MIT CEEPR

MIT Center for Energy and Environmental Policy Research

Working Paper Series

A Machine Learning Approach to Evaluating Renewable Energy Technology: An Alternative LACE Study on Solar Photo-Voltaic (PV)

BENNY SIU HON NG, CHRISTOPHER R. KNITTEL, AND CAROLINE UHLER

DECEMBER 2020

CEEPR WP 2020-021

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

A machine learning approach to evaluating renewable energy technology: An alternative LACE study on Solar Photo-Voltaic (PV).

by
Benny Siu Hon Ng
B.Eng. Engineering Systems and Design
Singapore University of Technology and Design, 2015

Submitted to the Institute for Data, Systems, and Society
and
Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degrees of
Master of Science in Technology and Policy
and
Master of Science in Electrical Engineering and Computer Science
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
May 2020

Author .....
Institute for Data, Systems, and Society
Department of Electrical Engineering and Computer Science
May 8, 2020

Certified by .....
Christopher R. Knittel
Professor, Applied Economics, MIT Sloan School of Management
Thesis Supervisor

Certified by .....
Caroline Uhler
Associate Professor, Department of Electrical Engineering and Computer Science
Associate Professor, Institute for Data, Systems, and Society
Thesis Reader

Accepted by .....
Noelle E. Selin
Associate Professor, Institute for Data, Systems, and Society, &
Department of Earth, Atmospheric, and Planetary Science
Director, Technology and Policy Program

Accepted by .....
Leslie A. Kolodziejski
Professor, Department of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

A machine learning approach to evaluating renewable energy technology: An alternative LACE study on Solar Photo-Voltaic (PV).

by
Benny Siu Hon Ng

Submitted to the Institute for Data, Systems, and Society
and the Department of Electrical Engineering and Computer Science
on May 8, 2020, in partial fulfillment of the
requirements for the degrees of
Master of Science in Technology and Policy
and
Master of Science in Electrical Engineering and Computer Science

Abstract

Currently, renewable technologies are often evaluated using the Levelized cost of electricity (LCOE), which is a measure of building and operating a generating plant over an assumed financial life and duty cycle. Naturally, instead of only measuring the cost, a more holistic approach would be to also assess the economical value of the renewable generating technology. One approach to this would be to measure the Levelized Avoided Cost of Electricity (LACE), which considers what it will cost the grid to generate electricity using renewable technology, amortized over its lifetime. However, estimating avoided cost can be challenging since it requires knowledge of how the renewable technology would perform in electricity generation, especially when taking into account a projected future period. Naturally this would have repercussions in policies adopting greater renewable technologies, further emphasising the importance of an adequate measure of evaluating renewable technology.

In this thesis, we explore several methods of evaluating alternative sources of energy, with an in-depth focus on a LACE evaluation of solar PV as an alternative source of electricity generation within CAISO market. Through experimentation of different variants of a recurrent neural network, an LSTM model was trained to predict 2016 electricity prices of all nodes within CAISO. The model achieved a Mean Absolute Scaled Error (MASE) of 0.761, outperforming a naive baseline using the Day-Ahead prices. Using the predicted prices, the LACE for solar PV was estimated and compared against the LACE computed with perfect knowledge of prices. Even though they had similar mean values, there was a significant difference in the variance. The effects of improvements in price prediction on the LACE was further explored. We found that the smaller the difference in the estimated LACE to the respective LCOE value, the greater the impact of improving price prediction performance; and was able to place an implicit value of an improvement of price prediction performance. Especially for policy and decision makers, this improvement in electricity price forecasting would directly translate to greater confidence when making the decision to switch a

solar PV alternative.

Thesis Supervisor: Christopher R. Knittel
Title: Professor, Applied Economics, MIT Sloan School of Management

Thesis Reader: Caroline Uhler
Title: Associate Professor, Department of Electrical Engineering and Computer Science
Associate Professor, Institute for Data, Systems, and Society

Acknowledgments

I would like to show my deepest gratitude to all everyone that have made my MIT journey such an enjoyable experience. I have had to go through very challenging experiences throughout this journey and I am glad to have been lucky enough to have be surrounded by many loving people, whose encouragement and company helped me learn and grow.

First and foremost, I wish to express my sincere appreciation to my supervisor, Professor Christopher Knittel, for his time and advice in helping me learn more about the energy space. As someone who was new to the sector but really looking to be more involved in energy and environment, I am grateful to be given a chance to be doing research under Professor Knittel, who's expertise in the area to me is second to none. His unique perspective in the various discussion topics as well as patience in helping me gain context in the energy space allowed me to craft this thesis, and hopefully allowed me to make some contribution to the energy and environment sector.

I would also like to thank my family's unwavering support throughout my life, but especially of my journey here at MIT. To my parents, who's sacrifice and encouragement allowed me to be here and develop at MIT, nothing I have today will be possible without your care and support. To my big bro and baby sis, from exercising together (#RT99) to helping me schedule my medical appointments, even setting up video sessions on our weekend visits to Yeye and Mama, you have allowed me to still 'be home' while being away. I really appreciate being close, even though I know it can be difficult with me being halfway across the world and a 12 hour time difference. To Darren, welcome again to the family, we often joke that you have replace me in the family, but I do gain comfort in knowing that my family has an extra person to count on, since I am not able to be there myself.

Last but not least, I would like to acknowledge my friends, whom I have had the pleasure of knowing, and having so much fun together. I would like to thank my best friends Shun Him and Ian, you guys have always been essentially family, and even more so here in MIT. We finally attained our Masters, through the countless parties we had. I have no doubt we will still be basking in our glory for many more years to count. To Bora, who was with me for all 3 years of my time here at MIT, we first met during the TPP open house back in 2017, and how time has flown since. I cannot thank you enough for all the help you have provided, from research and class advice, to 'discussions' we had during the El Clasico (we both know who REALly is the better team), even helping me get my first Data Science internship. I feel very lucky to be able to have met someone that shares so much interest and perspectives as me. To my boys at 'The Palace': What was the scene we said, and what a scene it was. Thanks for welcoming me to your group. I will always treasure the seshes we had, and tea we shared. Finally, to everyone else at TPP (special shout-out to Barb), it has been a pleasure meeting and getting to know all of you. I hope we continue to keep in touch and laugh about the good times we had.

I sincerely thank MIT for allowing me to develop and grow in the school, giving me the opportunity to meet passionate and intelligent from various backgrounds. Special thanks goes out to Professor Caroline Uhler, who has helped me build strong foundations in Computer Science and has kindly agreed to read my thesis.

I look forward to beginning the next stage in life, as I complete this thesis, with fresh perspectives and newly acquired knowledge, hoping to making an even greater impact to the world.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

According to "The future of solar energy: An interdisciplinary MIT study" [50], solar energy electricity is one of the few low-carbon technologies that has the potential for successful large scale implementation. In order to fulfill this potential, it is important to first evaluate solar technology compared to current methods of electricity generation. Currently, renewable technologies are often evaluated using the Levelized cost of electricity (LCOE), which is a measure of building and operating a generating plant over an assumed financial life and duty cycle [63]. Naturally, instead of only measuring the cost, a more holistic approach would be to also assess the economical value of the renewable generating technology. One approach to this would be to measure the Levelized Avoided Cost of Electricity (LACE), which considers the value to the grid from electricity generated by renewable technology, amortized over its lifetime. However, estimating avoided cost can be challenging since it requires knowledge of how the renewable technology would perform in electricity generation and the value of that electricity, especially when taking into account a projected future period [55].

In this thesis, the central aim is to answer the following questions: How would different forms of solar Photo-Voltaic (PV) such as roof-top solar, residential and utility scale panels, compare against conventional electricity generating sources? Can we better leverage Machine Learning and Optimization techniques, and measure the current benefits of solar technologies over its operating lifetime? Using data at an Independent Systems Operator (ISO)-level, with currently available electricity price data, as well as geographical information (such as weather and amount of sunshine), we will attempt to implement current state of the art machine learning techniques to forecast future electricity prices and geographic conditions. We will leverage on these forecast to determine how and where we can potentially position solar PV panels at a specified ISO region (e.g. California). Finally, under these conditions, we will measure the LACE of solar PV, among other renewable energy technologies, which can be used in discussions for greater policy support of solar PV as an alternative source of generation of electricity.

1.2 Thesis Outline

This thesis proceeds as follows: Chapter 1 introduces the motivation and provides the thesis outline. Chapter 2 provides a background on the current methods of evaluation renewable technology as well as introduction of current state of the art methods of Electricity Price Forecasting (EPF) used in the industry. Chapter 3 focuses on the computation of LACE on a Solar PV. This includes how LACE is computed and provides different LACE evaluation under different conditions, such different prediction performances of electricity prices under an expected solar output. Chapter 4 would then use the values computed and compare how improvements in price prediction would affect the valuation of solar. We will further explore if we can determine a break-even cost for the different conditions. Finally, we conclude with Chapter 5 where we will summarize the impact of improvement in EPF on a LACE evaluation of solar, and discuss further oppotunities and policy implication.

Chapter 2

Renewable Technology Evaluation

This chapter is intended to give the reader an overview of how different evaluation methodologies, specifically the LCOE and LACE, can be used to assess the economic competitiveness of renewable energy generation technologies, affecting investment and policy decisions [55].

2.1 How Evaluation of Renewable Energy can Affect Policy

In 2018, about 11% of the US energy consumption was generated by renewable energy with biomass, hydroelectric, wind, solar and geothermal making up most of the renewable energy sources [56]. However, the relative mix between the different sources of electricity generation changes and can largely be attributed to the relative cost and benefits between the various renewable technologies [18].

There are currently several methods to compare between power generation technologies [46]:

1. **Capacity vs Energy**
   We can compare between energy sources by looking at their capacity or energy discharge output. Capacity is measured in kilowatts (kW) or megawatts (MW) while energy is measured in kilowatts-hours (kWh) or megawatt-hours (MWh). The capacity refers to the maximum output an electricity generator can physically produce. Energy is the amount of electricity a generator produces over a specific period of time [30]. Although technology can be compared between their respective capacity and energy, typically, the cost of the technology affects its adoption [62]. This often leads to comparisons of capacity or energy together with its cost. For example, the costs/kW or costs/kWh.

2. **Initial Capital Cost**
   There are several ways to compare cost between renewable technologies. One of them would be to compare the initial capital cost. Higher capital cost is often associated with higher risk, which is an important factor for deterring investments in new technology [25]. This is not surprising since higher investment

cost would limit the opportunity to investors with sufficient upfront capital resource. Furthermore, it would take a longer time horizon to break even on the investment and reap its rewards [52]. In general, base load technologies will be more expensive to build but would have lower operational cost as compared to intermediate or peak load technologies that might have lower initial capital but higher operational cost [40].

Base load power sources are plants that operate to meet the minimum level of power demand at all times [15]. These plants generally produce power at a constant rate and are not designed to respond to peak demands or emergencies. Besides cost factors, these base-load technologies also characterized with low risk of interruption. Within US, coal and nuclear plants, together with hydropowered plants and geothermal plants (if available), are generally associated as base load power plants. Renewable energy sources such as wind and solar are typically examples of intermediate load plants due to their dependencies on natural weather conditions [42].

3. **Levelized Cost of Energy (LCOE)** As part of evaluating the overall competitiveness of different generating technologies, a more comprehensive view of the overall cost would require the operation cost of the respective technology to be taken into consideration, in addition to the initial capital cost. The LCOE is one such measure and represents the per-kilowatthour cost (in real dollars) of building and operating a generating plant over an assumed financial life and duty cycle [63, 55]. The U.S. Energy Information Administration (EIA) provides an annual report that determines the LCOE of key energy generating sources such as solar, wind, hydroelectric, coal, geothermal, biomass, nuclear [53]. Key inputs to calculating include capital cost, fuel cost, fixed and variable operations and maintenance (OM) cost, financing cost, and assumed utilization rate for each plant type [55]. The computation of the LCOE in the EIA annual report takes the following [54]:

$$LCOE = \frac{FCF * CC + OM_{fixed}}{\mathbb{E}(G_{hours})} + OM_{variable} + fuel \tag{2.1}$$

Where:

- $LCOE$ is the levelized cost of electricity, expressed in units of $/Megawatthour ($/MWh)
- FCF is the Fixed Charged Factor that annualized the capital cost, accounting for the weighted average cost of capital (return on debt and return on equity), Federal tax burden for the project, and the expected financial life of the project. This factor is estimated using a cash-flow model with the National Energy Modeling System (NEMS), and varies over time, based on changes on the cost of debt and cost of equity.
- CC is the Capital Cost. This is the initial investment per unit of capacity in the project, expressed in $/Megawatt ($/MW). Depending on the tech-

nology, this cost may be different due to several factors such as declining technology cost due to learning and adjustments from border economic factors.

- OM_{fixed} is the fixed O&M cost. This measures the annual expenditure per unit of project capacity for operations and maintenance, expressed in \$/MW/year. This cost is relatively constant, regardless of plant utilization levels, such as worker salaries and maintenance or refurbishment cost that are scheduled on a calendar basis (not operating-hour basis)
- $OM_{variable}$ is the variable O&M. This is the expenditure per unit of generation of operations and maintenance expressed in \$/Mwh. This expenditure includes cost that are relted to the operating hours of the equipment, such as consumable maintenance items and refurbishment cost that are scheduled based on operating hours.
- $\mathbb{E}(G_{hours})$ is the annual expected generation hours. This is determined by calculating the number of hours in a year that the plant is assumed to operate. For dispatchable genration such as coal, nuclear, or gas-fired plants, EIA calculates this based on an annual capacity factor that corresponds to the maximum annual availablity for that unit. For peak load units, this number is assumed to be 30% of annual capacity factor. For intermittent renewable sources, this calculation is based on loaction-specific resource availability.
- Fuel is the expenditure for fuel, expressed in \$/MWh. This is the product of the heat rate of the equipment and the fuel price. These cost represent the hourly average of the long-term fuel cost over the assumed financial life of the equipment (not just for the single year of estimation).

4. **Levelized Avoided Cost of Energy (LACE)** Instead of looking at the cost of the different energy generation technology, it is only natural to look at benefits between them. One way of doing so would be through consideration of the avoided cost, which measures what it would cost the grid to generate the electricity that is otherwise displaced by a different generation technology. The LACE can be seen as a proxy of the economic value gained of the alternative technology, summed over its financial life and converted to a stream of equal payments annually [55]. A simpler adaption of the LACE computation by the EIA takes the following form [54]:

$$LACE = \frac{\sum_{t=1}^{Y}(P_t * D_t) + C_p * C_c}{\mathbb{E}(G_{hours})} \tag{2.2}$$

Where:

- LACE is the levelized avoided cost of electricity, expressed in units of \$/MWh

- P_t is the marginal generation price at time t. This is also the cost of serving load to meet the demand at time t. While this price is typically determined by the variable cost (fuel cost plus variable O&M) of the most expensive generating unit that needs to be dispatched to meet energy demand. This price might also be impacted by the cost of meeting any environmental or portfolio policy requirements by the marginal generators. The marginal general price plays a significant role in the LACE calculation, which we will further explore in the remaining parts of the paper.
- D_t is the activity at time t and Y is the number of time steps in a year. In calculations used in this paper, t is measured in hours and Y would be approximately 8760 hours (in a year).
- C_p is the capacity payment. This is the value to the system of meeting the reliability reserve margin. It is the payment that would be required to incentives the last unit of capacity needed to satisfy a regional reliability reserve requirement.
- C_c is the capacity credit. This is the ability of the generating unit to provide system reliability reserves. For dispatchable units, which are generating units that can generating electricity at anytime of the day regardless of external factor conditions (like the weather). For intermittent renewable technologies (like solar or wind), the capacity credit is a function of the availability of the resource during peak load periods and the estimated probability of correlated resource -derived outages within a given region. An example of this would be the capacity credit of winde is the probability that if wind is not blowing within a certain region.
- $\mathbb{E}(G_{hours})$ is the annual expected generation hours. This is similar to the definition as described for LCOE.

There are many measures of evaluating these energy sources. Depending on which measure we use as our metric, the computed measure will compare differently between the different generating technologies and affect how we evaluate them. For example, among the different ways of comparison, the EIA have a greater focus on the LCOE and more recently the LACE to evaluate the economic value of different electricity generating technologies [53]. Their results are regularly used by the U.S Department of Energy Office of Scientific and Technical Information to help energy companies evaluate and finance industrial renewable energy measures [49]. There are also many different studies that uses the LCOE as a measure to build models. One of such examples was a model by Bruck, Sandborn and Goudarzi, which used the LCOE to evaluate wind as a generating technology for wind farms [11]. There is also another case where Clauser and Ewert uses the LCOE to compare geothermal electric energy to other primary electricity generating sources [13]. These studies have further implications towards policy discussion and will play a significant role in both state and federal energy efficiency programs [6]. As we aim to increase penetration of renewable energy generation, which measure is selected for comparison become even more important. Currently, we can see how this has already affected the formulation

and implementation of energy policy and it is important that we used relevant and appropriate evaluation methods and results.

2.2 LCOE as a Measure of Evaluation

The LCOE is currently still a popular metric of evaluation between different sources of electricity generation among both government as well as private sectors [55, 53, 6]. A big part for this is due to its simplicity of calculation and visual appeal [58]. Its widespread adoption further enhanced the LCOE as the preferred choice for comparison among energy stakeholders, as there is a common measure to evaluate between technology [3]. However, in more recent times, domain experts are finding that LCOE might not be as useful generally appropriate when considering unconventional resources like wind and solar [12]. This results in more decision makers to look at alternative metric (e.g. LACE) when it comes to comparing the alternative electricity sources [43]. There have been studies done comparing the LCOE against other measure of assessments highlight several issues of LCOE [3, 58, 53]:

1. LCOE only considers the cost to build and operate a plant and does not consider the value side of the energy generated. By not considering the revenue or value of electricity generated, this assumes that all technologies provide similar service. However, this is not necessarily the case and if we consider both the demand and supply of energy, we can already start to see how the value of electricity can differ. For example, the value of electricity might be higher at night when there is higher usage demand and if a clean alternate form of energy generation can provide electricity at this time, its value would be greater.

2. LCOE relies heavily on several assumptions such as the discount rate, inflation effects and the sensitivity of results to uncertainty in future commodity cost. With high variability of different commodity (e.g. fuel gas prices), there would be greater implications to the range of LCOE for the different technologies.

Taking these factors into consideration, even popular sources of LCOE, such as Lazard [36] and the Energy institute of The University of Texas at Austin [47], report LCOE numbers with detailed caveats and context. Figure 2-1 shows Lazard's LCOE as a range depending on various assumption.

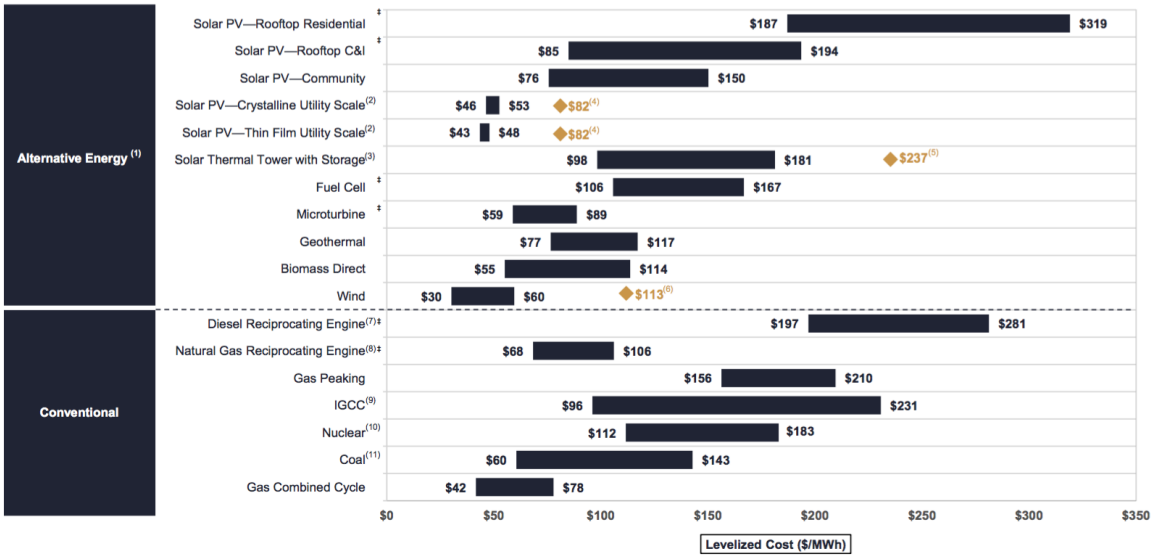

Figure 2-1: Adapted from Lazard's Levelized Cost of Energy Analysis - Version 11.0, showing the cost competitiveness of alternative electricity generation sources relative to conventional power sources [36].

2.3 What is Preventing LACE valuation of Renewable Energy

As we saw in the computation of the LACE for different renewable sources, this requires the marginal generation price as well as the amount of electricity that will be generated by the specific generation technology. This results in a more complex estimation compared to other measures, such as the LCOE; since the LACE is essentially an estimation of the value of the specific generation technology project to exist at a future date [53]. There are two main challenges with this: 1) predicting future electricity prices, and 2) estimating expected output of electricity generation with the technology in question.

