO region, coal/natural gas is normally on the margin during off-peak periods, and peak periods. In the NYISO region, natural gas/nuclear energy is on the margin during off-peak and natural gas is on the margin during peak periods. The effect of ramping constraints can be seen by the fact that some generators have non-horizontal bands, indicating a binding constraint in ramp-up or ramp-down.

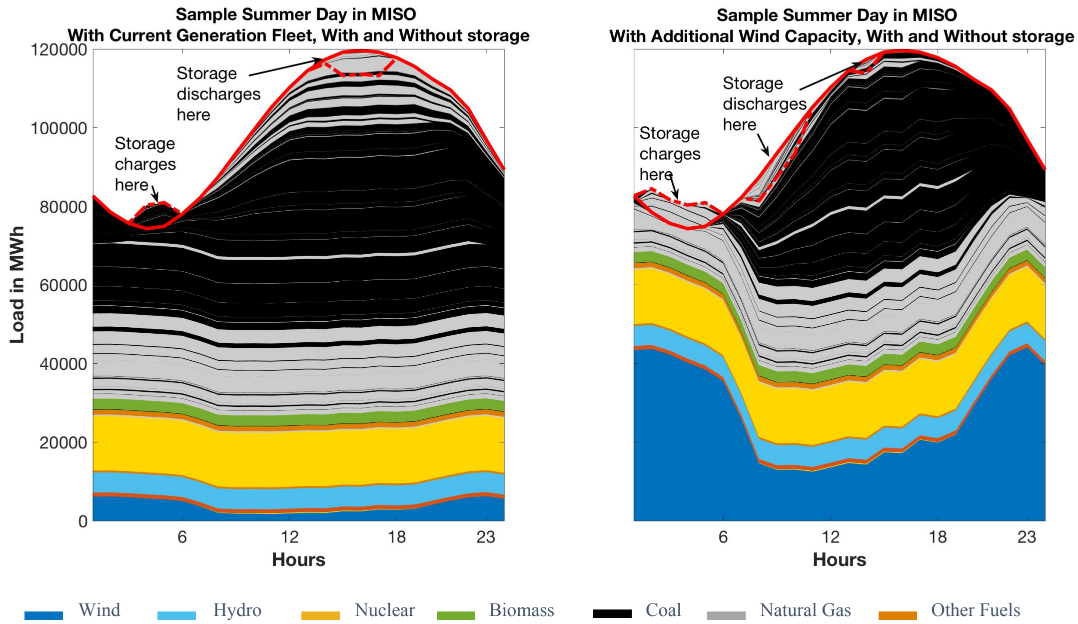

Fig. 17 Grid mix based on hourly dispatch of generators in Midcontinent ISO (MISO) on a sample day during summer, with and without storage – with current generation fleet, and with additional wind energy (output from economic dispatch model).

Each color band indicates the typical fuel of the dispatched power plants. The dotted outline represents the total load with storage operations, and the solid outline represents the total without storage operation. On the left, without additional wind energy, storage charges using additional coal energy at night and displaces natural gas plants during peak periods, thereby increasing emissions. On the same day on the right, with wind capacity of 75GW producing 2-4GWh of wind energy on a hot summer day, storage charges from more efficient combined cycle natural gas plants and displaces natural gas peaker plants.

3.4.1 Emissions from storage operation in NYISO and MISO

I model generator and storage operation and resulting $CO_{2eq.}$ emissions for addition of storage between 5-20% of the average load, which is 3,000MW-12,000MW in MISO and 1,000MW-4,000MW in NYISO regions. Fig. 18 shows storage-induced emissions for different storage charge rates between 4 and 24 hours. Total grid emissions induced by storage are sensitive to change in charge rates and round-trip efficiency of the storage. In the MISO region, storage-induced emissions increase both with increases in storage capacities and with more rapid charging rates. The annual emissions due to storage additions vary between 11,000 and 65,000 metric tonnes of $CO_{2eq.}$, depending upon the storage capacity and the charge rate.

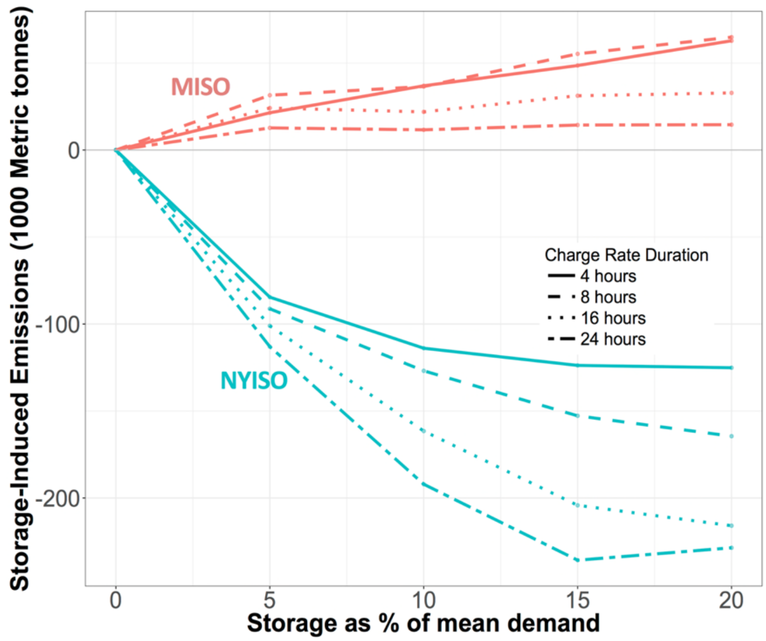

Fig. 18 Annual storage-induced emissions in New York ISO (NYISO) and Midcontinent ISO (MISO). The storage capacities are varied between 0%-20% of the average demand, which is 3,000MW-12,000MW in MISO, and 1,000MW-4,000MW in NYISO. The estimated storage-induced emissions are given in metric kilo tonnes of CO_{2eq}.

Fig. 19 illustrates why storage increases emissions in MISO, showing the emissions of each generator when in an economic dispatch order. Emissions increase when storage charges at night using coal plants and displaces natural gas power plants while discharging. The relatively high emissions of coal versus natural gas implies that storage operation, in this case, is increasing carbon emissions from the grid. In addition, because storage is a net consumer of electricity (due to losses), total electricity generation is increased in proportion to the quantity and operation of storage.

In the NYISO region, system emissions decrease with increases in storage capacities, and are sensitive to change in charge rates of the storage (Fig. 18). The annual net emissions due to storage additions range between -91,000 to -235,000 metric tonnes of $CO_{2eq.}$ for charge rates between 4 and 24 hours. Prior results by Hittinger and Azevedo [62] using a Marginal Emissions Factor method suggested small increases in emissions in NYISO as a result of storage additions. Performing sensitivity analysis with natural gas prices show that positive emissions in NYISO would be expected with higher natural gas prices, shown in Fig. 22 in the later sections, which was the case during the 2010-2012 period over which the Hittinger and Azevedo model operated.

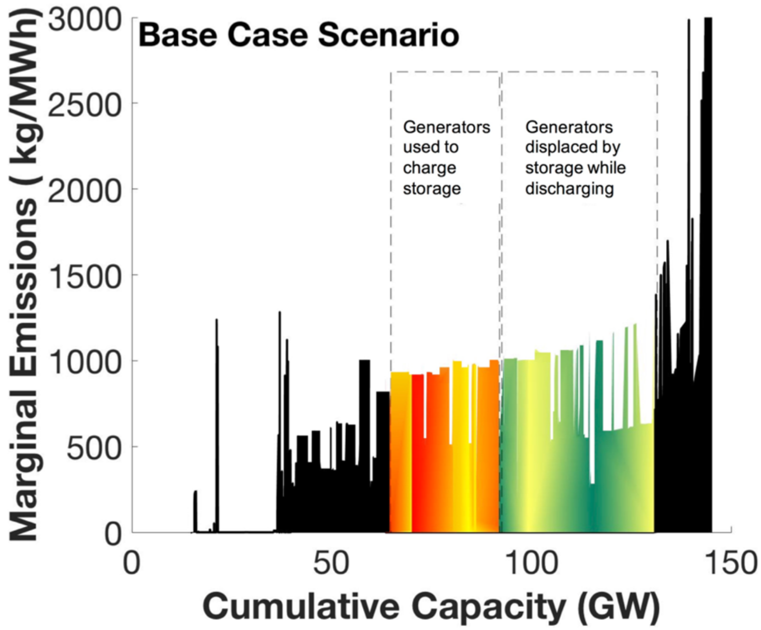

Fig. 19 Marginal emissions of Midcontinent ISO (MISO) generators, in order of economic dispatch. The graph shows a series of rectangles, with height equal to the plant's emissions rate and width equal to the capacity (GW) of the generator. The color gradient indicates the generators that are most often used to charge/discharge storage, over one year of operation (lighter color indicates more frequent use), shades of red refer to charging, shades of green refer to discharging. In MISO, storage-induced emissions increase due to charging when the marginal emissions of the grid are high (coal) and displaces generators that have lower emissions (natural gas).

3.4.2 Solar and wind capacity additions required to make storage carbon neutral

As of 2016, MISO has 14 GW of installed wind capacity and 290 MW of solar capacity [40],[83]. Results in Fig. 20 show that 20.6 GW of wind power (without solar) or 22.6 GW of solar (without wind) is estimated to be required in MISO before deployment of large storage (3000 MW/12,000MWh) results in zero increase in emissions. Modelling energy storage as a price taker slightly over-estimates these results to 23.5 GW of wind power or 29 GW of solar power. A lower quantity of wind is required (than solar) because of the higher availability: a 40% annual capacity factor compared to 16% annual capacity factor of solar in MISO.

Could the playing field be changed so that storage delivers more carbon benefits to the grid? One approach would be to abandon the economic arbitrage approach altogether, i.e. operating the storage for environmental rather than economic benefits. I do not consider this option here. Instead, I look at another approach consistent with the current economic operation of energy markets: addition of a carbon tax. A carbon tax increases the marginal cost of dirtier coal-based power plants, thus motivating cleaner power plants to operate more often (lower in the dispatch order), leading to decreasing storage-induced emissions. I thus investigate how a carbon tax would affect the total amount of wind/solar energy required to make net-zero emissions with bulk energy storage. I analyze two carbon taxes - \$20/metric tonne and \$30/metric tonne of $CO_{2eq.}$ emissions - using values consistent with valuations of the social cost of carbon from the EPA [84].

The amount of wind and solar capacities required in the grid reduces from around 18% of the grid generation capacity with no carbon tax to around 12% at \$20/metric tonne of $CO_{2eq.}$ emissions, or to 10% (the current capacity of wind+solar in MISO) at \$30/metric tonne of $CO_{2eq.}$ emissions.

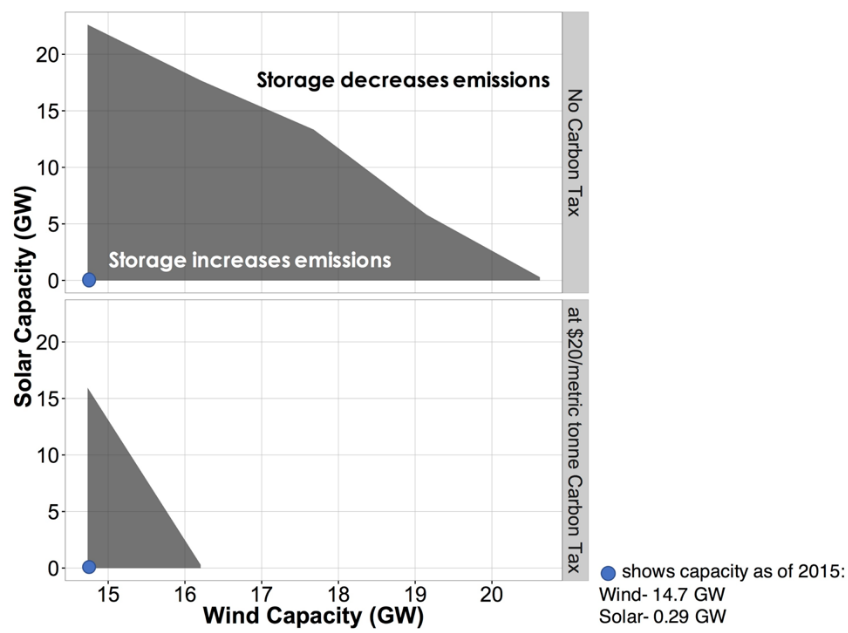

Fig. 20 Quantity of wind and solar required before storage-induced emissions are negative in MISO, at two different carbon taxes.
The figure compares between no carbon tax, and carbon taxes of \$20/metric tonne. Because of the effect on the dispatch order of generators, a carbon tax results in much lower quantities of required wind/solar to make storage emissions-reducing.

3.4.3 Emission factors of storage with addition of solar and wind

While the above results indicate that adding storage to the current MISO grid increases carbon emissions, it is not yet clear if these emissions are high or low compared to other generation sources. It is common to measure the carbon impacts of electricity generation technologies with emission factors, i.e. the carbon emitted per quantity of electricity produced, e.g. 400-500 kg/MWh for combined cycle gas plants [85] and 980-1300 kg/MWh for coal plants [40, 85]. To compare the storage-induced emissions with those from other electricity sources, I calculate an emission factor equal to the additional emissions induced by storage (kg $CO_{2eq.}$) divided by the total energy (MWh) delivered by storage to the grid. This storage-induced emissions factor is the normalized emissions associated with moving energy from off-peak to peak periods, and accounts for both storage losses and differences in peak and off-peak emissions rates.

The emissions due to storage operation in MISO are approximately 450 kg/MWh for 4-8 hours' charge rates and decrease as wind and solar are added. Storage –induced emissions factors for different wind and solar capacities is shown in the Appendix B section.

3.4.4 Sensitivity analysis- price-taker modeling versus the price-maker modeling

When the storage's effect on prices are neglected, arbitrage income is 45% lower than the expected from the original prices (Fig. 13). Therefore, profit maximization for large storage systems must consider the effect on marginal generation and clearing prices. When I ignore the effect on clearing prices and storage is optimized using a simple linear optimization, the model overestimates the total storage induced emissions by an average of 70% as compared to the price-maker model, depending upon the storage capacity. Price-taker modeling uses and dumps large capacities of the storage all at once, depending on the price signals. Therefore, a high surge in demand while charging and a steep decline in demand while discharging increases the overall emissions. During the charge phase, a surge in demand is met by the additional power plants, creating an increase in emissions. During the discharge phase, the demand differential increases immediately after the storage discharges all at once leading to an increase in peaker plants' generation to meet the steep gradient of change in demand, thereby increasing the emissions. Fig. 21 shows the difference in the estimation of emissions between the price-maker and price-taker modeling of the storage operation for 4 hours charge duration of the storage and 80% roundtrip efficiency in MISO and NYISO regions. In this case, storage induced emissions are 70% more than the price-maker modeling in MISO, and 35% more in NYISO region.

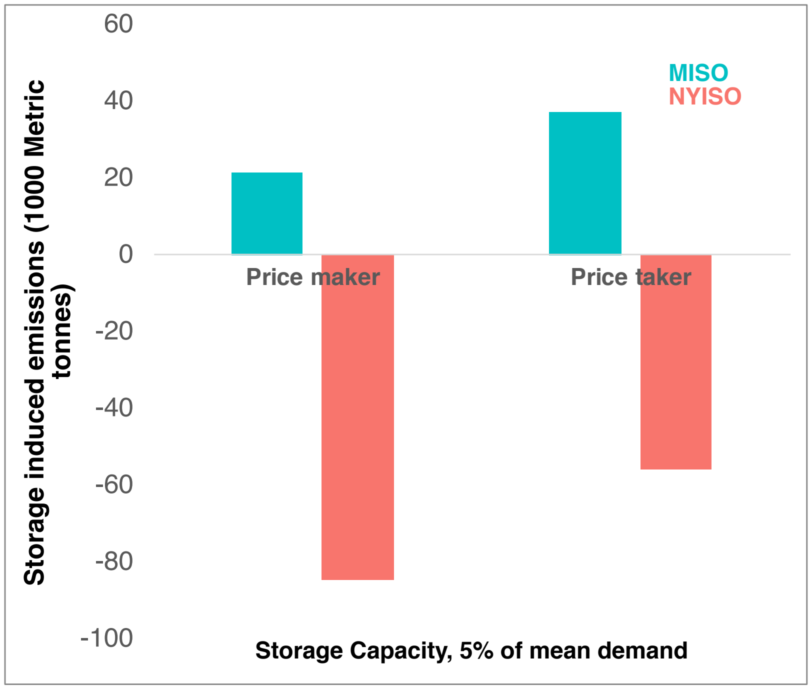

Fig. 21 Difference in annual storage induced emissions when large capacities of storage are considered a price-taker instead of price-maker in New York ISO (NYISO) and Midcontinent ISO (MISO).
The storage capacity considered is 5% of the average demand, which is 12GW in MISO, and 4GW in NYISO. The estimated storage-induced emissions are given in metric kilo tonnes of CO_{2eq}. The colors of the bars indicate the ISOs. Price-taker scenario over-estimates the annual storage induced emissions in MISO by 75% and in NYISO by 35%.

