A PRODUCTION SIMULATION TOOL FOR SYSTEMS WITH INTEGRATED PHOTOVOLTAIC ENERGY RESOURCES

BY

RAJESH BHANA

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Adviser:

Professor George Gross

ABSTRACT

Climate change awareness, the drive to sustainability and the push for energy independence have resulted in the wider utilization of renewable energy sources. Photovoltaic (PV) power is a renewable solar energy source that is increasing in use as its costs decrease. We have developed an assessment and planning tool to quantify the longer-term variable effects of large-scale PV energy production on power systems. The tool consists of a probabilistic model for the representation of the power output of the PV resources, at single or multiple sites on the network, and the incorporation of the model into an extended probabilistic simulation framework for the evaluation of the longer-term variable effects of the resources. We develop the probabilistic model in a manner that captures the time-dependent, variable and intermittent nature of the PV resources and incorporate the model into the extended framework in a way that captures the correlation between the chronological load and the uncontrollable PV output. Because typical daily PV output patterns vary markedly over a year, we construct the probabilistic model on a seasonal basis, and because the duration and magnitude of PV output changes from day to day, we scale the daily PV output patterns, both in time and magnitude, into scaled-output characterizations that allow the meaningful comparison of different days of the season. We then classify the seasonal set of daily PV scaled-output characterizations into pattern cluster sets of days with similar daily output patterns before re-scaling the pattern clusters into class sets of daily PV output representations. From these class sets, we approximate the conditional probability distributions of the PV output random variables conditioned on each pattern class. We extend the conventional probabilistic simulation framework to incorporate the PV resources by using these approximations of the conditional probability distributions. By incorporating the probabilistic PV output model into the extended framework, we are able to quantify the longer-term variable effects of a high penetration of large-

scale PV resources on a system in terms of reliability, economic-cost and environmental-impact metrics. To test the capability of the tool, we have applied the methodology on a variety of test system cases covering a wide span of load, system and resource characteristics. The application of the tool is useful in many areas of power system planning, including investment decision-making and policy formulation.

For Mum and Dad

ACKNOWLEDGMENTS

I would like to thank Professor George Gross for his direction in the research and the writing of this thesis. I also express my thanks to Yannick Degeilh for his helpful insights and discussion throughout the research process and acknowledge Nicolas Maisonneuve, whose previous research was instrumental in the undertaking of my own.

I would like to thank Fulbright New Zealand for making it possible for me to travel to the United States, immerse myself in American culture, further my education and undertake this research at the University of Illinois at Urbana-Champaign. I express my thanks to my many colleagues, past and present, in the university's Power and Energy research group, especially Christine Chen and Christopher Reeg. You are all great friends.

I am also grateful for the support of my former colleagues at Beca and my friends from the University of Auckland. In particular, I am grateful for the senior guidance of Kevin Allen, Stephen Salmon and Mark Andrews and for the continual friendship and advice of Christopher Rapson, Claudio Camasca and Saurabh Rajvanshi.

Thank you to my family in New Zealand. Being so far from you is difficult, but your love, support and never-ending encouragement make it possible. To my sister, Hema Bhana, you could not have been more helpful along the way.

Finally, I thank my lovely fiancée and her supportive family. Sheila McAnaney, you are amazing.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

In this thesis, we present the development of a probabilistic model for the representation of the power output of large-scale photovoltaic (PV) resources and its incorporation into a probabilistic simulation framework for the evaluation of the longer-term variable effects of PV resources on a power system. In this introductory chapter, we describe the motivation for this research, review the state of the art, provide an overview of our proposed methodology and present a detailed outline of the rest of the thesis.

1.1 Motivation

Climate change awareness, the drive to sustainability and the push for energy independence have resulted in wider utilization of renewable energy sources over the past decade [1]. Government incentives and mandates for renewable energy production, such as the renewable portfolio standards of various states in the United States [2], are some of the many indications that renewable resource penetration, particularly wind and solar resource penetration, will continue to deepen. PV power, though not currently as prevalent as wind power, is a solar energy source that continues to increase in utilization as its costs decrease [3]. The deeper penetration of renewable resources creates the need for assessment and planning tools to quantify the longer-term variable effects of renewable energy production on power systems.

Probabilistic tools are used throughout the electric power industry to aid in the assessment and planning of power systems with the various sources of uncertainty explicitly represented [4]. For example, the tools may be used to estimate fuel requirements or the emission of various effluents. Such tools use a probabilistic simulation framework to calculate the expected economic, environmental and reliability metrics that measure the variable effects over

a specified period. Conventional probabilistic simulation tools emulate systems with controllable generators that have controllable output and constant outage rates over the simulation period. PV resources are variable, uncontrollable, time-dependent, intermittent and not accurately predictable. As a result, they cannot be modeled in the same way as conventional resources in a probabilistic simulation framework. In Fig. 1.1, we present eight daily PV plant output patterns computed from real-world atmospheric measurements as described in Appendix A. Clearly, the output is variable and intermittent. The output patterns on March 2 and May 31 have high output with gradual output change, whereas the patterns on April 23 and May 1 change rapidly from minute to minute at similar magnitudes. The outputs on March 6 and May 3 have increased fluctuation at different times of the day and the outputs on March 7 and April 20 change rapidly at medium-to-low levels. In Fig. 1.2 we show the variation in sunrise times, sunset times and daily maximum power output for the same PV plant over the course of a year. It is clear that the variation exists throughout the year. When a PV plant is integrated into the grid, the variation of the plant's output affects the operation of the other controllable generators, and thus, simulation tools need to capture the time-dependent, variable and intermittent nature of the PV resources. In Fig. 1.3, we provide an example of a week of hourly load and PV resource output. We observe a positive correlation between the daily PV output and the chronological load. PV output only occurs during the day and this generally coincides with the higher load hours. For that reason, in addition to capturing the nature of the PV resources, the probabilistic simulation tools must also capture the relationship between the load and the PV output.

1.2 Overview of the state of the art

Researchers have developed many techniques for the modeling of PV resource output for longer-term planning. Many of these techniques, such as those described in [9, 10, 11, 12], focus on energy analysis, considering only the energy yield from the PV resource and not the system into which the resource is integrated. Some techniques, such as the methods described in [13, 14, 15], probabilistically model the load, the controllable resources and the PV re-

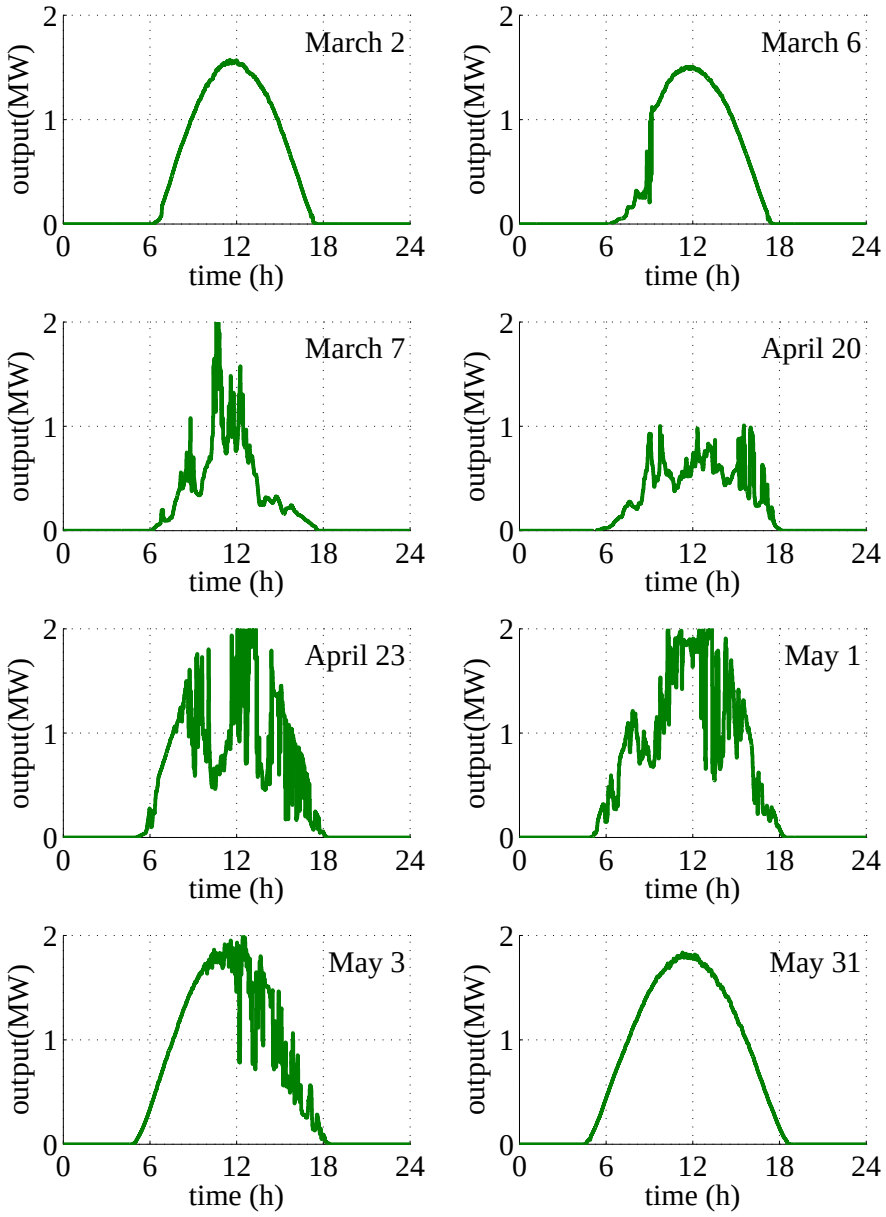
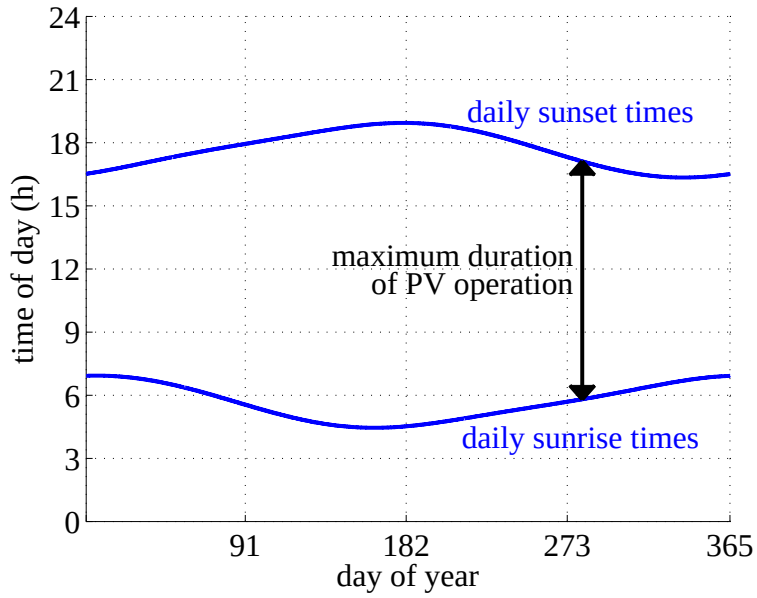
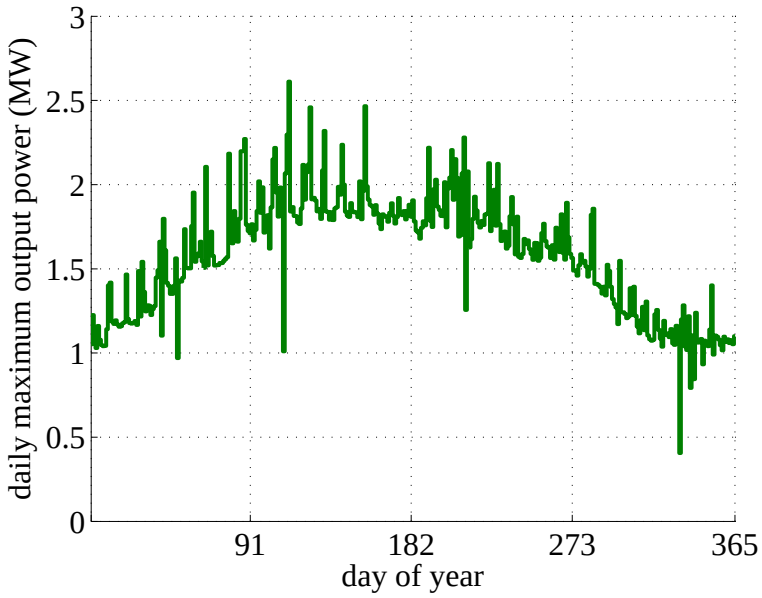


Figure 1.1: Daily PV plant power output patterns for eight days of the year, computed using one-minute resolution atmospheric data collected at Las Vegas, NV by NREL [5].



(a)



(b)

Figure 1.2: The variation of (a) sunrise/sunset times during a yearly cycle, and (b) maximum daily output over a year for a PV plant computed using one-minute resolution atmospheric data collected at Las Vegas, NV and published by NREL [5].

