SOLAR PHOTOVOLTAIC POWER INTERMITTENCY UNDER PASSING CLOUDS: CONTROL, FORECASTING, AND EMULATION

Thesis submitted in accordance with the requirements of
the University of Liverpool for the degree of Doctor in Philosophy by

XIAOYANG CHEN
(B. Eng. (Hons.), UoL)

Department of Electrical Engineering and Electronics
School of Electrical Engineering, Electronics and Computer Science

May 2021

I would like to dedicate this thesis to my mother YANG Hong.

We all miss you.

Acknowledgements

First of all, I would like to express my gratitude towards my supervisor Prof. Eng Gee Lim for offering me the opportunity to study in the university and facilitating all administrative matters. His professionalism and energy have been a great source of motivation and inspiration for me. I would also like to thank my co-supervisors Dr. Huiqing Wen and Prof. Lin Jiang for their support and suggestions during the entire period of my Ph.D. study.

Sincere thanks go to my external supervisor Dr. Yang Du for his valuable guidance, kindness, and patience during the years. His support and encouragement are the main drivers for me to complete my Ph.D. study. I will always be very grateful.

I would like to thank my colleagues, Chaoxian Wu, Haoran Wen, Bidan Zhang, and Jiaming Chang for their fruitful discussion and enlightenments on specific domains. It has been my pleasure to work with you. Furthermore, special thanks to my friends Zhe Jiang, Zhao Wang, Pengxiang Hao, Xiaoyu Lei, and Kai Bao for their companies and encouragement over the years.

Last but not least, I would like to extend my greatest gratitude to my family members for their continuous love throughout my life. Especially, I am deeply grateful to my father Weiquan Chen, who supports and encourages me to find my own path in life. Special thanks are reserved to my fiancee Lin Zheng, without your patience, comfort, and encouragement over the past four years, I would not have achieved this goal.

Abstract

Solar photovoltaic (PV) energy is becoming an increasingly vital source in electricity grids for energy harvesting. Inspired by the regulatory incentives and plummeting cost, the integration of utility-scale PV systems into the power grid is boosting. Nonetheless, due to the natures of cloud movements, PV system exhibits rapid power ramp-rates (RR) in the output profiles, which poses significant challenges for system operators to maintain grid transient stability. In this context, this thesis focuses on the management of cloud-induced solar PV intermittency. Three aspects for coping with solar intermittency are addressed, namely, control, forecasting, and emulation.

Firstly, from the control aspect, two predictive PV power RR control (PRRC) strategies are presented. To regulate system RRs, conventional methods are implemented either by active power curtailment (APC) or energy storage control (ESS). However, current APC method cannot deal with the ramp-down fluctuations, and the integration of an ESS is still costly. On this point, two innovative PRRC strategies are proposed, which are based on a solar nowcasting system. The first strategy does not require any ESS. With the prior knowledge of upcoming RRs, PV generation can be regulated before the actual shading occurs. The second strategy improves the conventional ESS method with minimal support of energy storage. The results show that both of the proposed strategies can effectively comply with RR regulations, and outperform the conventional methods.

Then, in terms of forecasting, an improved sensor network-based spatio-temporal nowcasting method is developed. The proposed nowcasting method overcomes the

shortcomings that typically associated with existing sensor network-based nowcasting methods, such as predictor mis-selection, inconsistent nowcasting, and poor model adaptability. The experimental results reveal that the proposed nowcasting method is more suitable for predicting system RRs. Subsequently, the operability of solar nowcasting for PRRC practice is demonstrated. To that end, temporal issues related to operational solar nowcasting are identified, and their effects on nowcasting and PV control performance are evaluated.

Lastly, from the emulation aspect, this thesis sets forth a partial shading emulator and a cloud shadow model, which can emulate the module-level responses of utility-scale PV systems under passing clouds. Based on the emulation tools, the characteristics of PV system RRs are comprehensively investigated across various system and cloud shadow attributions. The results indicate that a utility-scale PV system can frequently violate the RR limit imposed by grid operators. Hence, advanced RR control strategies should be essential for system operators to comply with the RR regulations.

List of Publications

Journal

J1. **X. Chen**, Y. Du, H. Wen, L. Jiang, and W. Xiao, "Forecasting-based power ramp-rate control strategies for utility-scale PV systems," *IEEE Trans. Ind. Electron.*, vol. 66, no. 3, pp. 1862-1871, 2019. (Q1, Impact factor: 7.515).