3.4.5 Sensitivity analysis- high natural gas prices

I perform sensitivity analysis for natural gas price by re-calculating all results at a higher price as compared to the base case natural gas price. From 2000-2016, the price of natural gas has varied between $2.5 and $11 per MMBtu, with an average projected value of natural gas price to be at $5/MMBtu till 2040 [86]. The base-case natural gas price used for the results above is the 2015-2016 natural gas price at $2.6/MMBtu and that is increased to $5/MMBtu in the high natural gas price scenario.

As seen in Fig. 22, at a higher natural gas price, total grid emissions induced by storage increase both in NYISO and MISO region at different storage charge rates between 4 and 24 hours. In the MISO region, storage-induced emissions increase by 40-60 times as compared to the base-case scenario, both with increase in storage capacities and with slower charging rates. The annual emissions due to storage additions vary between 800,000 and 4,000,000 metric tonnes of $CO_{2eq.}$, depending upon the storage capacity and charge rate.

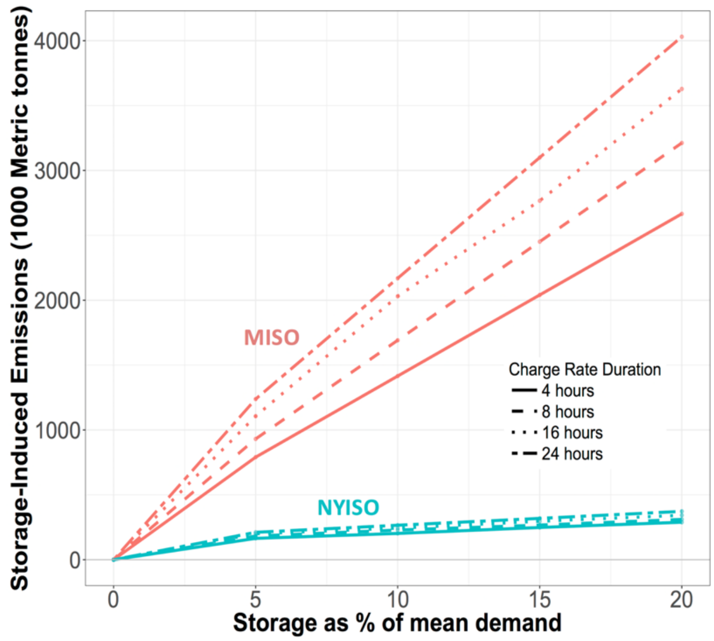

Fig. 22 Net change in annual emissions after adding storage in New York ISO (NYISO) and Midcontinent ISO (MISO) at a higher natural gas price of $5/MMBtu (compare with Fig. 6 showing base-case scenario of $2.6/MMBtu).
The storage capacities are varied between 5%-20% of the average demand, which is 3,000-12,000MW in MISO, and 1,000-4,000MW in NYISO. The estimated emissions are given in 1000 metric tonnes of $CO_{2eq.}$

Fig. 23 illustrates why MISO's storage-induced emissions increase by about 50 times for the high natural gas price scenario as compared to the base-case. In the high natural gas price scenario, the coal and natural gas are 'better sorted', and charging is almost completely met with coal generation while discharging almost universally displaces natural gas generation.
In the NYISO region, system emissions after adding storage increase between 150,000 and 370,000 metric tonnes of $CO_{2eq.}$ depending upon the charge rate and the storage capacity in the high natural gas price scenario. This contrasts with what is seen in the base-case scenario. Though NYISO has a very small percentage of coal based power plants (3%) [40], these plants are available on-margin when the storage charges during off-peak periods as compared to gas turbines during the peak period.

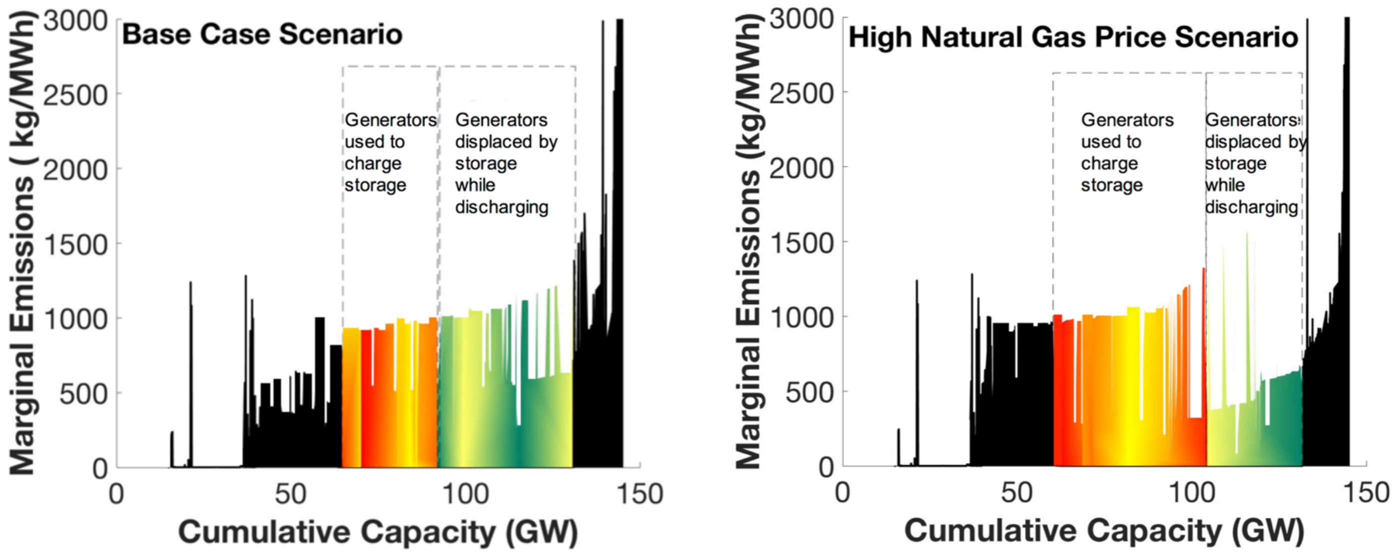

Fig. 23 Comparison of marginal emissions of Midcontinent ISO (MISO) generators, in the order of economic dispatch for the base case scenario and high natural gas price scenario.
This is essentially a series of rectangles, with height equal to the plant's emissions rate and width equal to the capacity (GW) of the generator. The color gradient indicates the generators that are most often used to charge/discharge, over one year of operation (lighter color indicates more frequent use). With high natural gas prices, storage is seen to increase emissions by 40 times since the difference between marginal emissions during charge-discharge phase are much higher than the base-case scenario.

The amount of wind and solar capacities required to de-carbonize the emissions induced by storage increase from 18% in the base-case scenario to about 35% of the grid mix at high natural gas prices (assuming no carbon tax), as shown in
Fig. 24 . At a carbon tax of $20/metric tonne of $CO_{2eq.}$ emissions, the amount of wind and solar required is 31% and at $30/metric tonne of $CO_{2eq.}$ emissions, it is 25% of the grid mix.

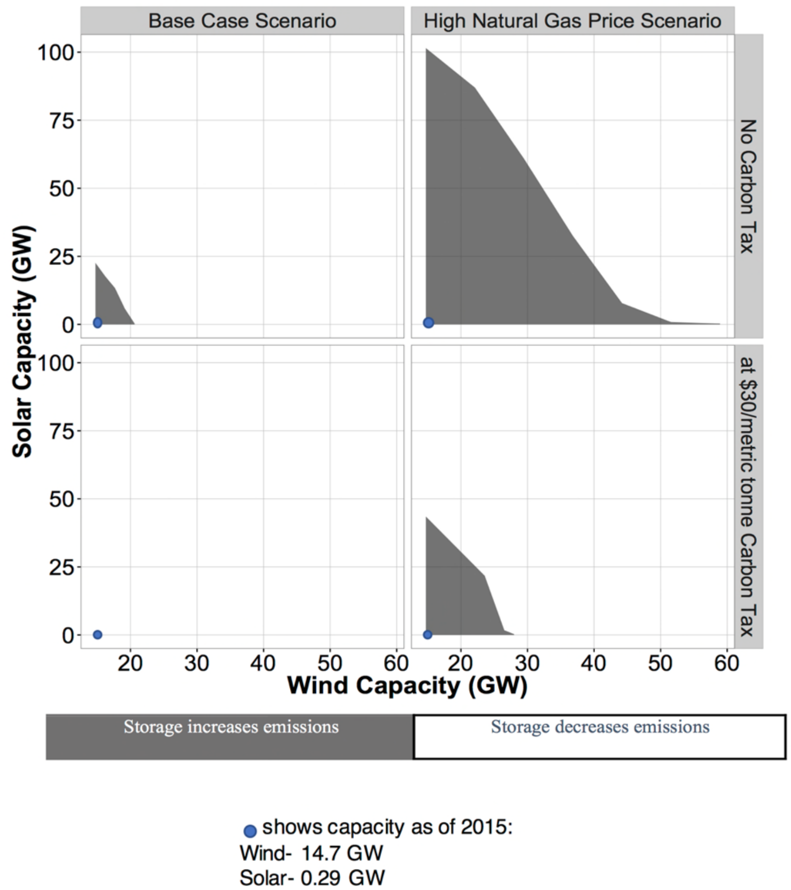

Fig. 24 Quantity of wind and solar required before storage-induced emissions are negative in MISO in the base-case scenario (at $2.6/MMBtu) and high natural gas price scenario (at $5/MMBtu).
The figure compares between no carbon tax, and carbon taxes of $30/metric tonne for both the scenarios.

3.5 Contribution to the literature and discussion

My contribution to the current chapter is developing a new modeling approach, considering the effects of large energy storage as a price-maker on the current electricity grids and estimating the change in emissions as intermittent renewables (wind and solar) are added to a grid and as natural gas prices vary. Also, this is the first study to examine how much wind and solar are needed to negate the emission affects from the arbitrage operation of the storage as a price-maker in a coal-heavy grid.

Most of the studies on energy storage till date assume bulk storage as a price-taker and ignore the effect of market electricity prices on storage. Results show that price-taker modeling approach does not consider the substantial alterations in demand patterns, prices, and dispatched generation, resulting in over-estimation of emissions by 70% from the large changes in the demand before and after storage operation. Also, the sub-optimal arbitrage operation results in 45% lower revenues to the storage. Though, price-maker approach is considered in some of the models in prior research, they either model assuming the storage operator to perfectly anticipate the price-load relationship _ or co-optimize the storage operation with the dispatch model [88]. The limitation of the first approach is that the representative functions between demand and price are challenging to construct and are not dynamic to the different market situations [88]. In the second case, co-optimizing the storage operation with the dispatch model is a robust approach but most of the planners and the storage operators have access to sophisticated dispatch models from the third-party vendors which are complicated to integrate any new storage additions into the model. This study integrates storage operation as a price-maker by using an iterative dispatch model in combination with the linear optimization of the storage. This method makes it easier to integrate the storage operation as a price-maker into the current or any sophisticated dispatch models. The approach developed in this work is the first method that allows faster and flexible integration of storage into any dispatch models and also attracted the storage operators after the publication of this model.

Results show that at today's low natural gas prices and grid mix, I find that energy storage operated under economic arbitrage reduces carbon emissions in NYISO and increases emissions in MISO. At higher gas prices, storage increases emissions in both NYISO and MISO by enabling coal power to substitute for natural gas. Emissions changes induced by storage are much larger in the higher natural gas scenario due to large-scale substitution of coal for natural gas. This implies that a rising natural gas price relative to 2015-2016 prices, as per EIA projections [86], could mean that grids dominated by coal may not see carbon benefits from storage without significant restructuring of their generation mix. For example, though wind capacity in MISO is steadily growing and is projected to reach 20GW of installed capacity by 2019 [89] (a growth rate of ~1GW/year), it is likely to be at least a decade before wind and solar capacities in total achieve 35% of the generation and induce storage emissions benefits under a high natural gas price scenario. The effect of natural gas price leaves the economic and environmental effects of storage at odds: I have found that storage-induced emissions that are zero or negative depend on the currently-low natural gas prices. However, storage providing energy arbitrage only makes financial sense if natural gas becomes more expensive, in which case energy storage will induce greater use of coal generation, increasing system emissions [90].

These results clarify the option space society faces with regards to storage. One choice is to accept increased carbon emissions in the short-term in some grids in order to achieve longer term benefits after more renewable energy is adopted. Arguments could be made justifying such a long-term perspective, but the current policy discourse does not frame the choice as a long-term one, instead asserting that storage delivers immediate benefits. Another option is to change the operation of storage to achieve environmental goals. For example, a storage system could be directly tied to a renewable generation plant to address intermittency. While technically possible, it is important to clarify the economic and environmental benefits delivered compared with alternative means of addressing intermittency, e.g. via flexible natural gas plants or improved transmission interconnection. A third option is to shift the economic context in which storage (and the grid) operates, e.g. a carbon tax, to ensure carbon benefits. I do not explore the benefits and cost of these three options here but assert there is a need for a clearer framing of societal expectations from storage.

Our study shows that levying a carbon tax could significantly reduce the solar/wind requirements before storage delivers carbon benefits. These requirements are largely dependent on the level of the tax. At current social cost of about $30/tonne, emission benefits from storage are plausible with the current installed wind capacity in MISO. On the other hand, if the natural gas prices start to increase, achieving wind capacities of 28 GW could take at least 10 years in MISO (though perhaps accelerated if such a carbon price is implemented), considering the current rate of projected growth in the absence of the Clean Power Plan [89]. Therefore, a reasonable carbon tax, set near the US EPA estimated social cost of carbon, without any support of other policies (such as the Clean Power Plan) would allow storage to deliver intended carbon benefits for MISO into the foreseeable future.

There are some encouraging outcomes in these results for energy storage and emissions. First, our analysis for NYISO illustrates that storage can be neutral or beneficial for emissions when it is routinely charged, during off-peak periods, with efficient combined cycle natural gas generation. In New York and similar grids (such as California), while storage is not expected to directly reduce emissions, it may indirectly mitigate carbon through improved integration of intermittent renewables. Second, the analysis for MISO illustrated that emissions due to storage additions is less related to the *quantity* of renewable energy in a system than the *curtailment* of renewable energy in a system. But the results show that any grid that has wind/solar curtailment of sufficient scale to be the primary source of charging energy for storage would experience excellent emissions benefits from storage. However, unless storage is predominantly charged with otherwise-curtailed renewable energy, its emissions benefits are likely to be neutral or negative.

Chapter 4: How Does Energy Storage Affect the Generation and Profit of Existing Generation Technologies?

Abstract

This work models the effects of economic operation of the energy storage on the generation and profits of the existing generation technologies. In this study, I first investigate how storage affects the likely generation of other fuels based on historical electricity prices and generation mix for the year 2016 in 22 eGRID subregions. In this case the storage is modeled as a price-taker, and its actions do not affect the market prices. To capture the change in profit, in the second part, I run a dispatch model with larger storage capacities up to 5GW in the New York ISO (NYISO), Midcontinent ISO (MISO), and California ISO (CAISO). In this case, storage distorts the current prices and is no longer a marginal operator in the system. First part captures the dynamics of the actual dispatch of current systems, and the second part captures the expected changes in the profit when storage to noticeably shift prices and dispatch. In the west and the Midwest region, storage increases the coal generation and decreases gas generation; in the eastern US region- storage operation increases the gas generation. In California, upstate New York and New England regions, storage increases the gas generation and decreases the coal generation. Second, when large capacities of storage act as price-makers, natural gas peaker plants lose the most profit in all the regions by more than 10%. Profits of coal increase in the Midwest region and profits of solar increase in California by 5%. Profits of all the generating units decrease in New York as natural gas offset by the storage most of the time sets the electricity prices.

4.1 Introduction and Literature Review

Energy storage comes with a plethora of benefits to electricity grids such as storing surplus electricity, providing reliability, stability, and emergency backup, enabling renewable integration, and demand-side management [91]. With its many advantages, policymakers are pushing for implementing energy storage as a comprehensive answer to not only improvise the quality of grid systems but also address the environmental challenges by enabling substantial renewable integration [13, 92]. Renewables are critical towards the decarbonization of the electricity sector, and it is essential to answer if operational strategies of energy storage effectively benefit renewables. Deployment of bulk energy storage system affects the system-wide costs, flexibility, generator operations, and GHG emissions depending upon the grid mix [20, 62, 80]. Therefore, before implementing policies or regulations for storage to support renewables, I must assess if storage is a net benefit/cost to society [93]. There is prior literature that argues otherwise that the economic operation of the storage has an impact on the emissions _, depending upon the grid mix across the U.S. But how the storage operation manifests the generation, and the profits of the other generating units, especially renewables across different regions in the U.S. have not been investigated. This chapter discerns the effects of storage on the profits and generation of renewables and other generating units from adding storage.

