IMPACT OF DISTRIBUTED PV GENERATION ON GRID OPERATION

A DISSERTATION
SUBMITTED TO THE INSTITUTE FOR COMPUTATIONAL
AND MATHEMATICAL ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Umnouy Ponsukcharoen
January 2017

Abstract

Distributed photovoltaic (PV) generation is growing rapidly due to renewable-energy policies, business innovation, and new technology. The variability and volatility of solar generation causes challenges when operating the grid. Given the current physical infrastructure and operational policies this results in limits on distributed PV generation. What the limit is to the penetration of PV is studied by analyzing the behavior of distributed PV generation and a model of the load.

We investigate how to represent the power output from distributed PV generation. We propose a reduced-form approach for constructing maximum PV power curves, which describe the maximum power from a group of PV systems at any time instance. Unlike a structural approach used typically by PV providers, the reduced-form approach requires no information about configuration or specification of each PV system in the group. We show how a grid operator can use maximum power curves in various applications such as finding a simple limit to PV installation, defining volatility distinctly from variability, and constructing a short-term PV power forecast system.

A key requirement is a model of the load. Analogous to the maximum power curve, a load reference curve is proposed with a simple definition and shown to improve the performance of a short-term load forecast. Lastly, a simulation of the grid operation using our forecast systems and a typical set of existing infrastructures is performed to measure the risk of demand-supply imbalance. The maximum level of distributed PV installation given an acceptable level of risk is established. Options of adding additional ancillary services and implementing PV curtailment to increase the maximum level of distributed PV installation are also investigated.

Preface

The grid

Electricity has been an essential commodity to deliver energy driving a modern civilization for more than a century. It is the backbone of all domestic, commercial, and industrial activities. Electricity transmission is not a simple task. Energy in the form of electricity is expensive to store, and the precise rate of the electric energy flow called **power** is required to operate electric appliances. Electricity transmission is not simply sending a portion of energy from a supplier to a consumer at an arbitrary instance like most other resources. The electricity transmission must match demand and supply at all instances. That is, at all times, the generating power from all generation units must equal to the consuming power, called **load**, from all consumers in the transmission network called the electric grid, or for short **the grid**. The grid can comprise a significant number of traditional power plants known as **generators** that supply a majority of energy, a large number of power consuming units known as loads from several sectors in the economy, and a small number of fast reacting reserves known as **ancillary services** that can match the power difference between generators and loads. The procedure to match demand and supply at any instance starts with **the day-ahead energy scheduling** where the load profile of the operating day is predicted and generators are scheduled to match their power with the predicted load profile. On the operating day, there is an **hour-ahead energy scheduling**, in which the load profile of the operating hour is predicted and additional generators are scheduled so that the combined power matches the updated predicted profile. At some time before the operating instance, the scheduling is no longer applied. In a nutshell, the collection of responsive generators and loads known as ancillary service is called to match the load at the operating instance. [1] Furthermore, the operation is complex due to the fact that these entities are distributed within a web of transmission lines. The problem is not just about matching the net demand and supply in the grid, but

[1] In fact, the prediction and scheduling in the sub-hour level are still running but limited to ancillary services. More detail of the grid operation procedure is described later in Chapter 9.

also about satisfying physical constraints of transmission. The mathematical formulation of balancing the grid under these constraints is called the power flow problem. An additional complication is that, even though the grid is large and interconnected, the grid is a dynamical system and sensitive to disturbances. It means that after a sudden change in a constraint, such as a transmission line is cut, the state of the grid might not evolve to a feasible steady state under a new set of constraints even though it exists. The mathematical formulation to solve this problem is known as stability analysis. Despite these challenges, the task of balancing demand and supply in the grid has been executed with only relatively few failures because both power consumption and generation are predictable and controllable.

The load

Understanding the behavior of the power consumption or the load in the grid is critical in the grid operation and planning. There are two key characteristics of the demand contributing to the higher cost of the grid operation: **variability** and **volatility**. The variability of the load can be found in the daily, weekly and seasonal basis. An extra load beyond the average load requires extra power supply, which is economically inefficient since the generator to serve the extra load is less often used. Furthermore, the highly variable load requires a power source with a higher rate of change in power known as **ramp rate**, which adds to the cost. If the load was deterministically known, matching the demand would be simple if generators with sufficient capacities and ramp rates are installed and the transmission constraints are satisfied. The cost of these generators depends solely on the variability of the load. In reality, there is uncertainty about when and how much power is consumed by different users in the grid. This causes the volatility of the load. Extra cost of ancillary services deployment comes from the volatility. However, the volatility of the load is small. The prediction error by existing forecast systems is also small.[2] The variability of the load emerges from dependent actions of consumers in the grid, while the volatility arises from independent decisions. The variation in the load pattern comes from daily variation in activity at roughly the same time such as opening a refrigerator to cook a dinner. However, since a large amount of people have slight variations for when to take an action, the variation in demand is smooth and the occurrences of a sudden significant change is negligible. Historically the load behavior has been well understood. The infrastructure and operational policies have been designed to match the needs of this type of behavior.

[2]The detail of the prediction accuracy is described in Chapter 7.

The distributed PV generation

In the last few decades, the threat of unsustainable and even harmful energy resources that supply electricity has been a growing concern. Many entities from scientists to policy makers have been seeking new ways to use clean and renewable resources. Photovoltaic (PV) technology is one of the candidates. **A photovoltaic (PV) system** comprises a semiconductor panel converting sunlight into direct current electricity and an inverter converting direct current to alternative current used in the grid. The deployment of PV systems has been increasing significantly due to supportive policies in many countries and several states in US. In 2015, more than one third of yearly PV installations is **distributed PV generation**, the PV deployment in which the PV systems are installed geographically near consumers scattered throughout the grid [11]. Business innovations and favorable regulatory policies attract homeowners and commercial entities to install PV systems on their roofs. Distributed PV has become the poster child of renewables providing not only clean energy but avoiding both the infrastructure to convert energy to electricity (no more eye sores such as cooling towers) and perhaps avoiding the need for large-scale transmission and its giant transmission towers. Since the distributed PV generation takes place where the power is consumed, one may view that the power supplied by distributed PV generation as a negative load. Hence, **the net load** in the demand side decreases, which may result in less use of unsustainable and harmful energy resources in the supply side. It seems that the problem of the current method of generation can be solved by introducing a certain amount of distributed PV generation.

The problem, nonetheless, cannot be solved easily as a new problem arises from the variable and intermittent nature of solar power. Even if we assume the most optimistic situation in which the weather and the panels always have an uninterrupted view of the sun, the power generated from the PV system changes drastically throughout a day. When the sunlight is blocked by clouds or shading from surroundings, the power from the PV system can sharply drop. Unlike mechanical based electricity generation, PV generation has no inertia that smooths the power output. As a result, any PV generation, including the distributed one, provides more volatile power than the current load from consumers and power from generators. If there is a high level of distributed PV generation in the grid, the characteristic of the net load will be significantly distorted from the normal load. Then the grid operator may not be able to balance with the current set of supply infrastructures. The current level of PV generation is not yet sufficient to cause grid operators a problem. With the high rate of increase of distributed PV deployment, clearly what the limits are to PV generation needs to be understood.

It is useful to have a model of the power output from a group of PV systems in the grid, which we term **the PV model**. An accurate PV model can help us to

answer critical questions: what is the limit on distributed PV generation under the current physical infrastructures and operational policies and how the grid operator can modify the infrastructure and policies to allow more distributed PV generation. The commonly used PV model for each individual PV system comprises multiple steps including finding solar irradiance in the sky, projecting solar irradiance to the panel, determining DC power output from the panel, and converting the power output compatible to the grid. [3] This structural model is not practical for the grid operators since it is difficult to keep track of so many factors for all PV systems in the grid. A more tractable model of the power output from a group of PV systems can be built based on historical PV power outputs from a sample of currently operating PV systems. We call this a **reduced-form approach**. In this work, we use this approach to build **the maximum power curve**, which describes the highest possible power output from a group of PV systems at any instance. Using the resulting maximum power curve as the core idea, we are able to answer critical questions regarding the impact of distributed PV generation on the operation of the grid.

Contributions

Contributions from this work are as follows:

In order to model the distributed PV generation of a large number of PV systems, it is not tractable for a grid operator to collect all factors to model each individual system. A reduced-form approach to model the distributed PV generation of a large number of PV systems requires no specific information of each PV system and is formulated using historical PV power output. A maximum PV power curve constructed from this approach is shown to be consistent and useful in many applications. One of the critical applications is to establish a limit on PV installation on the grid, given an assumption that the PV power cannot be larger than the load at any instance.

Using the maximum PV power curve, we are able to improve the definition of volatility and distinguish the difference between variability and volatility rigorously. It is worth noting that defining volatility to be relevant with the grid operation is not a trivial task. The volatility should be defined against a consistent reference at the correct sampling frequency.[4] These desirable properties are inherent in our definition. The notion of volatility is also shown to be closely linked with a short-term PV power output performance and the hour-ahead grid operation, while the notion of variability is embedded in the expected PV power curve derived from the maximum PV power curve and used in the day-ahead grid operation.

[3] More detail of the PV model is in Chapter 2.

[4] More detail on the consistency and correctness is in Chapter 6.

We show how using the maximum power curve as a reference improves the short-term forecast of the power output of a group of distributed PV systems. We have shown that a similar prediction technique can also be used in the short-term load forecast. Using a daily load curve of a previous day and a daily load curve of a previous week as references, we are able to improve the prediction performance of the load in the grid as well.

The last contribution is that we establish the limit of PV installation onto the grid, given an assumption that the grid can tolerate up to a certain level of risk of demand-supply imbalance. By day-ahead scheduling based on expected PV power curves and hour-ahead based on our forecast system, we are able to simulate the grid operation, assess the risk of demand-supply imbalance under different levels of PV installation, and identify the limit of PV installation. We are able to show that the current limiting factor to the level of PV installation is in the ancillary services, rather than generators in energy scheduling. From this, we are able to measure possible extra PV installation by adding ancillary services and adopting PV curtailment without modification on generators in energy scheduling.

Organization

The contents are organized as follows: Chapter 1 provides an in-depth explanation of how distributed PV generation can potentially impact the grid operation significantly. The key lesson is that the growth of PV generation is not solely by innate desire of consumers, but it is driven heavily by state policies. Then the rest of the contents are organized into three parts.

The first part, including Chapters 2-6, focuses on the distributed PV generation. Chapter 2 shows how the power output from a PV system can be modeled with numerous factors. The key issue is that the model here requires so many factors that it is not practical for a grid operator to track. Chapter 3 shows how one can apply the current model of the power output from a PV system to answer some problems of interest to PV owners and grid operators. The critical lesson here is that the optimal solutions for the optimization problems of interest to PV owners and grid operators are not necessarily aligned. Chapter 4 illustrates how we construct the maximum PV power curves, which describe the maximum power from a group of PV systems at any time instance. What distinguishes this approach from the model in Chapter 2 is that our reduced-form approach provides the consistent curve based on historical power output data from a relatively small sample of PV systems without additional numerous factors. Chapter 5 demonstrates some applications of maximum PV power curves from Chapter 4. Here we are able to establish the first bound on PV installation based on the fact that the PV power output should not be larger than the power consumption at any instance. Chapter 6 completes the first part and provides

a forecast system for distributed PV generation. Here the maximum PV power curve is shown to be useful in defining volatility isolated from variability. It also improves the prediction performance of the PV power forecast as well.

The second part, including Chapter 7 and 8, focuses on load modeling. Chapter 7 explains the nature of power consumption and how it can be accurately predicted. Without significant distributed PV generation added to the demand side, the load has relatively small variability and volatility. Chapter 8 provides an example of a short-term forecast system for load. We illustrate that the load can be predicted with a small error using **daily load curves**, which describe the power consumption in the grid at any time instance, from the previous day and the previous week. The prediction technique is similar to what we used in Chapter 6.

The third part, including Chapters 9 and 10, focuses on the grid with high distributed PV generation. Chapter 9 provides an overview of typical grid operations. Two critical components of the grid operations: energy scheduling and ancillary services are explained and linked to the concept of variability and volatility. In Chapter 10, the simulation of grid operations integrated with forecast systems in Chapters 6 and 8 is conducted to assess the risk of demand-supply imbalance. We estimate the limit of PV installation to the grid, and identify additional ancillary services and policies to improve the limit. The summary and remaining research questions are presented in Conclusions.

Acknowledgements

It is hard to mention all the people I met throughout my life who have helped and supported me and made me the person I have become. If you are not included please know that I am thankful for your support, love, and wisdom. I am grateful to have live a moment with you. There are some people who are so important that I hope I can highlight them without making others feel left out.

First, I would like to thank Walter Murray, my PhD adviser and my adviser for life. As many people know, he is like a living Wikipedia. The way he knows and juggles different subjects is impressive. What is more inspiring is his wisdom and kindness. I enjoyed our time when he advised me how to approach research problems and other academically related issues, and I cherished the times he taught me how to care about things around us and be a man of virtue for a family, a community, and even the world. What I learned from Walter are not only knowledge and skills, but also character. Walter, I hope one day I can be a great teacher like you.

Second, I would like to express my gratitude to my PhD thesis committee. John Weyant who read my thesis and taught me how researchers like us can contribute to the solution of climate change and energy problems. Robert Entriken who was also a reader and for whom I benefited from his extensive knowledge of the electrical grid. Yinyu Ye who was on my thesis defense and provided useful feedback during the exam.

Third, I would like thank Keak, my childhood friend. I have to admit that back then I was not likable, even worse than now, yet he still wanted to be my friend. What I remembered most about him is his passion to become an astronaut. It was such a hard dream to achieve but it drove him to study harder and to succeed in many other things. Seeing his example, I realized that we should put in our best effort to get a better education and to pursue our dream. Being friends with him is like a spark to a chain reaction of Physics Olympiad, education in the United States, and many more. Thank you for being with me.

Next, I would like to thank Fah, my brother. Throughout our lives, we fought over so many silly things like food and house chores. My relationship with my brother in early years was basically a competition. I lived under the pressure that my brother had aced in everything. As a result, I tried to study harder, play checkers better,

and eat more vegetables than he did. My perspective on him flipped when our father passed away. At that time it was hard to leave my mother behind and come to study in US, but my brother confidently said that he would take care of our mother for me. Ten years since then, he committed to what he said and has become the leader of our family. I am fortunate to have you as my brother.

Last but not least, I would like to thank my parents, Kumpol and Wilai. Our family is nowhere near being the ideal family. I have caused many troubles, made them worried, and neglected them many times. Regardless, they love me unconditionally. They raised me according to a simple principle: to provide me the opportunity to learn as much as I can without letting me starve. Both of my parents never finished grade school, but they raised me until I got a PhD. Dad, it is unfortunate that you are not physically with us at this precious moment, but you are still in our hearts. Thank you for what you had done for us and take a good rest until our family reunites again. Mom, you once said you were embarrassed to chat with others because you were not well educated. Please tell those people that you are even better than PhDs because you have brought one up. This is not my success, it is yours.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Highlights:

- Solar photovoltaic (PV) deployment, particularly distributed PV deployment in the United States, is growing and has yet to reach its maximum potential.
- The emergence of renewable-energy policies, business innovation, and technology drives the growth of distributed PV deployment.
- Although distributed PV deployment may give some benefits to Utilities, Utilities encounters challenges to control the grid caused by the variability and volatility of solar generation.
- Utilities should modify their physical infrastructure and operational policies to handle this growing distributed PV generation.

Energy is an essential resource to drive a modern civilization toward a better world for humanity. However, a threat of unsustainable and even harmful energy resources become apparent. Pollution can be in a local scale such as leaked radiation from nuclear fuels or smokes from kerosene, and be in a global scale such as greenhouse gas emission from fossil fuels. It is not difficult to project that these resources is going to depleted as well. Innovators from various sector from scientists to policy makers are working hard to find new ways to supply clean and renewable energy.

Solar photovoltaic (PV) is a clean and renewable energy resource with high growth and potential in recent years. While a scientific technology for PV is improving, several innovations in other sectors for supporting the PV deployment emerge. Globalization and verticalization of PV supply chains drives the cost of PV deployment down, while business models and policies facilitates the PV deployment. Recently, there are some business models and policies favoring a certain type of PV deployment called a distributed PV deployment, where PV systems are geographically distributed within the same area as the energy consumption.

Although solar PV deployment is a promising choice of clean and renewable energy resources, it is difficult to manage because of its nature of variability and volatility. A distributed fashion of generation adds more difficulty since distributed PV systems are generally not controlled by Utilities or grid operators. Utilities or grid operators need to modify their infrastructures and operational policies to handle the challenges from distributed generation.

In this chapter, we first review PV growth and potential with a focus on distributed PV deployment in the US. Next, we highlight key drivers for distributed PV deployment. Finally, we disseminate challenges from distributed PV generation to guide a further discussion in this study.

1.1 PV deployment growth and potential

1.1.1 Growth over time

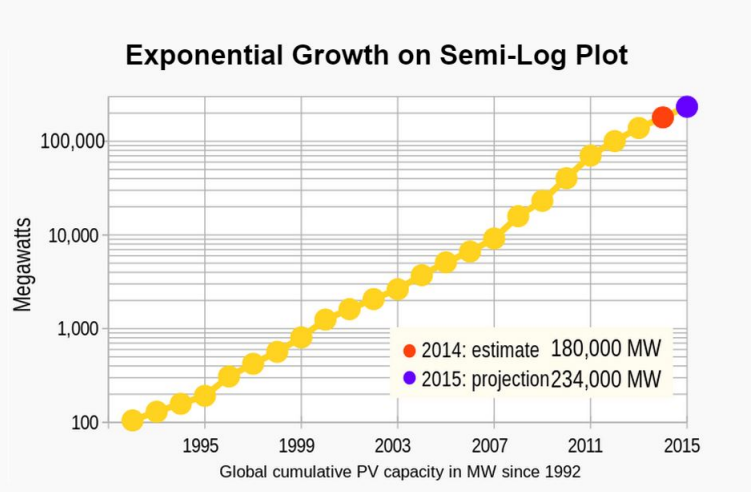


Figure 1.1: Worldwide exponential growth of PV deployment from 1992- 2014

PV deployment has grown significantly over time since 1992 as shown in Figure 1.1. By the end of 2014, the installed PV capacity was more than 177 GW, enough to supply 1% of the world's electricity consumption of 18,400 TWh [58]. As a quick reference of the rate of installation, in 2015, new PV systems are being installed in the United States at a rate of one every two minutes [48]. Note that a typical residential system size is 3-5 kW [27].

1.1.2 Growth over location

PV deployment growth is a global phenomenon. Leading nations in renewable energy adoption, including China, US, Germany, Japan and India, are now deploying PV at a fast pace. Even the Gulf States where fossil fuels are abundant, and countries in Africa, where the economy is relatively small, are following this trend. Environmental concerns from the use of fossil fuels, radioactive materials, and kerosene make PV deployment in these unlikely places preferable. In addition, in a large number of countries near the equator, the solar resource is relatively high, yet the PV adoption rate is low as shown in Figure 1.2. Therefore, the current global PV deployment is still far from the maximum potential.

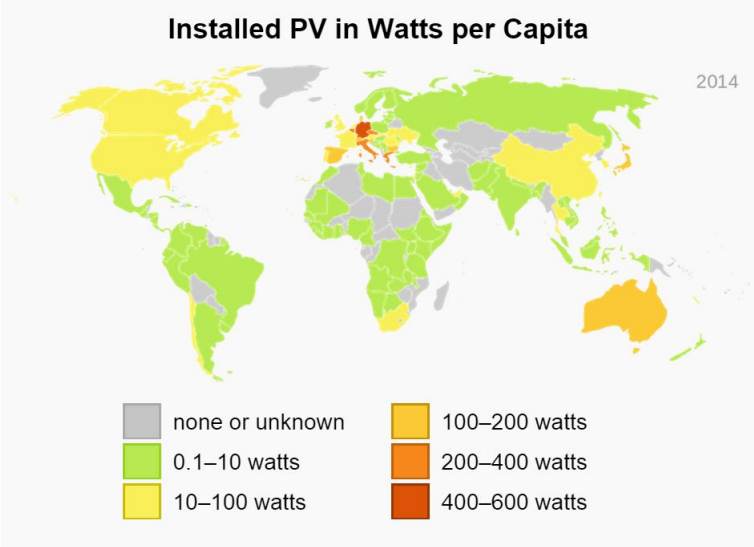


Figure 1.2: Installed PV in watts per capita in 2014

In the US, PV deployment is spreading throughout different geographical locations. Table 1.1 shows the solar capacity installed by January 2015 for the top 10 States. Two implications can be derived from this data. First, States on the list such as New Jersey and Massachusetts do not have ideal weather for solar PV generation compared to other states in the US. Their ranking must be due to extra financial incentives and legislation, which we explore in the next section. Second, there is high potential for new PV installations in many States. Many States with a large population and favorable weather such as Florida and Georgia are not leading users of PV generation.

Rank	States	Installed PV capacity (MW)[9]	Population (M)[14]	Installed PV in Watts per Capita
1	California	9,977	39.14	255
2	Arizona	2,609	6.83	382
3	New Jersey	1,451	8.96	162
4	North Carolina	953	10.04	95
5	Nevada	789	2.89	273
6	Massachusetts	751	6.79	111
7	Hawaii	447	1.43	313
8	Colorado	398	5.46	73
9	New York	397	19.80	20
10	Texas	330	27.47	12

Table 1.1: Top 10 US states by cumulative solar capacity installed in January 2015

1.1.3 Growth over sectors

It is common to define sectors in the PV market as residential, non-residential (including commercial, industrial, government and nonprofit) and Utility. Here we define **distributed** or **rooftop** PV generation as PV generations that not controlled by Utilities or grid operators. The common characteristic of distributed PV generation is that it is geographically distributed within the same area as the energy consumption. According to the 2014 SEIA and GTM US Solar Market Insight report, in 2014, the distributed PV installation was more than 2 GW and almost 40% of the overall PV market as shown in Figure 1.3. Note that the growth rates between distributed and non-distributed PV generation are different due to various reasons. One of favorable aspects toward the distributed PV generation, especially residential PV generation, is that it does not require a lengthy administrative process and decision making. The next section covers how other factors drive the growth in distributed PV deployment.

Note: Sectors as discussed here are all PV deployments connected to an electric grid. In fact, there is an additional sector called an off-grid PV deployment where the PV is not connected to an electric grid, likely because there is no electric grid there. The idea of an off-grid PV deployment is similar a use of mobile phones over land line phones which are associated with a high infrastructure cost. However, we do not discuss it in this study.

1.1.4 Estimate of the maximum potential in the US

The estimates of future distributed PV deployment and renewable resources vary depending on different methods and assumptions. Lopez et al. [30] had discussed

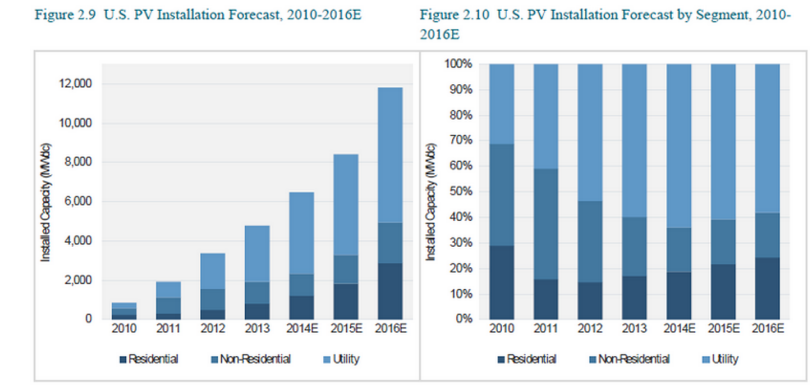


Figure 1.3: US PV installation forecast 2010-2016 from 2014 US Solar Market Insight report

four different types of assumptions: resource, technical, economic and market. Provided the assumptions are correct it may be this approach will give accurate results. However, their approach is itself based on prediction of other quantities such as the growth of suitable rooftops. Clearly if such prediction prove inaccurate so will the conclusion about PV growth estimate.

Lopez et al [30] estimates distributed PV deployment potential to be 664 GW by assessing rooftop area and assuming some technical specifications of PV panels. Denholm, Drury and Margolis [56] adds economic and market assumptions including the growth of roof's area and estimates distributed PV deployment potential to be 583 GW in 2030. In perspective, these are estimates are about 75 % of current peak load, which is about 750 GW [26]. These estimates exceed the US distributed PV deployment by 2014, which is less than 10 GW [12].

For California, Lopez et al [30] estimates PV distributed potential to be 76 GW. In 2012 Energy and Environment Economics Inc. adds a technical constraint from the electric grid called "no back flow" criterion and estimates distributed PV deployment potential to be 15 GW [28]. The estimate can be larger if we allow curtailment or storage technologies. In perspective, the peak load from 1998 to 2014 is about the order of 50 GW [16]. These estimates exceed the distributed PV deployment by 2014, which is less than 5 GW [12].

While the distributed PV generation in 2014 is relatively small, distributed PV generation in California and the US can play a significant role in power generation if the distributed PV generation keeps growing and reaches its full estimated potential.

1.2 Drivers for distributed PV deployment in the US

1.2.1 Policy

Federal policy, solar Investment Tax Credit (ITC) is one of the most important mechanisms for distributed PV deployment. The ITC allows a tax credit equal to 30 percent of a distributed PV projects costs. Starting in 2006, ITC is a tool for solar developers to attract tax equity investors and develop their financing models. Under a bill in December 2015, 30% Solar Investment Tax Credits for both residential and commercial projects was extended through the end of 2019, and then drops the credit to 26% in 2020, and 22% in 2021 before dropping permanently to 10% for commercial projects and 0% for residential projects [10].

In the State-level policy, Renewable Portfolio Standard (RPS) is a mechanism that obligates electricity supply companies to produce specified fraction of their energy from renewable energy sources. 33 States and DC have adopted RPS. RPS leads to Renewable Energy Certificates (RECs) - tradable, non-tangible energy commodities that represent proof that 1 MWh of electricity was generated from eligible renewable energy resource. Distributed PV providers may sell their credits to non-renewable electric supply companies and makes additional profit from RPS.

In California, the first highest cumulative PV capacity installed state in 2014, the state mandates a RPS of 33% by 2020. Figure 1.4 shows the mix of renewable power in California RPS includes wind, solar thermal, solar PV, small hydro, geothermal, and biopower. In 2013, three large Investor-Owned Utilities (IOUs) collectively served 22.7% of their 2013 retail electricity sales with renewable power. Note that here biopower is defined as biomass and biogas technologies. In 2014 California Public Utilities Commission (CPU) reports that solar PV is about 15% of total renewable resource mix and expected to contribute 40% of total renewable resource mix by 2020 [19].

In Arizona, the second highest cumulative PV capacity installed state in 2014, the state mandates a RPS of 15% by 2025 and 30% of renewable to be distributed generation. [17].

In New Jersey, the third highest cumulative PV capacity installed state in 2014, the State mandates a RPS of 22.5% by 2021 and 4.1% specifically for solar electric generation by 2028. [18].

Other than federal and state-level financial incentive-based policies, the administration policies for distributed PV generation are also crucial. Speer [66] discussed this issue extensively. A favorable technical and procedural processes for connecting a distributed generation system to the grid would facilitate the growth of distributed PV deployment. One of the successful practices is a **net energy metering** as shown

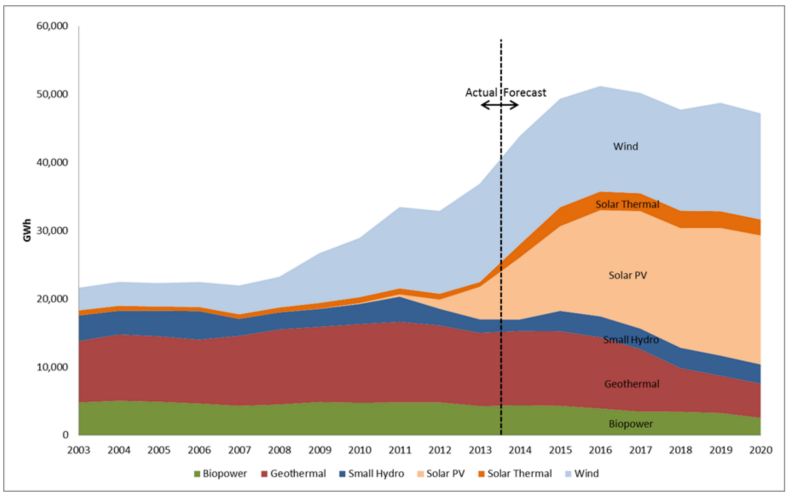


Figure 1.4: Renewable energy resource mix in California renewable portfolio

in Figure 1.5. Net energy metering is a billing procedure that allows a customer to give a surplus energy to the grid and consume energy from the grid when it is needed. A review of current and year-to-date charges and credits is sent monthly, and a bill of the net charges and credits is cleared annually. Note that the calculation of net charges and credits is simple with a fixed price of electricity. Under time-of-use pricing plan, the calculation of net charges and credits is complicated. This complex scheme is more preferable since it gives more benefit for distributed PV owners who want to help reducing the net load in a period with high demand. In 2015, 41 states have some level of mandatory net-metering policies, which allow solar and other generating customers to sell excess power back to the grid [48]. In California, the compensation rate for surplus power under the net energy monitoring program is fixed and based on a 12-month average of the market rate of energy, or roughly 3 or 4 cents per kilowatt-hour [35]. Note that a customer still needs to pay minimum service charge regardless of the amount of the net energy.

There are some legal and regulatory questions that we need clear answers in order to establish a strong ground for distributed PV generation practice. Such question include whether third-party PV entities should be regulated as a wholesale electric Utilities and whether customers who receives benefits from the net energy metering program should pay an extra fee. Legal and regulatory clarity would help the sustainability of this new industry.

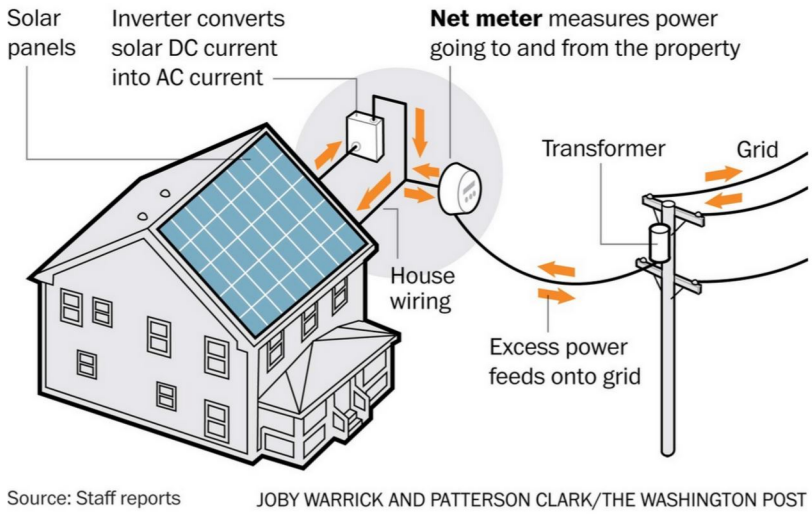


Figure 1.5: Net energy metering: the system that enables distributed PV surplus for energy credit or even cash for generating electricity

1.2.2 Business innovation

Direct investment by a home owner in PV installation requires a high upfront cost and the risk of high maintenance costs. Two recent business business innovations have been developed to address these issues and drive distributed PV deployment.

The first business innovation is third-party solar financing. Third-party solar financing companies usually offer two financing options over a term of about 20 years. The first option is a solar lease: a customer pays a fixed amount to rent a PV system regardless of the amount of electricity generated. Another option is a Power Purchase Agreement (PPA): a customer pays for all amount of electricity generated by a PV system with a rate competitive to local retail electricity prices. From the perspective of customers, these options reduce the high upfront cost and reduce risk. Note that operations and maintenance service may not be included in a solar lease option while such service is included in a PPA [66]. Hence, customers under a PPA option are relieved of possible system failures. From the perspective of solar financing companies, even though the companies pay the upfront cost of PV installation, a bundle of solar contracts becomes reliable assets. Tax credits and RECs from these credits can be traded with tax equity investors and Utilities.

Another business innovation is called “verticalizations”. Some solar companies go beyond their current roles in the solar PV industry and take over the whole chain. For example, SolarCity, previously focusing on the lower chain including financing and installation, has acquired a PV manufacturer named Silevo and aimed to build a large PV plant in Buffalo in 2014 [44]. On the other hand, Panasonic, which previously

focused on the upper chain including panels and electronics manufacturing, now provides end-to-end solution including financing along with operations and management [54]. In many ways this follows similar development in the auto industry where the finance arm become a major source of profit. The verticalization reduces the cost of a PV system by removing middleman commission and ensures solutions for possible system failures after installation as a part of service in the chain.

1.2.3 Technology

The SunShot vision study [68] illustrates that several efforts in research and development has reduced the PV system price. On the module side, technology and manufacturing process improvements have dramatically reduced the module price per watt over the last 30 years as shown in Figure 1.6. There are two ways to continue the improvement: increase the efficiency of the module and decrease the production cost. One opportunity to increase PV module efficiency would be to close the gaps between production, laboratory and theoretical (maximum) PV module efficiencies as shown in Figure 1.7. We expect to see further innovations to decrease the production cost when new PV technologies in laboratories are put in a production. Such examples in the past have included silicon slicing and large-scale thin film deposition.

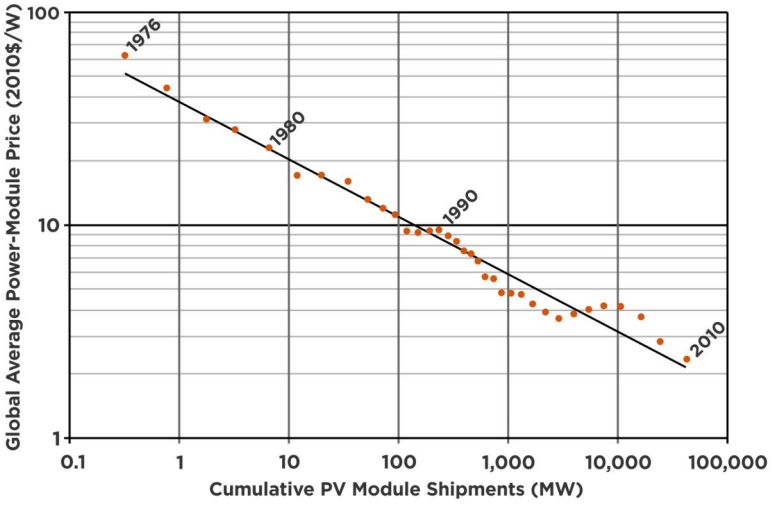


Figure 1.6: Decline in factory-gate PV module prices with increasing cumulative module shipments 1976-2010

In the balance-of-system (BOS) side, there is some room for improvement in research and development. Goodrich et al. [3] reported that the BOS price is higher than the module price for benchmarked 2010 distributed PV systems. Whereas

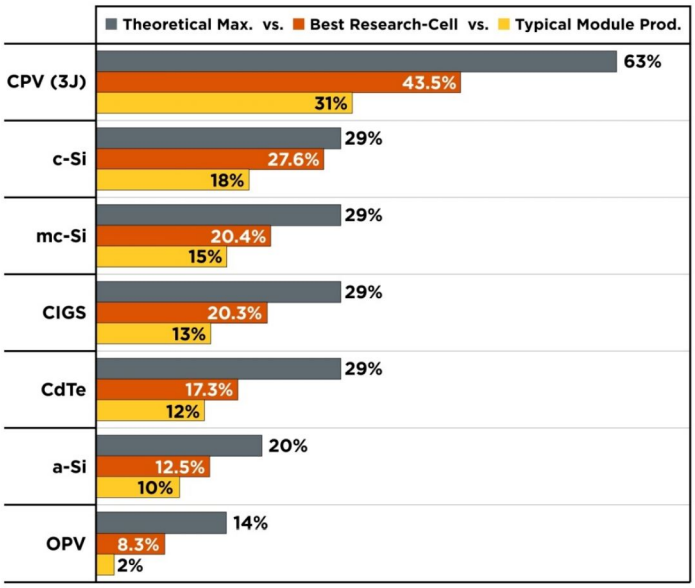


Figure 1.7: Production, laboratory, and theoretical (maximum) PV module efficiencies

blended average module pricing has dropped 79% between 2007 and 2014, BOS costs have been slower to respond, averaging between a 39% - 64% drop depending on the market sector and geography [50]. For the hardware component, the most expensive component is the inverter. Traditional inverters, also known as string inverters, have very high conversion efficiency but no capability of single panel power optimization. Recently, a new type of an inverter called a micro-inverter is adopted. Micro-inverters have single panel power optimization and allow for independent operation of the panel. However, micro-inverters have relatively low conversion efficiency and a high price. Research and development are needed to bring the cost of micro-inverters down. New technologies for racking and mounting materials are also important. Regarding the non-hardware component, some reduction in the BOS cost has been achieved by the verticalization in business innovation. Once the industry has matured, we expect that installation costs and supply chain costs will become cheaper and standardized as in other service industries.

1.3 Challenges in distributed PV generation

1.3.1 Difficulties for grid control

In traditional electric generation, distributed generation relieves line stress in the grid because the location of generation is near the location of consumption. Yet distributed generation is difficult for grid control. As there are more physical adjustments in the system, more computations and resources are required to control the grid. In distributed PV generation, the situation is even more challenging. First, distributed PV generation does not relieve line stress when sunlight is not available. The grid still needs to have lines of large capacity to provide back up unless there is generation or storage on site. Second, most current distributed PV generations do not have a centralized adjustment system. The grid control needs to view distributed PV generations as given parameters, not adjustable variables in planning and operation. Smart inverters, that is, inverters with bidirectional communications may partially solve the problem since a grid operator may take control over a group of PV systems when the curtailment of over-generation is necessary. Yet the deployment of smart inverters as ancillary services in grid operation is still small.

1.3.2 Coping with high variability and volatility

PV generation has high variability due to the character of available sunlight through a day and a year. Because this variability comes from the earth's self-orbit and its orbit around the sun, it has a consistent pattern and can be taken into account in regular planning such as day-ahead energy scheduling. A daily load profile can be estimated by counting PV generation as a negative load. Because the PV generation is highly variable and its peak generation is not aligned with peak load in general, it creates an irregular-shaped load profile where the net load rises steeply. In California, such a shape is known as a "duck curve" (Figure 1.8)[15]. This steep change of the net load requires complementary energy generation with a high ramp rate, which implies a higher generation cost.

Furthermore, PV generation has high volatility due to the weather. Cloud, fog, rain, and snow directly block sunlight from PV panels. Temperature, pressure, humidity and wind speed affect the performance of PV systems. Because this volatility comes from weather condition, it is difficult to predict and take into account in day-ahead energy scheduling. In 15-minute ahead energy dispatch, the grid operator needs to dispatch energy generation in accordance with uncertain PV production. Unlike the difficulty from variability, the rise and fall in PV generation may occur more sharply and more frequently. Utilities need to pay extra cost to counteract this volatile PV generation by adding more fast response energy generators and storage.

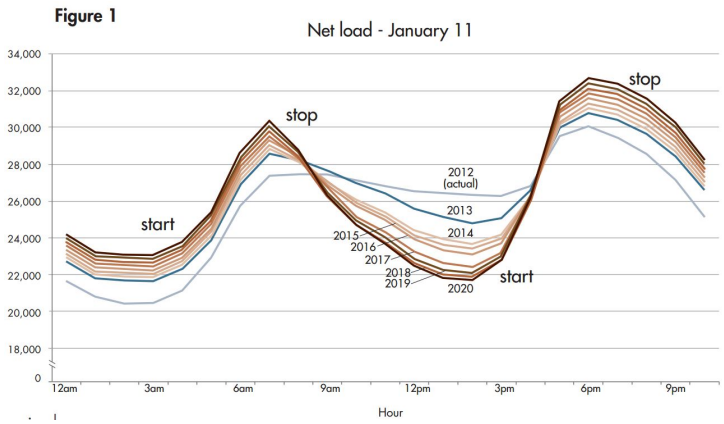


Figure 1.8: An example of net load (MW) in CA with projected solar and wind penetration from 2012 to 2020 according to CAISO. In this figure, a steep load rise occurs at around 5 pm

The worst case is to disconnect PV panels in a controlled manner and ramp traditional generators up and down to meet load demand instead.

The outlook is optimistic for improvement in distributed PV generation. Significant amount of energy storage can solve problems caused by high variability and volatility from PV generation. It also relieves line stress when sunlight is not available. This energy storage can be in the form of energy storage at a distribution substation, home energy storage, or even electric car batteries. For example, in 2014, PG&E started deploying energy storage at a distribution substation in compliance with the goal of California state of 1.3 GW by 2020 [60]. Several distributed PV providers started selling storage from their energy storage alliances or even building their own storage. Moreover, many researchers are investigating the concept of using fleets of electric cars as temporary storage, namely, vehicle-to-grid (V2G). It is worth noting that energy storage can serve as ancillary services in grid control as it is more reactive to signals from a power dispatcher compared to power turbine generators [8]. Yet the deployment of small energy storage as ancillary services may not be feasible in the near future because it requires hardware deployment, regulatory legislation, and multi-party commitment to share costs and benefits.

1.3.3 Relationship with Utilities

Distributed PV generation is beneficial to Utilities in three ways. First, as addressed above, distributed PV generation partially reduces line stress, which reduces the need for extra expensive transmission lines. Second, PV generation reduces power stress on summer days when demand is high and Utilities need to pay a high cost for extra

supply, which reduces the need for extra expensive power plants. Third, it helps Utilities to meet the Renewable Portfolio Standard (RPS) by trading Renewable Energy Credits (RECs). As a result, Utilities may avoid the cost of building new infrastructures.

However, distributed PV generation is also a threat to Utilities. First, it reduces Utilities sells less retail electricity as people use power from their own PV systems. In some sense, PV providers are "free-riding" Utilities infrastructure for their own profit. In addition, distributed PV owners can earn energy credit or even cash for the surplus energy they produce via a net energy metering. This means net energy metering customers are "free-riding" the infrastructure of the Utilities as well. Second, Utilities need to be responsible for coping with the variability and volatility of distributed PV generation, as explained above. This unfairness causes Utilities in many states to fight against current laws favorable to the distributed PV industry [42].

Weighing these costs and benefits, how should Utilities make a fair compromise with a distributed PV sector? One solution would be to add a surcharge to distributed PV providers and customers with net energy metering. Another solution would be to charge an additional infrastructure cost to every customer on the grid as every customer has shared infrastructure costs in the past. In 2015, more than 60 Utilities in the US proposed an increase of these fixed charges [48]. However, Utilities still need to solve an operational challenge: how to modify current physical instruments and operational policies so that customers can still meet their electricity needs given the issue of variability and volatility. These costs of physical and operational changes to accommodate high distributed PV generation in turn shapes how grid operators or Utilities do a cost-benefit analysis as well. In any event, distributed PV generation will be a significant part of a modern grid in many regions of the world. The grid operator will inevitably face challenges resulting from the variability and volatility of distributed PV generation.