J2. **X. Chen**, Y. Du, E. Lim, H. Wen, and L. Jiang, "Sensor network based PV power nowcasting with spatio-temporal preselection for grid-friendly control," *Appl. Energy*, vol. 255, pp. 113760, 2019. (Q1, Impact factor: 8.848).

J3. **X. Chen**, Y. Du, E. Lim, H. Wen, K. Yan, and J. Kirtley, "Power ramp-rates of utility-scale PV systems under passing clouds: Module-level emulation with cloud shadow modeling," *Appl. Energy*, vol. 268, pp. 114980, 2020. (Q1, Impact factor: 8.848).

J4. **X. Chen**, Y. Du, E. Lim, and H. Wen, "Towards the applicability of solar nowcasting: A practice of operational solar nowcasting for PV power ramp-rate control," *Appl. Energy*, 2021. (Status: Under review).

J5. H. Wen, Y. Du, **X. Chen**, E. Lim, H. Wen, L. Jiang, and W. Xiang, "Deep learning based multistep solar forecasting for PV ramp-rate control using sky images," *IEEE trans. Ind. Inform.*, vol. 17, pp. 1397-1406, 2021 (Q1, Impact factor: 9.112).

Conference

C1. **X. Chen**, Y. Du, and H. Wen, "Forecasting based power ramp-rate control for PV systems without energy storage," *IEEE International Future Energy Electronics Conference and ECCE Asia (IFEEC 2017 - ECCE Asia)*, pp. 733–738, 2017.

C2. **X. Chen**, Y. Du, W. Xiao and S. Lu, "Power ramp-rate control based on power forecasting for PV grid-tied systems with minimum energy storage," *IECON 2017 - Annual Conference of the IEEE Industrial Electronics Society*, pp. 2647-2652, 2017.

C3. C. Zhang, Y. Du, **X. Chen** and D. D. Lu, "Cloud motion tracking system using low-cost sky imager for PV power ramp-rate control," *IEEE International Conference on Industrial Electronics for Sustainable Energy Systems (IESES)*, pp. 493-498, 2018.

C4. C. Zhang, Y. Du, **X. Chen** and E. G. Lim, "Cloud motion forecasting and cloud base height estimation using two low-cost sky cameras," *IEEE Conference on Energy Internet and Energy System Integration (EI2)*, pp. 1-6, 2018. (Best paper award).

C5. J. Chang, Y. Du, **X. Chen**, E. Lim and K. Yan, "Forecasting based virtual inertia control of PV systems for islanded micro-grid," *Australasian Universities Power Engineering Conference (AUPEC)*, pp. 1-6, 2019.

Contents

List of Figures

List of Tables

Nomenclature

AnEn	Analog Ensemble
APC	Active Power Curtailment
CAISO	California Independent System Operators
CCF	Cross-Correlation Coefficient
CMV	Cloud Motion Vector
CSI	Clear-Sky Index
CSMV	Cloud Shadow Motion Vector
DNI	Direct Normal Irradiance
DST	Dynamic Spatio-Temporal
ECR	Energy Curtailment Ratio
ESS	Energy Storage System
FPPT	Flexible Power Point Tracking
FS	Forecast Skill
GFS	Global Forecast System

GHG	Greenhouse Gases
GHI	Global Horizontal Irradiance
kNN	k-Nearest Neighbors
Lasso	Least Absolute Shrinkage and Selection Operator
LR	Linear Regression
LSTM	Long Short-Term Memory Recurrent Neural Networks
MCP	Most-Correlated Pair
MPPT	Maximum Power Point Tracking
NAM	North American Mesoscale
nCRPS	Normalized Continuous Ranked Probability Score
nMAE	Normalized Mean Absolute Error
nMBE	Normalized Mean Bias Error
nPMAE	Normalized Peak Mean Absolute Error
NREL	National Renewable Energy Lab oratory
nRMSE	Normalized Root Mean Square Error
NWP	Numerical Weather Prediction
OLS	Ordinary Least Squares
PAPC	Predictive Active Power Curtailment
PDM	Peak Difference Minimization
PEC	Predictive ESS Control
PeEn	Persistence Ensemble

Pers	Persistence Model
PI	Prediction Interval
PICP	Prediction Interval Coverage Probability
PINAW	Prediction Interval Normalized Average Width
PRC	Power Reserve Control
PRRC	Power Ramp-Rate Control
PV	Photovoltaic
QR	Quantile Regression
RR	Ramp-Rate
RSR	Ramp Smoothing Rate
SOC	State of Charge
SP	Smart Persistence
SRP	Scenario-Recognizable Preselection
SRP-Enet	SRP-Based Elastic-Net
STC	Standard Test Conditions
WMO	World Meteorological organization
WRF	Weather Research and Forecasting