Economical operation of the storage reduces the wholesale electricity prices and also the generation of the peaker plants. Because of this, it reduces the profitability of all the generators and increases the generation of base-load power plants [80, 88]. A prior study by Denholm et. al. [95] shows that adding bulk energy storage increases the generation of the base-load marginal power plants such as coal and combined cycle units by 0.6% while decreasing the generation from the combustion turbines by about 1.5%. The literature on quantifying these effects of storage on the generation units across different regions and different grid mixes is scarce. This study investigates the effects of economic operation of storage on the net change in generation across 22 eGRID regions of the U.S, and the change in profits in three distinct grid mixes. Since the capacities of storage should be significant to affect the profitability of the other generating units, storage is considered a price-maker for determining the net change in profits.

This study is conducted in two parts to investigate the change in generation and the profits of the other generating units.

In the first part of the chapter, this study investigates how storage affects the likely generation from other fuels based on the actual electricity prices obtained from Independent System Operators (ISOs) and the current fuel mix in the system. It captures the dynamics of the actual operation but does not show the change in price or profit of the other generators as it is a marginally small capacity.

To answer the change in price/profit of other generators, this study uses a simulated dispatch model with larger storage capacities until 5GW capacities in the second part of the chapter. In this case, storage distorts the prevailing prices and is no more a marginal operator in the system. Also, the profit impacts from the storage could result in retirement of some of the power plants losing the most profit [80, 88], which is captured in the current study to account for change in market conditions from the integration of large storage systems.

Overall, the first part of the model captures the dynamics of the actual operation, and the second part captures the changes in the profit when storage is no more a marginal operator.

For modeling the first part, the study uses the actual electricity prices and the probability of marginal generators operating ('marginal generator technology factors') at a given time from 22 different eGRID regions. A linear programming model is used to optimize the storage operation from the clearing prices and the marginal generation factors provide information on the type of generators operating or displaced from storage operation.

In the second part - I calculate the change in profits with and without storage until 5GW capacities for three different regions- Midcontinent Independent System Operator (MISO), New York Independent System Operator (NYISO), and California Independent System Operator (CAISO). The storage operation as a price-maker is simulated using an iterative optimization of a storage operation with the dispatch model as described in section 3.2.3.

4.2 Method- Effect on the generation

In the first part, to study the implications of the storage operation on other generators, I use the actual clearing prices and the marginal generator factors to estimate the fuel mix used and displaced from the storage operation. The data sources for the actual prices across the 22 eGRID regions is provided in the Appendix C.

4.2.1 Marginal Generator Factor

"Marginal Generator Technology Factors" ('MF') broken by eGRID sub-region, season, and hour of the day, is likelihood of a marginal generator type operating at a given hour. For example, in MISO, during a typical summer day, a marginal increase in demand by 1MW during noon by storage could 85% likely come from coal, and 15% likely come from natural gas combined cycle power plant. The likelihood of a generator operating at a given time depends upon the fuel mix at that given hour, and the electricity clearing prices, which are published online in real-time by most of the ISOs [70, 96] in different eGRID regions.

To determine the marginal generator for the various subregions in any given hour, the study uses data from the EPA's Continuous Emissions Monitoring System (CEMS) for 2016 [97]. CEMS provides hourly emissions and generation from all thermal generating units greater than 25 MW, as well as data on primary fuel input. From the CEMS data, I aggregate to the plant level and build a new dataset tracking the change in generation by plant between one hour and the next. I then select only plants with more than a 5 MW increase or decrease in generation between any given two hours—these plants are said to be "on the margin" for each hour. Using this subset of marginal plants, I then aggregate plants by fuel class to determine the net amount of generation increase/decrease as shown in Eq. 42. Since there is an equal likelihood that the storage could charge or displace generators that are coming online or going offline, MF of each fuel *(f)* in a given hour *(t)* is calculated from the absolute values of change in generation by fuel (Eq. 43).

$$net\ generation_{f,t} = \sum_{p} increase_{p,f,t} - \sum_{p} decrease_{p,f,t} \qquad \text{Eq. 42}$$

$$MF_{f,t} = \frac{|net\ generation_{f,t}|}{\sum_{fuel}|net\ generation|} \qquad \text{Eq. 43}$$

Where: f- fuel type

t- hour

p- individual power plant

net generation – net generation on margin (MWh)

increase – Power plants increasing generation (MWh)

decrease – Power plants decreasing generation (MWh)

This aggregation provides us with the total change in generation in each hour (both increase and decrease) as well as the percentage of that change coming from coal, natural gas, biomass, and oil in each hour (MF).

There are some limitations with this data collection method. First, this data does not capture changes in generation from non-fossil generation, such as nuclear or renewable sources. In order to account for changes from renewables, I assume they are on margin when the electricity prices are almost zero or negative. In general, renewables are less likely to be on the margin, for example- renewables were on margin in California for 2% of the total hours in 2016 [98]. In addition, by assuming that the marginal generation in a given hour is a mix of the plants changing generation in that hour, I neglect plant operating constraints or physical system limitations which might otherwise could enable a power plant constantly operating without change as the marginal generator.

4.2.2 Net change in generation

Arbitrage operation of the storage charges when the prices are low and discharges when the prices are high. A small 3MW storage capacity with a round trip efficiency of 85% is assumed for the arbitrage operation. Storage is a price-taker in this case and its operation will not impact the marginal clearing prices or the profits of the power plant operators. Storage arbitrage operation is determined by linear optimization as described in section 3.2.3.

Fuel type used/displaced when the storage charges and discharges is determined from the MFs of the fuel types at a given hour. Based on the MFs and the storage operation, I determine the fuel mix of the marginal generators used by the storage during charge and discharge phase per unit of energy delivered from the storage (*'F'*) as shown in Eq. 45. Total energy delivered from the storage is the summation of total discharge from the storage *('E')* (Eq. 44). Because of the roundtrip efficiency losses, energy used by the storage is greater than the energy delivered from the storage. For example, for a roundtrip efficiency of 85%, storage uses 1.17MWh of energy to deliver 1MWh of energy. Net energy used for each fuel type per unit of the energy delivered from the storage *('NF')* is estimated from the difference of fraction of fuel used during the charge phase and the fraction of fuel used during the discharge phase per unit of energy delivered from the storage as shown in (Eq. 47).

$$E = -\sum_t P_{d,t} \quad \text{Eq. 44}$$

$$F_{c,f} = \frac{\sum_t MF_{f,t} * P_{c,t}}{E} \; (F_{c,f} > 1) \quad \text{Eq. 45}$$

$$F_{d,f} = \frac{\sum_t MF_{f,t} * P_{d,t}}{E} \; (F_{d,f} < 1) \quad \text{Eq. 46}$$

$$NF_{c,d,f} = (F_{c,f} + F_{d,f}) \quad \text{Eq. 47}$$

Where: f- fuel type

t- hour

d- discharge

c- charge

P – Energy delivered by storage in a given hour t (MWh/h)

p>1 – charge, p<1 – discharge

E – Total energy delivered from storage (MWh)

F – Fraction of fuel type used per MWh delivered energy

NF – Net fraction of fuel type used per MWh delivered energy

4.3 Method- Effect on profits

Small capacities of storage are price takers and do not have a significant effect on the electricity prices, or profits on the other generators as price takers. Therefore, for estimating the change in profits, large capacities of storage are assumed as a price maker.

I combine the economic dispatch model and arbitrage operation of the storage to estimate the profits before and after adding storage systems. Storage capacities up to 5 GW are considered to observe the effects on profits.

4.3.1 Economic dispatch model and storage operation model

An economic dispatch model determines the lowest-cost operation of generating facilities that can reliably meet a given demand within the generators' ramping constraints [99] and simulates the market clearing prices for electricity. These electricity prices are used in an optimization model to determine the schedule for the storage operation, considering the effects of large storage on electricity prices as shown in section 3.1.1.

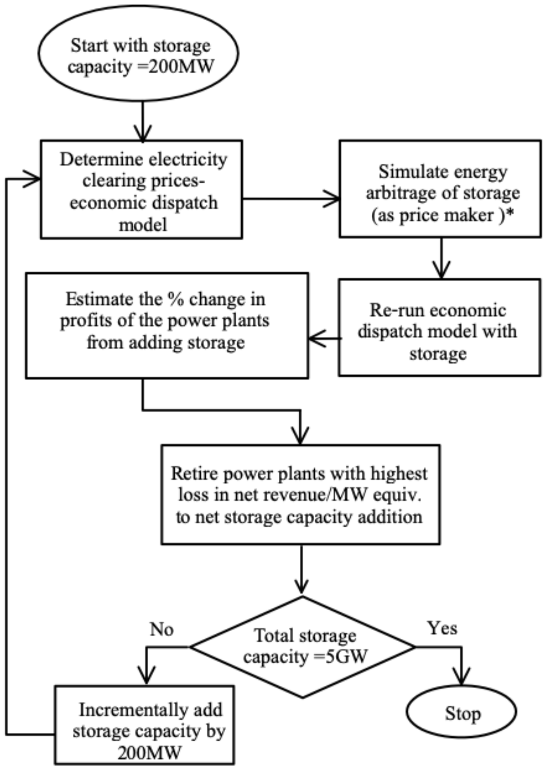

Fig. 25 Flowchart of methodology for evaluating percentage change in revenue of the power plants after adding storage.
The model produces a "no-storage" time series of prices, simulates storage operation, then calculates the percentage change in profits of the power plants with and without energy storage. Power plants that lose the highest net profit/MW after adding storage are retired. Capacity of the power plants retired is equiv. to net storage addition in each iteration. *For simulating storage operation, I use an iterative dispatch optimization which is shown separately in detail in Fig. 3.

The model accounting for the diversity of plant efficiencies in a region estimates total annual profits with and without storage for the power plants based on the plant's fixed cost ('FC'), marginal cost of operation ('MC'), and the clearing prices they receive ('I') for the power delivered in any given hour. The Marginal Cost ('MC') given in $/MWh is the summation of the fuel cost incurred per MWh and the variable O&M costs per MWh as shown in Eq. 4. The clearing prices of the electricity are simulated using the dispatch model with an objective to meet the demand in an given hour at the lowest cost possible (Eq. 6, Eq. 17].

The total generation ('G') of the power plants is estimated based on the hourly generation of each power plant ('P') over the year from the economic dispatch model. The variability of the renewables is accounted into the model by considering the hourly generation profiles of solar and wind energy across various locations from Wind Integration National Database (WIND) toolkit [42] and Eastern Solar Integration Data.

The power plants are then categorized into the primary fuel type used in their generation and the total profit and generation for each fuel type is estimated (Eq. 48**Error! Reference source not found.**). This process is repeated before and after adding storage and the net change in profit and generation are calculated as shown in Eq. 49**Error! Reference source not found.**.

$$R_f = \sum_{t,p} I_{p,f,t} * P_{p,f,t} - \ FC_{p,f} * N_{p,f} + \sum_{t,p} MC_{p,f,t} * P_{p,f,t} \qquad \text{Eq. 48}$$

Where: f- fuel type

t- hour

p- power plant

HR – heat rate (Btu/kWh)

Price – average spot price of fuel ($/MMBtu)

O&M- variable operations and maintenance cost ($/MWh)

P – Energy delivered in a given hour t (MWh/h)

I - Clearing price of the electricity ($/MWh)

R- Profit ($)

4.3.2 Net change in profits

The model estimates the net change in profit before and after adding storage as show in Eq. 49.

$$\Delta R_f = R_{f,ns} - \ R_{f,s} \qquad \text{Eq. 49}$$

Where: f- fuel type

R- Profit ($)

ns- no storage

s- after adding storage

For sensitivity analysis, a high wind capacity scenario of 70 GW is estimated for MISO, which is the most coal-heavy grid amongst MISO, NYISO, and CAISO.

4.3.3 Method for Retirement

Large capacities of storage alter both the dispatch stack and the wholesale electricity prices in the market. A steep decline in profits because of the low electricity prices could force retirements of the existing generating plants, especially peaker plants that are displaced by the storage operation [80, 88]. This refers to a state of market equilibrium, seen in competitive/de-regulated markets [88]. The model takes into account of these retirements as the storage is incrementally added into the system in each iteration.

After each iteration, the power plants losing the most profits from adding storage are retired and the total capacity of the power plants retired is equivalent to net storage added into the system (Fig. 25). This is assuming that there is enough reserve capacity in the power plant fleet and the storage capacity replaces the equivalent capacity of the plants losing the most profit. In the real world, any retirement decisions from the power plants are accepted by the ISOs only after ensuring that there is adequate reserve capacity to replace the existing retiring power plants [53, 100]. In our case, I assume that the storage replaces the retired capacity of the power plants to maintain the supply demand balance. For sensitivity analysis, results without the retirements are presented.

4.4 Results

The results section is organized as follows. I first present the effects of storage on the generating units, second – on the effects of storage on the profits of the generating units, assuming no retirements occur, and third- on the effects of storage on the profits of the generating units, assuming there are retirements of the loss-making power plants from the new entry of the storage.

4.4.1 Impact on generation from storage operation in 22 eGRID regions

This section discusses the type of fuels used and displaced when a storage capacity of 3 MW is added across the 22 eGRID regions in the U.S. In all the cases, the energy used for charging is always greater than the energy displaced because of the round-trip efficiency losses of the storage. This study assumes the round-trip efficiency loss of the storage to be 85%.

For the reader's understanding about the fuels used and displaced for delivering a MWh energy from energy storage, Fig. 26 illustrates the fuel types during the charge and discharge phase for a sample eGRID region CAMX (for California), and shows the net fuels used during the storage operation. Because of the efficiency losses in the storage operation, more energy is used to charge the storage than the energy displaced by the storage.

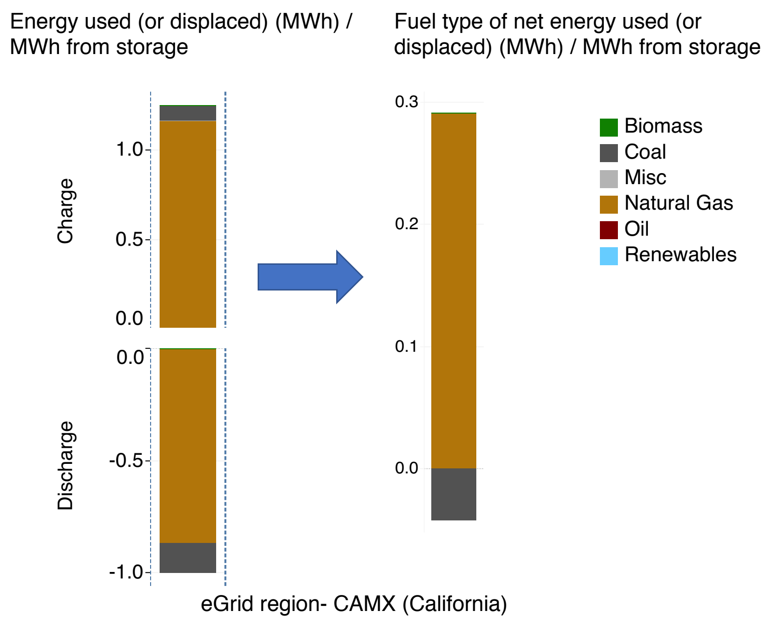

Fig. 26 Type of fuels used per MWh of energy delivered from the storage for a sample eGRID region CAMX covering California.
The leftmost figure indicates the energy consumed and displaced per MWh of energy from the storage. Negative x-axis values indicate the discharged energy from the storage and the positive values indicate the charge by the storage. Colors of the bars indicate the fuel mix during charge and discharge. The rightmost figure indicates the type of fuel used in net to deliver a MWh from the storage. The negative x-axis indicates the energy of the fuel type displaced and the positive x-axis indicates the energy of the fuel types used. Overall, storage in CAMX region in net uses 0.28MWh of natural gas and displaces 0.04 MWh of coal energy to deliver a MWh of energy.

For the results across the eGRID regions, all the regions can be broadly divided into the west- covering California, Arizona and the other western states, the Midwest and the east- covering most of the eastern coast of the U.S. (Fig. 27). Results show that the storage operation in most of the west - except in California, and in the Midwest consumes both coal-based energy and natural gas during the charge phase and displaces them both during the discharge phase. However, in net more coal-based energy is used to displace the natural-gas based energy. In most of the west and the Midwest, storage operation in net consumes an average of 0.3 MWh of coal-based energy and displaces 0.1 MWh of natural-gas based energy per unit MWh of energy delivered from the storage. On the other hand, in the eastern coast, a mix of coal and natural gas-based energy is used during the charge phase and the storage displaces a mix of coal, natural gas and oil. In net, on an average, 0.2 MWh of natural gas, 0.1 MWh of coal is used to displace 0.1 MWh of oil per unit MWh of energy delivered from the storage. Amongst all the regions, California (CAMX region), Upstate New York (NYUP region), and New England (NEWE region) in net consume more natural gas and displace coal during the discharge phase of the storage. The high concentration of renewables, hydro and natural gas during the off-peak hours pushes the usage of coal and natural gas peaker plants during the evening time in the peak-hours when the storage is most likely to charge, thus displacing these plants. Further information on the actual electricity grid mix seen in California on a sample day, illustrating the usage of the coal during the evening hours is given in Appendix C.