2.3.1 Future price prediction

Electricity price forecasting (EPF) has been tried for almost two decades [59] but has only achieved limited success. In general, across all the different electricity markets in the U.S., the electricity markets entails the following characteristics [4]:

1. high frequency,
2. non constant mean and variance (non stationary series),
3. multiple seasonality, both daily and weekly periodicity,
4. calendar effect (e.g. weekends and holidays),
5. high volatility,

6. high percentage of unusual prices (especially during period of high demand).

The electricity market is heavily dependent on a variety of factors, such as the weather, natural disasters, or technology advancements. Due to these characteristics, it makes it extremely difficult to predict electricity prices. In the subsequent sections, we will comparing the myriad of predictive models, comparing their strength and weakness, to determine in more detail what are some of the respective challenges of price forecasting. Without the ability to accurate determine what the future prices will be, the reliability of the LACE is undermined. This would have further ripple effects on the popularity and confidence of LACE as a metric for comparison within the electricity market.

2.3.2 Estimating electricity generation

There are currently several main generation sources within the electricity market that provides the supply required to meet the electricity demands of the market. Figure 2-2 shows the mix of generation sources within the U.S since 1950. While we see that there is increasing trend for electricity demand, we also notice that the contribution mix by the various energy sources changes. This is perhaps not surprising, due to advancement of technologies and environmental policies, to see greater contribution from renewable technology as a percentage of the over electricity generation. In more recent years, we can see from figure 2-3 wind has become a major driver of electricity generation and this has significant impact in evaluating how much benefits the alternative technology can bring. Similary, LACE uses electricity price as a measure of economic value and predicting this after incorporating the new technology. This has traditionally been difficult to determine as we have no past information on how this technology will do. Currently, early studies done by Inzunza and Knittel estimates how higher penetration of renewable technology can affect electricity output from the rest of generating sources and subsequent electricity prices [28]. We can leverage on these behaviours to better estimate electricity generation outputs and thus more realistic computation of the LACE for comparison in the subsequent chapters.

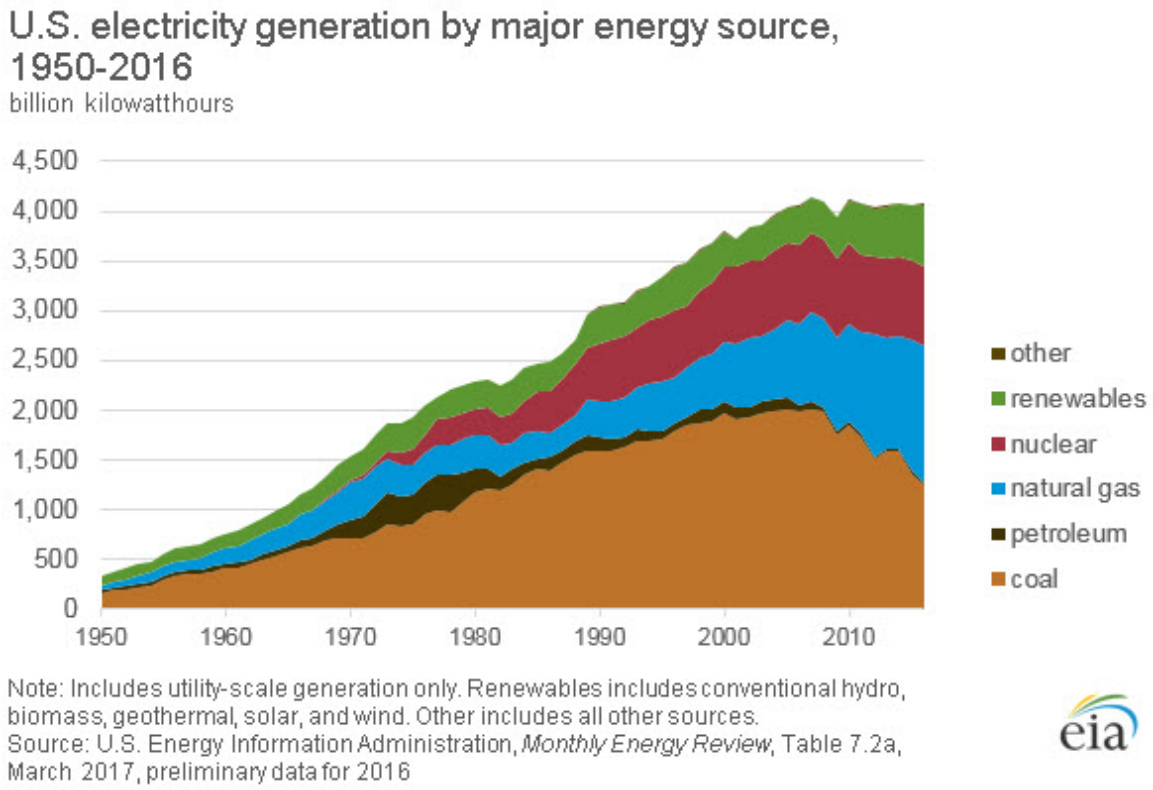

Figure 2-2: Graph showing the amount of electricity generated from various energy sources from 1950 to 2016. Green part indicates the generation by renewables, which has been increasing over time due to climate change and environmental pressure, as well as cost pressures [1].

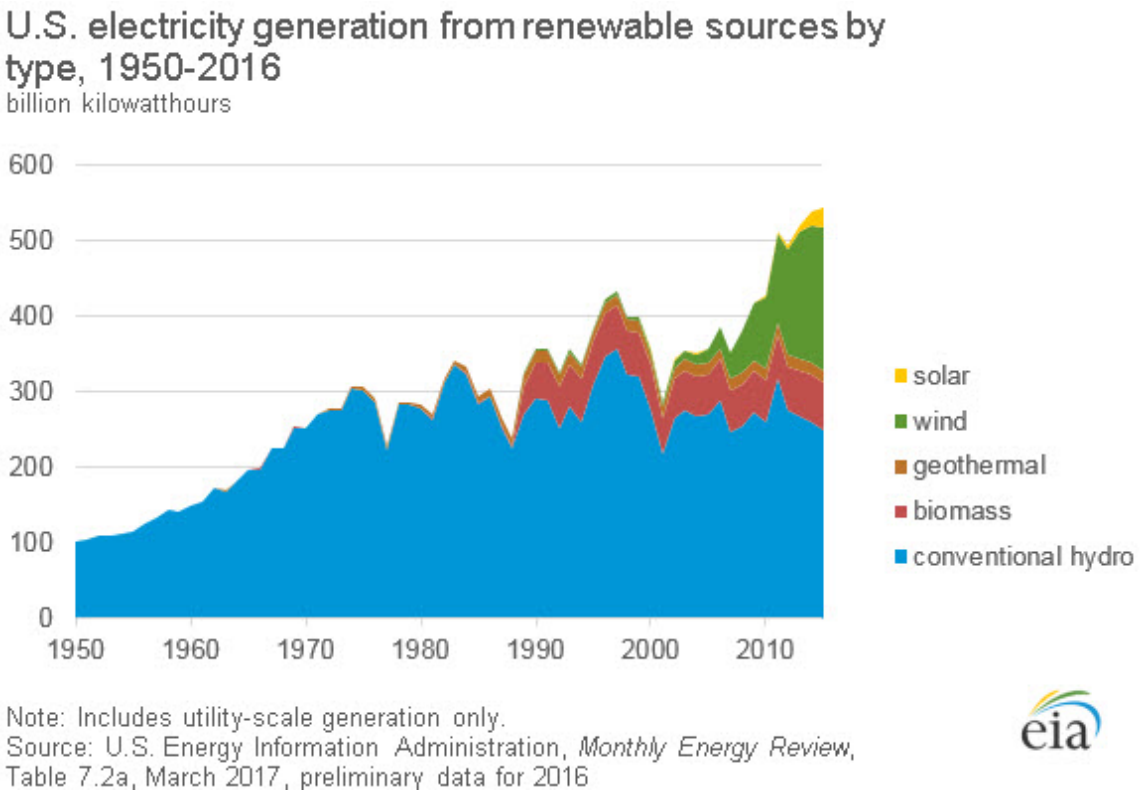

Figure 2-3: Graph showing that wind has been the major driver within the growth of new generation capacity [1].

2.4 Current Methods in Electricity Price Forecasting (EPF)

Finally, we introduce some of the current state of the art methods that are used in EPF to get a better understanding of how the different method types would can affect respective evaluation computations.

Based on a review conducted by Weron in 2014 [59], most of the EPF techniques originate from the fields of Econometrics and Electrical Engineering. Figure 2-4 shows details of some of these models which range from the following:

1) Multi-agent models:
   Examples include multi-agent simulation, equilibrium, game-theoretic models

2) Fundamental models:
   Structural models that include demand and supply modelling

3) Reduced-form models:
   Econometric models that can be quantitative or stochastic type in nature

4) Statistical models:

Econometric models that involves time-series analysis, technical analysis of financial and commodity markets

5) Computational Intelligence models:
Non-parametric or non-linear models which include traditional Machine Learning models as well as Neural Networks

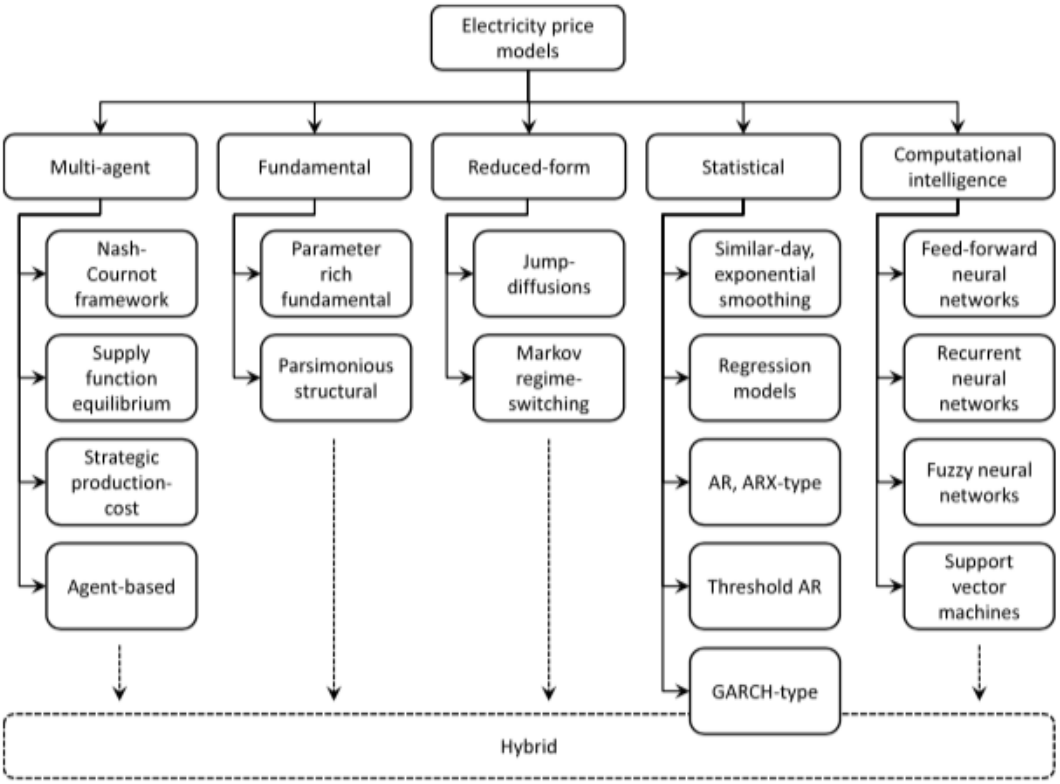

Figure 2-4: Taxonomy of electricity spot price modeling approaches [59].

2.4.1 Multi-agent models

As mentioned previously, the competitive electricity markets entails high levels of uncertainty and involves great interaction between many different parties from suppliers to consumers. This provides an opportunity for equilibrium (game theoretic) approaches, commonly viewed as generalizations of cost-based models, amended with strategic bidding considerations. These models are especially useful in predicting expected price levels in markets with no price history, but known supply costs and market concentration.

Figure 2-4 shows some examples of multi-agent approaches, which include: Nash-Cournot framework, Supply function equilibrium, Strategic production-cost models and Agent-based simulation models. These class of agent-based simulation techniques can address features of electricity markets that static equilibrium models ignore, and

are extremely flexible tools for the analysis of strategic behavior in electricity markets. However, they require the assumptions embedded in the simulation to be justified, both theoretically and empirically, such as who the players are, potential strategies, ways they interact and set payoffs. These are typically subjective, and can have a great impact on the results.

Finally, multi-agent models are generally more useful in understanding qualitative issues rather than quantitative results. They may provide insights as to whether or not prices will be above marginal costs, and how this might influence the playersoutcomes [59]. However, as in our case of using the predicted prices to determine the LACE, a more quantitative results is required and require electricity prices to be predicted with a high level of precision. Thus, we will opt for an alternative approach of prediction.

2.4.2 Fundamental models

Fundamental models, also known as structural models, are modelled to capture the basic physical and economic relationships present in the production and trading of electricity. They involve understanding the functional associations between fundamental drivers such as electricity loads, weather conditions, system parameters, etc. Often they rely on statistical, reduced-form that have the same foundations as computational intelligence techniques.

There are 2 main sub-classes of fundamental models (figure 2-4):

1. Parameter-rich fundamental models:
   Often developed as proprietary, in-house products and not publicly disclosed. This subclass model involves using specific features to explain and predict various outcomes (e.g. hydro inflow, snow and temperature conditions to explain spot price formation [16]).

2. Parsimonious structural models:
   The Parsimonious structural subclass models typically begins with an empirical analysis of market supply and demand curves. We then builds the spot price process by applying the inverse of the BoxCox transformation (which includes an exponential function as a special case) to an Ornstein-Uhlenbeck process (see Eqn 2.3 below). With the obtained jumpless spot price model which can exhibit spikes, we calibrates it to data from the respective electricity markets [5].

The Ornstein-Uhlenbeck process equation is as follows:

$$dX_t = (\alpha - \beta X_t)dt + \sigma dW_t, \tag{2.3}$$

where $X_t = D_t - \bar{D}_t$ is the horizontal stochastic deviations, D_t is the demand at time t, $\bar{D}_t$ is the mean demand at time t, β is the speed of mean-reversion, $\frac{\alpha}{\beta}$ is the long term mean-reversion level, σ is the volatility and dW_t are the increments of a standard Wiener process(i.e.,Brownian motion).

Fundamental models have high dependency on the availability of data that is fed into the model. Since typically fundamental data are collected over longer intervals (e.g. weekly or monthly), pure fundamental models are better suited for medium term prediction horizons (as apposed to short term prediction). Also, since we incorporate several stochastic fluctuations of the fundamental drivers, the specific assumptions made regarding the physical and economic relationship of the market place becomes extremely important. The price projections generated by the models are very sensitive to violations of these assumptions with the more detailed the model output, the more effort needed to adjust the parameters [59].

2.4.3 Reduced-form models

Reduced-form models are different from other models in that their main purpose is to replicate the main characteristic of daily electricity prices rather than to only make accurate predictions [59]. These models usually play a key role in price derivatives and risk management system.

There are a couple of main reduced-form models (figure 2-4). They are

1. Jump Diffusion model
   Within energy economics literature, they are known as special cases of general stochastic differential equation(SDE) for the increment of the (deseasonalized and detrended) spot electricity price (X_t):

   $$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t + dq(X_t, t), \tag{2.4}$$

   where dW_t are the increments of a standard Wiener process (i.e., Brownian motion) and $dq(X_t, t)$ are the increments of a pure jump process. There are other variations proposed by Albanese et al. [2] and Geman and Roncoroni [19], which utilize reduced-form models to model European and Bermudan electricity derivatives.

2. Markov regime-switching (MRS)
   MRS models takes into account consecutive spikes of prices, which jump diffusion models cannot. The idea underlying MRS is to represent the observed stochastic behavior of a (deseasonalized and detrended) spot price process X_t by L separate states or regimes with different underlying stochastic processes $X_t, j, j = 1, ..., L$. The switching mechanism between the states is assumed to be an unobserved (latent) Markovchain R_t governed by the transition matrix P containing the probabilities $p_{ij} = P(R_{t+1} = j|R_{t=i})$ of switching from regime i at time t to regime j at time $t + 1$:

   $$\boldsymbol{P} = (p_{ij}) = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1L} \\ p_{21} & p_{22} & \cdots & p_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ p_{L1} & p_{L2} & \cdots & p_{LL} \end{bmatrix}, \tag{2.5}$$

with $p_{ii} = 1 - \sum_{i \neq j} p_{ij}$.

Due to the Markov property, the current state R_t at time t depends only on R_{t1}. Although L regime models can be considered typically two or three regimes are sufficient to model the dynamics of electricity spot prices [32, 57].

Generally, the reduced-form models have been shown to perform relatively well for volatile or price spike forecast. However, they are not expected to predict hourly prices accurately and are heavily dependent on the price process (e.g. like marginal distributions at future time points, price dynamics, and correlations between commodity prices) chosen to predict daily electricity prices. This approach while useful for risk analysis and understanding of the general characteristic of price patterns [8], would not be as useful in the computation of LACE, which is sensitive towards having accurate future prices.

2.4.4 Statistical models

Price forecasting through statistical models are done via mathematical combinations of previous prices and/or previous or current values of related exogenous factors. Examples of these exogenous factors could be, but not limited to, production figures, weather variables, etc. Statistical models typically fall within 2 main categories:

1. Additive
   Predicted price is a linear sum (additive) of a number of components (features).

2. Multiplicative
   Predicted price is the product(multiplicative) of a number of components (features).

Figure 2-4 shows some popular examples of statistical models. Details of these methods will not be discussed here but highly encouraged to be read for understanding of specifics. In general, statistical models are popular due to the robustness of modelling many different problems [33, 35] and ease of interpretability of the model, allowing better understanding of price behaviours. This, however, requires huge dependency on the numerical efficiency of the algorithms employed, and also the quality of the data used for analysis. The ability to incorporate important fundamental factors,such as historical demand, demand and consumption forecasts, weather forecasts or fuel prices is extremely important but without any formal methods to know which factors to include. Furthermore, they are often criticized for their limited ability to model the (usually) nonlinear behavior of electricity price sand related fundamental variables [59]. Their performance are often compared against modern techniques in computational intelligence, to be further discussed.

2.4.5 Computational Intelligence

Computational Intelligence comes in many form in recent years but essentially describes models outside the realm of traditional statistical models. For this thesis,

specifically we are interested in Machine Learning and Artificial Intelligence field and will focus specifically on the Neural Networks.

Neural Network (NN) Architecture

A typical NN architecture consist of various neurons that form a model that contains various linear and non-linear combination. According to the universal approximation theorem by Cybenko in 1989 [14], even a basic feed forward NN with finite neurons can approximate any continuous functions within a subspace of $\mathbb{R}^n$. Figure 2-5 shows two common types of neural networks: 1) Feed Forward Networks (FFN) and 2) Recurrent Neural Networks (RNN).

NN models are excellent in modelling complex and non-linear problems, which are typical of EPF problems. Yet, at the same time, as in Markov model discussed above, the ability to model spiky, non-linear problems might not necessarily result in better forecasts [9]. It also requires large amounts of data but will be able to provide better interval and density forecasts than the linear models [59]. Currently, these methods have shown to produce promising results predicting electricity prices in the Pennsylvania-New Jersey-Maryland Interconnection (PJM) market [60]. In subsequent chapters, we will further discuss details of these methods, using them to predict electricity prices in the California Independent System Operator (CAISO) market.

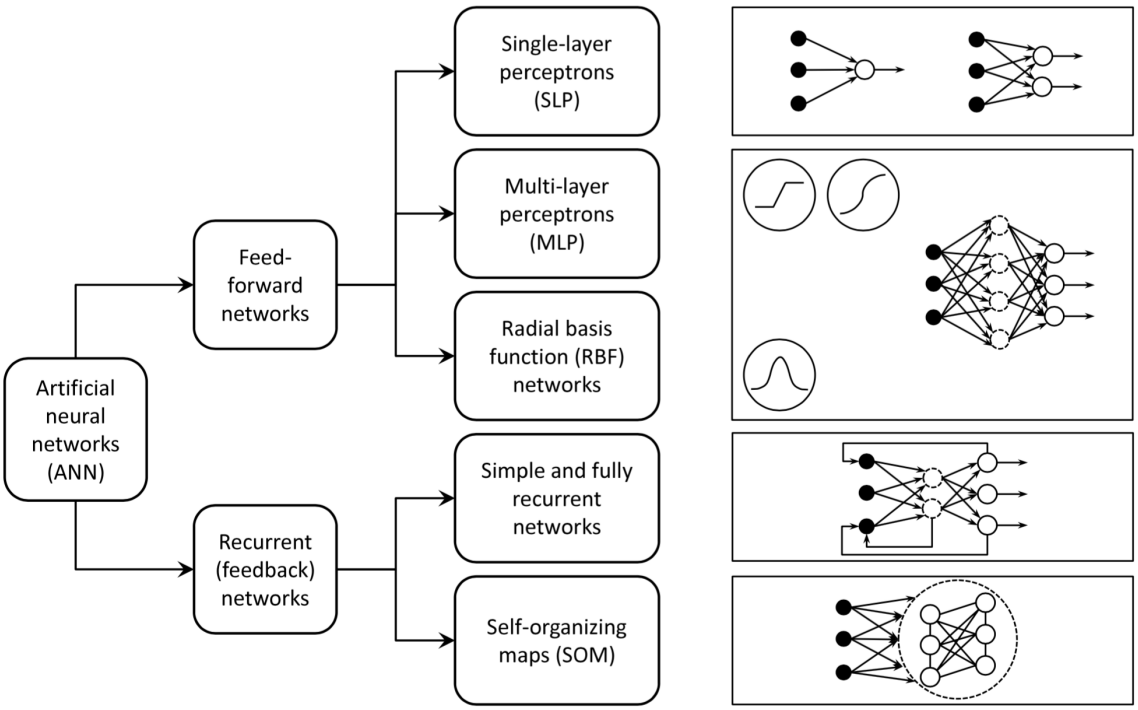

Figure 2-5: Taxonomy of Neural Network models, split mainly by their architecture [59]. The input nodes are denoted by filled circles, output nodes by empty circles, and hidden layer nodes with dashed-line circles.