Chapter 1

Introduction

1.1 Background

Electricity is vital for economic development and technological growth. It is a key factor in modern urbanization and industrialization, to the extent that economic growth is frequently measured in per capita power output of a country [1]. This ever-growing energy demand leads to an increasing need for electricity generation and distribution. However, globally, the reliance of electricity production is still on non-renewable pollution-causing fossil fuels. Approximately two-thirds of the global carbon dioxide emissions are from such fuel sources whose current share of energy production, if maintained, will inevitably lead to a significant rise in average global temperature and other catastrophes [2]. The World Meteorological Organization's (WMO) provisional statement on the State of Global Climate mentioned that the year 2019 witnessed one decade of unprecedented elevated global temperature, retreating glaciers and record high sea levels due to the greenhouse gas (GHG) emissions. The average global temperatures for the past five (2015-2019) and ten (2010-2019) years were the highest in recorded history.

Fortunately, the rapid development of renewable energies such as solar, wind, or tidal, has brought sweeping changes in the arena of energy generation and reveals the potential of clean and sustainable energy for the future. Moreover, policies enacted by international organizations and major players in world economy regarding carbon

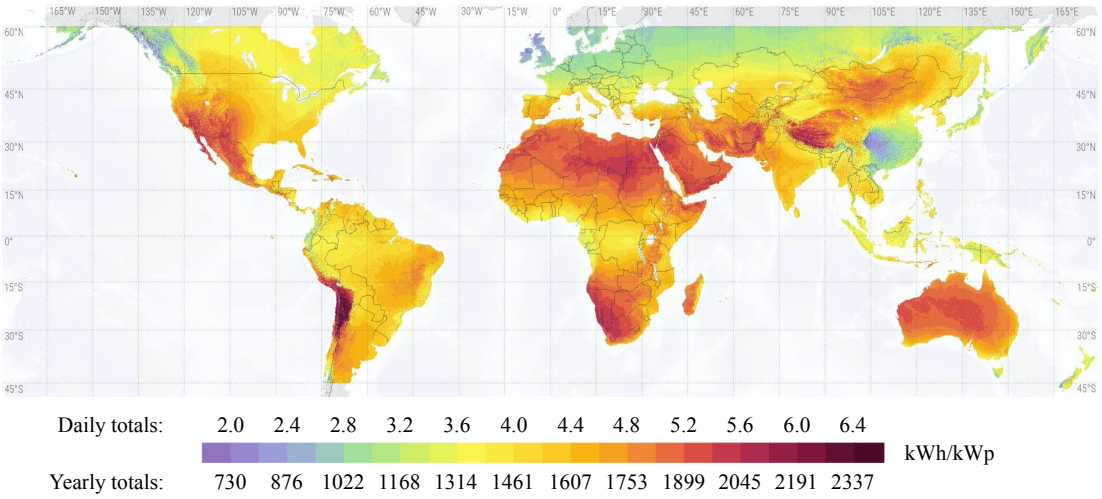


Figure 1.1: Long-term average of PV power potential. Countries located above 45°N or below 45°S latitudes show tremendous potential for harnessing solar energy. Data source: Solargis.

taxation have paved the way for the renewables [3]. For instance, the European Union has stated to reduce GHG emissions by 80% (from a 1990 baseline) and to produce 100% green energy from the renewables by 2050 [4]. In this context, the transition from fossil fuels to clean energy resources has become a global development trend.

Among the renewable energies, solar photovoltaic (PV) energy is considered as one of the most promising resources for energy harvesting, and gaining global popularity in recent years. With ever dropping levelized cost of electricity (LCOE),[1] the market size of solar PV is boosting [5]. Figure 1.1 shows the long-term average of PV power potential across the globe. By the end of 2019, over 115 GW PV systems are newly added to the global PV market, which has raised the cumulative installed capacity to well above 620 GW [6]. Nonetheless, as opposed to conventional energy source such as fossil fuels, solar energy is perceived as an inconsistent resource. With the increasing penetration of PV systems, there is a growing concern that the variable PV generation can strain the grid [7].