1.4 Conclusion

Distributed PV generation is growing fast due to government policy, technology, and finance innovation, and it is still far from the upper bound. However, despite its benefits, distributed PV generation introduces challenges to the grid operator by the particular nature of distributed and solar aspects of the generation. After all, distributed PV generation will have a limited but non-trivial role in an electric grid. A grid operator must know the maximum value of distributed PV generation, in which the current physical instruments and operational policy can still provide electricity to customers. Furthermore, the grid operator needs to understand how to modify physical instruments and operational policies to accommodate distributed PV generation beyond the capacity of the current system.

Part I

Distributed PV generation

Chapter 2

Overview of a PV System Model

Highlights:

- A model of power output from PV cells, modules, or arrays comprises multiple steps including finding solar irradiance in the sky, projecting solar irradiance on a plane, determining DC power output from panels, and converting power output from the PV system compatible to the grid.
- Modeling clear-sky solar irradiance is quite simple, while modeling actual solar irradiance is complicated. The process of data measurement from the ground or satellites is expensive, and the data from the two sources are significantly different.
- Projecting solar irradiance on a plane is quite straightforward unless we consider shading, soiling and, reflection loss.
- Determining DC power output from panels is difficult. The DC power output is not necessarily proportional to solar irradiance on a plane. The PV cell temperature plays a significant role in the power output as well. As we consider the system at the array level, module wiring is another significant factor to consider.
- Converting AC power output from DC power output is not straightforward when power loss in direct current wires is accounted.

Modeling the power output of a PV system is a complicated process as it involves a large number of factors. Some of them are well documented (e.g. system configuration) or consistent (e.g. the sun's position). However, many factors are poorly documented (e.g. the surrounding shading environment) and inconsistent (e.g. the weather pattern). Here we give an overview of the current model of a PV system. This overview demonstrates why a data-driven model is necessary in our problem.

2.1 Clear-sky solar irradiance models

Note: Most of the material in this section is derived from "Global Horizontal Irradiance Clear Sky models" by Reno, Hansen, and Stein (2012) [49].

Clear-sky solar irradiance is power received by sunlight per unit area on the earth, assuming there is nothing obscuring the sunlight such as clouds, haze, and smoke in the atmosphere. One can measure clear-sky solar irradiance in three ways: global horizontal irradiance (GHI), direct normal irradiance (DNI), and diffuse horizontal irradiance (DHI). There is a geometric relationship among these 3 kinds of irradiances:

$$GHI = DNI \cos(z) + DHI, \tag{2.1}$$

where z is the elevation angle of the sun. Several kinds of irradiance are necessary to derive the power received by an arbitrarily oriented plane known as a plane of array (POA) irradiance. There are several numbers of models for each kind of irradiance. The simplest model, as discussed by Reno, Hansen, and Stein (2012) [49], requires only the elevation angle of the sun z, which can be determined with two-dimensional geographical location and the time of year. The more complicated models allow the extraterrestrial radiation, I_0, to vary throughout the year. Yet these models still require only the two-dimensional geographic location and the time of year.

More complicated models consider the third geographic coordinate called elevation (h). Extra parameters representing optical atmospheric properties including air mass (AM) and the Linke Turbidity coefficient (TL) can be introduced. Since by definition air mass is the optical path length through the atmosphere where light is scattered and absorbed, it depends only on geometric parameters such as the elevation angle of the sun and the radius of the earth. However, since the Linke Turbidity coefficient is a property of a cloudless atmosphere relative to a water-and aerosol-free atmosphere, this coefficient depends on the climate of the location. The Linke Turbidity coefficient varies significantly throughout the year but the variation across years is negligible [23]. Note that there are even more complex models considering the variable Linke Turbidity coefficient which we did not discuss here.

Up to this point, we can model clear-sky solar irradiance at a location as a function of time of year only. Extra quantities including altitude (h) and monthly values of the Linke Turbidity coefficient (TL) at a given location can be used for more complicated models. Those quantities can be provided by a solar energy service provider such as Solar Energy Services for Professionals (SoDa) [33].

2.2 Actual solar irradiance

2.2.1 Data: ground measurement

The ground measurement data of actual solar irradiance can be represented in different ways. A collection of raw data spanning some years can be reported as is, which gives an opportunity for further investigation. However, the amount of data may be too large to comprehend for simple applications. Another way to report such data is to average the measurements at different points during of a year. This condensed measurement is called average year data. Even though this data is now easily understood, it diminishes the volatility of data. One other common practice to report such data is to select some data to present weather phenomena in a year while giving annual averages consistent with the overall data. This practice is called typical meteorological year (TMY) data. Procedures for developing the TMY data can be found in Wilcox and Marion (2008) [72]. According to their report, this data format is widely used by building designers and others for modeling renewable energy conversion systems but it should not be used for evaluating real-time of solar conversion systems as it is not designed to provide extreme values. Hence, this representation is not particularly useful for daily grid operation either.

Unfortunately, most publicly available and easily accessible datasets for ground measurement are in the form of either average or TMY data. In the United States, the National Solar Resource Data Center (NSRDB) run by NREL has two standard TMY data: TMY2 derived from 1961-1990 for 237 ground sites and TMY3 derived from 1991-2005 for 1454 sites. In Euporean countries, ESRA and PVGIS provide average and TMY data. The World Radiation Data Centre (WRDC), supposedly the central depository for solar irradiance data from ground measurement of the world, has incomplete data even for TMY data. The Measurement and Instrumentation Data Center (MIDC), slso run by NREL, provides historical raw data at 1-minute intervals, but there are records for only about 40 sites. Cooperation between organizations and transparency of raw data are necessary for grid operation study.

2.2.2 Data: satellite measurement

Data of actual solar irradiance measured by satellites is used in real time with more locations relative to ground measurement. For example, in the United States, SolarAnywhere, a satellite-based weather data service provides the 1-km resolution data for the NREL National Solar Radiation Database (NSDRB) [59]. Note that there are about 10 million square meters in the United States. Hence, satellite measurement is much more scalable than ground measurement.

One question to consider is how well satellite measurement and ground measurement agree. Bing et al. [41] compared SolarAnywhere datasets and ground datasets

over 7 solar monitoring sites with Rotating Shadowband Radiometers (RSR), measuring both GHI and DNI in Sacramento, Cailfornia, with 6 months of data. For a half of hour timeframe, the mean absloute error (MAE) in GHI is $6 - 11\%$. For DNI, MAE is $17 - 22\%$. Examining 8 years of data, Ineichen (2013) [40] conducted a comparison of both GHI and DNI from 18 solar monitoring sites across Europe. The best satellite irradiance product in the study called SolarGis gives the standard deviation in GHI of 17%. For DNI, the standard deviation is 34%.

Here one can see that even though actual solar irradiance from satellite measurement exhibits higher accuracy in terms of location, it yields more than 10 percent error compared to ground measurement. The use of satellite measurement consequently is a source of error when we try to complete the picture of a PV model.

2.3 Solar irradiance on a plane

In a PV model, once either clear-sky solar irradiance or actual solar irradiance is registered, it can be projected on a plane in order to compute the power output of a PV system. Such irradiance is called a plane of array (POA) irradiance. POA irradiance consists of three components:

$$I_{POA} = I_{beam} + I_{ground} + I_{diffuse} \tag{2.2}$$

I_{beam}, irradiance from the direct normal irradiance, is the most straightforward to compute. It equals DNI multiplied by cosine of angle of incidence. I_{ground}, irradiance from ground reflection, is more complicated than I_{beam}. It involves GHI, cosine of tilt angle of the surface, and a quantity of the ground called albedo. Albedo is close to 0 when the ground is dark and close to 1 when the ground is bright or reflective, as with metal. For example, the estimate of albedo is about 0.2 for an urban area and 0.8 for snow covered ground.

$I_{diffuse}$ is the most complex component of I_{POA}. There are several models based on three types of diffused irradiance: isotropic, circumsolar and horizon brightening. An isotropic component is an explicit function of DHI. A circumsolar component involves DNI, while a horizon brightening component is a function of GHI as well. Models for $I_{diffuse}$ may take several types into account. Some models add empirical correction and additional assumptions.

It is also worth noting that I_{POA} does not necessarily equal the power entering a light-to-electricity unit since there are possible shading, soiling, and reflection losses that block radiation.

Up to this point, we can model solar irradiance on a plane, but we need more information about our PV system and surrounding environment. Some quantities, including the panel orientations, may be known by solar service providers but not grid

operators. Other quantities, including ground quantity, shading and soiling, may not be known even by solar service providers. Without proper maintenance, the power quality from a PV system may decrease significantly. Incomplete knowledge of these factors contributes to error in a PV system model.

2.4 Conversion from light to DC power output

An object that converts light to electricity exhibits the following hierarchy: a cell, a module, and an array. The smallest unit, the cell, generates direct current (DC) electricity when photons excite electrons in the cell. The power generated can be derived from a characteristic IV curve of the cell as shown in Figure 2.1 [47]. The IV curve comes from a characteristic equation of an equivalent circuit of a solar cell. This is known as the 5-parameter model equation:

$$I = I_L - I_0\{\exp[\frac{q(V + IR_S)}{nkT}] - 1\} - \frac{V + IR_S}{R_{SH}}, \tag{2.3}$$

where I_L is a photogenerated current which is proportional to the input irradiance of the PV cell.

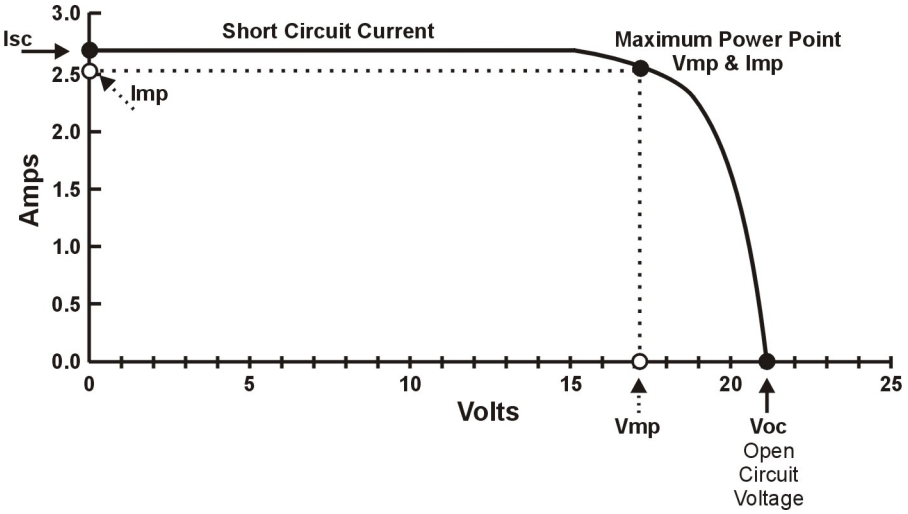


Figure 2.1: Characteristic IV curve of a solar cell.

There are 3 important points on the IV curve. The first point is the x-intercept $(V_{OC}, 0)$, where V_{OC} is defined as open-circuit voltage. The second point is the y-intercept $(0, I_{SC})$, where I_{SC} is defined as short-circuit current. The last point is the maximum-power point (V_{mp}, I_{mp}). This is where $P = IV$, the power generated by the cell, reaches maximum. In practice, the load is controlled in the circuit so that the

system reaches the maximum-power point. Such control is called Maximum-Power Point Tracking (MPPT). There is no closed form for the power output under MPPT.

The effect of input irradiance on the power output under MPPT is not linear. The effect of cell temperature, T, on the power output under MPPT is not linear either. In addition, the temperature significantly affects the diode saturation current I_0. Empirically, we describe the temperature effect using temperature coefficients for V_{mp} and I_{mp} which are constant for a range of $10 - 50°C$. The effect of the cell temperature to the power output under MPPT can then be described as

$$\frac{dP_{mp}}{dT} = V_{mp}\frac{dI_{mp}}{dT} + I_{mp}\frac{dV_{mp}}{dT}, \tag{2.4}$$

where $\frac{dI_{mp}}{dT}$ and $\frac{dV_{mp}}{dT}$ are temperature coefficients. Numerical values of temperature coefficients for different commercial modules at a standard condition are reported by King, Kratochvil, and Boyson [21].

It is common to model the power generated by a larger unit, a module, in similar fashion even though there are complicated issues including partial shading on a module and wiring of cells in a module. Modeling the biggest unit, an array, requires knowledge of a partial shading pattern. The wiring of modules in an array, that is, how many modules are connected in series and how many modules are connected in parallel, must be known. A model and simulation scheme is needed to provide an IV curve and to perform MPPT for an entire system [55]. Micro-inverters may increase the power output and reduce the optimization issue by individually optimizing each module. It is worth noting that even a small fraction of shading on an array may cause a large fraction of power loss. In other words, the power output from a PV system is fragile and sensitive to the shading environment.

Up to this point, we can model DC power output from a cell, a module, and an array, but we need more information about our PV system and its environment. Some quantities, including the module specifications and the wiring of modules, may be known by solar service providers but not grid operators. Some quantities, including the cell temperature and partial shading pattern, may not be known by even solar service providers. Even though one may know a weather condition such as air temperature, wind speed, humidity, and pressure, it is not straightforward to predict the cell temperature since the cell heating is not an instantaneous process. Incomplete knowledge of these quantities contributes to error in a PV system model.

2.5 DC-DC converters, DC-AC converters, and PV power output

Once we have DC power output from PV panels, a DC-DC converter boosts DC electricity to match voltage in the grid. Then DC-AC converter, namely inverter, converts DC electricity to Alternative Current (AC) electricity to match current type in the grid. During these processes, there is possible power loss. There is no formula to determine such loss. While DC-DC converter has very high peak efficiency of more than 99%, DC-AC converter has peak efficiency of 96%-98% and average efficiency of 95%-97%. There is also possible power loss as heat due to high DC currents in wires before the inverter. Such loss is difficult to estimate since it depends on modules wiring.

The use of micro-inverters play a role at this step as well. Micro-inverters do not only optimize the DC power output from individual modules but also convert power output from DC to AC right at individual modules as well. As a result, the power loss in wires are lower since power is now transported around the systems as AC rather than DC. The peak efficiency of micro-inverters is close to regular inverters but it is not clear if micro-inverters is cost efficient or not.

Up to this point, we can model the actual PV power output, but we need more information about electronics associated with the PV system. Some quantities including the converter efficiency may be known by solar service providers but not grid operators. Some quantities including the wiring loss may not be known by even solar service providers. Incomplete knowledge of these quantities contribute to the error in a PV system model.

2.6 Conclusion

Modeling the power output of a PV system is not simple. Even though most part of model is from well-grounded scientific theories and well-developed commercial software including PVsyst [62] and open source software including PV_LIB PV Performance Modeling collaborative (PVPMC) [61] have been made, it requires large number of factors. Some of them may be known by solar service providers but not grid operators. Even worse, some factors are not even known by solar service providers. The current PV system model is not practical in the view of grid operators since it is difficult to collect all factors for all PV systems in the grid. Since grid operators at the transmission level are only interested in the sum of power output from PV systems, we will build a data-driven model based on the power output data and well known factors instead.

Chapter 3

Applications of PV system modeling

Highlights:

- A structural approach in a PV system modeling has been used to address some questions on optimal layouts and orientations of PV systems, especially those of interest to PV owners.
- A simple example of an optimization problem of interest to PV owners is illustrated. The goal is to maximize energy generation in a year.
- Several examples of optimization problems of interest to grid operators are illustrated. The goals are to minimize a peak net load, to minimize a quadratic cost of net loads, and to minimize a quadratic cost of ramp rates from net loads.
- Solutions of interest to PV owners and grid operators are not necessarily the same. A feature that makes PV generation not appealing to grid operators is a high variability.

In previous chapters, we have seen the necessity to a reduced form approach to model power outputs from a group of PV systems in the form of maximum PV power curves. We also see some basic applications of maximum PV power curves. However, it does not mean a structural approach is not helpful for a grid operator. The structural approach has been used heavily by PV providers and PV owners to configure PV system layouts and orientations to achieve their goal, which is mainly maximizing PV energy generation. However, Utilities and grid operators have a different goal, which is minimizing a cost of serving a net load. This goal is not necessarily aligned with PV providers and owners.

Here we discuss several optimization problems structured from a structural approach of a PV system modeling. In the first section, we focus on problems of interest

to PV owners and PV providers. We show how to solve one of the problems in this category. In the second section, we focus on problems of interest to grid operators and Utilities. We solve a similar problem to the first section but with different objectives and see how optimal solutions differ.

3.1 Problems of interest to PV owners

In a simple term, PV owners and PV providers would like to generate as much energy from their PV systems as possible. Yet variables, constraints and objectives of optimization problems vary under different circumstances.

For variables, one can vary orientations including tilt angles and azimuth angles of panels to optimize some objective functions. If one consider more than a single panel, horizontal positions of panels can be another variable to consider, especially problems related to a piece of land with boundaries. To understand why positions of panels are important, imagine a problem of putting out PV panels on a very large piece of land. One can just lay all PV panels side by side on the ground without any tilt to maximize their energy generation. However, it is no longer obvious that such solution works for a piece of land with boundaries. One can tilt PV panels to get a similar amount of energy and let some shadow cast beyond the boundaries. However, one needs to be careful about shading from one panel to another. Horizontal positions of panels then becomes another variable to consider. Vertical positions of panels are also variables to consider as we put PV panels on the roof or high buildings. For example, putting panels at a high elevation may avoid shading from their surroundings.

For constraints, orientations and spacings can be limited especially in rooftop settings. In practice, one may rebuild a house roof to optimize tilt angles of panels but one may not renovate the whole house to optimize azimuth angles of panels. Locations of panels can be limited since some area may have some shading and some area may be reserved for other purposes including wiring materials and skylights.

For objectives, a simple objective of PV owners is to maximize a PV energy generation. However, the objective becomes more complex if we do a cost-benefit analysis. For example, one might add a cost of renovating roofs and a cost of maintenance into an objective function. An interaction with their house load and an electric rate structure can lead to a new objective function as well. Instead of maximizing a PV energy generation, one may minimize an electricity bill. This objective will be more appealing if the electric rate structure is dynamic and highly variable. As PV panels may replace other artifacts such as glass windows in a building, roofs in a parking lot, and even a surface on a road, we may need to convert a PV energy generation into a monetary value and add it into a bigger objective function, for example, minimizing the cost of a house renovation project.

With that said, an essential part in all optimization problems of interest to PV owners involves a calculation of PV power output that needs variables as discussed above in the calculation. Unlike a reduced form approach, a structural form approach allows us to describe the PV power output as a function of those variables, which will be a part of an optimization problem. Here we illustrate an example of how the structural approach of a PV system modeling can answer one of the simplest questions of interest to PV owners: given a single PV panel of a fixed size without any constraints on its orientation under a clear sky condition, what is the orientation such that the energy generation in a year is maximized?

In order to compute the power generation by a single PV panel under a clear sky condition, one can follow procedures as described in Chapter 2. First, we compute clear-sky solar irradiance and project it on a plane. This is where the information of the panel's orientation takes place. Then one can undergo an electronic process turning solar irradiance into DC and AC power output under the MPPT condition. An example of such calculation is shown in a PV_LIB PV documentation [73].

However, we cannot compute all steps without knowing a specification of a PV panel and the panel's temperature. We then assume that proportionality of irradiance values on a plane and the power outputs of a system to avoid complication from those unknown factors. As a result, an equivalent problem is to maximize the total insolation in a year on a single PV panel under a clear sky condition by varying an orientation of the panel.

As discussed earlier in the first section of Chapter 2, there are several choices of models for three components of irradiances: DNI, DHI and GHI. We decide to use a simple model called the Daneshyar-Paltridge-Proctor (DPP) model:

$$DNI = 950.2(1 - \exp(-0.075(90^\circ - z))) \tag{3.1a}$$

$$DHI = 14.29 + 21.04(\frac{\pi}{2} - z\frac{\pi}{180}) \tag{3.1b}$$

$$GHI = DNI\cos(z) + DHI, \tag{3.1c}$$

where $z = 90 -$ altitude is a zenith angle in degrees.

We compute a plane of array (POA) irradiance with three components.

$$I_{POA} = I_{beam} + I_{ground} + I_{diffuse}. \tag{3.2}$$

I_{beam}, irradiance from the direct normal irradiance, can be computed from

$$I_{beam} = DNI\cos(AOI), \tag{3.3}$$

where AOI is an angle of incidence. By trigonometry, the cosine value equals

$$\cos(AOI) = \cos(90 - \text{altitude})\cos(\text{array\_tilt}) + \sin(90 - \text{altitude})\sin(\text{array\_tilt})\cos(\text{azimuth} - \text{array\_azimuth}),$$

where 'altitude' and 'azimuth' here refer to the sun's altitude and azimuth angles. I_{ground}, irradiance from ground reflection, can be computed from

$$I_{ground} = GHI \cdot \text{albedo} \cdot \frac{1 - \cos(\text{array\_tilt})}{2}, \tag{3.4}$$

where albedo is a reflection property of the ground. To represent an urban area, we use albedo equal to 0.2. $I_{diffuse}$, diffused irradiance, can be easily computed if we select on an isotropic diffused irradiance. We use

$$I_{diffuse} = DHI \cdot \frac{1 + \cos(\text{array\_tilt})}{2}. \tag{3.5}$$

Up to this point, we can find solar irradiance and solar insolation in a year on a plane. As examples, we choose one geographic coordinate in California and one geographic coordinate in New Jersey. Figure 3.1 shows the insolation as a function of a panel orientation corresponding to those locations. The function, which is a large composition of functions, is too complicated to derive analytically. Yet the function are quite smooth with a single peak. Hence, we can find the optimal solution by a simple search. It turns out that for the location in CA, the optimal solution is (azimuth, tilt) = $(0 \pm 0.5, 31.3 \pm 0.1)$. For the location in NJ, the optimal solution is at (azimuth, tilt) = $(0 \pm 0.5, 37.1 \pm 0.1)$.

It is worthwhile comparing optimal insolations with insolations obtained from a typical heuristic: turn the panel facing south (array_azimuth = 0) with a tilt angle equal to the latitude of the panel's location. It turns out that for the location in CA, the insolation from the orientation according to the heuristic differs from the optimal insolation by 0.06 %. For the location in NJ, the insolation from the orientation according to the heuristic differs from the optimal insolation by 0.13 %. These results suggest that the rule of thumb gives an accurate way to assess the optimal orientation to maximize insolation in a year.

3.2 Problems of interest to grid operators

Here we explore a problem of minimizing a cost of serving net loads. At the surface, maximizing PV generation is minimizing net loads so the optimal solution of interest to grid operators should be the same. However, the cost of serving a net load is

(a) 'CA' data set

(b) 'NJ' data set

Figure 3.1: Maximum insolations in locations corresponding to our data set. The unit of the insolation is kilowatt-hours per year

not a simple linear function of a total net load consumption in a year. It turns out that the cost of serving a net load is complicated even though we ignore transmission constraints. Even at the level of a day-ahead planning, not a real-time operation where volatility plays a role in analysis, the cost associated with each daily load curve is a solution of an **economic dispatches** problem, where resources with different capacities, ramp rates, and costs are chosen to meet the daily load curve while keeping the total cost minimum. Resources with higher costs per generated power unit or per ramp rate unit will be chosen after resources with lower costs in the economics dispatches problem. Hence, the cost of serving a net load at each instance is a convex function of the load and the ramp rate at that instance where the derivative of a function is increasing.

Here we look at some simple problems that reflect some aspects of this cost of serving net loads. First, we try to minimize a peak net load of a year. Even though it sounds simplistic, it is one of the first goals for grid operators. We demystify the myth that the PV generation reduces a peak net load in this problem. Second, we try to minimize a quadratic cost function of instantaneous net loads. A quadratic function is chosen because it is a simple convex function where its derivative is increasing. Third, we try to minimize a quadratic cost function of instantaneous ramp rates of net loads.

3.2.1 Minimizing a peak net load

To illustrate our point, we consider the daily load curves of the whole grid for a year from the electric grid in California called CAISO and the electric grid in New Jersey called PJM in 2014. We observed that the original peak load is 43.48 GW in CAISO and 141.7 GW in PJM. We consider PV power curves generated by daily clear-sky irradiance curves generated by the structural approach described in the previous section. We assume that all PV systems have the same configurations operating at the STC. The differences between latitudes and longitudes within a geographical region of the grid are negligible. To keep the problem similar to the example in

(a) 'CA' data set

(b) 'NJ' data set

Figure 3.2: Net peak load in locations corresponding to our data set. The unit of the load is gigawatt

the previous section, we choose configuration angles as variables and keep installed capacities constant at 15% of the original peak load. This percentage is used in 2012 California's Rule 21 interconnection tariffs for distributed generations [20].

As in the previous problem, we can find an optimal solution by a simple search. Figure 3.2 shows the peak net load as a function of a panel orientation corresponding to the locations in CA and NJ. It turns out that an optimal solution is not unique. In addition, the optimal solution in the insolation optimization problem is also an optimal solution of this problem.

Note that the optimal net peak load differs from the original peak load by (43.48-41.5)/43.48 = 4.55 % in CAISO and (141.7-140.5)/141.7 = 0.8 % in PJM. These numbers are small given that we choose an installed PV capacity in the grid to be 15 % of the original peak load. Hence, there is nothing much one can do with configurations of PV panels to bring down the net peak load. This is because the occurrence of peak PV generation differs significantly from the peak load, which occurs in the evening. That is, the myth that PV generation reduces a peak load is not true in general.

3.2.2 Minimizing a quadratic function of net loads

Here we use the same set up as in the problem of minimizing a peak net load. However, the objective function is now a quadratic function of net loads. To be precise, we minimize

$$\sum_{i=0}^{N} p_i^2,$$

where p_i is an instantaneous net load and a time index i runs for all instances separated by 15 minutes interval starting from January 1st, 2014 00:00 to December 31st, 2014 23:45.

As in the previous problem, we can find the optimal solution by a simple search. Figure 3.3 shows the quadratic function of net loads as a function of a panel orientation

(a) 'CA' data set

(b) 'NJ' data set

Figure 3.3: A quadratic function of net loads in locations corresponding to our data set.

corresponding to the areas in CA and NJ. It turns out that for the location in CA, the optimal solution is (azimuth, tilt) $= (-9.5 \pm 0.5, 29.1 \pm 0.1)$. For the location in NJ, the optimal solution is at (azimuth, tilt) $= (-5.5 \pm 0.5, 36.9 \pm 0.1)$.

These optimal solutions have similar trends: the azimuth angles are slightly negative (facing east) and the tilt angles are slightly lower than the optimal tilt angles in the problem of maximizing insolation in a year. Note that the insolation here differs from the optimal insolation by 1-2413/2420 = 0.3 % in CA and 1-2351/2354 = 0.1 % in NJ.

3.2.3 Minimizing a quadratic function of ramp rates

Here we use the same set up as in the problem of minimizing a peak net load. However, the objective function is now a quadratic function of *ramp rates* of net loads. To be precise, we minimize

$$\sum_{i=1}^{N} (p_i - p_{i-1})^2,$$

where p_i is an instantaneous net load and a time index i runs for all instances separated by 15 minutes interval starting from January 1st, 2014 00:00 to December 31st, 2014 23:45.

As in the previous problem, we can find the optimal solution by a simple search. Figure 3.4 shows the quadratic function of ramp rates of net loads as a function of a panel orientation corresponding to the areas in CA and NJ. It turns out that for the location in CA, the optimal solution is (azimuth, tilt) $= (-178 \pm 0.5, 36.8 \pm 0.1)$. For the location in NJ, the optimal solution is at (azimuth, tilt) $= (180 \pm 0.5, 33.8 \pm 0.1)$.

These optimal solutions have a similar trend: the azimuth angles are close to ± 180 degrees (facing north). Note that the insolation here differs from the optimal insolation by 1-1148/2420 = 52.6 % in CA and 1-1050/2344 = 55.2 % in NJ.

(a) 'CA' data set

(b) 'NJ' data set

Figure 3.4: A quadratic function of ramp rates in locations corresponding to our data set.

Scenarios for CA	Azimuth	Tilt	% insolation difference
Max insolation	0	31.3	0
Heuristic	0	33.6	0.06
Min peak net load	0	31.3	0
Min quadratic net load	-9.5	29.1	0.3
Min quadratic ramp rate	-178	36.8	52.6
Scenarios for NJ	Azimuth	Tilt	% insolation difference
Max insolation	0	37.1	0
Heuristic	0	40.0	0.13
Min peak net load	0	37.1	0
Min quadratic net load	-5.5	36.9	0.1
Min quadratic ramp rate	180	33.8	55.2

Table 3.1: Comparison of orientations and difference relative to the maximum insolation from different scenarios in 2014

3.2.4 Comments

Table 3.1 summarizes tilt angles, azimuths angles, and differences relative to the maximum insolation from different scenarios.

From the table, one can see that the solutions of problems of interest to grid operators are close to the solution of a problem of interest to PV owners. The heuristic that a tilt angle of a solar panel should be equal to the latitude of its location and that a solar panel should face south works reasonably well. Note that there is no need to optimize the peak net load since the peak generation does not align the original peak load well. In order to optimize the quadratic cost of net loads, facing the panel slightly to the east is beneficial. Even though the load is generally higher in the evening, PV generation with slightly east facing helps reducing the cost of serving the net load during a smaller peak load in the morning.

The only problem that gives a solution against the interest to PV owners is the problem of minimizing the quadratic cost of ramp rates. Facing directions of PV panels are opposite and the total insolation in a year differ by a half. It means that PV generation severely increases ramp rates of net loads. Hence, the optimal solution of minimizing the quadratic cost of ramp rates is to avoid high PV generation. This shows a tension of adding PV generation into the grid even if we have not introduced the volatility in the discussion. While PV generation reduces the cost of net loads, it increases the cost of ramp rates due to its variability.

To be more realistic, for a grid operator's perspective, one should optimize a linear combination of both quadratic functions: net loads and its ramp rates. In general, we can look into a broader class of convex functions $f(p_i, |p_i - p_{i-1}|)$. In order to specify parameters and choices of functions, one need to have substantial amount of daily load curves and the associated cost of operation. Then one can use regression and construct a more reasonable cost function as an objective of a grid operator. Such task is left as a future work.

3.3 Conclusion

In this chapter, we illustrate some basic applications of a structural approach in a PV system modeling. As we have adjustable physical variables such as orientations and positions of PV panels, we may use a structural approach to derive optimization problems of interest to PV owners or the grid operators. From simple examples, we see that solutions of different interests may not be aligned. One can refine and extend the optimization problems to compromise benefits of both parties.

Here we conclude our discussion on a structural approach in a PV system modeling. In the next chapter we use the maximum PV power curve from the reduced form approach to build a short-term PV power forecast system.

Chapter 4

Maximum PV power curves

Highlights:

- A maximum power curve for a group of PV systems informs the grid operator of the maximum generated power at each time instance. It can be used to assess the maximum PV generation and the variability of a net load curve.
- Two approaches to find a maximum power curve are the structural approach and the reduced form approach. The reduced form approach is more practical and scalable. The concept that makes our model of a maximum power curve robust and scalable is a normalized maximum power curve. The normalized maximum power curve is essentially an envelope of a normalized power curve generated from actual power output measurements.
- To ensure that the definition of a normalized maximum power curve is consistent, we perform consistency checks under different circumstances for a normalized power curve, which is a precursor of a normalized maximum power curve.
- We illustrate how to construct maximum normalized power curves from normalized power curves using selection techniques and regressions.
- We show that the resultant maximum normalized power curves are consistent.

4.1 A maximum power curve for a group of PV systems

The maximum power curve for a group of PV systems is a graphical illustration of the variation in maximum power generation from the group to the grid. From a

grid operator's perspective, the maximum power curve for a group of PV systems in a substation can be used in a number of ways. The grid operator can estimate the maximum PV installation such that the generating power of a group of PV systems at any time does not exceed a certain limit. The grid operator can also assess variability in the PV generation from the maximum power curve.

The grid operator may construct the maximum power curve for a group of PV systems consisting of panels, an inverter, and other electronics to the grid if their factors is known. As seen in Chapter xx, the power output from a PV system can be modeled based on those factors. One may start with a clear-sky solar irradiance model, followed by an irradiance projection on a plane, a conversion to a DC power output, and a conversion to an AC power output to the grid. The result of using a clear-sky solar irradiance model over time yields a maximum power curve for a PV system. The sum of such maximum power curves represents the maximum power curve for a group of PV systems. We call this the **structural** approach. The problem with the structural approach is that it requires a large number of factors to be identified. As the PV generation grows in a distributed fashion, the grid operator is likely to have difficulty in monitoring all factors for each PV system in the grid.

An alternative approach is to construct a maximum power curve for a group of PV systems. We aggregate power outputs from each PV system and build the maximum power curve by selecting several curves with a high power generation and performing regression. We call this a **reduced-form** approach. A grid operator may track power outputs for each PV system with help from solar PV providers or monitoring companies. Then the grid operator can use the reduced form approach without worrying about numerous unknown factors. A comparison between the structural approach and the reduced-form approach of the maximum power curve construction is illustrated in Figure 4.1.

Scalability by the capacity of the PV installation is an essential feature of a maximum power curve model. The current PV generation capacity is low compared to both the total load and the potential for new PV installations. Consequently, it is possible that the maximum power curve for a current group of current PV systems may not be applicable to a system with high PV penetration. However, comparing the power curve for increasing subsets of current installations, we show that the power curve quickly converges. This establishes the consistency in defining a normalized power curve. With this normalized power curve, we can scale it in order to estimate future limitations on the capacity of PV installations.

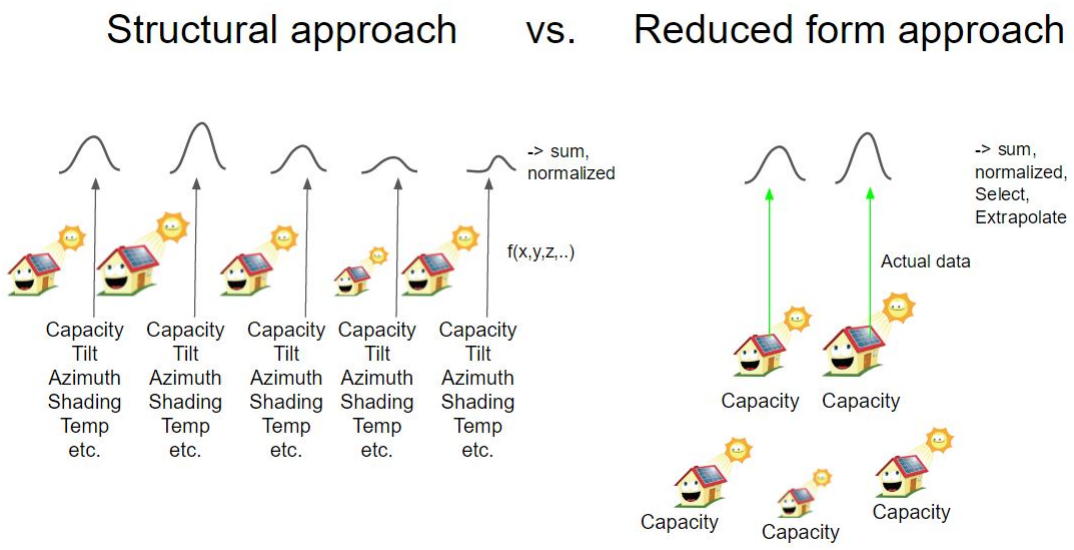


Figure 4.1: Comparison between the structural approach and the reduced-form approach of a maximum power curve construction

4.2 Issues with normalized maximum power curves for a group of PV systems

The normalized maximum power curve from the reduced form approach is attractive since it requires only power output data from a sample of PV systems. It also requires only the total installed capacity to scale for a bigger group of PV systems in the same geographical region. Some issues to consider when constructing the normalized maximum power curve are addressed below:

- The idea of linear scalability relies on the assumption that the power generation from a group of distributed PV systems is proportional to its total installed capacity. Such assumption must be justified. We justify the assumption in Section 4.5

- The precursor of a normalized maximum power curve is a normalized power curve derived from actual power outputs from an existing group of PV systems. We need to ensure that the normalized power curve is consistent as we apply our model to predict what will happen in the future. To be precise, we would like to ensure that data from different installation years gives consistent normalized power curves. Furthermore, we would like to ensure that data from different

installation dates give consistent normalized power curves. We perform these tasks in Section 4.6

- A number of PV systems in a group required to produce a consistent normalized power curve is not immediately known. This issue is therefore addressed in Section 4.7.

- As the actual power output is not always at the maximum, it is not straightforward to find a normalized maximum power curve from a normalized power curve. Simple smoothing and regression do not work. Therefore we perform methods of selection, regression, and adjustment on normalized power curves to obtain maximum power curves as described in Section 4.9.

- To validate our approach, we checked the consistency of normalized maximum power curves generated from two random samples of power outputs of PV systems in a same geographical area. This is described in Section 4.10.

4.3 Definition of consistency measurement

In this section, we examine the concept of consistency. An issue that needs to be addressed is how to measure consistency between normalized power curves. First, we can focus at each time instance and compare the power readings from different curves. Then we must determine a statistic of all time instances in a time interval including the mean of an average absolute deviation and the maximum of an average absolute deviation, where the average absolute deviation at a time instance from C normalized power curves is defined as

$$\sum_{j=1}^{C} |p_j - \frac{p_1 + ...p_C}{C}|.$$

If $C = 2$, the average absolute deviation equals to one half of the average absolute difference $|p_1 - p_2|$. Since normalized power curves have values of the order of 1.00, we may concretely define that normalized power curves are consistent if the mean of an average absolute deviation in power outputs is less than 0.05.

The current measure has a pitfall when we want a measure of consistency over a long period of time, for example, the consistency of normalized power curves for different years. Given that weather patterns from various years may differ substantially, the previous measure on power readings may give a large value even though the normalized power curves are consistent. Hence, we propose a second way to measure consistency between normalized power curves. One can focus on the total energy

generation from different curves. Then we report the percent of the average absolute deviation in an energy generation over a time period, which is defined as

$$\frac{\sum_{j=1}^{C} |E_j - \frac{E_1+...E_C}{C}|}{\frac{E_1+...E_C}{C}} \times 100\%.$$

Note that an energy generation E refers to the area under the normalized power curve. The concept of a percentage is needed for an easy comparison because the energy generation depends on the length of the time interval. We may define that normalized power curves are consistent if the percentage of the average absolute deviation in an energy generation is less than 5%.

4.4 PV output data

4.4.1 Data set

The data set for building a normalized maximum power curve model for a group of PV systems has two components: the installed capacity of each PV system in a group and the actual power outputs from each PV system in the group. We received data from 2 groups of PV systems collected by a solar PV monitoring company. The locations of PV systems in each group are from the same zip code. The grid operator's perspective may concern a group of PV systems under the same substation rather than the same zip code. Still we assume the same characteristics. One data set is collected from PV systems in one particular zip code in CA. Another data set is collected from PV systems in one particular zip code in NJ.

Each data set contains two parts:

1. `metadata` consists of information including the installed capacity of each PV system. Each row of the metadata contains information about each PV system including its ID, latitude, longitude, zip code, timezone, and installed capacity in kilowatts (kW). Since the installed capacity is confidential information, it is reported as a range of 0-1 kW, 1-2 kW, 2-3 kW, 3-5 kW, 5-10 kW, 10-20 kW, 20-50 kW, 50-100 kW, or `NaN`. This makes normalization of power curves difficult. We address how to estimate the installed capacity of each PV system in the data set in Subsection 4.4.3.

2. `timeseriesdata` consists of time series of power outputs from each PV system. Each time series contains information about an instance of power reading for a PV system at a given Universal Time Coordinated (UTC) time with a frequency of 15 minutes. Each system appearing in `metadata` corresponds with a time series in `timeseriesdata`. As the data collection has a 15-minute frequency, we expect to have a continuous stream of data every 15 minutes starting from the installation date. However, this may not be the case as the monitoring system is sometimes inactive.

The time series of power output from each PV system starts on either January 1st, 2013 or the PV system's installation date. All these time series end on December 31st, 2014.

4.4.2 Data cleaning

In order to process data and construct normalized maximum power curves correctly in an organized manner, we clean data as follows:

1. We remove data from systems with an installed capacity greater than 20 kW as we are only interested in residential-size PV systems.

2. We convert the UTC time stamps in `timeseriesdata` to be local time stamps.

3. Since time stamps in raw data may not end with 00:00, 15:00, 30:00, or 45:00 minutes, we adjust time stamps so that all time stamps end with these increments. We do such rounding in order to ensure there are enough power readings at each time stamp.

4. For each PV system, we record the first time stamp that appears in `timeseriesdata` in `metadata` as the first monitoring time of a system.

5. In `timeseriesdata`, we only select entries registered after sun rises and before sun sets. We bundle multiple time series into a single table where each row represents a time stamp from January 1st, 2013 to December 31st, 2014 with a 15-minute frequency, and each column represents each PV system. Since we have identified the first monitoring time of each PV system, we are able to sort columns based on this value. We assign `NaN` to cells with no power output.

At this point, it may appear that we can generate normalized power curves for a group of PV systems. The value of normalized power curves at each time stamp is simply the row sum of `timeseriesdata` divided by the sum of the installed capacities of PV systems that have power outputs at that time stamp. The difficulty in this calculation is that we do not know the exact value of the installed capacity of each component. In the next subsection, we show how to calculate an estimate for the installed capacity.

4.4.3 Installed capacity estimates

As addressed above, `metadata` does not contain the exact value of an installed capacity of each PV system due to a confidentiality constraint. Instead, the installed capacity of a PV system is given as a range. If we assume that the installed capacity

Range	Midpoint	CA estimate	NJ estimate
0-1	0.5	0.64	0.60
1-2	1.5	1.56	1.55
2-3	2.5	2.49	2.52
3-5	4	3.96	4.10
5-10	7.5	6.76	7.49
10-20	15	12.81	12.94

Table 4.1: Ranges, midpoints, and estimates for the installed capacity for bins as appeared in `metadata`. All values are in kW.

is uniformly distributed in a range, then we can estimate any installed capacities in the range to be equal to the midpoint of the range. For example, each PV system with a 10-20 kW range has an installed capacity of $(10+20)/2 = 15$ kW.

In order to find a better estimate of the installed capacity, we may look into other sources of data where information about installed capacities of PV systems is available. NREL's Open PV project [46] has a data set of installed capacities of PV systems over time and region. To obtain large enough yet relevant samples, we selected data of PV systems in each state of interest from 2008 to 2012, which is the most recent year of data. Histograms of installed capacities up to 20 kW in CA and NJ are shown below.