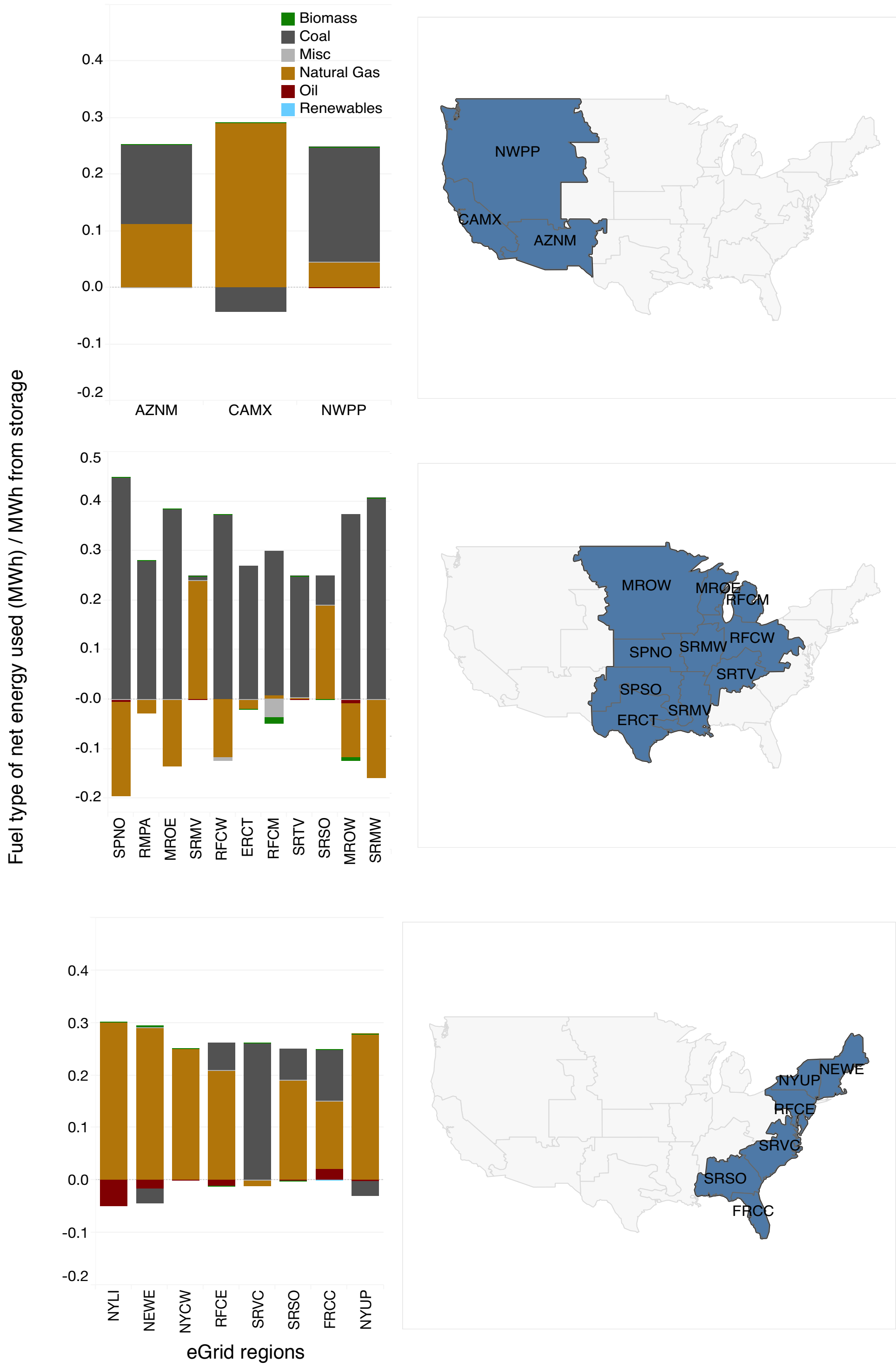

Fig. 27 Fuel type of net energy used per MWh of energy delivered from the storage.
The figures in the leftmost column indicate the energy consumed and displaced per MWh of energy from the storage. Y-axis represents the different eGRID regions. Colors of the bars indicate the fuel type. The negative x-axis indicates the energy of the fuel type displaced and the positive x-axis indicates the energy of the fuel types used. The right most column of figures indicates the highlighted eGRID regions for which the values are plotted.

4.4.2 Impact on profits from storage operation as a price-maker without retirements

Percentage change in profit before and after adding storage is estimated with respect to profits and without any retirements in CAISO, MISO and NYISO regions (Fig. *28*).

In CAISO, as the storage from 0-5GW is added, profits of all the generating units decrease from the decrease in electricity clearing prices as the NGCT and oil are offset from the storage. Profits decrease for oil based power plants decrease by almost 100% as the capacity of storage is increase in the grid, followed by base-load power plants- coal, and biomass by 55%, followed by NGCT by 35%, NGCC by 20%, hydro and nuclear by 5%, and wind and solar by less than 0.5%. Because of the high operations and maintenance cost for the coal, the percentage decrease in profits is higher than the NGCT power plants.

In NYISO, largest percentage decrease in profits is seen for the NGCT, dropping to 100% at 5GW of storage in the system. The grid mix during the peak demand hours is supplied by a mix of NGCT, NGCC, and <1% of oil. Therefore, any offset demand by the storage operations displaces a large percentage of NGCT and some NGCC. The base-load power plants are NGCC, hydro, nuclear, and less than 5% of coal, whose generation increases from the storage operation. Yet, the decrease in profits from the offset of peaker plants is bigger than the increase in profits from generating for storage. Overall, the percentage decrease in profits for NGCC is 15%, coal is 2%, and all other fuel is less than 1%.

In MISO, the capacity of the current power plant fleet is 3 times higher than the other two ISO regions. Therefore, much larger capacities of storage till 12 GW (15% of the demand) is considered for the analysis of profits without retirements. As the storage from 0-12GW is added into the system; the largest percentage decrease is for coal from 0-2%. All the other generating units have decrease in profits by less than 1%. A large percentage of coal in the generation mix (>70%) cause the price differentials to be much flatter in the MISO region than in the CAISO, or NYISO. Therefore, charged energy from the storage is much lower in the MISO (0.25 TWh), than in the CAISO (4.25 TWh), or in the NYISO (0.85 TWh) regions (Table 8).

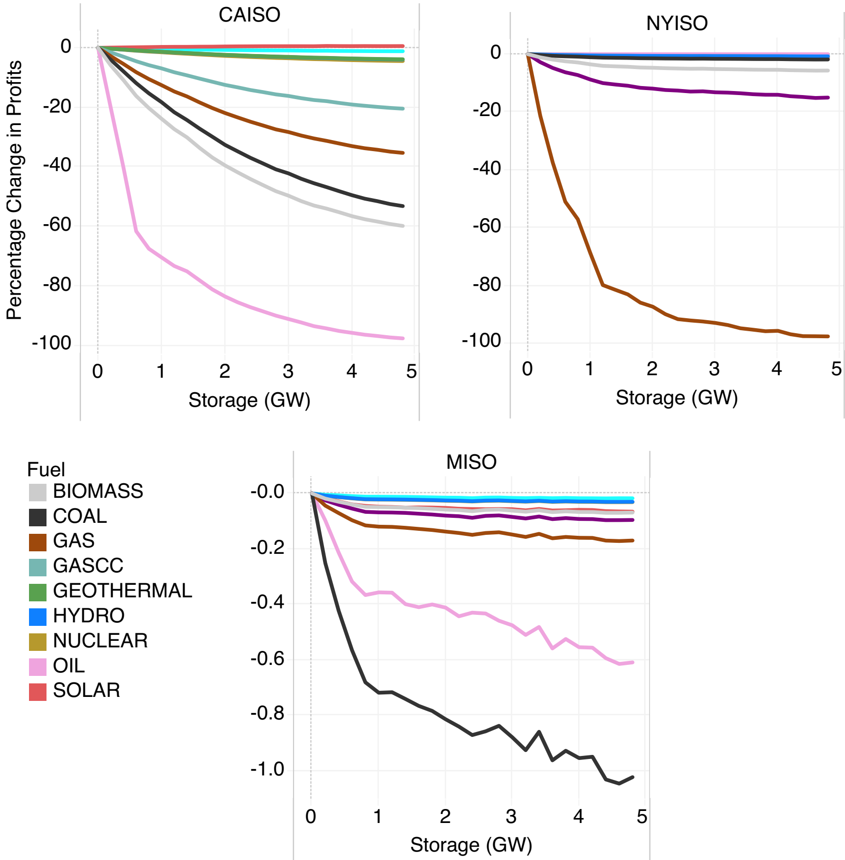

Fig. 28 Annual percentage change in profit before and after adding storage as a price-maker in Midcontinent Independent System Operator (MISO) and New York ISO (NYISO) without any retirements. X-axis represents the storage capacity in GW, Y-axis represents the percentage change in profit after adding storage, and the colors represent generation technologies.

Table 8. Total annual energy charged and discharged by 5GW storage in CAISO, MISO, and NYISO.

ISO	Annual energy charged by 5GW storage (TWh)	Annual energy discharged by 5GW storage (TWh)
CAISO	5	4.25
NYISO	1	0.85
MISO	0.3	0.25

4.4.3 Retirements

After each iteration, the power plants losing the most profits after adding storage are retired and the total capacity of the power plants retired is equivalent to the net storage added into the system. As shown in Fig. *29*, most of the retired capacities in CAISO and MISO are peaker plants -NGCT and coal power plants. This is because the storage operation offsets the peaker plants, thereby decreasing the most profits to them as compared to any the other power plants operating in the system. A small percentage of baseload power plants such as coal, and biomass retire from the high operational costs and decrease in electricity clearing prices after adding the storage. In NYISO, the loss-making plants are a mix of NGCC, oil, NGCT, and biomass. The peaker plants in this region are a mix of >70% NGCT power plants and ~25% NGCC power plants and <1% oil power plants. Therefore, baseload power plants – NGCC retire in NYISO as the profits offset by storage operation during the peak periods are much higher than the profits for NGCC when the storage charges.

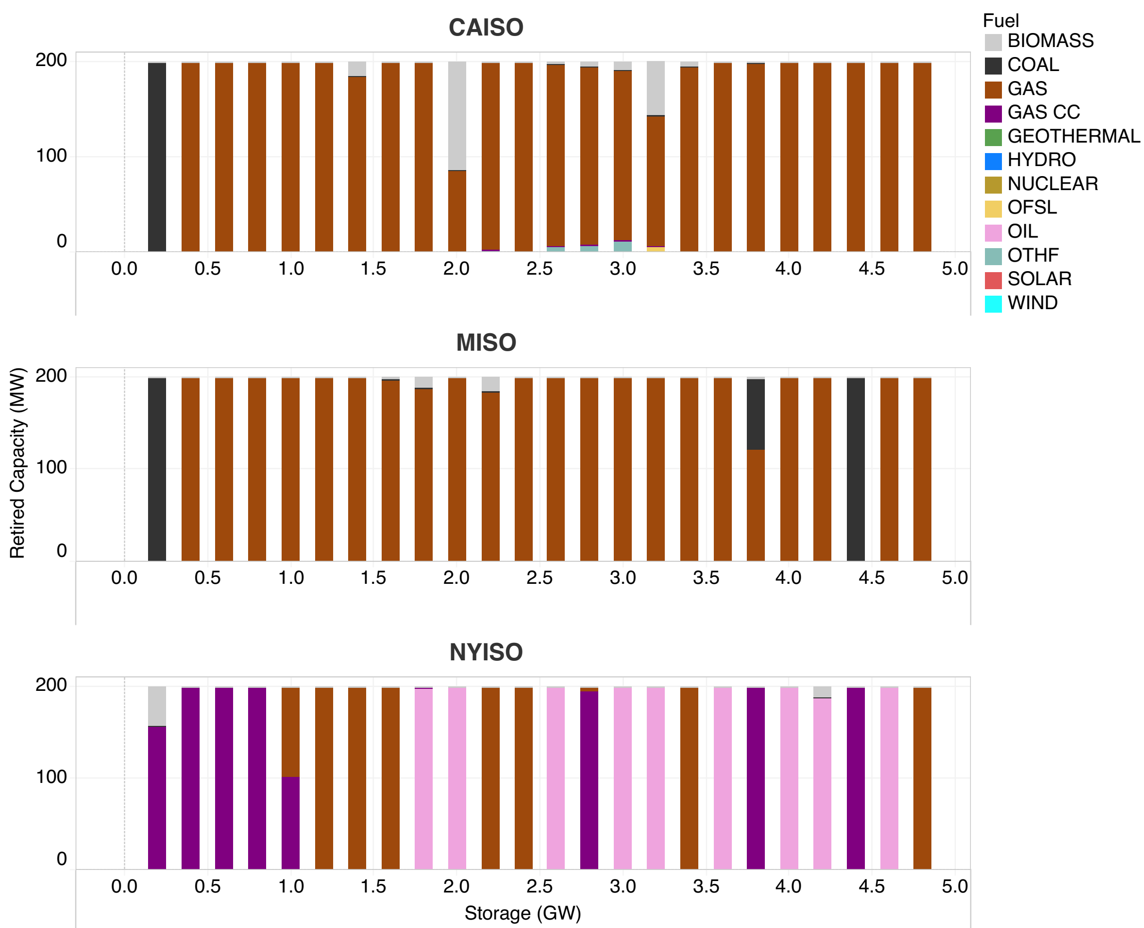

Fig. 29 Fuel mix of the retired power plants from adding incremental storage capacities.
X-axis represents the storage capacity in GW and Y-axis represents the total capacity of power plants retired after every incremental addition of storage capacity-200MW. Bar colors represent the fuel type. Each row of plots represents the ISO region. The ISO regions considered are California ISO (CAISO), Midcontinent ISO (MISO), and New York ISO (NYISO).

4.4.4 Impact on profits from storage operation as a price-maker with retirements

Percentage change in profit before and after adding storage is estimated with respect to profits without storage in CAISO, MISO and NYISO regions (Fig. *28*). Percentage change in profits is affected by the clearing prices set in the market and the economic retirement of power plants from adding storage.

In CAISO, as the storage from 0-5GW is added, a positive increase in profits is observed for solar from 0-4%, NGCC, nuclear, biomass, and hydro from 0-2%, and wind from 0-1%. When the storage discharges at the high electricity prices, peaker plants using NGCT are displaced and consistently lose profit, with a decrease to 6% at 5GW storage capacity. Profits of coal and oil drop to 10% and 6% respectively at 400MW of storage capacity. Coal continues to see negative profits, as much of the coal-based energy in CAISO is used during the peak hours with NGCT, when the storage discharges energy into the grid. Economic retirements are mostly seen for the peaker plants using NGCT, followed by biomass and coal. Though the clearing prices of the market go down because of this, additional NGCC plants during the charge phase of the storage increase the clearing prices of the market, thereby increasing the overall profits for the other plants.

In MISO, as the storage from 0-5GW is added into the system, the largest percentage decrease from 0-35% is observed for NGCT. Most of the NGCT plants in MISO are used during the peak demand for a faster ramping up, thereby attracting the discharge of the storage during these large price differentials in the system. Coal has the highest percentage increase in profit to 4%, followed by NGCC and biomass to 2% and other renewables at almost 0% because of the drop-in clearing prices from displacing of the NGCT plants in the system.

In NYISO, largest percentage decrease in profits is seen for the NGCT, dropping to 60% at 5GW of storage in the system. The grid mix during the peak demand hours is supplied by a mix of NGCT, NGCC, and 1% of oil. Therefore, depending on the resultant grid mix after the retirements, the average clearing prices of the market fluctuate depending upon the mix with cheaper NGCC on the margin versus NGCT on the margin. Therefore, profit change of the fuels is 'noisier' than the other regions. The percentage change in profits from the storage for the renewables is 0.5% and for the coal is 1.5%.