[1]LCOE is the expected net present value of energy generation over the lifetime, offset by the system costs in unit money per unit of produced energy

1.1.1 Soalr Resource Intermittency

Due to the natures of sun movement and cloud transitions, solar resource exhibits both long-term and short-term variabilities. Figure 1.2 depicts the variation of global horizontal irradiance (GHI) during a typical day. In a coarse sense, solar resource follows the diurnal cycle of sun movement, with peak appearing at solar noon and zero from sunset to sunrise (see Figure 1.2(a)). In a finer view, the frequent cloud coverage gives rise to the intermittent irradiance fluctuations (see Figure 1.2(b)), which brings more uncertainties to the daily solar pattern. In this context, the short-term solar variability—the "noise" caused by passing clouds—is termed as solar *intermittency*.

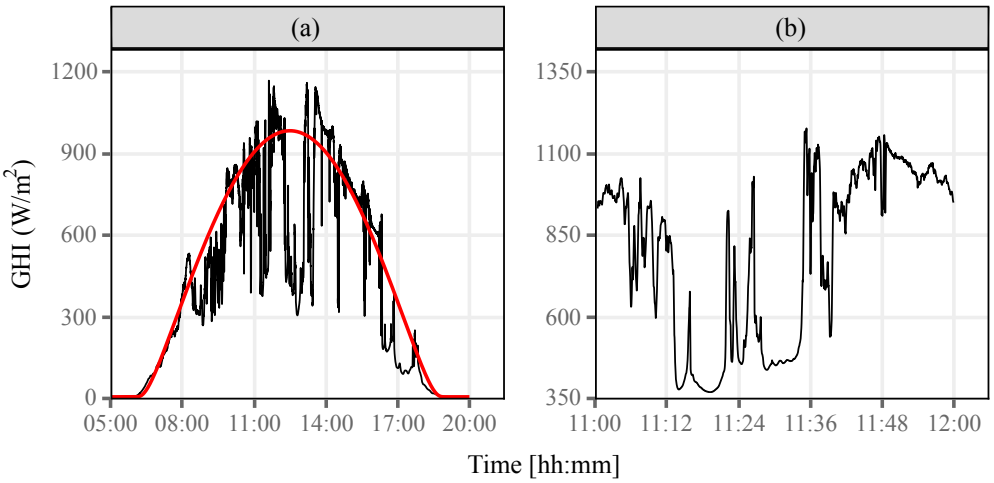


Figure 1.2: (a) Time series plot for GHI measured on 2011 May 15 at Oahu, Hawaii (black line). The red line shows the clear-sky GHI computed by the Ineichen-Perez clear-sky model. (b) A zoomed view for 11:00-12:00.

In general, the long-term solar variability due to sun movement is precisely predictable and only causes notable PV power variations over timescales of hours. Provided with the solar position, target location, local time, and other meteorological variables (e.g., extraterrestrial irradiance, aerosols, water vapor, etc.), a clear-sky model can be developed to fully capture the diurnal trend in solar irradiance [8]. To give perspective, Figure 1.2(a) demonstrates an example of retrieving clear-sky GHI based on the Ineichen-Perez clear-sky model (see the red line) [9]. Owing to the high predictability of the daily

solar cycle, grid operators have effectively exploited a series of system operations to handle the peak-valley characteristics of PV generation, such as demand side management or reserve capacity planning [10].

In contrast to long-term solar variability, the cloud-induced solar intermittency is more challenging to manage, which impacts the resultant PV yield in timescales of seconds. To quantify solar intermittency, the term *ramp-rate* (RR) is commonly adopted as a measure of irradiance/PV power change rate [11]. When a utility-scale PV system or accumulated high penetration is considered, the large power RRs can seriously damage the grid transient stability, leading to voltage flicker, frequency deviations, even blackouts [12]. Nonetheless, as PV system RRs are highly dependent on both the endogenous factors that configure the PV systems (e.g., array arrangement, system size, and orientation, etc,) and exogenous factors that are relevant to cloud properties (e.g., velocity, size, or intensity of the cloud shadows), the characteristics of RRs can vary from plant to plant and day to day, which makes it difficult for grid operators to plan adequate grid reserves [13]. In this sense, solar intermittency is becoming a key barrier to promote high PV integration in modern power systems.

1.1.2 Power Ramp-Rate Control of PV Systems

In order to address the adverse effects of intermittent PV power generation that continue to be integrated into the power grid, PV *power ramp-rate control* (PRRC) has been imposed by grid operators in different countries. Given a predefined RR limit, PRRC aims to constrain the PV output power change rate to a certain RR level, thereby

Table 1.1: An overview of several typified PRRC standards in different countries.