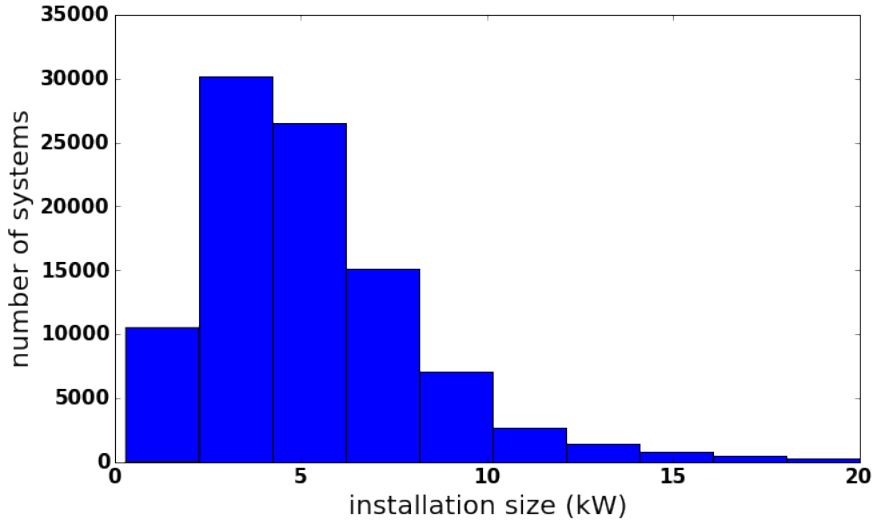
It is worth noting that the installed capacities of distributed PV systems in CA are smaller than the installed capacities in NJ. This is likely due to the fact that the weather in CA is more favorable to the use of PV systems than in NJ. As a result, a PV system with a smaller installed capacity is sufficient in CA.

Next we divide the queried data into bins corresponding to ranges in `metadata` and compute the mean of the installed capacity in each bin. The mean value will be used as an estimate for the installed capacity in our analysis. Ranges, midpoints of the ranges, and estimates of the installed capacity for each state are summarized in Table 4.1.

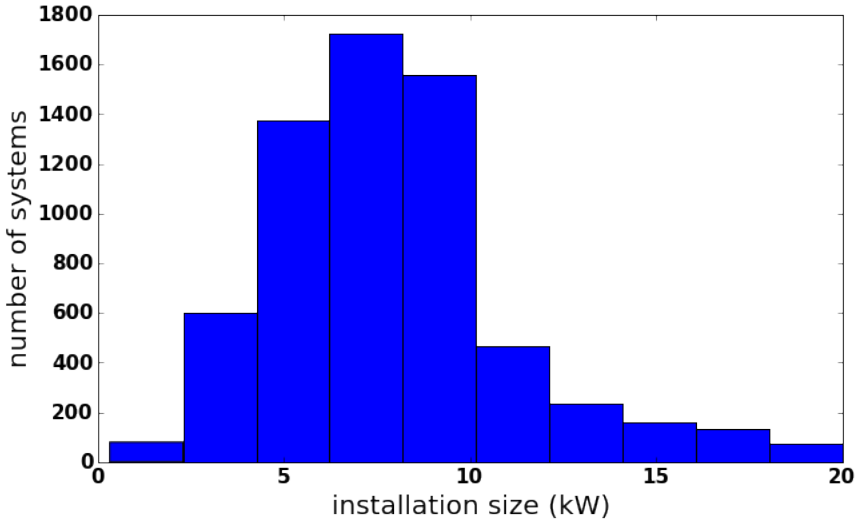
With this estimate, we can compute normalized power curves for a group of PV systems. In the next section we see that this estimate is at least as compatible as a midpoint when we test the assumption on linear scalability.

4.5 A test of a linear scalability assumption

We solve the first issue regarding the assumption of linear scalability. We illustrate by our data that the power generated by a group of PV systems is proportional to



(a) CA



(b) NJ

Figure 4.2: Histograms of installed capacities up to 20 kW in two states in 2008-2012

its total installed capacity. From `timeseriesdata`, we construct a new table for PV systems with a size of 5-10 kW and another table for PV systems with a size of 10-20 kW. For each table, one can compute a time series of the mean power by the number of components. If the assumption on linear scalability is true, when we plot values of time series from the two tables against each other, we expect to see a proportional relationship in the graph. Note that such a proportional relationship occurs when there is a sufficient number of components in each table. For the data set in CA, we plot a data point for each instance only when there are more than 10 power readings in each table. For the data set in NJ, we plot a data point for each instance only when there are more than 100 power readings in each table. The required number for the data set in CA is fairly low due to small number of systems with size of 10-20 kW in the data set. Still, we see proportional relationships from both data sets in CA (Figure 4.3a) and NJ (Figure 4.3b).

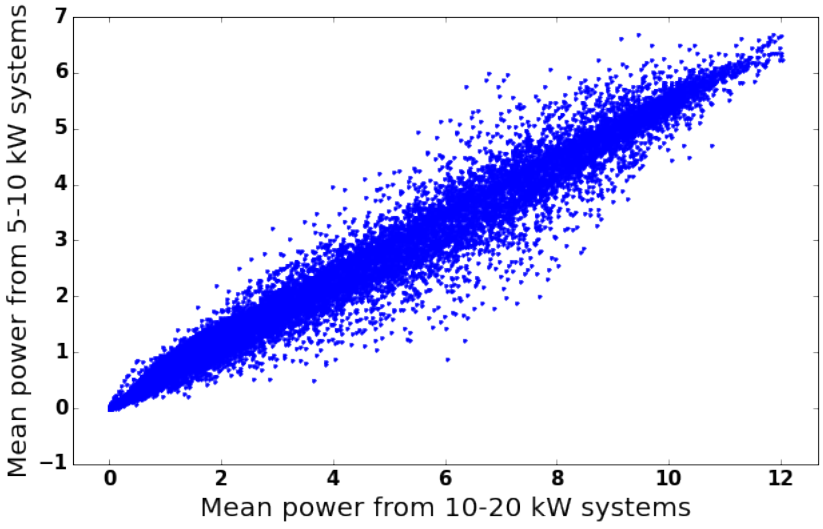
The plot of the data set in CA demonstrates that the y-intercept from a linear regression is equal to -0.002, in other words, very close to zero. The slope is equal to 0.61. The ratio of our installed capacity estimates for PV systems in a 5-10 kW bin and a 10-20 kW bin is equal to $6.76/12.81 = 0.53$, while the ratio of midpoints between two bins is equal to 0.5.

The plot of the data set in NJ shows that the y-intercept from a linear regression is equal to 0.07, in other words, very close to zero. The slope is equal to 0.54. The ratio of our installed capacity estimates for PV systems in a 5-10 kW bin and a 10-20 kW bin is equal to $7.49/12.94 = 0.58$, while the ratio of midpoints between these two bins is equal to 0.5.

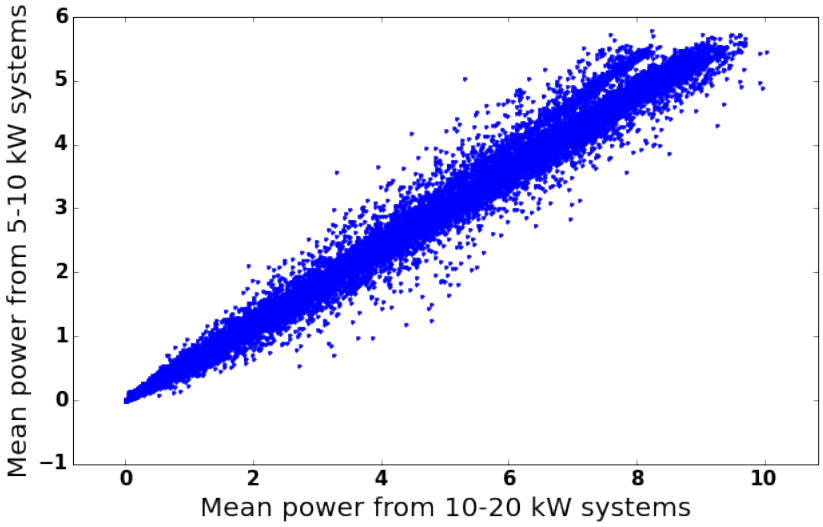
These results confirm that our assumption is reasonable. In addition, they also imply that our installed capacity estimates in each bin are the same or even more accurate relative to the midpoints.

4.6 Consistency checks for normalized power curves

As addressed in the list of issues with normalized maximum power curves for a group of PV systems, we first need consistent normalized power curves. We need to ensure that the normalized power curves are consistent as we apply our model to indicate what will happen in the future. Two questions must be answered in order to ensure that our model will work for the future, at least for a following few years. First, if data are collected from different years, are the normalized power curves still consistent? Second, if data are from PV systems with different installation dates, are the normalized power curves still consistent? There are two factors to consider: the year of data collection and the installation date of PV systems.

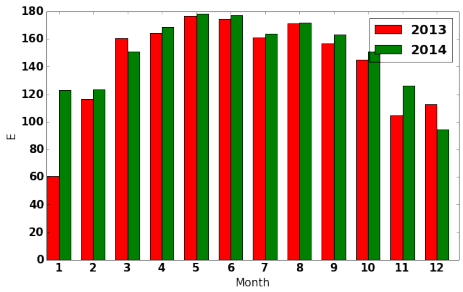


(a) CA

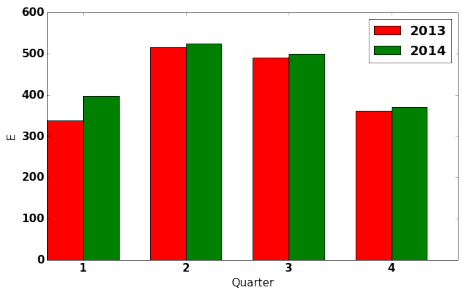


(b) NJ

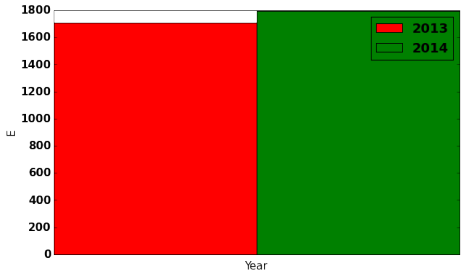
Figure 4.3: Scatter plots of the mean power by PV systems from 2 bins: 10-20 kW and 5-10 kW



(a) Monthly



(b) Quarterly



(c) Yearly

Figure 4.4: Energy generation from normalized power curves from the data set in CA for 2013 and 2014

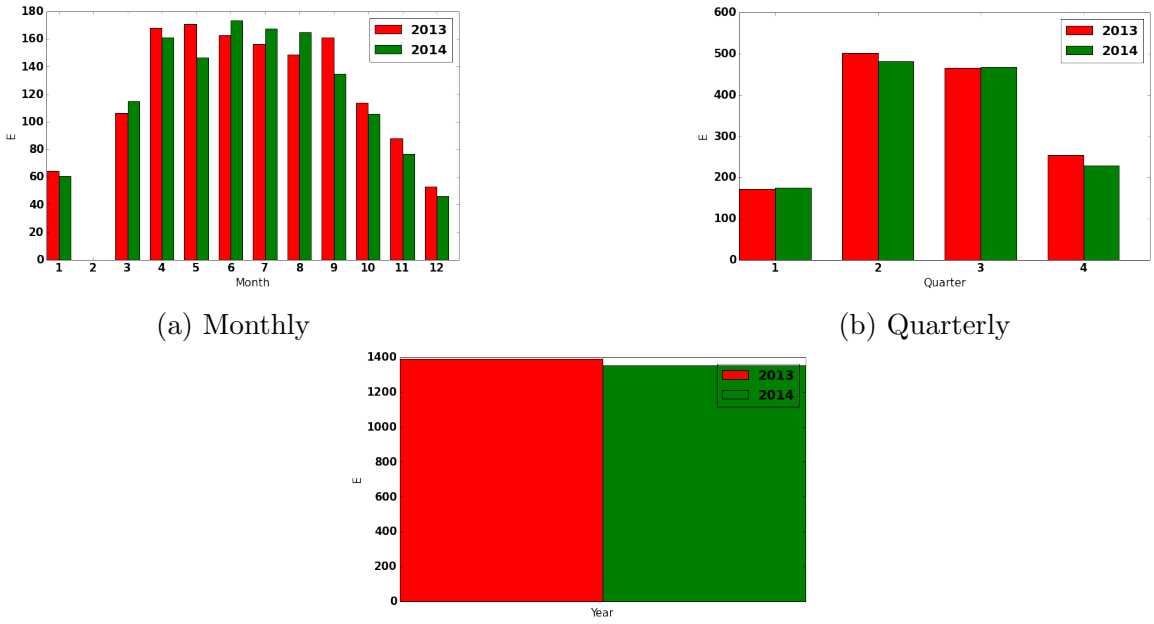
4.6.1 Consistency of data from different years

In order to investigate consistency of data from different years, we construct normalized power curves for different years. In order to control another factor, we build all normalized power curves from the same group of PV systems. In our data set, we constructed normalized power curves for 2013 and 2014 from the PV systems with the first monitoring dates before January 1st, 2013.

When we compare normalized power curves from 2013 and 2014, we immediately see that they are different since the weather varies from year to year. Hence the mean of the average absolute deviation in power is not a good measure for the consistency in this situation. The better measure is the percentage of an average absolute deviation in energy generation. First, we compare monthly energy generations, quarterly generations, and yearly energy generations between 2 years as shown in Figure 4.4 and 4.5.

Note: we reject daily normalized power curves in this analysis if some readings of power outputs on the curves are missing.

For the data set in CA, we build normalized power curves and calculate energy generations from 2 systems that pass our condition. From the data set in CA shown in Figure 4.4, one can calculate that the average absolute deviation in monthly energy



(a) Monthly

(b) Quarterly

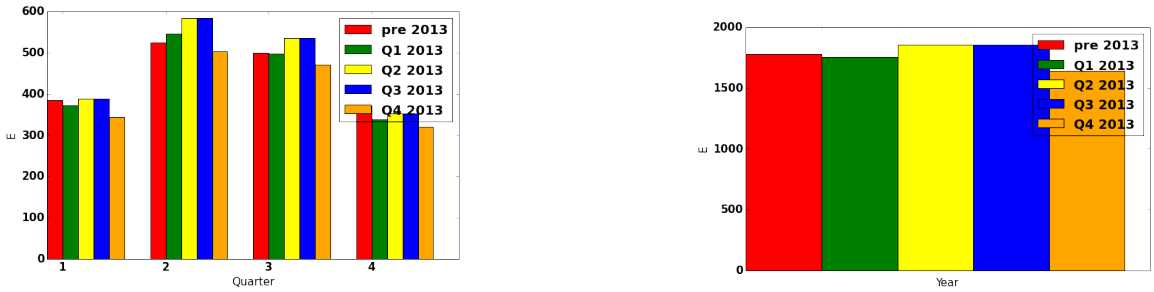
(c) Yearly

Figure 4.5: Energy generation from normalized power curves from the data set in NJ for 2013 and 2014

generations in 2013 and 2014 can be as large as 34 percent in January. The average absolute deviation in quarterly energy generations can be as large as 8 percent in the first quarter. The average absolute deviation in yearly energy generations is about 2 percent.

For the data set in NJ, we build normalized power curves and calculate energy generations from 5 systems that pass our condition. From the data set in NJ as shown in Figure 4.5, one can calculate that the average absolute deviation in monthly energy generations in 2013 and 2014 can be as large as 9 percent in September. The average absolute deviation in quarterly energy generations can be as large as 2 percent in the second quarter. The average absolute deviation in yearly energy generations is about 1 percent.

It is worth noting that, even though the number of systems in this investigation is relatively small, the deviation of energy generation from year to year is low from the perspective of yearly generation. From two independent data sets, we see that there is no particular direction of change from one year to another. Hence, we have consistency over different years of data collection. This result also implies that the degradation of PV systems is negligible when considered year by year.



(a) Quarterly (b) Yearly

Figure 4.6: Energy generation from normalized power curves from the data set in CA for 5 groups of PV systems categorized by installation dates

4.6.2 Consistency of data from different installation dates

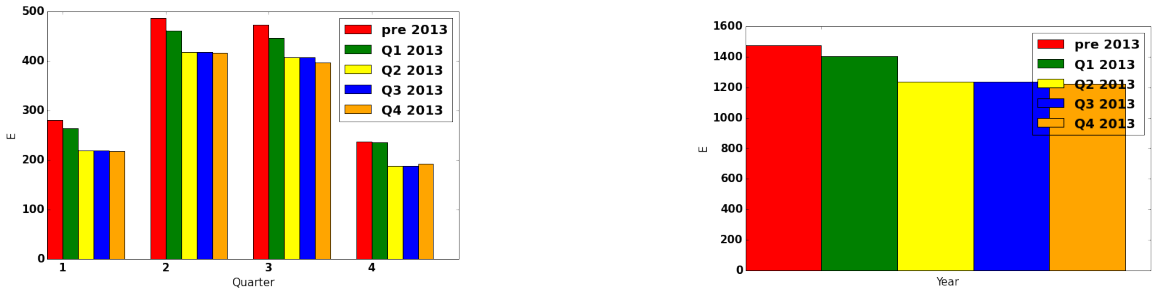
In order to investigate consistency of data from different installation dates, we build and compare normalized power curves from PV systems with different installation dates. In order to control the date of data collection, we build all normalized power curves from data gathered during the same time period. In our data set, we can build and compare normalized power curves for 2014 from 5 groups of PV systems based on their first monitoring dates: before 2013, the first quarter of 2013 (Q1 2013), Q2 2013, Q3 2013 and Q4 2013.

As in the previous subsection, we compare energy generations to verify the consistency. To avoid overwhelming visual representation of the data, we consider only quarterly and yearly energy generations.

For the data set in CA, we build normalized power curves and calculate energy generations from 5 groups categorized by installation dates (pre-2013: 2 components, Q1 2013: 3 components, Q2 2013: 4 components, Q3 2013: 8 components, Q4: 65 components. From the data set in CA shown in Figure 4.6, one can calculate that the average absolute deviation in quarterly energy generations can be as large as 5 percent in the second quarter. The average absolute deviation in yearly energy generations is about 4 percent.

For the data set in NJ, we build normalized power curves and calculate energy generations from 5 groups categorized by installation dates (pre-2013: 5 components, Q1 2013: 40 components, Q2 2013: 11 components, Q3 2013: 81 components, Q4: 99 components. From the data set in NJ as shown in Figure 4.7, one can calculate that the average absolute deviation in quarterly energy generations can be as large as 11 percent in the fourth quarter. The average absolute deviation in yearly energy generations is about 8 percent.

It is worth noting that the deviation of energy generation from PV systems with different installation dates is quite high. This result might be influenced by the low



(a) Quarterly (b) Yearly

Figure 4.7: Energy generation from normalized power curves from the data set in NJ for 5 groups of PV systems categorized by installation dates

number of components. While the data set in CA does not suggest any particular trend, the data set in NJ reveals a downward trend of energy generation. Both the concept of learning-by-doing and the process of deterioration of a PV system suggest that more recently installed PV systems should have higher energy generation per an installed capacity over time. On the other hand, one might argue that PV should be installed in a place with a higher yield first and a place with a lower yield later. This concept known as "taking low hanging fruit first" suggests that more recently installed PV systems should have lower energy generation per an installed capacity over time. The result from the data set in NJ provides evidence for the last argument.

Due to the low number of components in the data set, it makes sense to combine data with different years of data collection and different PV system installation dates. As a consequence, we need to accept that the normalized power curves generated by such data may exhibit an inherited deviation. In the future, if data over several years is collected and there is a higher number of PV systems at various time periods, we should re-investigate this issue.

4.7 The number of PV systems needed for normalized power curves to be consistent

One of the issues from our model for maximum power curves is about a sufficient number of PV systems to produce consistent normalized power curves. If too few PV systems are used to construct a normalized power curve, the normalized power curve will be not be a good representative of current and future collections of PV systems in a region. We need to find a minimum number of PV systems to produce a consistent normalized power curve within the average absolute deviation in yearly energy

generation of one percent. Note that such the minimum number is not necessarily universal for all regions.

From each data set, we sample 10 independent subsets of data from a chosen number of PV systems. We construct normalized power curves for each subset and calculate the yearly energy generation for each power curve. We then measure the average absolute deviation in yearly energy generations of those 10 subsets. As we increase the chosen number of PV systems, we expect to see a smaller deviation. In order to control this experiment, we build all normalized power curves from data collected in the same time period and sampled from systems installed before the starting point of the time frame. In our data set, we select that time period to be the entire year of 2014.

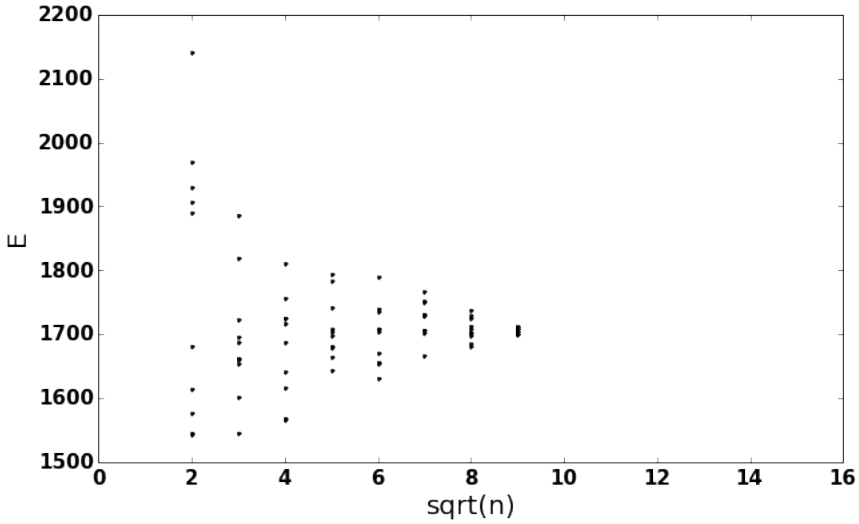
For the data set in CA, 82 PV systems pass the criteria. For the data set in NJ, 237 PV systems pass our criteria. Figure 4.8 shows yearly energy generations of 10 samples for each number of PV systems. It is clear that as the number of PV systems used to generate a normalized power curve increases, the difference of yearly energy generations among samples decreases.

Table 7.1 shows the percentage of the average absolute deviation in yearly energy generations from a different number of PV systems. We need only 49 systems to achieve the deviation of less than 1 percent for the data set in CA while we need 100 systems to obtain the deviation of less than 1 percent for the data set in NJ. The data set in NJ requires a larger minimum number of systems for two reasons. First, PV systems in the data set in CA may be more homogeneous than PV systems in the data in NJ. To achieve the same level of accuracy, we need a higher number of PV systems to produce consistent normalized power curves from a less homogeneous area. Second, the weather associated with the data set in CA may be more volatile than the weather associated with the data set in NJ set. To achieve the same degree of accuracy, we require a higher number of PV systems to produce consistent normalized power curves for more volatile weather.

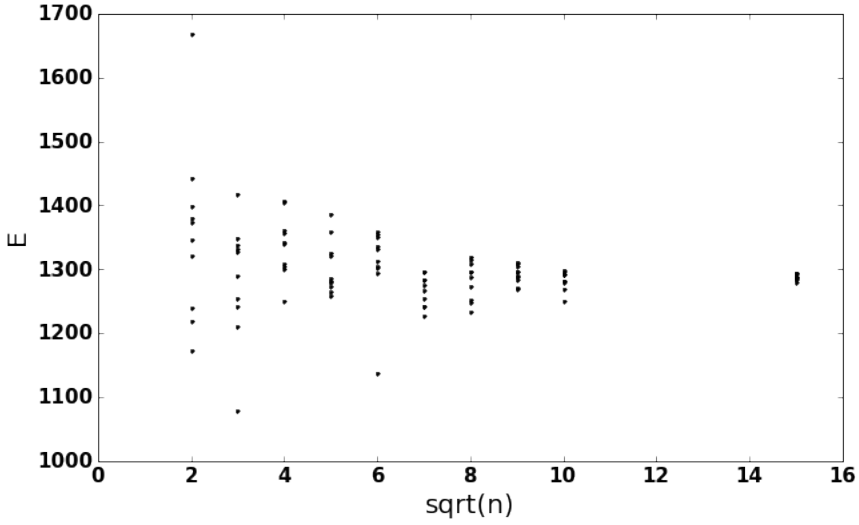
4.8 Heuristics to produce consistent normalized power curves

From previous sections, we have seen how to construct normalized power curves of a group of PV systems. We have analyzed several issues that might affect the consistency of normalized power curves. We now determine the heuristics to produce consistent normalized power curves as follows:

1. Clean the data so that time series of power outputs from PV systems are aligned. Select a part of the time series associated with daylight times only. Make sure that the installed capacity of each PV system is known or well estimated.



(a) CA



(b) NJ

Figure 4.8: Yearly energy generation derived from a different number of PV systems. Each number has 10 samples. The horizontal axis shows the square root of the number of PV systems.

The number of components	CA	NJ
4	5.3	6.1
9	4.2	7.5
16	3.9	1.8
25	2.4	3.1
36	2.0	1.7
49	0.8	1.5
64	0.9	1.3
81	0.2	1.3
100	-	0.9
225	-	0.3

Table 4.2: The percentages of the average absolute deviation in yearly energy generation from different numbers of PV systems.

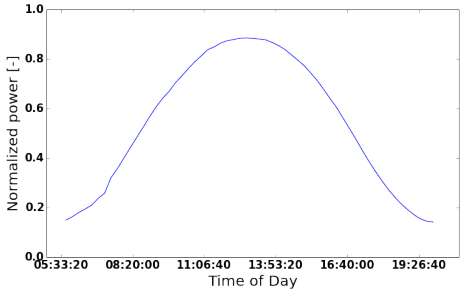
2. Determine a point on a normalized power curve by summing all power outputs and dividing them by the sum of installed capacities associated with the power readings at that instance. As a rule of thumb, the number of PV systems should be at least 50 to achieve the possible deviation of about 2 %.

3. Check if there is a need to distinguish data from different years of collection or from PV systems with various installation dates. However, due to a limited amount of data, we may treat them equally and accept a possible deviation of about 8 %.

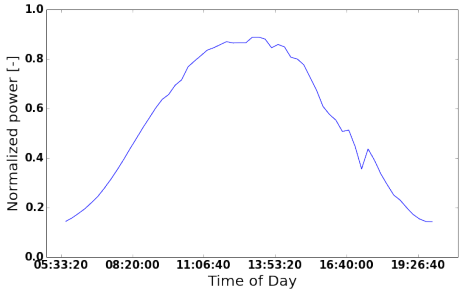
4.9 Maximum normalized power curve construction

We construct maximum normalized power curves from normalized power curves derived from power outputs from a group of PV systems. To get an insight of how to construct them, we consider several daily normalized power curves generated from the data set in CA and the data set in NJ using heuristics described in the previous section. Some examples are shown in Figure 4.9 and Figure 4.10.

After examining a number of daily normalized power curves, we found that for most of the days in a year, a normalized power curve is not likely a maximum normalized curve since there are some spikes on the curve. However, on some days, a normalized power curve resembles a simple bell shape with a high daily energy generation relative to neighboring days. The simple bell shape may not be generally true

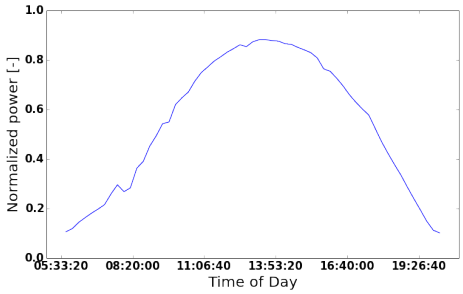


(a) Well-behaved normalized power curve, on 2014/6/24

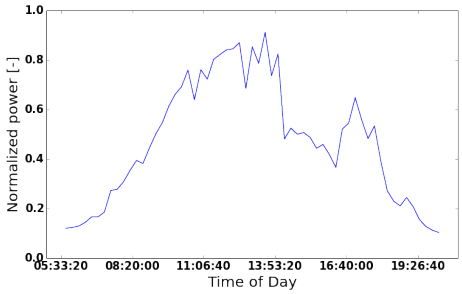


(b) Typical normalized power curve, on 2014/6/21

Figure 4.9: Daily normalized power curves generated from the data set in CA



(a) Well-behaved normalized power curve, on 2014/6/24



(b) Typical normalized power curve, on 2014/6/21

Figure 4.10: Daily normalized power curves generated from the data set in NJ

for power curves from a single PV system, but it is common for power curves from a large number of PV systems. Note that we are interested only in power outputs between sun rise and sun set. Hence we make an assumption:

Assumption: A daily maximum normalized power curve resembles a simple bell-shape curve with a high daily energy generation relative to neighboring days. The domain of a maximum normalized power curve should be a time interval between sun rise and sun set. The range of a maximum normalized power curve should be non-negative.

We call actual power curves that resembles daily maximum normalized power curves as **well-behaved normalized power curves**. A candidate function that fits the assumption of the daily maximum normalized power curve is a linear combination of B-splines. If all normalized power curves constructed from the data are maximum normalized power curves, we can use a regression to find such a linear combination of B-splines for all the days in a year. The problem is that most normalized power curves are not ideal.

Still, there is a way to construct maximum normalized power curves for all the days in a year. First, we select some well-behaved daily normalized power curves, which are likely to be maximum normalized power curves. Then we may use a regression to find a linear combination of B-splines that fit a well-behaved daily normalized power curve. Once we obtain regression coefficients for well-behaved days, we can extrapolate regression coefficients to neighboring days as well. With this process, we have linear combinations of B-splines as maximum normalized power curves for all the days. Lastly, we may need to scale or shift B-splines to ensure that the resulting linear combinations are indeed higher than normalized power curves from the data.

Note: One may model a normalized power curve before sun rise and after sun set, then augment it with our model that focuses on power readings between sun rise and sun set. However, we do not consider such an issue in this study.

Before we proceed to the selection process of well-behaved daily normalized power curves, a regression of B-splines, and an extension of regression coefficients. There is a problem of generalization under different domains. As the season changes, the sun rises and sets progressive different times. As a result, B-splines generated from different days are not compatible to each other. This makes the extrapolation of regression coefficients impossible unless we transform the domain of a PV power curve and find a systematic way to bridge models between various days.

4.9.1 Transformation of the domain of a daily power curve

In order to make a model of daily maximum power curves generalizable, transformation of its domain is necessary. From sun rise and sun set each day, we need a domain common to all days. We define s as a negative cosine of an angle on a facial plane

of an observer facing the south. For each time instance between a sun rise and a sun set, one may find a Cartesian coordinate of the sun on the unit sphere. One may then find an associated s using trigonometry:

$$s = \frac{y}{y^2 + z^2}, \tag{4.1}$$

where (x, y, z) is a Cartesian coordinate of the sun where the positive x-axis points to the south and the positive y-axis points to the east. It turns out that $s = -1$ corresponds to a sun rise and $s = 1$ corresponds to a sun set regardless of the day of the year. This gives the same domain for all daily power curves.

4.9.2 Regression of daily normalized power curves with B-splines

With the transformation addressed above, we can run a regression on daily normalized power curves with B-splines. We first construct B-splines by specifying the degree of basis functions, the number of inner knots, and the locations of knots. Next we see how the constructed curve fits with well-behaved daily power curves selected by a human. The coefficient of determination (R^2) is utilized to measure the fitness. We experiment with several degrees of basis functions, numbers and locations of inner knots. It turns out that a set of basis functions of the degree 3 with 8 inner knots distributed as shown in Figure 4.11 gives the best fit.

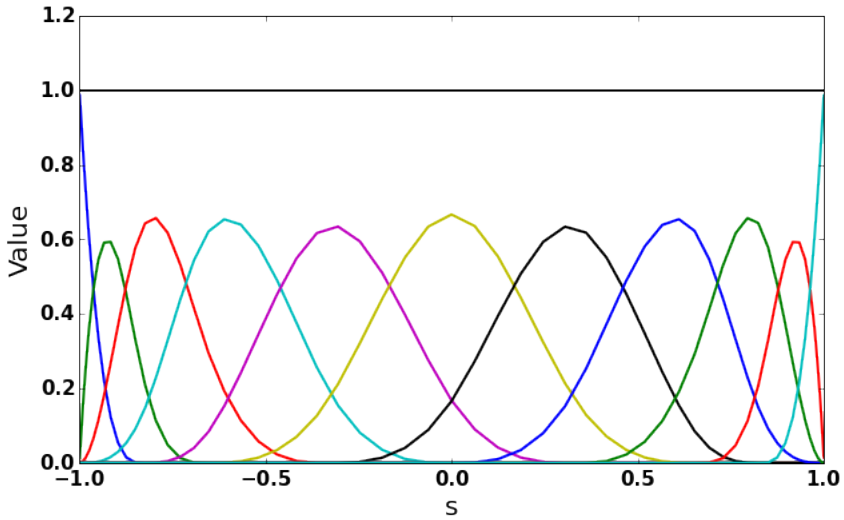
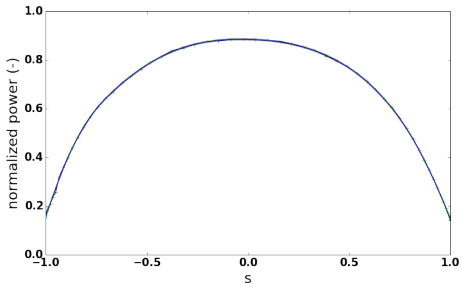
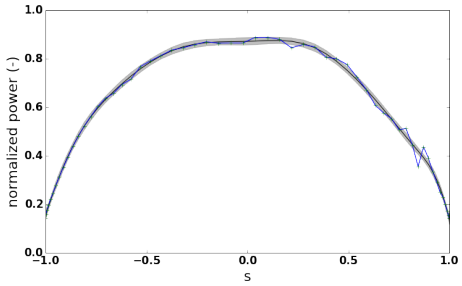


Figure 4.11: Basis functions for a regression including B-splines of the degree 3 with 8 inner knots. There are 11 basis functions in total.

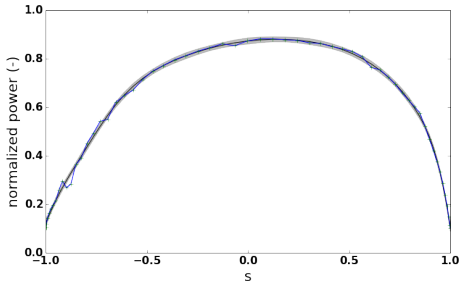


(a) A well-behaved normalized power curve, on 2014/6/24. $R^2 = 0.9998$

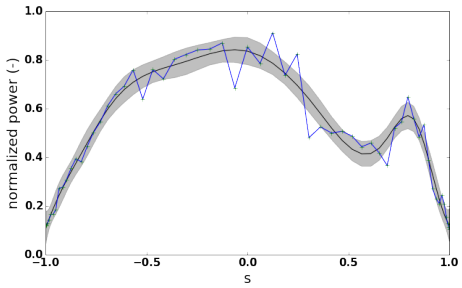


(b) A typical normalized power curve, on 2014/6/21. $R^2 = 0.997$

Figure 4.12: Actual (blue) and fitted (black) daily normalized power curves generated from the data set in CA



(a) A well-behaved maximum normalized power curve, on 2014/6/24. $R^2 = 0.998$



(b) A typical normalized power curve, on 2014/6/21. $R^2 = 0.959$

Figure 4.13: Actual (blue) and fitted (black) daily normalized power curves generated from the data set in NJ

Figures 4.12 and 4.13 below compare actual power curves and fitted power curves along with 95% confidence intervals. For well-behaved power curves, R^2 can be as high as 0.995. Linear combinations of B-splines for well-behaved power curves are indeed similar to each other and resemble bell shapes. However, linear combinations of B-splines for other power curves tend to have low values of R^2 and do not exhibit bell shapes. These clues lead to an automatic process for selecting well-behaved maximum power curves.

4.9.3 Automation of a well-behaved power curve selection

One of the key steps in modeling maximum normalized power curves for all days in a year is to find select well-behaved power curves from all normalized power curves generated from the power output data. Such a task can be tedious if done by a human. However, one can develop an algorithm to select well-behaved power curves. The selection process consists of two stages. The first stage is based on the information of the B-splines regression in the previous subsection. The second stage is based on a comparison of daily energy generations with neighboring days.

The first stage

From the B-splines regression, we propose two criteria to filter out power curves that are not well behaved as follows:

The first criterion is that, if the goodness of the fit of a daily power curve is low, the power curve is not a well-behaved maximum power curve. Here we use $R^2 = 0.95$ as a cut-off. This criterion comes from the assumption of maximum normalized power curves.

The second criterion is that, if a daily power curve does not have only one critical point, which is also the local maximum, the power curve is not a well-behaved maximum power curve. This criterion comes from the assumption of a bell-shaped curve.

In order to determine critical points, finding a derivative of a linear combination of B-splines of the degree 3 is helpful. Since finding derivatives directly from a recursion relation is a tedious task, we discovered a way to find a derivative of such functions based on numerical tryouts. We know that a function defined between two knots $x \in [k_i, k_{i+1})$ has a form:

$$f(x) = a_i x^3 + b_i x^2 + c_i x + d_i.$$

We may use a regression function to predict values for 4 distinct points, say $\{k_i, \frac{2}{3}k_i + \frac{1}{3}k_{i+1}, \frac{1}{3}k_i + \frac{2}{3}k_{i+1}, k_{i+1}\}$. Next, one can solve a system of 4 linear equations with 4 unknowns $\{a_i, b_i, c_i, d_i\}$. Then the derivative of the function is

$$f'(x) = 3a_i x^2 + 2b_i x + c_i$$

This leads to a simple algorithm:

Note: since our construction of splines has non-negative coefficients and zero value at endpoints, if there is one critical point then such a critical point is the maximum. There is no need to evaluate the second derivative.

One may propose another rule: if the peak of a daily power curve is an outlier compared to the peaks of daily power curves from neighbor days, then the power

Algorithm 1 The first stage filtering algorithm

1: For each interval $[k_i, k_{i+1})$ between knots of B-splines, predict values for $\{k_i, \frac{2}{3}k_i + \frac{1}{3}k_{i+1}, \frac{1}{3}k_i + \frac{2}{3}k_{i+1}, k_{i+1}\}$.
2: Use the values to construct a system of 4 linear equations and solve for $\{a_i, b_i, c_i, d_i\}$.
3: Find roots of the quadratic equation: $f'(x) = 3a_ix^2 + 2b_ix + c_i = 0$. Count the number of roots that lie inside the interval. This value is the number of critical points in the interval.
4: Count the total number of critical points of the function. If the total number does not equal 1, we filter the daily power curve out.

curve is not a well-behaved power curve. However, this criterion is not practical as the daily power curves after a domain transformation do not have sharp peaks. Hence we do not apply this criterion.

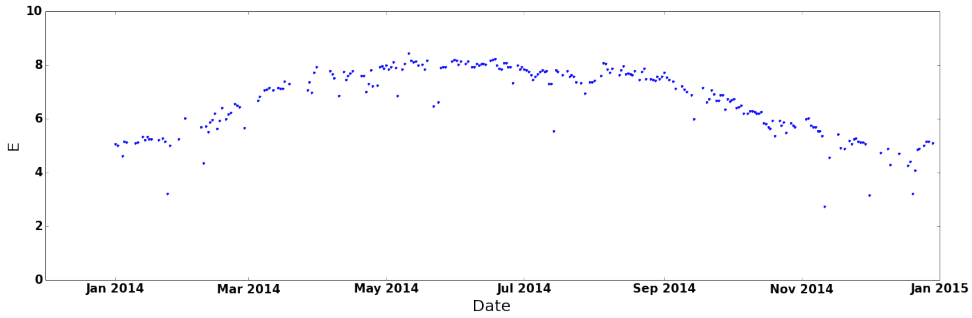
After applying the algorithm explained above, we found that, for the data set in CA, 35% of daily power curves remained. For the data set in NJ, 20% of daily power curves remained.

After the first stage, we filtered out some of the daily power curves that are not well-behaved power curves. However, the remaining daily power curves in the data set were not necessarily well-behaved power curves. For example, a daily power curve with a low energy generation due to uniform cloud cover may exhibit a bell shape and may pass the first stage. We need a more elaborate procedure to continue filtering.

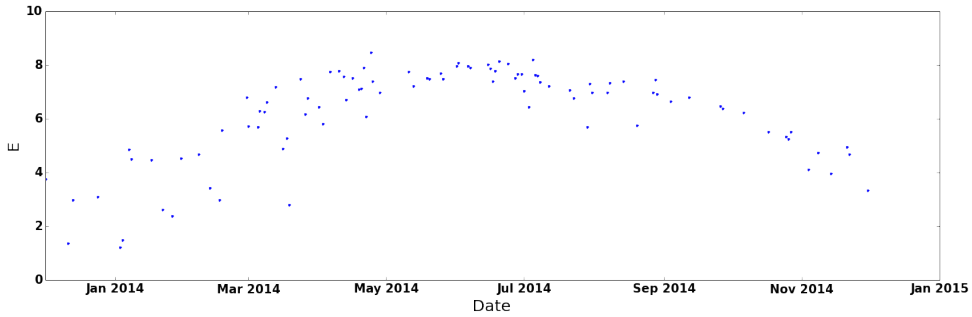
The second stage

To start the second stage, it is useful to look at daily energy generations from all daily power curves that pass the first stage. Figure 4.14 shows those curves. Note that there were not many points before 2014 since too few PV systems existed to generate sufficient data at that time.

Since we have justified that the variation between years is negligible, we can overlay the data from multiple years into a single year window. We convert the date into a variable called Day Before a Winter solstice (DBW). We set the Winter solstice of each year to be 0 and assign the other days with negative integers. Figure 4.15 shows the daily energy generations with the overlay. The plot shows seasonal variation from the Winter Solstice (DBW = -365) to the Summer Solstice in the middle of the plot, and back to the Winter Solstic ($DBW = 0$). In this representation, it is clear that there is an envelope contributed from well-behaved daily power curves and there are outliers below the envelope of each data set.

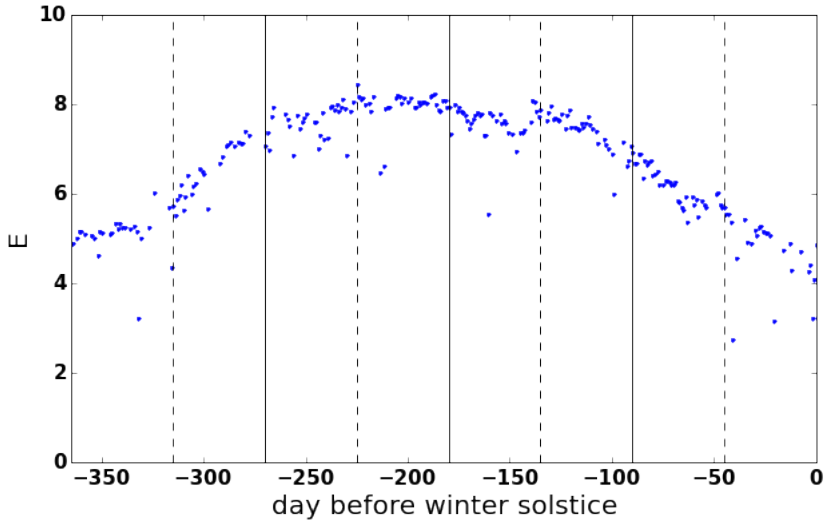


(a) CA

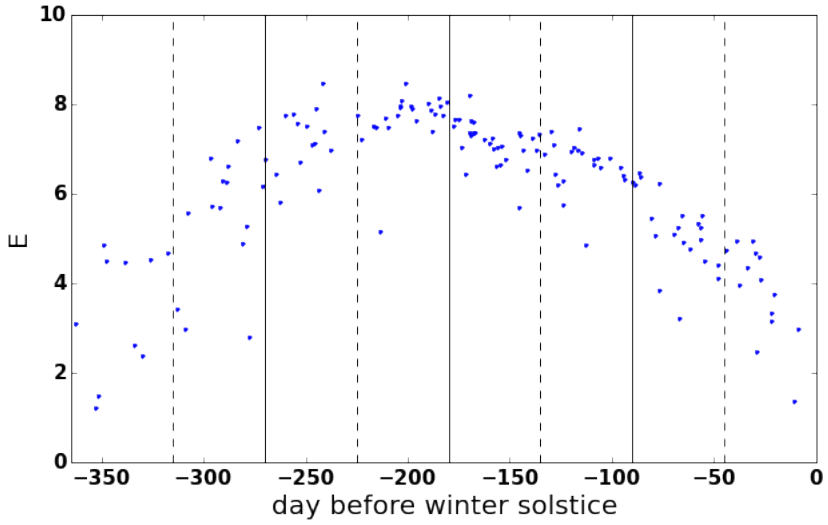


(b) NJ

Figure 4.14: The daily energy generations of all daily power curves that pass the first stage.