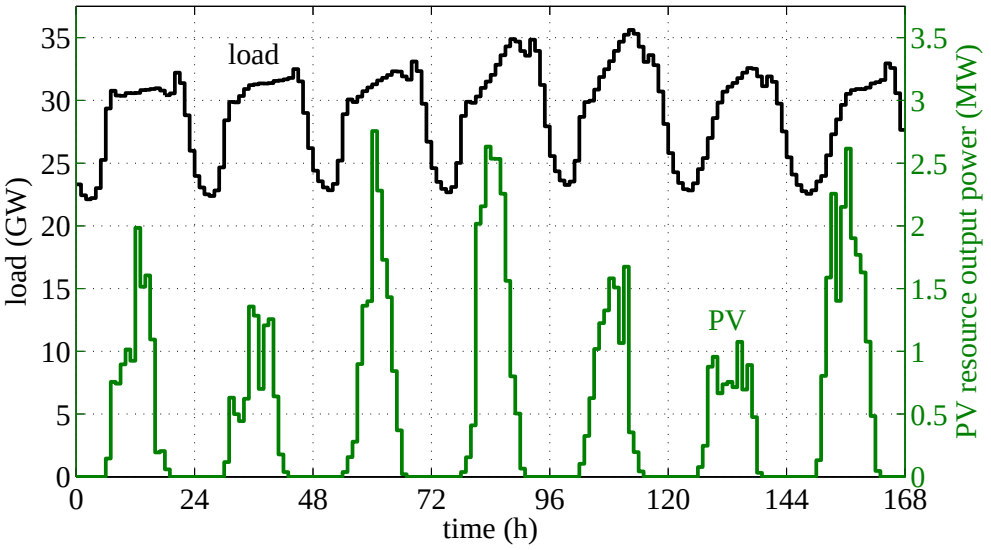


Figure 1.3: One week of ERCOT hourly load [6] and PV output for a plant computed using one-hour resolution atmospheric data collected at Abilene, TX and by NREL [7] and NOAA [8].

source output, but do not consider the time-dependence of the PV output and its interplay with the chronological load. The methods in [16, 17, 18], for example, capture the relationship between time-dependent renewable resource output and the load, but do so using Monte Carlo methods that are computationally demanding when simulating larger power systems. Complex PV output characterizations, such as those described in [19], are also computationally demanding; and techniques for the prediction of PV output, such as those described in [20, 21], are useful only for short-term operational decisions. Such methods are inappropriate for longer-term planning. The approach presented in [22] allows us to determine probabilistically the variable effects of time-dependent wind resources, but it is not suitable for the modeling of PV resources, whose output duration varies from day to day. The electric power industry recognizes the need for new methods to gauge the impacts of uncontrollable renewable energy sources [23], and this need will intensify as renewable energy penetration deepens.

1.3 Overview of the proposed methodology

We propose a probabilistic PV resource output model (PPVOM) that captures the seasonal and diurnal characteristics of PV resource output, when integrated into a power system, and thus explicitly represents the time-dependent, variable and intermittent nature of the resources. We also propose the computationally tractable incorporation of this PPVOM into an extended probabilistic simulation framework, which captures the interplay between the PV resource output and the chronological load. Unlike the prediction of short-term power output, our PPVOM takes the average of many possible outcomes over a longer period and – unlike the long-term prediction of energy production – captures shorter-term changes in PV resource output. As a result, our methodology is best suited for application to longer-term system planning. The incorporation of our PPVOM into the probabilistic simulation framework, while maintaining the necessary time-dependence, is crucial to obtaining the realistic measures of the variable effects. Consequently, we employ an incorporation methodology based on that described in [22], in which a wind resource output model is incorporated into the extended probabilistic simulation framework.

We construct the PPVOM in three separate steps. In step one, we scale a seasonal set of daily PV output patterns, both in time and magnitude, into characterizations that allow the meaningful comparison of the output patterns on different days of the year. In step two, we classify the scaled-output characterizations into distinct clusters that contain similar output patterns. In step three, we scale the scaled-output characterizations of each cluster into a class set of PV output representation patterns with appropriate magnitudes, durations and resolutions. We approximate the class-conditioned, cumulative probability distributions of the power output random variables from the class sets of re-scaled PV output representations. We extend the conventional probabilistic simulation framework [4] to incorporate the power output random variables in a manner that maintains the time-dependence of the PV resource output and its ability to be correlated to the chronological load. The extension of the framework is accomplished by the effective application of conditional probability concepts to the probabilistic representation of the load random variable. We apply the extended framework to compute the variable effects of interest in the form of system metrics. After we com-

pute the metrics for each pattern class, we utilize conditional probability to determine the seasonal system metrics.

1.4 Contents of the remainder of the thesis

We organize the thesis as follows. In Chapter 2, we describe the construction of the PPVOM. We first describe the collection of the input dataset and then detail the steps of scaling, clustering and re-scaling of the daily PV output patterns for the approximation of the power output random variables. In Chapter 3, making detailed use of conditional probability concepts, we describe the extension of the probabilistic simulation framework to incorporate the PPVOM. In the same chapter, we describe how we apply the extended framework for a case study.

In Chapter 4, we provide the details, the results and our analyses of a number of representative simulation case studies, which demonstrate the capability of our proposed methodology. We then provide concluding remarks and outline future research topics in Chapter 5.

In Appendix A, we describe the instantaneous PV output model we use to convert atmospheric data into PV output data. In Appendix B, we describe the computation of benchmark PV output for the scaling and re-scaling processes. Finally, in Appendix C, we detail the mechanics of our PV output scaling and re-scaling algorithms.

CHAPTER 2

THE PROBABILISTIC PHOTOVOLTAIC RESOURCE OUTPUT MODEL

In this chapter, we present the development of a probabilistic model for the large-scale PV resources sited at a single or multiple sites and integrated into a power system: the PPVOM. With this model, we capture the time-dependent, variable and intermittent nature of the PV resources. We describe the three development steps: the scaling of daily PV output patterns for meaningful comparison of the output on different days of the year, the classification of daily scaled-output patterns into pattern clusters and the re-scaling of pattern clusters into pattern classes for the approximation of the class-conditioned distribution functions of the PV output random variables. We begin the development with a description of the input to the probabilistic model.

2.1 Data needs for the construction of the probabilistic model

PV output patterns vary significantly from season to season. For example, cloudy days with lower PV output may occur frequently in the Fall season and rarely in the Summer season. As such, we build the probabilistic PV output model on a seasonal basis and the initial step is the collection of a seasonal data set. For a system with PV resources at S sites, we consider PV resource power output data where M values are used to describe the output at a single site over a midnight-to-midnight period comprised of M equal-duration sub-periods. Let $u_{s,m}^{(d,a)}$ denote the average PV resource output power at site s over sub-period m of day d in year a. We construct the vector $\underline{u}_s^{(d,a)} \triangleq [u_{s,1}^{(d,a)}, u_{s,2}^{(d,a)}, \ \ldots \ , u_{s,M}^{(d,a)}]^T \in \mathbb{R}^M$ to rep resent the PV output pattern at the site s over day d in year a. We create such vectors for each site $s = 1, 2, \ \ldots \ S$, and we then construct the super-vector

$\underline{u}^{(d,a)} \triangleq [\underline{u}_1^{(d,a)^T}, \underline{u}_2^{(d,a)^T}, \ \ldots \ , \underline{u}_S^{(d,a)^T}]^T \in \mathbb{R}^{(S \cdot M)}$ to represent the midnight-to-midnight PV output at all the S sites on day d in year a. We construct the seasonal, input data set $\mathcal{U}$ corresponding to all the days in the same season of interest that constitute the set $\mathcal{D}$ and for all the years available denoted by $\mathcal{A} = \{a_1, a_2, \ \ldots \ a_A\}$. We define

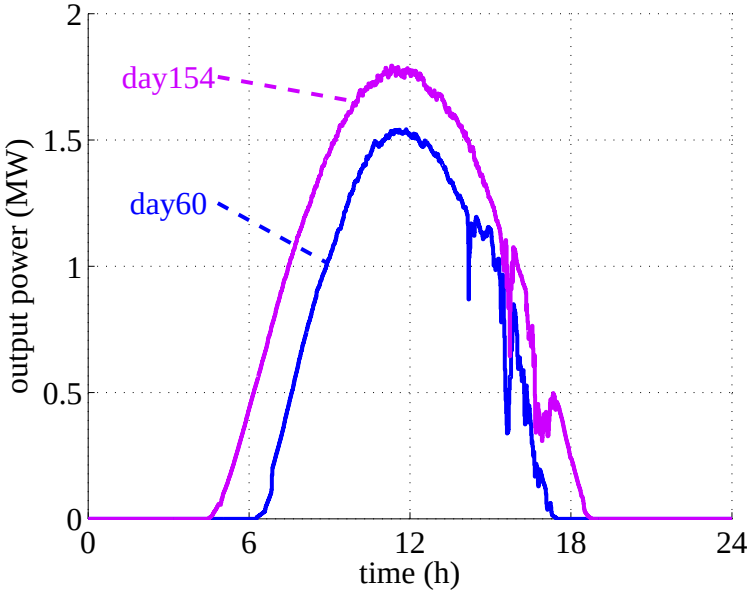
$$\mathcal{U} = \{\underline{u}^{(d,a)} : a \in \mathcal{A}, \ d \in \mathcal{D}\} \ . \tag{2.1}$$

In general, the larger the data set $\mathcal{U}$, the better the ability to represent the power output of the multi-site PV resources. As PV resource output measurement data are scarce, or non-existent in the case of new or future installations, we make use of past atmospheric data and an instantaneous PV resource output model to approximate the output power of PV plants. In Appendix A, we provide the details for the instantaneous model we use in our work. Having assembled the input data set, we proceed with the scaling process described in the next section.

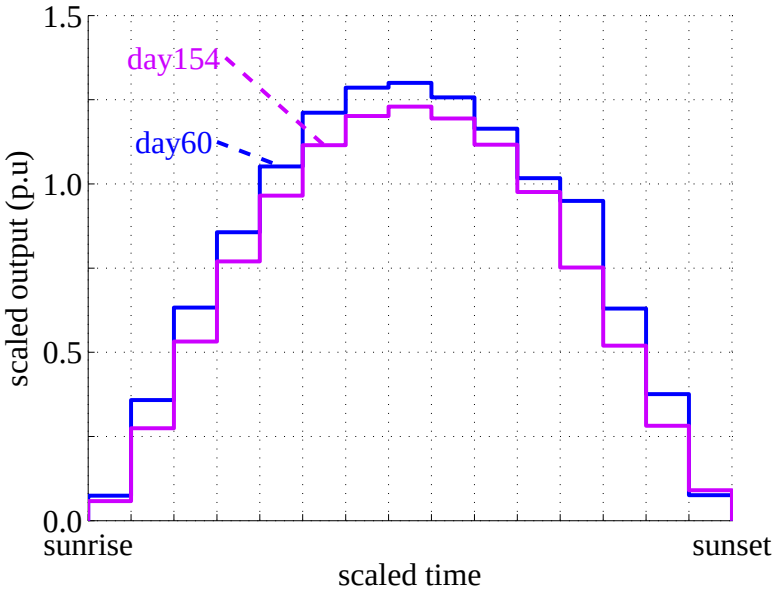
2.2 Scaling of the daily PV output patterns

The careful examination of PV output data for a specified season using multi-year observations leads to the realization that distinct daily output patterns are discernible. However, days with similar output patterns may have significantly different output magnitudes and are associated with distinct sunrise times and sunset times. For example, the PV output patterns of the two days depicted in Fig. 2.1a appear to be similar, notwithstanding their different magnitudes, sunrise times and sunset times. Indeed, the variation of the sunrise and sunset times, and hence the daily PV output durations, can be considerable over a season, as Fig. 1.2a clearly indicates. The maximum daily output over a year, an example of which is presented in Fig. 1.2b, fluctuates significantly from day to day but has a discernible pattern over the year. We develop a scaling algorithm to convert the midnight-to-midnight power output patterns into scaled (per-unit) sunrise-to-sunset output characterizations that allow the meaningful comparison of different days in the season. The application of the scaling algorithm results in the ability to discern the similar patterns in a very effective way. For example, its application

to the power output patterns in Fig. 2.1a results in the similar scaled-output patterns in Fig. 2.1b.



(a)



(b)

Figure 2.1: Daily PV plant (a) power output patterns for two days of the year in the same season, computed using one-minute resolution atmospheric data collected at Las Vegas, NV and published by NREL [5], and (b) their corresponding scaled-output characterizations obtained by application of the scaling algorithm.

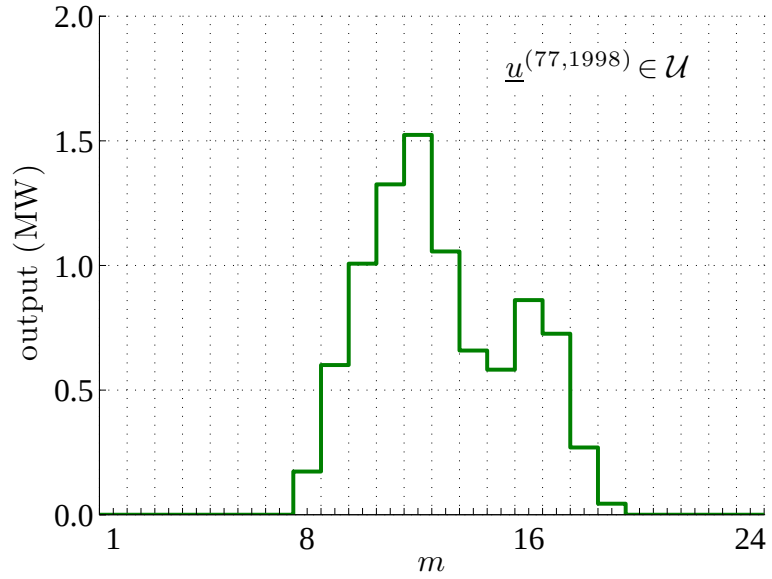
For our scaling algorithm, we use a uniform time grid, which consists of a specified number J scaled sub-periods from sunrise to sunset, to represent the daily PV output for each day. Clearly, as the number of scaled sub-periods is fixed, the scaled sub-periods are shorter on days with shorter PV output duration and longer on days with longer PV output duration. We modify the clearness index approach, widely employed by atmospheric scientists [24], to scale the PV output magnitudes. The clearness index is the value of measured solar radiation divided by the benchmark calculated extraterrestrial radiation [24] and thus can be considered a "per-unit" measure. Typically, we observe distortion in clearness indices at times close to sunrise and sunset. We use computed clear-sky radiation values and an instantaneous power output model to compute a benchmark PV output pattern for the day and site of interest. However, rather than dividing the measured PV output magnitudes by their corresponding benchmark output magnitudes, we divide the measured magnitudes by the maximum value of the benchmark PV output pattern over the day. In this way, we eliminate the distortion problems associated with clearness indices. In Appendix B, we describe the computation of the clear sky-radiation values and the subsequent computation of the benchmark PV output that we use in our work.