Chapter 3

LACE Evaluation of Solar PV

3.1 Introduction

Previously we saw how the commonly used LCOE measure does not consider the value side of energy generated by the different technologies. LACE would then be a good alternative, despite its difficulties, as it measures the value of energy the generation source produces. In this chapter, we estimate electricity prices and respective potential output, to compute the LACE of CAISO nodal data and evaluate solar as an electricity generating source. The results are then discussed in the next chapter where we will gain an in depth understanding on how sensitive the LACE is towards predicted prices.

3.2 Prediction of CAISO Prices Methodology

3.2.1 Predicting electricity prices with Neural Networks

Based on prediction methodology developed by Wong et al. [61, 60] used to predict electricity prices in the PJM market, we will similarly predict the electricity prices in the CAISO market. We begin by detailing the methods used below:

1. **Feed Forward Networks (FFN)**

   FFN are the most basic architecture types of a NN. A basic FFN passes information in a forward fashion, and typically consist of an input layer, usually the size of the feature vector $x_{\text{input}} \in \mathbb{R}^d$, an intermediate layer with hidden nodes, which takes the output of the input layer, puts the input through an affine function, followed by an activation function, usually a sigmoid or tanh (see Figure 3-1 below for an illustration). The output layer then takes all the intermediate layer inputs and performs another affine function transformation [26]. For classification problems, the output layer usually is a softmax layer, which applies the normalized exponential function to the outputs of the hidden layer. For regression problems such as EPF, the output layer does the affine transformation and ends with 1 output (e.g. the next day RT price). The FF

networks are trained by back-propagation, which is a gradient descent method. For continuously valued functions and regression/classification tasks, one can update the parameters of the system such that they minimize a specified loss function. There is a range of loss function that can be selected, depending on the specific prediction problem. For our specific EPF problem, an ℓ^1-type loss function (Eqn. (3.1)) was chosen for backpropagation of training error. One reason for this is due to the L1 loss generally being more robust and not as affected by outliers as compared to other loss functions. For example, an ℓ^2 loss might try to adjust the model according to outlier values, even at the expense of other samples. The ℓ^1 loss was also selected as the metric for optimization of hyperparameters based on error with a validation data set, to be discussed further in the paper.

The ℓ^1 loss function is given by:

$$\text{loss}(\boldsymbol{t}, \boldsymbol{y}) = \frac{1}{n} \sum_i |t_i - y_i|, \tag{3.1}$$

where $\boldsymbol{t}$ is the predicted prices of the model and $\boldsymbol{y}$ is the target prices from the training or validation set (depending on whether the model is being trained or tuned, respectively).

Equation 3.2 represents the mathematical representation of the input vector x_t, hidden layer h_t, and output layer o_t and the connections between the different types of nodes in matrix vector form, namely:

$$\begin{aligned} h_t &= \sigma(W^{x,h} x_t), \\ o_t &= W^{h,o} h_t, \end{aligned} \tag{3.2}$$

where $W_{a,b}^{x,h}$ represents the weight multiplied used for the affine transformation of the input vector $h_t^{(b)} = \sum_a W_{x,h}^{(a,b)} x_t^{(a)}$.

2. **Recurrent Neural Networks (RNN)**

An RNN is a class of neural network which the connection between units form a directed cycle. By keeping track of a learning state of the RNN as data are passed through this cycle, RNN models can exhibit "memory-like" behavior and dynamically learn sequences of data to capture certain temporal dynamics, as well as exhibit dynamic temporal behavior. In sequence modeling problems, RNNs tend to far outperform traditional feedforward MLPs, since they allow for information across values of inputs to be stored in the network. That is, information from the past is stored in the system, and affects the future outputs of the model.

Figure 3-2 depicts an illustration of the RNN. We can see through a set of "context units" how the RNN is able to maintain and update the state network. The nodes in the hidden layer are connected to the context layers, which are

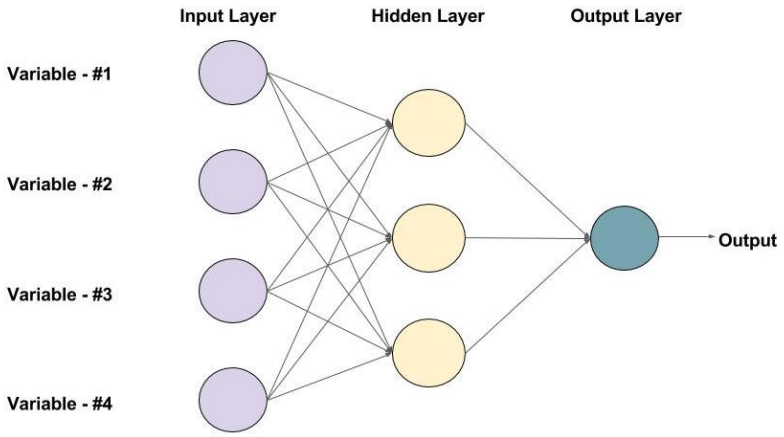

Figure 3-1: Example of simple feedforward network, adapted from [22].

then connected to the output layer. As a result, each of the neurons in the hidden layer processes both the external input signals and signals from the feedback.

Equation 3.3 shows the vector representation of a basic RNN. With input vectors x_t and hidden state h_t, the updates follow the following equations:

$$h_t = \sigma(W^{h,h} h_{t-1} + W^{h,x} x_t), \tag{3.3}$$

where $\sigma(\cdot)$ is a non-linear activation function. The additional parameters learnt in RNNs include the 2 weight matrices $W^{h,h}, W^{h,x}$.

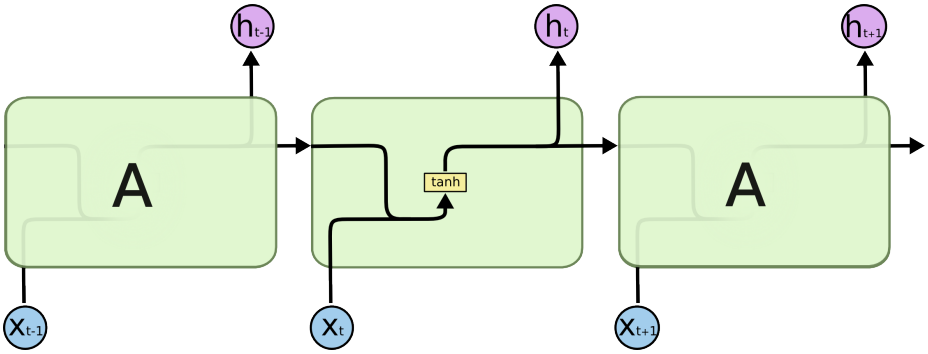

Figure 3-2: The internal workings of RNN cells [41]. In the above diagram, each line carries an entire vector, from the output of one node to the inputs of others. In addition, the figure depicts tanh as the activating function, but any non-linear activation function can be similarly represented.

Gated Recurrent Unit (GRU)

GRUs are an extension of RNNs, where there is a gating mechanism in the RNN. With input vectors x_t and hidden state h_t, the updates follow the following

equations:

$$
\begin{aligned}
g_t &= \sigma(W^{g,h} h_{t-1} + W^{g,x} x_t) \\
h_t &= (1 - h_t) \odot h_{t-1} + g_t \odot tanh(W^{h,h} h_{t-1} + W^{h,x} x_t),
\end{aligned}
\tag{3.4}
$$

where $\sigma(\cdot)$ is the sigmoid function. The additional parameters learnt in GRUs include the 2 weight matrices $W^{g,h}, W^{g,x}$ in addition to those learnt in a basic RNN.

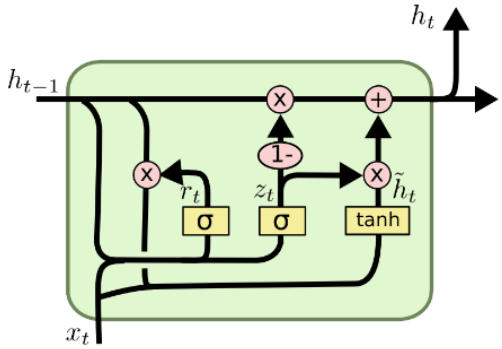

Figure 3-3: The internal structure of a GRU cell [41]. In the above diagram, each line carries an entire vector, from the output of one node to the inputs of others. The pink circles represent pointwise operations, like vector addition, while the yellow boxes are learned neural network layers.

Long-Short Term Memory (LSTM)

LSTMs are further generalizations of RNNs, where there are multiple gating mechanisms, including the forget gate, the input gate and the output gate, and were introduced in 1997 by Hochreiter [24] and in 1994 by Bengio [7].

With input vectors x_t and hidden state h_t, the updates follow the following equations:

$$
\begin{aligned}
f_t &= \sigma(W^{f,h} h_{t-1} + W^{f,x} x_t) \\
i_t &= \sigma(W^{i,h} h_{t-1} + W^{i,x} x_t) \\
o_t &= \sigma(W^{o,h} h_{t-1} + W^{o,x} x_t) \\
c_t &= (f_t) \odot c_{t-1} + i_t \odot tanh(W^{c,h} h_{t-1} + W^{c,x} x_t) \\
h_t &= o_t \odot tanh(c_t),
\end{aligned}
\tag{3.5}
$$

where $\sigma(\cdot)$ is the sigmoid function. The additional parameters learned in LSTMs include the 8 weight matrices $W^{f,h}, W^{f,x}, W^{i,h}, W^{i,x}, W^{o,h}, W^{o,x}, W^{c,h}, W^{c,x}$.

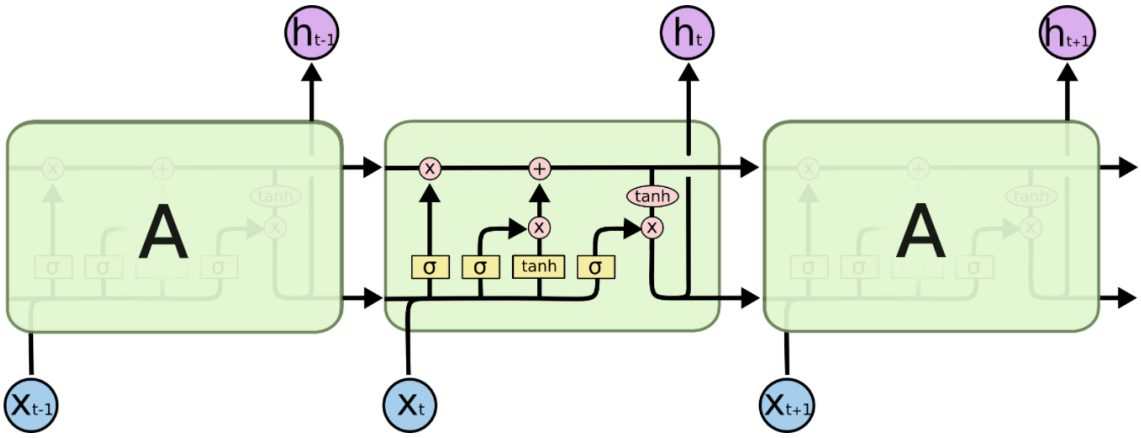

Figure 3-4: The internal structure of an LSTM cell, which features mechanisms to forget [41]. In the above diagram, each line carries an entire vector, from the output of one node to the inputs of others. The pink circles represent pointwise operations, like vector addition, while the yellow boxes are learned neural network layers.

Previous experiments found the basic RNN to be able to predict PJM electricity prices well [60]. Using a similar prediction model architecture, we predict electricity prices within CAISO, and use these values to estimate the annual LACE of solar. Further details of the prediction model will be further elaborated later in the paper at Section 3.3.2.

3.2.2 Solar PV expected output

In chapter 2, we saw how the expected output from solar radiance will form a key input in the LACE computation. This chapter will involve estimating the expected output with the following considerations:

1. Assume that current nodal locations within the CAISO are converted to solar generation. This would allow us to use current location and nodal quantity information as inputs to estimate pv output levels.

2. Based on past 5 year radiance profiles (at current nodal locations), estimate the maximum amount of radiance to be converted to electricity. Part of this calculation will involve assuming a 1-axis tracking array of solar panels pointed at the direction to convert the Global Horizontal Irradiance (GHI) at each location into solar energy.

3. Understand what are current state of the art efficiency of solar panels to determine the energy loss from solar radiance in pv generation.

Solar Capacity Factor

Capacity factor is a measure of how much energy is produced by a plant compared with its maximum output. It is measured as a percentage, generally by dividing

the total energy produced during some period of time by the amount of energy the plant would have produced if it ran at full output during that time [38]. The nature of solar power depends on the availability of solar energy at specific locations. Naturally, different areas would experience different intensity of sunlight throughout the day and no energy production at night. This presents a significant downtime of energy production and can affect the capacity factor of solar as a PV generator [51].

Figures 3-5 below shows the capacity factor for a modeled nodal PV by Brown and O'Sullivan [10] across the different markets from 2010 to 2017. Specifically for CAISO, the modeled nodal PV capacity factor has been relatively stable throughout the 5 years, averaging at about 30%. A stable capacity factor suggest that the general irradiance levels within the CAISO has been consistent, with no significant changes in irradiance patterns. This would become important as we estimate the potential output power at the various nodal location later in the paper.

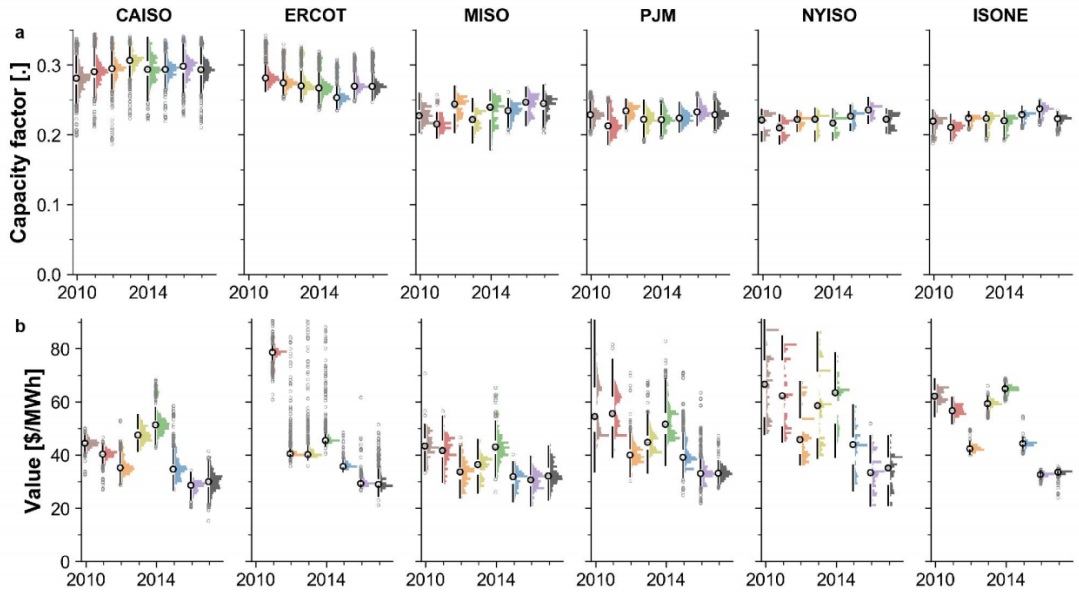

Figure 3-5: Modeled nodal PV capacity factor for a 1-axis tracking array (a) and yearly average energy value (b) sorted by ISO and year [10].

Estimating nodal solar pv output

The output power of a PV generator at the maximum power point (MPP) is obtained from equation 3.6 below [48, 44]:

$$P_{PV} = P_{PV,STC} \times \frac{G_j}{G_{T,STC}} \times [1 - \gamma \times (T_j - 25)] \times N_{PVs} \times N_{PVp}, \tag{3.6}$$

where $P_{PV}, P_{PV,STC}, G_j, \gamma, T_j, N_{PVs}, N_{PVp}$ are the generator output power at the MPP, the rated PV power at the MPP and STC, the irradiance level node j at STC, the power temperature coefcient at MPP, the cell temperature at node j, and the number of modules in series and in parallel that composed the generator respectively.

For our calculations, we will assume Standard Test Condition (STC) measure conditions at $T_{STC} = 25$C, $G_{T,STC} = 1000W/m^2$, and wind speed of 1 m/s. $G_{T,STC}$ refers to the approximate energy density at the Earth's surface for a surface perpendicular to the Sun's rays at sea level on a clear day.

In this thesis, since our focus is on the impact of predicted electricity prices on LACE, we will, for simplicity, estimate the expected Solar PV output with the following assumption: $T_j = 45.5$ for all node locations, $\gamma = 0.043\%/C$, $N_{PVs} = 72, N_{PVp} = 1$ and $PV_{PV,STC} = 165$ as provided by Riffonneau et al. [48]. For the initial analysis, we will be assuming the radiance profile of 2016 at the respective nodal locations.

3.2.3 Data

Electricity price

The California Locational Marginal Pricing (LMP) data at reported nodes were obtained from the CAISO [39] using pyISO, a Python library which facilitates data acquisition from the different ISO application programming interfaces (APIs). Note that the CAISO footprint does not cover the entire state of California and only nodes with serially complete LMP availability for a given calendar year were utilized. The original dataset included 5735 LMP nodes with hourly DA/RT prices between 2009 and 2017. The geographic locations[1] of the nodes were also extracted for reference to weather information at the respective nodal locations. For each location, there was a separate CSV file with information on the date, time (in 5 minute interval), and the LMP price (dollars) for each interval - both Real-Time(RT) and Day-Ahead (DA). We will be focusing on the RT LMP since that will be the eventual price used for LACE computation.

Meteorological data

Meteorological data used are taken from the National Solar Radiation Database Physical Solar Model (NSRDB PSM) [23]. This included information on the global horizontal irradiance (GHI, W/m^2), direct normal irradiance (DNI, W/m^2), diffuse horizontal irradiance (DHI, W/m^2), surface air temperature (°C), and surface wind speed at 2 m height (m/s). For each nodal location, meteorological data are derived from satellite observations and are available on a 4km by 4km grid across the continental United States, at 30 min resolution for historical data from 1998 to 2017. The meteorological data for a given timestamp are assumed to remain constant until the

[1] Approximate locational data were derived from publicly available data on nodes [29]. Exact node location data such as latitude and logitude, considered Critical Energy Infrastructure Information, are confidential and not for distribution, and subject to FERC regulations.

next timestamp; for example, the reported irradiance, wind speed, and temperature between 10:00 and 10:30 are assumed to have the same value throughout. Historical meteorological data at each nodal location was then combined with other nodal information (e.g. LMP) before being used for further calculations.

Data processing

All data processing and analysis was carried out using Python. The data were split into 3 sets: training, validation and test data, which will be used during the implementation of our price prediction model. PyTorch was used for implementation of neural networks models for price prediction. PyTorch is a Python library that handles tensor computation and includes a package for automatic differentiation, which can be run on both CPUs and GPUs. PyTorch makes use of dynamic computational graphs, as opposed to static computation graphs such as those implemented by TensorFlow [45].

In this thesis, much of the development was carried out with testing on CPUs but for large-scale runs a cluster equipped with GPUs was used. The speed-up associated with this was typically about 6 times. This was an extremely significant improvement since in this thesis there was substantial hyper-parameter tuning involved, which required grid searches over large spaces of hyper-parameters. The code used in the following modeling was based on the code done in a previous work on price forecasting in Wong et al. [61].

3.3 LACE Computation

In this thesis, we will compute a basic version of the LACE by adapting equation 2.2 from chapter 2, such that we can easily attain an estimate of the LACE at the various nodal locations quickly by reducing the computation time. Although the actual LACE would involve further considerations (e.g. tax subsidies and capacity payment), these factors are typically consistent throughout the region and thus would not significantly affect the result when we are comparing variation of the LACE estimates between different models of price prediction. The equation used will be as follows:

$$LACE_s = \frac{\sum_{t=1}^{Y}(P_t * D_t)}{\mathbb{E}(G_{hours})} \tag{3.7}$$

Where:

- $LACE_s$ is the levelized avoided cost of electricity, expressed in units of \$/MWh per m^2 of the solar array.
- P_t is the marginal generation price ($\$/kWh$) at time t as define previously.
- D_t is the activity at time t, which is also the expected electricity output generated in kWh.

- t is measured in minute intervals and Y is the number of time steps in a year. In our computations, Y would be approximately add up to 525600 minutes (in a year).
- $\mathbb{E}(G_{hours})$ is the annual expected generation hours in MWh as defined previously.

3.3.1 Base case prediction

We first compute the annual LACE for respective CAISO nodes with the initial assumption that we have perfect information of the marginal prices for electricity in a particular year. In other words, we will assume that we are able to forsee the hourly electricity marginal prices all CAISO nodes in 2016 and use these prices to determine the LACE.
Figure 3-6 shows what the LACE values of nodes within the CAISO using marginal price data from 2016.