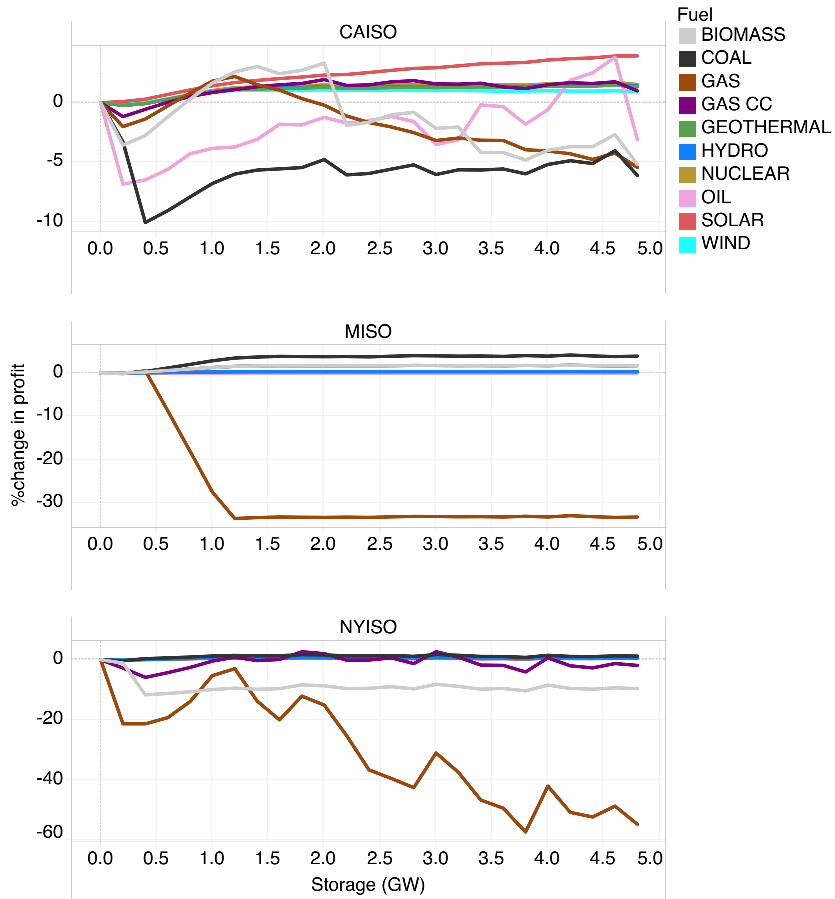

Fig. 30 Annual percentage change in profit before and after adding storage as a price-maker in Midcontinent Independent System Operator (MISO), New York ISO (NYISO), and California ISO (CAISO). X-axis represents the storage capacity in GW, Y-axis represents the percentage change in profit after adding storage, and the colors represent generation technologies.

If the wind capacity in MISO is increased by 5 times the current capacity to 70GW, the percentage increase in profit for wind is 1% as compared to 0.25% at the current capacity of 14GW. Yet the percentage change in profit is lower than the coal which at 4%, or biomass at 2%. This is because the clearing prices set by the coal are lower than the NGCC, whose generation on margin is offset by the larger percentage of wind in the mix.

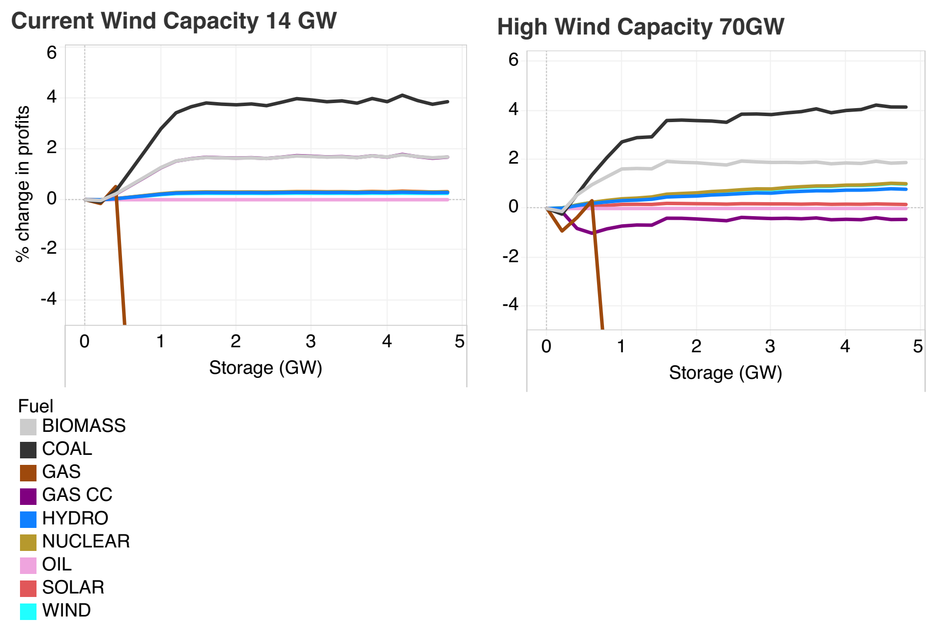

Fig. 31 Percentage change in profit with and without storage in Midcontinent ISO (MISO) as wind capacity is increased from current 14GW to 70GW.
X-axis represents the wind capacity in GW and Y-axis represents the percentage change in profit after adding storage w.r.t profit without storage and added wind capacity. Line colors represent generation technologies. The profits of NG CT decrease up to 40% in both the cases and is not shown to a complete scale in the figure to highlight the changes in profits of other technologies.

4.5 Contribution to the literature review, and discussion

Reviewing prior literature on the effects of storage arbitrage on the other generating units, Lueken et al. [80] shows that the market revenues and the generation of the power plants are affected by the storage operation, assuming storage as a price-taker and focusing on the PJM region. Similarly, Zamani et al. [88] calculated the effects of storage as a price-maker on the revenues and the generation of the units in the Alberta region. The study concludes that the generation from the base-load power plants- coal and combined cycle increases but the revenues of all the power plants decrease because of a steep decrease in the wholesale electricity prices. However, these results are sensitive to the grid mix and to the assumption that the electricity market does not adjust to the entry of the new players (in this case- storage) without any retirements of the loss-making power plants [88]. Also, the study is constrained to Alberta electricity market, not necessarily representing the spectrum of grid mixes seen in the U.S.

Prior work has established that, depending on the grid mix and the operating strategy of the storage, in most cases, storage increases the generation of the base-load power plants and decreases the revenues of the generating units, especially peaker plants. However, the studies so far do not quantify the sources of generation used and displaced by the storage or the effects of revenues on other generating units in different grid mix scenarios. The contribution of this work is to quantify the source of generation used and displaced by the storage operation across all the eGRID region in the U.S., based on the real electricity data. It is also the first analysis on the effects of storage operation as a price-maker on the profits of other generating units, covering the spectrum of possible grid mixes seen in the U.S.- from coal-heavy grid to renewables-heavy grid. The storage is considered a price-maker to quantify the profits of the other generating units, as the wholesale electricity prices and in turn profits are not widely affected by storage being a price-taker [88]. The second part of the analysis is carried out in three different grid mixes- solar-heavy (CAISO), coal-heavy (MISO), and natural gas-heavy (NYISO) grid mix. The results from this study informs the planners and the storage operators on the effects of storage on the generation units all across the U.S., and also on the profits of renewables in different grid mixes. The results from this study could help the regulators to effectively strategize the storage operations in places where it is a net social benefit rather than a net social cost.

The results show that the change in generation from the energy storage is largely dependent upon the grid mix when the capacities are small and marginal. Non-marginal, large capacities of storage alter the electricity prices, thus affecting not only the peaker plants, but also renewables that depend on these prices for the profit. Though, the generation of NGCC and coal power plants increase in these instances, the prices are not high enough to improve the profits of the renewables. This is observed for both current wind capacities and high wind capacities in the Midwest region. However, this trend is reversed in California, as it has renewables on par with the natural gas in the system. These results clarify that the profits of the renewables are largely dependent on the clearing prices set by the fossil fuels. Displacement of these fuels by storage impacts the profits of the renewables. One choice is to accept the decreased profits in the short-term in some grids in order to achieve longer term benefits after more renewables are adopted. Another choice is to create other major profit streams for the renewables, apart from the current energy markets highly dependent on the fossil fuels.

Chapter 5: Conclusions, Recommendations and Future Work

The world needs renewables for addressing the sustainability challenges and must continue to transition towards cleaner energy in this age of de-regulated energy markets and declining renewable technology costs. Operationally, the major limitation of solar and wind technologies is the variability and the weather dependency. Currently, the two major technologies that complement this variability are storage and natural gas. Many traditional models are forecasting a large percentage of natural gas in the future, and the storage operation in the existing grids to support the growth of renewables. This dissertation studies the near-term implications of natural gas and storage on the wind and solar and provides valuable insights if these deductions and assumptions of the traditional models are relevant to the changing dynamics of the electricity sector, and the results show otherwise.

> The overarching research question of this study is: *Does stochastic forecasting of the future grid for different risk preferences of the market enable more renewables additions over cheaper natural gas and does the assumption of promoting storage help in the growth of the renewables and in decarbonizing electricity grid?*

Chapter 2 introduces a proof-of-concept stochastic model that includes uncertainty and risk as core elements. Grid build-out now depends on a distribution of system costs; a genetic algorithm is used for co-optimization. Two objective functions are considered: “risk-neutral”, which optimizes to minimize average system cost and “risk-averse”, which optimizes to minimize average of the top 5% of costs (also called 95% Conditional Value at Risk (CVaR)). The results from this study show that risk-averse scenario does not increase mean system cost but adds significantly more wind (~ 20GW) and solar capacity (~15 GW) by 2050 compared to the risk-neutral objective. These results corroborate prior work showing that electricity system costs can be surprisingly inelastic to renewable adoption, which from the modeling perspective is like a ‘flat bowl’ of the cost-optimization space i.e., a small increase in net system costs packs more renewables, in our case shown through a small degree of risk aversion.

These results contrast the results from deterministic optimization used in capacity expansion, which assumes decisions are made assuming perfect information, i.e. parameter values are fixed and there is one grid-build out that minimizes system costs. Deterministic capacity expansion undermines the risks of higher costs and future uncertainties, and the capital investments once made are sunk. Most of the modeling efforts in the U.S. are centered at government agencies and national laboratories. The results from their future capacity expansion models are used by utilities, ISOs, and policymakers for planning future grid infrastructure. My results suggest that accounting for future input uncertainties and risk preferences show that the risk-averse scenario has almost the same mean NPV as the risk-neutral scenario but has more renewables. Also, minimizing for deterministic inputs does not necessarily produce optimized results when subjected to long-term input uncertainty. Also, risk-averse scenario has the least emissions of all the scenarios, while deterministic scenario results in the highest emissions.

Chapter 3 and 4 study the effects of storage operation on the electricity grid and on the generation and profits of other generating units, especially renewables. Chapter 3 models the deployment of large, non-marginal quantities of energy storage and wind and solar power to determine their combined effects on grid system emissions. Two different grid environments are analyzed: a coal-heavy grid (Midcontinent ISO) and non-coal grid (New York ISO). In this chapter, a new modeling approach is introduced, considering the effects of large energy storage as a price-maker on the current electricity grids and estimating the change in emissions as intermittent renewables (wind and solar) are added to a grid and as natural gas prices vary. Results show that the emissions from economic operation of the storage are highly sensitive to the natural gas prices and the coal capacity in the grid. Therefore, low cost; efficient natural gas generation is important to realize emissions reductions with storage under economic arbitrage. Adding storage operated to maximize profit in the MISO region will not be carbon neutral until wind or solar power reach around 18% of the generation capacity. A major caveat in this study is that it only considers the economic arbitrage of the storage. Different operation patterns for storage could realize higher carbon reductions. For example, a carbon price on emissions from generators would shift operation to make energy storage carbon neutral even with current wind and solar capacities.

The modeling approach of considering storage as a price-maker in chapter 2 is easy to integrate into current simple or sophisticated dispatch models. This is the first method that allows faster and flexible integration of storage as a price-maker and can be used by the large storage operators, energy modelers, and policymakers to quantify the environmental implications of large storage capacities in different grid situations.

In Chapter 4, I build a model that examines the effects of arbitrage operation of the storage on the profits and generation of the other generating units. In all the regions, the profits of the all the fuels decrease from adding storage, including the renewables. However, considering economic retirements from the entry of a new player into the market (storage), coal has a small increase in profits by 2% and all the other fuels continue to see a decline in profits in NYISO and MISO. The decrease in profits of the other generating units is because of the offset/retirements of the peaker NGCT plants that set the market clearing prices. On the other hand, in CAISO, the profits for renewables increase from the increase in electricity clearing prices set by the NGCC plants to meet the additional demand from the storage charging. Without this additional demand, the grid operates using clean energy sources at much lower electricity prices.

In most of the Midwest regions, storage increases the generation of coal and displaces natural gas. But, in the east and the west, storage increases the generation of the natural gas and displaces coal/oil, except in California, New York Upstate, and New England. Here, storage increases the generation of natural gas and displaces coal-based energy. The results from this study on the impacts of storage on the generation across the U.S. could be a useful tool to the policymakers and the decision makers to analyze the effects of storage in any given region. This could help the regulators to effectively strategize the storage operations in places where it is a net social benefit rather than a net social cost.

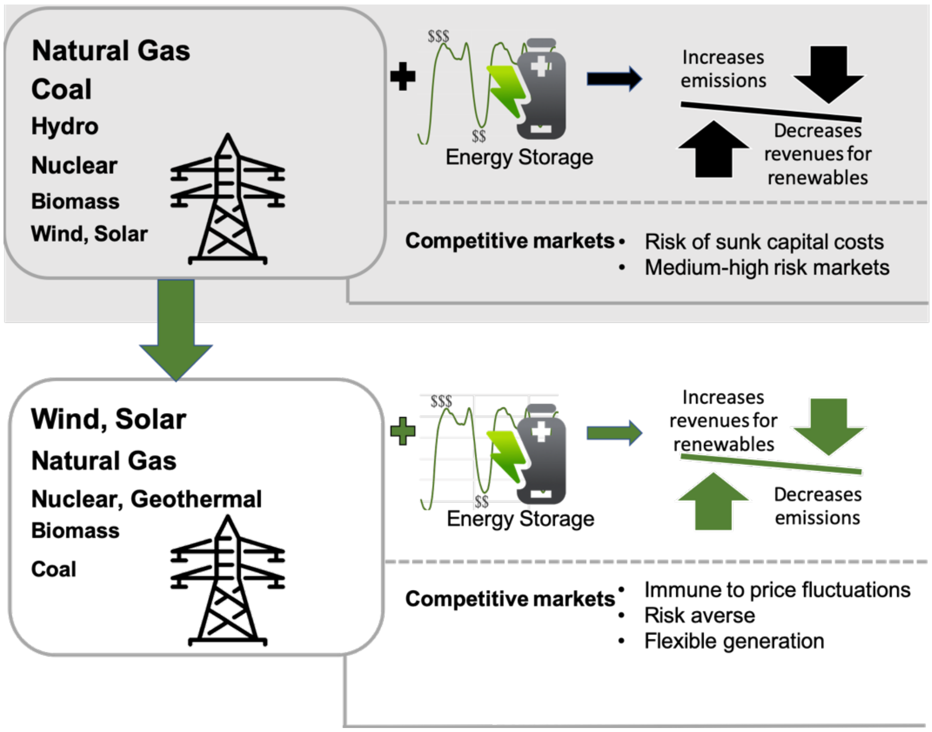

Fig. 32 Overview and conclusion of the dissertation.

The key takeaways from this dissertation are:

- Electricity system costs are insensitive to renewable adoption and more wind and solar can be installed with a small increase in total system costs through addition of a small degree of risk aversion.
- The emissions from economic operation of the storage are highly sensitive to natural gas prices and the coal capacity in the grid. Increase in natural gas prices from the current prices increases the storage induced emissions by 50 times in a coal-heavy grid (e.g. Midwest region).
- Profits of renewables from adding storage on grid vary based on the grid mix. In a coal-heavy or natural gas-heavy grid mix like in the Midwest or New York regions, the profits for renewables decrease. In a grid mix like in California with more than 40% energy from the clean energy fuels, profits for renewables increase from adding storage.
- Storage operation by economic arbitrage increases the generation of electricity by coal in the Midwest region, and natural gas in the west and east.

5.1.1 Policy Implications

Decision makers, utility planners and policy makers should consider uncertainty in inputs as core part of their analysis which could yield a lot more renewables in the future grid. Most of the current models do not adequately explore the optimization space but result in fewer renewables a fixed minimization point. Also, risk aversion attitude in the electricity markets could have led to more renewables than the predictions in the recent past, which should be considered as part of the future planning.

Policymakers should be cognizant of the operational implications of storage on the system and accept the near-term increase in emissions and profit losses to the renewables for longer term benefits when more renewables are added into the system. Arguments could be made justifying such a long-term perspective, but the current policy discourse does not frame the choice as a long-term one, instead asserting that storage delivers immediate benefits. It is also important to clarify the economic and environmental benefits delivered compared with alternative means of addressing intermittency, e.g. via flexible natural gas plants or improved transmission interconnection. Another option is to shift the economic context in which storage (and the grid) operates, e.g. a carbon tax, to ensure carbon benefits. I do not completely explore the benefits and cost of these three options in this study but assert there is a need for a clearer framing of societal expectations from storage.

5.2 Limitations

Sophisticated electricity system elements, such as transmission constraints, startup time, maintenance time and spinning reserves, are not included in the modeling done in this dissertation. However, adding in such detailed structures is unlikely to change the qualitative lessons learned here. Also, given more computational capacity, the modular nature of the model allows replacement with sophisticated dispatch models without major changes to the modeling framework. There are plans underway to make the Regional Energy Deployment System (REEDS) Model from National Renewable Energy Laboratory (NREL) [26] into an open-source model and this could be used in the future studies in place of the current dispatch model for a better representation of the transmission constraints.

The caveat with the storage operation model is that it does not include the other operations of the storage, except for the economic arbitrage. Though, 88% [11] of the current storage operation is economic arbitrage, energy storage offers other benefits such as a backup energy supply, frequency regulation, demand response, ramping, transmission quality support, etc., these benefits are not quantified in the model. Further research is required to analyze and quantify these benefits for the storage.