We may view the scaling process to be a mapping $\alpha_s(\cdot,\cdot)$ of the vector $\underline{u}_s^{(d,a)}$ representing the power output on day d in year a at site s into the scaled-output characterization vector $\underline{y}_s^{(d,a)} \in \mathbb{R}^J$:

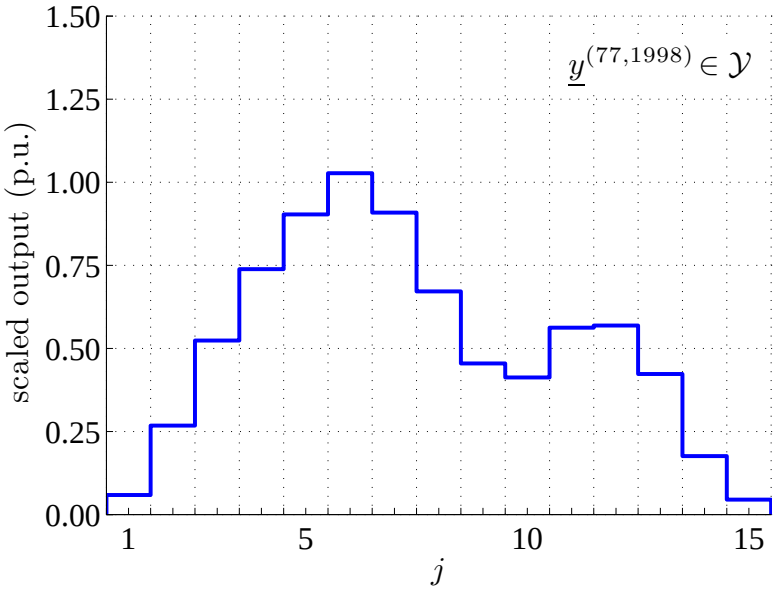
$$\underline{y}_s^{(d,a)} = \alpha_s(\underline{u}_s^{(d,a)}, d) \; . \tag{2.2}$$

We provide an illustrative example, in Fig. 2.2, of the application of this mapping to the single Abilene, TX, site for the day March 18, 1998. The detailed mechanics of the mapping are described in Appendix C.

The scaled-output characterization $\underline{y}_s^{(d,a)} \triangleq [y_{s,1}^{(d,a)}, y_{s,2}^{(d,a)}, \; \ldots \; , y_{s,J}^{(d,a)}]^T \in \mathbb{R}^J$ describes a scaled-output power pattern, where $y_{s,j}$ is the scaled power output at site s for the scaled sub-period j. We construct the super-vector $\underline{y}^{(d,a)} \triangleq [\underline{y}_1^{(d,a)^T}, \underline{y}_2^{(d,a)^T}, \; \ldots \; , \underline{y}_S^{(d,a)^T}]^T \in \mathbb{R}^{(S \cdot J)}$ to characterize the scaled-outputs at the S sites with PV resources on day d of year a. For the S sites we construct the vector scaling function $\underline{\alpha}(\cdot,\cdot) \triangleq [\alpha_1(\cdot,\cdot), \alpha_2(\cdot,\cdot), \; \ldots \; , \alpha_S(\cdot,\cdot)]^T$ and we apply it to $\underline{u}^{(d,a)}$, as depicted in Fig. 2.3, to obtain the super-vector $\underline{y}^{(d,a)}$, i.e.,



(a)



(b)

Figure 2.2: Example of a PV (a) output pattern ($M = 24$), and (b) corresponding scaled-output pattern ($J = 15$), for a PV plant sited at Abilene, TX on March 18, 1998 (day 77 of of the year).

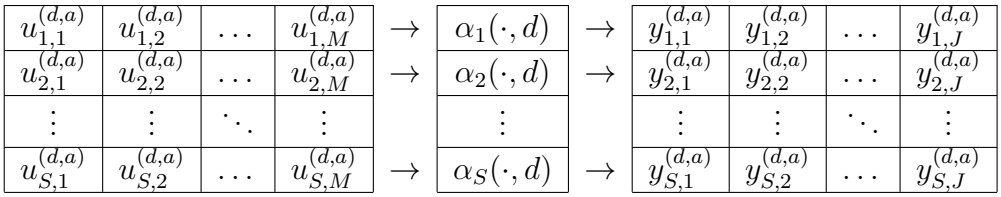


Figure 2.3: The scaling of a super-vector $\underline{u}^{(d,a)}$ into the scaled-output super-vector $\underline{y}^{(d,a)}$ using the scaling functions for the S sites.

$$\underline{y}^{(d,a)} = \underline{\alpha}(\underline{u}^{(d,a)}, d) \ . \tag{2.3}$$

We apply (2.3) to perform the scaling process on the seasonal set $\mathcal{U}$ of daily PV power output patterns for the set of days $\mathcal{D}$ in each year $a \in \mathcal{A}$ to obtain the seasonal set $\mathcal{Y}$ of the daily scaled-output characterizations, where

$$\mathcal{Y} = \{\underline{y}^{(d,a)} : \underline{y}^{(d,a)} = \underline{\alpha}(\underline{u}^{(d,a)}, d),\ \underline{u}^{(d,a)} \in \mathcal{U},\ d \in \mathcal{D},\ a \in \mathcal{A}\} \ . \tag{2.4}$$

We work with the scaled super-vectors in $\mathcal{Y}$ to classify the days with similar patterns into clusters.

2.3 Classification of the scaled daily PV output patterns

We use a clustering scheme to classify the seasonal set of daily scaled-output characterizations into distinct clusters. After extensive experimentation, we selected the K-means algorithm with a Euclidean distance measure [25] as most appropriate for this purpose. We deploy the K-means clustering algorithm on the set $\mathcal{Y}$ to construct the K scaled-output pattern clusters $\mathcal{R}_1, \mathcal{R}_2, \ \ldots \ , \mathcal{R}_K$. We illustrate the clustering process conceptually in Fig. 2.4. Each cluster $\mathcal{R}_k$ is a set of similar daily, scaled-output characterizations and we associate the probability π_k computed as the fraction of daily output patterns classified in the set $\mathcal{R}_k$ of the total elements in the scaled seasonal set $\mathcal{Y}$:

$$\pi_k = \frac{|\mathcal{R}_k|}{|\mathcal{Y}|} \ . \tag{2.5}$$

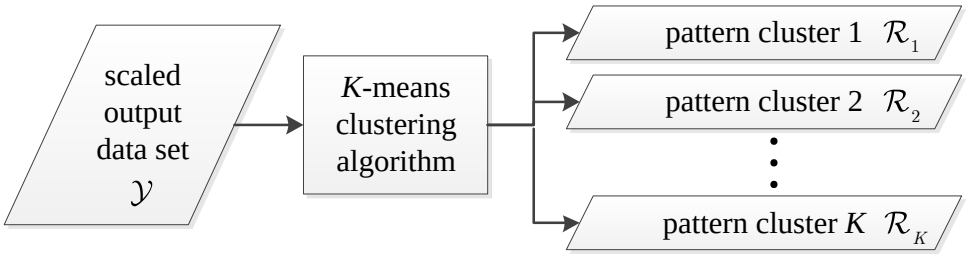


Figure 2.4: The clustering process classifies the elements of $\mathcal{Y}$ into K distinct pattern clusters.

The clustering algorithm disaggregates $\mathcal{Y}$ into the K non-overlapping pattern clusters $\mathcal{R}_k$ with the properties

$$\mathcal{R}_k \cap \mathcal{R}_{k'} = \varnothing,\ k \neq k' \text{ and } \mathcal{Y} = \bigcup_{k=1}^{K} \mathcal{R}_k\ .$$

The pattern clusters $\mathcal{R}_k$ contain the characterizations of the similar daily patterns of the doubly-scaled output vectors at the S sites. To be useful in deploying such outputs when PV resources are integrated into the power system, they need to be judiciously re-scaled. We describe the re-scaling process in the next section.

2.4 Re-scaling of the scaled PV output pattern clusters

In this section, we describe how we re-scale the scaled daily PV output characterizations in the pattern clusters $\mathcal{R}_k$ for use in the extended probabilistic simulation framework. The re-scaled representations must have the appropriate magnitudes and be expressed on the 24 hour midnight-to-midnight time scale. We use an equal-duration re-scaled sub-period representation of the 24-hour period from midnight-to-midnight. We denote the PV output at site s for the H re-scaled sub-periods by the vector $\underline{p}_s \triangleq [p_{s,1}, p_{s,2},\ \ldots\ , p_{s,H}]^T \in \mathbb{R}^H$. We construct the super-vector $\underline{p}$ for the output at the S sites with $\underline{p} \triangleq [\underline{p}_1^T, \underline{p}_2^T,\ \ldots\ , \underline{p}_S^T]^T \in \mathbb{R}^{(S \cdot H)}$.

We may view the re-scaling as the inverse of the scaling function described

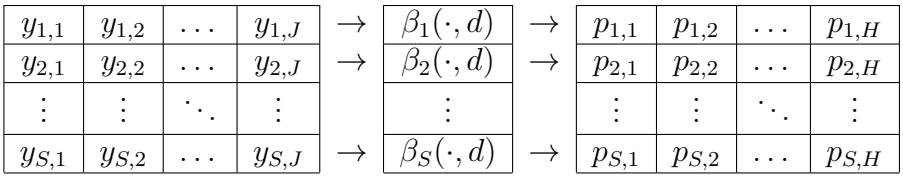


Figure 2.5: The re-scaling of a super-vector $\underline{y}$ into the re-scaled output super-vector $\underline{p}$ for day d using the re-scaling functions for the S sites.

in Section 2.2. The idea is to take an element $\underline{y} \in \mathcal{R}_k$ and transform it into an element $\underline{p} \in \mathbb{R}^{(S \cdot H)}$. Let $\underline{\beta}(\cdot, d)$ denote the mapping of $\underline{y} \in \mathcal{R}_k$ into the vector $\underline{p}$ for use to represent a day $d \in \mathcal{D}$, i.e.,

$$\underline{p} = \underline{\beta}(\underline{y}, d) \; . \tag{2.6}$$

As depicted in Fig. 2.5, the vector function $\underline{\beta}(\cdot, d)$ has S components $\beta_s(\cdot, d)$ with one for each site $s = 1, 2, \; \ldots \; , S$. The day d has an associated output duration for each site in accordance with sunrise/sunset times. We provide an illustrative example, in Fig. 2.6, of the application of the re-scaling function for the day May 15 at the single Abilene, TX, site. The detailed mechanics of the single-site re-scaling function are described in Appendix C.

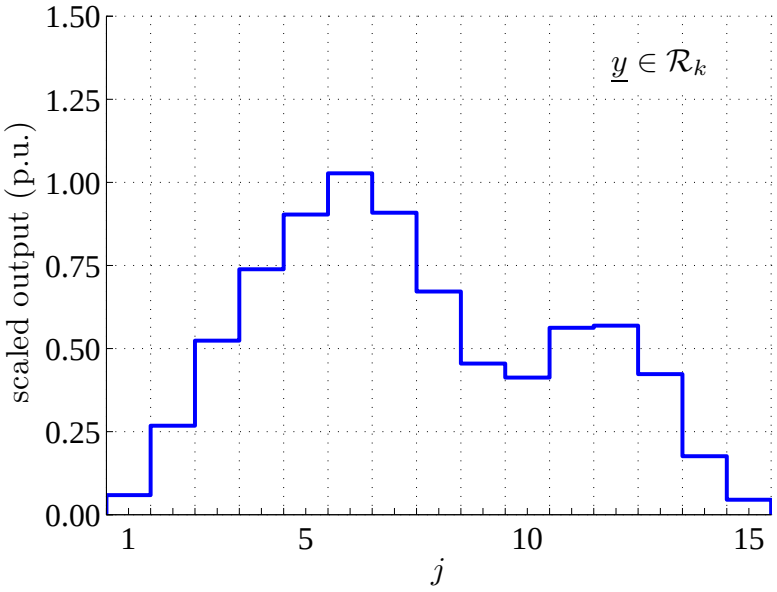
We associate an output class $\mathcal{P}_k^{(d)}$ of re-scaled PV output representation super-vectors with each pattern cluster $\mathcal{R}_k$ for the day d. We assemble

$$\mathcal{P}_k^{(d)} = \{\underline{p} : \underline{p} = \underline{\beta}(\underline{y}, d), \; \forall \underline{y} \in \mathcal{R}_k\} \; . \tag{2.7}$$

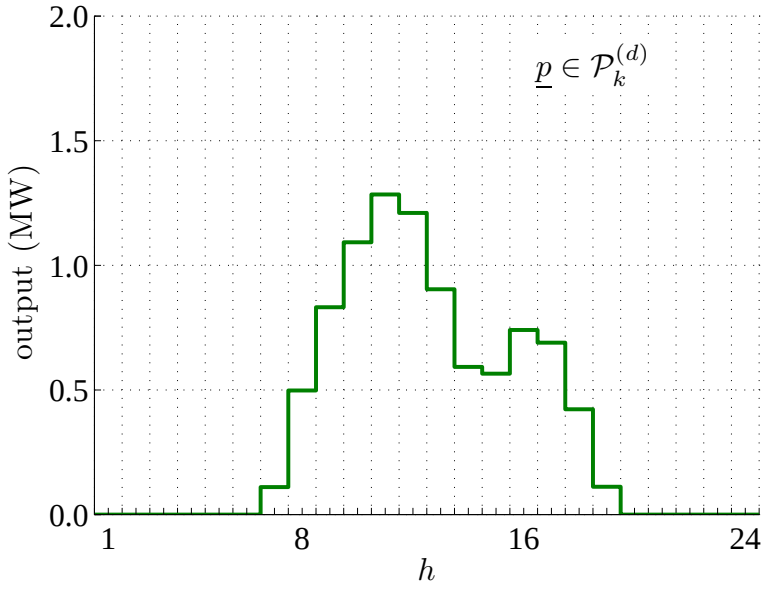
We make use of the set $\mathcal{P}_k^{(d)}$ to approximate the distributions of the PV output power random variable conditioned on the output class k, as we describe in the next section.