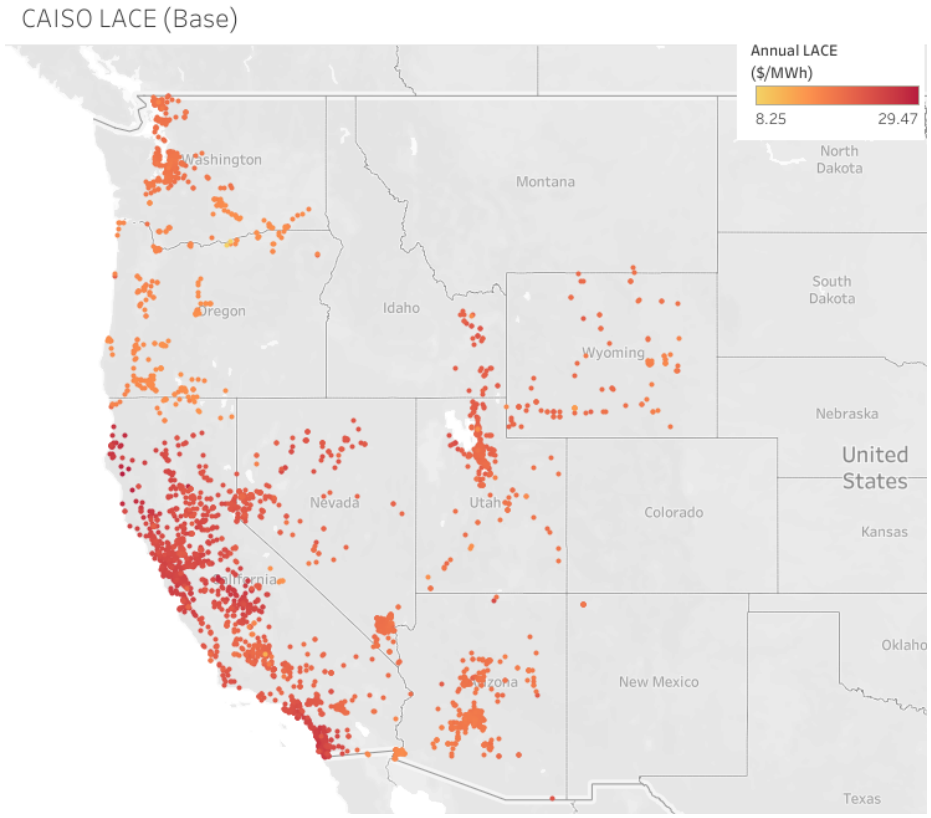

Figure 3-6: Annual LACE for CA ISO nodes assuming 2016 marginal prices. LACE values are calculated for each m^2 of solar panel at current nodal locations.

Here we can see how the coastal areas of California seems to have higher annual LACE per m^2 of solar. As we look to decarbonize electricity within the U.S., replacing the electricity generation at these node locations to solar first will allow us to reap greater benefit. Typically, we will like to assess the potential of solar generating sources at a particular site prior and hence this kind of perfect information on

electricity prices might not always be possible. However, this will serve as a useful benchmark for comparison to the LACE computed with predicted prices from our NN model.

3.3.2 NN prediction

For our proposed prediction model, we will be predicting 2016 lmp prices for all nodes, using historical price data. Since this will be one model to predict prices for all nodes within the CAISO region, the trade off is that the model will be generalized to predict prices for nodes within the CAISO region, without necessarily tuned to individual nodes. We will have more data available for training of the model, this typically improving prediction performance; but should not expect the model to learn nuances of individual nodes. A regular RNN architecture was initially selected for training but we will experiment with the GRU and LSTM architectures as well in the hyperparameter tuning phase.

With the California LMP price data described earlier, the data were transformed into model inputs and expected prediction outputs, where the inputs are formed with a sequence of previous x timesteps of Real-Time (RT) LMP prices before the prediction point. The output y would then be the RT price at the current point in time. The training data were used to perform the initial training of the model while the validation data were used to tune the model hyperparameters. The final model would then utilize the test set to determine the performance of the model and to determine how well it is able to predict electricity prices.

Model performance evaluation metric

In the training phase, backpropagation of the different variations of the RNN model, as well as the optimization of hyperparameters, used the ℓ^1-type loss function (Eqn. (3.1)) discussed above. To determine the performance of the model, the mean absolute scaled error (MASE) was selected as a standard and more relevant measure of time series prediction accuracy. The MASE was calculated using Eqn. (3.8),

$$\text{MASE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|t_i - y_i|}{|t_i - \tilde{t}_i|}, \tag{3.8}$$

where the error of each prediction is scaled by a baseline "naive" prediction and then averaged. The baseline prediction $\tilde{t}_i$ was selected to be the respective node's Day-Ahead (DA) marginal price. The DA price, released 24 hours prior to the RT price, is part of the ISO's the Day-Ahead Energy Market that lets market participants commit to buy or sell wholesale electricity one day before the operating day, to help avoid price volatility [31]. The DA price should in theory encompass information from previous periods as well as market expectations and bidding results of the RT prices making this a reasonable baseline for comparison.

With MASE, values less than one are an improvement over predictions using the mean of the data and zero represents a perfect prediction. This statistic is a transpar-

ent way of understanding a models success at improving predictions over an already available baseline [27]. Another popular performance metric would be the Mean Absolute Percentage Error (MAPE), a measure of prediction accuracy that measures the percentage error from the actual value. In our case, the MASE was selected instead such that we remove any directional bias when interpreting the performance of the prediction error. The MASE further allows for easy interpretation of the model performance compared to a specified baseline.

Hyperparameter tuning

To increase the prediction power of the NN model, we initialized the model with hyperparameters from the PJM model implemented by Wong [60] and performed tuning on selected model hyperparameters. The purpose of hyperparameter tuning is to find the optimal parameters that will provide the model with the lowest loss function. We will then select these parameters as our final model parameters. Hyperparameter tuning was done through a basic grid search algorithm, varying each of the parameters sequentially while fixing the rest of the parameters, with the validation data set to determine the parameter that provided the lowest loss value. In addition, we further experimented with different variations in the model architectures by including LSTM and GRU cells to determine which model architecture would train the best.

The following parameters for model variation were selected for tuning:

1. **Window size:** Number of prior timesteps incorporated to predict target price value
2. **Hidden layers:** Number of hidden layers
3. **Hidden dimension:** Dimension of cell and hidden state of LSTM cell
4. **Batch size:** Number of observation used in a batch
5. **Epoch:** Number of training iteration with input data

For training the model, the learning rate of 0.01 was set, since the model was training well and produced decent results. The initialized hyperparameter values were selected as follows:

1. **Window size:** 26
2. **Hidden layers:** 2
3. **Hidden dimension:** 10
4. **Batch size:** 225
5. **Number of Epochs:** 80

Performance of Varying Window Size

The window size determines how much prior data are fed into the model to predict the target price. In previous work by wong [61], the hourly window size to predict PJM prices was found to be around 26 hours, or roughly one day prior, for better prediction. With this in mind, the model was tuned over a range of window sizes around 24 hours, using previous CAISO RT data.

Figures 3-7, 3-8, 3-9 show the validation loss for the GRU, RNN and LSTM model respectively, as the model was trained over a ranged from 1 to 40 window sizes. In general, all 3 models showed increasing run times as the window sizes increased. This is not surprising since with larger window sizes, the input data dimension increases with the window size and in turn increase the weights dimension of the initial layer that have to be updated during the training of the model. However, this increased run time does not necessarily translate to a lower validation loss. For the GRU, RNN and LSTM, the window size with the lowest validation loss is 36, 8 and 24 respectively. Both the GRU and LSTM model represent a more complex architecture with additional gating characteristics that enable them to model longer sequences. This could explain why the optimal window sizes was larger for the GRU and LSTM model compared to the regular RNN model. Interestingly, we were able to see some patterns of lower validation loss around window sizes of 24 and 36. In the context of the hourly data, this is coincidentally 1 and 2 days respectively, perhaps indicating that information in daily periods might be more valuable to predicting RT prices.

Figure 3-7: Loss over various window size for GRU model.

Figure 3-8: Loss over various window size for RNN model.

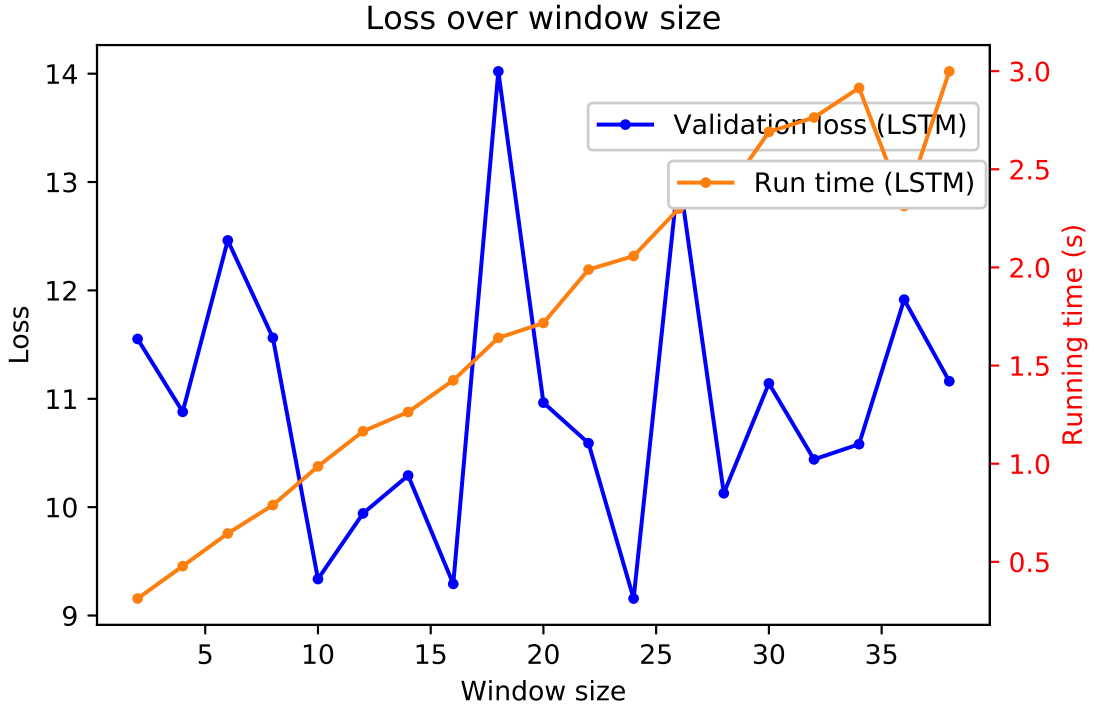

Figure 3-9: Loss over various window size for LSTM model.

Performance of Varying Hidden Layers

The number of hidden layers represent the number of layers between the input and output layers. For the GRU and LSTM model, this indicates the number of GRU and LSTM cells to be included in the model. Overall, a greater number of hidden layers results in a deeper model and would and thus have more parameters to be tuned during training. Looking at figures 3-10, 3-11 and 3-12, this is exactly what we see with longer training times as the number of hidden layers increase.

Comparing the validation loss, the optimal GRU cells seem to be 4 while for the RNN and LSTM models 1 layer was the best. While a deeper model allows for greater potential for the model to adjust its gradients to fit specific data, it runs the risk for the gradients to become unstable and potentially lead to vanishing or exploding gradients and result in poorer results. In this case, lesser number of layers might be better, which is what we observe for the RNN and LSTM models.

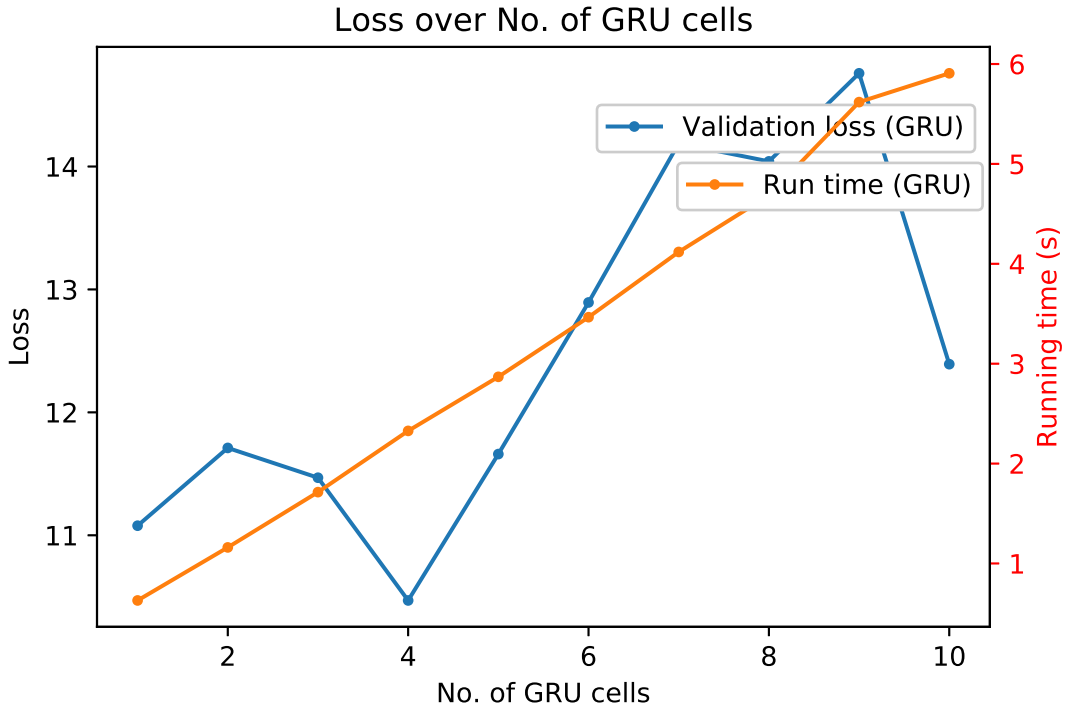

Figure 3-10: Loss over different number of hidden layers for GRU model.

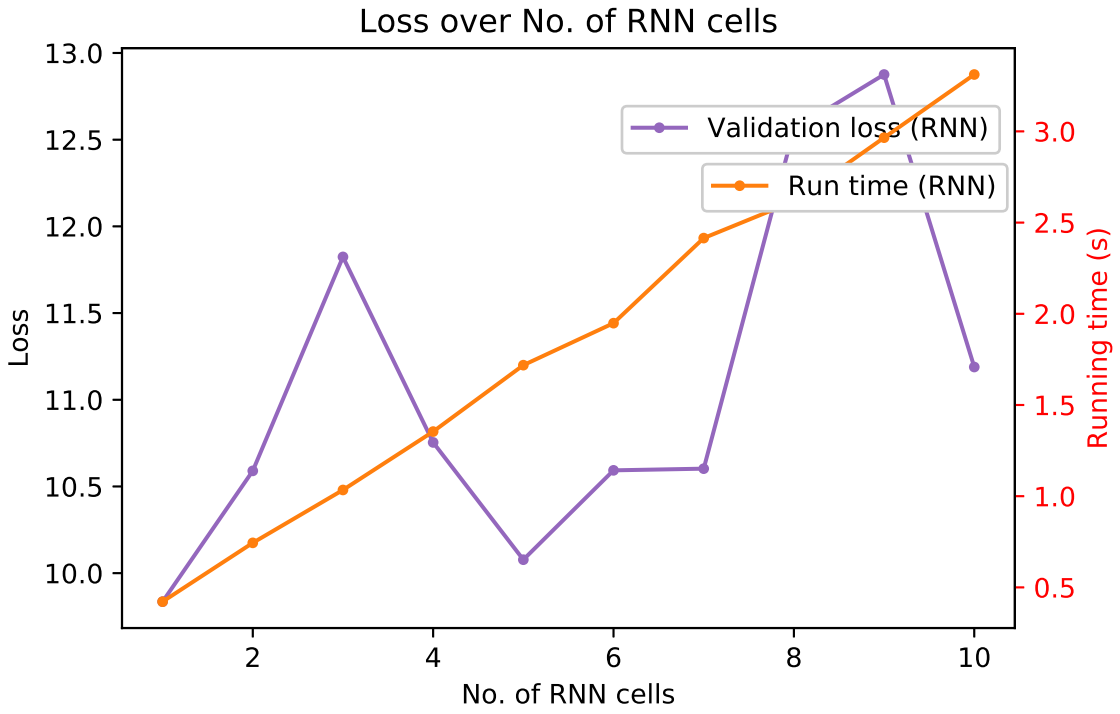

Figure 3-11: Loss over different number of hidden layers for RNN model.

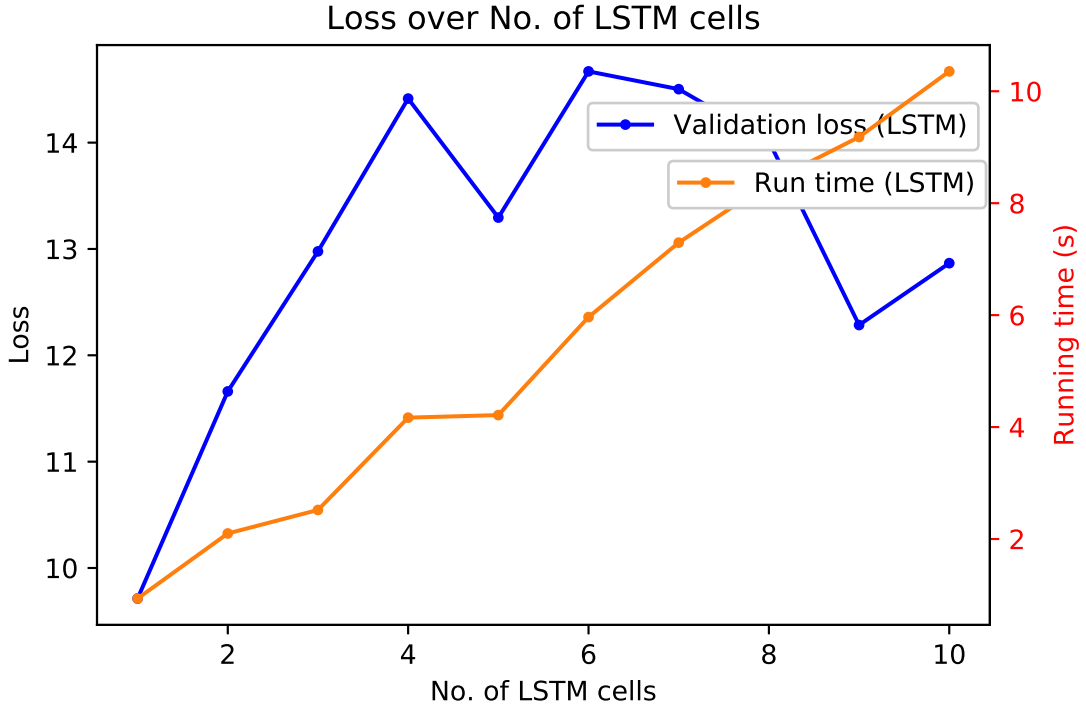

Figure 3-12: Loss over different number of hidden layers for LSTM model.

Performance of Varying Hidden Layer Dimensions

For each hidden layer, the hidden layer dimensions is another parameter that can provide greater complexity for the NN models. Instead of adding complexity in term of the length of the NN, a higher dimension add more depth in each NN layer. Not surprising, when we increase the hidden dimension from 1 to 150, we see that the computation power increases proportionately (as seen from the training time for all 3 models in figures 3-13, 3-14 and 3-15).

Generally, it is recommended [17] that the hidden dimensions should be at least the same as the inputs dimensions so that there is sufficient flexibility for the model to learn. For the GRU, RNN and LSTM model, the hidden dimension that provided the lowset validation loss was 140, 120 and 60 respectively. This was trained with using base parameters of window size of 26, and batch sizes of 225 and 2 hidden layers. This could explain a relatively larger hidden dimension that allowed the model to have sufficient complexity to learn previous price sequences. We further notice that the LSTM model seems to have a much lower optimal dimension. This could be due to the nature of an LSTM cell having greater gating functions, specifically the forget gate, and hence might not utilize the entire sequence of input data at once. With a much lower computation time to train with lower hidden dimension, a lower optimal would be valuable as we train our model on larger data sets.