5.2.1 Future Research

Technology subsidies and Renewable Portfolio Standard (RPS) targets on emissions' performance: The current study focuses on the uncertainties in demand and natural gas prices for the future capacity expansion. However, future work should expand on the uncertainties in meeting RPS standards and with several technology subsidies. While RPS specifies target years for minimum adoption levels for renewables, public subsidies and utility costs depend on the trajectory through which targets are met. The model should be improved to allow exploration of issues such as the effect of subsidies/changing policies over the choice of new generations and address the probability of sustainability challenges that could be addressed.

Resilient Infrastructure as part of stochastic modeling: With the inevitable climate change, I see an increase in the number of natural disasters. Further research should be done to explore if the current technologies and policies on renewable energy can be modified with reasonable cost to deliver the benefits during the disasters and to understand if the benefits justify the cost under uncertainty.

Impacts of other energy storage operations: Current research explores the storage implications on the renewables and the environment. Further analysis should be done on strategically leveraging the storage potential to tap renewables in the world of cheap natural gas and reduce the dependency on the fossil fuels. With the inevitable climate change, this is important to address resiliency too. Also, stochastic modeling of the electricity markets can be extended to provide better estimates on the value of storage in the de-regulated energy markets.

References

1. EPA: Sources of Greenhouse Gas Emissions, https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions, (2014)
2. IPCC: Climate change 2014- mitigation of climate change ; Working Group III contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge Univ. Press, New York, NY (2014)
3. EIA: Electric Power Monthly, http://www.eia.gov/electricity/monthly/pdf/epm.pdf, (2016)
4. EIA: Preliminary Monthly Electric Generator Inventory (based on Form EIA-860M as a supplement to Form EIA-860), https://www.eia.gov/electricity/data/eia860M/(2016)
5. Electric Choice: Map of Deregulated Energy Markets (Updated 2018) – Electric Choice, https://www.electricchoice.com/map-deregulated-energy-markets/
6. Bushnell, J.B., Mansur, E.T., Saravia, C.: Vertical Arrangements, Market Structure, and Competition: An Analysis of Restructured US Electricity Markets. Am. Econ. Rev. 98, 237–266 (2008). doi:10.1257/aer.98.1.237
7. EIA: Annual Energy Outlook, http://www.eia.gov/outlooks/aeo/pdf/0383(2017).pdf, (2017)
8. NCSL: State Renewable Portfolio Standards and Goals, http://www.ncsl.org/research/energy/renewable-portfolio-standards.aspx (2018)
9. Jason Channell: citigroup-renewables-and-natgas-report.pdf, http://www.ourenergypolicy.org/wp-content/uploads/2013/04/citigroup-renewables-and-natgas-report.pdf (2013)
10. EIA: Annual Energy Outlook 2019 with projections to 2050. 83 (2019)
11. Sandia National Laboratories: DOE Global Energy Database, https://www.energystorageexchange.org (2017)
12. Bloomberg NEF: New Energy Outlook 2018, https://bnef.turtl.co/story/neo2018 (2018)
13. California Energy Commission: AB 2514 - Energy Storage System Procurement Targets from Publicly Owned Utilities, http://www.energy.ca.gov/assessments/ab2514_energy_storage.html

14. Masachusetts Gov: Energy Storage Target, https://www.mass.gov/service-details/energy-storage-target
15. Senator Guzzone: Income Tax Credit - Energy Storage Systems. (2017)
16. New York Sate Gov: Governor Cuomo Announces New York Energy Storage Roadmap to Achieve Nation-Leading Target of 1,500 Megawatts by 2025 to Combat Climate Change, https://www.governor.ny.gov/news/governor-cuomo-announces-new-york-energy-storage-roadmap-achieve-nation-leading-target-1500
17. Governor Phil Murphy: Office of the Governor | Governor Murphy Signs Measures to Advance New Jersey's Clean Energy Economy, https://nj.gov/governor/news/news/562018/approved/20180523a_cleanEnergy.shtml (2018)
18. Aas, D., O'Boyle, M.: PART 2 – REGULATORY ALTERNATIVES. 60
19. EIA: Availability of the National Energy Modeling System (NEMS) Archive, https://www.eia.gov/outlooks/aeo/info_nems_archive.php (2017)
20. Arbabzadeh, M., Johnson, J.X., Keoleian, G.A., Rasmussen, P.G., Thompson, L.T.: Twelve Principles for Green Energy Storage in Grid Applications. Environ. Sci. Technol. 50, 1046–1055 (2016). doi:10.1021/acs.est.5b03867
21. National Conference of State Legislation: State Renewable Portfolio Standards and Goals, http://www.ncsl.org/research/energy/renewable-portfolio-standards.aspx
22. World Bank: Carbon Pricing Dashboard | Up-to-date overview of carbon pricing initiatives, https://carbonpricingdashboard.worldbank.org/
23. IEA: WEO Model, https://www.iea.org/weo/weomodel/ (2018)
24. EIA: World Energy Projection System Plus: Overview. 35 (2017)
25. Blair, N., Zhou, E., Getman, D.: Electricity Capacity Expansion Modeling, Analysis, and Visualization: A Summary of Selected High-Renewable Modeling Experiences. Renew. Energy. 39 (2015)
26. NREL: Regional Energy Deployment System (ReEDS) | Energy Analysis | NREL, https://www.nrel.gov/analysis/reeds/ (2019)
27. Creutzig, F., Agoston, P., Goldschmidt, J.C., Luderer, G., Nemet, G., Pietzcker, R.C.: The underestimated potential of solar energy to mitigate climate change. Nat. Energy. 2, 17140 (2017). doi:10.1038/nenergy.2017.140

28. Fripp, M.: Switch: A Planning Tool for Power Systems with Large Shares of Intermittent Renewable Energy. Environ. Sci. Technol. 46, 6371–6378 (2012). doi:10.1021/es204645c
29. Hunter, K., Sreepathi, S., DeCarolis, J.F.: Modeling for insight using Tools for Energy Model Optimization and Analysis (Temoa). Energy Econ. 40, 339–349 (2013). doi:10.1016/j.eneco.2013.07.014
30. Schröder, A.: An electricity mrket model with genertion capacity expansion under uncertinty. Energy Syst. 5, 253–267 (2014). doi:10.1007/s12667-013-0106-0
31. Ehrenmann, A., Smeers, Y.: Generation Capacity Expansion in a Risky Environment: A Stochastic Equilibrium Analysis. Oper. Res. 59, 1332–1346 (2011). doi:10.1287/opre.1110.0992
32. Patankar, N., de Queiroz, A.R., DeCarolis, J.F., Bazilian, M.D., Chattopadhyay, D.: Building conflict uncertainty into electricity planning: A South Sudan case study. Energy Sustain. Dev. 49, 53–64 (2019). doi:10.1016/j.esd.2019.01.003
33. Bertocchi, M. ed: Stochastic optimization methods in finance and energy: new financial products and energy market strategies. Springer, New York (2011)
34. Yue, X., Pye, S., DeCarolis, J., Li, F.G.N., Rogan, F., Gallachóir, B.Ó.: A review of approaches to uncertainty assessment in energy system optimization models. Energy Strategy Rev. 21, 204–217 (2018). doi:10.1016/j.esr.2018.06.003
35. DeCarolis, J., Daly, H., Dodds, P., Keppo, I., Li, F., McDowall, W., Pye, S., Strachan, N., Trutnevyte, E., Usher, W., Winning, M., Yeh, S., Zeyringer, M.: Formalizing best practice for energy system optimization modelling. Appl. Energy. 194, 184–198 (2017). doi:10.1016/j.apenergy.2017.03.001
36. Lee, K.Y.: Generation expansion planning in a competitive environment using a genetic algorithm. In: IEEE Power Engineering Society Summer Meeting,. pp. 1169–1172 vol.3 (2002)
37. Banzhaf, W.: Genetic Programming for Pedestrians. In: Proceedings of the 5th International Conference on Genetic Algorithms. pp. 628–. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1993)
38. EIA: Updated Capital Cost Estimates for Utility Scale Electricity Generating Plants, http://www.eia.gov/outlooks/capitalcost/pdf/updated_capcost.pdf, (2013)

39. NREL: Cost-Benefit Analysis of Flexibility Retrofits for Coal and Gas-Fueled Power Plants, http://www.nrel.gov/docs/fy14osti/60862.pdf
40. EPA: eGRID, https://www.epa.gov/energy/egrid
41. Rubin, E.S., Azevedo, I.M.L., Jaramillo, P., Yeh, S.: A review of learning rates for electricity supply technologies. Energy Policy. 86, 198–218 (2015). doi:10.1016/j.enpol.2015.06.011
42. Draxl, C., Clifton, A., Hodge, B.-M., McCaa, J.: The Wind Integration National Dataset (WIND) Toolkit. Appl. Energy. 151, 355–366 (2015). doi:10.1016/j.apenergy.2015.03.121
43. National Renewable Energy Laboratory (NREL): Solar Power Data for Integration Studies, http://www.nrel.gov/grid/solar-power-data.html, (2006)
44. EIA: Assumptions to the Annual Energy Outlook 2017. 29 (2017)
45. Williams, E., Hittinger, E., Carvalho, R., Williams, R.: Wind power costs expected to decrease due to technological progress. Energy Policy. 106, 427–435 (2017). doi:10.1016/j.enpol.2017.03.032
46. Ranjbari, L., Bahar, A., Aziz, Z.A.: Stochastic Models of Natural Gas Prices. 20
47. Keles, D., Genoese, M., Möst, D., Fichtner, W.: Comparison of extended mean-reversion and time series models for electricity spot price simulation considering negative prices. Energy Econ. 34, 1012–1032 (2012). doi:10.1016/j.eneco.2011.08.012
48. EIA: Henry Hub Natural Gas Spot Price (Dollars per Million Btu), https://www.eia.gov/dnav/ng/hist/rngwhhdm.htm
49. MISO: Daily Regional Forecast and Actual Load, https://www.misoenergy.org/Library/MarketReports/Pages/MarketReports.aspx
50. Adler, D.: Genetic algorithms and simulated annealing: a marriage proposal. In: IEEE International Conference on Neural Networks. pp. 1104–1109 vol.2 (1993)
51. Mathworks: Genetic Algorithm, https://www.mathworks.com/discovery/genetic-algorithm.html (2019)
52. Mills, A.D., Wiser, R.H., Seel, J.: Power Plant Retirements: Trends and Possible Drivers. (2017)
53. MISO: 2017 MISO State of the Market Report.pdf, https://www.potomaceconomics.com/wp-content/uploads/2018/07/2017-MISO-SOM_Report_6-26_Final.pdf (2017)

54. EPA: eGRID2014 Technical Support Document, https://www.epa.gov/energy/egrid2014-technical-support-document (2014)
55. Temoa Project: The TEMOA Project | Tools for Energy Model Optimization and Analysis, http://www.temoaproject.org/
56. Craig, P.P., Gadgil, A., Koomey, J.G.: What Can History Teach Us? A Retrospective Examination of Long-Term Energy Forecasts for the United States. Annu. Rev. Energy Environ. 27, 83–118 (2002). doi:10.1146/annurev.energy.27.122001.083425
57. EIA: Annual Energy Outlook 2016 Early Release, http://www.eia.gov/outlooks/aeo/ (2016)
58. Denholm, P., Hand, M.: Grid flexibility and storage required to achieve very high penetration of variable renewable electricity. Energy Policy. 39, 1817–1830 (2011). doi:10.1016/j.enpol.2011.01.019
59. Georgilakis, P.S.: Technical challenges associated with the integration of wind power into power systems. Renew. Sustain. Energy Rev. 12, 852–863 (2008). doi:10.1016/j.rser.2006.10.007
60. CAISO: Advancing and Maximizing the Value of Energy Storage Technology, https://www.caiso.com/Documents/Advancing-MaximizingValueofEnergyStorageTechnology_CaliforniaRoadmap.pdf, (2014)
61. Lin, Y., Johnson, J.X., Mathieu, J.L.: Emissions impacts of using energy storage for power system reserves. Appl. Energy. 168, 444–456 (2016). doi:10.1016/j.apenergy.2016.01.061
62. Hittinger, E.S., Azevedo, I.M.L.: Bulk Energy Storage Increases United States Electricity System Emissions. Environ. Sci. Technol. 49, 3203–3210 (2015). doi:10.1021/es505027p
63. Siler-Evans, K., Azevedo, I.L., Morgan, M.G.: Marginal Emissions Factors for the U.S. Electricity System. Environ. Sci. Technol. 46, 4742–4748 (2012). doi:10.1021/es300145v
64. Sioshansi, R., Denholm, P., Jenkin, T., Weiss, J.: Estimating the value of electricity storage in PJM: Arbitrage and some welfare effects. Energy Econ. 31, 269–277 (2009). doi:10.1016/j.eneco.2008.10.005
65. Das, T., Krishnan, V., McCalley, J.D.: Assessing the benefits and economics of bulk energy storage technologies in the power grid. Appl. Energy. 139, 104–118 (2015). doi:10.1016/j.apenergy.2014.11.017

66. EPRI: Midwest Independent Transmission System Operator (MISO) Energy Storage Study, http://www.epri.com/abstracts/Pages/ProductAbstract.aspx?ProductId=000000000001024489, (2012)
67. NYISO: Power Trends 2016: The Changing Energy Landscape, http://www.nyiso.com/public/webdocs/media_room/publications_presentations/Power_Trends/Power_Trends/2016-power-trends-FINAL-070516.pdf, (2016)
68. EPA: Summary of the Energy Policy Act, https://www.epa.gov/laws-regulations/summary-energy-policy-act
69. MATLAB. The MathWorks Inc.,
70. MISO: Prices, https://www.misoenergy.org/MarketsOperations/Prices/Pages/Prices.aspx
71. NYISO: Integrated Real-Time Data, http://www.nyiso.com/public/markets_operations/market_data/load_data/index.jsp (2018)
72. EIA: Cost and Performance Characteristics of New Generating Technologies, Annual Energy Outlook 2017, https://www.eia.gov/outlooks/aeo/assumptions/pdf/table_8.2.pdf, (2017)
73. EIA: Power Plant Operations Report. (2016)
74. Uranium Information Center: The Economics of Nuclear Power, http://www.uic.com.au/nip08.htm, (2000)
75. EIA: Data-Petroleum and other Liquids, https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=F000000__3&f=A, (2016)
76. IEA: Gas to Coal competition in the U.S. Power Sector, https://www.iea.org/publications/insights/insightpublications/CoalvsGas_FINAL_WEB.pdf, (2013)
77. Matek, B., Gawell, K.: The Benefits of Baseload Renewables: A Misunderstood Energy Technology. Electr. J. 28, 101–112 (2015). doi:10.1016/j.tej.2015.02.001
78. Kalam, A., King, A., Moret, E., Weerasinghe, U.: Combined heat and power systems: economic and policy barriers to growth. Chem. Cent. J. 6, S3 (2012). doi:10.1186/1752-153X-6-S1-S3
79. Li Rui, Chen Laijun, Zhao Bo, Wei Wei, Liu Feng, Xue Xiaodai, Mei Shengwei, Yuan Tiejiang: Economic Dispatch of an Integrated Heat-Power Energy Distribution System with a Concentrating Solar Power Energy Hub. J. Energy Eng. 143, 04017046 (2017). doi:10.1061/(ASCE)EY.1943-7897.0000472

80. Lueken, R., Apt, J.: The effects of bulk electricity storage on the PJM market. Energy Syst. 5, 677–704 (2014). doi:10.1007/s12667-014-0123-7
81. McConnell, D., Forcey, T., Sandiford, M.: Estimating the value of electricity storage in an energy-only wholesale market. Appl. Energy. 159, 422–432 (2015). doi:10.1016/j.apenergy.2015.09.006
82. Bradbury, K., Pratson, L., Patiño-Echeverri, D.: Economic viability of energy storage systems based on price arbitrage potential in real-time U.S. electricity markets. Appl. Energy. 114, 512–519 (2014). doi:10.1016/j.apenergy.2013.10.010
83. Tom Falin: MISO Wind Capacity Credit Calculation, http://www.pjm.com/~/media/committees-groups/committees/pc/20160310/20160310-item-14b-miso-wind-calculation.ashx, (2016)
84. EPA: The Social Cost of Carbon, https://www.epa.gov/climatechange/social-cost-carbon
85. EIA: Electric Power Annual 2015, http://www.eia.gov/electricity/annual/pdf/epa.pdf, (2016)
86. U.S. Energy Information Administration: Annual Energy Outlook, (2016)
87. Sioshansi, R.: Emissions impacts of wind and energy storage in a market environment. Environ. Sci. Technol. 45, 10728–10735 (2011). doi:10.1021/es2007353
88. Zamani-Dehkordi, P., Shafiee, S., Rakai, L., Knight, A.M., Zareipour, H.: Price impact assessment for large-scale merchant energy storage facilities. Energy. 125, 27–43 (2017). doi:10.1016/j.energy.2017.02.107
89. MISO: Wind Forecasting Review, https://www.misoenergy.org/Library/Repository/Meeting%20Material/Stakeholder/BOD/Markets%20Committee/2015/20150826/20150826%20Markets%20Committee%20of%20the%20BOD%20Item%2006%20Wind%20Forecasting.pdf, (2015)
90. Hittinger, E., Lueken, R.: Is inexpensive natural gas hindering the grid energy storage industry? Energy Policy. 87, 140–152 (2015). doi:10.1016/j.enpol.2015.08.036
91. Pearre, N.S., Swan, L.G.: Technoeconomic feasibility of grid storage: Mapping electrical services and energy storage technologies. Appl. Energy. 137, 501–510 (2015). doi:10.1016/j.apenergy.2014.04.050
92. Spector, J.: Maryland Passes First-of-a-Kind Tax Credit for Residential and Commercial Storage, https://www.greentechmedia.com/articles/read/maryland-passes-tax-credit-for-residential-and-commercial-energy-storage