2.5 Approximation of the conditional distributions of the PV output random variable

The re-scaled power output super-vectors in the set $\mathcal{P}_k^{(d)}$, which represent the possible realizations of daily output patterns at the S sites on day d, may be considered to be representative samples of the distribution of the multi-dimensional hourly PV output random variables (RVs). From these realizations, we approximate the joint cumulative distribution function (CDF)



(a)



(b)

Figure 2.6: PV (a) scaled-output characterization ($J = 15$), and (b) corresponding re-scaled power output representation ($H = 24$), for a PV plant sited at Abilene, TX for the day d of May 15 (day 135 of the year).

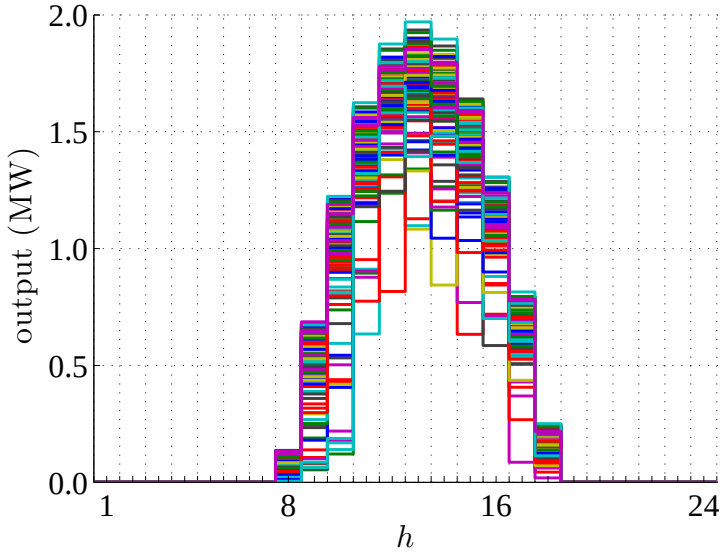
of the multi-dimensional RVs of the power output at the S sites for the H re-scaled sub-periods conditioned on being in the pattern output class k. In Fig. 2.7, we present an example of the set $\mathcal{P}_k^{(d)}$ and its corresponding approximated conditional CDF for the pattern class k, further conditioned on site 1 and re-scaled sub-period 16. We are unable to present the full joint CDF pictorially as there are $S \cdot H$ dimensions.

2.6 Choice of the time-resolution of the scaled-output

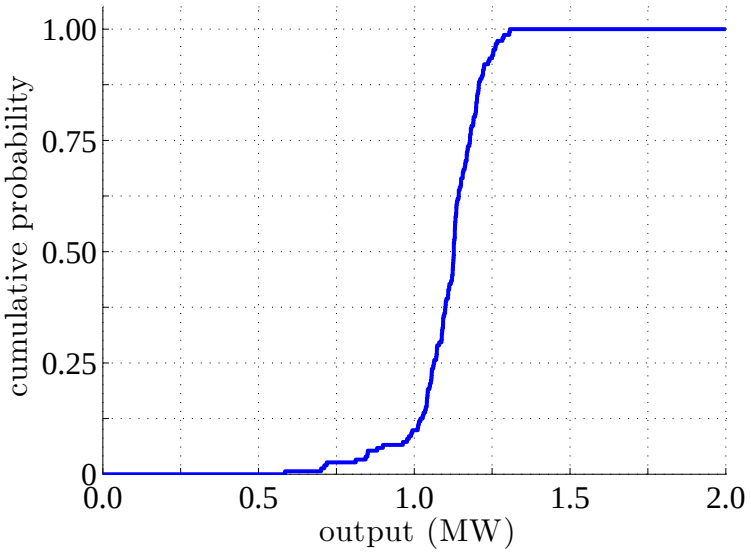
The parameter J, i.e., the number of values we use to characterize a daily scaled-output pattern at each site, plays a significant role in capturing the details of the PV outputs. The parameter dictates the time-resolutions of the scaled-output characterizations, and so we must choose J to be sufficiently high so as to limit the error associated with the scaling and re-scaling processes. We present the sensitivity of the scaling and re-scaling process error to J for two different input/output time-resolutions in Fig 2.8. To perform this sensitivity study, we scale and then re-scale one year of input data, from and to the same time-resolution scales, i.e., $M = H$, and compute the average absolute difference of the re-scaled output representations to the input PV output power patterns as a percentage of the average PV output value. Clearly, for a finer time resolution, i.e., higher values of M and H, a higher value of J is required for the scaled-output characterizations. We observe that as J increases, the scale and re-scale error decreases.

2.7 Summary

In this chapter, we describe the construction of a probabilistic model for the output of PV resources at a single or multiple sites integrated into a power system: the PPVOM. The PPVOM is built on a seasonal input data set of daily PV output patterns. The construction process involves three main steps. In step one, we scale the daily PV output patterns, both in time and magnitude, into scaled-output characterizations so that we may compare the output on different days of the year. In step two, we classify the scaled-output characterizations into K pattern clusters. In step three, we re-



(a)



(b)

Figure 2.7: An example of (a) a set $\mathcal{P}_k^{(d)}$ and (b) the associated CDF of the power output variable conditioned on site 1 and re-scaled sub-period 16.

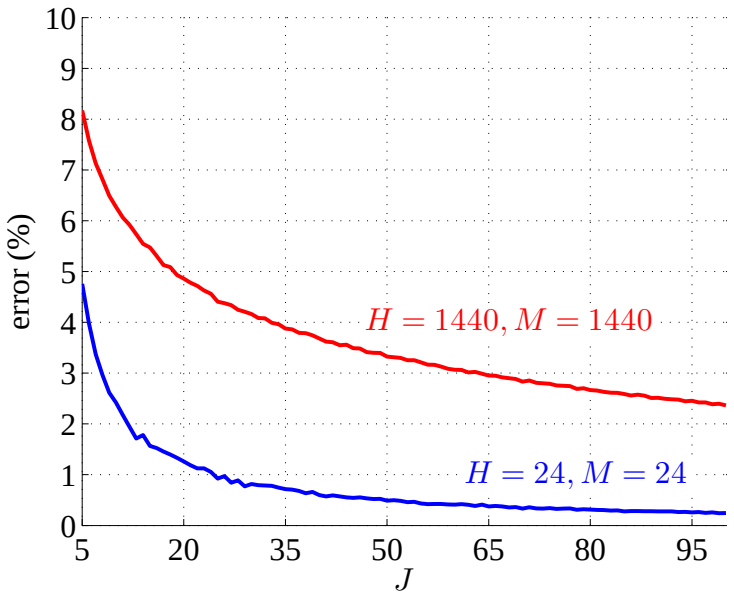


Figure 2.8: Sensitivity of the scale and re-scale error to the number of scaled sub-periods J for H and M equal to 24 (hourly input/output resolution) and H and M equal to 1440 (one-minute input/output resolution).

scale the elements of each pattern cluster into the power output elements of the corresponding output class for each day of interest. We use these pattern classes to approximate the conditional distributions of the PV output random variables that define the daily pattern. In the next chapter, we describe the extension of the probabilistic simulation framework and the incorporation of these class-conditioned CDFs of the PV output RVs into the creation of the extended framework.

CHAPTER 3

THE EXTENDED PROBABILISTIC SIMULATION FRAMEWORK

Probabilistic simulation is widely used to evaluate the power system variable effects – measured in terms of the appropriate economic, environmental and reliability metrics. Simulation tools do not, however, have the capability to effectively represent time-varying resources such as PV plants with their highly volatile and intermittent outputs. In this chapter, we focus on the extension of the probabilistic simulation framework to allow the incorporation of the PPVOM so as to explicitly represent these resources with time varying outputs. The modifications of the conventional probabilistic simulation framework [4] are carried out in a manner similar to that for incorporating a wind model [22]. Thus, we describe the extension of the framework with a focus on these aspects associated with the incorporation of the PPVOM. We first describe the extension of the conventional framework and then describe its application to run case studies of systems with integrated PV resources.

3.1 Extension of the probabilistic simulation framework

The evaluation of the longer-term metrics requires that the seasonal characteristics of all the variables over the entire study period be appropriately captured. We partition the study period into W non-overlapping simulation periods, during which we assume uniform characteristics prevail for all load and controllable resource variables. The simulation periods need not necessarily have the same duration. Each simulation period w has its own set of committed units, resource characteristics and load characteristics and is comprised of Ω_w sub-periods. We define $\mathcal{W}_w \triangleq \{1, 2, \ \ldots \ \Omega_w\}$ to be the sub-period index for simulation w and we perform probabilistic simulation of each of the simulation periods.

The mechanics of the simulation framework involves the convolution of the load RV with the available capacity RVs of the controllable generation units as described in Section 4.1 of [26]. We proceed with this methodology and approximate the CDF of the load RV from samples of historical load. If a sample l_γ represents the load during sub-period γ, we approximate the CDF of the system load RV $\underset{\sim}{L}$ from the set of load samples $\mathcal{L}$, where

$$\mathcal{L} = \{l_\gamma : \gamma \in \mathcal{W}_w\} \ . \tag{3.1}$$

In order to capture the time-dependence of the PV output and its correlation to the chronological load, both the simulation and PV output must have commensurate time resolution. Thus, the sub-periods of the simulation must have the same duration as each re-scaled sub-period h of the re-scaled PV output representations we describe in Section 2.4. The time-resolution thus specifies the granularity of the simulation: we cannot capture any phenomena with durations shorter than the resolution. We define for each simulation period w the subsets $\mathcal{W}_w^{(h)}$ for each sub-period $h = 1, 2, \ \ldots \ , H$, which are the subsets of the indices in $\mathcal{W}_w$ that correspond to the sub-period h of the midnight-to-midnight day. We also define H subsets $\mathcal{L}^{(h)}$ of the set $\mathcal{L}$, where

$$\mathcal{L} = \{l_\gamma : \gamma \in \mathcal{W}_w^{(h)}\} \ . \tag{3.2}$$

The following properties hold for the sets $\mathcal{W}_w^{(h)}$ and the set $\mathcal{W}_w$:

$$\begin{aligned}
&\mathcal{W}_w^{(h)} \subset \mathcal{W}_w, \ h = 1, 2, \ \ldots \ , H \ , \\
&\mathcal{W}_w = \bigcup_{h=1}^{H} \mathcal{W}_w^{(h)} \text{ and} \\
&\mathcal{W}_w^{(h)} \cap \mathcal{W}_w^{(h')} = \varnothing, \ h \neq h' \ .
\end{aligned}$$

We approximate the CDF $\mathrm{F}_{\underset{\sim}{L}|h}$ of the RV $\underset{\sim}{L}$ conditioned on the sub-period h from the load samples in the subset $\mathcal{L}^{(h)}$ and using conditional probability we restate the CDF $\mathrm{F}_{\underset{\sim}{L}}$:

$$
\begin{aligned}
\mathrm{F}_{\underset{\sim}{L}}(l) &= \Pr\{\underset{\sim}{L} \leq l\} \\
&= \Pr\{\underset{\sim}{L} \leq l \text{ in every sub-period } h\} \\
&= \sum_{h=1}^{H} \Pr\{\underset{\sim}{L}_w \leq l \mid \text{sub-period } h\} \Pr\{\text{sub-period } h\} \\
&= \frac{1}{H} \sum_{h=1}^{H} \mathrm{F}_{\underset{\sim}{L}|h}(l) \; . \qquad (3.3)
\end{aligned}
$$

We approximate the joint CDF $\mathrm{F}_{\underset{\sim}{P}|k,h}$ of the system PV output RV $\underset{\sim}{P}$ for each simulation conditioned on pattern class k and sub-period h from the elements of the super-vectors in sets $\mathcal{P}_k^{(d)}$ corresponding to that sub-period. We choose d to be the middle day of the simulation period (in the case of two middle days, we arbitrarily choose one).

In the studies we present in this thesis, we assume the power system has adequate and lossless transmission and thus perform single-node simulation. Thus, in approximating the joint CDF $\mathrm{F}_{\underset{\sim}{P}|k,h}$, for sub-period h, we sum the PV outputs $p_{s,h}$ at all the sites $s = 1, 2, \ \ldots \ , S$ sites to compute the total PV output:

$$
p_h = \sum_{s=1}^{S} p_{s,h} \; . \qquad (3.4)
$$

However, the reader should note that multi-node analysis, with associated multi-node load information and without applying (3.4), may be performed by modeling the PV output and load in the same manner as we describe above, conditioning on both the sub-periods h and sites s.

We capture the effects of the time-dependent PV resources, by computing the load that the controllable resources must supply, i.e., the load minus the PV output, with explicit recognition of the time-dependent nature of the load and the PV output. In doing so, we assume that all PV resource production will be used to serve the load, and thus we are able to incorporate the PPVOM prior to the convolution with the available capacity RVs of the controllable generation units. Through the convolution of $\underset{\sim}{L}$ conditioned on sub-period h and the negative of $\underset{\sim}{P}$ conditioned on pattern class k and sub-period h, we are able to approximate the CDF $\mathrm{F}_{\underset{\sim}{C}|k,h}$ of the RV $\underset{\sim}{C}$, the RV of

the load that is demanded from the controllable generation units, conditioned on pattern class k and the sub-period h. The expression of the "net-load" RV $\underset{\sim}{C}$ for pattern class k and sub-period h results in an expression analogous to (3.3):

$$\mathrm{F}_{\underset{\sim}{C}|k}(c) = \frac{1}{H} \sum_{h=1}^{H} \mathrm{F}_{\underset{\sim}{C}|k,h}(c) \ . \tag{3.5}$$

For each $\mathrm{F}_{\underset{\sim}{C}|k}$, where $k = 1, 2, \ \ldots \ , K$ we proceed with the conventional simulation framework methods to find expected system metrics for the simulation period. We obtain seasonal metrics by summing the pattern-class probability-weighted expected values computed in the simulation.

3.2 Application of the extended probabilistic simulation framework

In general, the variation of load patterns from week to week is slight and, in addition, there are periods of multiple weeks over which generation units are not committed due to planned maintenance and/or their economics. Consequently, we simulate representative weeks, each of which represents the loads and the committed units for multiple weeks of the year. We weigh the metrics evaluated in the simulation by the number of weeks represented.