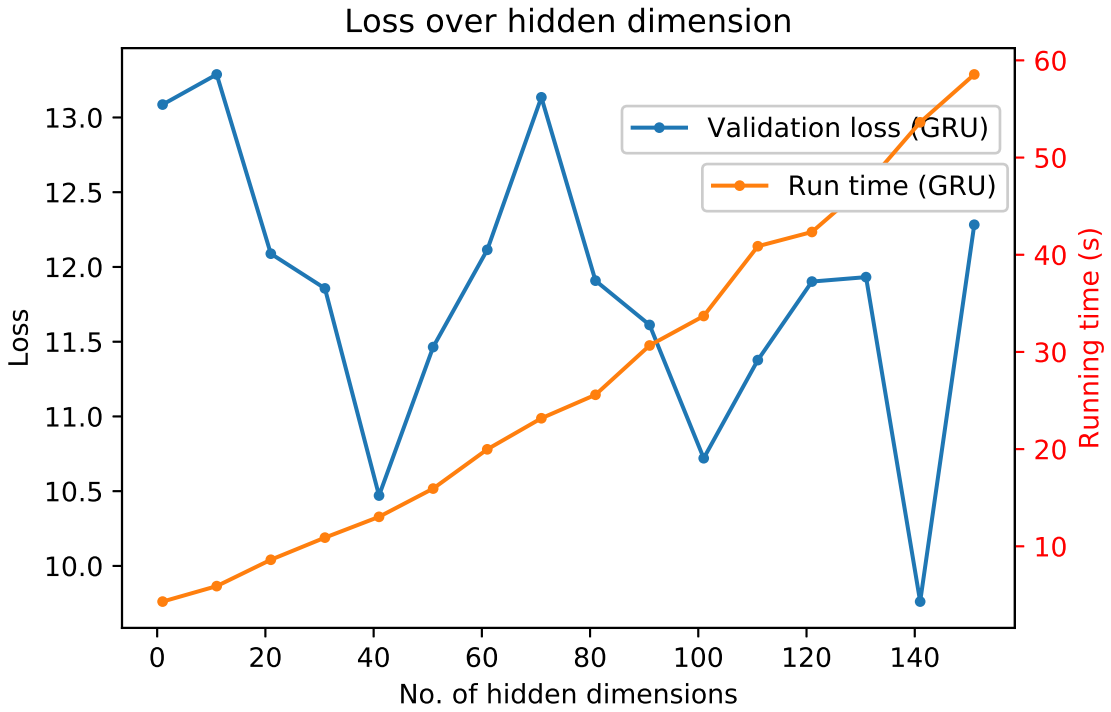

Figure 3-13: Loss over various dimensions of hidden layers for GRU model.

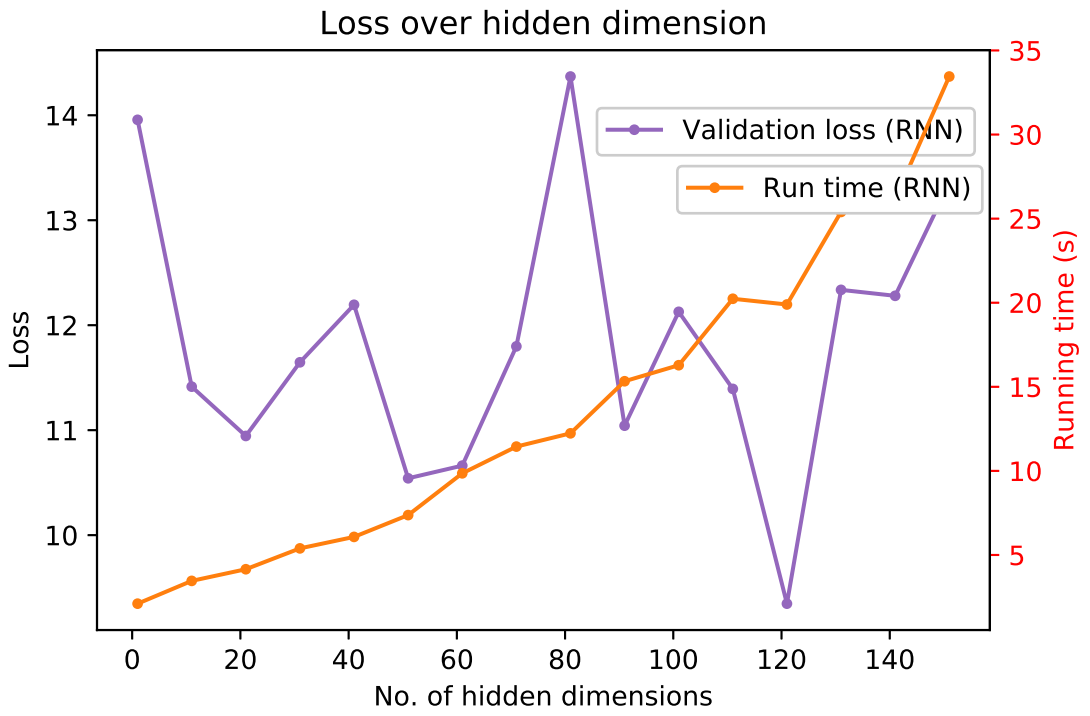

Figure 3-14: Loss over various dimensions of hidden layers for RNN model.

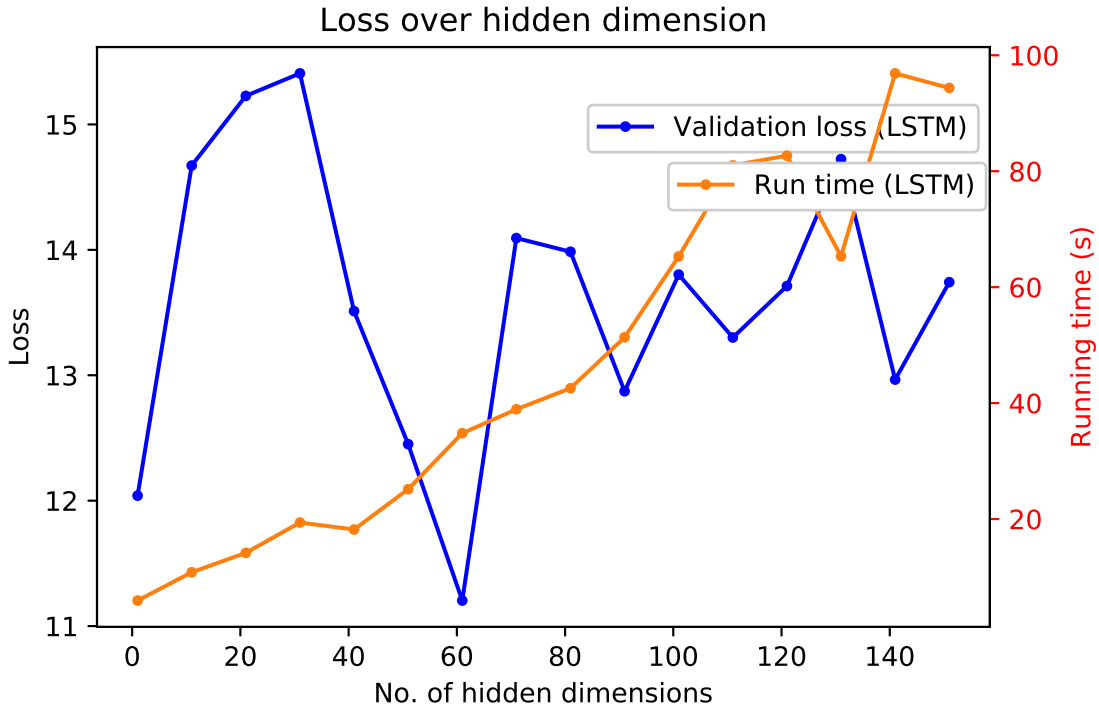

Figure 3-15: Loss over various dimensions of hidden layers for LSTM model.

Performance of Varying Batch Sizes

The batch size determines the number of samples that will be propagated throughout the network. As seen in figures 3-16, 3-17 and 3-18, increasing the batch size from 25 to 500 in intervals significantly decreases the running time of training the model. With a larger batch size, more data sequences are being utilized at once to train. Hence, with the same number of data points, less computation is required.

At the same time, comparing these figures (Fig. 3-16, 3-17 and 3-18), we see huge variation in validation losses as different models learn better with different batch sizes. One hypothesis for this could be that the batch size allows the model to learn a generalized gradient according to the batch input it is provided. If these samples were split into batches with very different characteristics in each batch, during each run the model will find weights that fit each unique sample, making drastic adjustments during each training epoch that might not converge and thus not generalize well across all node prices. Since the weights updated process is different for different NN architectures, the optimal batch sizes would also be unique to the respective model. The batch size with the lowest validation loss for the GRU, RNN and LSTM is 275, 50, and 300 respectively.

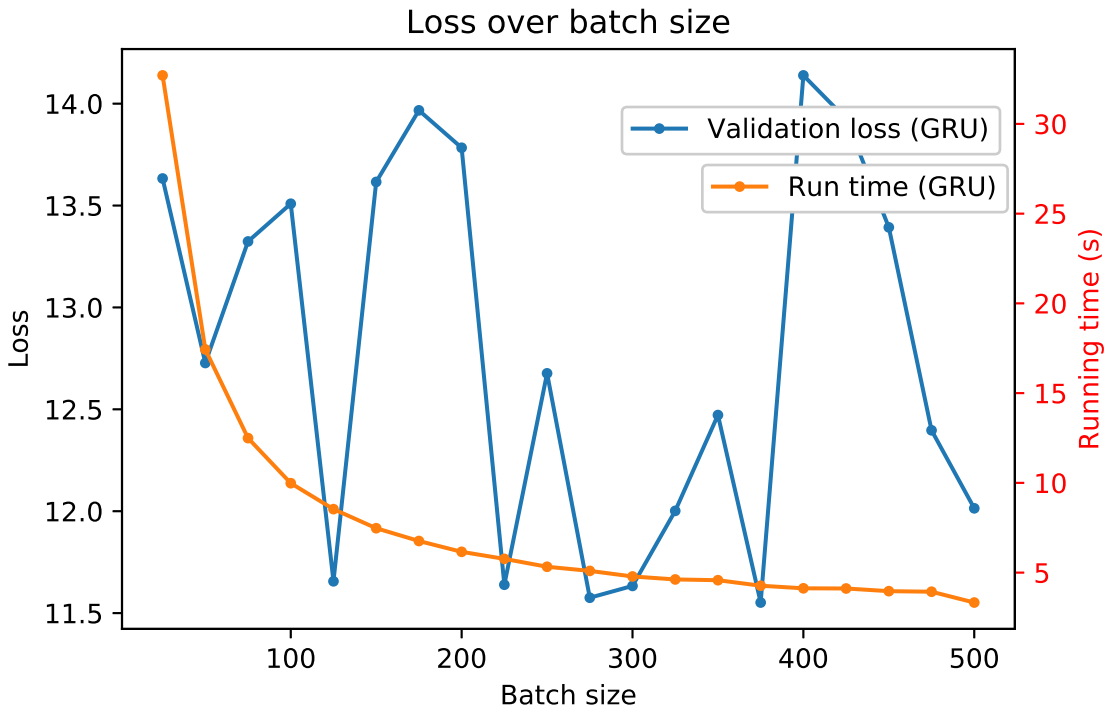

Figure 3-16: Loss over batch sizes for GRU model.

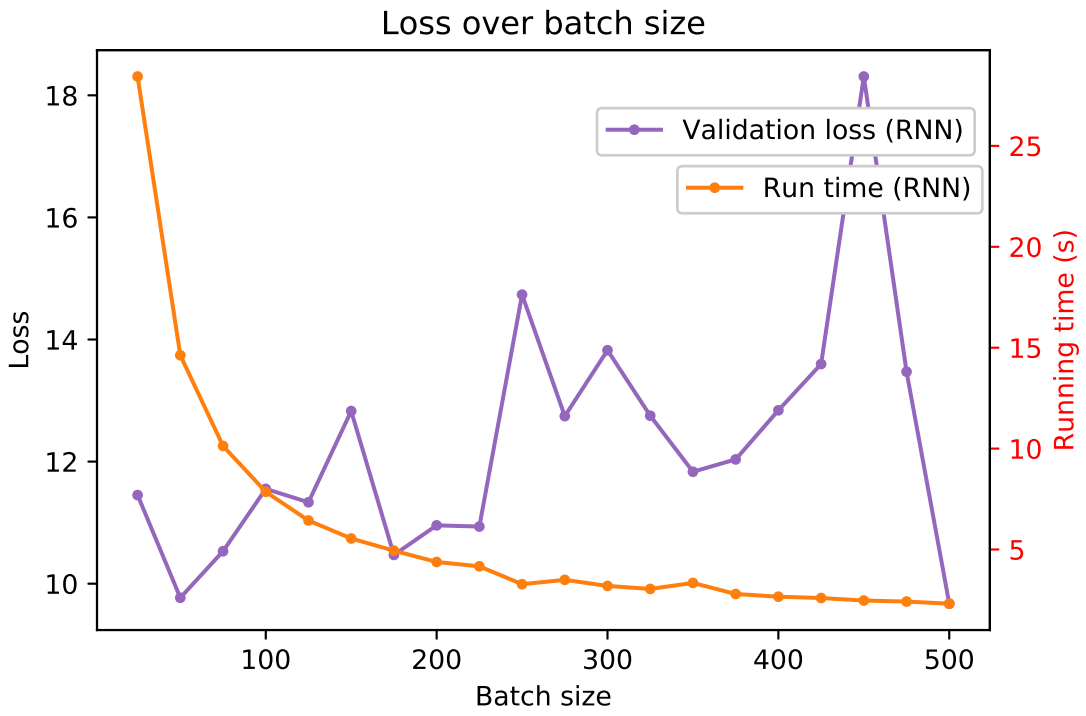

Figure 3-17: Loss over batch sizes for RNN model.

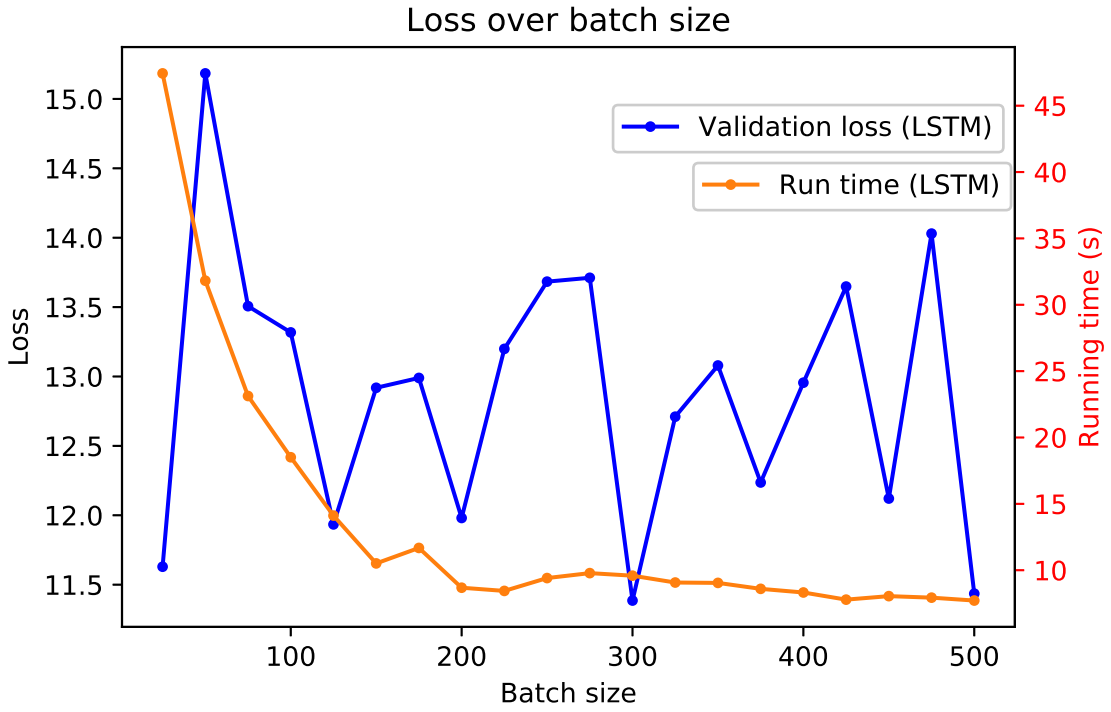

Figure 3-18: Loss over batch sizes for LSTM model.

Early Stopping Criteria

Early stopping can be a useful regularization tool to prevent overfitting [20]. All models were trained over various epoch numbers to attain the training and validation loss. Looking at the results illustrated in figure 3-19, an early stopping criteria was identified for the different models by observing the lowest validation loss. Looking at rate of training loss decreases, the LSTM model seemed to converge at a faster rate. The stopping criteria (epoch number) for the GRU, LSTM and RNN models were 40, 28 and 35 respectively. After which, the validation loss seemed to stabilized, a sign of overfitting of the training set. This result was incorporated into the final model accordingly, with the overall performance discussed later in the report.

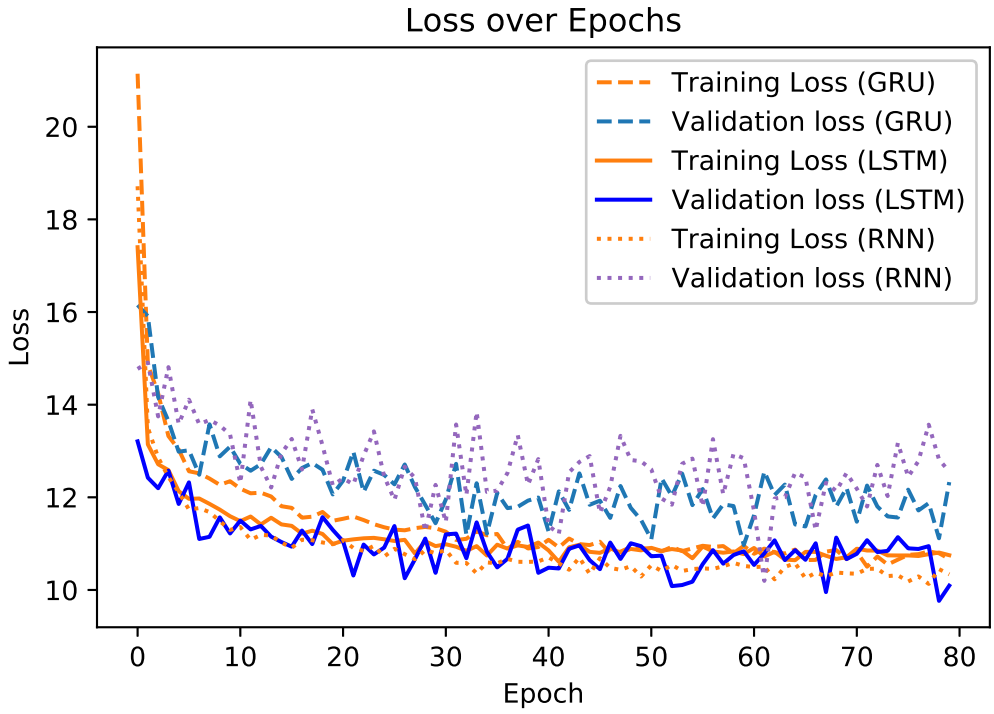

Figure 3-19: Loss over epoch for LSTM, GRU and RNN model

Optimal Parameters

Using the validation loss as our metric for comparison among the various hyperparameter configurations, we see in figure 3-20 that while taking a longer runtime compared to the GRU and RNN model, the LSTM model outperforms them with a significantly lower validation loss. One explanation for this superior performance can be attributed to the unique gating systems of an LSTM cell. When properly trained with sufficient data, the input, forget and output gate of the LSTM cell is able to model longer term sequences, emphasizing on key entries while forgetting less significant data prior to the prediction point.

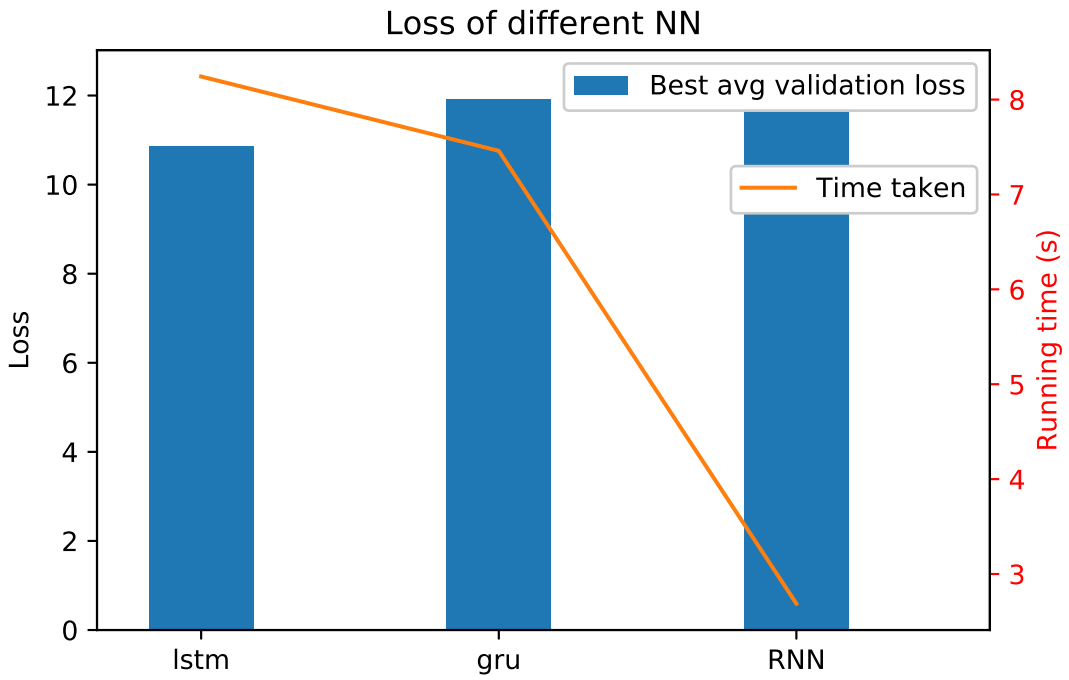

Figure 3-20: Validation Loss comparing the LSTM, GRU and RNN models with the standard hyperparameter setting.