Region	Upward RR	Downward RR
Puerto Rico	10%/min	10%/min
Germany	10%/min	10%/min
Denmark	100 kW/s	100 kW/s
Hawaii	2 MW/min	2 MW/min
Ireland	30 MW/min	No Req.

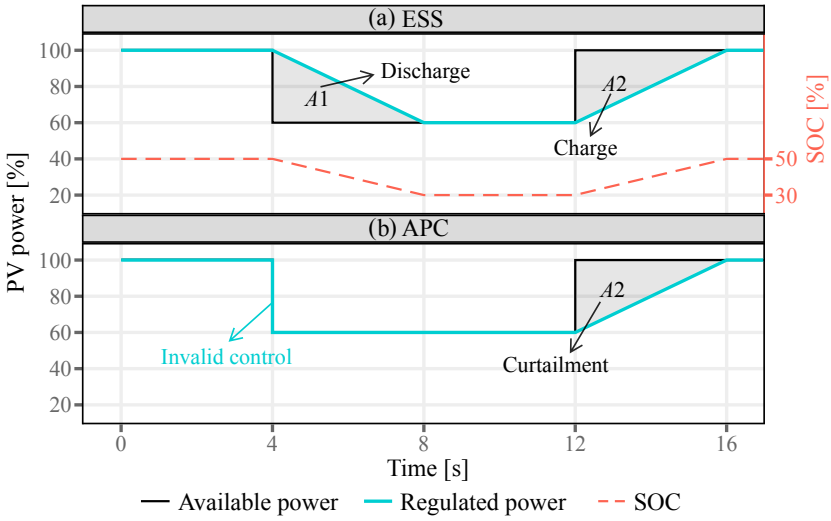


Figure 1.3: An example RR scenario over a sudden 40% power drop follows a recovery. Assuming a RR limit of 10%/s, PRRC is implemented based on (a) ESS, and (b) APC, respectively.

reducing the fluctuation in the PV power injection. Table 1.1 summarizes several typified RR standards introduced in [14–17]. To comply with PRRC requirement, PV output power is typically mitigated by utilizing the energy storage system (ESS) [18–20], or active power curtailment (APC) [21–23]. Figure 1.3 illustrates the control principles of the two PRRC methods.

ESS is considered as the most straightforward way to achieve PRRC. As shown in Figure 1.3(a), when PV power decreases faster than the RR limit, ESS releases energy (discharge) to compensate the sudden power drop (area A_1). When PV power increases faster than the RR limit, ESS stores energy (charge) to suppress the rising PV generation (area A_2). Besides, to ensure the ESS operability, additional state of charge (SOC) control is mandatory to maintain a SOC reference around 50% [24]. Most generally, the ESS storage requirement needs to be determined from the worst case that the PV generation decreases from full production to almost none within a very short period. In this regard, the ESS power capacity should be designed equal to the rated power of the PV system, and a doubled energy capacity is required [25]. Although integrating ESS can effectively regulate PV system RRs, the high installation and maintenance cost is

still hindering its large-scale application.

Since the utilization of ESS introduces additional costs, the APC method provides a self-supporting solution to PRRC. The principle of APC is based on the flexible power point tracking (FPPT) of the PV converters [26]. Different from the conventional PV system operations that working under maximum power point tracking (MPPT), FPPT allows users to flexibly modify the PV operating point to meet different operation commands. Fig. 1.3(b) depicts the PRRC implementation with APC. Typically, the application of APC is limited at the power ramp-up side, where the FPPT is activated to curtail the surplus energy of A_2. For the ramp-down RRs, however, APC becomes invalid as no extra energy source can be used to compensate for the power drop. On this point, one possible solution is to combine the APC and ESS methods [27]. In that, the ESS is merely used to handle the ramp-down RRs, and the required ESS energy capacity can be halved. Nonetheless, even with a halved capacity, the resultant financial burden of deploying an ESS for utility-scale PV systems can still be heavy.

1.1.3 Solar Forecasting

Besides leveraging advanced control operations for PV systems, forecasting of the solar irradiance and projected PV generation is becoming an increasingly important aspect to cope with solar variabilities, which brings solar PV one step closer to being "grid-friendly" [28]. In the field of energy meteorology, the two topics, namely, solar irradiance forecasting and PV power forecasting, are jointly known as *solar forecasting* [29]. For the former, solar forecasters are interested in forecasting the irradiance components, such as GHI, or direct normal irradiance (DNI). For the latter, the forecast GHI or DNI needs to be further converted into the solar-generated power, e.g., using an irradiance-to-power conversion model [30]. Due to the limitation of input data resolution, solar forecasting methods are conventionally categorized into three classes: 1) day-ahead forecasting, 2) intra-day forecasting, and 3) intra-hour forecasting [31]. However, with the recent advances in sensing technologies, data at higher spatial and temporal scales become available. Hence, in this thesis, solar forecasting methods are classified into four

categories, with the fourth class of methods being intra-minute forecasting.