As we discuss in Chapter 2, the magnitude and duration of daily PV output depends significantly on the day of the year and, accordingly, the re-scaling process we describe in Section 2.4 re-scales according to a day d of the year. For each representative week we re-scale the class sets of scaled daily PV output pattern characterizations according to the middle day of the period being represented (not necessarily the middle day of the representative week). Clearly, a representative week must represent a continuous period, i.e., the representative week may not represent two or more discontinuous periods of time. The period must be limited in order to limit the error associated with using the middle day instead of simulating each day individually. We limit the representation sub-period to three weeks in the Spring and the Fall seasons, during which the daily PV output duration changes faster, and five

weeks in the Winter and the Summer seasons.

3.3 Summary

In this chapter, we discuss the extension of the probabilistic simulation framework so as to extend its capability to explicitly represent PV resources. We also describe the application of the framework using the concept of representative weeks and discuss the matter of selecting the appropriate day for the re-scaling process. In doing so, we have completed the presentation of the construction of our tool.

CHAPTER 4

SIMULATION STUDIES OF SYSTEMS WITH INTEGRATED PHOTOVOLTAIC RESOURCES

In this chapter, we present a set of representative simulation results to demonstrate the capability of the extended probabilistic simulation methodology that incorporates the PPVOM we developed. We use a modified version of the IEEE Reliability Test System (RTS) [27] to carry out the studies presented. We describe the test system characteristics, the simulation set-up and provide simulation results and analysis.

4.1 Details of the system test cases

We use a modified version of the RTS to perform our simulations and the modifications we introduce result in a system with a scaled version of the controllable resource mix, by generation type, of the ERCOT system [28]. We maintain the RTS-specified forced outage rates, heat rates and maintenance requirements and use the most recent EIA fuel cost and emission rate data [29], as summarized in Tables 4.1 and 4.2. We scale the 2007 ERCOT hourly load data [6] to result in an annual peak load of 2,850 MW, the peak level specified in the RTS [27]. We provide general characteristics of the modified test case in Table 4.3.

We construct the seasonal PPVOMs given by (2.1) to (2.7) using atmospheric measurement data for the five years from 1998 to 2002 collected at five different sites (Abilene, Amarillo, Austin, Corpus Christi and Houston) in Texas. The hourly-resolution ($M = 24$) atmospheric data consists of the NREL solar radiation data [7] and the NOAA temperature and wind speed data [8]. We partition the data according into the Winter, Spring, Summer and Fall seasons defined in Table 4.4. In Table 4.5, we list the geographic and installation configuration parameters used in the instantaneous output modeling of the integrated PV plants. The PV installations are Kyocera

Table 4.1: Characteristics of the controllable generator resource mix of the modified RTS.

no. of units	unit size (MW)	minimum capacity (MW)	forced outage rate	scheduled maintenance (weeks/year)	heat rate (Btu/kWh)	fuel type
4	20	10	0.02	2	16,000	gas
4	20	10	0.10	2	12,000	gas
6	80	20	0.02	3	17,107	coal
6	100	30	0.04	3	11,000	gas
4	160	60	0.04	4	12,000	coal
3	200	70	0.05	4	11,000	gas
2	350	140	0.08	5	10,200	coal
1	400	200	0.12	6	12,751	nuclear

Table 4.2: Fuel data.

fuel type	cost ($/MMBtu)	CO_2 emission rate (lb/MMBtu)
gas	4.43	117
coal	1.88	208
nuclear	0.82	0

Table 4.3: General characteristics of the modified RTS.

total installed capacity	3,580 MW
annual peak load	2,850 MW
total annual energy demand	14.632×10^6 MWh

.

Table 4.4: Day indices for the season partitions of the PV output data.

season	indices of days
Winter	$1, 2, \ldots, 60, 335, 336, \ldots, 366$
Spring	$61, 62, \ldots, 151$
Summer	$152, 153, \ldots, 243$
Fall	$244, 245, \ldots, 334$

Table 4.5: Geographic and installation configuration parameters of the PV plants.

parameter	PV plant site				
	Abilene	Amarillo	Austin	C.Christi	Houston
longitude (°)	−99.71	−101.90	−97.74	−97.63	−95.09
latitude (°)	32.47	34.99	30.29	27.88	29.56
altitude (m)	530	1068	213	6	6
time zone (h)	−6	−6	−6	−6	−6
slope angle (°)	32.47	34.99	30.29	27.88	29.56
azimuth angle (°)	0	0	0	0	0

KC200GT PV Solar Cell arrays with Fronius IG2000 inverters, where parameters are given in [30] and [31].

We report on a variety of test cases that serve to illustrate the capability of the extended probabilistic simulation methodology for evaluating power systems with integrated PV resources. In Case A (the base case), we simulate the system with no PV resources. In Case B, we simulate a single PV resource integrated into the system: a 500 MW nameplate capacity plant at Corpus Christi. In Case C, we simulate the modified RTS with multiple integrated PV resources: 100 MW nameplate capacity plants sited at Abilene, Amarillo, Austin, Corpus Christi and Houston. In Case D, we simulate the modified RTS with the largest coal plant (350 MW) de-committed and 345 MW nameplate capacity PV plants added at each of the five sites listed in Case C. We summarize the list of test cases in Table 4.6

Table 4.6: Key characteristics of the system test cases.

system case	key characteristics
A (base case)	existing resource mix without PV resources
B	existing resource mix with a 500 MW PV plant at Corpus Christi
C	existing resource mix with a 100 MW PV plants at Abilene, Amarillo, Austin, Corpus Christi and Houston
D	existing resource mix with a 350 MW coal plant de-committed and 345 MW PV plants at Abilene, Amarillo, Austin, Corpus Christi and Houston

4.2 Setup of the simulation studies

We run simulation studies for each test system case using a study period of one year, so as to capture the various seasonal characteristics of the study. We partition the year into the same four seasons that we use for the PV resource output data and perform a number of representative week simulations for each of the seasons. In Table 4.7, we summarize the sixteen representative week simulations we perform.[1] In line with the load data, and hence the PV output representation data, we use a one-hour resolution for all simulations.

We assume adequate and lossless system transmission and thus perform a single-node analysis as described in Section 3.1. We model each conventional generator with a two-state, full or zero capacity model and, using a loading list approach, we specify a unit commitment schedule that satisfies the maintenance requirements, maximum/minimum conventional generator available capacity levels and a reserves margin equal to 15% of the peak load for the representative period plus 10% of the total nameplate capacity of the system-integrated PV resources. In Fig. 4.1, we provide a plot of the peak/minimum loads, maximum/minimum committed availability and reserve margins used for each of the representative week simulations for Case B.

For each season, we construct four ($K = 4$) PV output pattern clusters $\mathcal{R}_k$ with 128 ($J = 128$) sub-periods used to characterize the daily output pattern at a single site. We then construct PV output patterns with an hourly resolution ($H = 24$), in line with the resolution of the load data, for

[1] We represent the effect of the extra day of a leap year ($d = 366$) by adding a quarter-day (1/28 week) to the period represented in simulation number two.

Table 4.7: Representative week simulation periods.

season	sim. number	week simulated	period (weeks)	represented day indices	middle day
Winter	1	49	4.14	$335, 336, \ldots, 363$	349
	2	2	4.89	$364, 365, 366,$ $1, 2, \ldots, 32$	15
	3	7	4	$33, 37, \ldots, 60$	46
Spring	4	10	3	$61, 62, \ldots, 81$	71
	5	13	2	$82, 83, \ldots, 95$	88
	6	16	2	$96, 97, \ldots, 109$	102
	7	17	3	$110, 111, \ldots, 130$	120
	8	20	3	$131, 132, \ldots, 151$	141
Summer	9	24	4	$152, 153, \ldots, 179$	165
	10	28	5	$180, 181, \ldots, 214$	197
	11	33	4.14	$215, 216, \ldots, 243$	229
Fall	12	36	3	$244, 245, \ldots, 264$	254
	13	40	2	$265, 266, \ldots, 278$	271
	14	41	2	$279, 280, \ldots, 292$	285
	15	45	3	$293, 294, \ldots, 313$	303
	16	47	3	$314, 315, \ldots, 334$	324

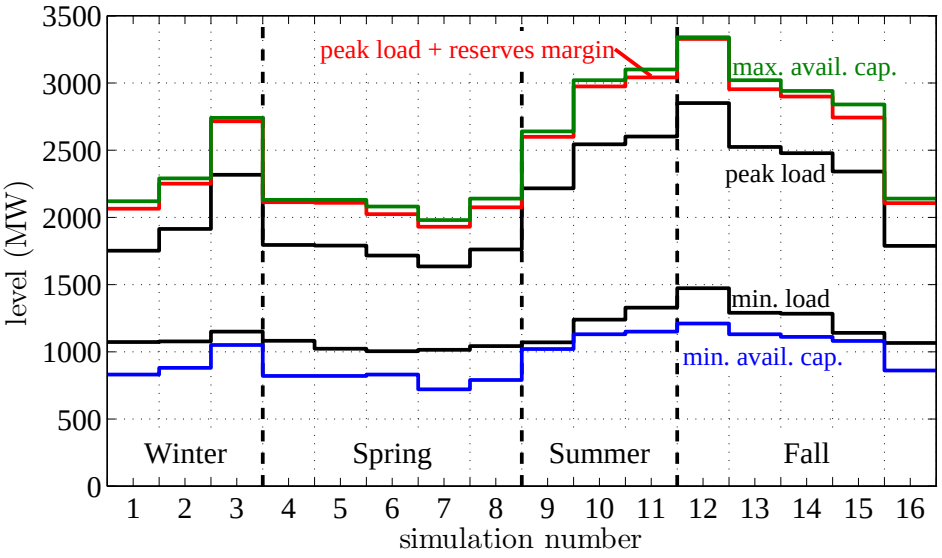


Figure 4.1: The peak/minimum loads, maximum/minimum committed available capacities and peak load plus reserve margin for the representative week simulations of Case B.

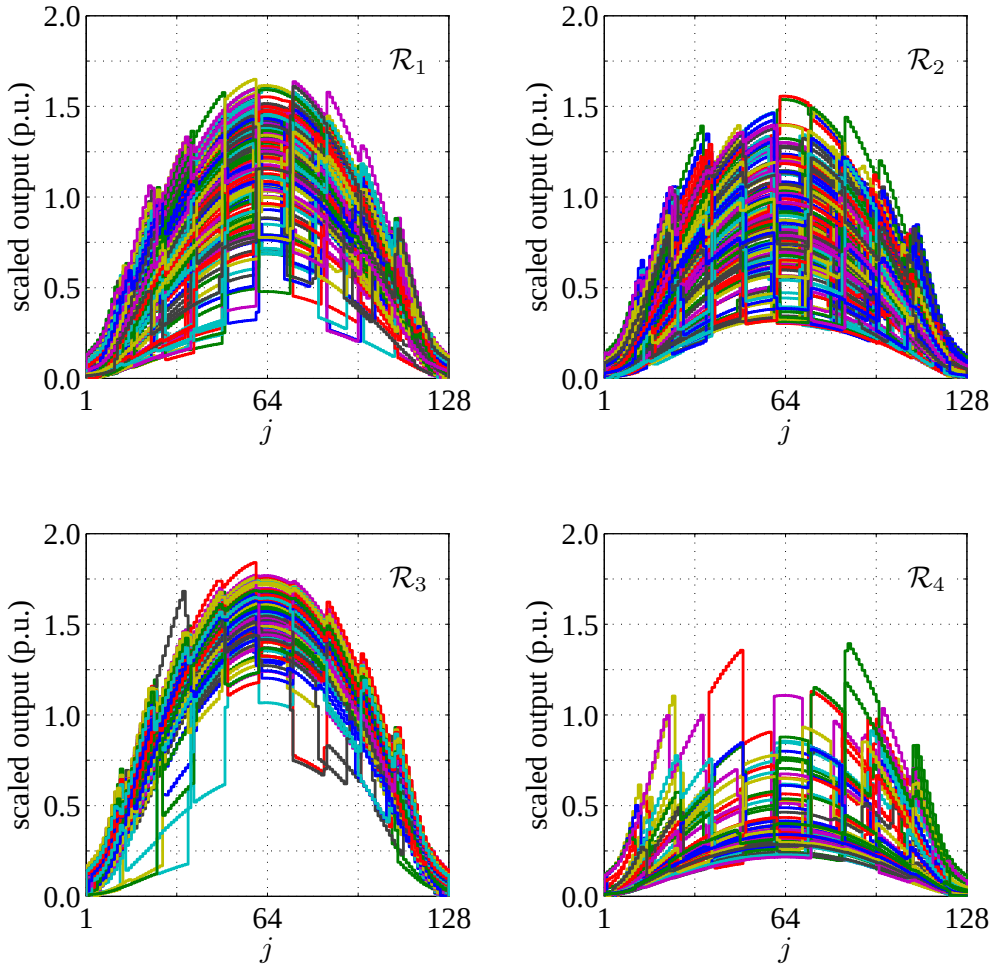


Figure 4.2: Case B, Summer season, pattern classes $\mathcal{R}_k$

the middle day d of the represented period specified for each representative week simulation. In Fig. 4.2, we provide plots of the super-vectors of the four output pattern classes $\mathcal{R}_k$ computed for the Summer season simulation of Case B. We order the clusters from largest to smallest and using (2.5), for each scaled-output cluster of each season, we compute the probability π_k, as listed in Table 4.8 for Case B.

We proceed with the representative week probabilistic simulations for each test system as described in Chapter 3. We compute the various economic, environmental and reliability metrics – the expected production costs, the expected CO_2 emissions, the loss of load probability (LOLP) and the expected unserved energy (EUE) – as described in Chapter 4 of [26].

Table 4.8: PV pattern class probabilities for Case B.

pattern class probability	season			
	Winter	Spring	Summer	Fall
π_1	0.376	0.297	0.465	0.308
π_2	0.332	0.259	0.278	0.279
π_3	0.152	0.242	0.207	0.240
π_4	0.139	0.202	0.050	0.174

Table 4.9: Case B expected production cost simulation results for the Summer season.

expected weekly production costs (10^6 \$)		representative week simulation no.		
pattern class k	prob. π_k	9	10	11
1	0.465	15.27	16.33	17.04
2	0.278	15.38	16.45	17.15
3	0.207	15.34	16.40	17.11
4	0.050	15.53	16.60	17.30
probability-weighted totals		15.33	16.39	17.10

4.3 Simulation results and analysis

We begin our presentation of the test case system results with an overview of the expected production cost calculation for Case B. In Table 4.9, we summarize the probability weighting of the expected production costs for Case B in the Summer season. In a similar fashion, we then compute the seasonal averages shown in Table 4.10. Finally, through seasonal probability weighting we determine an annual metric.