93. Akhil, A.A., Huff, G., Currier, A.B., Kaun, B.C., Rastler, D.M., Chen, S.B., Cotter, A.L., Bradshaw, D.T., Gauntlett, W.D.: DOE/EPRI 2013 Electricity Storage Handbook in Collaboration with NRECA. 340 (2013)
94. Goteti, N.S., Hittinger, E., Williams, E.: How much wind and solar are needed to realize emissions benefits from storage? Energy Syst. 1–23 (2017). doi:10.1007/s12667-017-0266-4
95. Denholm, P., Jorgenson, J., Hummon, M., Jenkin, T., Palchak, D., Kirby, B., Ma, O., O'Malley, M.: The Value of Energy Storage for Grid Applications. (2013)
96. NYISO: Pricing Data, http://www.nyiso.com/public/markets_operations/market_data/pricing_data/index.jsp
97. EPA: Air Markets Program Data | Clean Air Markets | US EPA, https://ampd.epa.gov/ampd/
98. CAISO: 2016AnnualReportonMarketIssuesandPerformance.pdf, http://www.caiso.com/Documents/2016AnnualReportonMarketIssuesandPerformance.pdf
99. Li, M., Zou, X., Wang, W., Niu, Y., Liu, J.: Economic dispatch of wind-thermal power system with MW and ramp rate dependent generator costs. In: 2016 IEEE/PES Transmission and Distribution Conference and Exposition (T D). pp. 1–5 (2016)
100. NYISO: 2018 State of the Market Report for the New York ISO Markets, https://www.nyiso.com/documents/20142/2223763/2018-State-of-the-Market-Report.pdf/b5bd2213-9fe2-b0e7-a422-d4071b3d014b?t=1557344025932 (2018)

Appendix A

A1. eGRID data

Emissions & Generation Resource Integrated Database (eGRID) [1] from Environmental Protection Agency (EPA) provided information about the existing set of power plants which include their technology and cost specifications. The complete dataset can be downloaded from the url: https://www.epa.gov/energy/emissions-generation-resource-integrated-database-eGRID and Table S1 below illustrates the sample of the data used in the stochastic capacity expansion model in chapter 2. The data used from the eGRID for the model are existing fleet of power plants in an ISO region, fixed costs, operating costs, heat rates, and plant level emissions data.

Table S1. Sample eGRID data used in the dissertation.

Plant name	DOE/EIA ORIS plant or facility code	eGRID subregion acronym	Plant associated ISO/RTO Territory	Plant primary fuel	Plant primary coal/oil/gas/ other fossil fuel category	Plant capacity factor	Plant nameplate capacity (MW)	Combined heat and power (CHP) plant adjustment flag: 1 = Yes	CHP plant electric allocation factor	Plant annual net generation (MWh)	Plant annual NOx total output emission rate (lb/MWh)	Plant ozone season NOx total output emission rate (lb/MWh)	Plant annual SO2 total output emission rate (lb/MWh)	Plant annual CO2 total output emission rate (lb/MWh)	Plant annual CH4 total output emission rate (lb/GWh)	Plant annual N2O total output emission rate (lb/GWh)	Plant annual CO2 equivalent total output emission rate (lb/MWh)	Plant annual Hg total output emission rate (lb/GWh)	Plant nominal heat rate (Btu/kWh)
PNAME	ORISPL	SUBRGN	ISORTO	PLPRMFL	PLFUELCT	CAPFAC	NAMEPCAP	CHPFLAG	ELCALLOC	PLNGENAN	PLNOXRTA	PLNOXRTO	PLSO2RTA	PLCO2RTA	PLCH4RTA	PLN2ORTA	PLC2ERTA	PLHGRTA	PLHTRT
Agrium Kenai Nitrogen Operations	54452	AKGD		NG	GAS		21.6	Yes										-	
Alakanuk	57053	AKMS		JF	OIL	0.0990	2.6	Yes	1.0000	2,244	33.1	33.1	5.2	1,630.6	67.7	13.5	1,636.184	-	10,241
Allison Creek Hydro	58982	AKMS		WAT	HYDRO		6.5											-	
Anchorage 1	75	AKGD		NG	GAS	0.0320	121.4			33,642	0.0	0.0	0.0	1,504.7	28.5	2.9	1,506.2	-	12,873
Angoon	7462	AKMS		DFO	OIL	0.1020	1.9	Yes	1.0000	1,693	34.7	34.7	3.1	1,765.3	71.5	14.3	1,771.3	-	10,809
Aniak	7182	AKMS		DFO	OIL	0.1500	2.0			2,630	35.1	35.1	3.0	1,690.6	68.5	13.7	1,696.3	-	10,351
Annex Creek	62	AKMS		WAT	HYDRO	0.7810	4.0			27,380				-	-	-	-	-	
Auke Bay	7250	AKMS		DFO	OIL		36.2											-	
Aurora Energy LLC Chena	79	AKGD		SUB	COAL	0.6970	31.3	Yes	0.9046	191,062	-	-	5.3	3,401.5	385.7	56.1	3,427.0	-	15,903
Barrow	7173	AKMS		NG	GAS	0.2790	20.3	Yes	1.0000	49,637	0.0	-	0.0	1,770.4	33.4	3.3	1,772.1	-	15,146
Battery Energy Storage System	57583	AKGD		MWH	OTHF		27.0											-	
Beaver Falls	6580	AKMS		WAT	HYDRO	0.9620	5.4			45,508				-	-	-	-	-	
Beluga	96	AKGD		NG	GAS	0.3710	374.4			1,216,481	3.5	3.8	0.0	834.8	23.5	2.3	836.0	-	7,142
Bernice Lake	6292	AKGD		NG	GAS	0.0260	76.7			17,447	7.3	7.3	0.0	2,602.6	49.1	4.9	2,605.1	-	22,265
Bethel	6566	AKMS		DFO	OIL	0.3800	12.6	Yes	1.0000	41,963	32.6	32.6	2.8	1,602.6	64.9	13.0	1,608.0	-	9,812
Black Bear Lake	7752	AKMS		WAT	HYDRO	0.5360	4.5			21,145				-	-	-	-	-	
Bradley Lake	7367	AKGD		WAT	HYDRO	0.3920	126.0			432,575				-	-	-	-	-	
Centennial	7112	AKMS		DFO	OIL	-	3.3			(39)	-	-	-	-	-	-	-	-	(57,462)
Chester Lake	7168	AKMS		WAT	HYDRO	0.6810	1.0			5,966				-	-	-	-	-	
Chevak	6311	AKMS		DFO	OIL		1.0											-	
Cooper Lake	6291	AKGD		WAT	HYDRO	0.1890	19.4			32,050				-	-	-	-	-	
Craig (AK)	421	AKMS		DFO	OIL	0.0080	4.4			306	52.0	51.8	4.7	2,652.7	107.4	21.5	2,661.6	-	16,242
Delta Power	56325	AKGD		DFO	OIL	-	23.1			(121)	-	-	-	-	-	-	-	-	(8,339)
Delta Wind Farm	58511	AKGD		WND	WIND	0.2700	1.9			4,500				-	-	-	-	-	
Dillingham	109	AKMS		DFO	OIL	0.1970	10.7			18,459	30.9	30.9	2.8	1,571.5	63.6	12.7	1,576.8	-	9,622
Dutch Harbor	7502	AKMS		DFO	OIL	0.2050	26.0			46,622	29.4	29.4	2.7	1,497.2	60.6	12.1	1,502.275	-	9,167
Eielson AFB Central Heat & Power Plant	50392	AKGD		SUB	COAL	0.2330	33.5	Yes	0.1568	68,264	0.0	0.0	2.8	1,217.0	138.0	20.1	1,226.2	-	5,690
Eklutna Generation Station	58989	AKGD		NG	GAS		171.0											-	
Eklutna Hydro Project	77	AKGD		WAT	HYDRO	0.2140	44.4			83,110				-	-	-	-	-	
Elim	57060	AKMS		JF	OIL	0.0850	1.6	Yes	1.0000	1,198	32.8	32.8	5.1	1,615.8	67.1	13.4	1,621.4	-	10,149
Emmonak	6314	AKMS		DFO	OIL	0.1120	3.5	Yes	1.0000	3,425	32.2	32.2	2.8	1,583.8	64.1	12.8	1,589.1	-	9,697
ESS Battery	58405	AKMS		MWH	OTHF		3.0											-	
Eva Creek Wind	57935	AKGD		WND	WIND	0.3330	24.6			71,779				-	-	-	-	-	
Fairbanks	6286	AKGD		RFO	OIL	0.0190	42.2			7,085	22.0	21.6	23.5	3,888.0	157.6	31.5	3,901.0	-	23,827

Appendix B

B1. Summary of Storage Services in the United States

Global Energy Storage Database developed by Sandia National Laboratories, supported by Department of Energy, US is an open source, up-to date information on grid connected energy storage projects [2]. It has 22 detailed categories of storage services- energy arbitrage being one of the services. Apart from energy arbitrage, we categorized the other 21 types of storage services into 7 major types illustrated in the Fig. S1. Those are storage for residential purposes, reserve capacities, for integrating renewables to grid (given as renewable energy support), ramping, power quality, power backup, frequency regulation, and demand response. Overall, out of 24 GW of storage capacity in the US, 21 GW provide arbitrage services.

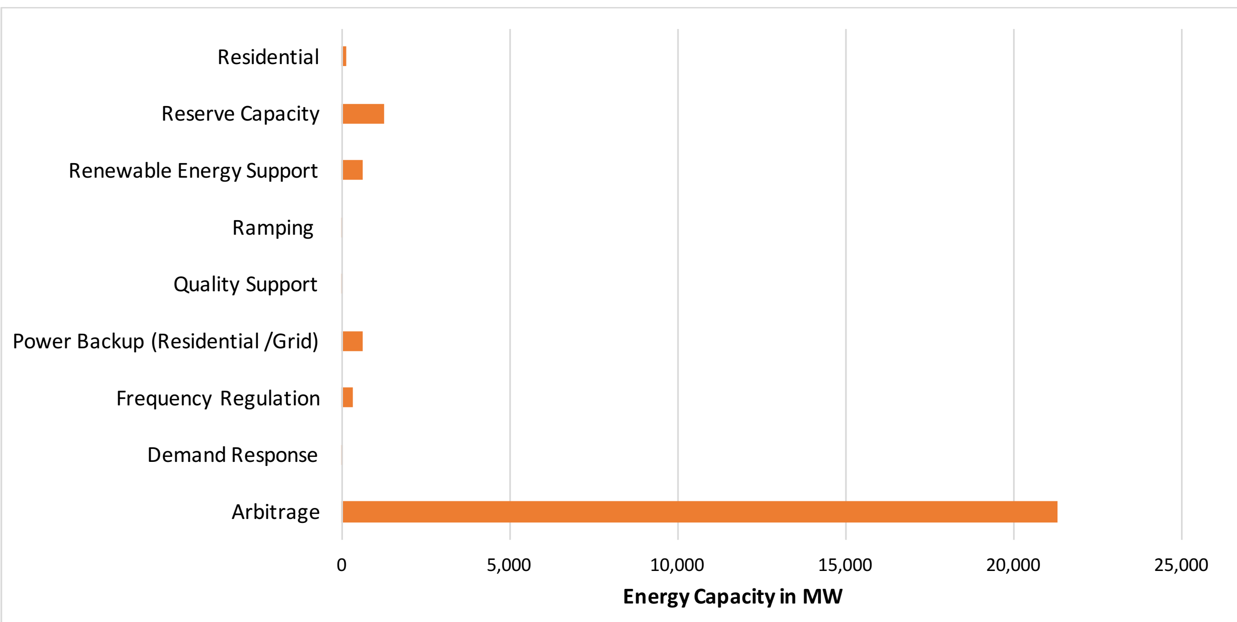

Fig. S1 Total energy storage capacities of different services offered by storage facilities in the US. Y-axis represents the different services provided by the storage, and X-axis represents the total capacities of these services in MW.

B2. Summary of data sources used in economic dispatch model

Table S2. Summary of data sources used in economic dispatch model. *-Detailed data sources of fuel costs are given in Table 1 in the main text.

Parameters	Database Sources
Electricity load demand	Real-time market data available from NYISO and MISO [3,4]
Power plants data	eGRID database [1]
Variable O&M cost at plant level	EIA [5]
Carbon Tax	EPA [6]
Fuel costs	EIA Electricity database*
Hourly wind variation	Eastern Wind Integration Dataset [7]
Hourly solar variation	Eastern Solar Integration Dataset [8]

B3. Variations in Wind and Solar energy

The average hourly variations of the wind and solar energy in MISO region across the 30 potential sites chosen from Wind Integration National Database (WIND) Toolkit [7] and Eastern Solar Integration Database [8] by NREL respectively is as shown in the Fig. S2.

The screenshot of the potential sites of wind energy on the map as seen on the NREL Wind Prospector interface based on the Eastern Wind Dataset [9] is shown in the Fig. S3. Out of all the points, 30 potential locations, 2 from each state under the Midcontinent ISO (MISO) are chosen. Most of the locations have capacity factor of wind greater than 0.4.

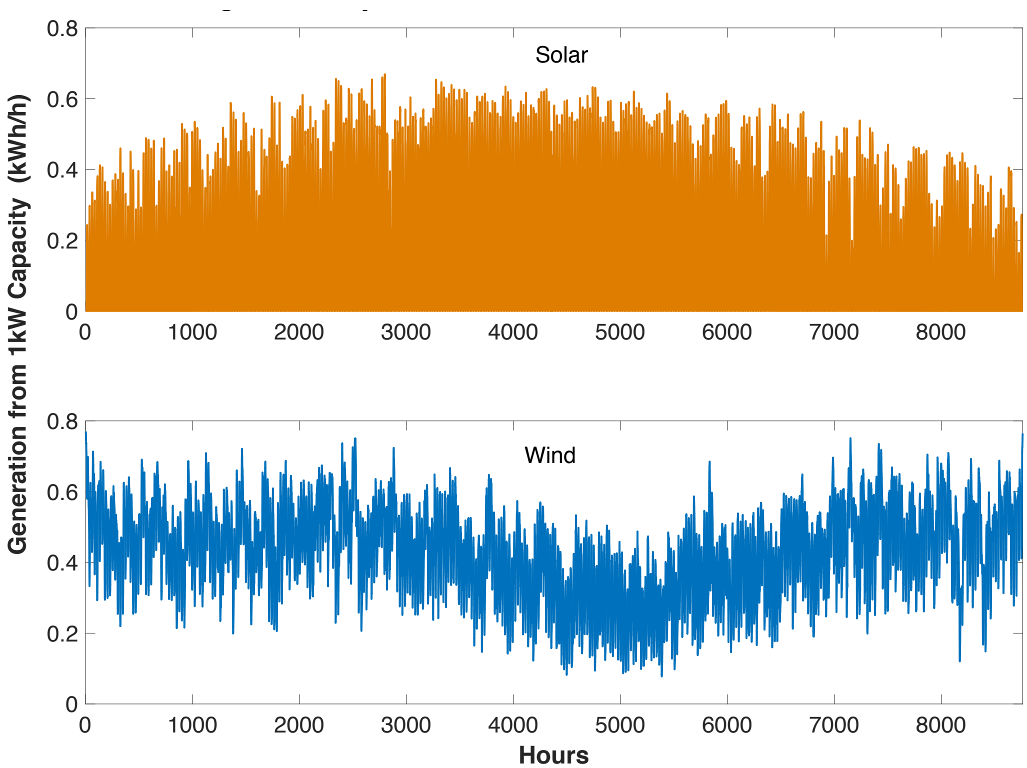

Fig. S2 Average Variability of the Wind and Solar Energy across 15 potential sites chosen in the MISO region. The variability is shown for a sample 1kW capacity system to understand the system output/kW/hour given in kWh/hour.