Day-ahead solar forecasting

Day-ahead solar forecasting is normally used for power system economic planning and unit commitment. It covers forecast horizons up to 48 h, depending on when forecasts are issued [32]. In this case, numerical weather prediction (NWP) is often applied, which takes a detailed physical description of the atmosphere into consideration and directly simulates the irradiance fluxes at multiple levels in the atmosphere. Generally, the performance of a NWP model can highly rely on the user's knowledge on the physics options. Furthermore, as most of the NWP models are not adapted specifically for solar forecasting purposes, the NWP forecasts are commonly biased. As a result, statistical post-processing such as the application of model output statistics and Kalman filtering are often needed to empirically adjust the NWP output [33].

Intra-day solar forecasting

Intra-day solar forecasting covers a forecast horizon from 1 to 6 h and is often applied to load following. Apart from NWP, satellite imaging is widely used for generating forecasts at this timescale [34]. The satellite imaging-based approaches estimate solar irradiance using cloud images captured by instruments onboard geostationary satellites. Both physical and statistical models can be used to map the satellite images to irradiance [35]. To complete the forecast, cloud motion vector (CMV) fields between two consecutive images are identified through cloud tracking algorithms, such as block matching [36] or optical flow [37]. Subsequently, by projecting the CMV field 1-step ahead, the areal irradiance forecasts can be obtained. One issue with satellite imaging-based approaches is the spatial-inhomogeneous systematic bias potentially embedded in the derived irradiance data [38]. On this point, site adaptation is considered as a must prior to using these approaches [39].

Intra-hour solar forecasting

Intra-hour solar forecasting covers forecast horizons from a few minutes to 1 h. It is important for grid operators to optimally schedule spinning reserves and demand response [40]. For timescales smaller than 1 h, the main factor causing changes in solar irradiance is the presence of local clouds. At this stage, ground-based sky imagers are frequently utilized. The sky imager is literally a bottom-up approach that uses a wide angle (fish-eye) lens or curved mirror to project the full sky hemisphere onto a finite range. An irradiance sensor is often deployed to map the image pixels to irradiance values. The methodology for making forecast with sky imagers largely resembles the satellite-based techniques, namely, given a deduced CMV, the irradiance fields are propagated forward in time resulting in a final forecast. It should be noted that the forecast horizon of a sky imager is typically restricted to ∼30 min due to the residence time of clouds over the field of view. Besides the upper bound, the lower bound on the forecast horizon of a sky imager is resulted from the circumsolar glare, which renders forecast horizon shorter than 2 min inaccessible [41].

Intra-minute solar forecasting

Intra-minute solar forecasting, also known as *solar nowcasting*, is a new subdomain of solar forecasting. It was not until the early 2010s, when modern solar forecasting is considered to start, that the first papers on solar nowcasting appeared [29]. As it provides forecasts at shorter time horizons, i.e., a few seconds to 1 min, solar nowcasting has gained recognition in coping with solar intermittency and facilitating real-time PV system control [42]. In contrast to intra-day or intra-hour forecasting that is based on CMV measurement, intra-minute solar forecasting typically relies on the detection of local cloud shadow motion vector (CSMV) fields. Most notably is the use of shadow cameras [43], and wireless sensor network [44].

Whereas sky imagers take a bottom-up approach in generating irradiance maps over an area, shadow cameras consider a top-down approach, by taking photos of the ground from an elevated position below the clouds (e.g., 87 m above ground). With

additional ground-based irradiance measurements, the detected shadow maps then can be used to infer the future irradiance conditions. At present, the studies on shadow cameras are quite rare; only a handful of works are available [45]. In this context, the forecasting aspects of shadow camera applications are still hypothetical. However, as shadow cameras bypass cloud geolocation, they present great potential of advantages over sky imagers.