In Table 4.11, we summarize the various annual metrics we compute from the simulation of the four system cases (Case A to Case D) and provide the reduction in metrics relative to the simulation of the base case (Case A). In Table 4.12, we provide a breakdown, for each system case, of the expected energy computed, measured in the percentage of the total expected demanded energy (14.632×10^6 MWh).

Both Cases B and C have a 500 MW, 17.54% of peak load, integrated PV resource capacity; and the simulated operation of the systems results in

Table 4.10: Case B expected production cost simulation result summary.

expected weekly production costs (10^6 \$)	season (season-probability)			
simulation no.	Winter	Spring	Summer	Fall
1/4/9/12	12.33	12.23	15.33	15.85
2/5/10/13	14.33	12.48	16.39	15.75
3/6/11/14	12.62	11.85	17.10	14.55
-/7/-/15	-	10.50	-	12.62
-/8/-/16	-	11.85	-	13.07
probability-weighted totals	13.17	11.72	16.29	14.25

Table 4.11: Annual metrics for Case A to Case D.

metric	system case			
	Case A	Case B	Case C	Case D
LOLP	0.0337	0.0164	0.0163	0.0347
LOLP reduction (%)	-	51.3	51.6	−2.9
EUE (10^3 MWh)	48.8	21.4	21.3	47.4
EUE reduction (%)	-	56.2	56.4	2.9
expected production costs (10^6 \$)	745.48	723.35	723.57	653.91
prod. cost reduction (%)	-	3.0	2.9	12.3
expected CO_2 emissions (10^6 ton)	12.21	11.83	11.84	11.37
CO_2 emission reduction (%)	-	3.1	3.1	6.9

Table 4.12: Annual expected energy breakdown for Case A to Case D as a percentage of the 14.632×10^6 MWh of expected demanded energy.

expected energy (%)	system case			
	Case A	Case B	Case C	Case D
unserved	0.33	0.15	0.15	0.31
supplied by controllable generation	99.67	95.74	95.77	84.67
supplied by PV resources	-	4.12	4.08	15.02

energy penetrations (percentage of total energy production) of 4.12% and 4.08%, respectively. These two systems have reliability significantly superior to that of the base case, and we attribute this to the correlation of the daily PV resource output to the daily system load. Because the probabilistic simulation framework simulates the controllable generators with constant output capabilities and constant outage rates over the entire simulation period, for a system with only conventional generation resources, outages are most likely to be computed for the higher load hours of the day. Therefore, when we integrate PV resources that provide additional capacity at these higher load hours, the improvement in reliability is greater than the percentage energy provided by the PV resources. In a separate study of wind resources, multiple widely dispersed wind resources resulted in smoother total wind resource output patterns and reliability improvements superior to that of a single large resource [22]. However, for Cases B and C, we observe little difference in their reliability metrics, despite the dispersement of the PV resources in Case C. From this observation, we conclude that the PV resource outputs at the various sites in Case C are highly correlated. This issue is of significant interest to system planners who might hope to improve the system reliability through disperesed PV resources.

We note that for each of the three cases with integrated PV resources, the reduction in both expected production costs and expected emissions are lower than the energy penetration level. Although the simulated operation does result in the displacement of the most expensive of the additional generation blocks, the most expensive and polluting, minimum-capacity blocks of the conventional generation units are not displaced by the PV resources, and the higher reserves levels require a greater number of units to be committed.

In our simulation of Case D, the energy provided by one of the two largest coal units in the base case is displaced by the energy provided by 1,725 MW of PV resources at the five sites.[2] The de-commitment of the 350 MW coal unit reduces the average available capacity and, hence, the system reliability over the study period. Conversely, the substantial PV resources help to raise the reliability level by providing additional capacity during the higher diurnal load hours. The simulation results indicate a slight net increase in LOLP and a slight net decrease in EUE with respect to the base case.

[2]We take the Case A commitment schedule and de-commit the first (350 MW) coal unit on the loading list.

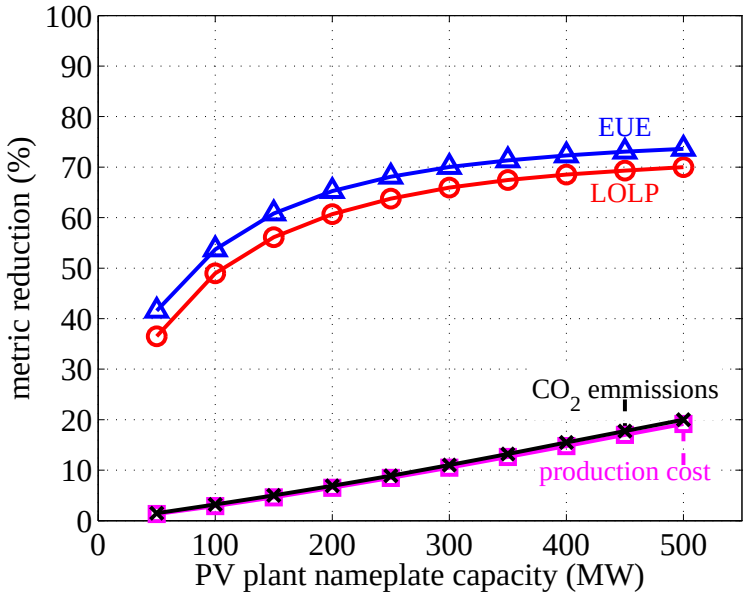


Figure 4.3: Sensitivity study of the system metrics to the nameplate capacity of PV plants installed at Abilene, Amarillo, Austin, Corpus Christi and Houston.

We study the sensitivity of various metrics to the installed PV capacity by varying the nameplate capacity of the PV resources in Case C from 50 MW to 500 MW in 50 MW increments and simulating the various modified cases. In Fig. 4.3, we provide a plot of the reduction in simulation metrics for the various modified cases. At the nameplate capacities studied, we see a trend of diminishing returns for the reliability metrics and almost linear production-cost and CO_2 emission reduction. We see a similar trend to that of the reliability metrics in Fig. 4.4, which shows the change in the effective load carrying capability (ELCC). We define the ELCC as the load increment that a system can support after a change in its resource mix while maintaining the same reliability level [32]. These observations help to validate our methodology because we expect diminishing returns with increased uncontrollable and intermittent resources.

The representative simulation studies we have presented showcase only a small subset of the key observations and metrics in which system planners, investors and regulators may be interested. There is also room for the extension of the tool to model other solar resources as we discuss in the next chapter.

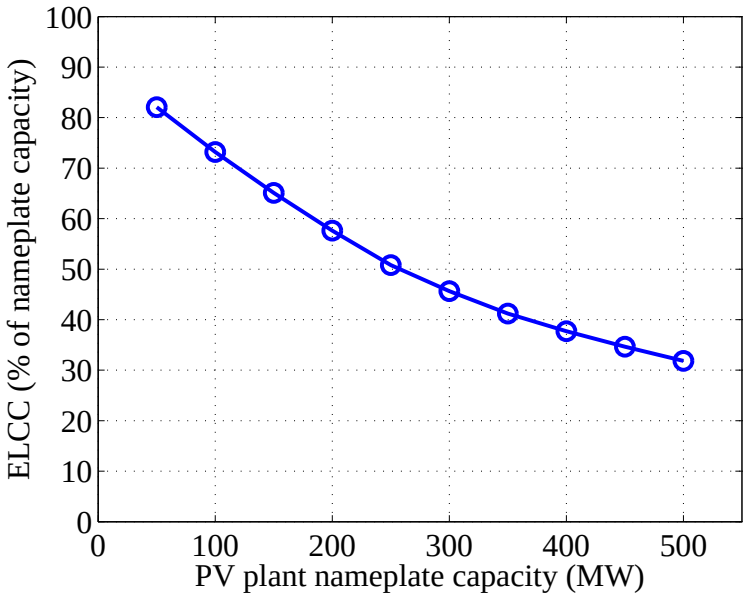


Figure 4.4: Sensitivity of the ELCC to the nameplate capacity of PV plants installed at Abilene, Amarillo, Austin, Corpus Christi and Houston.

CHAPTER 5

CONCLUSION

In this thesis, we present the development and testing of a planning tool to quantify the variable effects of integrated, utility-scale PV resources on power systems over the longer-term. Based on historical data and using a scaled-output pattern-class approach, we construct the PPVOM in a manner that captures the time-varying, intermittent and uncontrollable nature of the resources in addition to capturing seasonal variation. We extend the widely used probabilistic simulation framework so as to incorporate the output of the PPVOM and through the application of conditional probability concepts, we are able to maintain the time-dependence between the load and PV resource random variables and thus obtain realistic results. The extended probabilistic simulation framework gives us the capability to determine reliability indices and expected economic and environmental costs. This capability, which we demonstrate by way of a number of example case studies, is very useful in many areas of power system planning, including investment decision-making and policy analysis.

There is great potential to further the research we have undertaken. Because utility-scale PV resources are often located far from load centers, the incorporation of transmission constraints, using multi-node analysis, is of great interest. Another challenge is the modeling of concentrated solar power (CSP) resources. CSP modeling includes the generally complex modeling of the thermal storage that gives the operators a degree of control and thus results in variable power output durations. Looking further ahead, as distributed PV resource data becomes available, an opportunity arises in the development of distributed PV resource models.

APPENDIX A

THE TIME-VARYING, INSTANTANEOUS PHOTOVOLTAIC OUTPUT MODEL

We make use of a time-varying, instantaneous PV output model (TVIPVM) to approximate the PV resource output at a specified site. The model is comprised of several components that we have combined for use as the TVIPVM in our research. In this appendix we provide the detailed construction of the model. We also provide details for the computation of the clear-sky radiation values, which are integral to the computation of the benchmark output, which we describe in Appendix B. The reader should note that the notation we use in this appendix does not relate to the notation used in the rest of the thesis.

The inputs to the TVIPVM are the set of atmospheric measurements, as listed in Table A.1, measured at time t on day d of the year. In addition, we also input the parameters associated with the site location, resource technology and site installation configuration of the PV resource. The output P_{ac} of the TVIPVM is an estimate of the instantaneous power output to the grid.

A.1 Clear-sky radiation

This section describes the modeling of clear-sky beam solar radiation that is used to create benchmark PV electricity production data. The reader should note that the data produced can also be used in the modeling of solar thermal

Table A.1: Atmospheric measurement input to the TVIPVM.

symbol	atmospheric measurement
G_h	global-horizontal solar radiation
G_b	beam direct solar radiation
T_a	ambient air temperature
v_w	wind speed

production.

It is assumed that the sun emits radiation at a constant rate. The solar constant G_{sc} is the energy from the sun per unit time received on a unit area of a surface perpendicular to the direction of propagation of the radiation at mean earth-sun distance outside the atmosphere and is approximated at 1367 W/m^2 [24]. As the earth-sun distance varies, we model the extraterrestrial radiation incident on a plane normal to the radiation G_{on}, on day d of the year by

$$G_{on} = G_{sc}(1 + 0.033 \cos \frac{360d}{365}) \ . \tag{A.1}$$

The zenith angle θ_z is the angle between vertical and the line to the sun. The air mass ratio m is the ratio of the mass of the atmosphere through which it passes to the mass it would pass through if the zenith angle was zero. We approximate the relationship between these two variables by

$$m = \frac{1}{\cos \theta_z} \ . \tag{A.2}$$

Direct beam radiation is the solar radiation received from the direction of the sun. For an atmospheric transmittance τ_b, we approximate the clear-sky beam radiation on a surface normal to the beam radiation G_b by

$$G_b = G_{on}\tau_b \ , \tag{A.3}$$

and the horizontal surface clear-sky beam radiation G_{bh} by

$$G_{bh} = G_{on}\frac{\tau_b}{m} \ . \tag{A.4}$$

We compute the atmospheric transmittance τ_b for beam radiation by

$$\tau_b = a_0 + a_1 \exp(-a_2 m) \ , \tag{A.5}$$

where the constants a_0, a_1 and a_2 are computed by

$$a_0 = r_0(0.4327 - 0.00821(6 - A)^2) \ , \tag{A.6}$$

$$a_1 = r_1(0.5055 + 0.00595(6.5 - A)^2) \text{ and} \tag{A.7}$$

$$a_2 = r_2(0.2711 + 0.01858(2.5 - A)^2) \ , \tag{A.8}$$

Table A.2: Atmospheric transmittance calculation constants for different climate types.

climate type	r_0	r_1	r_2
default	1.00	1.00	1.00
tropical	0.95	0.98	1.02
mid-latitude summer	0.97	0.99	1.02
sub-arctic summer	0.99	0.99	1.01
mid-latitude winter	1.03	1.01	1.00

where A is the altitude of the collector in meters and the constants r_0, r_1 and r_2 depend on the climate type, as listed in Table A.2.

Diffuse radiation is solar radiation that does not come directly from the sun. This radiation is mostly that which is scattered by air molecules, dust and water vapor.

A.2 Solar collector

Global collector radiation is the sum of the beam, diffuse and reflected radiation on a collector. The global radiation on a solar collector is the most significant input to a PV cell model. In this section, we present the model we use to determine the global radiation on a flat panel solar collector and describe the relationship between the model and time.

A.2.1 Radiation direction

To determine the direction of beam radiation on a surface, we must know the various angles and their relationships to one another. Table A.3 lists and defines the pertinent angles, as depicted in Figure A.1, for a flat plate solar collector.