Finally, complemented with other optimal hyperparameter values discussed, the LSTM model is thus the selected model to predict the next hour RT price within the CAISO. The final parameters selected are as follows:

1. **Model:** LSTM
2. **Window size:** 24
3. **Hidden layers:** 1
4. **Hidden dimension:** 60
5. **Batch size:** 300
6. **Learning Rate:** 0.01
7. **Epoch:** 28

Final Model Results

Table 3.1 shows a summary of results using the best experimented parameters with the various prediction models.The predictions were tested using a hold out test data set, using the MAPE and MASE as our metric for comparison. With a MASE of

0.761, the LSTM model outperforms the naive baseline of using the DA price. However, when we look at the MAPE, even the best LSTM model does not seem to perform well. This could be due to the fact that there are relatively frequent and significant spikes in RT prices within the CAISO market, which would skew the towards a larger percentage error when we try to predict the prices.

Technique	LSTM	GRU	RNN
MAPE (%)	25.8	30.16	32.76
MASE	0.761	0.867	0.872
Validation Loss	10.85	11.9	12.39

Table 3.1: Summary of results, showing the best results from the implemented model versus other baseline methods computed on the same data

lmp	Mean	STD	Min	Max
Actual	24.93	57.89	-1162.99	2498.82
Predicted	22.21	10.81	-263.12	505.9

Table 3.2: Summary statistics comparing actual and predicted prices. While the mean values are close, the standard deviation of the actual prices are much greater in magnitude compared to the predicted prices.

Figure 3-21 provides a plot of predicted values (from the final LSTM model) against the actual prices of nodes within CAISO. Generally, we observe that the model produces predicted prices that are able to capture the direction variation of the prices. With hourly fluctuation of prices, being able to follow this fluctuation and in some cases even get accurate predictions is a considerable achievement and will be useful as we use to compute the LACE at the respective nodes. However, for extreme spike in the prices, such as lmp prices that go above $200, the model prediction is pretty conservative and does not capture these prices well. Table 3.2 further highlights this. We notice that while the mean of the actual and predicted lmp (all the nodes) are relatively close in value, the difference in standard deviation of the actual prices are much higher; with much lower minimum prices and much higher maximum prices. Appendix A provides a more granular breakdown of the summary statistics on a nodal and hourly level in tables A.1 and A.2 respectively.

Fortunately, in the context of determining the LACE of solar generation, the impact of not being able to capture these sharp spike in prices remains low. There are two main reasons for this, firstly, in the context of solar generation, if there were errors in price prediction in the night (where there is no expected energy output), this would not affect the LACE computation. Figure 3-22 show an example of a node within the CAISO market. We observe that majority of the price spikes from October 2016 to Janurary 2017 occur at night, mitigating some of the impact of the prediction

error. Secondly, during the calculation of the annual LACE, the expected value is averaged over the year, smoothing out the error over the year and further reducing the effects of the prediction error on the eventual LACE computation. In the later part of the paper, we will compute the LACE based on predicted prices from the LSTM model and compare how different the determined LACE is from the base case.

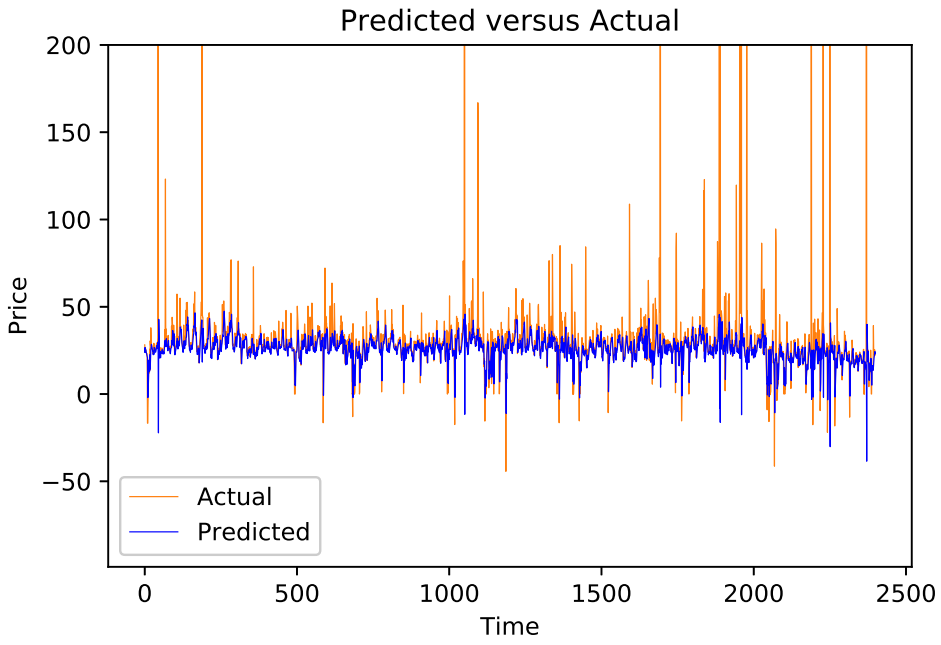

Figure 3-21: Prediction vs Actual plots on test set. This model was trained on past RT information from nodes within CAISO.

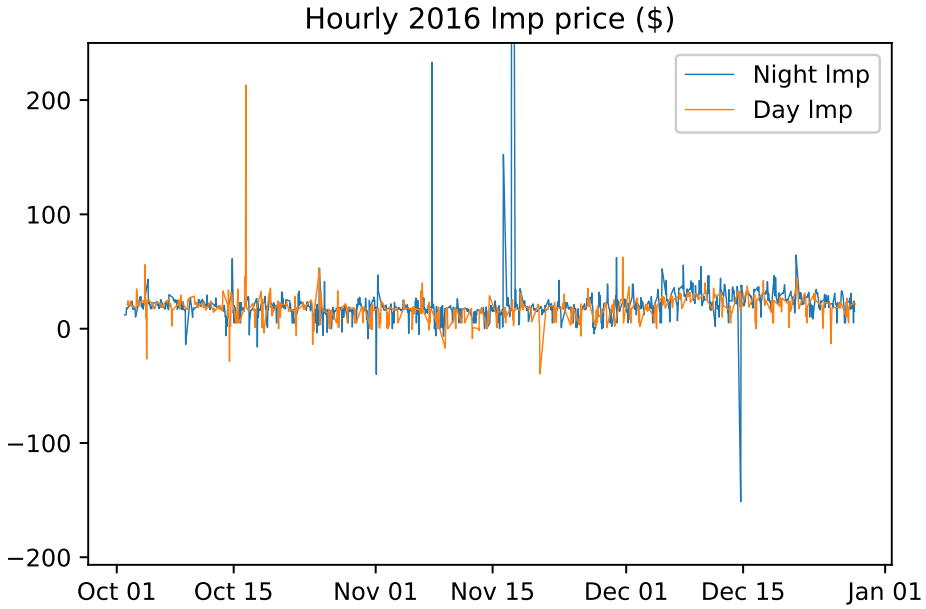

Figure 3-22: Hourly lmp price from one example node in CAISO in 2016. Prices have been differentiated to prices in the day and the night.

Model Extensions

In order to further improve our prediction power of our NN models, there are several extension to the model that can be further explored. From the feature engineering persepctive, we could include greater contextual and relevant information such as the weather or information on the electricity demand in the region, since these would directly affect the prices of electricity. Other features that would be interesting, particularly in the electricity market, would be the fact that nodes within the region are connected via power lines and thus can have a more direct effect on each other. If we could include price information (both RT and DA prices) of neighboring nodes, we should expect significant improvement in prediction performance.

From a modeling perspective, there might also be additional value by deploying an ensemble method by combining the predictions from various model and averaging them. Examples of additional models that can be include might include Auto-Regressive Moving Average (ARMA) type models or Kalman filtering and smoothing procedures, which are popular time-series models that are able to better identify significant price spikes [21, 34].

Finally, we might also consider training individual models to predict prices of each node. This would enable the model to learn any nuances in price behavior of each node and potentially lead to better prediction performance. However, the trade off here is that for each node, we would have much less data available, which is essential in the training of deep NN models.

NN LACE results

Figure 3-23 shows the LACE values of nodes within the CAISO, using the predicted hourly prices from our final LSTM model. At initial glance, the general LACE value distribution throughout the nodes seem to be similar to the base case from figure 3-6, with greater annual LACE of solar per m^2 around the coastal areas of California. This is encouraging from a decision making and policy implementation perspective, since we are able to quickly identify regions where a solar transition would be more valuable.

Comparing the difference in the computed LACE from the LSTM model against the base case, however, we do see some differences in the final computed LACE. Figure 3-24 shows the difference in the LACE of respective nodes in the CAISO, with orange nodes indicating greater LACE values from prediction model and blue nodes reflecting smaller LACE values anticipated by the LSTM model. The darker the shade of the node (both orange or blue), the bigger difference in LACE determine between the base case and NN model.

The difference in LACE values between the base case and the NN model is calculated as shown in equation 3.9 below. Since both cases used the same solar radiance availability assumptions, a negative difference would indicate that the NN model had overall predicted prices that are higher than the base case. Vise versa, a positive difference would indicate that generally, predicted prices were lower compared to the base case. Under prediction would result in a lower valuation of solar at those locations. This might result in policy makers deciding that we should not switch to solar, when it is actually more economical to do so. Conversely, and over prediction of prices might falsely present solar as a valuable switch, which in fact might not be, at least not in those exact node locations. Fortunately, most of the nodes analyzed in this paper do not exhibit drastic LACE values betwen the base case and NN model. In fact, out of 5735 nodes shown in figure 3-23, less than 10% have difference of more than $\$1.5/MWh$ per m^2 of solar array. In the next chapter, we will go in further detail in understanding how the difference in prediction prices from the actual prices, translate to the difference in LACE computation from the base case. At the same time, We will also discuss potential policy effects of having a deviation from the actual LACE.

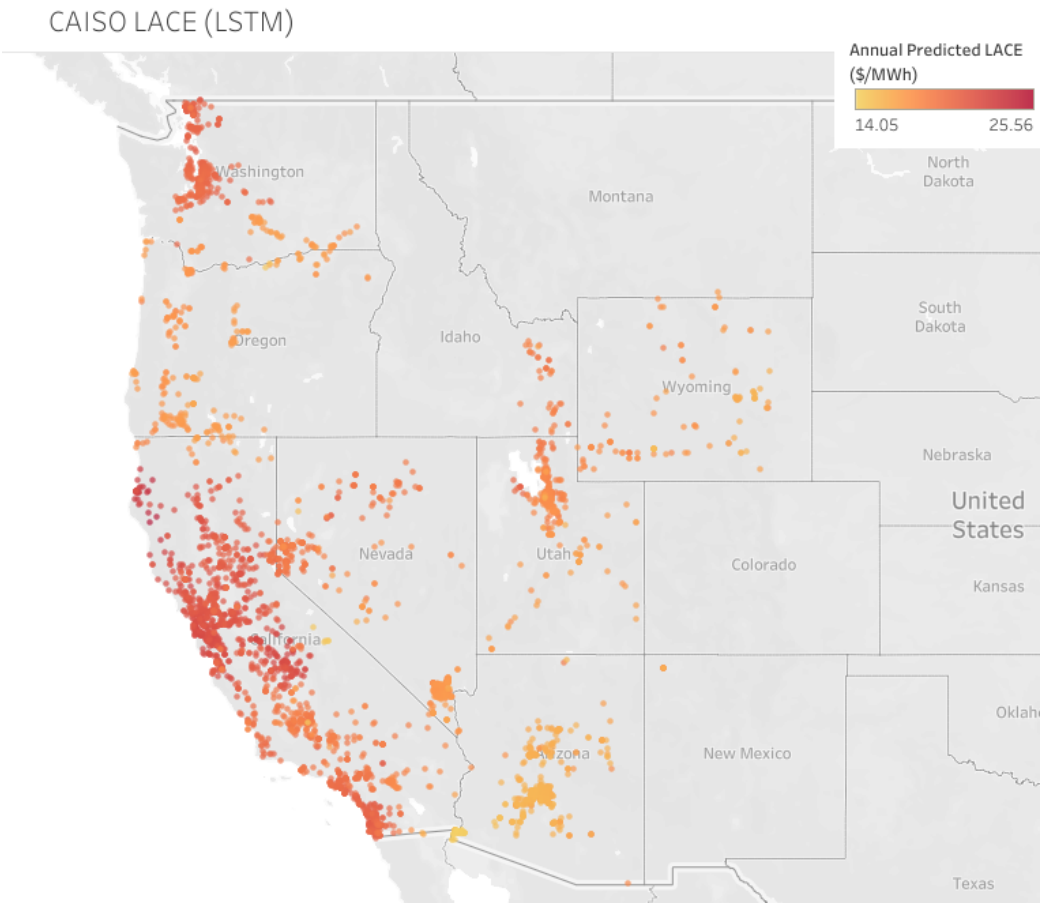

Figure 3-23: Annual LACE for CA ISO nodes using predicted prices for 2016 marginal prices. The predicted prices were determined using LSTM model trained with previous hourly RT data. LACE values are calculated for each m^2 of solar panel at current nodal locations.

$$\text{LACE}_{\text{diff}} = LACE_{base} - LACE_{NN}, \tag{3.9}$$

where $LACE_{base}, LACE_{NN}$ represents LACE computations from the base and NN case respectively, using the LACE formula as described in equation 3.7.

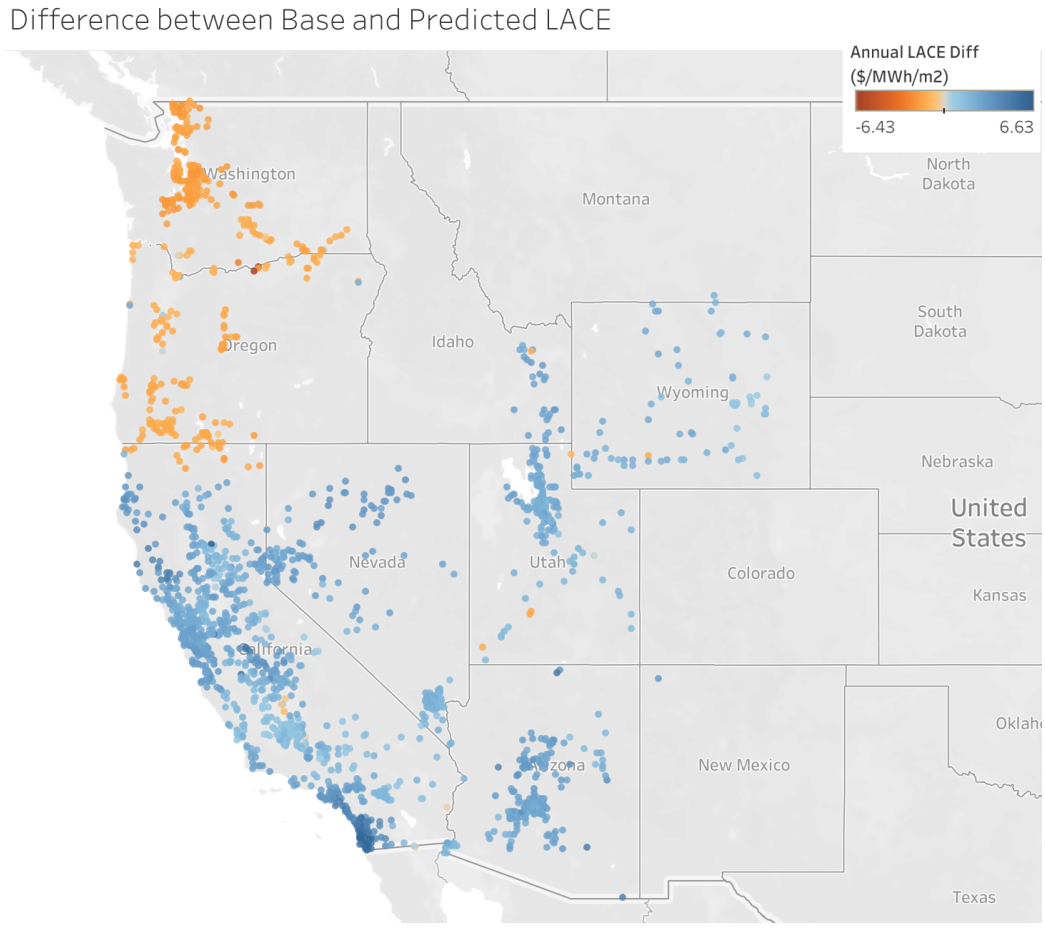

Figure 3-24: Difference between the annual LACE determined from the base case versus the NN model. Darker nodes indicate greater variation in LACE. Negative prices indicate overestimation of LACE using the NN model. Values are calculated for each m^2 of solar panel at current nodal locations.

Chapter 4

LACE Evaluation and Break Even Performance

4.1 Impact of Prediction accuracy on LACE

We have seen how there can be some variation in the evaluation of solar between having perfect price information against prices generated from predictive models. The objective of this chapter is thus to gain a better understanding of the importance of having accurate predictive models, especially in the context of a LACE evaluation of alternative generating electricity sources. We begin by analyzing the prediction errors produced by the model in greater detail, and go in-depth by finding out how this would affect the respective LACE computation as we vary the prediction error. Finally, we will discuss, in relation to the cost of solar PV, how this would affect any breakeven decisions.

4.1.1 Estimating price prediction errors

Referring back to figure 3-21, we observed that for each time step i in the test set, there is a prediction error (e_i) such that $e_i = t_i - y_i$, where t_i and y_i are the actual and predicted prices at time i respectively. e_i is hypothesized to follow a normal distribution with mean 0 and some variance σ^2 and is normalized to a standard normal according to the below equation:

$$z_i = \frac{e_i - \bar{e}}{s}, \tag{4.1}$$

where e_i are the prediction errors at time i, $\bar{e}$ is the mean of the prediction errors and s is the sample standard deviation of the prediction errors. The cumulative distribution (CDF) of z_i is then plotted against a standard Normal CDF to compare the 2 distributions.

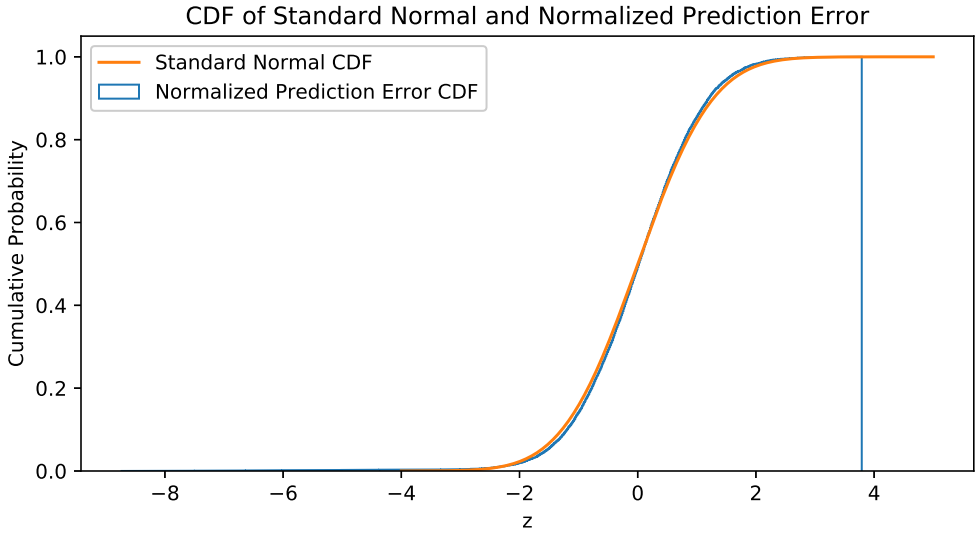

Figure 4-1: Plot of normalized prediction errors and standard normal CDF from the final LSTM model. CDF looks relatively close to standard normal, however, has a much longer left tail.

Figure 4-1 shows a plot of the CDF of a standard normal against the normalized prediction error from our LSTM model. We see that compared to a normal CDF, despite generally looking like a standard normal distribution, the normalized prediction error seemed to have a much longer left tail, going way beyond -4 on the x axis. This effect is more pronounced when we look at the QQ plot of the standardised prediction errors at figure 4-2, where we see a significant drop off before -2. The QQ plot, also known as the Quantile-Quantile Plot, is a graphical technical that shows if 2 sets of data come from populations with a common distribution [37]. In our case, we are comparing the normalized prediction errors against a theoretical standard normal distribution. If the normalized errors did indeed come from the same distribution, we expect the points to fall approximately along the 45-degree reference line. The greater the deviation from this line, the less likely our normalized prediction errors are from a normal distribution. Finally, we performed Kolmogorov-Smirnoff (K-S) test to further validate our results. With n = 2000, the K-S statistic reported was 0.02707 with a p-value of 8.572×10^{-7}. With a low p-value, we can safely reject the null hypothesis that the two distributions are the same, indicating that perhaps the prediction errors do not follow a normal distribution.

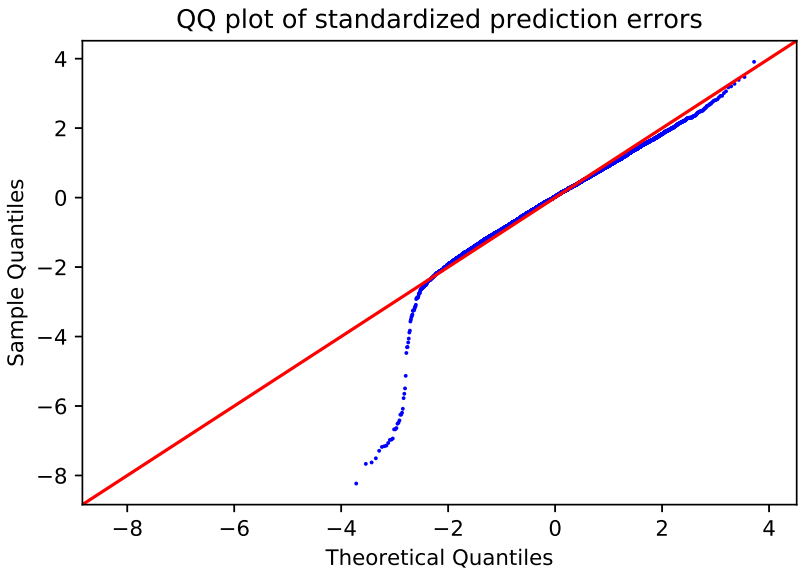

Figure 4-2: QQ plot of the normalized prediction error from the final LSTM model. Points are roughly straight and close to the 45 degree line between -2 and 4. However, points curve off in the extreme parts, some indication that the data have more extreme values than would be expected if they truly came from a Normal distribution.