Instead of providing an explicit view of shadow movements, sensor network-based approaches derive the cloud shadow information indirectly. As clouds propagate over a sensor network, the frequent cloud coverage and opening show sequential peaks in sensor readings. Consequently, CSMVs can be identified from the spatio-temporal correlations among the sensors [46]. The final forecasts are then produced by advecting the CSMVs into the future. The main advantage of sensor network-based solar nowcasting is that it overcomes the challenges typically associated with camera-based approaches, such as cloud-height estimation or pixel-to-irradiance conversion. Moreover, it is easy to be generalized and adapted to different PV systems, as one can flexibly design or adjust the network layout [47]. Nonetheless, forecast horizons of using a sensor network are often limited by the size of the network. Therefore, the correspondence between spatial scale of the network and forecast horizon requires further study.

As a brief summary, Figure 1.4 illustrates the spatial and temporal applicability of current solar forecasting techniques, alongside with noting the associated forecasting purposes.

1.1.4 Emulating and Characterizing Solar PV Ramp-Rates

Another important aspect to tackle solar PV intermittency is to characterize PV system RRs during cloud transitions. As discussed very briefly in Section 1.1.1, the RR of PV systems depends on both *endogenous* and *exogenous* factors. Endogenous factors refer to the inherent attributions of a PV system, such as capacity, array arrangement, or system orientation. On the other hand, exogenous factors are more relevant to the properties of cloud shadows, i.e., shadow velocity, thickness, and size. To study PV

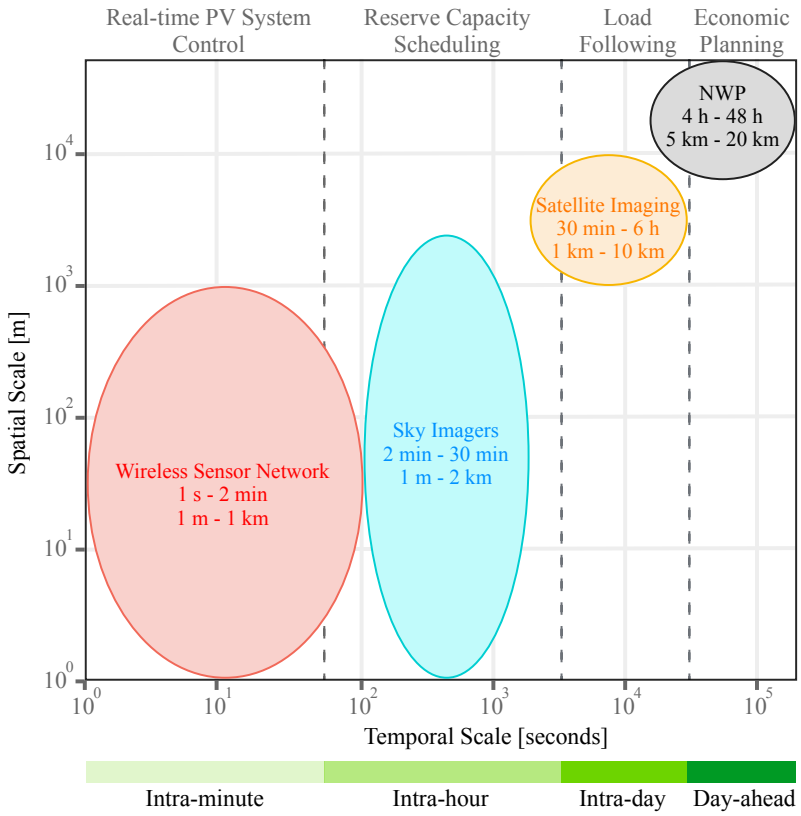


Figure 1.4: Spatial and temporal scale coverages for standard solar forecasting techniques. The potential applications of different forecasts are indicated on the top of the figure .

system RRs, current practice can be categorized into three classes—measurement-based methods, simulator-based methods, and analytical methods.

As the name suggests, measurement-based methods aim to directly use data from practical PV plants to evaluate RRs. Several typified works consist of using 5 min data from 100 PV sites (243 kW in total) in Germany [48], 10 s and 1 min data from a 4.6 MW PV system in the United States [49], and 1 s data from six PV plants ranging from 1 to 9.5 MW in Spain (18 MW in total) [50]. While a majority of these measurement-based works are meant to analyze the effects of endogenous factors (e.g. system geographic dispersions) on RRs, the exogenous factors are scarcely discussed. In order to match with PV fluctuations, exhaustive sensing of cloud shadow transitions both in spatial and temporal scales is desired. Unfortunately, such high-frequency measurements are

typically unavailable for a practical PV system. Even if the measurements were made possible, the analyses would be site-specific. Consequently, the presented results would be difficult to generalize.