(a) CA



(b) NJ

Figure 4.15: The daily energy generations of all daily power curves that pass the first stage with a new variable DBW and overlay.

In order to filter out those outliers, we introduce the concept of the "excessive downhill". A data point in the plot is in an excessive downhill if there exists a higher point in the plot such that the line segment connecting them has a slope higher than a selected threshold. We filter out a power curve if it is associated with a point with an excessive downhill in the daily energy generation plot. Using the data plotted in Figure 4.15, we select a threshold to be the highest slope on the envelope, which is about $(8-6)/(360/8) = 2/45 \approx 0.05$, at the first half of any given spring. This condition guarantees that any point that we eliminate deviates excessively from the envelope. The filter at this stage can be summarized as follow:

Algorithm 2 The second stage filtering algorithm

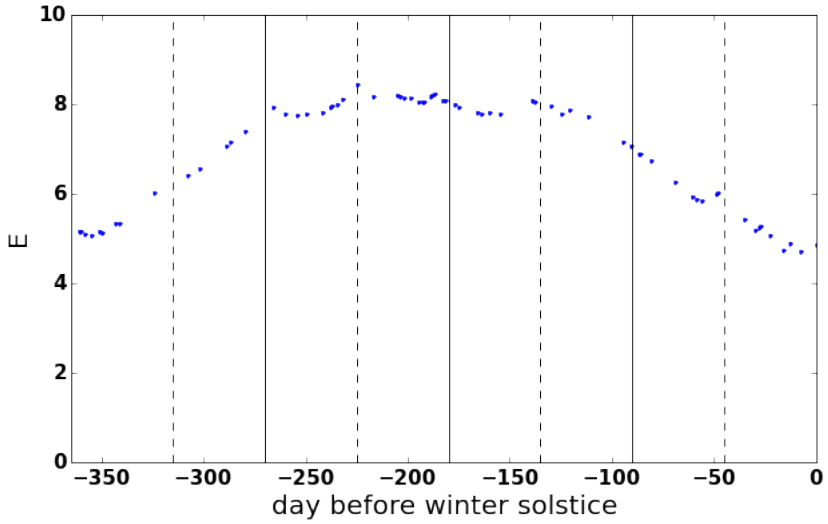
1: Calculate daily energy generation for each normalized daily power curve and plot it as function of a Day Before Winter solstice (DBW).
2: For each point i in the plot, compute the slope of the line segment connecting it with a higher point j.
3: If the slope is higher than the threshold 0.05, filter the daily power curve associated with the point i out. If not, repeat for all higher points j.

Figure 4.16 shows the daily energy generations after the second stage selection. After the second stage, we found that, for the data set in CA, 9% of daily power curves remained. In the data set in NJ, 3% of daily power curves remained. The daily power curves after the second stage are used to generalize the maximum power curves for all the days in a year.

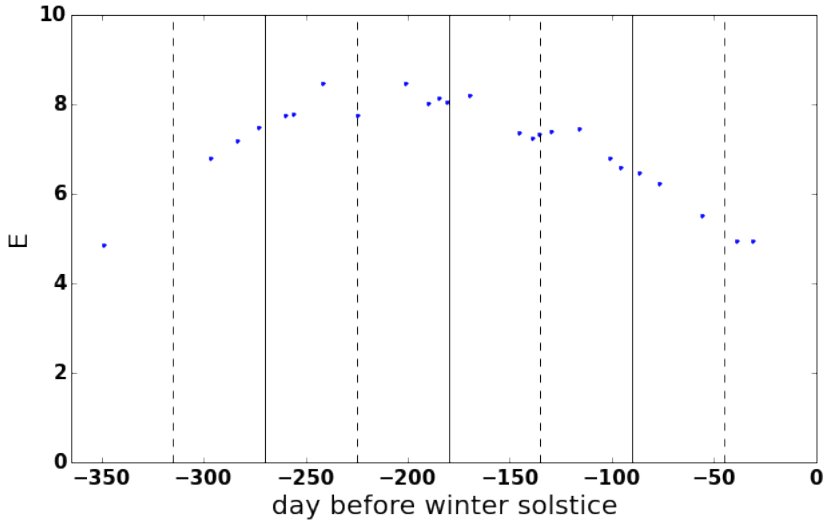
4.9.4 Generalization of models from well-behaved power curves

Now that we have a set of well-behaved normalized power curves, we may extrapolate the regression functions from well-behaved power curves as maximum normalized power curves for all the days examined. To be exact, for each selected power curve, one can extract the regression coefficients of its linear combination of B-splines. Each coefficient can be treated as a function of DBW as shown in Figure 4.17 and Figure 4.18. We run the regression and use the regression function on coefficients to determine the coefficients for all maximum normalized power curves.

In Figure 4.17 and Figure 4.18, even though we have a sophisticated selection process for well-behaved normalized power curves, there are some outliers with low values in coefficients. One way to decrease the significance of outliers in the regression is to weigh outliers differently. A quantile regression can be used. While the method of least squares aims for a conditional mean of a response variable given the value of a predictor variable, a quantile regression aims for a quantile of the response variable given a value of a predictor variable. Since we want to use the quantile regression



(a) 'CA' data set



(b) 'NJ' data set

Figure 4.16: The daily energy generations of all daily power curves that pass the second stage with a new variable DBW and overlay.

function as a maximum power curve, we seek the τth quantile, where τ is close to 1. However, the quantile regression with a high value of τ is usually too sensitive to anomalous high data points. Another way to remove the outliers from the regression is to select points using the excessive downhill technique to data points first, then perform the quantile regression with $\tau = 0.5$.

Various choices of regression functions on the coefficients can be applied. In this case, we use splines again. Periodic splines are appropriate because our domain is a set of all the days in a year, and therefore periodic. However, another family of periodic functions, Fourier series, may be employed. After some experiments we found that the Fourier series up to the third harmonic gives a good fit. The black lines in Figure 4.17 and Figure 4.18 show the quantile regression of coefficients and predicted coefficients for all the days using the Fourier series. With these coefficients, we can now generalize B-spline models to construct maximum power curves for all the days.

4.9.5 A correction for maximum normalized power curves

After obtaining coefficients of B-splines, we are able to build maximum normalized power curves for all the days. Figure 4.19 and Figure 4.20 are examples of daily maximum normalize power curves from the data sets in CA and NJ. Figure 4.21 shows the daily energy generations of maximum power curves in comparison with all daily power curves. In some instances from both plots of power readings and daily energy generations, our proposed maximum values are lower than the actual values from the data. Further investigation shows that, for the data set in CA, 17% of points are higher than the corresponding points in the daily maximum normalized power curves. For the data set in NJ, there are 11% of such points.

In order to get a true maximum power curve, one can modify a proposed maximum power curve by scaling and shifting. The modification should make the change as small as possible. We may write this as an optimization problem:

$$\begin{aligned} \underset{a\geq 1, b\geq 0}{\text{minimize}} \quad & \sum_{i=1}^{N} [am_i + b - m_i] \\ \text{subject to} \quad & p_i \leq am_i + b; i = 1, \ldots, M, \end{aligned}$$

where a is a scaling factor, b is a vertical shift, m_i is a value on an original maximum power curve, p_i is a value of on a normalized power curve, N is the number of points in all maximum power curves, and M is the number of points where the value in the original maximum power curves is lower than the value in the actual power curves. The optimal solution of such a problem is also the optimal solution of

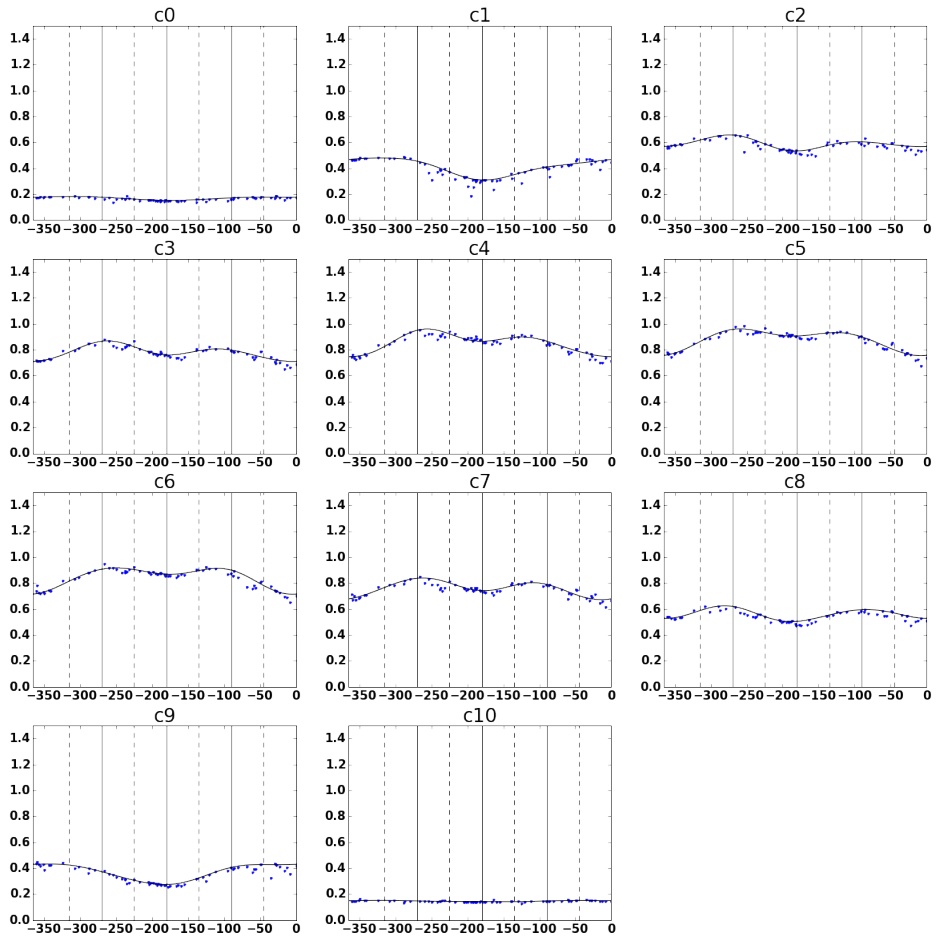


Figure 4.17: Quantile regression for coefficients of maximum power curves of the data set in CA

$$\begin{aligned} \underset{a\geq 1, b\geq 0}{\text{minimize}} \quad & a\sum_{i=1}^{N} m_i + Nb \\ \text{subject to} \quad & am_i + b \geq p_i; i = 1, \ldots, M, \end{aligned}$$

This is a linear programming problem with the optimal solution of $a = 1$ for all data sets. It turns out that the optimal solution for all data sets has $a = 1$. That is, the least perturbation is done by shifting only. For the data set in CA, the optimal solution is to shift maximum power curves up by 0.0862 without scaling. For the data set in NJ, the optimal solution is to shift up by 0.1458 without scaling. Once these shifts are applied, all points of maximum power curves become higher

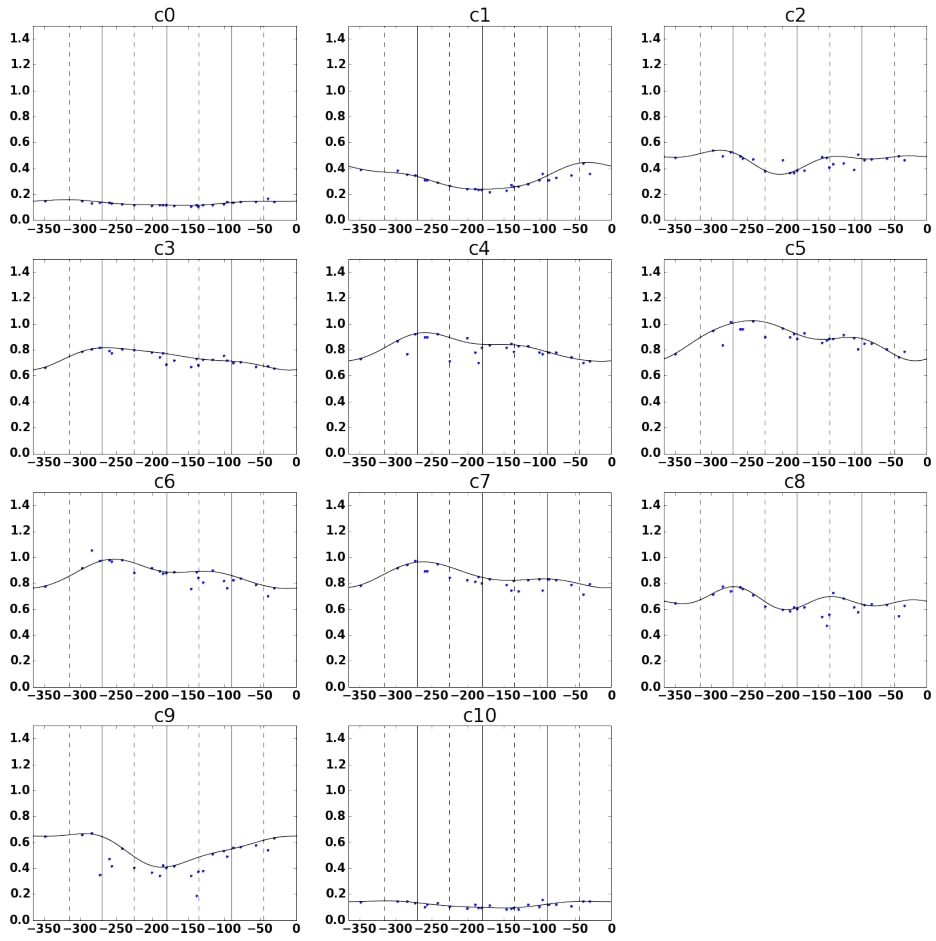
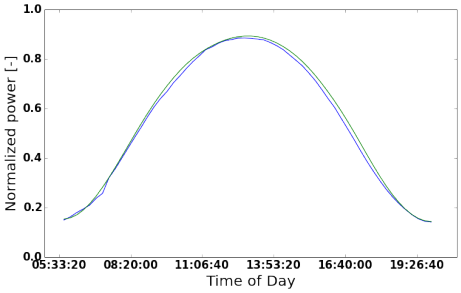


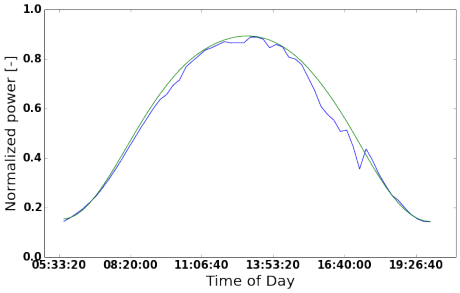
Figure 4.18: Quantile regression for coefficients of maximum power curves of the data set in NJ

than corresponding points in power curves from our data set. That is, we obtain legitimate maximum normalized power curves.

It should be noted that the correction by shifting has a downside as it implies that there is a sharp rise and drop in power occur at a sun rise and a sun set. In order to solve this issue, we can impose a non-uniform but smooth shifting and scaling to the power curve. However, we do not pursue the task in this study.

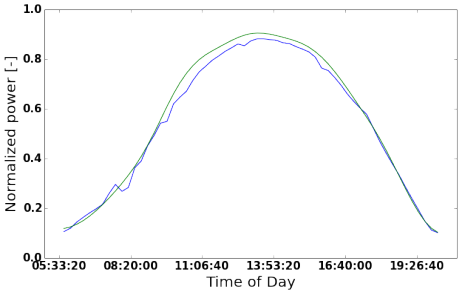


(a) A well-behaved normalized power curve, on 2014/6/24

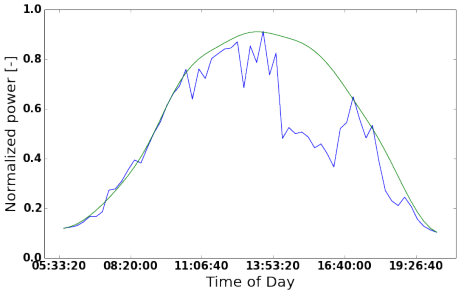


(b) A typical normalized power curve, on 2014/6/21

Figure 4.19: Daily normalized power curves and proposed maximum power curves from the data set in CA

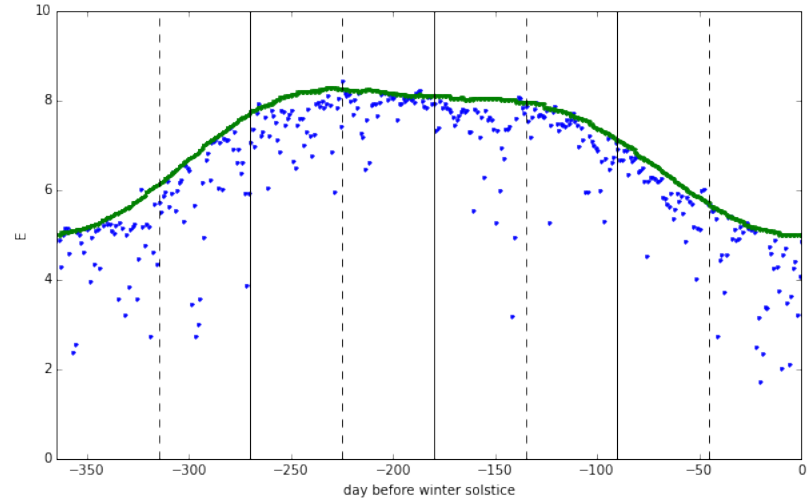


(a) A well-behaved normalized power curve, on 2014/6/24

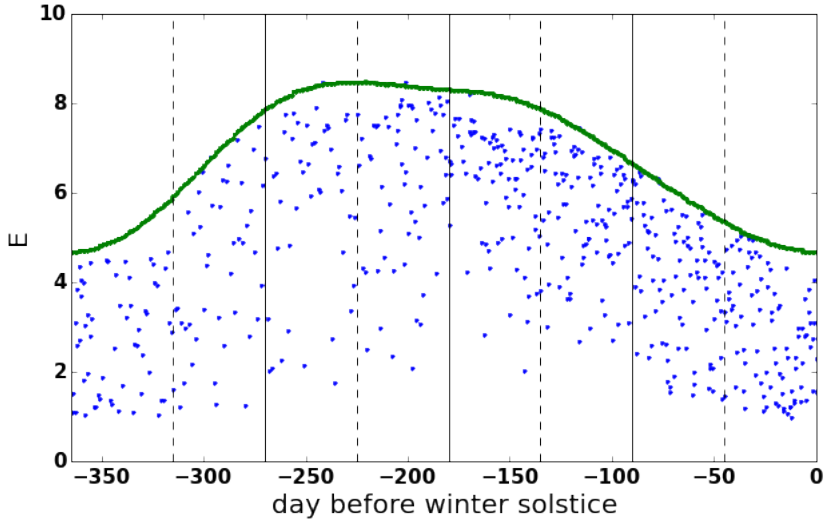


(b) A typical normalized power curve, on 2014/6/21

Figure 4.20: Daily normalized power curves and proposed maximum power curves from the data set in NJ



(a) CA



(b) NJ

Figure 4.21: Daily energy generations of proposed maximum power curves in comparison with all daily power curves

4.10 Consistency of maximum normalized power curves

To complete our work on maximum normalized power curves, we investigate the consistency of our end product. We test if maximum normalized power curves generated from two samples of PV systems in the same geographical area are different.

We randomly split PV systems in the data set into two groups. Next we generate normalized power curves from power output data, select well-behaved power curves and construct maximum normalized power curves as described in the previous sections. Finally, we compare maximum normalized power curves from two sets using both consistent measures: the mean of the average absolute deviation in power and the average absolute deviation in yearly energy generation. Since the number of PV systems in the data set in CA is low, we perform the test for the data set in NJ only.

For the data set in NJ, we split the PV systems into two sets and generate two maximum normalized power curves. We found that the mean of the average absolute deviation in power equals 0.0135. The average absolute deviation in yearly energy generation equals 1.6 percent. Hence we may conclude that maximum normalized power curves constructed by our method are consistent.

4.11 Conclusion

We define what a maximum power curve for a group of PV systems is and how it is useful for grid operators. We propose a reduced form approach to construct maximum power curves for a group of PV systems based on time series of power outputs and its total installed capacity only. We show the detail of the construction process and consistency checks from the initial product, which is a set of daily normalized power curves produced from actual data, to the final product, which is a set of daily maximum normalized power curves for all the days in a year.

In the next chapter, we show how a grid operator may use the maximum PV power curves to estimate the maximum PV penetration into the grid. In a subsequent chapter, we show that the use of the maximum PV power curves can improve an accuracy of a PV power forecast.

Chapter 5

Applications of maximum PV power curves

Highlights:

- Maximum PV power curves give a clarification on different notions of a PV capacity.
- Maximum PV power curves lead to a useful concept of a performance ratio, a key variable in an hour-ahead forecast.
- Expected PV power curves can be constructed from maximum PV power curves and the mean of performance ratios. Such curves are useful in a day-ahead energy scheduling.
- A maximum PV penetration to a grid can be computed from maximum PV power curves based on a set of simple assumptions. A more realistic estimate requires further investigation of power flow constraints and a system to cope with a demand-supply mismatch.

In Chapter 4, we have constructed maximum PV power curves, which illustrate a variation in the maximum power generation from a group of PV systems. Instead of a structural approach, we used a reduced form approach, where we construct consistent daily normalized power curves from power output data of PV systems and build the maximum normalized power curves without a need of extra information except the installed capacities of a group of PV systems. In this chapter, we show some basic applications of maximum PV power curves. First, we show how maximum PV power curves clarify a distinction between an installed capacity, a maximum generated capacity, and an expected generated capacity. Second, we introduce a performance ratio and an expected PV power curve derived from maximum PV curves. Finally,

we show how maximum PV power curves establish a maximum installed capacity of PV systems in correspondence with a load pattern.

5.1 Different notions of PV capacity

There are number of ways to report how much power a group of PV systems can be generated. The most common way to report is an **installed capacity**, also known as a nameplate capacity, a rated capacity, or a nominal capacity. An installed capacity of an individual PV cell is rated under a Standard Test Condition (STC) [34]. That is, the panel is under a light intensity of 1000 W/m^2, with a spectrum similar to sunlight hitting the earth's surface at the latitude 35° N in the summer (airmass 1.5), and the temperature of the cell being 25° C. The highest measurement while varying the resistive load on the module between a maximum and a minimum resistance is the installed capacity. This installed capacity divided by the incident light power defines the panel's efficiency. For a group of PV systems, one may define an installed capacity as a sum of installed capacities from each panel in the group.

Even though the installed capacity is useful to compare efficiencies of PV panels fairly, it is not quite useful by itself in term of an actual power generation. The immediate drawback of the installed capacity is that it does not match the actual power generated by a group of PV systems since each panel operates in an environment different from the STC. In a grid operator's point of view, the maximum or the average power are useful than the power measured at a certain condition. In fact, the word 'installed' used in the PV industry can be misleading since the generator installed capacity typically refers to the maximum power from a generator in the grid operation [2]. Hence, we define more useful ways to report the power including a peak capacity, an maximum average power production, and an average power production.

A **peak capacity** of a group of PV systems is the maximum power a group of PV systems can generate. Given a set of maximum normalized power curves of a year, the peak capacity can be written as

$$\text{peak capacity} = C \max_{d=1,..,365} \max_{i=1,..,96} P_d^{max}(i)$$

where C is the installed capacity of a group of PV systems and $P_d^{max}(i)$ is the value of the maximum normalized power curve of the day d at the time instance i. Since the time interval is 15 minutes, there are 96 instances in a day. For a group of PV systems with an installed capacity of 1 kW, the peak capacity is not necessarily equal to 1 kW. This number is more informative for the grid operator than the installed capacity since it gives a more accurate picture of the maximum power that a group of PV systems can generate at an instance in a year. If we know all information about each PV system in the group, we may use a structural approach to compute the

peak capacity. However, as addressed in the previous chapter, it is easier to obtain the maximum PV power curve along with the peak capacity using a reduced form approach. Table 5.1 shows two peak capacities from two groups of PV systems in the data sets from CA and NJ. Note that both of them are slightly higher than 1 kW.

An **maximum average power production** of a group of PV systems is an average of an ideal power from a group of PV systems over a year. Given a set of maximum normalized power curves of a year, the average maximum capacity can be written as

$$\text{maximum average power production} = \frac{C}{365 \times 96} \sum_{d=1}^{365} \sum_{i=1}^{96} P_d^{max}(i)$$

where C is the installed capacity of a group of PV systems and $P_d^{max}(i)$ is the value of the maximum normalized power curve of the day d at the time instance i. Since the time interval is 15 minutes, there are 96 instances in a day. This number is more informative for the grid operator than the installed capacity since it gives a more accurate picture of the maximum power that a group of PV systems can generate over a year. Table 5.1 shows two maximum average power production values from two groups of PV systems in the data set from CA and NJ. Note that both of them are significantly less than 0.5 kW. This is because we also count time instances without sunlight in the average in order to give a fair comparison if compared with other conventional energy generators.

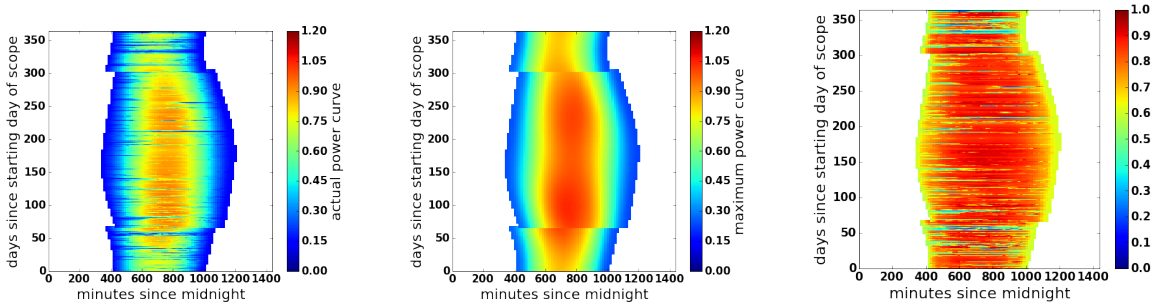
An **average typical capacity** of a group of PV systems is an average of an actual power from a group of PV systems over a year. Given a set of actual normalized power curves of a year, the average typical capacity can be written as

$$\text{average power production} = \frac{C}{365 \times 96} \sum_{d=1}^{365} \sum_{i=1}^{96} P_d(i)$$

where C is the installed capacity of a group of PV systems and $P_d(i)$ is the value of the actual normalized power curve of the day d at the time instance i. Since the time interval is 15 minutes, there are 96 instances in a day. While an average maximum capacity is informative in extreme case problems such as a maximum penetration assessment, an average typical capacity is informative in general behavior problems such as a PV generation's value assessment. Table 5.1 shows two average power production values from two groups of PV systems in the data sets from CA and NJ in 2014 only. Note that the average power production is significantly less than the maximum average power production

Power measure (kW)	CA	NJ
Installed capacity	1.0000	1.0000
Peak capacity	1.0431	1.1494
Maximum average power production	0.3346	0.3600
Average power production (2014)	0.2641	0.1979

Table 5.1: Different power measures for a group of PV systems from two data sets



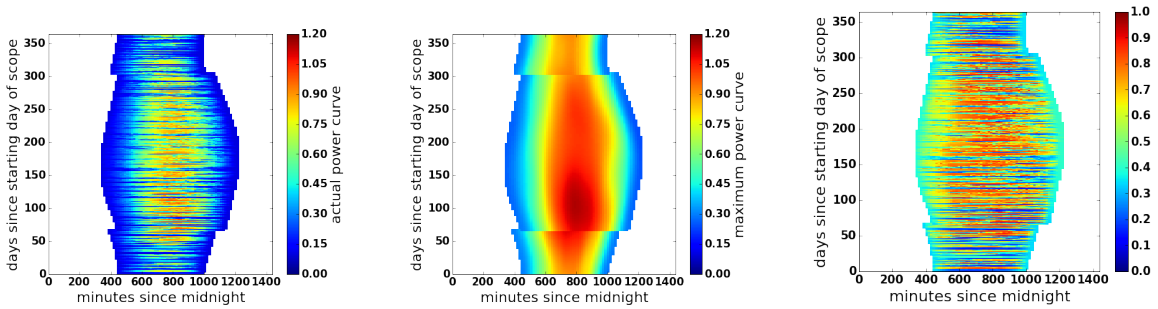
(a) Actual output (b) Maximum output (c) Performance ratio

Figure 5.1: Actual power outputs, maximum power curves, and performance ratios for the data set in CA in 2014

5.2 Performance ratios and expected PV power curves

In order to find how well a group of PV systems generate at a given instance or a time period, it makes sense to have a definition relative to the maximum value. This leads to the definition of **performance ratios**. A performance ratio at an instance is defined as an actual power output divided by a value on maximum power curve at that instance. Figure 5.1 and 5.2 show actual power outputs, maximum power curves, and performance ratios in two dimensional grids for the data sets in CA and NJ in 2014. The horizontal axis represents a time of a day, and the vertical axis represents a day of a year.

An immediate application of the performance ratio is expected PV power curves. In a long-term planning or a day-ahead energy scheduling, suppose that a grid operator does not have any prior about recent PV output data and weather data, how would the grid operator construct an expected PV power curve, which describes an average power output from a group of PV systems over a day? The expected PV power curve should reflect a PV power's variability but not a volatility. One may construct a simple expected PV power curve by scaling a maximum PV power curve



(a) Actual output (b) Maximum output (c) Performance ratio

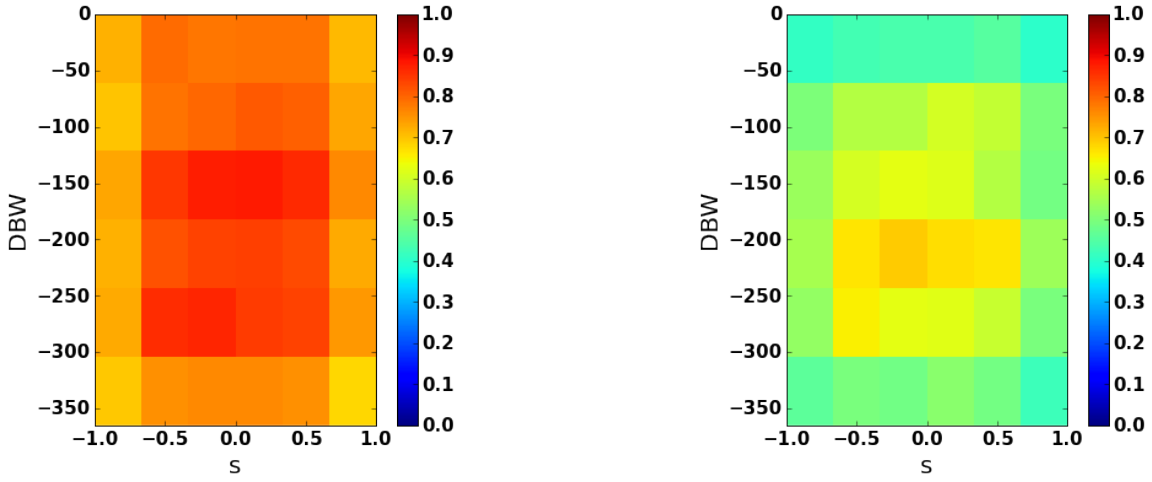
Figure 5.2: Actual power outputs, maximum power curves, and performance ratios for the data set in NJ 2014

by a factor. A suitable factor is the mean of performance ratios. From the data set in CA in 2014, the mean performance ratio equals 0.7641. From the data set in NJ in 2014, the mean performance ratio equals 0.5282.

However, if one look closely to the two dimensional grid of performance ratios, there is a non-homogeneity in values by average. Performance ratios near sun rises and sun sets tend to be lower than performance ratios in the middle of day. Performance ratios in a winter tend to be lower than performance ratios in a summer. Figure 5.3 shows the mean performance ratio for different buckets of (s, DBW) where s is a parameter associated with a time of a day from a sun rise ($s = -1$) to a sun set ($s = 1$), while DBW is a parameter associated with a day of a year from a day after a winter solstice ($DBW = -364$) to the next winter solstice ($DBW = 0$).

With this set of mean values, one can refine an expected PV power curve by scaling each part of a maximum PV power curve with a different factor depends on a time of a day and a day of a year. Such factor is the mean value on each bucket. In order to avoid a discontinuity, one can use an interpolation from mid points in the (s, DBW) grid to define a continuous function of (s, DBW). A simple interpolation is a linear piecewise function on triangular meshes. We also impose periodic boundary conditions on the domain. Once we have such a continuous function, we multiply it with the maximum PV power curve to obtain the expected PV power curve. Figure 5.4 and 5.5 show examples of daily expected power curves (dashed green lines) in a comparison with maximum power curves (solid green lines) and actual power curves (blue lines). Such refinement makes an expected PV power curve to have a different shape from its corresponding maximum PV power curve, but to be consistent with the actual PV power pattern by average.

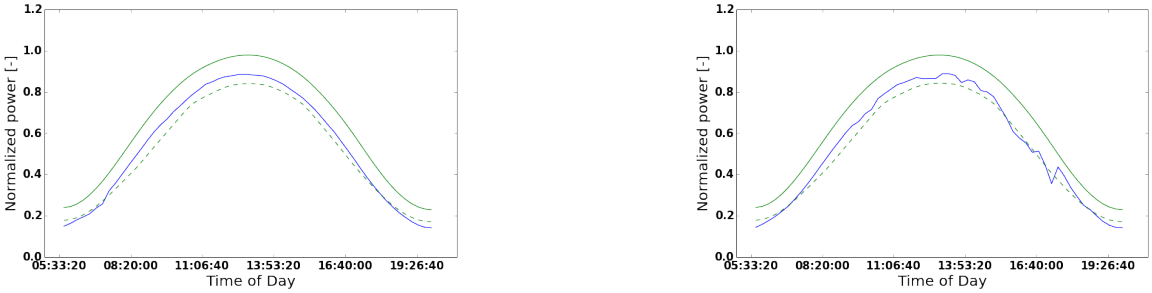
Another application of performance ratios is a hour-ahead PV power forecast. In a hour-ahead energy scheduling, a grid operator needs to predict the actual PV power an hour ahead. It turns out that predicting performance ratios first, instead



(a) CA

(b) NJ

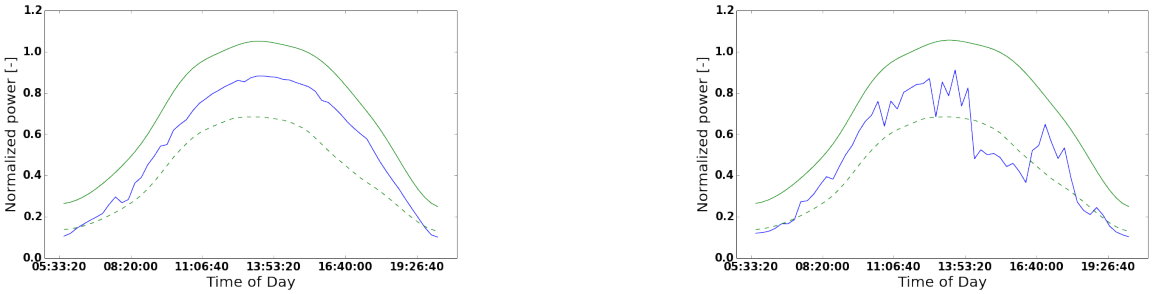
Figure 5.3: Mean performance ratios for each bucket of (s, DBW) in 2014



(a) Well-behaved normalized power curve, on 2014/6/24

(b) Typical normalized power curve, on 2014/6/21

Figure 5.4: Daily normalized power curves generated from the data set in CA



(a) Well-behaved normalized power curve, on 2014/6/24

(b) Typical normalized power curve, on 2014/6/21

Figure 5.5: Daily normalized power curves generated from the data set in NJ

of predicting power outputs directly, gives a higher prediction accuracy. The detail of a forecast system utilizing performance ratios is given in a subsequent chapter.

5.3 A maximum PV penetration

One of basic questions about a PV penetration is that how much an installed capacity of PV systems one can put into a grid. It sounds simple, but such question requires more details in order to give a meaningful answer. In this section, we will dissect this question by configuring details on the scope and assumptions of the problem until we have a reasonable answer.

5.3.1 Limitation of data

Throughout the grid integration analysis in this study, we assume that all PV systems in the grid behave similarly to the PV systems as we have in our data set. The data on PV generation comes from a particular area in a state where the data on power consumption comes from the whole grid which may cover more than a single state. Hence, our approach to combine two sets of data in the grid integration analysis is significantly different from what the grid operator should do in practice. This is due to the limitation of data. In this study, the focus is to show the procedures rather than the absolute results. To highlight the distinction between the example and the real operation, we call the combination of the PV data from a zip code in CA and the load data of a whole grid where the zip code is located as **the western example**. We call the combination of the PV data from a zip code in NJ and the load data of a whole grid where the zip code is located as **the eastern example**.

5.3.2 The first scenario

In the first scenario, suppose that the whole grid has infinite storage capacity with a complete control on a ramp rate. Suppose that the grid operator can transfer the power from one point to another in the grid without any constraints. Then one can serve a whole energy consumption in the grid using an energy generation by PV systems. Hence, one can install from zero up to the point where the total energy generation by PV systems matches the total energy consumption in the grid. To balance both demand and supply in a yearly basis, we consider the total energy consumption and the PV generation for a year.

Let's consider the total energy consumption for a year. The grid of California is operated by a grid operator called CAlifornia Independent System Operator (CAISO). From hourly load data of the whole grid in 2014 under CAISO, one can sum all values over time and get the total energy consumption of 228.1 TWh. This energy

consumption is equivalent to have a constant load of 26.04 GW for a year. A grid of New Jersey is a part of bigger grid operated by a grid operator called PJM. From hourly load data of the whole grid in 2014 under PJM, one can undergo a similar process and get the total energy consumption of 797.0 TWh. This energy consumption is equivalent to have a constant load of 90.98 GW for a year.

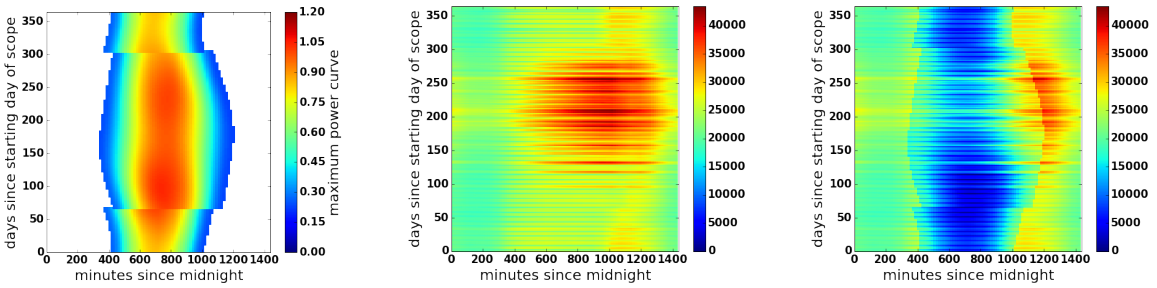
Now consider the total PV generation for a year. We use our maximum normalized PV power curves and our expected normalized PV power curves to represent the total PV generation in the grid. In an ideal case, we would be able to generate the total energy equal to the area under maximum normalized PV power curves for a year, multiplied by an installed capacity. In a realistic case, we should be able to generate the total energy equal to the area under expected normalized PV power curves for a year, multiplied by an installed capacity. For the data set in CA in 2014, the area under maximum power curves over a year equals 2.931 MWh and the area under expected power curves over a year equals 2.287 MWh. For the data set in NJ in 2014, the area under maximum power curves over a year equals 3.153 MWh and the area under expected power curves over a year equals 1.707 MWh.

Since we know both information, we can establish the maximum installed capacity for the first scenario by dividing the total energy consumption by the area under normalized PV power curves. For the western example, the maximum installed capacity equals to 77.82 GW in the ideal case and 99.74 GW the realistic case. For the eastern example, the maximum installed capacity equals to 252.8 GW in the ideal case and 466.9 GW in the realistic case.

5.3.3 The second scenario

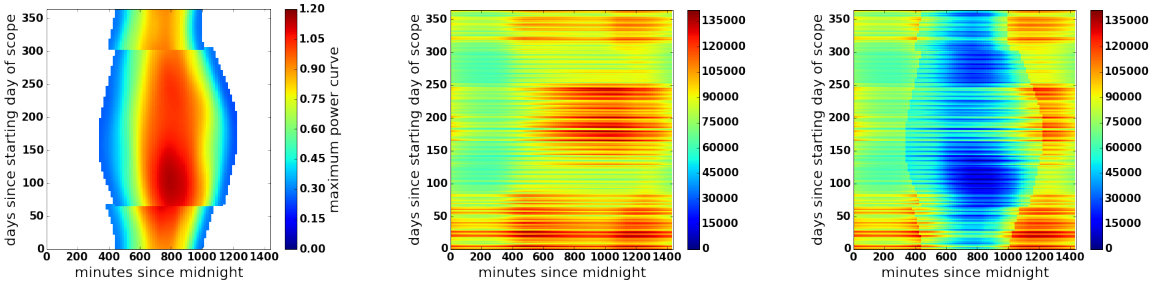
In the second scenario, suppose that the whole grid has no storage capacity at all. We still suppose that the grid operator can transfer the power from one point to another in the grid without any constraints. Then at any time, the total power generated PV systems cannot be larger than the total load in the grid. Hence, one can install from zero up to the point where there exists an instance where the total power from PV systems equals to the total load of the grid. In other words, if a PV generation is interpreted as a negative load, one can install from zero up to the point where there exists an instance where the net load reaches zero.