We approximate the declination angle δ, by

$$\delta = 23.45 \sin\left(360 \frac{284 + d}{365}\right) . \tag{A.9}$$

The relationship between various angles are well established [33]. We determine the angle of incidence θ by

Table A.3: Pertinent angles.

symbol	name	definition
ϕ	latitude	angular geographic location relative to the equator; north positive
δ	declination	angular position of sun relative to the equator at solar noon; north positive
β	slope (tilt)	angle between the collector plane and the horizontal
γ	surface angle	angle between the local meridian and the normal to the collector surface as measured on a horizontal plane; west positive
ω	hour angle	angular displacement of the sun relative to the local meridian due to the rotation of the earth on its axis (15° per hour); afternoon positive
θ	incidence angle	angle between beam radiation and normal to the collector surface
θ_z	zenith angle	angle between vertical and the line to the sun
α_s	solar altitude angle	angle between the horizontal and the line of the sun
γ_s	solar azimuth angle	angle between projection of sun on the horizontal plane and south; west positive

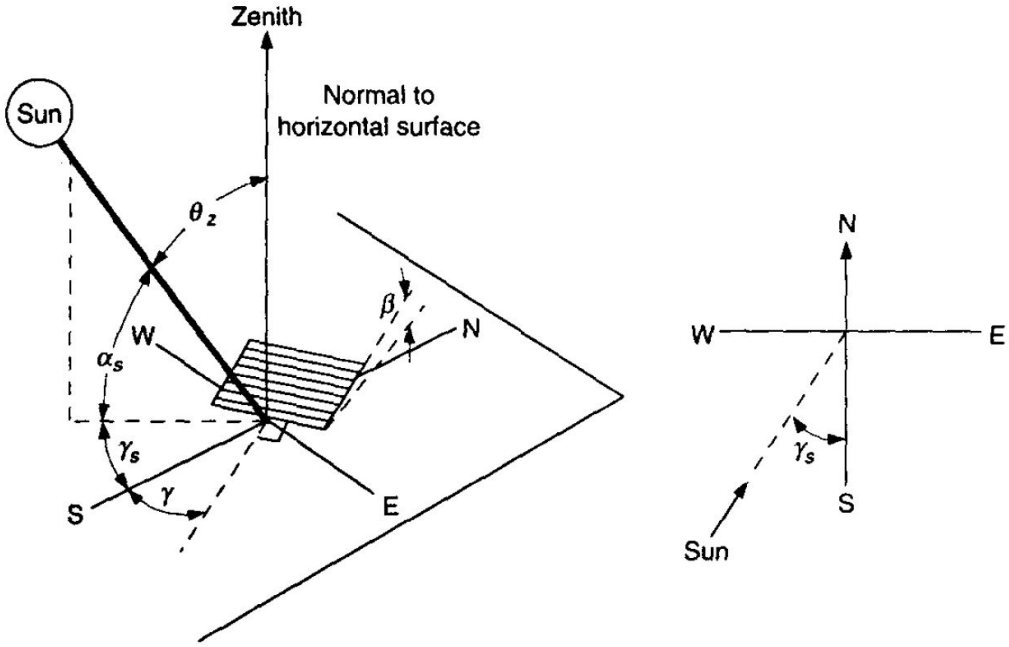


Figure A.1: Pertinent angles for a flat plate solar collector [24].

$$\begin{aligned}\cos\theta = &\sin\delta\sin\phi\cos\beta - \sin\delta\cos\phi\sin\beta\cos\gamma \\ &+ \cos\delta\cos\phi\cos\beta\cos\omega + \cos\delta\sin\phi\sin\beta\cos\gamma\cos\omega \\ &+ \cos\delta\sin\beta\sin\gamma\sin\omega \ . \end{aligned} \tag{A.10}$$

A.2.2 Solar time

The computation of θ using (A.10) depends on the hour angle ω, which is the angular displacement of the sun relative to the local meridian due to the rotation of the earth on its axis. This angle relates directly to solar time t_{so}. Solar noon is the time at which the sun crosses the meridian and ω equals zero. We compute t_{so}, in minutes, by

$$t_{so} = 720 + 4\omega \ . \tag{A.11}$$

Standard time t_{st} and solar time t_{so} are offset due the difference in longitude of the collector L_c, and the longitude of the standard time meridian L_{st}. An additional correction E, which is a function of the day of the year, is required to take into account perturbations in the earth's rotation [33]. We compute t_{st} by

$$t_{st} = t_{so} - 4(L_{st} - L_c) - E \ , \tag{A.12}$$

where E is given by

$$\begin{aligned} E = 229.2(&0.000075 + 0.001868\cos B - 0.032077\sin B \\ &- 0.014615\cos 2B - 0.04089\sin 2B) \ , \end{aligned} \tag{A.13}$$

and B is given by

$$B = (n-1)\frac{360}{365} \ . \tag{A.14}$$

Of particular interest are the times at sunset and sunrise. The sunset hour angle ω_s, which is equal to the negative of the sunrise hour angle, is a function

Table A.4: Radiation measures.

symbol	radiation name
G_{sc}	solar constant
G_{on}	normal extraterrestrial
G_b	direct beam
G_{bh}	beam horizontal
G_{dh}	diffuse horizontal
G_h	global horizontal
G_{bc}	beam collector
G_{dc}	diffuse collector
G_{rc}	reflected collector
G_c	global collector

of the latitude and the declination angle. We compute ω_s by

$$\omega_s = -\tan\phi \tan\delta \ , \tag{A.15}$$

and then compute sunrise and sunset times by (A.11) and (A.12). We also use (A.11) and (A.12) in reverse to find the hour angle ω for a specific time t_{st}.

A.2.3 Collector radiation

Given the global horizontal radiation, direct beam radiation, the time and the day of the year, we are able to determine the total radiation incident on the collector. Table A.4 lists the various radiation measures we use. We determine the global (or total) collector radiation G_c by the following:

$$G_{dh} = G_h - G_b \cos\theta_z \ , \tag{A.16}$$

$$G_{bc} = G_b \cos\theta \ , \tag{A.17}$$

$$G_{dc} = G_{dh}(\frac{1+\cos\beta}{2}) \ , \tag{A.18}$$

$$G_{rc} = \rho(G_b \sin\alpha_s + G_{dh})(\frac{1+\cos\beta}{2}) \text{ and} \tag{A.19}$$

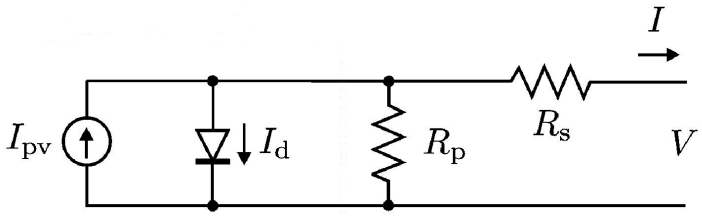


Figure A.2: Single diode model circuit diagram [30].

$$G_c = G_{bc} + G_{dc} + G_{rc} \,, \tag{A.20}$$

where ρ is the ground reflectance coefficient. For collectors with the ability to track the sun, modified versions of (A.17) to (A.19) exist, as described in [34].

A.3 The photovoltaic array

There exist many models for the behavior of a PV array. The single diode model has a balance of accuracy and simplicity and is widely used in output estimation [24, 34]. Figure A.2 shows the circuit diagram of the single diode model we use. In this section, we describe the use of the single diode model to estimate PV array power output.

A.3.1 Radiation spectra

PV cells generate current by using the energy of a specific band of the solar radiation spectrum to free electrons in a doped semiconductor material, such as silicon. As a consequence, the magnitude of power output from a PV array is dependent upon both the magnitude and spectrum of the incident radiation.

The attenuation and scattering of radiation at different frequencies varies, depending on the makeup of the atmosphere through which the radiation travels. In this model, however, it is assumed that the distribution of the spectrum of the solar beam radiation at earth level is constant, irrespective

of its magnitude. This assumption allows the use of the single diode model, which directly uses incident solar radiation magnitudes.

A.3.2 The power output curve

The power output curve, for different values of voltage V of an array at a given collector radiation G_c and cell temperature T_c are determined by

$$I = I_{pv} - I_0[\exp(\frac{V + R_s I}{V_T a}) - 1] - \frac{V + R_s I}{R_p} \text{ and} \tag{A.21}$$

$$P = VI \ , \tag{A.22}$$

using the parameters listed in Table A.5. These parameters, though not explicitly provided by array manufacturers, may be calculated from the data sheet information normally provided [30]. Many of the parameters are functions of collector radiation and cell temperature as per the following:

$$V_T = \frac{N_s k T_c}{q} \ , \tag{A.23}$$

where N_s is the number of cells connected in series, k is the Boltzman constant ($1.3806503 \times 10^{-23}$ J/K) and q is the electron charge ($1.60217646 \times 10^{-19}$ C);

$$I_{pv} = (I_{pv,n} + K_I(T_c - T_{c,n}))\frac{G_c}{G_{c,n}} \ , \tag{A.24}$$

where $G_{c,n}$, $I_{pv,n}$ and $T_{c,n}$ are the nominal collector radiation, cell current and cell temperatures and K_I is the short circuit current/temperature coefficient;

$$I_0 = \frac{I_{sc,n} + K_I(T_c - T_{c,n})}{\exp(\frac{V_{oc,n} + K_V(T_c - T_{c,n})}{aV_T}) - 1} \ , \tag{A.25}$$

where $I_{sc,n}$ and $V_{oc,n}$ are the short circuit current and open circuit voltage at nominal conditions and K_V is the open circuit voltage/temperature coefficient; and

$$R_p = \frac{R_{p,n} G_{c,n}}{G_c} \ , \tag{A.26}$$

where $R_{p,n}$ is the nominal parallel resistance.

Table A.5: Parameters for the single diode model.

symbol	parameter
I_{pv}	PV current
I_0	diode saturation current
a	diode ideality factor
V_T	thermal voltage
R_s	series resistance
R_p	parallel resistance

A.3.3 Maximum power point tracking

Once the power output curve for an array is known, the operating point needs to be determined. In our research we assume a maximum power point tracking (MPPT) system is in place for the array. This system adjusts the operating voltage V, such that the power output is maximized. In order to determine this maximum power point, the derivative of the power with respect to voltage is set to zero, i.e.,

$$\begin{aligned}\frac{dP}{dV} &= I_{pv} - I_0[\exp\left(\frac{V+R_sI}{V_Ta}\right) - 1] \\ &\quad - VI_0\exp\left(\frac{V+R_sI}{V_Ta}\right) - \frac{2V+IR_s}{R_p} \\ &= 0\ ,\end{aligned} \tag{A.27}$$

and solved simultaneously with (A.21) to find the maximum power point voltage and current and the associated maximum dc power output P_{dc}. These equations are nonlinear and require an iterative solution method.

A.3.4 Cell temperature

As described above, the temperature T_c of the cells within a PV array affects the efficiency of the array. We use the ambient air temperature T_a and wind speed v_w at a site to calculate the back surface module temperature T_m by

$$T_m = T_a + G_c\exp(b_1 + b_2v_w)\ , \tag{A.28}$$

where b_1 and b_2 are empirically determined constants. We then compute T_c

by

$$T_c = T_m + \frac{G_c}{G_{c,n}} \Delta_T \ , \tag{A.29}$$

where Δ_T is the empirically determined constant for the cell-to-back-surface-module temperature difference.

A.4 Inverter

The final computation for the TVIPVM entails the modeling of the action of the inverter. The inverter is used to convert the dc power produced by the solar array into ac power so that it can be injected into the power grid. We use the inverter model developed by Sandia National Laboratories [31] to compute the output power P_{ac} by

$$P_{ac} = \frac{P_{ac,n}}{c_1 - c_2} - c_3(c_1 - c_b))(P_{dc} - c_2) + c_3(P_{dc} - c_2)^2 \ , \tag{A.30}$$

where $P_{ac,n}$, c_1 , c_2 and c_3 are empirically determined constants that depend on the properties of the inverter. The determination of these constants is described in detail in [31].

A.5 Model summary

In this appendix, we present the TVIPVM, a model we use to approximate the output power of a PV plant from time-stamped atmospheric measurements. We also present the clear-sky radiation computation we use in the development of benchmark PV output, which we discuss in Appendix B.

APPENDIX B

COMPUTATION OF BENCHMARK PHOTOVOLTAIC POWER OUTPUT

We make use of benchmark PV output patterns in both the scaling of PV output patterns and the re-scaling of scaled PV output patterns. The initial step in the computation of the benchmark output pattern is the determination of the benchmark atmospheric conditions for the site s. We then use these conditions and the TVIPVM described in Appendix A to find the benchmark output pattern $\underline{b}_s$ for each day d of the year.

For the specified site s and day d, we first compute the sunrise time t_r and the sunset time t_f, as described in Section A.2.2. We partition the radiation period (t_r to t_f) into N equal sub-periods. For the mid-point of each of the N sub-periods, we compute the necessary benchmark atmospheric conditions: G_b, G_h, T_a and v_w. We compute the clear-sky beam direct radiation G_b and global horizontal radiation G_h as described in Section A.1. We compute the ambient air temperature T_a and the wind speed V_w by taking the 31-day moving average, centered on day d, of the historical daily atmospheric data we source from NOAA [8]. For the mid-point of each of the N sub-periods, we compute the power output $b_{s,n}$ using the TVIPVM. We construct the benchmark PV output vector $\underline{b}_s \triangleq [b_{s,1}, b_{s,2}, \ \ldots \ , b_{s,N}]^T \in \mathbb{R}^N$ to represent the output power over the N equal sub-periods from t_r to t_f.

APPENDIX C

THE PHOTOVOLTAIC OUTPUT SCALING AND RE-SCALING ALGORITHMS

In this appendix, we describe the scaling and re-scaling algorithms we have developed to perform the scaling and re-scaling mapping functions described below.

We apply the PV output scaling mapping $\alpha_s(\cdot,\cdot)$, for site s, to convert the midnight-to-midnight, PV output pattern $\underline{u}_s$ into the sunrise-to-sunset, scaled-output characterization $\underline{y}_s$. We apply the mapping to compute $\underline{y}_s \in \mathbb{R}^J$ from $\underline{u}_s \in \mathbb{R}^M$ for a site s, where d is the day of the year upon which the measurements of $\underline{u}$ were made, i.e.,

$$\underline{y}_s = \alpha_s(\underline{u}_s, d) \; . \tag{C.1}$$

In Section C.2, we detail the steps of the scaling algorithm we have developed to perform the mapping described in (C.1). In the scaling algorithm, we use the benchmark PV power output vector $\underline{b}_s$ that describes the benchmark output pattern at site s on day d of the year. We compute the sunrise-to-sunset, benchmark PV output $\underline{b}_s \in \mathbb{R}^N$ for day d, as described in Appendix B, in addition to computing the sunrise time t_r and sunset time t_f (in hours), as per the method described in Section A.2.2. It is against this benchmark $\underline{b}_s$ that we scale $\underline{u}_s$.