However, when we look at the actual prices of nodes in 2016, we realise that the long left tail of prediction errors could be explained by huge spike patterns within the CAISO market. Figure 4-3 shows an example of 2016 lmp prices from one node. Here we defined prices about $400 as "spike" prices, which are highlighted in blue. As mentioned previously, the current model is not great at predicted these price spikes and based on our definition of prediction errors, naturally this would result in some extremely large negative errors that, after normalizing, result in the long left tails in the distribution. In order to get a better understanding of the generalized characteristics of prediction errors, we will ignore the "spikes" for now and analyze the distribution of the remaining prediction errors with a similar methodology as performed previously; to see if we can approximate our errors to a normal distribution. We will then assume that all prediction would follow these same characteristics (even for "spiked" prices) to generate prediction errors and study how the LACE changes with increasing errors.

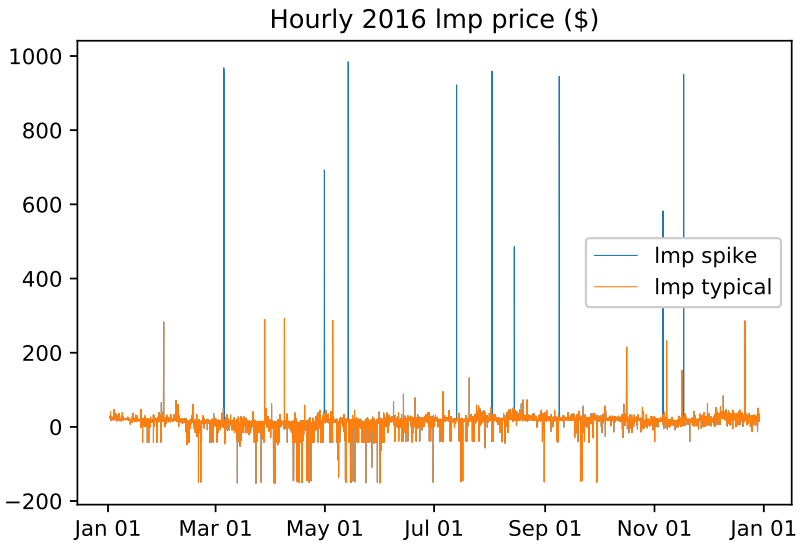

Figure 4-3: Hourly lmp price from another example node in CAISO in 2016. Prices above $400 are considered "spike" prices and are highlighted in blue.

After obtaining the prediction errors of prices below $400 and normalizing the errors using equation 4.1, we generate the QQ plot and the CDF of the standardized prediction errors compared against a standard normal distribution. Figures 4-4 and 4-5 depict the respective QQ plot and CDF graphs.

The points within the QQ plot (fig 4-4) now almost entirely lie along the 45-degree reference line, showing characteristics that the prediction errors are now much closer to that of a standard normal distributions. The CDF plot of the normalized prediction errors against the standard normal further validates this and we see that the distributions are now very similar. Finally, the K-S test statistics is now 0.01048 with a p-value of 0.22156, making it statistically less certain that the 2 distribution are different. It is safe to establish that the prediction errors from the final LSTM model follow a normal distribution with some variance, σ^2.

In the next section, we will assume that all prediction errors for electricity price forecasting, under different prediction techniques, will also follow a standard normal distribution and use this to get a better understanding of how the prediction accuracy would impact the computation of LACE and how it relates to the respective breakeven point.

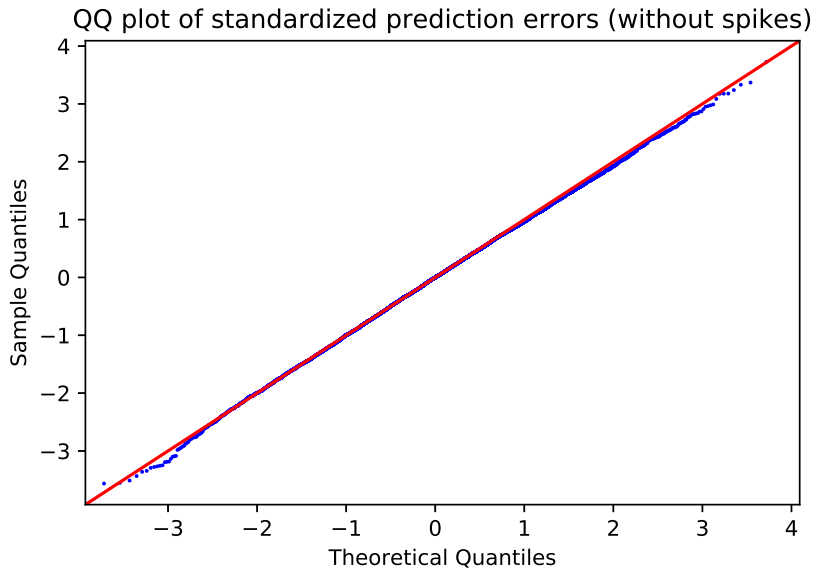

Figure 4-4: QQ plot of the normalized prediction error from the final LSTM model without considering data with significant spikes. Prediction errors reflects much more like standard normal after normalization, although slight curving off still exist at the tails.

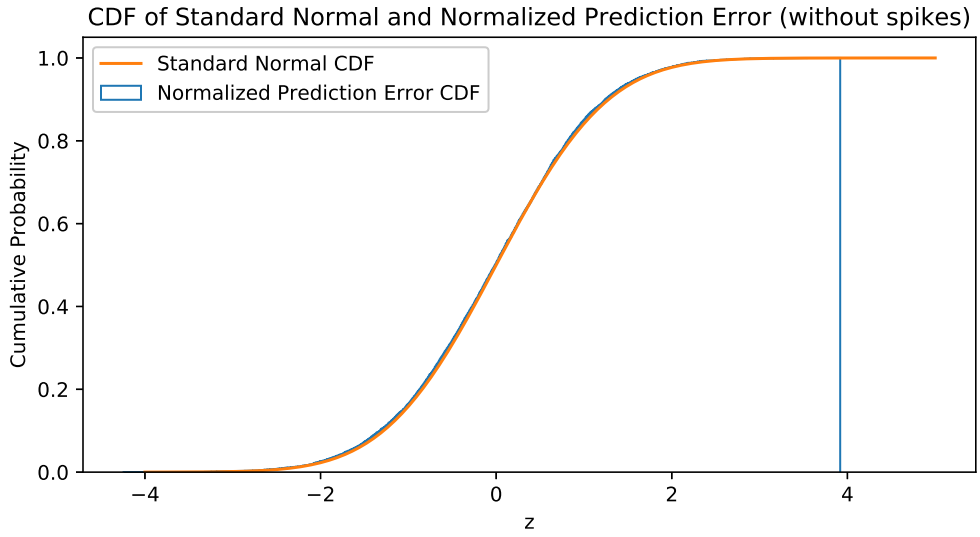

Figure 4-5: Plot of normalized prediction errors (without considering prices with significant spike) and standard normal CDF from the final LSTM model. CDF resembles much like a CDF of a standard normal.

4.1.2 Simulation of Prediction Errors

We randomly sampled nodes from CAISO to study the effects of the LACE as the prediction power varies. The LACE is computed using equation 3.7 with prices that include predicted errors drawn from a Standard Normal distribution with mean 0 and variance σ^2. The σ was varied from a range of 0 to 20 to simulate prediction errors. σ with a value of 0 refers to having no error (perfect knowledge), while $\sigma = 20$ equates to having a bad prediction model with huge variance in its prediction. For each σ, the LACE as well as its respective MAPE was computed 2000 times to simulate the expected LACE and respective error rate from a given prediction model.

Figures 4-6, 4-7 and 4-8 below show examples of selected nodes where the computed LACE for each node was plotted against the MAPE of the prediction. The blue line indicates the expected LACE with perfect knowledge and the light blue area denotes the standard deviation of the LACE computed from the simulation. Notice that while all of them had different annual LACE values, they exhibit similar characteristics as the MAPE increases. Overall, as the MAPE of the prediction increases, the variance of LACE computed increases as well. This is not surprising since we expect that with increasing prediction uncertainty, the LACE would proportionately increase in its variance, since the computation of the LACE is linearly related with the predicted price.

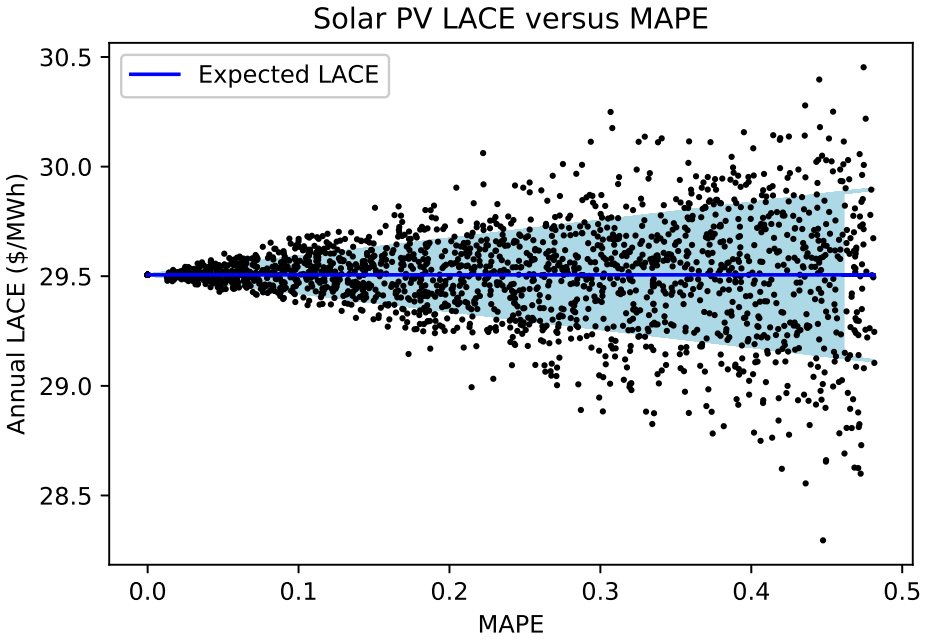

Figure 4-6: Plot of simulated LACE versus Mean Absolute Percentage Error (n = 2000) for an sampled node. The blue line indicates the expected LACE based on 2016 predicted prices, the red line indicates the estimated LCOE [55], and the light blue area denotes the standard deviation of the simulated LACE computation.

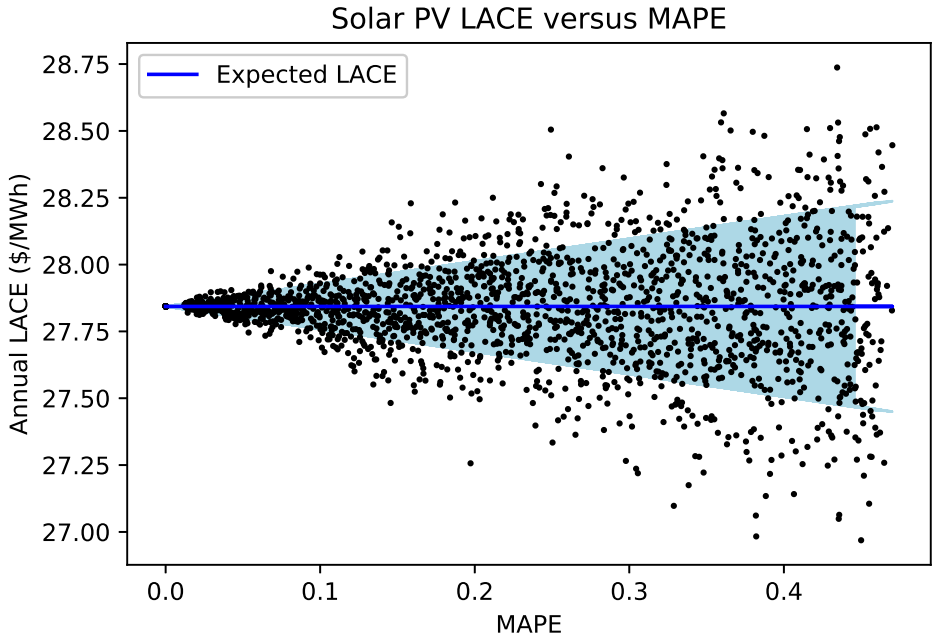

Figure 4-7: Plot of simulated LACE versus Mean Absolute Percentage Error (n = 2000) for another sampled node. The blue line indicates the expected LACE based on 2016 predicted prices, the red line indicates the estimated LCOE [55], and the light blue area denotes the standard deviation of the simulated LACE computation.

However, note that in our computation, we are using the simple LACE calculation of equation 3.7, which in reality might include further benefits from tax policies that might be dependent on the geographic location of the node. This might not always be the best assumption, as we will see later in the thesis. Currently, we also assume that even at the "spiked" prices, the prediction error still follows a normal distribution. In both scenarios, while we expect different the variation to affect each node differently, we can be certain that the variance of the values will continue to increase with increasing prediction errors. With many different alternative sources of energy, the certainty of our evaluation can be the difference in deciding if we switch to solar as our alternative.

4.2 Breakeven Accuracy Performance

4.2.1 Solar PV breakeven LACE

The breakeven point occurs when the the benefits of switching to solar PV is equal to the cost of implementation. So far, we have discussed potential annualised benefits of solar PV within the CAISO region using simplified LACE ($LACE_s$) as our metric. Recall that in chapter 2, we mentioned the LCOE as a extremely popular method of a cost-based evaluation of alternative energy. As such, we will determine the breakeven point as follows:

$$
\begin{aligned}
BE_{solar} - LCOE_{solar} &= 0 \\
BE_{solar} &= LCOE_{solar},
\end{aligned}
\tag{4.2}
$$

where BE_{solar} is the breakeven point for solar, $LCOE_{solar}$ is the annualised solar PV.

According to the Annual Energy Outlook 2020 report by the EIA, the estimated annual LCOE for 2025 would be about $\$27.56/MWh$ [55]. For our example node illustrated in figure 4-8, we see that the expected LACE, when we have 0 MAPE (able to predict prices perfectly), lies above the estimated LCOE (depicted with by the red line), indicating that for this node, we expect to gain greater financial benefits from switching to a solar alternative.

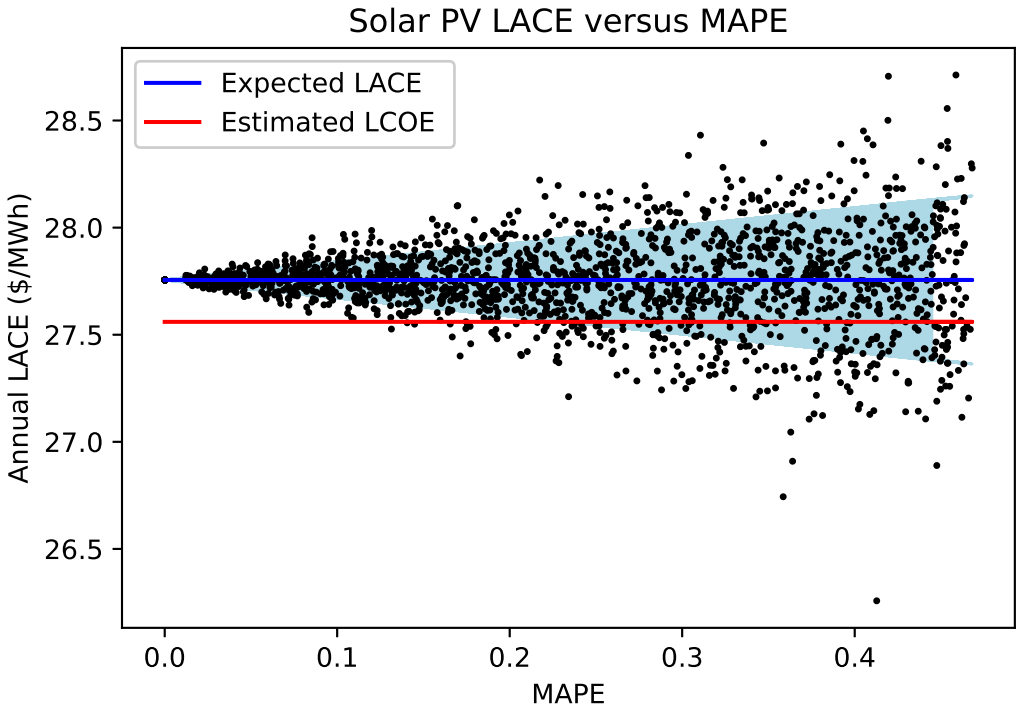

Figure 4-8: Plot of simulated LACE versus Mean Absolute Percentage Error (n = 2000) for the same node in figure 4-3. The blue line indicates the expected LACE based on 2016 predicted prices, the red line indicates the estimated LCOE [55], and the light blue area denotes the standard deviation of the simulated LACE computation.

We compare the $LACE_s$ computed previously with the LCOE from the Annual Outlook 2020 report and found that less than 0.5% of nodes in CAISO, based on 2016 LMP profiles, would break even at todays LCOE for solar PV. This is consistent with values provided by Sullivan and Brown [10] where an estimated of less than 2% of

nodes within U.S electricity markets (excluding ERCOT 2011 data) would break even at the upfront system cost of solar PV, depending on the year.

One of the key assumptions in our computation of the $LACE_s$ is that it does not include any additional capacity payment or capacity credit. If incorporated, we would expected additional economic benefits of switching to solar, further increasing the LACE value. This could provide initial explanation on why we found less nodes above the breakeven point. At the same time, the effects of the assumption is further exacerbated, especially when coupled with prediction uncertainty of electricity prices. When we compare the $LACE_s$ using the predicted prices from our LSTM model, we actually find that none of the nodes met the breakeven point. This would have further repercussions in policies since this seems to indicate that a solar transition would not be worth it, even as cost of solar systems continue to fall. Under such an assumption, a further exploration in the relationship between prediction accuracy and the respective breakeven values (of the respective nodes) can be valuable.

4.2.2 Prediction accuracy and breakeven LACE

When we use the $LACE_s$ computed from the predicted prices from our LSTM model, the maximum value of \$25.56/MWh resulted in none of the nodes passing the breakeven point. Clearly, the accuracy of predicting prices would play a significant role in deciding if it would be a viable switch.

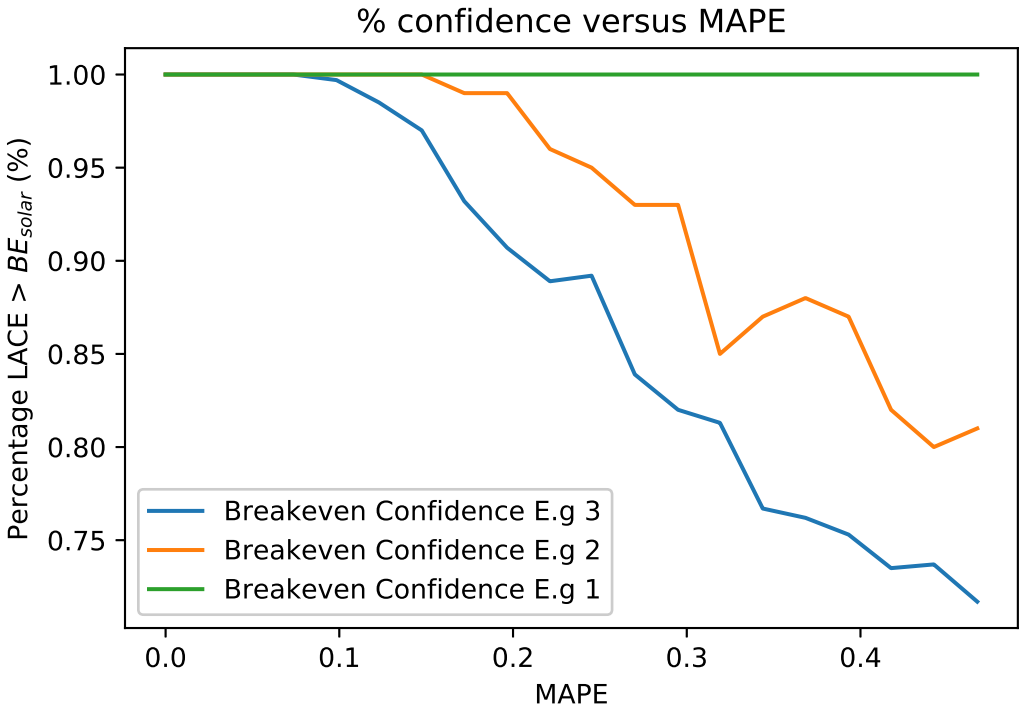

Figure 4-9: Confidence plots from previous simulation examples. E.g 1, 2 and 3 refer to examples from figures 4-6, 4-7 and 4-8 respectively

.