In order to find the maximum installed capacity, we need to compare maximum normalized PV curves and load curves for a whole year. The maximum installed capacity is just the minimum of ratios of load values and normalized PV power values. Figure 5.6 and show two dimensional representation of normalized PV power curves, load curves, and net load curves at the maximum penetration. For the western example, the maximum installed capacity equals to 21.62 GW. For the eastern example,



(a) Normalized max PV (b) Load in GW (c) Net load in GW

Figure 5.6: The maximum normalized PV power, the load, and the net load at the maximum penetration for the western example in 2014

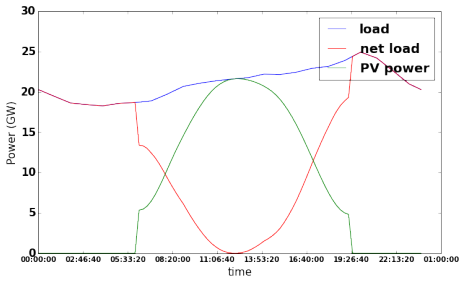


(a) Normalized max PV (b) Load in GW (c) Net load in GW

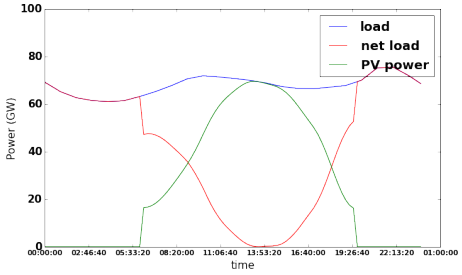
Figure 5.7: The maximum normalized PV power, the load, and the net load at the maximum penetration for the eastern example in 2014

the maximum installed capacity equals to 60.53 GW. With this amount of installation, PV systems contributes only a fraction of the total energy consumption. For the western example, the maximum PV installed capacity contributes 27.8 % of the yearly energy consumption, and for the eastern example, the maximum PV installed capacity contributes 23.9 % of the yearly energy consumption. The maximum penetration of PV systems in the second scenario drop significantly from the first scenario because the surface of the two dimensional PV generation curve do not completely align with the surface of the two dimensional load curve. From the net load color map in Figure 5.6 and 5.7, one can see that the load in the evening and the summer is so high that the PV generation cannot accommodate. Figure 5.8 shows the load, PV power, and net load curves on the date that determines the limit, which is April 20, 2014 in both cases.

It is worth noting that using maximum power curves is too stringent as we currently have a system to cope with a mismatch between a demand and a supply at



(a) Western



(b) Eastern

Figure 5.8: Load, PV power and the net load curves under the maximum penetration limit on 2014/4/20

Power measure (GW)	Western	Eastern
Average load	26.04	90.98
Peak load	43.48	141.7
Max PV capacity - 1st scenario (ideal)	77.82	252.8
Max PV capacity - 1st scenario (realistic)	99.74	466.9
Max PV capacity - 2nd scenario	21.62	60.53
Yearly energy measure (TWh)	**Western**	**Eastern**
Total demand	228.1	797.0
PV generation - 1st scenario (both cases)	228.1	797.0
PV generation - 2nd scenario	63.41	190.5

Table 5.2: A comparison of the power and the energy of the load demand and the PV generation from different scenarios in 2014

any instance. Hence, the more optimistic maximum installed capacity can be higher if the system to cope with the mismatch has a high capacity.

5.3.4 Comments

Table 5.2 summarizes key figures from the hourly load demand in 2014 in the western and the eastern examples, along with the power and the energy generation of PV systems under two scenarios. It is worth noting that the first scenario is very unlikely in both western and eastern examples. However, such scenario analysis may be applicable to isolated and small grids.

There are some disagreements with the future grid and the second scenarios. On the one hand, the second scenario holds an overly optimistic assumption that the power can be transferred from one point to another in the grid without any

constraints. The old power distribution infrastructures are not suitable for power flows back to substations. Even though an update is made so that the back flow is possible, the power flow must follow specific rules. The problem of matching a demand and a supply under constraints is called a power flow problem. Hence, it is too optimistic to claim such a high installed capacity limit based on the second scenario since the system may violate the power flow constraints in the grid. However, we will not address this issue in the study.

On the other hand, the second scenario holds an overly pessimistic assumption that mismatches between a demand and a supply is not allowed in any time instance. In fact, the current grid can handle some amount of mismatches between a demand and a supply by various instruments including an ancillary service, an energy storage, and a curtailment. This allows more installations of PV systems. In addition, it is unlikely that the actual PV power output reaches the maximum as quoted in the maximum power curve. We may construct a better reference than the maximum power curve once we understand the volatility of the PV power curve in order to assess the maximum installed capacity. Hence, it is too pessimistic to claim such a low installed capacity limit based on the second scenario. We will discuss how to refine the maximum penetration with extra information in a subsequent chapter.

5.4 Conclusion

In this chapter, we illustrate some basic applications of maximum PV power curves generated by a reduced form approach. With support from maximum PV power curves, we are able to distinguish different notions of PV capacities. Using maximum PV power curves, we can define performance ratios and construct expected PV power curves, which are useful for a day-ahead planning. Finally, we can use maximum power curves to estimate the maximum PV penetration to the grid with a set of simple scenarios.

In subsequent chapters, we show more sophisticated applications of maximum PV power curves in an hour-ahead forecast. Using maximum PV power curves and performance ratios, we construct a system for an hour-ahead forecast, which is useful in hour-ahead energy scheduling. A performance analysis of the forecast system informs us about a capacity necessary to accommodate errors in the prediction. Once we know such a capacity, we are able to refine our procedure to estimate the maximum PV penetration and give a more realistic estimate of the maximum PV penetration to the grid.

Chapter 6

A forecast system for distributed PV generation

Highlights:

- A short-term PV power forecast is important in a grid operation under high PV penetration.
- A formal definition of volatility is introduced and validated. It gives a more informative illustration of a forecast system's performance, which leads to a procedure to manage potentially many complementary prediction models in a system.
- We investigate prediction models in three features: a use of indirect prediction via performance ratios, a set of predictors, and size of training data.
- We give an example of a forecast system where several prediction models are assembled. We give a protocol to switch between models depending on information available at an instance and volatility of the past hour.

As a part of grid operation, a short-term load forecast is essential to matching energy demand and supply. An ability to forecast distributed PV power in a short term is also essential as it is a part of net loads. Even though the variability of PV generation and the regular load are related by common factors such as time of a day and day of a year, since we model the net loads under different levels of PV penetration, it makes sense to forecast the PV generation and regular loads separately instead of forecasting the net load with a single model. Here we see that a concept of performance ratios with respect to maximum PV curves becomes useful in a short-term PV power forecast system.

We introduce a formal definition of volatility that is different from the concept of variability. With this definition, the volatility leads to our understanding of a forecast

system's performance. Then we test several prediction models with three features: a use of indirect prediction via performance ratios, a set of predictors (previous power outputs and other variables), and a size of training data as shown in Figure 6.1. Finally we illustrate how volatility informs us how to switch back and forth between an ensemble of several models in order to maximize the performance of a forecast system.

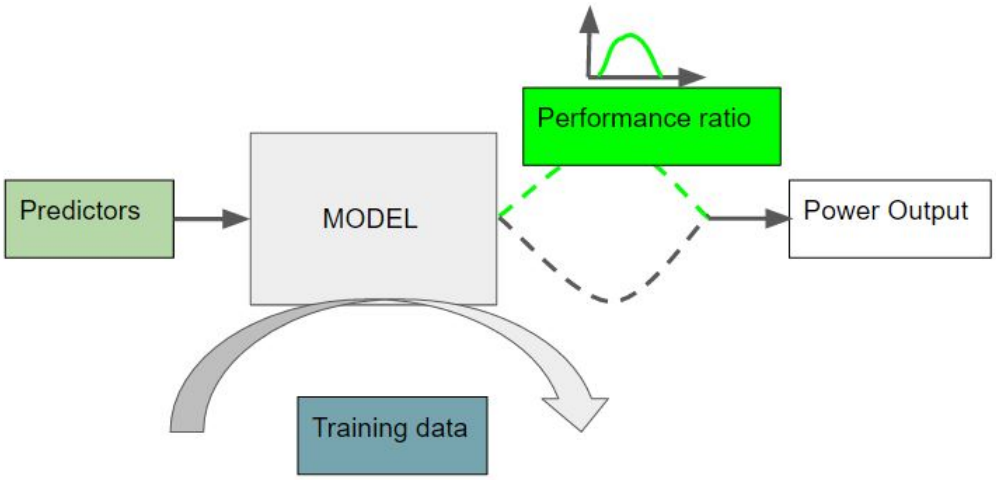


Figure 6.1: Three features of a short term PV power forecast model

6.1 Volatility

6.1.1 How to measure volatility

Given a daily PV power curve, we would like to give a formal definition of volatility. A basic idea is that the volatility should capture variation in power output that cannot be explained by daily variation due to the sun's position. We use a term **variability** separately for a variation due to the sun's position. The variability is already captured in a shape of a PV maximum power curve. Hence, the volatility of a PV power curve from uniformly performing PV systems should be small. In addition, volatility should have units and magnitudes that are useful to grid operators. We then want the volatility to be additive and scalable upon the installed capacity of PV systems. To summarize, the definition of the volatility should satisfy the following properties.

1. Volatility of any part of well behaved PV power curves should be small. Similarly, volatility of any part of maximum PV power curves should be small.

2. Volatility of a daily PV power curve with a low energy generation should be small.

3. Volatility of a union of two sections of a PV power curve should be equal to the sum of values of each section's volatility.

4. Volatility of a scaled PV power curve should be equal to the volatility of a normalized PV power curve multiplied by the scaling factor.

Following are several definitions of volatility in the literature that do not satisfy at least one of these desirable properties. Hoff and Perez (2010) [37] define volatility by the sample standard deviation of changes in the generated power values, or ramp rates, between neighboring time stamps:

$$V = \frac{1}{C^{group}} SD\{\sum_{n=1}^{N} \Delta P_n\},$$

where $SD\{x\}$ is the sample standard deviation of the quantity x. ΔP_n is a difference of PV outputs between neighboring time stamps of the n-th system. N is the number of PV systems in the group and C^{group} is the total installed capacity of a group of PV systems. The sample standard deviation is computed from a set of neighboring time-stamp pairs in a given time interval. Using our definition of normalized power curves, we may rewrite this as

$$V = SD\{\Delta p\},$$

where now Δp is a difference of PV normalized outputs between neighboring time stamps.

This definition has a serious shortcoming since the volatility of well behaved or maximum PV power curves is deemed highly volatile. Goldstein et al (2013) [6] rectified this shortcoming by considering the difference between changes in the generated power values and the expected changes in the generated power values. In our formulation, it is equivalent to

$$V = SD\{\Delta p - E(\Delta p)\},$$

where $E(\Delta p)$ is the expected changes in the generated power values. Goldstein et al compute such values by averaging Δp for each month. This definition also has a shortcoming since the volatility of a daily PV power curve with low energy generation is deemed volatile. We need a new definition to overcome this issue.

One way is to consider higher-order differences rather than first-order difference. Higher-order differences correspond to higher-order derivatives. From previous chapters, we know that maximum PV power curves and expected PV power curves can be

approximated by splines of degree-3 polynomials. Hence, we expect that the fourth derivatives of those curves are negligible. This leads to a derivative-based definition of volatility:

$$V = SD\{\Delta^4 p\},$$

where $\Delta^4 p_i = p_{i+2} - 4p_{i+1} + 6p_i - 4p_{i-1} + p_{i-2}$ and p_i refers to a power reading at time stamp i. By this definition, we will have low values of volatility for both well behaved PV power curves and PV power curves with low energy generation. Sudden changes in power outputs give high values of the fourth order differences and makes high values of volatility. However, this definition still has a shortcoming: it does not have units and magnitudes that are useful to grid operators. The first-order difference corresponds to the ramp rate, the second-order corresponds to the acceleration rate, but the fourth-order difference is difficult to interpret.

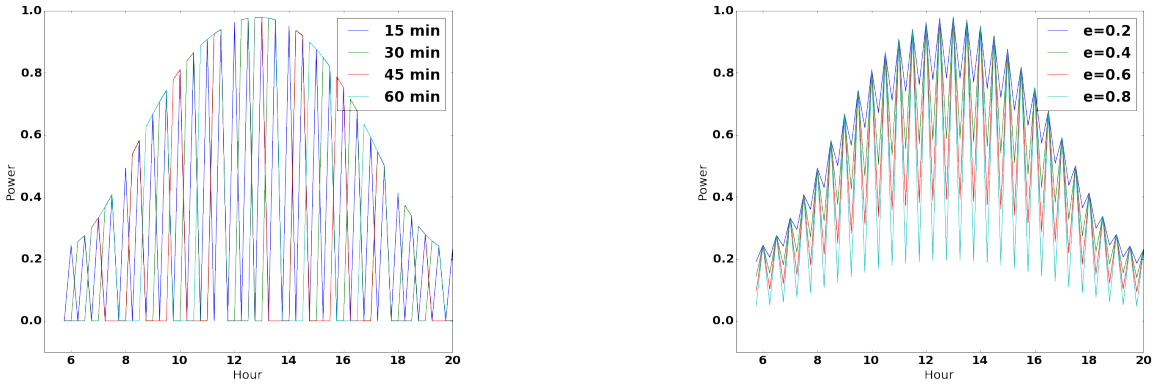
We may define volatility in a more sensible way using normalized maximum PV power curves and performance ratios as

$$V = \sum |P_i - r_{i-1} p_i^{max} C^{group}| = C^{group} \sum |p_i - r_{i-1} p_i^{max}|,$$

where P_i (p_i) are power outputs of (normalized) power curves power curves at time stamp i, and r_{i-1} is a performance ratio at the previous time stamp $i-1$. Note that there are several differences between this definition and previous versions. First, the definition uses a sum instead of a standard deviation. Hence, this definition satisfies property 3. Second, the definition has an explicit factor C_{group} in the equation. Hence, this definition satisfies property 4 in the check list. Finally, this definition of volatility looks into differences between actual power outputs and hypothetical power outputs if the latest performance retains. Any power curves that persist on a fixed performance have a low volatility. Hence, this definition satisfies property 1 and 2.

To simplify the discussion, we can define normalized volatility as $v = V/C^{group} = \sum |p_i - r_{i-1} p_i^{max}|$. We can convert it back if needed by scaling with the total installed capacity of a group of PV systems.

In the following subsections, we investigate further the volatility under this definition. In the first two subsections, we illustrate quantitatively that our volatility has the desirable properties addressed earlier by measuring the volatility of artificial data sets where a human assessment of volatility would be obvious. In addition, we show that the normalized volatility measures from two distinct geographical areas are likely to be different, while the normalized volatility measures from two neighboring geographical areas are likely to be the same. In the next two subsections, we show how the volatility depends on two quantities: a number of PV systems in a calculation and a frequency of PV power readings. In the last section, we investigate how the volatility depends on the time of day and the date.



(a) Set I

(b) Set II

Figure 6.2: Normalized power curves from different sets

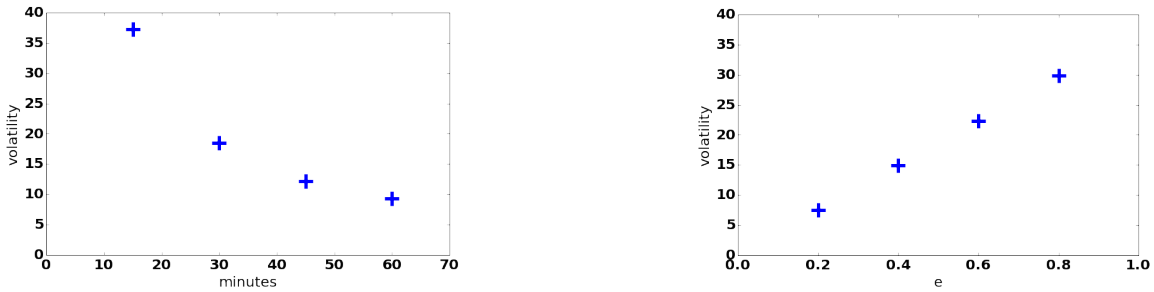
6.1.2 Demonstration of consistency using artificial data

In order to show quantitatively that our definition of volatility indeed gives reasonable values, we measure values of volatility from different artificial PV power curves, sort them according to these values, and compare their ranks with how humans rank those PV power curves according to our intuition of volatility. We consider two sets of different PV power curves:

1. Set I: a set of power curves with alternating values between zero and maximum every 15, 30, 45 and 60 minutes.

2. Set II: a set of power curves with alternating values such that a performance ratio oscillates between $1-e$ and 1 every 15 minutes. Here we use $e = 0.2, 0.4, 0.6$ and 0.8.

We select a maximum PV power curve on a day in Summer 2014 from the data set in CA. Then we construct two sets of different PV power curves from the power curve shown in Figure 6.2. The normalized volatility of each curve is shown in Figure 6.3. In the set I, judging by a human interpretation, a power curve with a higher frequency of oscillation has a higher value of volatility. Our measure agrees with this fact. Finally, in the set II, judging by a human interpretation, a power curve with a higher amplitude of oscillation has a higher value of volatility. Again, our measure agrees with this fact. Hence, all different power curves here support the validity of this definition of volatility.



(a) Set I (b) Set II

Figure 6.3: Values of volatility from different power curves

6.1.3 Variation of volatility with geographical areas

Another way to show by numbers that our definition of volatility indeed gives reasonable values is that, we measure values of normalized volatility from various geographical areas. We expect that the normalized volatility measures from two distinct geographical areas are likely to be different, while the normalized volatility measures from two neighboring geographical areas are likely to be the same.

To compare distinct geographical areas, we measure values of normalized volatility of daily power curves from the data set in CA and the data set in NJ. Figure 6.4 compares the values with a scatter plot where each point represents a day in a year. We found that there is no correlation between values of volatility from two data sets. The value of volatility in the data in NJ is generally higher than the value of volatility in the data set in CA. This agrees with the fact that the weather in New Jersey is less predictable than the weather in California.

To compare neighboring geographical areas, we use another set of power readings data in a neighboring zip code. Figure 6.5 compares the values of volatility from an original data set and a neighboring data set using a scatter plot. Black diagonal lines in the scatter plots are guidelines for equal values of volatility. We found that, unlike Figure 6.4, there is a high correlation between values of volatility from two neighboring data sets. This agrees with our expectation.

6.1.4 Variation of volatility with the number of PV systems

Here we show how the volatility depends on a number of PV systems in a calculation. Similar to our study of how a higher number of PV systems leads to a convergent power curve for a group of PV systems, we expect to see that a higher number of PV systems leads to a convergent value of volatility for a daily power curve as well. In order to show this quantitatively, we consider the data set in NJ from 241 systems with a capacity less than 20 kW. First, we use data to generate normalized maximum

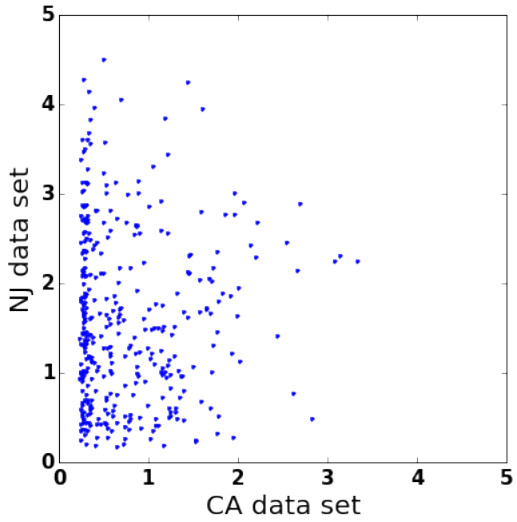
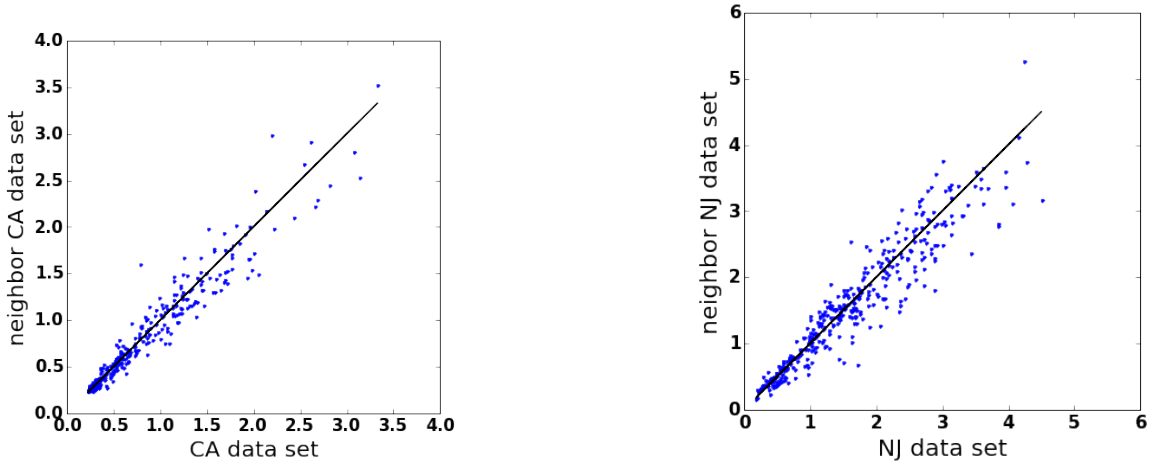


Figure 6.4: Comparison between values of volatility from daily power curves from two data sets

PV power curves as a reference for performance ratios computation. Then we group data into following sets:

- Set A-1: data from the 1st to the 10th system
- Set A-2: data from the 11th to the 20th system
- Set B-1: data from the 1st to the 20th system
- Set B-2: data from the 21st to the 40th system
- Set C-1: data from the 1st to the 40th system
- Set C-2: data from the 41st to the 80th system
- Set D-1: data from the 1st to the 80th system
- Set D-2: data from the 81st to the 160th system
- Set E-1: data from the 1st to the 160th system

Note that sets with same letters but different numbers are mutually exclusive. In addition, the union of sets with same letters equal to a set with the next letter and the number 1. We generate normalized power curves from these sets and compute the values of normalized daily volatility. Figure 6.6 shows the average values of volatility



(a) CA (b) NJ

Figure 6.5: Comparison between values of volatility from daily power curves in neighboring data sets

from all daily power curves in 2014 for different sets. Each blue dot represents a value for each set, and black lines connect averages of two groups with the same number of systems. One can see that the volatility of a union of two sets is always less than or equal to an average of the values of volatility from individual sets. In addition, as a number of systems increases, a volatility decreases and converges. These observations agree with our expectation.

6.1.5 Variation of volatility with the frequency of PV power readings

Here we show how the volatility depends on a frequency of PV power readings in a calculation. For the data sets in CA and NJ, the frequency of PV power readings is 15 minutes, which is common in solar PV monitoring practices. However, a volatile change in a PV generation is likely at a higher frequency. We look for PV generation data with a higher frequency and found a 5-month long data of a single PV system with a frequency of one minute. We call this data set as the Stanford data set.

From the Stanford data set, one may calculate a value of volatility for each daily PV power curve from power readings every 1, 2, 5, 10, 15, 30 and 60 minutes. Each line in Figure 6.7 shows a value of daily volatility of a PV power curve at different frequencies. As the frequency of PV power readings increases, the value of volatility increases. In addition, as the frequency of PV power readings increases, the gap between values of volatility of less volatile and more volatile days expands. Such an

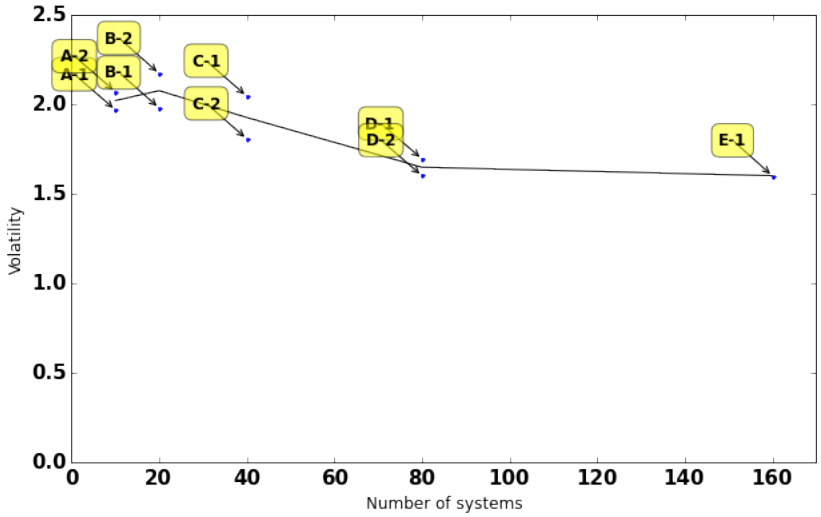


Figure 6.6: Average values of volatility from all daily power curves in 2014 generated from different number of systems

increase comes from the refinement in measuring a smooth change in power outputs. It is analogous to measure a length of a smooth curve with a smaller ruler. As we use a smaller ruler, we may increase the chance of observing changes in a smaller time period. As the size of the ruler goes to zero, according to our intuition of a physical process over time, we expect our measurement of power output's volatility converges to a finite value.

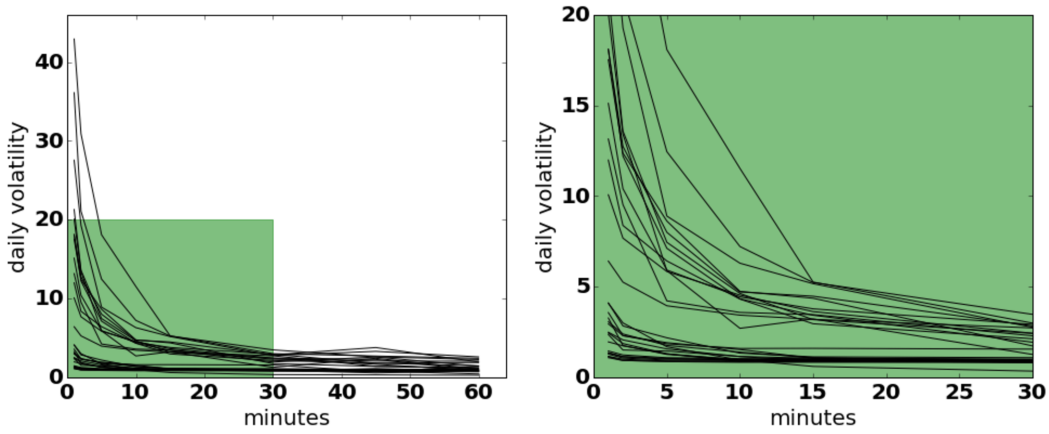


Figure 6.7: Values of daily volatility of different daily PV power curves with different frequency of power readings.

The question is what should be a sufficient frequency so that the value of volatility captures a reasonable amount of sudden changes that are problematic for a grid operator. There is no clear answer but one may guess by looking at a value of daily volatility divided by a number of readings. As a frequency decreases to zero, we expect that the value of daily volatility divided by the number of readings decreases to zero since the daily volatility converges to a finite value. Figure 6.8 shows such smooth decreasing trend for all days when the frequency is below 5 minutes. The transition from frequency higher than 5 minutes is not smooth and even increasing. This is because, as the frequency decreases, changes in the smaller time period in power outputs are discovered. We do not see it below 5 minutes. Hence, we claim that a frequency of 5 minutes is sufficient to capture a reasonable amount of volatility in power outputs. This frequency can be different for a different data set.

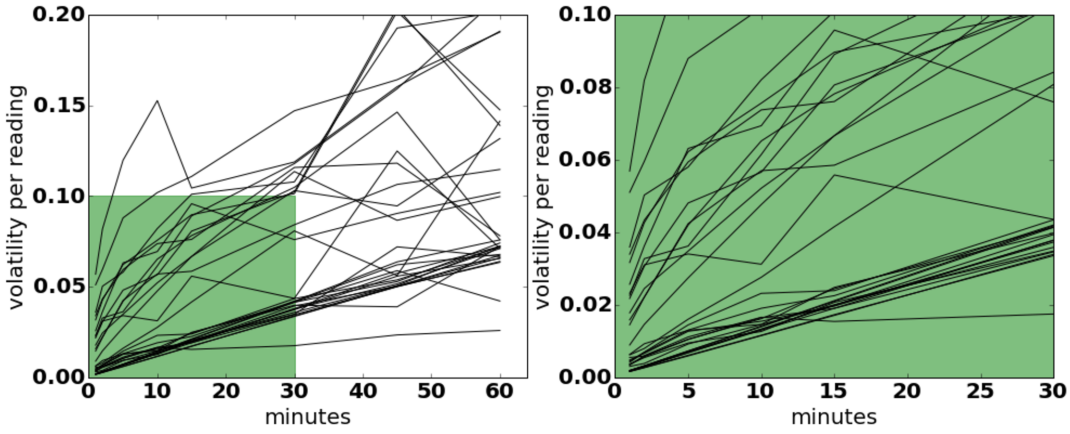
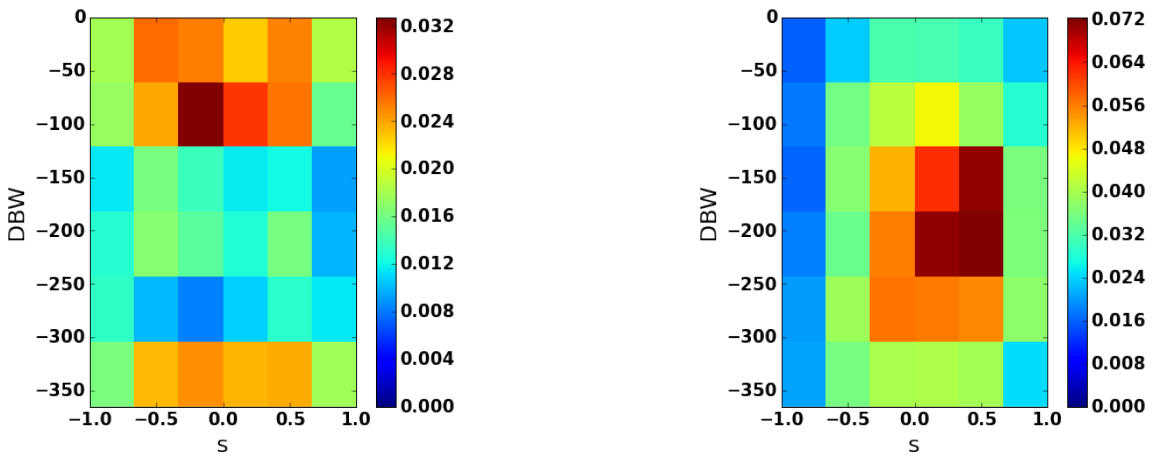


Figure 6.8: Values of daily volatility of different daily PV power curves divided by numbers of power readings.

6.1.6 Variation of volatility with the time of a day and the day of a year

It is worthwhile noting how the volatility depends on the time of a day and the date in the measurement. We can measure values of volatility of each time stamp and consider them in buckets as we have done with performance ratios. Figure 6.9 shows the mean volatility for different buckets of (s, DBW) where s is a parameter associated with the time of a day from a sun rise ($s = -1$) to a sun set ($s = 1$), while DBW is a parameter associated with the date from a day after a winter solstice ($DBW = -364$) to the next winter solstice ($DBW = 0$). In both data sets, the volatility near sunrises and sunsets tends to be low, this is because the PV power



(a) CA

(b) NJ

Figure 6.9: Mean value of the volatility for each bucket of (s, DBW) in 2014

generation itself is low. However, two data sets have different patterns of volatility over days of a year. For the data set in CA, the volatility in the winter is higher than the volatility in the summer. For the data set in NJ, on the contrary, the volatility in the summer is higher. This information is useful when we are planning on ancillary services in an hour-ahead operation.

6.2 Short-term prediction models

We investigate some short-term prediction methods for PV power outputs. There are several features of prediction models that we can investigate. In the first feature, we investigate whether a prediction of power via a prediction of performance ratios is better than a direct prediction of power. In the second feature, we investigate what is the optimal set of predictors used in the model. In the last feature, we investigate what is a sufficiently good size of training data.

6.2.1 Definition of performance of prediction models

We frequently use a concept of a prediction model's performance. There are many ways to construct a measure of this concept from differences between predicted power outputs and actual power outputs. Two simple measures are the mean and the maximum of those differences. In some circumstances, a measure with more detail is required in order to develop a meaningful policy based on the forecast performance[1].

[1]An example of the policy is the PV curtailment in Chapter 10

Since we have a formal definition of volatility of power outputs over a time period, we may break down a model's performance according to the level of hourly volatility. First, for each time stamp, we calculate a value of volatility over an hour before the time stamp (most recent 3 time stamps). Then we use a prediction model and calculate a difference between a predicted power output and a true power output at that time instance. After that, we set up bins with different ranges of volatility and calculate the mean and the maximum of differences in each bin. We call the outcome of this procedure a **prediction performance curve**. The prediction performance curve gives a broader ground for comparison and gives an insight on how to switch back and forth between prediction models in order to maximize the performance of a forecast system using an ensemble of several models.

As an example of a prediction power curve construction, consider the **zeroth method**, which is to predict the power output from the day-ahead scheduling without any update. As addressed in the chapter on applications of maximum PV power curves, we may use an expected PV power curve as a power curve in a day-ahead scheduling. In this simple short-term prediction method, we predict that the power output in the next time instance, which is the next 15 minutes in this data set, is equal to the power output as the expected PV power curve. That is, we do not use actual ongoing power outputs to train the model at all.

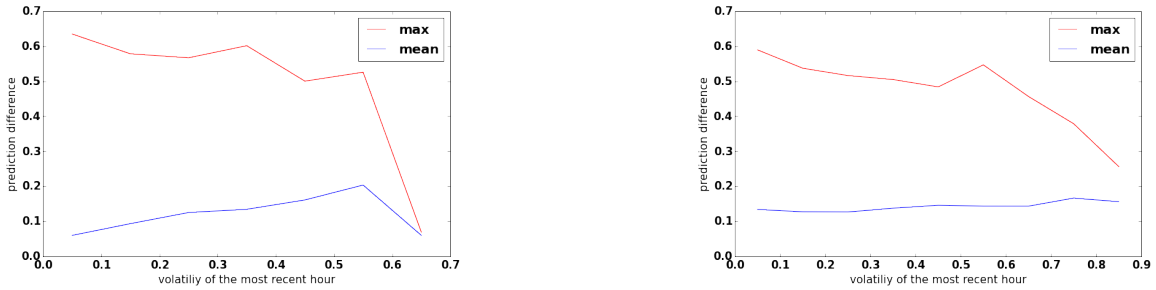
We test this prediction method with the data sets in CA and NJ in 2014 and compute performance measures. It turns out that the mean and the maximum of prediction differences are equal to 0.0659 and 0.6349 for the data set in CA, and 0.1323 and 0.5896 for the data set in NJ. Figure 6.10 shows prediction performance curves of two data sets. Here one can see that the mean of prediction differences correlates with the volatility of the most recent hour. An outlier, which is the last point in the data set in CA is likely because there is a small number of values in a (0.6,0.7) bin.

It is worth noting that, under this prediction method, the maximum prediction differences likely occur when the power output is extremely low. Such occurrence can happen even when the volatility of the recent hour is low because we never change our prediction from the expected value. Hence, we do not observe a strong correlation between maximum prediction difference and the volatility of the recent hour.

In the next subsection, we will adopt these prediction model's performance measures to compare prediction models with various features.

6.2.2 A prediction of power via a prediction of performance ratios

Our end goal of a short-term PV power forecast system is obviously to predict the value of a PV power from a group of PV systems in a future time instance. However,



(a) CA (b) NJ

Figure 6.10: Prediction performance curves for the zeroth method with data sets in 2014

one can predict a future actual power, p, directly, or one can first predict a future performance ratio, r, to obtain p. We compare performances of a prediction model predicting a future actual power directly and a prediction model first predicting a future performance ratio. To avoid complexity from other features of prediction models, we will focus on three simple methods that do not require a training set:

- Method 1: a persistent method [2] $x_{i+1} = x_i$
- Method 2: an extrapolation method $x_{i+1} = x_i + (x_i - x_{i-1}) = 2x_i - x_{i-1}$
- Method 3: a moving average $x_{i+1} = \frac{1}{2}(x_i + x_{i-1})$

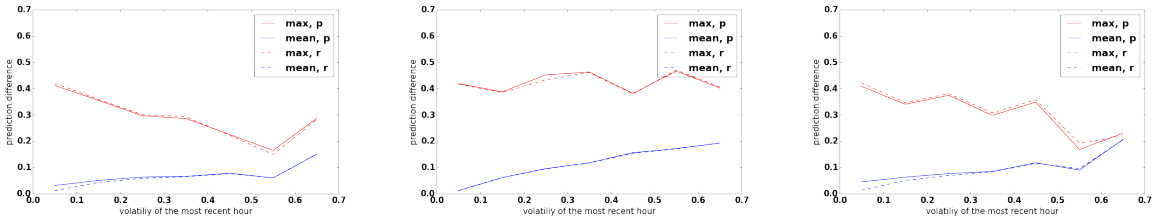
Here x can be either p or r. Note that we add an extra condition $0 \leq p_{i+1} \leq p_{max,i+1}$ or $0 \leq r_{i+1} \leq 1$ to prevent overestimated and underestimated prediction. This overestimated prediction can be prevented because of the existence of maximum power curves. Table 6.1 compares the mean and the maximum of prediction differences using different variables among three methods addressed above.

From the table, one can see that the models predicting a performance ratio, r, give significantly lower mean prediction differences in general. However, there is no significant difference between maximum prediction differences. From prediction performance curves in Figure 6.11 and 6.12, one can see that predicting a performance ratio (dash lines) reduces the mean prediction differences especially when a volatility of a recent hour is low. It does not clearly reduce the maximum prediction differences. Based on these observations, we conclude that, to achieve a better prediction performance, at least in terms of the mean prediction differences, one should first predict a future performance ratio, r, then convert it to an actual power.

[2]The term "persistent" illustrates the idea of the method where the predicted value is the same as the current value. This concept is similar to the persistence forecast discussed in the literature but the detail of the methods are slightly different [31, 67].

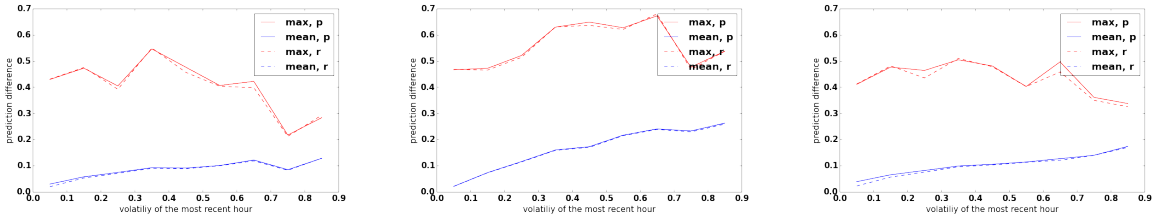
CA	Predict p directly	First predict r then p
Method 1	(0.034, 0.412)	(0.016, 0.420)
Method 2	(0.021, 0.466)	(0.020, 0.470)
Method 3	(0.048, 0.408)	(0.019, 0.422)
NJ	Predict p directly	First predict r then p
Method 1	(0.043, 0.546)	(0.035, 0.549)
Method 2	(0.050, 0.672)	(0.050, 0.680)
Method 3	(0.052, 0.505)	(0.039, 0.510)

Table 6.1: Value comparison of performance measure (the mean prediction difference, the maximum prediction difference) for three different forecasting methods for CA and NJ data sets.



(a) Method 1 (b) Method 2 (c) Method 3

Figure 6.11: Prediction performance curves of the data set in CA for three methods with different variables to predict.

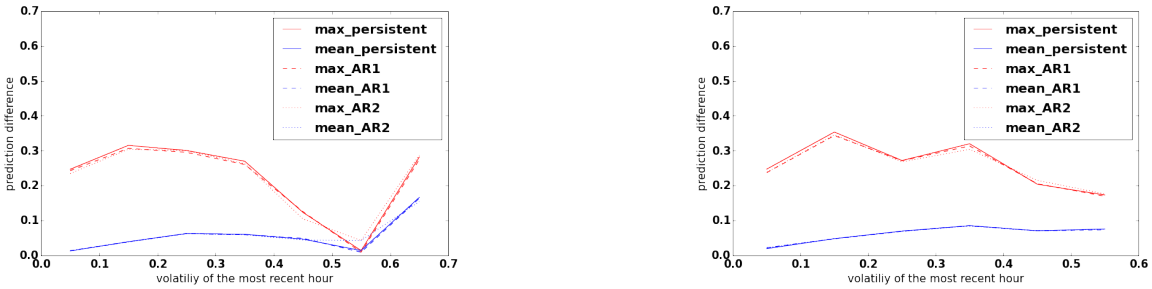


(a) Method 1 (b) Method 2 (c) Method 3

Figure 6.12: Prediction performance curves of the data set in NJ for three methods with different variables to predict.

CA	mean	max
Persistent	0.018	0.315
AR(1)	0.019	0.307
AR(2)	0.018	0.304
NJ	mean	max
Persistent	0.032	0.354
AR(1)	0.033	0.343
AR(2)	0.033	0.346

Table 6.2: Value comparison of performance measure from different simple methods



(a) 'CA' data set

(b) 'NJ' data set

Figure 6.13: Prediction performance curves for simple methods

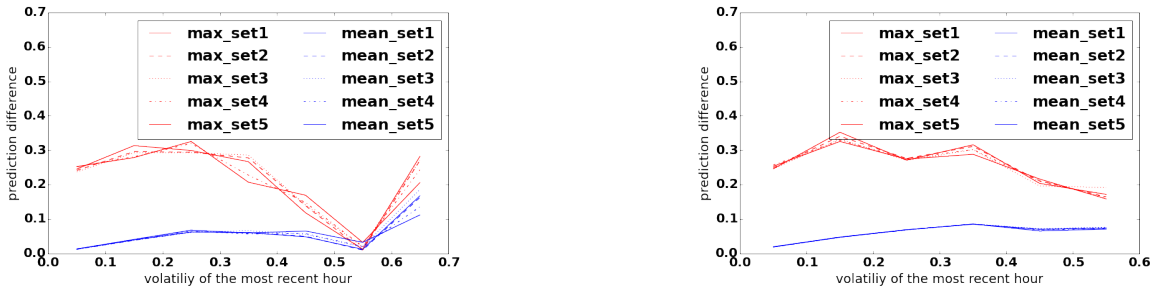
6.2.3 Predictors of prediction algorithms

We investigated whether models with different predictors gives significantly better performance than models based on previous power outputs only. First, we consider three simple methods that rely on previous power outputs only including a persistent method used in the previous subsection, an auto-regression of order 1 (AR(1)), and auto-regression of order 2 (AR(2)).

To make a fair comparison, we measured the performance of all methods on the test set consisting of data on the last quarter of 2014 and used the remaining data as a training set if needed (the persistence method does not use any training set).

In Table 6.2 and Figure 6.13, we compared performances of three simple algorithms: persistent, AR(1) and AR(2). One can see that auto-regressive models do not perform better than the persistent model. There are two implications. First, previous power outputs other than the most recent one may not be necessary. Second, we may need a non-linear model for a better performance.

To improve the prediction performance, we may add additional predictors into the model including s and DBW of the time instance we want to predict. We may



(a) CA (b) NJ

Figure 6.14: Prediction performance curves for the SVR with different sets of predictors

also add some weather information of the recent time instance in the model as well. Each quantity in weather information is normalized so that the minimum value in the data is zero and the maximum value in the data is one. DBW is also normalized so that the range is $[-1, 0]$. There is no need to normalize s since its range is already between -1 and 1. We may add V, the hourly volatility of the recent hour into the model as well. This leads to five sets of predictors as follow:

- $\{r_{i-1}\}$
- $\{r_{i-1}, s, DBW\}$
- $\{r_{i-1}, s, DBW, V\}$
- $\{r_{i-1}, s, DBW\} \cup$ Weather
- $\{r_{i-1}, s, DBW, V\} \cup$ Weather

Due to the complexity of how different predictors might affect the prediction, a non-linear model is required.