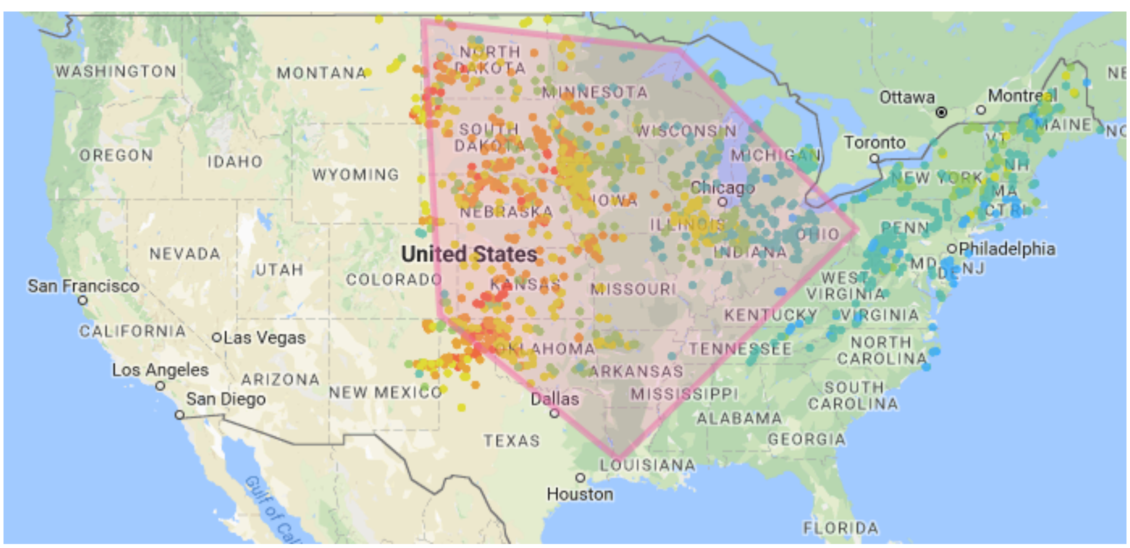

Fig. S3 Screenshot of the potential sites of wind energy on the map as seen on the NREL Wind Prospector interface, based on the Eastern Wind Integration Dataset. The pink shaded region indicates the states under Midcontinent ISO (MISO). The color gradient of the dots indicates the capacity factor of the wind power plants- Green being the lowest (0.032) and red being the highest (0.472). The average capacity factor of most of the sites in MISO is 0.4.

B4. Emissions factors of storage with addition of solar and wind

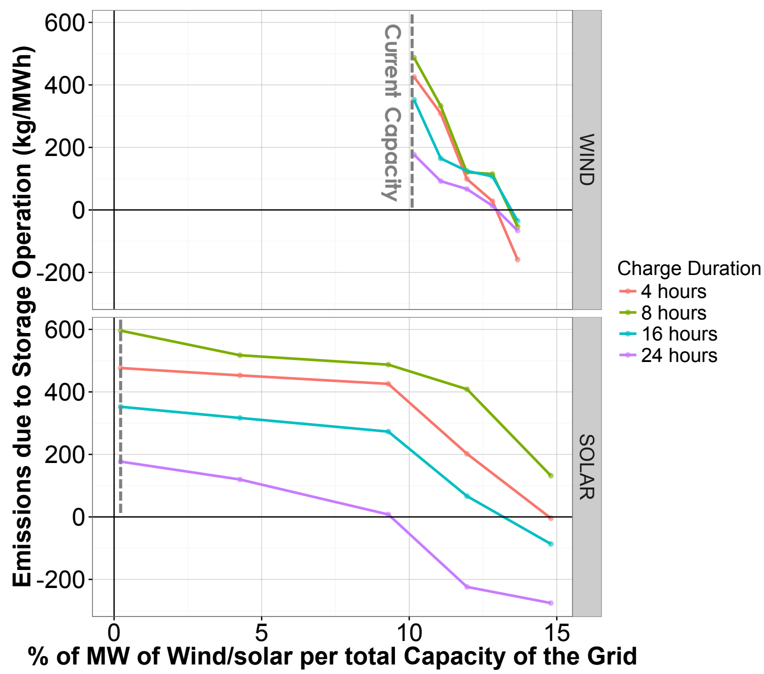

Fig. S4 Change in emissions per delivered electricity from storage with the addition of wind/solar energy on the grid in the Midcontinent ISO (MISO). CO_{2eq}. emissions/MWh decrease as wind/solar are added to the grid and for slower charging rates.

Appendix C

C1. Data sources for the clearing prices across all the eGRID regions

Nodes evaluated for pricing information were based off geographical location. Representative nodes in mostly-central locations to the relevant eGRID sub-region had their annual data for the year of 2016. Regions where zones needed to be averaged were done so by using the demand in the zones for a weighted average in the region. Demand data for zones where taken from the same pricing information authority. Information on exact nodes used for data collection may be found in Table S2. Multiple node/zone names denote an average across the nodes/zones being used to calculate hourly price data.

Table S2: Pricing Data resource guide.

eGRID Sub-region	Node/Zone Name(s)	Source ISO
NEWE	.Z.MAINE, .Z.NEWHAMPSHIRE, .Z.VERMONT, .Z.CONNECTICUT, .Z.RHODEISLAND, .Z.SEMASS, .Z.WCMASS, .Z.NEMASSBOST *	NEISO
NYUP	CAPITL, CENTRL, DUNWOD, GENESE, HUD VL, MHK VL, MILLWD, NORTH, WEST	NYISO
NYLI	LONGIL	NYISO
NYCW	NYC	NYISO
RFCE	AECO, PPL, PENELEC, BGE, JCPL, METED, PSEG, PEPCO	PJM
RFCW	APS, AEP, ATSI, DUK	PJM
ERCT	AEN, CPS, HOUSTON, LCRA, NORTH, RAYBN, SOUTH, WEST *	ERCOT
SPSO	SPPSOUTH_H	SPP
SPNO	SPPNORTH_H	SPP
SRVC	Weighted average of selected. **	MISO
SRTV	Weighted average of selected. **	MISO
SRMW	ILLINOIS.HUB	MISO
SRMV	ARKANSAS.HUB	MISO
MROW	MINN.HUB	MISO
MROE	WPS.WPSM Load Zone	MISO
RFCM	MICHIGAN.HUB	MISO
NWPP	Weighted average of selected. **	CAISO
FRCC	Weighted average of selected.	MISO

SRSO	Weighted average of selected. **	MISO
AZNM	GENE_2_N001	CAISO
CAMX	Average of selected. ***	CAISO
RMPA	SPRINGCR_LNODED1	CAISO

*Used demand weighted average.

**Used demand weighted average of interface prices within the given area (demand in interfaces obtained from EIA).

***Average of prices in default load aggregate points (DLAP).

Load data was obtained from:

https://www.eia.gov/realtime_grid/?src=data#/data/graphs?end=20160725&start=20160625&frequency=Daily&regions=008

C3. Real-time electricity coal generation in CAISO

A sample real time coal generation in CAISO is showed to illustrate that the coal usage in this region peaks during the evening hours [10]. Therefore, storage operation in net displaces coal while discharging during evening hours in California.

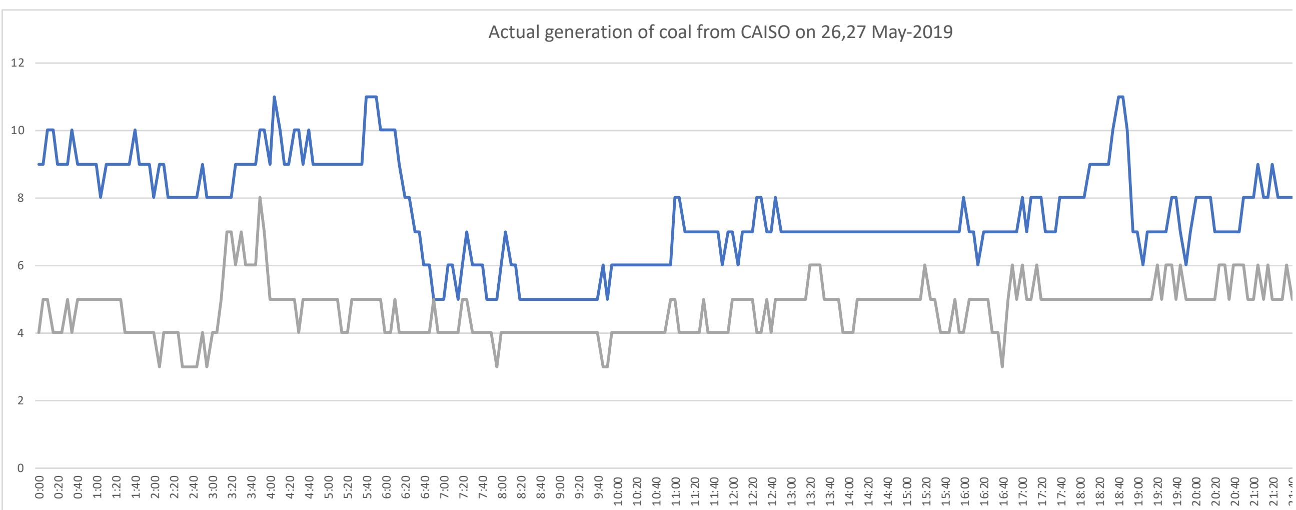

Fig. S5 Real time coal generation mix taken from CAISO website.

C4. Impact on generation from storage operation as a price-maker

In this section, Fig. S6 shows the change in generation resulting from the large storage system of capacity 5 GW in CAISO, MISO and NYISO regions. The change in generation is caused by two parameters, first- storage operation and second- economic retirement of the power plants because of the profit loss from the addition of bulk energy storage systems.

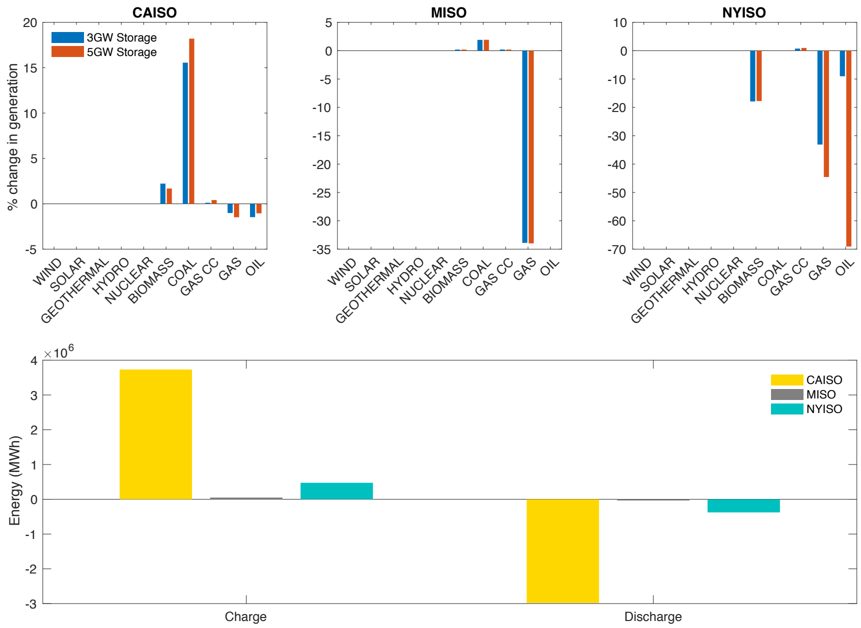

Fig. S6 Top two figures indicate an annual change in generation before and after adding storage per installed capacity of generation technology, expressed in MWh/MW-year.
X-axis represents the generation technology, Y-axis represents the change in generation/MW-year, and colors of the bar indicate different storage capacities varied from 4GW – 8GW. The top left figure is for Midcontinent ISO(MISO) in the Midwest region and the top right figure is for New York ISO(NYISO) in the NY region. Note that the y-axis range is different for MISO and NYISO. The bottom center figure indicates the annual energy consumed/discharged by a 5GW storage capacity. X-axis represents the charge (>0)/discharge (<0) state, Y-axis represents the energy in GWh, and the colors indicate the region.

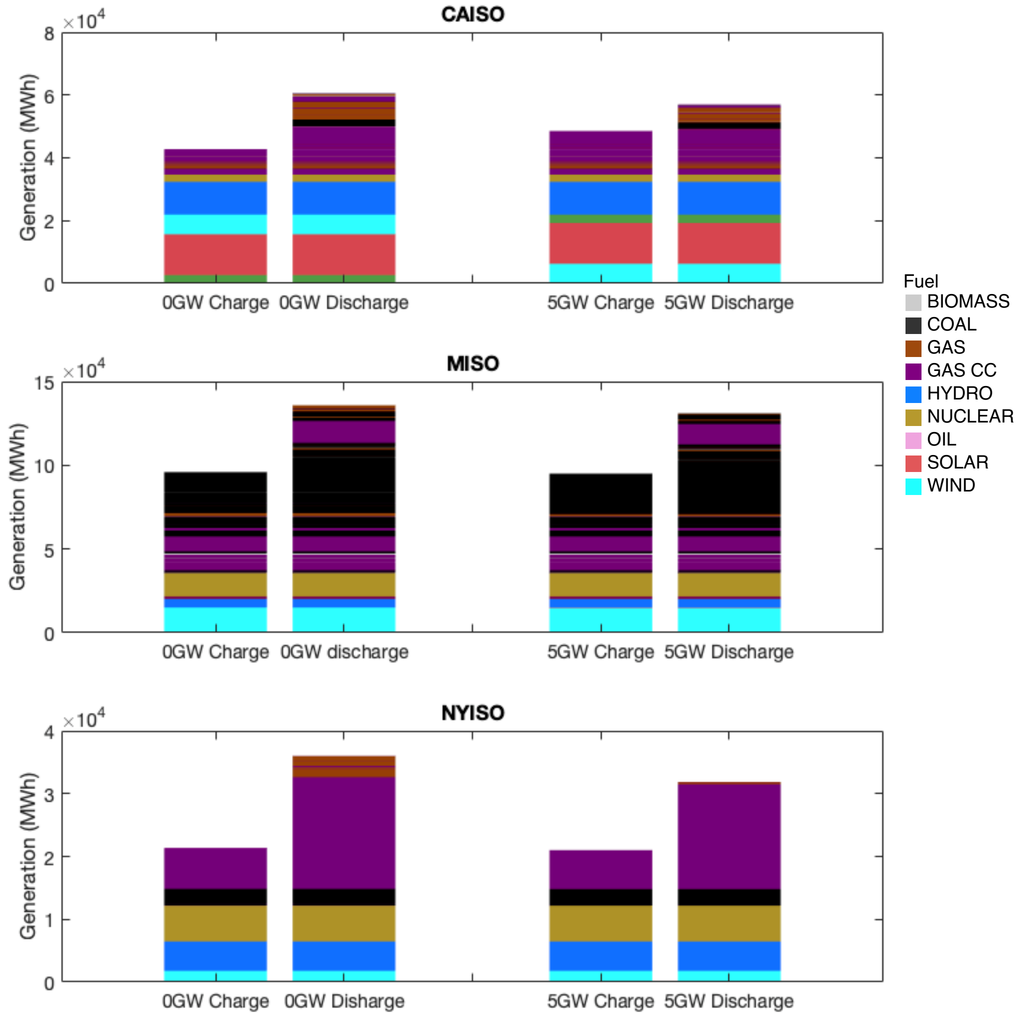

Fig. S7. Difference in dispatch stacks with and without storage, during the hours when storage charges and discharges.

Two sample hours on a typical summer day is chosen when storage charges and discharges. X-axis represents charge and discharge state of a storage capacity, Y-axis represents the generation in MWh, colors represent the generation technologies dispatched in the order to meet the demand. Labels represent the clearing price of the electricity at that given hour. The topmost figure is for Midcontinent ISO (MISO) and the bottom most figure is for New York ISO (NYISO).

Appendix References

1. Sandia National Laboratories, and Strategen Consulting: DOE Global Energy Database. US Department of Energy (2017). http://www.sandia.gov/ess/doe-global-energy-storage-database/ (accessed on 12- Nov 2016)
2. MISO: Historical Regional Forecast and Actual Load- Summary, Market Reports. MISO (2015), https://www.misoenergy.org/Library/MarketReports/Pages/MarketReports.aspx (accessed on 18- June 2016)
3. NYISO: Integrated Real-Time Load Data, NYSO (2015), http://www.nyiso.com/public/markets_operations/market_data/load_data/index.jsp (accessed on 18- June 2016)
4. EPA: Emissions & Generation Resource Integrated Database (eGRID), US EPA (2014), https://www.epa.gov/energy/eGRID (accessed on 3- Apr 2017)
5. EIA: Cost and Performance Characteristics of New Generating Technologies, Annual Energy Outlook 2017. EIA (2017), https://www.eia.gov/outlooks/aeo/assumptions/pdf/table_8.2.pdf, (accessed on 10- Aug 2017)
6. EPA: The Social Cost of Carbon, Environmental Protection Agency (2016), https://19january2017snapshot.epa.gov/climatechange/social-cost-carbon_.html
7. Draxl, C., Clifton, A., Hodge, B.-M., McCaa, J.: The Wind Integration National Dataset (WIND) Toolkit. Appl. Energy. 151, 355–366 (2015). doi:10.1016/j.apenergy.2015.03.121
8. National Renewable Energy Laboratory (NREL): Solar Power Data for Integration Studies. NREL (2006), http://www.nrel.gov/grid/solar-power-data.html (accessed on 20 April 2017)
9. Draxl, C., B.M. Hodge, A. Clifton, and J. McCaa.: Overview and Meteorological Validation of the Wind Integration National Dataset Toolkit, (2015)