Figure 4-9 depicts a plot of our previous 3 example nodes, showing how confident we are that solar PV is a profitable switch as the performance (in terms of MAPE) of our prediction decreases. Here, we see that the difference in LACE from the breakeven point inversely relates to the significance of prediction accuracy. The closer the LACE to the BE_{solar}, the greater the importance of a more accurate model. In example 1 (figure 4-9), we are confident that of always being above the BE_{solar} point no matter the performance of our model. However, for examples 2 and 3, we see how our confidence levels drop drastically with increasing MAPE. Focusing on the worst case example of node 3, we observe an approximate change of 1% decrease of confidence for every % increase in MAPE.

Currently, our LSTM model provided a MAPE of about 0.25. At this MAPE performance, many of the simulated prediction (from figure 4-8) would fall below the breakeven line. With 2000 simulations, this amounts to about 15% of the computed LACE lie below the BE_{solar}. In this case, a maximum MAPE of 10% can be accommodated if we would like to conclude with certainty, since all simulated LACE values at least greater or equal to the breakeven point.

Chapter 5

Conclusion

In this thesis, we studied the impact of the utilizing machine learning techniques in the evaluation of renewable energy as a source of electricity generation. This chapter aims to summarize key findings of the paper as well as discuss potential further research opportunities as well as policy implications.

5.1 Summary of Key Findings

Focusing on electricity generating nodes within CAISO market, we explored solar PV as an alternative generating source of electricity. Key findings of the paper can be summarized as follows:

1) Currently there are several different methods of evaluating alternative sources of energy. For these different methods, we saw how they can have very different perspective to decision makers and how policies are set. While a cost approach is currently most popular and is easy to determine, it ignores almost completely the benefits potential of alternatives. LACE provides this benefits perspective and a more holistic evaluation can be made when we look at both the LACE and the LCOE together.

2) Electricity prices at different generating nodes provide information of the potential value of generating electricity at that specific nodes. However, these prices are typically volatile, making it extremely uncertain to predict. We discussed various EPF methods that included Multi-agent models, Fundamental models, Reduced-form methods and Computational Intelligence models. Using state of the art computational intelligence models in a Neural Network model, we pursued a machine learning approach to predict prices for 2016, outperforming a naive baseline of using day-ahead prices. For the CAISO price of 2016, an LSTM model was found to be the best model after performing hyperparameter tuning. While having very promising results, there were still a lot of potential for further improvements in prediction performances as well as better understanding of features that contribute to the CAISO electricity prices.

3) With greater investment in prediction performance, we will be able to better evaluate solar PV. Especially when the LACE is close its cost, the investment towards better prediction directly translates into greater certainty in deciding whether we should switch to solar PV. From an economic perspective, this becomes more important to policy and decision makers so that we can achieve both environment and economic goals.

5.2 Further Research and Conclusion

In the end, policy and decision makers are interested in making more informed choices to determine if it is economically viable to switch to a renewable alternative. Throughout the paper, we have been evaluating electricity nodes within CAISO using the $LACE_s$ without potential subsidies and other benefit considerations (e.g. capacity payment or capacity credit). These considerations can be included to determine a more accurate evaluation of the LACE.

From a price prediction perspective, we can conduct further research in the improvement of prediction performance. This can be done through several measures. Firstly, we can improve the training of our NN model through careful feature engineering. This could involve including more exogenous variables such as weather information or geographic information. Next, we could also improve the models selected for training. We mentioned how a potential ensemble method of various prediction models, such as a combination of classical machine learning methods with deep learning techniques, can further improve the prediction power of electricity prices. An improvement in these 2 main areas would already drastically improve the estimation of potential future value of electricity, which is a important input of the LACE evaluation.

We saw how for majority of the nodes, a switch to solar might not seem to be a viable option in the short term, even when we use the base case of perfect knowledge and amid decreasing cost of solar. One possible reason could be that these nodes might be suited for other alternatives (other than solar PV) instead. Although this research focused on solar PV, a similar methodology can be applied on other renewable sources such as wind and hydro to reach a similar decision point. We can then compare between different forms of alternative sources to decide which technology we should switch to.

In conclusion, research showed that while there is going to be variability when attempting to evaluating the potential benefits of solar PV, this variability can be reduced with ongoing improvements in machine learning techniques. Through this machine learning approach in evaluating solar PV, we were able to better understand the relationship between the performance of machine learning prediction, and the value solar PV. Comparison of the LACE and LCOE metric for solar PV will provide easy economical comparison, not just for solar PV but across alternative renewable sources as well. This can improve our decision making in terms of switching away from current electricity generation sources to alternatives as we work towards meeting current climate goals with larger scale implementation of renewable alternatives.

Appendix A

Additional Tables

Node ID	Actual lmp mean	Actual lmp STD	Actual lmp min	Actual lmp max	Predicted lmp mean	Predicted lmp STD	Predicted lmp min	Predicted lmp max
0	28.49	65.23	-295.43	1100.26	24.24	11.81	-56.58	174.42
1	26.45	61.14	-297.84	1078.34	23.12	11.06	-56.98	178.11
2	28.49	65.23	-295.43	1100.26	24.24	11.81	-56.58	174.42
3	28.49	65.23	-295.43	1100.26	24.24	11.81	-56.58	174.42
4	28.49	65.23	-295.43	1100.26	24.24	11.81	-56.58	174.42
5	28.49	65.23	-295.43	1100.26	24.24	11.81	-56.58	174.42
6	28.49	65.23	-295.43	1100.26	24.24	11.81	-56.58	174.42
7	28.49	65.23	-295.43	1100.26	24.24	11.81	-56.58	174.42
8	28.49	65.23	-295.43	1100.26	24.24	11.81	-56.58	174.42
9	28.49	65.23	-295.43	1100.26	24.24	11.81	-56.58	174.42
10	28.55	65.57	-295.43	1096.01	24.25	11.89	-56.58	174.42
11	28.55	65.57	-295.43	1096.01	24.25	11.89	-56.58	174.42
12	28.55	65.57	-295.43	1096.01	24.25	11.89	-56.58	174.42
13	16.44	34.95	-154.29	991.16	17.86	9.29	-49.28	85.75
14	16.42	34.92	-154.07	990.32	17.84	9.28	-49.24	85.64
15	16.53	35.11	-156.63	997.16	17.94	9.33	-49.54	85.82
16	16.53	35.11	-156.63	997.16	17.94	9.33	-49.54	85.82
17	19.68	23.75	-151.37	953.42	20.22	7.42	-29.63	57.8
18	27.83	61.6	-268.99	1080.6	24.59	11.03	-73.29	186.1
19	27.83	61.6	-268.99	1080.6	24.59	11.03	-73.29	186.1
20	21.46	53.58	-237.82	986.94	18.96	9.26	-96.31	171.13
21	21.46	53.58	-237.82	986.94	18.96	9.26	-96.31	171.13
22	26.02	58.95	-639.99	1060.08	23.42	11.25	-135.39	172.78
23	23.61	59.7	-157.9	1005.47	20.52	9.91	-38.2	196.55
24	23.61	59.7	-157.9	1005.47	20.52	9.91	-38.2	196.55
25	28.39	62.05	-157.51	1091.3	24.8	10.95	-40.36	186.48
26	28.39	62.05	-157.51	1091.3	24.8	10.95	-40.36	186.48
27	19.66	23.7	-151.23	950.84	20.2	7.41	-29.6	57.71
28	19.66	23.7	-151.23	950.84	20.2	7.41	-29.6	57.71
29	29.18	68.64	-285.61	1108.74	24.6	12.92	-60.35	170.39
30	29.18	68.64	-285.61	1108.74	24.6	12.92	-60.35	170.39
31	19.78	23.91	-153.25	960.07	20.31	7.49	-30	58.08
32	27.48	63.7	-290.28	1092.47	23.96	11.73	-55.65	172.54
33	27.48	63.7	-290.28	1092.47	23.96	11.73	-55.65	172.54
34	26.5	61.86	-288.44	1161.76	23.18	11.17	-55.25	167.7

Table A.1: Abbreviations (1/2): 2016 lmp stats from example nodes in CAISO. Prediction values reflect statistics from predicted prices of the final LSTM model.

Hour	Actual lmp mean	Actual lmp STD	Actual lmp min	Actual lmp max	Predicted lmp mean	Predicted lmp STD	Predicted lmp min	Predicted lmp max
0	20.34	10.82	-380.93	629.61	22.04	7.9	-263.12	154.85
1	20.32	28.24	-543.53	1051.86	21.50	7.26	-108.68	109.7
2	17.97	12.96	-355.87	533.8	20.19	7.88	-169.74	199.76
3	19.64	27.41	-179.64	569.38	19.47	7.63	-81.84	224.32
4	18.38	17.11	-1019.13	1843.86	19.49	7	-72.83	99.24
5	19.87	10.47	-975.21	246.81	19.60	7.6	-208.31	159.44
6	26.99	28.93	-418.65	563.03	20.78	6.81	-178.09	89.9
7	27.61	57.8	-974.68	1299.43	23.70	7.61	-88.39	97.45
8	21.93	56.96	-998.99	2498.83	23.18	9.78	-177.58	96.46
9	17.08	25.42	-346.88	1196.21	19.93	10.87	-184.29	165.49
10	17.92	27.37	-975.69	1196.08	18.73	11.45	-66.5	282.6
11	24.9	81.71	-1007.31	1556.52	18.83	11.36	-187.37	316.97
12	19.14	49.61	-982.69	1721.39	18.80	12.54	-209.11	505.98
13	27.22	94.64	-1163	1331.18	18.91	12.33	-256.15	152.27
14	24.51	79.37	-987.84	1543.89	19.69	13.17	-217.17	206.27
15	26.41	79.12	-588.66	1695.13	20.07	14.78	-214.36	264.29
16	29.46	92.65	-622.93	1572.7	20.73	13.4	-228.85	265.85
17	25.97	33.17	-789.56	1428.17	22.37	12.4	-171.74	164.75
18	41.88	98.26	-482.51	1411.95	25.010	10.61	-162.27	247.18
19	39.59	102.17	-637.62	1264.72	28.37	11.01	-96.92	277.14
20	32.46	53.39	-611.28	1567.39	28.97	10.28	-172.26	243.07
21	34.19	68.72	-763.67	1208.5	29.03	8.98	-112.17	203.94
22	24.27	14.19	-750.09	438.16	28.48	9.31	-203.88	179.33
23	20.52	14.49	-982.23	760.03	25.12	8.01	-234.33	145.01

Table A.2: Abbreviations (2/2): 2016 Hourly lmp stats from all nodes in CAISO nodes. Prediction values reflect statistics from predicted prices of the final LSTM model

Bibliography

[1] Electricity in the United States - Energy Explained, Your Guide To Understanding Energy - Energy Information Administration. `https://www.eia.gov/energyexplained/index.cfm?page=electricity_in_the_united_states#tab2`. Online; accessed 2018-03-17.

[2] Claudio Albanese, Harry Lo, and Stathis Tompaidis. A numerical algorithm for pricing electricity derivatives for jump-diffusion processes based on continuous time lattices. *European Journal of Operational Research*, 222(2):361–368, 2012.

[3] J. Aldersey-Williams and T. Rubert. Levelised cost of energy a theoretical justification and critical assessment. 124:169–179.

[4] N. Amjady and M. Hemmati. Energy price forecasting - problems and proposals for such predictions. 4(2):20–29.

[5] M. T. Barlow. A Diffusion Model For Electricity Prices. *Mathematical Finance*, 12(4):287–298, October 2002.

[6] C. Baskette, B. Horii, E. Kollman, and S. Price. Avoided cost estimation and post-reform funding allocation for california's energy efficiency programs. 31(6):1084–1099.

[7] Y. Bengio, P. Simard, and P. Frasconi. Learning long-term dependencies with gradient descent is difficult. *Trans. Neur. Netw.*, 5(2):157–166, March 1994.

[8] Fred Espen Benth, Jrat altyt Benth, and Steen Koekebakker. *Stochastic Modeling of Electricity and Related Markets*. Number 6811 in World Scientific Books. World Scientific Publishing Co. Pte. Ltd., January 2008.

[9] Marie Bessec and Othman Bouabdallah. What causes the forecasting failure of Markov-Switching models? A Monte Carlo study. Econometrics 0503018, University Library of Munich, Germany, March 2005.

[10] Patrick R. Brown and Francis M. O'Sullivan. Spatial and temporal variation in the value of solar power across united states electricity markets. 121:109594.

[11] Maira Bruck, Peter Sandborn, and Navid Goudarzi. A levelized cost of energy (LCOE) model for wind farms that include power purchase agreements (PPAs). 122:131–139.

[12] Chris Namovicz. Assessing the economic value of new utility-scale renewable generation projects.

[13] Christoph Clauser and Markus Ewert. The renewables cost challenge: Levelized cost of geothermal electric energy compared to other sources of primary energy review and case study. 82:3683–3693.

[14] G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4):303–314, Dec 1989.

[15] W El-Khattam and M. M. A Salama. Distributed generation technologies, definitions and benefits. 71(2):119–128.

[16] A. Eydeland and Krzysztof Wolyniec. Energy and power risk management. 01 2003.

[17] Farhad Malik. What are hidden layers. `https://medium.com/fintechexplained/what-are-hidden-layers-4f54f7328263`. Online.

[18] Frank A. Felder and Ruthanne Haut. Balancing alternatives and avoiding false dichotomies to make informed u.s. electricity policy. 41(2):165–180.

[19] Helyette Geman and A. Roncoroni. Understanding the Fine Structure of Electricity Prices. Post-Print halshs-00144198, HAL, April 2006.

[20] Federico Girosi, Michael Jones, and Tomaso Poggio. Regularization Theory and Neural Networks Architectures. *The MIT press*, 7(2):219–269, 1995.

[21] Dana E. Goin and Jennifer Ahern. Identification of spikes in time series. 8(1).

[22] Vikas Gupta. Understanding feedforward neural networks, 2017.

[23] Aron Habte, Manajit Sengupta, and Anthony Lopez. Evaluation of the national solar radiation database (NSRDB): 1998-2015.

[24] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[25] Anthony John Hughes, Havelock Brewster, and Commonwealth Secretariat. *Lowering the Threshold: Reducing the Cost and Risk of Private Direct Investment in Least Developed, Small and Vulnerable Economies*. Commonwealth Secretariat. Google-Books-ID: 1dkf4hlPe0cC.

[26] Rob J. Hyndman and George Athanasopoulos. Forecasting: principles and practice. `https://www.otexts.org/fpp`. Online; accessed 2018-04-07.

[27] Rob J. Hyndman and Anne B. Koehler. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22:679 – 688, 2006.

[28] Andres Inzunza and Christopher R. Knittel. The impacts of renewable energy on the economic value of energy storage in california. 2019.

[29] ISO California. Market price maps. http://www.caiso.com/PriceMap/Pages/default.aspx. Online.

[30] ISO New England. Capacity vs. Energy: A Primer. https://www.iso-ne.com/about/what-we-do/in-depth/capacity-vs-energy-primer. Online.

[31] ISO New England. Day-Ahead and Real-Time Energy Markets. https://www.iso-ne.com/markets-operations/markets/da-rt-energy-markets/. Online.

[32] Joanna Janczura and Rafal Weron. An empirical comparison of alternate regime-switching models or electricity spot prices. MPRA Paper 20546, University Library of Munich, Germany, February 2010.

[33] Chang-il Kim, In-Keun Yu, and Y.H Song. Prediction of system marginal price of electricity using wavelet transform analysis. *Energy Conversion and Management*, 43:1839–1851, 09 2002.

[34] Christopher R. Knittel and Michael R. Roberts. An empirical examination of restructured electricity prices. 27(5):791–817.

[35] Siem Jan Koopman, Marius Ooms, and M. A. Carnero. Periodic seasonal reg-arfima-garch models for daily electricity spot prices. *Journal of the American Statistical Association*, 102:16–27, 02 2007.

[36] Lazard. Lazard's Levelized Cost of Energy Analysis - Version 11.0, 2017.

[37] NIST SEMATECH. Quantile-Quantile Plot. https://www.itl.nist.gov/div898/handbook/eda/section3/qqplot.htm. Online.

[38] NREL. Solar Energy and Capacity Value. https://www.nrel.gov/docs/fy13osti/57582.pdf. Online.

[39] OASIS. California ISO Open Access Same-time Information System (OASIS). http://oasis.caiso.com/mrioasis/logon.do. Online.

[40] PennState College of Earth and Mineral Sciences. 3.1.1 fixed cost concepts for power generation | EBF 483: Introduction to electricity markets.

[41] Christopher Olah. Understanding LSTMs. http://colah.github.io/posts/2015-08-Understanding-LSTMs/. Online; accessed 2019-04-14.

[42] PennState Department of Energy and Mineral Engineering. 9.1. base load energy sustainability | EME 807: Technologies for sustainability systems.

[43] William Pentland. Levelized cost of electricity: Renewable energy's ticking time bomb?

[44] PowerScout. How to Calculate Solar Panel Output. https://powerscout.com/site/how-to-calculate-solar-panel-output. Online.

[45] PyTorch. PyTorch. https://github.com/pytorch/pytorch.github.io, 2017.

[46] Renewable Energy World. How to Compare Power Generation Choices. https://www.renewableenergyworld.com/2009/10/29/how-to-compare-power-generation-choices/. Online.

[47] Joshua D. Rhodes, Carey King, Grcan Gulen, Sheila M. Olmstead, James S. Dyer, Robert E. Hebner, Fred C. Beach, Thomas F. Edgar, and Michael E. Webber. A geographically resolved method to estimate levelized power plant costs with environmental externalities. 102:491–499.

[48] Yann Riffonneau, Bacha Seddik, Franck Barruel, and Stphane Ploix. Optimal power flow management for grid connected pv systems with batteries. *Sustainable Energy, IEEE Transactions on*, 2:309 – 320, 08 2011.

[49] B. Saunders, L. Zarker, and S. Parker. Renewable-energy resource planner: a guide for evaluating and financing industrial renewable energy measures.

[50] R. Schmalensee, V. Bulovic, R. Armstrong, C. Batlle, P. Brown, J. Deutch, H. Jacoby, R. Jaffe, J. Jean, R. Miller, F. O'Sullivan, J. Parsons, J.I. Prez-Arriaga, N. Seifkar, R. Stoner, and C. Vergara. The future of solar energy: An interdisciplinary MIT study, 2015.

[51] Sunmetrix. What is capacity factor and how do solar and wind energy compare? https://sunmetrix.com/what-is-capacity-factor-and-how-does-solar-energy-compare/. Online.

[52] Ronald J. Sutherland. Market barriers to energy-efficiency investments. 12(3):15–34.

[53] U.S. Energy Information Administration. EIA Annual Outlook 2020. https://www.eia.gov/outlooks/aeo/electricity_generation.php. Online.

[54] U.S. Energy Information Administration. Levelized cost of electricity and levelized avoided cost of electricity methodology supplement. page 4, 2013.

[55] U.S. Energy Information Administration. Levelized cost and levelized avoided cost of new generation resources in the annual energy outlook 2019. page 20, 2019.

[56] U.S. Energy Information Administration. U.S. energy consumption by energy source, 2018. 2019.

[57] Karakatsani Nektaria V and Bunn Derek W. Fundamental and Behavioural Drivers of Electricity Price Volatility. *Studies in Nonlinear Dynamics & Econometrics*, 14(4):1–42, September 2010.

[58] Laura Malaguzzi Valeri. INSIDER: Not all electricity is equaluses and misuses of levelized cost of electricity (LCOE).

[59] Rafa Weron. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, 30(4):1030 – 1081, 2014.

[60] S. H. Wong. Valuing energy storage in electricity grids: A machine learning approach, 2017.

[61] S. H. Wong, B. Ng, and J. B. Montgomery. EPF using recurrent neural methods, 2017.

[62] Irene M Xiarchos. Factors affecting the adoption of wind and solar-power generating systems on u.s. farms: Experiences at the state level. page 45.

[63] Tarn Yates and Bradley Hibberd. Levelized cost of energy. 2012.

MIT CEEPR

MIT Center for Energy and Environmental Policy Research

Since 1977, the Center for Energy and Environmental Policy Research (CEEPR) has been a focal point for research on energy and environmental policy at MIT. CEEPR promotes rigorous, objective research for improved decision making in government and the private sector, and secures the relevance of its work through close cooperation with industry partners from around the globe. Drawing on the unparalleled resources available at MIT, affiliated faculty and research staff as well as international research associates contribute to the empirical study of a wide range of policy issues related to energy supply, energy demand, and the environment.

An important dissemination channel for these research efforts is the MIT CEEPR Working Paper series. CEEPR releases Working Papers written by researchers from MIT and other academic institutions in order to enable timely consideration and reaction to energy and environmental policy research, but does not conduct a selection process or peer review prior to posting. CEEPR's posting of a Working Paper, therefore, does not constitute an endorsement of the accuracy or merit of the Working Paper. If you have questions about a particular Working Paper, please contact the authors or their home institutions.

MIT Center for Energy and Environmental Policy Research
77 Massachusetts Avenue, E19-411
Cambridge, MA 02139
USA

Website: ceepr.mit.edu

MIT CEEPR Working Paper Series is published by the MIT Center for Energy and Environmental Policy Research from submissions by affiliated researchers.

For inquiries and/or for permission to reproduce material in this working paper, please contact:

Email ceepr@mit.edu
Phone (617) 253-3551
Fax (617) 253-9845