First, we look at a Support Vector Regression (SVR)[63]. In Table 6.3 and Figure 6.14, we compare performances of the SVR with different sets of predictors after tuning SVR parameters C and e. From both data sets, adding s and DBW in the set of predictors improve overall performance. Adding weather information and V does not improve the performance for the data set in CA, but it improves the performance for the data set in NJ. Note that the main benefit of complex models with more predictors is to reduce the maximum prediction difference rather than the mean prediction difference.

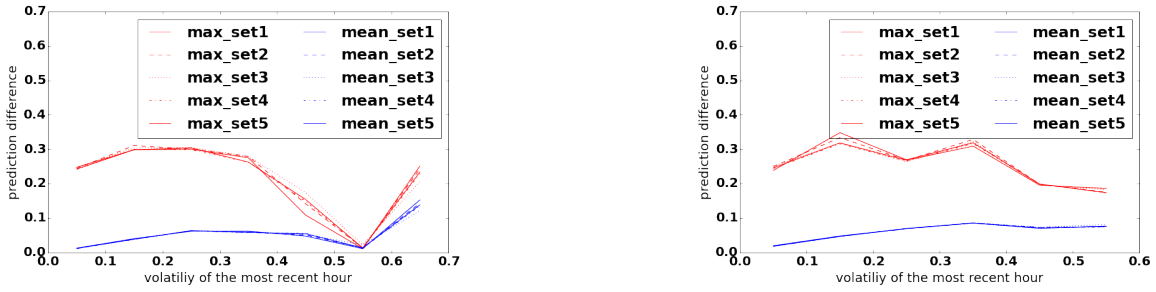
Next, we look at a Gradient Boosted Regression (GBR)[64]. In Table 6.4 and Figure 6.15, we compare performances of the GBR with various set of predictors after

CA	mean	max
$\{r_{i-1}\}$	0.018	0.313
$\{r_{i-1}, s, DBW\}$	0.017	0.296
$\{r_{i-1}, s, DBW, V\}$	0.017	0.293
$\{r_{i-1}, s, DBW\} \cup$ Weather	0.019	0.319
$\{r_{i-1}, s, DBW, V\} \cup$ Weather	0.019	0.325
NJ	mean	max
$\{r_{i-1}\}$	0.032	0.352
$\{r_{i-1}, s, DBW\}$	0.031	0.340
$\{r_{i-1}, s, DBW, V\}$	0.031	0.332
$\{r_{i-1}, s, DBW\} \cup$ Weather	0.031	0.329
$\{r_{i-1}, s, DBW, V\} \cup$ Weather	0.031	0.325

Table 6.3: Value comparison of performance measure from the SVR with different sets of predictors

tuning GBR parameters including a number of estimators, a depth, and a learning rate. From both data sets, adding s and DBW in the set of predictors improves overall performance. Adding weather information does not always improve the performance for the data set in CA, but it improves the performance for the data set in NJ. We observe again that the main benefit of models with more predictors is to reduce the maximum prediction difference rather than the mean prediction difference.

At this point, we see that the GBR performs better than the SVR and other simple models in general. With a different set of predictors of each data set, we can bring both the mean and the maximum prediction differences down by about 10 percent



(a) CA (b) NJ

Figure 6.15: Prediction performance curves for the GBR with different sets of predictors

CA	mean	max
$\{r_{i-1}\}$	0.018	0.300
$\{r_{i-1}, s, DBW\}$	0.017	0.307
$\{r_{i-1}, s, DBW, V\}$	0.017	0.307
$\{r_{i-1}, s, DBW\} \cup$ Weather	0.017	0.302
$\{r_{i-1}, s, DBW, V\} \cup$ Weather	0.018	0.304
NJ	mean	max
$\{r_{i-1}\}$	0.032	0.348
$\{r_{i-1}, s, DBW\}$	0.031	0.333
$\{r_{i-1}, s, DBW, V\}$	0.031	0.331
$\{r_{i-1}, s, DBW\} \cup$ Weather	0.030	0.318
$\{r_{i-1}, s, DBW, V\} \cup$ Weather	0.030	0.318

Table 6.4: Value comparison of performance measure from the GBR with different sets of predictors

relative to the persistent model. In the next subsection, we use the GBR method to study the effect of training-data size.

6.2.4 Size of training data

In a **batch** approach, we train a prediction model from fixed training data, then we apply that prediction model to any data point we would like to predict. This is the approach we have used in previous subsections. However, in practice, there are new data points that we can use to improve our prediction model. This raises a question of what should be the size of training data in this on-line approach. One may keep adding to training data while doing prediction, or one may limit training data to a certain number of most recent data points.

To answer this question, we start with the best model we have at hand. For the data set in CA, we use a GBR method to predict the next performance ratio first based on r_{i-1}, s, DBW, and V. Then we obtain the actual load from the ratio. For the data set in NJ, we have a similar model weather information is included as predictors as well. We consider five schemes for an on-line approach:

- There is no limit on size of training data. Keep adding new data points of a day in the data set and retrain every day.
- Limit training data to be at most 12 months of data. For the remaining scheme, we add new data points of a day, delete data points of the oldest day, and retrain every day.

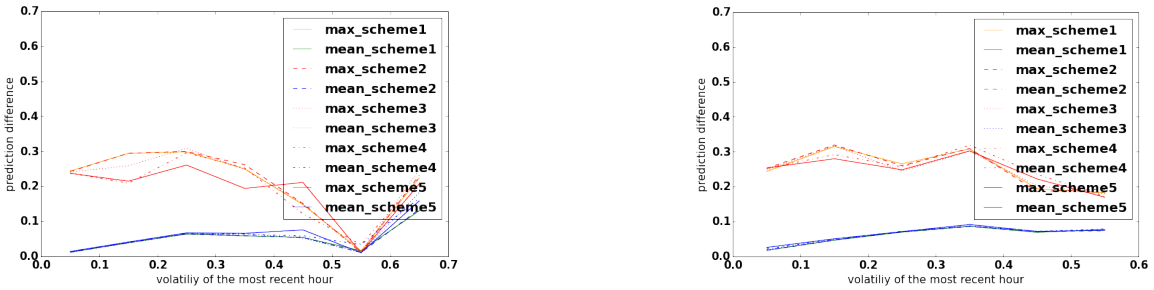
CA	mean	max
No limit	0.017	0.298
12 months	0.017	0.301
6 months	0.017	0.309
3 months	0.017	0.307
1 month	0.019	0.259
NJ	mean	max
No limit	0.030	0.313
12 months	0.030	0.319
6 months	0.032	0.314
3 months	0.033	0.320
1 month	0.037	0.300

Table 6.5: Value comparison of performance measures from an on-line approach with different sizes of training data

- Limit training data to be at most 6 months of data.
- Limit training data to be at most 3 months of data.
- Limit training data to be at most 1 months of data.

All schemes are applied to the same test set, which is data of the last quarter in 2014. We choose to update the training data every day instead of every new data point (15 minutes) because the GBR is computationally expensive. In practice, an update can be done for every new data point since there is enough time between arrivals of new data points. It might be worthwhile updating more frequently, but we did not investigate this option.

In Table 6.5 and Figure 6.16, we compare performances of on-line prediction models with various sizes of training data. From both data sets, adding new data points of a day in the data set without deleting gives the best mean performance measure. Even though there is no observable clear trend on the maximum prediction difference, the mean prediction difference is monotonically decreasing as the training-size data increases. It means that the optimal amount of data to produce the best performed model is more than 1 year of data. Larger size of training data is not investigated because our data is limited.



(a) CA (b) NJ

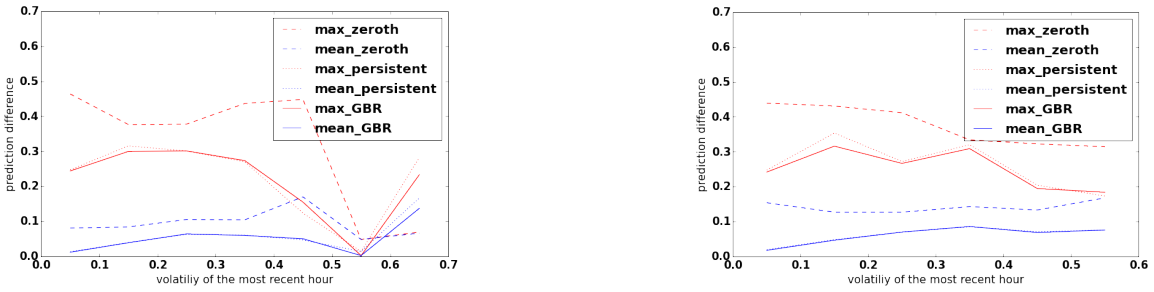
Figure 6.16: Prediction performance curves from an on-line approach with different sizes of training data

6.3 A forecast system: ensembles of prediction models

In the previous section, we have found that for each prediction model the performance depends on a use of indirect prediction via performance ratios, a set of predictors, and the size of training data. We have established from our two data sets that it is good to first predict a performance ratio then obtain an actual power. We also established that the size of data should be as large as possible in our data set. For a set of predictors, for the data set in CA we use a performance ratio of the most recent instance, s, DBW and the hourly volatility of the most recent hour. For the data set in NJ, we include weather information as well. It is worthwhile comparing prediction performance curves of three algorithms: the zeroth method, the persistent method, and the GBR with predictors and sizes of training data as described above. Figure 6.17 shows that the GBR performs better than other methods except for the data set in CA where the volatility of a recent hour is higher than 0.6. Due to a low number of data points in which the volatility of a recent hour is higher than 0.6, it is difficult to conclude that such trend is generalizable. Still, we will assume that the GBR performs worse than a simple method like the zeroth method when the volatility of a recent hour is higher than 0.6. This leads to a protocol to switch between the GBR and the zeroth method when the volatility of a recent hour passes 0.6.

To conclude our short-term forecast system of PV power output, its procedure is listed below.

For the data set in CA, since weather information is not needed, we just need to keep track of GBR models with two sets of predictors: the 2nd and the 3rd sets. For the data set in NJ, since we use weather information, we keep track of GBR models with four sets of predictors: the 2nd, the 3rd, the 4th and the 5th sets. We follow these instructions, start a batch training with all data except the last quarter

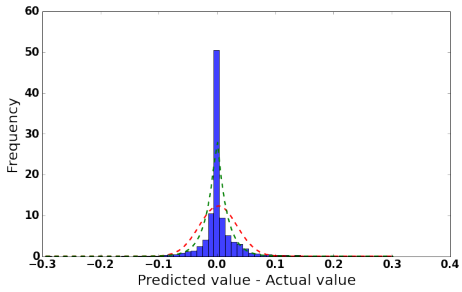


(a) CA (b) NJ

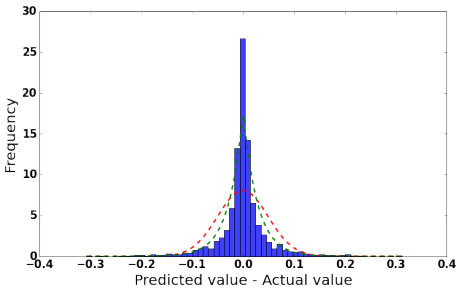
Figure 6.17: Prediction performance curves from the zeroth method, the persistent method, and the GBR

Algorithm 3 Short-term forecast system for PV power output

1: At sun rise, there is no performance ratio to use as a predictor it the first place, so the zeroth method is used.
2: From the next time stamp to the following hour, there is no volatility of a recent hour as a predictor, so the GBR model without volatility is used.
3: After that, if the volatility of a recent hour is less than 0.6, the GBR model with a full set of predictors as discussed earlier is used, otherwise use the zeroth method is used.
4: If weather information is not available, the GBR model without weather information is used.
5: After sun set, add data points of the day into a training set and retrain all models.



(a) 'CA' data set



(b) 'NJ' data set

Figure 6.18: Distribution of prediction differences from a short-term forecast system. The histogram shows actual data. A red line shows a fit of a normal distribution and a green line shows a fit of a Laplace distribution.

of 2014, and perform an on-line training with data from the last quarter of 2014. If we consider basic performance measures, we found that the mean prediction difference equals to 0.0165 for the data set in CA and 0.0281 for the data set in NJ. The maximum prediction difference equals to 0.2981 for the data set in CA and 0.3157 for the data set in NJ. To get a sense of performance relative to the actual power, one may compute the mean power output in the test set. It turns out that the mean power output in the test set is 0.5225 in the data set in CA and 0.3860 for the data set in NJ. It means that the mean error is at most 9 percent, while the maximum error can be as large as 82 percent.

To tell how likely a bad prediction will happen, a distribution of prediction differences is needed. Figure 6.18 shows a distribution of differences between predicted power outputs and actual power outputs on the test set with this short-term forecast system of PV power output. One can see that the distribution of such differences is a Laplace distribution rather than a normal distribution. One can fit a Laplace distribution and get the best fit with parameters $(\mu, b) = (-0.0001, 0.01652)$ for the data set in CA and $(\mu, b) = (-0.0002, 0.0281)$ for the data set in NJ. The bias of prediction is negligible. From this, one can calculate how likely a bad prediction will happen in general when using this short-term forecast system.

6.4 Conclusion

After an investigation in a short-term PV power forecast system, we found that there are some generalized results across different data sets. First, our concept of performance ratios is useful to constructing a good model. Next, our definition of volatility is helpful as a means of communicating the performance of a prediction models and a parameter to toggle between several prediction models in an ensemble.

However, there are several aspects of the system that needs to be customized for each data set such as adding weather information as predictors in a model. We need more data to give a clear procedure on the size of training data and to understand what really happens when the volatility is high.

Part II

Load Modeling

Chapter 7

Overview of load modeling

Highlights:

- Load pattern is simple compared to PV generation pattern. It depends on number of factors including time, location, weather, and human behavior.
- The load variability and volatility is critical to the cost of grid operation. Higher variability means higher cost of generation with high ramp rate. Higher volatility means higher cost of generation to correct differences between forecasted and actual load.
- There have been new techniques and improvement in load forecasting over the years. However, there are new challenges in the load model. Distributed PV generation as a part of net load is one of the challenges.

Modeling load is a necessary task for energy scheduling and grid operation. At the level of substation and beyond, modeling load is relatively simple. While load at the level of household is volatile, aggregation of loads at the level of substations and the grid is likely to be smooth and predictable. Despite significant differences from day to day, modeling load at substation is accurate provided contingencies do not arise. Here we will give an overview of the load pattern at substation and the grid. We will give an overview of different types of forecast needed in general operation and discuss a relevant forecast in our study in detail.

7.1 Energy and electricity consumption by sector

In order to model the load accurately, it is useful to understand where energy consumption and electricity consumption take place. Figure 7.1 [1] shows US energy consumption by sector, while Figure 7.2 [1] shows US electricity consumption by

sector in 2012. The consumption under building operations comes from about half residential and half commercial operations [24]. Traditionally, a schedule for electricity consumption from all sectors, especially residential operations, is quite rigid and bumpy because of human behavior. However, with new technologies including smart appliances in building operations and automated controls in the industry sector, a schedule for electricity consumption can be more flexible and flat. It is worth noting from two figures that even though the transportation takes more than a quarter of energy consumption, it is less than one percent of electricity consumption. A large potential for an electrification of the transportation sector is still available. Automated transportation system and energy storage is electric vehicles are potential technologies that make a schedule of electric consumption more flexible and flat as well. However, these technologies are not currently mature and widely used enough to alter the schedule of electricity consumption or the load pattern.

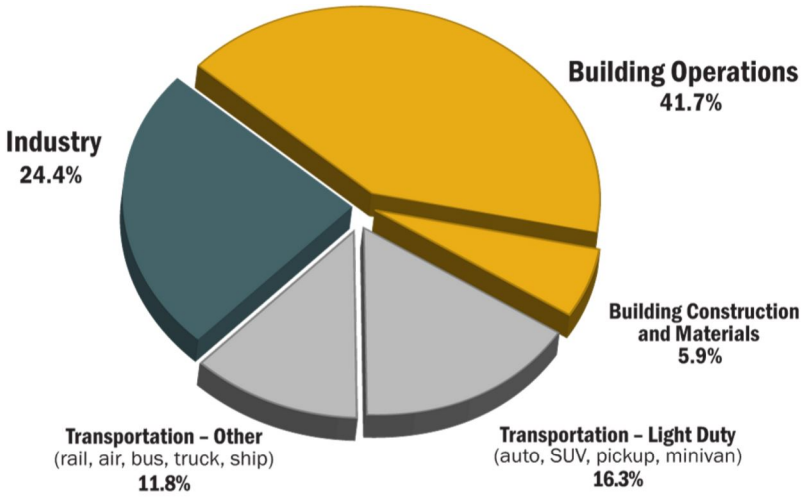


Figure 7.1: US Energy Consumption by Sector.

7.2 Load pattern

In order to understand a pattern of load, one may start plotting load with respect to time. Such plot for 24 hours is known as **daily load curve**. The shape of such curve depends on types of loads (residential, non-residential), time of the year, and geographical areas. The daily load curve can be drawn from single household to entire area under the control of an independent system operator (ISO). Here Figure 7.3 [57] shows an example of daily load curves in different seasons of PJM, an ISO in the east

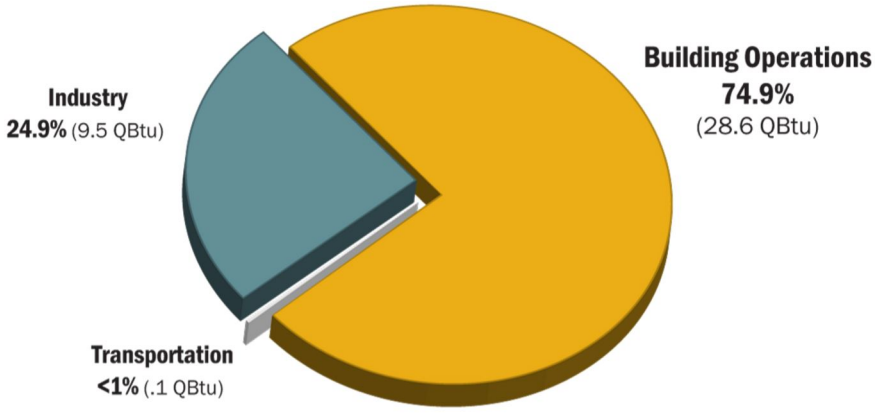


Figure 7.2: US Electricity Consumption by Sector.

coast. The orange curves in Figure 7.4 [25] shows an example of daily load curves in different seasons of CAISO, an ISO in the west cost.

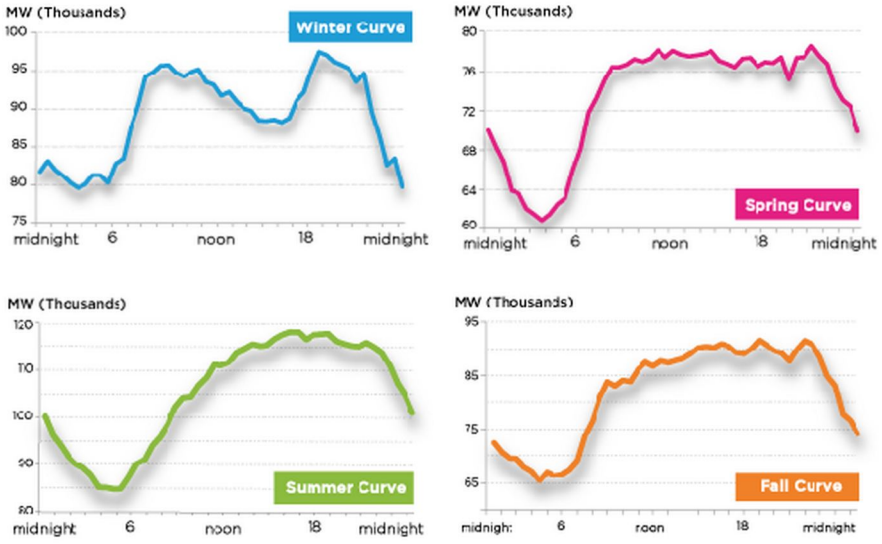


Figure 7.3: Daily Load Curve Example from PJM under different seasons.

In daily load curves from both areas of the US, one can see a pattern. First, in a daily basis, the load peaks from 7 AM to 10 PM. The load increases with a high ramp rate in the morning. The highest peak usually occurs in evening. Next, on a seasonal basis, the load peaks in the summer, is smaller in Winter, and is smallest in Fall and Spring. This is likely because weather during summer and winter in the US

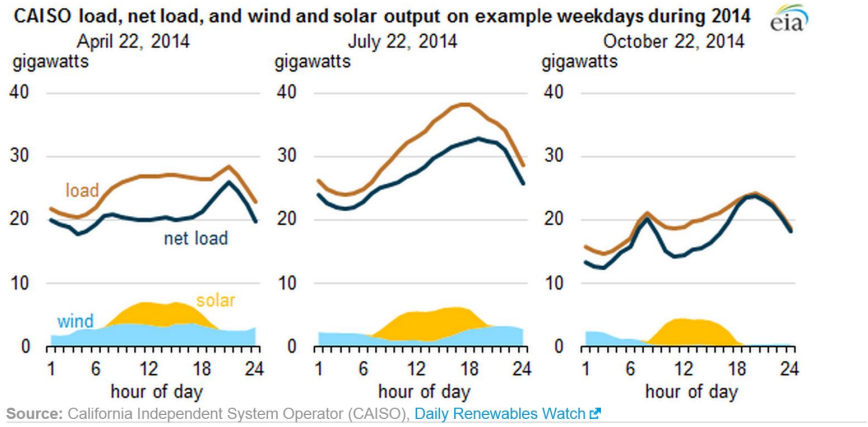


Figure 7.4: Daily Load Curve Example from CAISO under different seasons.

is so uncomfortable that people needs to use cooling or heating system. The load in winter tends to smaller than the load in summer because cooling systems mainly use electricity while heating systems mainly do not use electricity.

Another pattern of load, which cannot be seen from figures above, is in weekly basis. The load peaks during weekdays. We expect to see the load change significantly during holidays and major events as well.

The load pattern is highly dependent on the location. Weather, human behaviors, and human response to the weather change at a location attributes to the load pattern. As a result, unlike a PV system model, there is no generic formula to determine the load pattern. Historical load data in the location, weather data and time are only ingredients to build an accurate model to forecast the load.

7.3 Load variability is bad

Load variability is bad for grid operators in two ways. First, imagine an ideal scenario where the daily load curve is constant. In order to supply power the load, one needs only a power plant with capacity just equal to the load. Now suppose that the daily load curve has variability: one hump in the early morning and one peak in the evening. Suppose that this new daily load curve still has the same under the curve. That is, the total energy load over a day stays the same. Yet the same power plant cannot cope with the daily load curve. One needs to build a new power plant or upgrade the old plant to match the peak load. This leads to extra cost.

Next, imagine a similar load curve with but higher variability: the hump and peak has a shorter width. Comparing with the previous load curve, this daily load curve requires the same capacity and total energy. Yet the same power plant may not be

able to cope with such a daily load curve due to a high ramp rate. The current power plant may not increase or decrease fast enough to serve the energy need. One may need to build a new power plant or upgrade the old plant that is capable of reacting faster. Again, this leads to extra cost.

Because of load variability, the grid operator requires energy supply with higher capacity and ramp rate. This increases the cost of grid's operation. It is worthwhile noting that high the penetration of distributed PV power changes the daily load curve significantly and introduce more variability in term of the ramp rate as shown, for example, in Figure 7.5 [25].

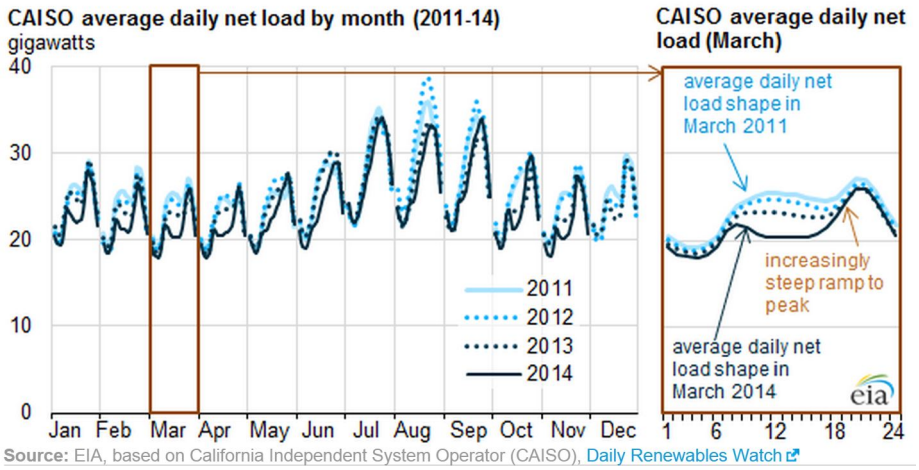


Figure 7.5: Average Daily Load Curve from 2011 to 2014

The load variability is not the only issue since the actual daily load curve can be different from the expected daily load curve. We will address this issue in the next section.

7.4 Load volatility is bad

Given perfect foresight of the daily load curve, one can plan day-ahead scheduling of energy supply to match the load perfectly. In the day of operation, demand and supply balances out perfectly. In reality, no one has the perfect foresight of the daily load curve. In the day of operation, one may perform an hour-ahead forecast and schedule extra energy supply or load to adjust demand and supply to be as close as possible. The hour-ahead energy scheduling involves highly responded instruments, which is more expensive than power plants in day-ahead scheduling. Within the hour of operation, one may perform a "real-time" forecast and schedule extra energy supply or load to readjust. The cost of real-time energy scheduling is even more expensive.

Now if the load has no volatility, we expect that with a horizon of 1 day we can predict the load perfectly and pay no cost for extra hour-ahead scheduling and real-time adjustment. If the load has low volatility, we expect that with a horizon as 1 hour we can predict the load well and pay no cost for extra real-time adjustment. However, if the load has high volatility, we expect to pay cost for both hour-ahead scheduling and real-time adjustment. Note that the cost has two dimensions: capacity and operational cost. The capacity cost is associated with maximum error in forecast and the operational cost is associated with mean absolute error in forecast.

Because of load volatility, the grid operator requires extra energy supply and ancillary service. This increases the cost of grid's operation. It is worth noting that PV power introduces both short term and long term volatility due to the unpredictability of weather.

7.5 Load forecast: types, history, trend, and accuracy

One can categorize load forecast based on forecast horizon: (1) long-term load forecasting (LTLF) with a forecast horizon of years or decades (2) short-term load forecasting (STLF) with a forecast horizon of days, hours or minutes.

The report by Tao Hong [38] is a comprehensive review of load forecast from Thomas Edison's time to present. Alfares and Nazeeruddin (2002) [4] gave a literature survey and classified methods for load forecast. Singh et al (2013) [7] gave a more compact and up-to-date classification for forecast techniques. As we are interested in grid's operation rather than planning, we will give a brief overview of a long-term load forecast and discuss more on a short-term load forecast.

In the pre-PC period, long-term load forecasting is conducted for the whole US with limited information and techniques. In the post-PC period, more detailed forecasting about when, where and how much the load growth will occur is conducted known as spatial load forecasting. The forecast can be done by regression analysis on load trend, simulation to reproduce historical load data and to project the future load, or the mixture of both. With new techniques from artificial intelligence, the mean absolute percent error (MAPE) for year-ahead forecast can be as low as 2% using just historical loads data [39]. In longer horizon, we still need simulations with additional adjustments. In 2009, California energy commission included detail of energy efficiency and electric vehicles in the long-term forecast, disaggregated the forecast of the State load into more than 10 planning areas, and suggested to model load for each Utility [45]. Segmentation is a useful concept in order to improve the accurarcy of a forecast.

In the pre-PC period, short-term load forecasting utilized just charts and tables. Temperature and day type are key elements in the model similar to the post-PC period. However, the techniques used were limited to simple regression and similar day method. In the post-PC period, regression techniques are further sophisticated such as multiple regression or localized regression. More techniques from artificial intelligence including artificial neural networks, fuzzy logic, and support vector machines are introduced but not well received by Utilities because they are "black-box" techniques. Time series analysis performs accurately but for only very short terms.

Recently, Utilities start to be open to more non-traditional techniques for short term load forecasting. IEEE Working Group on Energy Forecasting (WGEF) organized the Global Energy Forecasting Competition 2012 (GEFCom2012). Some of their goals are to improve the forecasting practices of the Utility industry and to bridge the gap between academic research and industry practice. Out of the top 10 teams 7 had at least one non-parametric method as a part of their prediction model. The ensemble, a procedure to combine prediction from multiple methods, is used as well [69]. The introduction of non-traditional techniques and the ensemble forecast will be crucial in the future short term load forecasting as we have data with more variety and granularity.

Another important subject in short-term load forecasting is hierarchical load forecasting. Traditionally, most STLF jobs are conducted using system level data only. Because of issues with correlations and extra work, a bottom-up forecast from smaller geographic regions may not be worthwhile. However, there are several attempts to predict the load at a smaller level first as different segments of the grid have different attributes. In the GEFCom2012, participants were encouraged to forecast the load at zonal level before forecasting the load at the system level. Fan, Methaprayoon, and Lee (2009) [65] claimed that there exists an optimal region partition under diverse weather and load conditions so that more accurate forecasts for aggregated system load is achieved. In our problem, the ability to predict the load at the substation level is necessary. The practice of hierarchy load forecasting will be even more crucial once we have access to smart meter information.

Note: The load modeling at a substation level with application in forecast have been investigated at least since 1988 [36]. In 2012, Western Electricity Coordinating Council (WECC) Modeling and Validation Work Group attempted to build a load model for each distribution substation for 12 climate zones and 5 seasonal conditions with the detail of electrical end-uses under the substation especially air-conditioners [51]. Unfortunately, the purpose is for dynamic simulations and outage explanations rather than the load forecast. We should seek a new load model at substation with applications in forecast and aggregate them in order to predict the load at system level.

It is difficult to justify what is the most accurate method in short term load forecasting. It depends on the load size, forecast horizon, additional predictors allowed in the model, and even time/location where data is collected. The measure of accuracy is also an issue. A common practical measure is the mean absolute percent error (MAPE) but another relevant measure is the maximum absolute percent error. The table below shows some examples of hourly day-ahead models and their accuracies. Unfortunately, the most equal ground competition GEFCom2012 used neither of these measures.

Note: the models with temperature tends to assume that they have perfect foresight on temperature, which is not true in real day-ahead practice.

In the future the load forecast task is more challenging. Renewable generation, which will be counted as negative loads in the grid is an issue which we will address in this study. Coupling between the load and the electricity price is also an issue especially when the pricing is highly variable over a day. Such an event is likely to occur when volatile renewable generation has high penetration. In addition, new kind of loads such as electric vehicles and smart appliances may be an issue as well. However, we will not address them in this study.

7.6 Conclusion

Modeling load is relatively simple compared to modeling PV. The variability and volatility of load pattern are essential features as it affects additional cost of grid operation. The load forecast, especially short term load forecast of our interest, closely links to the volatility of load pattern. There are variety of techniques and improvement in accuracy over years but new challenges are coming in term of new energy policies, new types of loads, and renewable generation. In subsequent chapters we address how to establish variability, volatility and forecast practice for distributed PV generation, which in turn contributes to better practice in net load modeling.

Name	Techniques	Size (GWh)	Additional predictors	Time/ Location	MAPE
ANNSTLF [5]	Neural networks	5-10	24 hourly weather parameters for the previous day and forecasts for the coming day, 7 day types of a week	1997 / Various locations in US and Canada	2 %
MATLAB [22]	Neural networks, bagged regression trees	10	Dry bulb temperatures, holiday/ weekend flag	2008 / New England	1-2%
Fan et al [65]	SVR with region partition	5	Dry bulb temperatures, anomalous days flag	March-April 2007 / Midwest	2.5-3.5%
Charytoniuk et al [71]	Neural networks without long historical load data. Customer categories partition	20	Temperatures, aggregate demands of the groups of surveyed customers of different categories	July pre 2000 / Unknown	6.3%

Table 7.1: The percentage of average absolute deviation from a mean energy generations from different number of components

Chapter 8

A forecast system for load

Highlights:

- A short-term load forecast at the grid is highly accurate compared to a short-term PV power forecast.
- Due to a consistent load pattern, a methodology to construct a reference load curve, which is analogous to a maximum PV power curve in the PV power forecast, is simple.
- We investigate prediction models for loads in three features: a use of indirect prediction via load ratios, complexity of predictors, and a size of training data.
- It is not necessary to apply a concept of volatility and a forecast system where several prediction models are assembled. It is because one current load forecast model performances significantly better a forecast system of PV power.

As a part of grid operation, a short-term load forecast is essential to match energy demand and supply. In Chapter 7, we discussed a history, a trend and the accuracy of short-term load forecast models. While a short-term load forecast model can be complex, a simple model from the load pattern, as we noted in Chapter 7 is accurate.

Similar to short-term PV power forecast models, one can construct a set of reference load curves, a combination of daily load curves from a previous day, a previous week, and a previous year. The reference load curve is analogous to a maximum PV power curve. This curve leads to a definition of a ratio between an actual load and a reference load, which is analogous to a performance ratio. We then apply the same analysis as in PV forecast models. That is, we test several prediction models with three features: a use of indirect prediction via load ratios, complexity of predictors (previous loads and other variables), and a size of training data. We see that even with a simple model, a short-term load forecast model is relatively accurate compared to a short-term PV power forecast model.

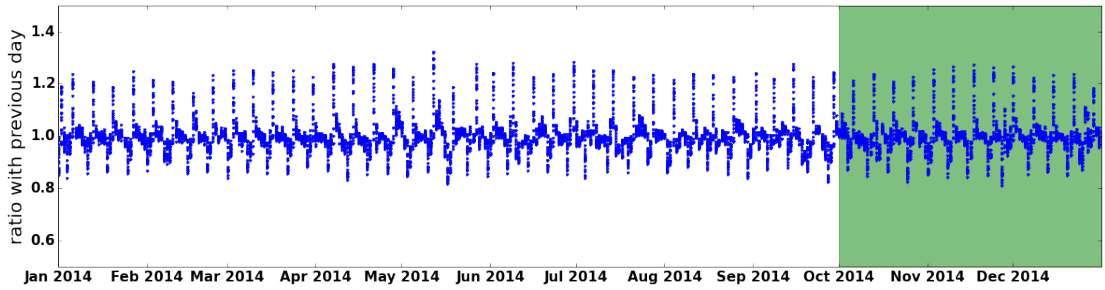
8.1 Load pattern and a reference curve

As we have shown earlier, a maximum PV curve is useful in many ways. It gives an expected PV curve, which informs a grid operation the amount of generated PV power for a next day energy planning. It also leads to a better forecast model via a concept of a performance ratio. It is good to have a smooth curve that represents an ideal variation of a quantity that we want to predict. For PV power, such a curve is a maximum PV curve. For loads, the existence of such curve is not obvious since individual loads can be volatile due to human behavior. However, it turns out that loads at substation and hence the grid are well-behaved. A daily load curve from the previous day, previous week, and previous year tend to have a similar shape as a daily load curve on a current day despite the magnitude of the load changing markedly from day to day. This comes from our observation on load patterns in an Chapter 3.

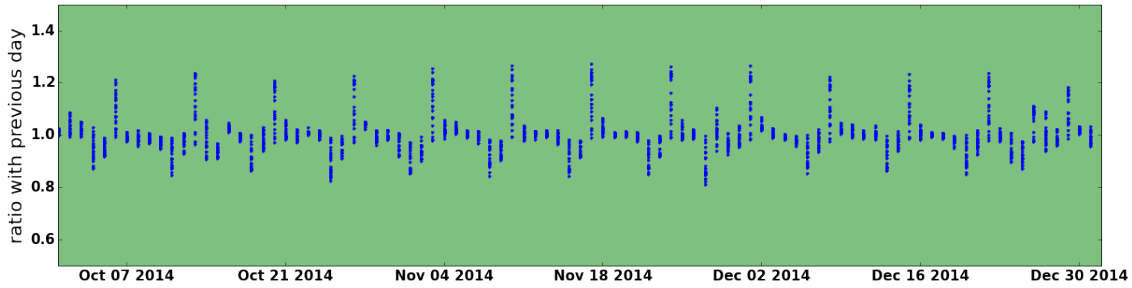
To illustrate this observation, we analyze daily load data from CAISO and PJM for 2014. For each hour instance, we compute a ratio of a current load and a load of a previous day. We collect these ratios and plot their distribution for each day. Figure 8.1 provides the distribution of such ratios in the whole year of 2014 (a, c) to see possible yearly pattern and the last quarter of 2014 (b, d) to see other possible patterns in a smaller timescale. One can see that most of ratios over a year are between 0.9 and 1.1. It just means that the load does not usually change much from one day to another. However, there are days where most of ratios are significantly different than one. In addition, some ratios on those days can be as small as as 0.8 and as large as 1.3. It means that for those days, the average load is significantly different from the previous day, and the profile of daily load curve is significantly different as well. It turns out that those days correspond to Mondays and Saturdays. A drastic difference between two consecutive days across weekdays and weekends discovered here agrees with our discussion earlier on a load pattern.

From this observation, it makes more sense to compare a current load and a load of a previous week to avoid comparison between a weekday and a weekend. Figure 8.2 shows the distribution of ratios between a current load and a load of the previous week in 2014. The outliers on Mondays and Saturdays vanish. Overall, ratios within a same day tend to cluster together. That is, a shape of a daily load curve is similar to the same day of the previous week. However, some ratios shift collectively away from one. It means that the magnitude of a load may be significantly different from the same day of the previous week. Note that the occurrence of collective shifts is on a weekly basis rather than a seasonal basis. It means that such occurrence is likely to be random rather than a consistent variation due to a seasonal change.

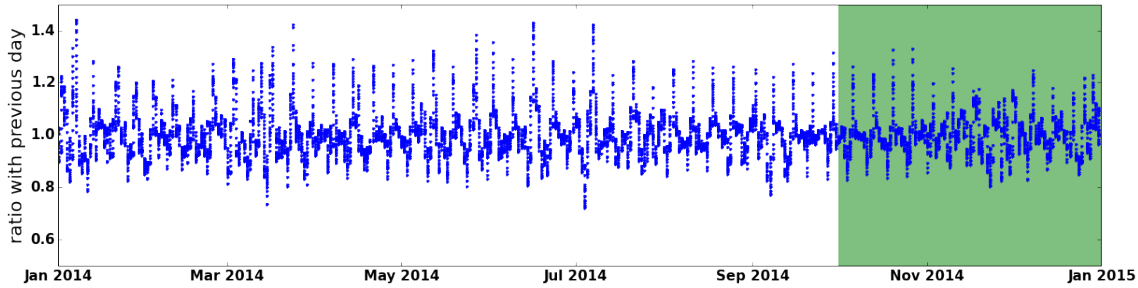
Another comparison one can make is to compare a current load and a load of a previous year. To avoid an effect of a comparison between a weekday and a weekend, we align data to compute a ratio such that the first day of 2014, which is Wednesday,



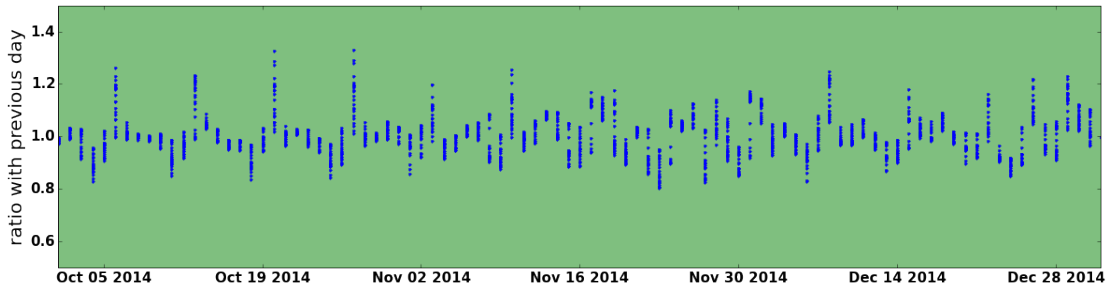
(a) CAISO - the whole year



(b) CAISO - the last quarter

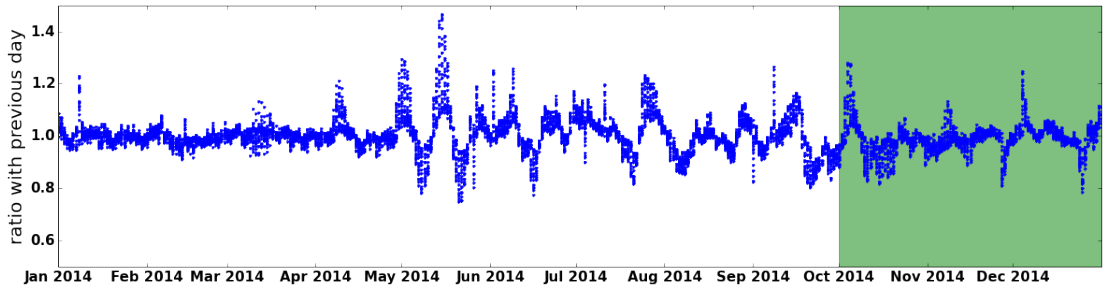


(c) PJM - the whole year

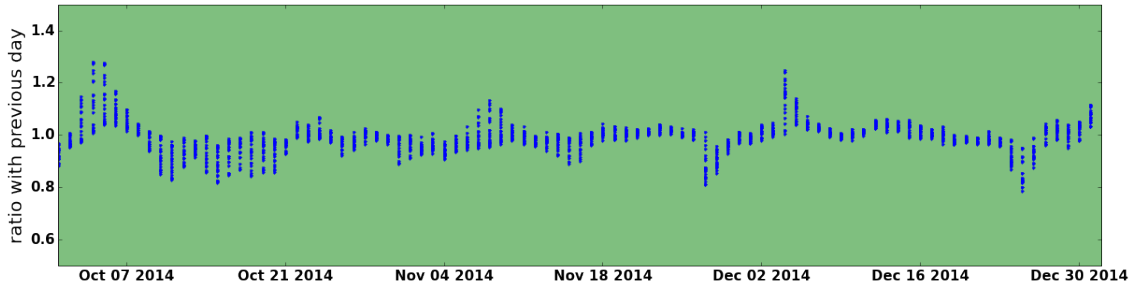


(d) PJM - the last quarter

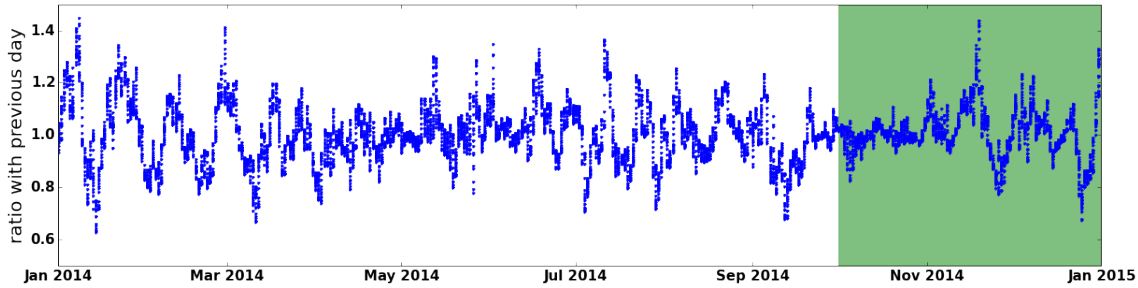
Figure 8.1: Ratio between a current load and the load from the previous day in 2014. Each point represents an hour instance.