We also apply the re-scaling mapping $\beta_s(\cdot,\cdot)$, for site s, to convert a sunrise-to-sunset, scaled-output characterization $\underline{y}_s$ into a midnight-to-midnight, re-scaled power output representation $\underline{p}_s$. We apply the algorithm to compute $\underline{p}_s \in \mathbb{R}^H$ from $\underline{y}_s \in \mathbb{R}^J$, where d is the day of the year for which $\underline{p}$ is being created, i.e.,

$$\underline{p}_s = \beta_s(\underline{y}_s, d) \; . \tag{C.2}$$

In Section C.3, we detail the steps of the re-scaling algorithm we have developed to perform the mapping described in (C.2). For the re-scaling algo-

rithm, we use the benchmark PV power output vector $\underline{b}_s$ that describes the benchmark output pattern at site s on day d of the year.

The scaling and re-scaling algorithms require care in the allocation of the energy output over a sub-period on the input time-scale to one or more sub-periods on the output time-scale. We allocate this energy according to the shape of the benchmark output pattern. As a result, both the scaling and re-scaling algorithms employ a benchmark integration algorithm, which we detail in Section C.1.

C.1 Benchmark integration algorithm

We apply the benchmark integration algorithm to calculate the benchmark energy output between the times τ_1 and τ_2. We use Δ_4 to denote the time-resolution of the benchmark output $\underline{b}_s$. The algorithm proceeds as follows.

Step B0: Define $\Delta_4 = (t_f - t_r)/N$, $\tau_r = \max\{t_r, \tau_1\}$, $\tau_f = \max\{t_f, \tau_2\}$.
Step B1: Set $n = 1$.
Step B2: If $(t_r + n\Delta_4) < \tau_r$, continue; else, go to Step B4.
Step B3: $n = n + 1$; return to Step B2.
Step B4: Set $a = 0$, $\tau = \tau_r$, $\tau_4 = t_r + n\Delta_4$.
Step B5: If $\tau_4 \leq \tau_f$, continue; else, go to Step B9.
Step B6: $a = a + b_{s,n}(\tau_4 - \tau)/\Delta_4$.
Step B7: $n = n + 1$; set $\tau = \tau_4$.
Step B8: Set $\tau_4 = t_r + n\Delta_4$.
Step B9: If $\tau_4 \leq \tau_f$, continue; else, go to Step B13.
Step B10: $a = a + b_{s,n}$.
Step B11: $n = n + 1$; set $\tau = \tau_4$.
Step B12: Set $\tau_4 = t_r + n\Delta_4$; return to Step B9.
Step B13: If $n \leq N$, continue; else, go to Step B15.
Step B14: $a = a + b_{s,n}(\tau_f - \tau)/\Delta_4$.
Step B15: $a = a\Delta_4$, benchmark-integration complete.

C.2 Scaling algorithm

We apply the scaling algorithm to perform the mapping described in (C.1). We use Δ_1 to denote the time-resolution of PV output pattern $\underline{u}_s$ and Δ_2 to denote the time-resolution of the scaled-output characterization $\underline{y}_s$. The algorithm proceeds as follows.

Step S0: Define $\Delta_1 = 24/M$, $\Delta_2 = (t_f - t_r)/J$,
$\eta = \max\{b_n : n = 1, 2, \ \ldots \ , N\}$.
Step S1: Set $m = 1$.
Step S2: If $m\Delta_1 < t_r$, continue; else, go to Step S4.
Step S3: $m = m + 1$; return to Step S2.
Step S4: Set $\underline{y}_s = \underline{0}$, $t = t_r$, $t_1 = m\Delta_1$, $j = 1$.
Step S5: If $j \leq J$, continue; else, go to Step S25.
Step S6: Set $t_2 = t_r + j\Delta_2$.
Step S7: If $t_1 \leq t_2$, continue; else, go to Step S15.
Step S8: Set $\tau_1 = t$, $\tau_2 = t_1$.
Step S9: Go to Step B0; set $a_1 = a$.
Step S10: Set $\tau_1 = (m-1)\Delta_1$, $\tau_2 = m\Delta_1$.
Step S11: Go to Step B0; set $a_2 = a$.
Step S12: $y_j = y_j + (a_1/a_2)u_m$.
Step S13: $m = m + 1$; set $t = t_1$.
Step S14: Set $t_1 = m\Delta_1$.
Step S15: If $t_1 \leq t_2$, continue; else, go to Step S19.
Step S16: $y_{s,j} = y_{s,j} + u_m$.
Step S17: Set $t = t_1$, $m = m + 1$.
Step S18: Set $t_1 = m\Delta_1$; return to Step S15.
Step S19: Set $\tau_1 = t$, $\tau_2 = t_2$.
Step S20: Go to Step B0; set $a_1 = a$.
Step S21: Set $\tau_1 = (m-1)\Delta_1$, $\tau_2 = m\Delta_1$.
Step S22: Go to Step B0; set $a_2 = a$.
Step S23: $y_{s,j} = y_{s,j} + (a_1/a_2)u_m$.
Step S24: Set $t = t_2$, $j = j + 1$; return to Step S5.
Step S25: $\underline{y}_s = \underline{y}_s(\Delta_1/\eta\Delta_2)$; scaling complete.

C.3 Re-scaling algorithm

We apply the re-scaling algorithm to perform the mapping described in (C.2). We use Δ_2 to denote the time-resolution of the scaled-output characterization $\underline{y}_s$ and Δ_3 to denote the time-resolution of the re-scaled output representation $\underline{p}_s$. The algorithm proceeds as follows.

Step R0: Define $\Delta_2 = (t_f - t_r)/J$, $\Delta_3 = 24/H$,
$\eta = \max\{b_n : n = 1, 2, \ \ldots \ , N\}$.
Step R1: Set $h = 1$.
Step R2: If $h\Delta_3 < t_r$, continue; else, go to Step R4.
Step R3: $h = h + 1$; return to Step S2.
Step R4: Set $\underline{p}_s = \underline{0}$, $t = t_r$, $t_3 = h\Delta_3$,$j = 1$.
Step R5: If $j \leq J$, continue; else, go to Step R24.
Step R6: Set $t_2 = t_r + j\Delta_2$.
Step R7: If $t_2 \leq t_3$, continue; else, go to Step R10.
Step R8: $p_{s,h} = p_{s,h} + y_j$.
Step R9: Set $t = t_2$; $j = j + 1$; return to Step R5.
Step R10: If $t_3 \leq t_2$, continue; else, go to Step R18.
Step R11: Set $\tau_1 = t$, $\tau_2 = t_3$.
Step R12: Go to Step B0; set $a_1 = a$.
Step R13: Set $\tau_1 = (j - 1)\Delta_2$, $\tau_2 = j\Delta_2$.
Step R14: Go to Step B0; set $a_2 = a$.
Step R15: $p_{s,h} = p_{s,h} + (a_1/a_2)y_j$.
Step R16: $h = h + 1$; set $t = t_3$.
Step R17: Set $t_3 = h\Delta_3$; return to Step R10.
Step R18: Set $\tau_1 = t$, $\tau_2 = t_2$.
Step R19: Go to Step B0; set $a_1 = a$.
Step R20: Set $\tau_1 = (j - 1)\Delta_2$, $\tau_2 = j\Delta_2$.
Step R21: Go to Step B0; set $a_2 = a$.
Step R22: $p_{s,h} = p_{s,h} + (a_1/a_2)y_j$.
Step R23: Set $t = t_2$, $j = j + 1$; return to Step R5.
Step R24: $\underline{p}_s = \underline{p}_s(\eta\Delta_2/\Delta_3)$; re-scaling complete.

REFERENCES

[1] BP, "Statistical review of world energy 2011," June 2011, BP p.l.c., London, United Kingdom. [Online]. Available: http://www.bp.com/sectionbodycopy.do?categoryId=7500&contentId=7068481

[2] DSIRE, "Database of state incentives for renewables & efficiency," Mar. 2011, North Carolina Solar Center, Raleigh, NC. [Online]. Available: http://dsireusa.org/

[3] S. Teske et al., "Solar generation 6," Feb. 2011, European Photovoltaic Industry Association, Brussels, Belgium. [Online]. Available: http://www.epia.org/publications/photovoltaic-publications-global-market-outlook.html

[4] A. Wood and B. Wollenberg, *Power Generation, Operation and Control*, 2nd ed. New York, NY: John Wiley & Sons, Inc., 1996, ch. 8, pp. 264–327.

[5] NREL, "Measurement and instrumentation data center," Dec. 2010, National Renewable Energy Laboratory, Golden, CO. [Online]. Available: http://www.nrel.gov/midc/unlv/

[6] ERCOT, "Hourly load data archives," May 2011, Electric Reliability Council of Texas, Austin, TX. [Online]. Available: http://www.ercot.com/gridinfo/load/load_hist/

[7] NREL, "National solar radiation database - 1991-2005 update," May 2007, National Renewable Energy Laboratory, Golden, CO. [Online]. Available: http://rredc.nrel.gov/solar/old_data/nsrdb/1991-2005/hourly/list_by_state.html

[8] NOAA, "National climatic data center," May 2011, National Oceanic and Atmospheric Administration, Silver Spring, MD. [Online]. Available: http://cdo.ncdc.noaa.gov/dlyp/

[9] F. Mavromatakis et al., "Modeling the photovoltaic potential of a site," *Renewable Energy*, vol. 35, no. 7, pp. 1387–1390, July 2010.

[10] L. Pratt and D. King, "The effect of uncertainty in modeling coefficients used to predict energy production using the Sandia Array Performance Model," in *35th IEEE Photovoltaic Specialists Conference*, June 2010, pp. 2718–2723.

[11] F. Safie, "Probabilistic modeling of solar power systems," in *IEEE Annual Reliability and Maintainability Symposium*, Jan. 1989, pp. 425–430.

[12] I. Colak et al., "Determining suitability of locations for installation of solar power station based on probabilistic inference," in *IEEE Ninth International Conference on Machine Learning and Applications*, Washington, DC, Dec. 2010, pp. 714–719.

[13] H. Bayem et al., "Probabilistic study of the maximum penetration rate of renewable energy in an island network," in *IEEE 2009 Powertech Conference*, June 2009, pp. 1–5.

[14] D. Lei et al., "Probabilistic load flow analysis considering power system random factors and their relevance," in *2011 Asia-Pacific IEEE Power and Energy Engineering Conference (APPEEC)*, Mar. 2011, pp. 1–4.

[15] A. Keane et al., "Capacity value of wind power," *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 564–572, May 2011.

[16] D. Halamay and K. Ted, "Monte Carlo analysis of the impacts of high renewable power penetration," in *2011 IEEE Energy Conversion Congress and Exposition (ECCE)*, Sep. 2011, pp. 3059–3066.

[17] F. Ruiz-Rodriguez et al., "Probabilistic load flow for radial distribution networks with photovoltaic generators," in *2011 IEEE International Conference on Power Engineering, Energy and Electrical Drives*, May 2011, pp. 1–6.

[18] A. da Silva et al., "Long-term probabilistic evaluation of operating reserve requirements with renewable sources," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 106–116, Feb. 2010.

[19] T. Soubdhan et al., "Classification of daily solar radiation distributions using a mixture of Dirichlet distributions," *Solar Energy*, vol. 83, no. 7, pp. 1056–1063, July 2009.

[20] E. Lorenz et al., "Irradiance forecasting for the power prediction of grid-connected photovoltaic systems," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 2, no. 1, pp. 2–10, Jan. 2009.

[21] L. Martin et al., "Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning," *Solar Energy*, vol. 84, no. 10, pp. 1772–1781, Oct. 2010.

[22] N. Maisonneuve and G. Gross, "A production simulation tool for systems with integrated wind energy resources," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2285–2292, Nov. 2011.

[23] M. Milligan, "Methods to model and calculate capacity contributions of variable generation for resource adequacy planning (IVGTF1-2): Additional discussion," National Renewable Energy Laboratory, Atlanta, GA, Tech. Rep. NREL/PR-5500-50355, Jan. 2011.

[24] J. Duffie and W. Beckman, *Solar Engineering of Thermal Processes*, 3rd ed. Hoboken, NJ: John Wiley & Sons, Inc., 2006, pp. 3–146.

[25] G. Gan et al., *Data clustering: Theory, algorithms, and applications*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 2007.

[26] N. Maisonneuve, "A production simulation tool for systems with integrated wind energy resources." M.S. thesis, Dept. Elec. and Comp. Eng., Univ. Illinois, Urbana-Champaign., 2009.

[27] C. Grigg et al., "The IEEE reliability test system - 1996," *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 1010–1020, Aug. 1999.

[28] EIA, "Annual electric power generator data - EIA-860 files," Nov. 2010, Energy Information Administration, Washington D.C. [Online]. Available: http://www.eia.gov/cneaf/electricity/page/eia860.html

[29] EIA, "Electric power annual," Nov. 2010, Energy Information Administration, Washington D.C. [Online]. Available: http://www.eia.gov/cneaf/electricity/epa/epa_sum.html

[30] M. Villalva et al., "Comprehensive approach to modeling and simulation of photovoltaic arrays," *IEEE Transactions on Power Systems*, vol. 24, no. 5, pp. 1198–1208, May 2009.

[31] D. King et al., "Performance model for grid-connected photovoltaic inverters," Sandia National Laboratories, Albuquerque, NM, Tech. Rep. SAND2007-5036, Sep. 2007.

[32] L. Garver, "Effective load carrying capability of generating units," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-85, no. 8, pp. 910–919, Aug. 1966.

[33] M. Iqbal, *An Introduction to Solar Radiation*. New York, NY: Academic Press, Inc., 1983.

[34] G. Masters, *Renewable and Efficient Power Systems*. Hoboken, NJ: John Wiley & Sons, Inc., 2004.