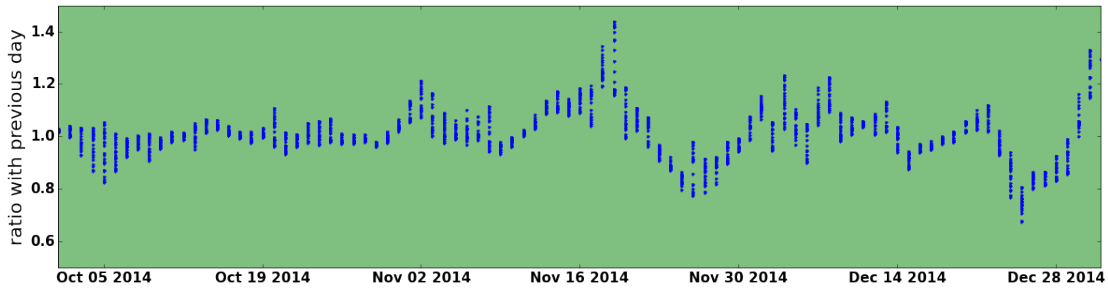
(a) CAISO - the whole year



(b) CAISO - the last quarter



(c) PJM - the whole year



(d) PJM - the last quarter

Figure 8.2: Ratio between a current load and the load from the previous week in 2014. Each point represents an hour instance.

is paired with a first Wednesday of 2013 and so on. Figure 8.3 shows the distribution of ratios between a current load and a load of the previous year in 2014. Similar to Figure 8.2, ratios within a same day tend to cluster together but not as tight as appeared in Figure 8.2. That is, a shape of a daily load curve of a current day is similar to one of the previous year, but it is more similar to one of the previous week. Note that some ratios shift collectively away from one. The occurrence of collective shifts also looks random.

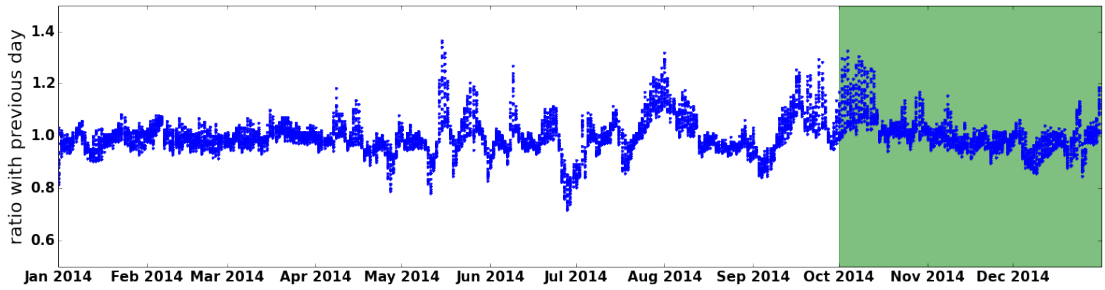
From comparisons with a previous day, a previous week, and a previous year, it is difficult to pick the most sensible reference curve for a comparison since load curves from different times have different characteristics. One may try to build a single set of reference curves by patching daily load curves from a previous week on Mondays and Saturdays with daily load curves from a previous day on other days of a week. However, it is difficult to make a seamless patching without drastically altering original load curves. It is also important not to destroy a sharp variation in daily load curves during the construction of reference load curves. Hence, we decide to build three hour-ahead load forecast models from three unprocessed reference curves (day, week, and year) and find the best way to combine forecast results from them in Section 8.3.1.

8.2 Comments on volatility of load curves

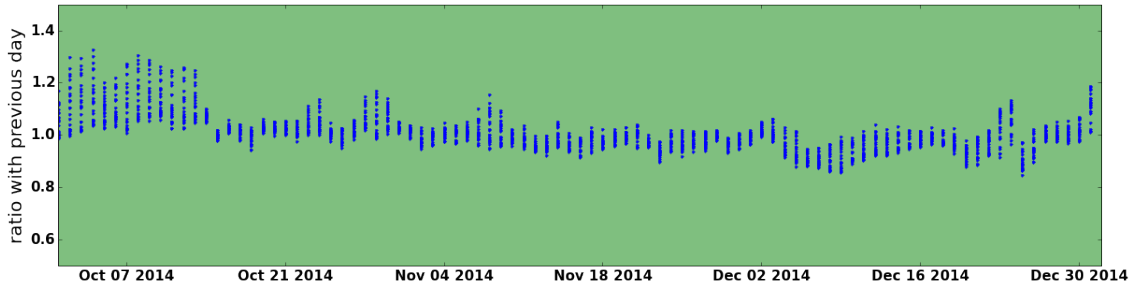
Since we have constructed a reference load curve, it seems natural to apply our definition of volatility on daily PV power curves into daily load curves. The problem is that our reference load curve is not necessarily smooth, especially when we use only a daily load curve of a previous day as a reference curve. If the daily load curve of a previous day is not smooth, then, according to our definition, the volatility of a daily load curve may be high even though the daily load curve is relatively smooth. If we want to apply this persistent definition of volatility, we need to sophisticate the construction of a reference load curve. However, as shown in the next section, our construction of reference load curves gives a much better performance than a forecast system of PV power. For a purpose of prediction accuracy, it is not necessary to refine our construction of reference curves and apply the notion of volatility for load curves at this point.

8.3 Short-term prediction models

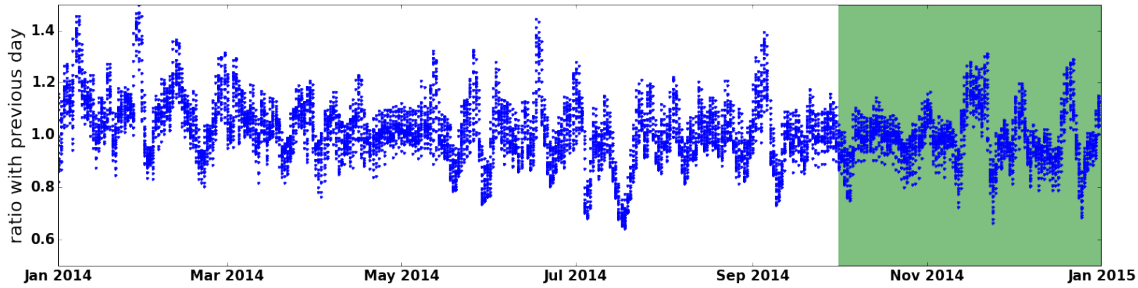
Similar to models for PV power outputs, we considered three features of models for loads. In the first features, we investigated whether a prediction of power via a prediction of load ratios is better than a direct prediction of load. In the second



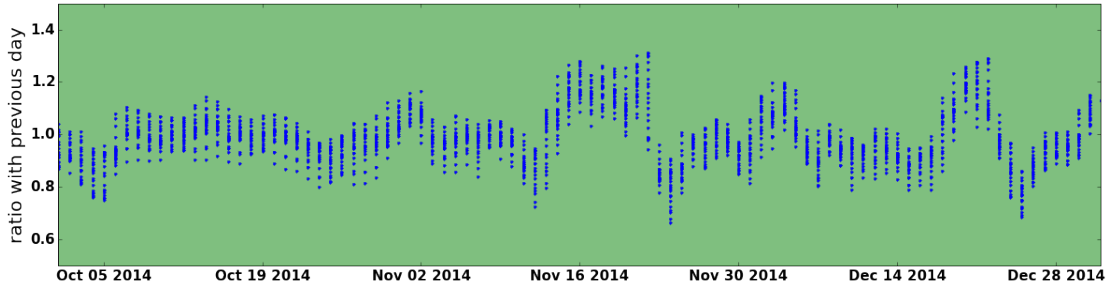
(a) CAISO - the whole year



(b) CAISO - the last quarter



(c) PJM - the whole year



(d) PJM - the last quarter

Figure 8.3: Ratio between a current load and a load from the previous year in 2014. Each point represents an hour instance.

CAISO	p	r_{day}	r_{week}	r_{year}
Method 1	(846,3478)	(245,2418)	(253,1707)	(336,2111)
Method 2	(539,2658)	(290,2574)	(307,2319)	(448,3250)
Method 3	(1231,4611)	(302,2904)	(309,2003)	(385,2811)
PJM	p	r_{day}	r_{week}	r_{year}
Method 1	(2367,9951)	(774,7153)	(688,6700)	(1697,8677)
Method 2	(1573,8349)	(574,8093)	(512,6917)	(1414,11360)
Method 3	(3448,12830)	(1094,9476)	(970,8646)	(2282,10068)

Table 8.1: Value comparison of performance measure (the mean prediction difference, the maximum prediction difference) from different methods predicting different variables

feature, we investigated what is the optimal set of predictors used in the model. Lastly, we investigated what is a sufficiently good size of training data.

To make a fair comparison among all models, we measured the performance of all methods on the test set consisting of data on the last quarter of 2014. To get a sense of a model's accuracy, one may compute the mean load in the test set. It turned out that the mean load in the test set is 24614 MW in the CAISO data set and 86469 MW in the PJM data set.

8.3.1 A prediction of power via a prediction of load ratios

Our end goal of a short-term load forecast system is obviously to predict the load at a future time instance. However, one can predict a future actual load, p, directly, or one can first predict a load ratio, r, to obtain p. We compare performances of a prediction model predicting a future actual power directly and a prediction model first predicting a future load ratio. To avoid complexity from other features of prediction models, we will focus on three simple methods that do not a require training set:

- Method 1: a persistent method $x_{i+1} = x_i$
- Method 2: an extrapolation method $x_{i+1} = x_i + (x_i - x_{i-1}) = 2x_i - x_{i-1}$
- Method 3: a moving average $x_{i+1} = \frac{1}{2}(x_i + x_{i-1})$

Here x can be either p or r. Table 8.1 compares the mean and the maximum of prediction differences using different variables among three methods addressed above.

From the table, one can see that the models first predicting a load ratio, r, gives significantly lower mean prediction differences in general. Among models first predicting a load ratio, the model based on daily load curves from a previous year gives

CAISO	r_{day}	r_{week}	combined
Method 1	(245,2418)	(253,1707)	(203,1598)
Method 2	(290,2574)	(307,2319)	(251,1765)
Method 3	(302,2904)	(309,2003)	(248,1799)
PJM	r_{day}	r_{week}	combined
Method 1	(774,7153)	(688,6700)	(606,5133)
Method 2	(574,8093)	(512,6917)	(451,6586)
Method 3	(1094,9476)	(970,8646)	(856,6402)

Table 8.2: Value comparison of performance measure (the mean prediction difference, the maximum prediction difference) from individual models and the combined model

a significantly high error. In the CAISO data set, the model based on daily load curves from a previous day performs well in term of the mean prediction difference, while the model based on daily load curves from a previous week performs well in term of the maximum prediction difference. In the PJM data set, the model based on daily load curves from the previous week performs well on both measures. Since models based on values from a previous day and a previous week both perform quite well, Here we propose to combine the models by averaging their prediction output. The performance with this new model is shown in Table 8.2. Here one can see that combining results from two models based on different reference curves improves both performance measures in both data sets. Hence, we conclude that, to achieve a better prediction performance, one should build two models, predict load ratios, r_{day} and r_{week}, then obtain actual loads and average them.

8.3.2 Predictors of prediction algorithms

We investigated if models with different predictors gives significantly better performance than models based on previous loads only. We considered three simple methods that relies on previous loads only including a persistent method used in the previous subsection and several auto-regression models.

Table 8.3 compares the performances of four simple algorithms: persistent, AR(1), AR(2) and AR(3).

One can see a general trend that as the number of previous ratios increases, the mean prediction difference decreases. However, adding the third ratio does not yield a significant change. It is not clear how the maximum prediction difference behaves.

To improve the prediction performance, we may add additional predictors into the model including an hour number from the local time h from 0 to 23 and a date number d from 0 to 365 of an instance we want to predict. We may also add some

CAISO	Performance measure
Persistent	(203, 1598)
AR(1)	(200, 1546)
AR(2)	(194, 1584)
AR(3)	(193, 1573)
PJM	Performance measure
Persistent	(606, 5133)
AR(1)	(608, 5132)
AR(2)	(415, 5729)
AR(3)	(412, 6804)

Table 8.3: Value comparison of performance measure from different simple methods

weather information of the recent time stamp in the model as well. Each quantity in weather information is normalized so that the minimum value in the data is zero and the maximum value in the data is one. h and d are also normalized so that the range is $[0, 1]$. This leads to six sets of predictors as follow:

- $\{r_{i-1}\}$
- $\{r_{i-1}, r_{i-2}\}$
- $\{r_{i-1}, h, d\}$
- $\{r_{i-1}, r_{i-2}, h, d\}$
- $\{r_{i-1}, r_{i-2}\} \cup$ Weather
- $\{r_{i-1}, r_{i-2}, h, d\} \cup$ Weather

Due to the complexity of how different predictors might affect the prediction, a non-linear model is required.

First, we considered Support Vector Regression (SVR). Table 8.4 compares performances of the SVR with different sets of predictors after tuning SVR parameters C and e. It turns out that, for the CAISO data set, the best set of predictors is $\{r_{i-1}, r_{i-2}, h, d\}$. For the PJM data set, the best set is $\{r_{i-1}, r_{i-2}h, d\} \cup$ Weather.

Next, we considered Gradient Boosted Regression (GBR). Table 8.5 compares performances of the GBR with various set of predictors after tuning GBR parameters including a number of estimators, a depth, and a learning rate. It turns out that, for the CAISO data set, the best set of predictors is $\{r_{i-1}, r_{i-2}, h, d\}$. For the PJM data set, the best set is $\{r_{i-1}, r_{i-2}, h, d\} \cup$ Weather.

CAISO	Performance measure
$\{r_{i-1}\}$	(199, 1529)
$\{r_{i-1}, r_{i-2}\}$	(195, 1548)
$\{r_{i-1},\ h,\ d\}$	(189, 1485)
$\{r_{i-1}, r_{i-2}, h, d\}$	(188, 1468)
$\{r_{i-1}, h, d\} \cup$ Weather	(191, 1548)
$\{r_{i-1}, r_{i-2}, h, d\} \cup$ Weather	(188, 1552)
PJM	Performance measure
$\{r_{i-1}\}$	(609, 5157)
$\{r_{i-1}, r_{i-2}\}$	(413, 5703)
$\{r_{i-1},\ h,\ d\}$	(611, 5033)
$\{r_{i-1}, r_{i-2}, h, d\}$	(411, 5673)
$\{r_{i-1}, h, d\} \cup$ Weather	(623, 6699)
$\{r_{i-1}, r_{i-2}, h, d\} \cup$ Weather	(418, 5513)

Table 8.4: Value comparison of performance measure from the SVR with different sets of predictors

CAISO	Performance measure
$\{r_{i-1}\}$	(201, 1540)
$\{r_{i-1}, r_{i-2}\}$	(190, 1573)
$\{r_{i-1},\ h,\ d\}$	(183, 1238)
$\{r_{i-1}, r_{i-2}, h, d\}$	(182, 1250)
$\{r_{i-1}, h, d\} \cup$ Weather	(183, 1250)
$\{r_{i-1}, r_{i-2}, h, d\} \cup$ Weather	(182, 1351)
PJM	Peformance measure
$\{r_{i-1}\}$	(615, 5510)
$\{r_{i-1}, r_{i-2}\}$	(417, 4525)
$\{r_{i-1},\ h,\ d\}$	(586, 5154)
$\{r_{i-1}, r_{i-2}, h, d\}$	(397, 4383)
$\{r_{i-1}, h, d\} \cup$ Weather	(588, 5199)
$\{r_{i-1}, r_{i-2}, h, d\} \cup$ Weather	(415, 4364)

Table 8.5: Value comparison of performance measure from the GBR with different sets of predictors

It may be seen that the GBR performs better than the SVR and other simple models in general. With a different set of predictors of each data set, we bring both the mean and the maximum prediction differences down significantly. It is worth noting that adding weather information is not distinctively beneficial. Since our weather data is for a specific zip code, it may not relevant to the load data for an entire balancing area. Furthermore, the effect of weather to the load may be already captured in reference load curves. We used the GBR method with the set $\{r_{i-1}, r_{i-2}, h, d\}$ both data sets to study the effect of a size of training data in the next subsection.

8.3.3 Size of training data

The size of training data is the last feature of a prediction model. Similar to a forecast model for PV outputs, we compare performances from the best model we have at hand with different sizes of training data in an on-line approach. One may keep adding training data while doing prediction or one may limit training data to a certain number of most recent data points. For boyh CAISO and PJM data sets, we use a GBR method to predict the next load ratio first based on $\{r_{i-1}, r_{i-2}, h, d\}$. Then we obtain the actual load from the ratio. We consider four schemes for an on-line approach:

- There is no limit on size of training data. Keep adding new data points of a day in the data set and retrain every day.
- Limit training data to be at most 6 months of data. Adding new data points of a day, delete data points of the oldest day, and retrain every day.
- Limit training data to be at most 3 months of data.
- Limit training data to be at most 1 months of data.

All schemes are applied to the same test set, which is data of the last quarter in 2014. The training data was updated daily instead of hourly because the GBR is computationally expensive. In practice, an update can be done for every new data point since there is enough time between arrivals of new data points. It might be worthwhile updating more frequently, but we did not investigate this option in this study.

In Table 8.6, we compare performances of on-line prediction models with various sizes of training data. From both data sets, adding new data points of a day in the data set without deleting gives the best performance. It means that the optimal amount of data to produce the best performed model is more than 6 months of data. Larger size of training data is not investigated because of our data is limited.

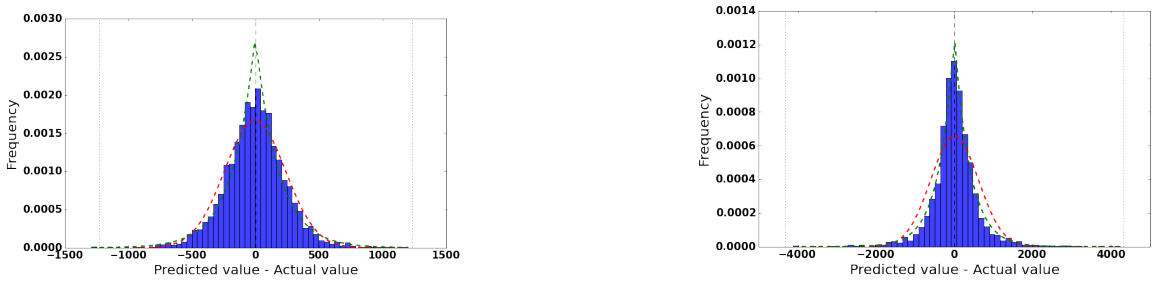
CAISO	mean	max
No limit	180	1269
6 months	199	1583
3 months	205	1659
1 month	205	2438
PJM	mean	max
No limit	396	4340
6 months	525	7477
3 months	569	7265
1 month	569	7374

Table 8.6: Value comparison of performance measures from an on-line approach with different sizes of training data

8.4 Comments on a forecast system

In the previous section, we have found that for each load prediction model the performance depends on a use of indirect prediction via load ratios, a complexity of predictors, and a size of training data. We have established from our two data sets that it is good to first predict a load ratio then obtain an actual power. We also established that the size of our data set should be as large as possible. For a complexity of predictors, in both data sets, we use load ratios from two most recent instances, a date number and an hour number. It is worth noting again that weather information, especially the temperature, does not improve the performance of a prediction model. It might be because the weather data is for a specific local area, while the load data is for an entire balancing area. Another possible reason is that the effect of weather to the load is already captured in reference load curves.

In a task of forecasting a load, we do not construct ensembles of prediction models with a protocol to switch between models as we do not apply a concept volatility used in PV power outputs into loads. As we have addressed earlier, our rough call of reference load curves may cause an incorrect measure of volatility. Even though we do not have ensembles of prediction models, a single prediction model works well. Recall that the mean load in the test set is 24614 MW in the CAISO data set and 86465 MW in the PJM data set. From the best prediction model, the mean prediction difference equals 180 MW in the CAISO data set and 397 MW in the PJM data set. The maximum prediction difference equals 1293 MW in the CAISO data set and 4366 MW in the PJM data set. It means that the mean error is about 1 percent of the mean load in the test test, while the maximum error is at about 6 percent of the mean load in the test set. These numbers are about 10 times smaller than PV power



(a) CAISO (b) PJM

Figure 8.4: Distribution of prediction differences from a short-term forecast model. The histogram shows actual data. A red line shows a fit of a normal distribution and a green line shows a fit of a Laplace distribution.

prediction. Furthermore, in the CAISO data set, there are only one instances where the error is higher than 5 percent of the mean load in the test set. In the PJM data set, there is only one instance where the error is higher than 5 percent.

To tell how likely a bad prediction will happen in detail, a distribution of prediction differences is needed. Figure 8.4 shows a distribution of differences between predicted loads and actual loads on the test set with this short-term forecast model of loads. The vertical dotted lines show values equivalent to 5 percent of the mean load in the test set. One can see that the distribution of such differences is a Laplace distribution rather than a normal distribution. One can fit a Laplace distribution and get the best fit with parameters $(\mu, b) = (-0.4, 175)$ for the CAISO data set and $(\mu, b) = (4, 397)$ for the PJM data set. The bias of prediction is relatively small. From this, one can calculate how likely a bad prediction will happen in general when using this short-term forecast model.

8.5 Conclusion

After an investigation in a short-term load forecast system, we found that applying procedures from a short-term PV power forecast system is useful. A concept of load reference curves and load ratios is useful to construct a good model. A simple use of load reference curves makes our definition of volatility inapplicable, yet an accuracy of the forecast model based on such reference curves is high. One may refine a concept of load reference curves to be more formal if needed. In addition, we need more data to establish an optimal size of training data.

Part III

Grid under high distributed PV generation

Chapter 9

Overview of grid operations

Highlights:

- A definition of a grid operator and its role are explained.
- A grid operator achieves demand-supply balance using different stages of operation. It requires energy scheduling to prepare supply as accurate as possible, and ancillary services planning to correct a demand-supply imbalance.
- We define and categorize ancillary services. We also introduce a rule of ancillary services requirement relates to a risk of demand-supply imbalance.

The main goal of a electric grid operation is to supply electric energy from generators to customers that satisfies operational limits and minimizes both the cost of operation and the risk of degradation of the infrastructure. In a grid operation, an electric power produced by generators called a **generated power** or **supply** must be equal to an electric power required by consumers called a **load** or **demand** at all times. What makes the grid operation complicated is that there is a large number of non-linear constraints coming from operational limits in generators, transmission networks, distribution networks, and customer units. The reliability of components in the grid also makes the grid operation complicated. Another source of complication is an uncertainty in loads. Once a high amount of distributed PV generation is added to the demand side as a negative load, the **net load**, which is a summation of a regular load and a negative load, become even more uncertain.

In this chapter, we define a grid operator and its actions as needed in our study. We focus on a demand-supply balance and stages of operation to ensure the balances using energy scheduling and ancillary services. We further discuss the ancillary services requirement, its relation with a risk of demand-supply imbalance, and different categories of ancillary services.

9.1 Definition of a grid operator

A grid operator is an entity who controls an electricity transmission system to keep it in balance and operates a wholesale electric market in order to minimize the cost of operation. In the US, these two functions are not necessarily executed by the same entity. In this section, we clarify how the two functions are executed in US.

In terms of the electricity transmission system, a basic operating unit of the electric power industry is called a **balancing authority** controlling a balancing area or a control area. Each control area may have a different number of **nodes** representing points where two or more elements in the grid are connected. Some nodes can be physical transmission substations where generators and loads are connected to the transmission lines. Some nodes can be just intersection points in the transmission network. There are more than 100 balancing authorities under North American Electric Reliability Corporation (NERC). Balancing authorities are united to improve the reliability of the bulk power system together as NERC regional entities. A map of regions and balancing authorities is shown in Figure 9.1 [53]. A total of eight NERC regional entities are distributed throughout the Northern America electric grid, which is actually comprised of three smaller grids, called interconnections as shown in Figure 9.2 [52]. The responsibility of balancing authorities is to continuously maintain load-interchange-generation balance within balancing authorities and to help the entire interconnection regulate and stabilize the frequency. A balancing authority is not responsible for managing retail or distribution-level power systems. More importantly, a balancing authority does not have a direct role to minimize the cost of operation.

In terms of the wholesale electricity market, a basic supplier of the electric power market is called a **Utility**. In a traditional wholesale electricity market, **vertically-integrated utilities** own the generation, transmission and distribution systems used to serve electricity consumers. In order to facilitate a competition for electricity generation using bid-based markets, an Independent System Operator (ISO) or Regional Transmission Organization (RTO) is formed to operate the transmission system independently. With this arrangement, Utilities are just generators or distribution units. Currently there are 10 main wholesale electricity markets in the US. A map of those markets are shown in Figure 9.3. Note that these market regions are not aligned with balancing regions. While major sections of the country operate under more traditional operating structures, two-thirds of the nation's electricity load is now served under ISOs or RTOs [32].

In our definition, a grid operator must function as a balancing authority and a market. Hence, we count ISOs, RTOs and some vertically-integrated Utilities who balances its control area as grid operators. Balancing Authorities without an authority to settle the cost of operation, Utilities without ability to control the grid at

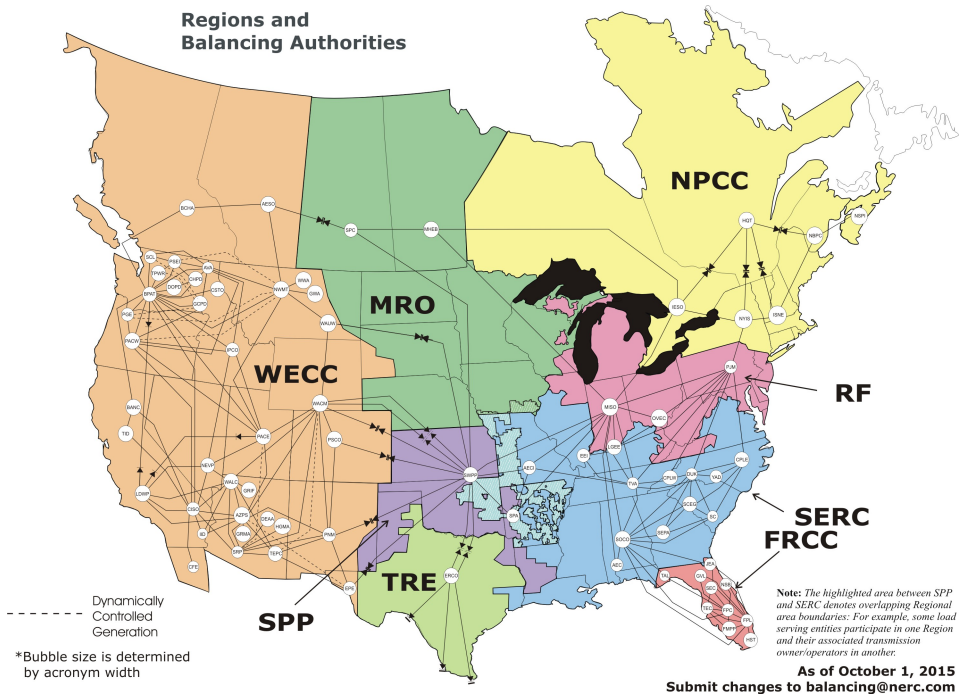


Figure 9.1: Balance authorities in US, 2015.

a transmission level, and market operators with no control on the transmission and generation are not grid operators.

9.2 Actions by a grid operator

In order to operate a power system, different entities make decisions on time scales covering fifteen order of magnitude prior to real-time operation as shown in Figure 9.4 [13]. In this section, we will focus only scheduling and demand-supply balance within a control area from a time scale of at most one day without concerning stability and congestion within the balancing area.

At the time scale of a day before the real-time operation, a grid operator conducts a **day-ahead scheduling** by determining how much power is needed and which **generators** are scheduled to serve during each hour of the operating day. In CAISO, a day-ahead scheduling is made from a bidding market opening 7 days prior to the operating day and closing the day prior to the operating day. A complete schedule is published at 1 PM prior to the operating day. Note that the time resolution of a day-ahead scheduling is one hour, but it can be further refined to account for 20-minute ramps between hours.

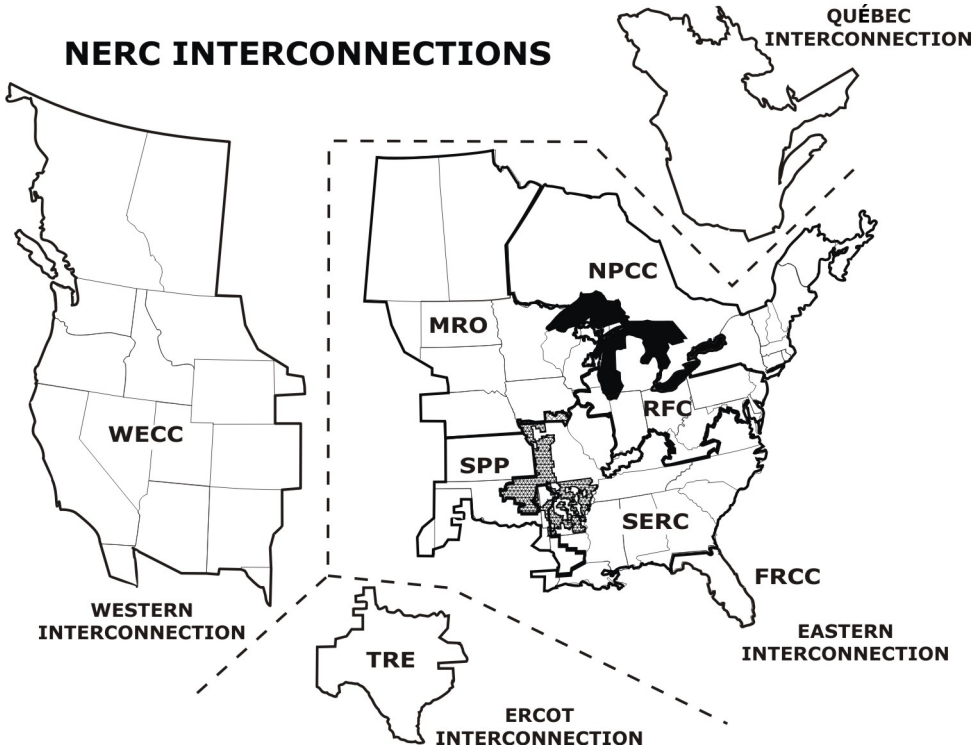


Figure 9.2: NERC Interconnections

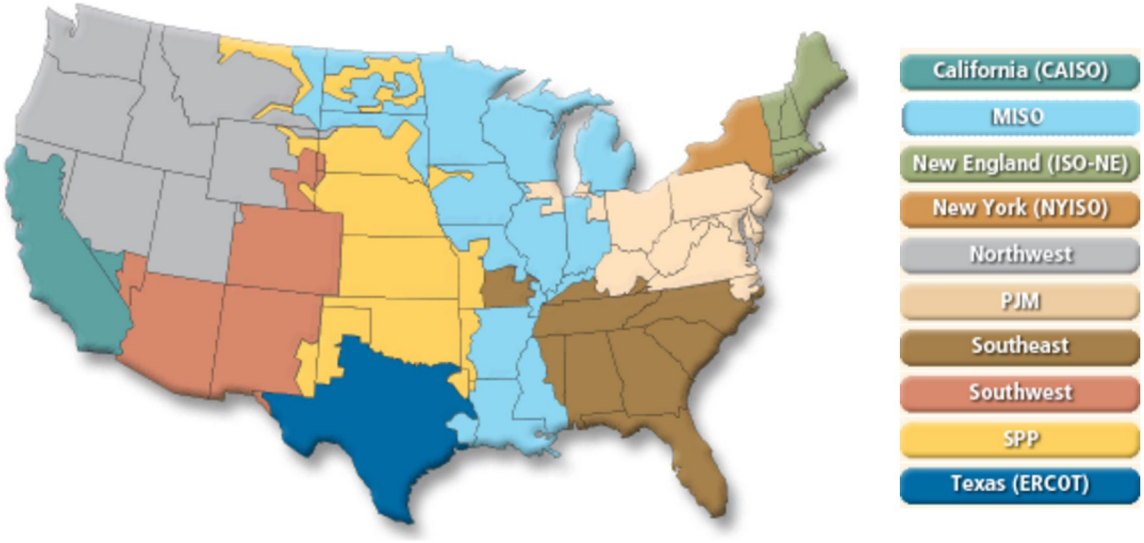


Figure 9.3: 10 main wholesale electricity markets in US.

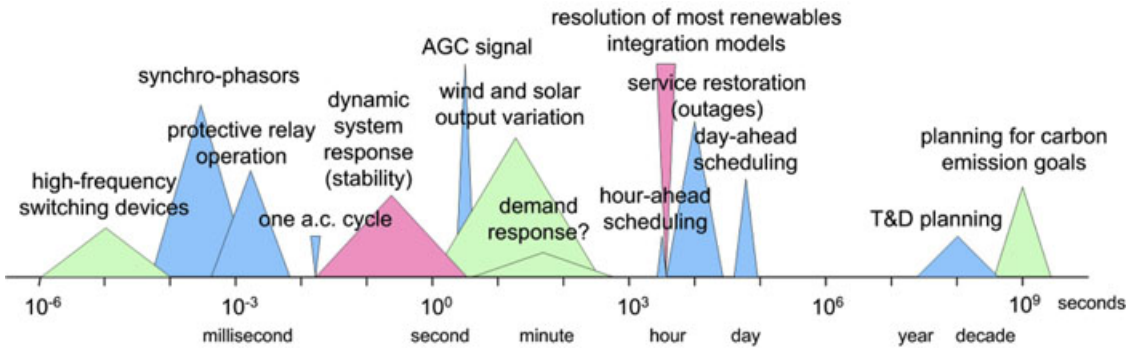


Figure 9.4: Schematic of time scales involved in power system planning and operations

At the time scale of an hour before the real-time operation, a grid operator conducts an **hour-ahead scheduling**. It is sometimes called misleadingly **real-time scheduling**. A grid operator conducts an hour-ahead scheduling by determining how much power is needed and which generators are scheduled to serve during each 5-minute block of the operating hour. In CAISO, an hour-ahead scheduling is made from a bidding market opening 1 PM prior to the operating day and closing 75 minutes prior to the operating day. A complete schedule is published 45 minutes prior to the operating hour.

In the next time scale before the real-time operation, a grid operator dispatches an **ancillary service** called **operating reserves** or **contingency reserves**, in order to generate power to the grid within minutes. A bidding market is not practical at this point. The grid operator just dispatches this type of reserves in operating 5-minute blocks.

At the real-time operation, a grid operator dispatches an ancillary service called **regulating reserves** in order to match the demand within seconds. When a demand and a supply are out of balance, Area Control Error (ACE) increases, which is measurable via a change of the system frequency and other metering. The grid reacts to this change using Automatic Generation Control (AGC) that controls regulating reserves to bring a balance back.

The operation of a grid operator from a day before to the real time is summarized graphically in Figure 9.5.

It is worthwhile noting that, since ancillary services are dispatched close to or right at the operating instance, the market to decide how much capacity is needed and which services are chosen is conducted a day prior to the operating day. Since there is no specific amount of power assigned until the last minutes or seconds, the capacity and its ramp rate are important factors when choosing ancillary services. Other than traditional resources, renewable energy resources paired with storage and demand responses can function as fast responsive ancillary services.

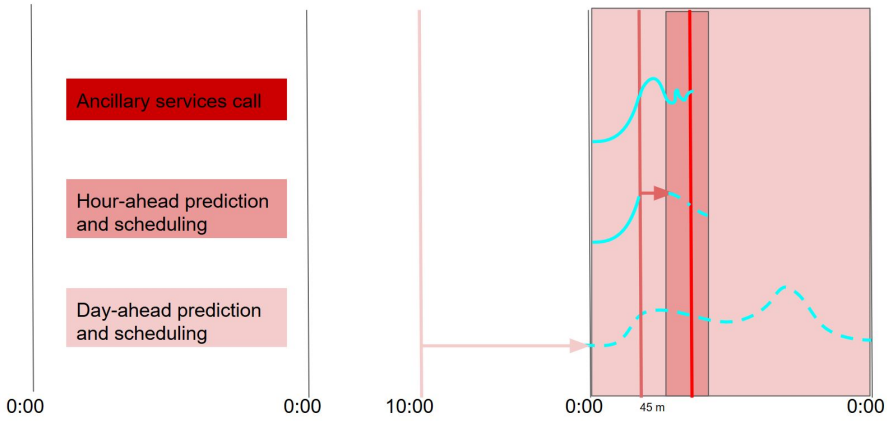


Figure 9.5: Timeline of actions by a grid operator

A grid operator needs to adjust their actions when high levels of distributed PV generation are introduced to the grid in term of both energy scheduling and ancillary services planning. In energy scheduling, as we have seen in a previous chapter, the variability of PV generation alters the shape of the net load drastically which requires generators with high ramp rates. In ancillary services planning, the volatility of PV generation turns into a high uncertainty of net load prediction, which requires ancillary services with larger capacity and faster response. For energy scheduling, preparing generators to cope with variability of net loads has an additional cost yet manageable since it can be scheduled ahead and the mismatch between demand and supply can be corrected by ancillary services later. Ancillary services planning is discussed in the next section.

9.3 Ancillary services

9.3.1 Ancillary services requirement

As addressed the previous section, an ancillary service is essential to maintain the power balance within the grid. As a result, NERC directs some ancillary service requirements and several ISOs further define them. Other than reserves and regulations, there are additional required ancillary services for other purposes including reactive power and voltage control, real-power loss compensation, and power recovery when the system collapses. However, we are not concerned with ancillary services for those purposes. Here we are concerned with the ancillary services requirement as discussed in Ela et al [29].

For operating reserves, the NERC Disturbance Control Standard (DCS) requires that each balancing area have enough reserve to cover its largest credible contingency, which equals to the capacity of the largest generating unit in the area. In addition, most regions require that at least 50% of reserves are **spinning reserves**, also called **synchronous reserves** and **responsive reserves**, which is on-line, synchronized, and ready to respond within 10 minutes. Note that the definition of spinning reserves from different literature can be different [74]. Some examples of spinning reserves are pumped-storage hydroelectricity stations and gas turbines. The rest of reserves are **non-spinning reserves**, also called **non-synchronous reserves** and **supplementary reserves**, which is off-line and ready to respond within from 10 to 30 minutes. Some examples of non-spinning reserves are combustion turbines. Again, both quantity (power) and speed (ramp rate) are key quantities to decide which reserves will be used. Ortega-Vanquez [70] summarized spinning reserve requirements from different power systems. In CAISO, the rule for spinning reserve requirement is complicated. A rough estimate is about 3% of the peak load of the day. In PJM, the rules for spinning reserve requirement depends on the region. A rough estimate is about 1.5 % of the peak load of the day.

For regulating reserves, the NERC Control Performance Standard (CPS) directs how much and what type of regulation to procure. Each ISO adjusts its requirements based on extra information on time of a day, day of a week, and season. The regulation is required to be fully available within 5 minutes, which is an economic dispatch interval. Note that the regulation requirement may be uni-directional or bi-directional, meaning that both upward and downward capacity must be made available. Ela et al summarized regulation requirements from different ISOs. In CAISO, the rule of regulation requirement has a floor of 350 MW, which can be adjusted based on different factors. Such a minimum is about $350/45832 = 0.8\%$ of the peak load of the year. In PJM, the rule for regulation requirement is based on 1% of the peak load during the peak hours and 1% of the valley peak during off-peak hours.

In conclusion, an ancillary requirement for spinning reserves, non-spinning reserves, and regulation adds up to roughly 4-7% of the peak load of the year. Whatever the rule for ancillary service requirement is, it should be set based on reliability of the individual scheduled generators and the accuracy of the load forecast. An amount of PV generation, which can be viewed as a negative load, clearly affects the accuracy of the net load forecast and hence the ancillary services requirement.

9.3.2 A risk of demand-supply imbalance and a rule of ancillary services requirement

In order to set a reasonable rule for an ancillary services requirement, especially when volatile resources like PV generation are included, it makes sense to establish a risk

of demand-supply imbalance. If we assume that there is no risk of energy delivery failure from both scheduled supply and ancillary services, then we can consider only a probability distribution of demand-supply imbalance between the net loads and the scheduled generation at an instance. Ancillary services available at that instance gives an interval in which an imbalance in such interval is acceptable since available ancillary services can re-balance it. A probability that a demand-supply imbalance falls outside the interval provided by available ancillary services defines the risk of demand-supply imbalance.

A rule of ancillary services requirement should assign ancillary services enough to keep a risk of demand-supply imbalance to a certain threshold. For example, if we check every hour whether a demand-supply imbalance beyond ancillary service capability, and we expect to find the imbalance less than once in a year, we need a risk of demand-supply imbalance to be smaller than $1/(24 \times 30 \times 12) \approx 0.01\%$.

9.3.3 Comments on Type I and Type II ancillary services

In the discussion of a grid operation, we categorize ancillary services by their responsiveness into operating reserves and regulating reserves. We can also categorize ancillary services by active capability to generate or consume power into **Type I** and **Type II** services. A Type I service is an ancillary service that active puts power into the grid or takes power out of the grid. As a result, operating a Type I service does not require sophisticated scheduling as long as there is a sufficient capacity. An example of Type I services is a gas turbine generator. On the other hand, a Type II service is an ancillary service that shifts demand and supply existing in the grid to coincide. One may call Type II service as a load shifting service. As a result, operating a type II service requires sophisticated scheduling with constraints on the maximum stored energy and the maximum shift time. Some examples of Type II services are a flywheel, an energy storage, and a demand response service. The majority of ancillary services in the current grid is Type I services, but with a small operational cost and a decreasing up-front cost, Type II services may take a larger share in the future.

Note that in energy scheduling, we tend to use generators, such as fossil fuel power plants, to serve the load. However, a grid operator may schedule Type II instruments to reduce consistent variation in the net load as well. For example, one may assign a thermal storage in a day-ahead energy schedule to flatten the daily net load curve in addition to traditional power plants. This kind of use does not function as an ancillary service.

9.4 Conclusion

In this study, we discussed how a grid operator achieves a demand-supply balance within its control area using energy scheduling and ancillary service. With high distributed PV generation, the variability and volatility of the net load increases, leading to a need of better instruments. For energy scheduling, a grid operator needs generators with higher ramp rates or energy storage to cope with the variability of daily load curves. For ancillary services planning, a grid operator needs ancillary service units with higher capacities to cope with the volatility of the net load. If there is no instrument updates, we expect that the risk of demand-supply imbalance increases as the distributed PV generation increases.

In the next chapter, we assess the risk of a demand-supply imbalance under different levels of distributed PV generation. We investigate options of ancillary services and PV curtailment required to keep the risk of a demand-supply imbalance down to an acceptable amount.

Chapter 10

Risk management of demand-supply imbalance

Highlights:

- The risk of a demand-supply imbalance under high distributed PV generation can be assessed using simulation based on our forecast system of the net load given a set of generators and ancillary services.
- Based on our data set, the risk of demand-supply imbalance increases as installed PV capacity in the grid increases. In order to maintain an acceptable risk, the maximum level is 15 percent of the peak load for the western example and 9 percent of the peak load for the eastern example.
- In the simulation, the limiting factor to the maximum level of PV installation is volatility rather than variability. Hence, the ancillary services and PV curtailment are primary options to reduce the risk when the level of PV installation increases.
- Adding Type I ancillary service or adding Type II ancillary service with a simple heuristic can increase the maximum percentage of installed PV capacity by 6-8% of the peak load without an extra modification on generators in day-ahead and hour-ahead scheduling.
- Adopting the PV curtailment with a policy based on the normalized volatility of PV generation allows extra PV installation of about 2% of the peak load.

A grid operator delivers electric energy from generators to customers reliably using two actions: scheduling and ancillary services control. With high distributed PV generation, the net load that a grid operator needs to serve becomes more variable and volatile. In order to keep the risk of a demand-supply imbalance down to an

acceptable level, new infrastructures such as generators and ancillary services are required. Operational procedures including Type II services control and PV power curtailment are also relevant.

In this chapter, we first assess the risk of a demand-supply imbalance under a different level of distributed PV generation with a typical set of infrastructures including generators and Type I ancillary services. Then, we determine specifications of new ancillary services in order to reduce the risk of a demand-supply imbalance down to an acceptable level. We investigate how PV curtailment can help reduce the risk as well.

10.1 Limitation of data

We want to simulate how a grid operator controls its balancing area using scheduling and ancillary service control under a high PV installation according to procedures described in the previous chapter. However, some limitations of data prevents us from doing this. First, our data is not frequent enough. The typical resolution that a grid operator needs to actively manage is 5 minutes. Our data on the PV generation, however, has a resolution of 15 minutes, while our data on the load and the weather information has a resolution of 1 hour. Hence, we cannot fully simulate the grid operation at the same level as the actual grid operation. Still, we are able to simulate both a day-ahead scheduling and an hour-ahead scheduling without any problem. We may calculate differences between actual net loads and generating powers, and compare them with capacities of available ancillary services at the beginning of hours. In an ideal situation, we may do calculations and comparisons at the beginning of all 5-minute blocks. In our situation, we just do calculations and comparisons with a frequency of 15 minutes. We expect that such a sampling frequency is sufficient.

Other than the issue with the frequency of data, the aggregation of data is also a limitation. While each load data set from CAISO or PJM is aggregated from the entire balancing area, the PV generation and weather data are not. As described earlier, each PV generation data set in CA or NJ is aggregation from PV systems clustered in a certain zip code, which is comparable to a PV generation under a substation, not the entire balancing area. The weather data is also specific to a certain zip code. In an ideal situation, we may build each model for each substation or zip code, and aggregate results from the models. In our situation, we just use our PV generation and weather data as if it is sampled from the entire balancing area. Our approach to combine two sets of data in the grid integration analysis is significantly different from what the grid operator should do in practice. Once again, to highlight the distinction between the example and real operations, we call the combination of the PV data from a zip code in CA and the load data of a whole grid where the zip code is located **the western example**. We call the combination of the PV data from a zip code

Infrastructure	Power capacity	Ramp rate capacity
Generators (day-ahead)	Unlimited	+5% of the peak ramp rate
Generators (hour-ahead)	Unlimited	+10% of the peak ramp rate
Ancillary services	5% of the peak load	Unlimited

Table 10.1: Specification for infrastructure in the simulation

in NJ and the load data of a whole grid where the zip code is located **the eastern example**.

10.2 Simulation setup

Before explaining the simulation procedures, we address set-ups on infrastructures, forecast systems, and how to measure the risk. We need to make various assumptions based on our best knowledge in the grid operation in general. None of the assumptions critically alter the conclusion. In fact, many assumptions are on the side that favors PV integration.

10.2.1 Infrastructure

In the simulation, we set up a typical set of infrastructures such as generators and ancillary services. We assume the following specifications. In a day-ahead market, the combined ramp rate of all generators in the balancing area is 5% higher than the maximum ramp rate of the load in a year. In an hour-ahead market, the combined ramp rates for all generators in the balancing area is 10% higher than the maximum ramp rate of load in a year. For ancillary services, the amount of Type II services is negligible. Combined capacity of ancillary services at all times is 5% of the peak load of a year. Ancillary services are fast enough to correct any level of imbalance within their combined capacity at all times. These specification is summarized in Table 10.1.

10.2.2 Forecast system

In the simulation, we set up a forecast model of the load and the PV power for both a day-ahead scheduling and an hour-ahead scheduling. For a day-ahead scheduling, we predict a daily net load curve of the operating day with a resolution of one hour. Since the schedule is finalized at 1 PM prior to the operating day, only PV power, load, and weather data before that time are available for forecasting a daily net load curve. For a daily PV power curve, we use an expected PV power curve as a predicted PV power curve. Hence, there is no issue if recent data is not available. For a daily

load curve, one may use a combination of daily load curves of a previous week and a previous day. However, the daily load curve of a previous day is not available yet. Hence, we choose daily load curves of a previous week instead.

For an hour-ahead scheduling, an hourly net load curve of the operating hour should be predicted with a resolution of five minutes. However, since the finest resolution of available PV data is 15 minutes, we decide to consider the system at the resolution of 15 minutes. Since the schedule is finalized at 45 minutes prior to the operating hour, only PV power, load, and weather data before that time are available for forecasting net load in the hour block. We decided to use only data from at least 60 minutes prior to the operating hour. For a PV power prediction, our forecast system for the next time stamp is not relevant here since a gap between consecutive data points is 15 minutes. We need a forecast system that predicts the PV power in the next 60, 75, 90, 105, and 120 minutes. To simplify the forecast system, we built a model to predict the PV power in the next 60 and 120 minutes. Then we predict the PV output for other instances by interpolation. For the load prediction, our forecast system of the next time stamp is relevant since a gap between consecutive data points is 1 hour. Still, we want to predict the load in the next two hours as well in order to complete the hour block. Hence, we need a forecast system that predicts the load in the next 60 and 120 minutes. Then we predict the load for other instances by interpolation. We built forecast systems using the same features as we used in the last sections of Chapters 6 and 8.

10.2.3 Risk measurement

In order to measure the risk of a demand-supply imbalance, we first count the number of instances where the demand-supply imbalance exceeds the ancillary service capacity at each instance. Then we calculate the risk equal to that number divided by the number of instances in the period of consideration. Ideally, the period of consideration should be multiple years. However, in this simulation with the limited amount of data, the period is the last 3 months of 2014. Since the time interval is 15 minutes, we are accounting $3 \times 30 \times 24 \times 4 = 8640$ total instances. We define an acceptable risk of demand-supply mismatch to be 0.05 %. This risk allows about $3 \times 30 \times 24 \times 4 \times 0.05\% \approx 4$ imbalance instances in 3 months period. Note that if we sample instances every hour instead, the number of instances should be 1 instance in 3 months period or 4 instances in 1 year. It is equivalent to 4 loss of load events (LOLE) [43] per year.

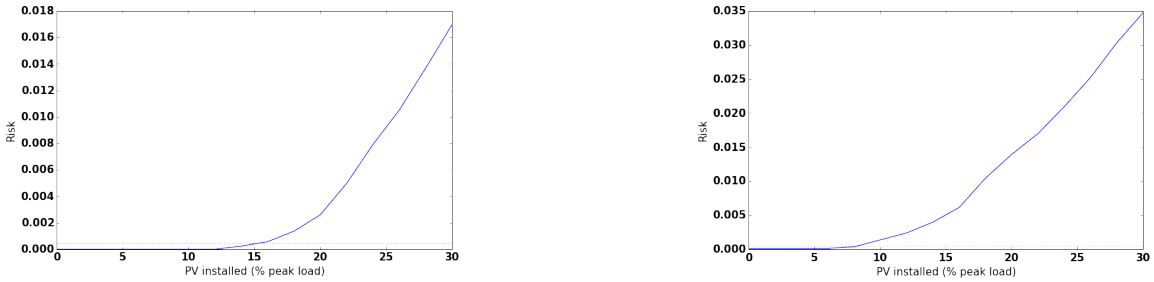
10.3 Simulation procedure

The simulation procedure to assess the risk of a demand-supply imbalance is listed below.

1. Define the simulation period. In this study, it is the fourth quarter of 2014.
2. Identify the peak load and the maximum hourly ramp rate of load from the most recent year before the simulation period. In this study, we identify them from the load data from fourth quarter of 2013 to the third quarter of 2014.
3. Identify the maximum hourly ramp rate capacity of generators in both a day-ahead scheduling and an hour-ahead scheduling. Identify the capacity of ancillary services.
4. Assign the PV capacity as some percentage of the peak load.
5. Build a forecast system for a day-ahead scheduling from the data before the simulation period. In this study, we use data of the first three quarters of 2014.
6. Build a forecast system for an hour-ahead scheduling from the data before the simulation period. In this study, we use data of the first three quarters of 2014.
7. For each day in the simulation period, predict a daily net load curve. Then assign generators under a day-ahead scheduling without exceeding the maximum hourly ramp rate limit. Record when the ramp limit is reached.
8. At each hour in that day, predict the net load of an entire hour block for the next hour. Readjust generators accordingly. Make sure that the addition of generators does not exceed the maximum ramp rate limit in an hour-ahead scheduling. Record when the ramp limit is reached.
9. Every 15 minutes, measure the difference between the net load and the assigned generation from scheduling. Record when such difference is higher than the capacity of ancillary services.
10. Once the simulation is completed, compute the risk of a demand-supply imbalance as explained in the previous subsection.

10.4 The risk of demand-supply imbalance

Figure 10.1 shows the risk of demand-supply imbalance from the simulation in two regions as a function of the installed PV capacity in the grid. The risk of demand-supply imbalance increases as the installed PV capacity in the grid increases. For the western example, the grid with the installed PV capacity greater than 15 percent of the peak load has a significant risk of demand-supply imbalance. For the eastern example, the grid with the installed PV capacity greater than 9 percent of the peak load has a significant risk of demand-supply imbalance.



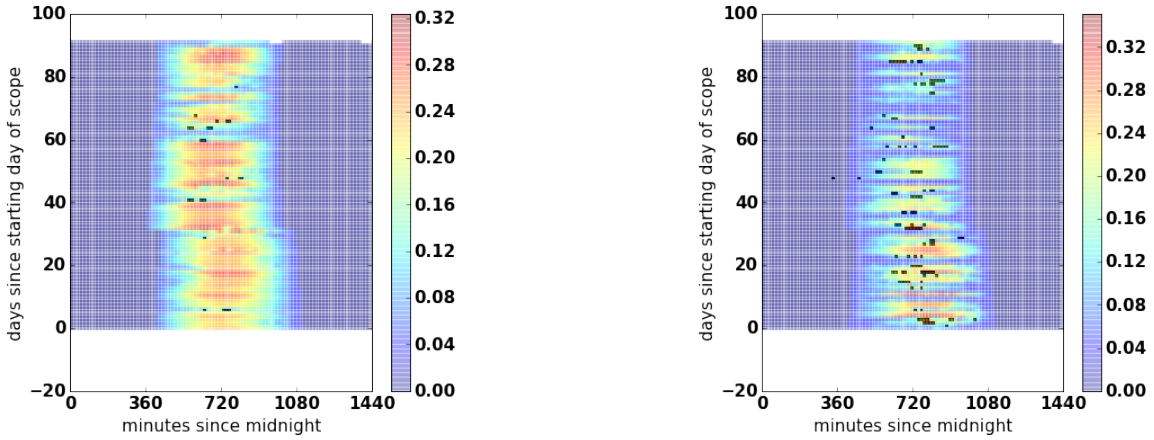
(a) Western (b) Eastern

Figure 10.1: Risk of demand-supply imbalance given a typical set of infrastructures

Here are some statistics of interest for the grid operator. For the western example, with the limit of 15 percent of the yearly peak load, PV systems provide 6.5% of the total energy demand in the simulation period. The maximum ratio between the actual PV power output and the load at any instance in the simulation equals 0.24. For the eastern example, with the limit of 9 percent of the yearly peak load, PV systems provide 4% of the total energy demand in the simulation period. The maximum ratio between the actual PV power output and the load at any instance in the simulation equals 0.16. Here one can see that there are various references one can use to communicate: the peak load, the total energy demand, and the actual load at each instance. Most of values in this paper make a reference to the peak load. This agrees with how the grid operator sizes the ancillary services capacity. However, in the operation with the high level of PV penetration, other types of references may be more informative.

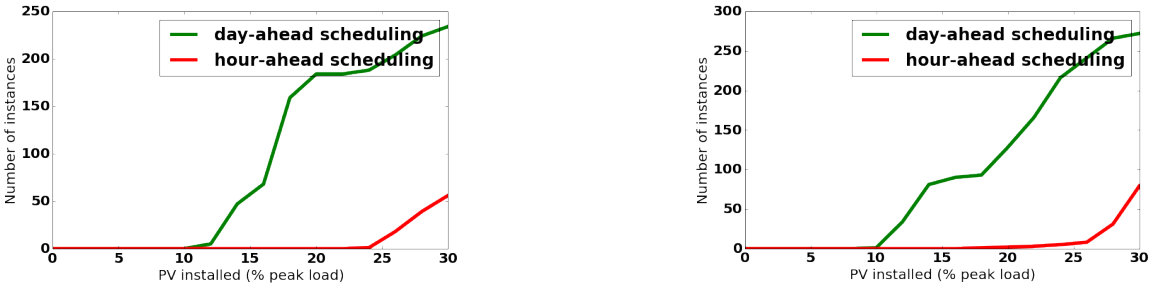
To see the distribution of demand-supply imbalances, one may plot when the demand-supply imbalances occur from the simulation. Figures 10.2 shows the distribution of demand-supply imbalances when the installed when the installed PV capacity is 20 percent of the peak load. Black dots show imbalances while the rainbow map is a ratio of the actual PV generation and the actual load. It is worth noting that it is possible that imbalance occurs even the level of PV penetration is 0%. However, in our test data set, it occurs only once for the Eastern example and none for the Western example. One may see from the figures that the demand-supply imbalance due to PV generation tends to happen from the late morning to the early afternoon.

To understand how variability and volatility in PV generation contribute to the risk of demand-supply imbalance, we look at the number of 15-minute instances when the maximum ramp rate of generators is reached in a day-ahead scheduling and an hour-ahead scheduling as a function of the installed PV capacity in the grid as shown in Figure 10.3.



(a) Western (b) Eastern

Figure 10.2: The distribution of demand-supply imbalances over time domain example where the level of PV penetration is 20% of the peak load



(a) Western (b) Eastern

Figure 10.3: Number of instances when the maximum ramp rate of generators is reached

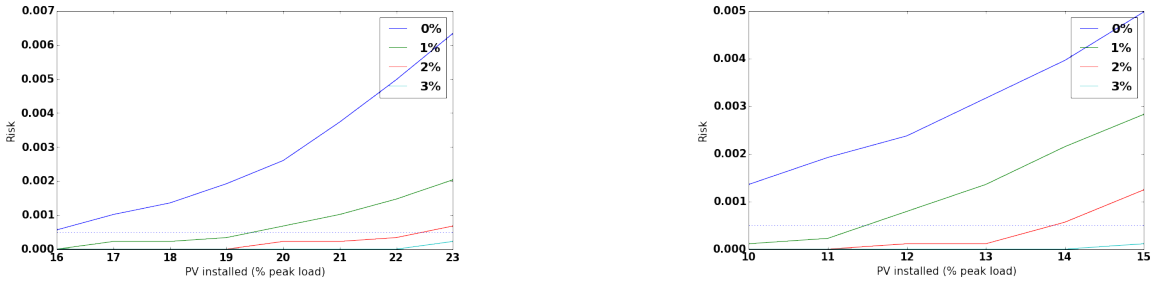
In the western example, an instance where the maximum ramp rate of generators is reached in the day-ahead scheduling occurs when the installed PV capacity is 10 percent of the peak load. However, there is no incidence where the maximum ramp rate is reached in the hour-ahead scheduling. Furthermore, the risk of demand-supply imbalance is acceptable. It means that the typical infrastructure can handle the variability of PV generation at this level of PV installation. The risk of demand-supply mismatch induced by the volatility of PV generation grows larger than the acceptable value when the installed PV capacity is 15 percent of the peak load. The risk induced by variability emerges when the maximum ramp rate is reached in the hour-ahead scheduling. It happens when the installed PV capacity is higher than 23 percent.

In the eastern example, an instance where the maximum ramp rate of generators is reached in the day-ahead scheduling occurs when the installed PV capacity is 10 percent of the peak load. At that time, there is no incidence where the maximum ramp rate is reached in the hour-ahead scheduling. However, the risk of demand-supply imbalance is higher than the acceptable value. It means that the typical infrastructure can handle the variability but not volatility of PV generation at this level of PV installation. The risk induced by variability emerges when the maximum ramp rate is reached in the hour-ahead scheduling. It happens when the level of installed PV is higher than 15 percent.

The observation of risk induced by variability and volatility is helpful in the risk reduction. In the western example, when the level of PV installation is in the range of 15 to 23 percent of the peak load, the grid operator may add ancillary services to reduce the risk due to the volatility. In the eastern example, when the level of PV installation is in the range of 10 to 15 percent of the peak load, the grid operator may provide additional ancillary services to decrease the risk without modifying generators. Such risk reduction by additional ancillary services is discussed in the next two sections.

10.5 Risk reduction by adding Type I ancillary services

A simple way to reduce the risk of demand-supply imbalance is to add more Type I ancillary services. In a typical infrastructure, most of available ancillary services are Type I. In the previous section, we set all ancillary services to be Type I with a capacity of 5 percent of the peak load. In this section, we explore the risk of demand-supply mismatch given different amounts of Type I ancillary services when the level of PV installation is 15-23 percent of the peak load in the western example and 10-15



(a) Western

(b) Eastern

Figure 10.4: Risk of demand-supply imbalance given a different capacity of additional Type I ancillary services

percent of the peak load in the eastern example. The same simulation was executed but the additional capacity of Type I ancillary services was varied from 0 to 3 percent.

Figure 10.4 shows the risk of demand-supply imbalance from the simulation as a function of the installed PV capacity in the grid given different amounts of Type I ancillary services. Each line represents each value of Type I ancillary services capacity relative to the peak load. In the western example, increasing the capacity of Type I ancillary services by 3 percent of the peak load reduces the risk of demand-supply imbalance down to the acceptable value when the installed PV capacity is 23 percent of the peak load. In the eastern example, increasing the capacity of Type I ancillary services by 3 percent of the peak load reduces the risk of demand-supply imbalance down to the acceptable value when the installed PV capacity is 15 percent of the peak load.

10.6 Risk reduction by adding Type II ancillary services

Another procedure to reduce the risk is to add Type II ancillary services. In the previous section, we set all ancillary services to be Type I with a capacity of 5 percent of the peak load. In this section, we explore the risk of demand-supply imbalance given a different amount of Type II ancillary services when the level of PV installation is 15-23 percent of the peak load in the western example and 10-15 percent of the peak load in the eastern example. We need to modify the simulation by adding Type II ancillary services and a control as follows:

1. At the beginning of the simulation period, the initial available energy of Type II ancillary service is at 50% of its total capacity.

2. Every 15 minutes, measure the difference between the net load and the assigned generation from scheduling.

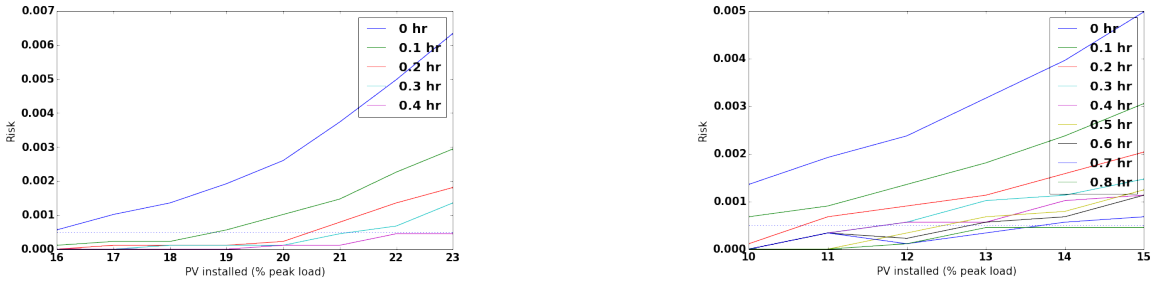
3. If possible, counterbalance the difference using Type II units. Change the amount of stored energy in the Type II units accordingly. if the capacity of Type II units is not enough, use Type I units. Record when combined ancillary services cannot counterbalance the difference.

4. After each counterbalance, prevent the Type II units from reaching either 0% or 100% of the capacity using type I units if possible. When the stored energy is below 20% or above 80% of the capacity, trigger the Type I units to use their remaining capacity after the counterbalance process to push the stored energy back to 50% of the capacity. Note that this adjustment may take several instances after the trigger.

Note that the capacity of Type I ancillary services is set at 5 percent of the peak load as it is in a typical infrastructure. The capacity of Type II ancillary services varies from 5 percent of the peak load multiplied by 0 hour to 5 percent of the peak load multiplied by 0.8 hours.

Figure 10.5 shows the risk of demand-supply imbalance from the simulation as a function of installed PV capacity in the grid given different amounts of Type II ancillary services. Each line represents each capacity value of the Type II ancillary services. In the western example, increasing the capacity of Type II ancillary services to 5 percent of the peak load multiplied by 0.4 hours reduces the risk of demand-supply imbalance down to the acceptable value when the installed PV capacity is 23 percent of the peak load. In the eastern example, increasing the capacity of Type II ancillary services to 5 percent of the peak load multiplied by 0.8 hours reduces the risk of demand-supply imbalance down to the acceptable value when the installed PV capacity is 15 percent of the peak load.

Adding Type II ancillary services has an advantage over adding Type I ancillary services because it may not require extra energy to execute demand-supply balancing. While a load shifting in demand response induces no energy loss, an electrical storage may lose few percentage of energy during charging process. If Type II ancillary services are not deployed, the Type I ancillary services will be called at all instances. However, if Type II ancillary services are deployed, the Type I ancillary services will be called less often. For example, in the simulation for the western example with additional Type II capacity of 5 percent of the peak load multiplied by 0.4 hours, the type I ancillary services are called only 1639 instances out of $3 \times 30 \times 24 \times 4 = 8640$ instances when the installed PV capacity is at 23 percent of the peak load. This implies that the operational cost of Type I ancillary services is likely reduced while the capital cost of type I ancillary services remains.



(a) Western

(b) Eastern

Figure 10.5: Risk of demand-supply imbalance given a different capacity of Type II ancillary services

A question one might ask is whether it is possible to limit the risk of demand-supply imbalance using the Type II ancillary services only. The answer is no. Because the bias of any predictive model of the net load is small yet non-zero, the energy in type II ancillary services to accommodate the bias eventually hits either zero or one hundred percent of its total capacity. There must be a Type I ancillary service to actively offset such the bias. Hence, there should be an optimal combination of Type I and Type II services that minimizes the overall cost of accommodating high distributed PV generation. There should also be a better control policy of those services than our current control procedure.

The optimal combination of infrastructures and control policy is not an easy problem to solve. One of the reasons is that the true operation cost of Type II services is not fully known. For instance, it is commonly known that frequent charge-discharge switching and fully discharging down to zero can shorten the lifetime of some energy storage devices. However, there is yet no quantitative function for how these operations decrease the lifetime of an energy storage device and increase the overall cost. To avoid frequent charge-discharge switching, one may constrain the number of cycles that the storage can be switched. Another heuristic is to assign one group of electric storage responsible for charging and another group for discharging for a certain period such as 12 hours. To avoid fully discharging, one may constrain minimum level of discharge in the operation. These procedures may provide a legitimate control policy, but there is no guarantee of optimality. A similar situation arises with Type I ancillary services since the frequency of their use may change significantly from the current use. The current cost function of Type I ancillary services, consisting of just the capital cost and the fuel cost, should be reformulated as well. A more rigorous formulation of the problem and its solution should be investigated once the complete cost function of both Type I and Type II ancillary services is known, but that is beyond the scope of this study.

Western	Volatility of the most recent hour				
Level (% Peak load)	0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	> 0.4
13	0	0	0	0	0
14	1	0	1	0	0
15	1	0	1	0	0
16	2	1	2	0	0
17	3	1	3	2	0
Eastern	Volatility of the most recent hour				
Level (% Peak load)	0-0.1	0.1-0.2	0.2-0.3	0.3-0.4	> 0.4
7	0	0	0	0	0
8	0	0	0	2	0
9	0	1	0	3	0
10	2	2	1	5	1
11	2	2	3	6	3

Table 10.2: The number of demand-supply imbalance instances associated with different ranges of normalized volatility from various levels of PV installation. The rows represent levels of PV installation, and the columns represent ranges of normalized volatility.

10.7 Risk reduction by PV curtailment

Other than adding ancillary services, the grid operator may curtail PV generation in order to reduce the risk of demand-supply imbalance. It is known that the volatility is closely related to the prediction performance, and the effect of prediction errors scales with the amount of PV capacity connected to grid. A simple strategy is to disconnect some PV panels when the volatility is high in order to prevent a potential imbalance. This strategy leads to a couple of questions. What should be the threshold of normalized volatility to trigger the PV curtailment and how much capacity should be curtailed?

To answer those questions, one should consider the distribution of normalized volatility measured when a demand-supply imbalance occurs. Table 10.2 shows the number of demand-supply imbalance instances associated with different ranges of the normalized volatility from various levels of installed PV capacity. The numbers in Table 10.2 are counted from the simulation with a single set of data of 3 months. One may see that the distribution as a function of volatility in this table is not quite smoothly monotonic. To obtain a better distribution, one may find a probabilistic model of each component in the grid operation and perform a Monte-Carlo simulation

to produce a large sample and average the number of demand-supply imbalances in each entry of the table. However, we will not perform this task in the study.

Since an acceptable risk is defined to be 0.05 %, for 3 months of our simulation period, we want to keep the number of such instances less than $0.05\% \times 3 \times 30 \times 24 \times 4 \approx 4$. We can control the total number of instances by operating the PV systems at a high level of penetration when the volatility is small, and operating the PV systems at a low level of penetration when the volatility is large. For example, in the western example, one may install PV systems with the installed capacity being 17 % of the peak load. If the volatility is lower than 0.1, the grid operator does not curtail any PV generation. Then the PV capacity is curtailed down to 15 % of the peak load whenever the normalized volatility is higher than 0.1. In this way, the total number of instances from this policy highlighted in green becomes $3 + 0 + 1 + 0 + 0 = 4$.

It is worth noting that implementing PV curtailment is challenging. First, in order to disconnect and reconnect PV systems to the grid, the two-way automatic communications between panels and the grid operators must be mandatory. Second, the grid operator needs to pay PV system owners since the curtailment becomes an ancillary service for the grid operation. Finally, PV curtailment by disconnecting too many PV systems may lead to significant PV energy loss. The grid operator needs to coordinate extra generators in their hour-ahead scheduling to make up the loss.

10.8 Conclusion

In this chapter, we illustrate the risk assessment of demand-supply imbalance under different levels of distributed PV generation. With a typical set of infrastructures, the grid can operate within an acceptable risk when the installed PV capacity is up to 15% relative to the peak load for the western example, and 9% for the eastern example. By adding Type I ancillary services or Type II ancillary services coupled with a simple control heuristic, one may increase the PV installation within an acceptable risk. We found that the grid may have the PV capacity up to 23 % relative to the peak load for the western example and 15% for the eastern example. Beyond this level, an extra modification on generators in a day-ahead and and hour-ahead scheduling is required. Another way to increase the maximum level of PV installation without increasing the risk is to implement the PV curtailment with a policy based on the normalized volatility of PV generation. It allows some extra PV installation of about 2% of the peak load. Still, there are several challenges to implement the PV curtailment. An indirect control such as real-time pricing is another way to manage the risk of demand-supply imbalance. Overall, any policy has diminishing returns as we try to increase the level of PV installation into the grid. The economical aspect of these policies is left as an open question.

Conclusions

Distributed PV generation is growing fast due to government policy, technology, and finance innovation, and it is still far from the upper bound. However, despite its benefits, distributed PV generation introduces challenges to the grid operator due to the particular nature of distributed and solar aspects of the generation. Distributed PV generation will have a limited but non-trivial role in an electric grid. The major achievement of this thesis is to model distributed PV generation in a manner that enabled us to estimate the limit of PV generation under the current physical structure and operational policy.

To set up a framework, we reviewed the current PV system modeling in the first part of this study. We found the notion of the maximum power curve is useful in estimating the bound of distributed PV installation and other applications. However, that it is not practical to construct the maximum power curve based on the structural approach. The data-driven model based on the power output from a sample of currently installed PV systems is more useful in this problem. We constructed the maximum power curve using time series of power outputs and its total installed capacity only. From a set of daily maximum normalized power curves for all the days in a year, we established the first upper bound of distributed PV generation into the grid by comparing it with a set of daily load curves of the entire grid. It is worth noting that this upper bound emerges without consideration of the volatility aspect of distributed PV generation.

In order to take the volatility aspect of distributed PV generation into consideration, we looked into a grid operation in the daily and hourly time-scale. Since this operation involves a forecast, we built an example of forecast systems for both PV generation and load. In the first part of study, we showed that the maximum PV power curves can be used to improve the accuracy of a PV power forecast. In addition, we clearly defined the volatility of distributed PV generation and confirmed that the volatility correlates with the prediction performance. In the second part of the study, we reviewed examples of load modelling and constructed a simple forecast model of load. With a reference curve from historical data, which is analogous to the maximum PV power curve, we obtained a desirable prediction performance even without additional information. In the third part of this study, we simulated the grid operation

comprising of energy scheduling, ancillary services control, and the net load forecast systems. By varying the amount of installed PV systems in the grid, we measured the probability that the demand-supply imbalance cannot be coped using the current infrastructures. We then established the second upper bound of distributed PV generation into the grid by limiting the probability. Furthermore, our result shows that the current limiting factor to the level of PV installation is in the ancillary service rather than generators in energy scheduling. In order to increase PV installation, we have a choice of curtailing PV generation or adding additional ancillary services. The modification on generators in energy scheduling is not required.

Other than establishing the limit of distributed PV generation to the grid under a typical set of infrastructures and suggesting how to increase the limit, our work can be helpful in planning and policy making in many ways. First, from a simple scheme to minimize the cost of grid operation based on the shape of expected net load curves in Chapter 3, policy makers may regulate or recommend installers to put solar panels in the direction that minimizes the cost of grid operation. The policy makers may encourage or discourage installing solar panels for different areas based on this optimization scheme as well.

Second, with expected net load curves and hourly operation analysis, an electricity market can propose a fair real-time price. Since the volatility and variability are distinguishable and affect different mechanisms, namely energy scheduling and ancillary services, the market can set a part of the price to reflect the cost of generators based on the expected net load curves, and another part of the price to reflect the cost of ancillary services based on the volatility measure. This pricing can motivate consumers to use electricity at the right time where the overall cost of electricity is the cheapest.

Third, once the limit of PV installation is known, the policy makers may discourage extra installation by expiring some incentives. If the level of PV systems still increases, the grid operator can modify the grid infrastructure as we suggested in Chapter 10. The grid operator may pursue the curtailment procedure based on the volatility as described in Chapter 10 directly or use the price mechanism to encourage PV owners with storage to store the power when it is volatile instead of putting it into the grid. It should be noted that the policy makers may mandate that an electrical storage needs to be equipped with any additional PV system beyond the limit.

It is worth noting that, in order to provide a legitimate policy, there might be a need to check the consistency of the estimated distributions over the solar output and load over time as structural factors like climate change will likely change both distributions. The re-estimation of both distributions should be conducted regularly anyway as the system configuration and characteristic of the energy customer base change. Accounting climate change explicitly might be helpful especially in a longer range capacity planning.

There are two possible directions worth further investigation. The first direction is a detailed analysis where the network constraint is taken into consideration. In this work, the distributed PV power analysis in this study is modeled from a group of PV systems equivalent to the area under a substation, while the load analysis in this study is modeled from a whole grid. In practice, we suggest that both distributed PV power and load should be analyzed with our approach for each substation. An individual net load analysis can then be conducted to get a better estimate for maximum PV installation if a back flow is not allowed. Once analysis for each substation is conducted, one may find a more realistic estimate for maximum PV installation by taking the power flow equations into account. The concern here is not only the balance of the demand and the supply of the whole grid, but also transmission constraints within the grid itself. As the net load is equal to the normal load subtracted by the distributed PV generation at each node, the net load becomes the source of uncertainty in the power flow equation. One may assess the risk of the grid failure from either demand-supply imbalance or breaking constraints based on the simulation as we have done in this work. Then the limit of distributed PV installation is the amount corresponding to the maximum tolerable amount of risk. Another approach is to turn the problem into a contingency analysis where the net load ranges between extreme values. The limit of PV installation is then the maximum amount such that the set of power flow equations with extreme values of the load is still feasible. Due to the limited access on the load data under a substation, we were not able to pursue this direction.

The second direction is a cost-benefit analysis of adding distributed PV generation into the grid. In this study, we tend to find a solution favorable to PV industry without causing a major risk to the grid operation. However, it may not be eventually beneficial to all customers in the grid to have a large amount of distributed PV generation. As we push additional PV installations to the grid, the grid operator needs to provide extra ancillary services or curtailment policy in order to reduce the risk of demand-supply imbalance. In this work, we did not answer if it is profitable to modify the grid to accommodate the increased PV installation. If it is profitable, the next question is how to modify and operate the grid with the least cost. Such the question involves a dynamic programming problem to minimize the cost of operation. However, the cost function of Type II reserves, especially an electric storage, is complicated and unsettled. Even the cost function of Type I reserves under high volatility is under question, not to mention the cost of curtailment execution which involves the loss in benefit of PV owners. The problem remains unsolved until we are able to systematically quantify these costs.

Distributed PV generation is a choice of sustainable energy resource. However, it cannot be the only the sole supplier. As addressed in this study, the variability and the volatility nature become limiting factors of how much distributed PV generation

can be deployed to the grid. In order to have more distributed PV generation or any other intermittent renewable resource, one should look at a bigger picture than the electricity sector. The whole landscape of energy should be taken into consideration. Nowadays, water heating systems in household and in traditional turbine power plant can be integrated to alleviate the variability and the volatility from intermittent renewable resources. In the future, we hope the electrification of the transportation sector may be a part of solution as well. With a holistic view and a systematic management, the energy industry can be both sustainable and reliable.

Bibliography

[1] Inc./Architecture 2030 2030. Why the building sector?

[2] US Eenergy Information Administration. Glossary: Generator nameplate capacity (installed).

[3] Ted James Alan Goodrich and Michael Woodhouse. *Residential, Commercial, and Utility-Scale Photovoltaic (PV) System Prices in the United States: Current Drivers and Cost-Reduction Opportunities*. 2012.

[4] Hesham K. Alfares and Mohammad Nazeeruddin. Electric load forecasting: Literature survey and classification of methods. *International Journal of Systems Science*, 23(1):23–34, 2002.

[5] Dominic Maratukulam Alireza Khotanzad, Reza Af'khami-Rohani. Annstlf - artificial neural network short-term load forecaster - generation three. *IEEE Transactions on Power Systems*, 13(4), 1998.

[6] Shawn Kerrigan Andrew Goldstein, Alexander Thornton. Volatility and deviation of distributed solar.

[7] S. Khatoon Md. Muazzam D. K. Chaturvedi Arunesh Kumar Singh, Ibraheem. An overview of electricity demand forecasting techniques. *National Conference on Emerging Trends in Electrical, Instrumentation Communication Engineering*, 3(3), 2013.

[8] Energy Storage Association. Energy storage: facts and figures.

[9] Solar Energy Industries Association. 2014 top 10 solar states.

[10] Solar Energy Industries Association. Impacts of solar investment tax credit extension.

[11] Solar Energy Industries Association. Solar industry data.

[12] Solar Energy Industries Association. Solar market insight report 2014 q4.

[13] Seth Blumsack. Regional transmission organizations.

[14] US Census Bureau. Population estimates. state totals: Vintage 2015.

[15] California Indepedent System Operator (CAISO). What the duck curve tells us about managing a green grid.

[16] California Independent System Operator (CAISO). California iso peak load history 1998 through 2014.

[17] N.C. Clean Energy Technology Center. Arizona's renewable portfolio standard.

[18] N.C. Clean Energy Technology Center. New jersey's renewable portfolio standard.

[19] California Public Utilities Commission. *Renewable Portfolio Standard, Quarterly Report 4th quarter 2014*. 2005.

[20] Rachel Peterson. California Public Utilities Commission. California distribution system interconnection standards: Proposed technical advances and policy considerations relevant to proposed rulemaking to address small generator interconnection procedures, 2012.

[21] Jay A. Kratochvil David L. King and William E. Boyson. Temperature coefficients for pv modules and arrays: Measurement methods, difficulties, and results. *Sandia National Laboratories. Presented at 26th IEEE Photovoltaic Specialists Conference*.

[22] Ameya Deoras. Electricity load forecasting using neural networks.

[23] Abdanour Irbah Djelloul Djafer. Estimation of atmospheric turbidity over ghardaa city. *Atmospheric Research, Elsevier*, 128:76–84.

[24] US EIA. Delivered energy consumption by sector, annual energy outlook 2015.

[25] US Energy Information Administration (EIA). Increased solar and wind electricity generation in california are changing net load shapes.

[26] U.S. Energy Information Administration (EIA). Noncoincident peak load by north american electric reliability corporation assessment area, 2003 - 2013.

[27] Gary Kuzkin. Schneider Electric. Solar systems for existing residential installations.

[28] Energy and Inc. Environment Economics. *Technical Potential for Local distributed Photovoltaics in California*. 2012.

[29] Nivad Navid J. Charles Smith National Renewable Energy Laboratory Erik Ela, Brendan Kirby. Effective ancillary services market designs on high wind power penetration systems. 2012.

[30] Anthony Lopez et al. *US Renewable Energy Potential Analysis: A GIS-based analysis, NREL technical report*. 2012.

[31] Pelland et al. Photovoltaic and solar forecasting: State of the art. *IEA PVPS Task 14, Subtask 3.1 Report IEAPVPS T1401: 2013*, 2013.

[32] Federal Energy Regulatory Commission (FERC). Electric power markets: National overview.

[33] Solar Energy Services for Professionals. Solar energy services for professionals.

[34] Fraunhofer-Institut fr Solare Energiesysteme ISE. Stc measurements.

[35] Pacific Gas and Electric Company. Understand net energy metering (nem) and your bill.

[36] E. Handschin and Ch. Dornemann. Bus load modelling and forecasting. *IEEE Transactions on Power System*, 3(2):627–633, 1988.

[37] Perez R Hoff, T. Quantifying pv power output and variability. *Solar Energy*, 84(10), 2010.

[38] Tao Hong. Load forecasting case study. *Eastern Interconnection States Planning Council (EISPC) in response to the NARUC solicitation NARUC-2014-RFP042DE0316*, 2015.

[39] Wei-Chiang Hong. Electric load forecasting by support vector model. *Applied Mathematical Modelling*, 33(5):24442454, 2009.

[40] Pierre Ineichen. Long term satellite global, beam and diffuse irradiance validation. *Energy Procedia, Elsevier*, 48:15861596.

[41] Obadiah Bartholomy Thomas Vargas Tom Hoff Richard Perez James Bing, Pramod Krishnani.

[42] The Washington Post Joby Warrick. Utilities want solar unplugged.

[43] Kevin Carden Johannes P. Pfeifenberger. Resource adequacy requirements: Reliability and economic implications. *The Brattle group and Astrape Consulting*, 2013.

[44] The Wall Street Journal. Solarcity buys silevo, a module maker.

[45] Chris Kavalec and Tom Gorin. California Energy Commission. *California Energy Demand 20102020, Adopted Forecast*. 2009.

[46] National Renewable Energy Laboratory. The open pv project.

[47] Folsom Labs. Module model.

[48] Katie Lynch. Quick summary: The 50 states of solar.

[49] Joshua S. Stein Matthew J. Reno, Clifford W. Hansen. *Global Horizontal Irradiance Clear Sky Models: Implementation and Analysis*. 2012.

[50] GTM Research MJ Shiao. Pv balance of systems 2015: Technology trends and markets.

[51] Western Electricity Coordinating Council Modeling and Validation Work Group. *Composite Load Model for Dynamic Simulations*. 2012.

[52] North American Electric Reliability Corporation (NERC). Nerc interconnections.

[53] North American Electric Reliability Corporation (NERC). Regions and balancing authorities.

[54] Panasonic Corporation of North America. End-to-end solar.

[55] Hiren Patel and Vivek Agarwal. Matlab-based modeling to study the effects of partial shading on pv array characteristics. *IEEE TRANSACTIONS ON ENERGY CONVERSION*, 23(1):302–310, 2008.

[56] Easan Drury Paul Denholm and Robert Margolis. *Solar Deployment System (SolarDS) Model: Documentation and Results, NREL technical report*. 2009.

[57] PJM. How energy use varies with the seasons.

[58] International Energy Agency Photovoltaic Power Systems Programme. Snapshot of global pv 1992-2014, 2015.

[59] Clean Power Research. Solaranywhere: Overview.

[60] GTM Research. Pge takes energy storage to the distribution substation.

[61] Clifford Hansen Robert W. Andrews, Joshua S. Stein and Daniel Riley. Introduction to the open source pv lib for python photovoltaic system modelling package. *Sandia National Laboratories*.

[62] PVsyst SA. Pvsyst photovoltaic software.

[63] scikit-learn developers (BSD License). Epsilon-support vector regression.

[64] scikit-learn developers (BSD License). Gradient boosting for regression.

[65] Kittipong Methaprayoon Shu Fan and Wei-Jen Lee. Multiregion load forecasting for system with large geographical area. *IEEE transactions on industry applications*, 45(4), 2009.

[66] Bethany Speer. *Residential Solar Photovoltaics: Comparison of Financing Benefits, Innovations and Options*. 2012.

[67] Mike Taylor Steven Letendre, Miriam Makhyoun. Predicting solar power production: Irradiance forecasting models, applications and future prospects. *Solar Electric Power Association*, 2014.

[68] US Department of Energy Sunshot Initiative. *SunShot Vision Study*. 2012.

[69] Shu Fan Tao Hong, Pierre Pinson. Global energy forecasting competition 2012. *International Journal of Forecasting*, 30(2):357–363, 2014.

[70] Miguel Angel Ortega Vanquez. *Optimizing the Spinning Reserve Requirements*. 2006.

[71] Wei-Jen Lee Mo-Shing Chen P. Kotas P. Van Olinda W. Charytoniuk, E.D. Box. Neural-network-based demand forecasting in a deregulated environment. *IEEE Transactions on Industry Applications*, 36(3):893 – 898, 2000.

[72] S. Wilcox and W. Marion. *Users Manual for TMY3 Data Sets: A Technical Report*. 2008.

[73] Antonio T. Lorenzo Joshua S. Stein William F. Holmgren, Robert W. Andrews. Pvsc 2015.

[74] Daniel Kirschen Yann Rebours. What is spinning reserve?