As an alternative to exploring practical PV systems, a wide range of studies are based on computer-aided simulators (mostly Matlab/Simulink) [51–54]. The main advantage of using a simulator is that it provides a flexible simulation layout that allows users to easily adjust system configurations. Among the studies, the simulation resolution is typically set on the module-level (one PV module is the basic simulation unit) to retain sufficient accuracy of system behaviors under partial shading. However, due to heavy computational burdens, the module-level simulations on the whole are limited on evaluating the performance of a single or several PV arrays. In the case of simulating a utility-scale PV system that contains tens even hundreds of arrays, the simulation resolution needs to be downscaled, e.g., to string-level (all the modules in one string are assumed to receive identical irradiance) to relieve the computational complexity. As a consequence, the simulation accuracy decreases, and significant errors can rise especially for partial shading analysis [55]. Besides, the shadings considered in these studies are typically randomly assigned and assumed to be static. In another word, the shadow patterns are generated without fully considering the natural shapes and dynamics of a real shadow coverage. Although several works have devoted to deriving cloud shadow models that can reproduce the spatial diversities of cloud natures, such as using fractals [56], imaging systems [57], and random noise [58], the model applicability of interacting with PV system simulations are seldom presented.

Instead of explicitly simulating PV systems and cloud shadows, several studies also focus on the analytical modeling of PV power fluctuations. In [59], an analytical model to quantify the output variability resulting from an ensemble of PV systems, or a PV fleet is developed. The model demonstrates the ensemble variability of a PV fleet as a function of the number of PV systems and the dispersion factor. However, the model is only applicable when the PV systems within the fleet are identical and equally-distributed. Moreover, it requires the measurements of a single PV system as

the input, which may hinder its practical application [60]. In order to simplify the modeling process, the wavelet variability model has been introduced in [61], which can simulate the production of an arbitrary fleet of PV system given the endogenous PV system information and measurements from a single irradiance point sensor. The wavelet model, however, on the whole is more akin to a statistical, or machine learning approach that provides one-stop estimations of ensemble PV generation with limited irradiance data as input, thus the relevant cloud shadow information is left as a black box. In this regard, the model is typically utilized for studying the effects of endogenous factors on RRs, but inferior to assess the cloud-induced fluctuations [62].

1.2 Research Motivation

As aforementioned, solar intermittency has become a major concern for grid operators to maintain steady power system operations. From the technical aspects, PRRC of PV systems provides an immediate solution to counteract the intermittent PV generation. Nonetheless, as the large-scale application of ESS is not yet being commercially available, and the APC method is limited at the power ramp-up side, existing PRRC practice still show deficiencies in efficiently complying with the RR regulations. Therefore, PRRC methods that can handle both upward and downward RR violations while minimizing the reliance on ESS are strongly demanded. On this point, solar forecasting may provide a remedy for improving current PRRC operations. Since clouds are the primary source of large PV power fluctuations, accurate forecasts of cloud motions and resultant PV generation may facilitate a more effective implementation of PRRC.

In terms of solar forecasting, referring to in Figure 1.4, the application of PRRC—a function of real-time PV power control—lies within the scope of intra-minute forecasting (or solar nowcasting). Whereas the applicability of shadow cameras on solar nowcasting is still considered as underdeveloped, nowcasting with a sensor network becomes the first priority for PRRC practice. In this regard, a reliable sensor network-based solar nowcasting method that can cooperate with PRRC needs to be investigated. In addition,

as a forecast always involves several operational time parameters such as forecast horizon, forecast resolution, or forecast updating rate, the effect of these time parameters on PRRC application should be evaluated as well.

On the other hand, emulating and characterizing PV system RRs under passing clouds are also of high importance in order to manage solar intermittency. Since PV system RRs vary with different PV configurations and cloud shadow properties, having the prior knowledge of PV system RRs may offer a possibility to design the PV system in a site-specific way, so that the effect of local cloud movements can be minimized, and PRRC can be implemented more effectively. However, as current studies still show deficiencies in comprehensively investigating PV system RRs rising from both endogenous and exogenous factors, the characteristics of PV system RRs are still not well evaluated. In this sense, modeling approaches to mimic the dynamic performance of PV system RRs during cloud transitions should be explored.

1.3 Research Objectives

In light of the above motivation, the *final goal* of this thesis is to facilitate the mitigation of solar PV intermittency arising from passing clouds. To this end, three aspects, namely, *control*, *forecasting*, and *emulation* of cloud-induced PV power fluctuations are addressed. More specifically, the following research questions are considered:

- Is it possible to improve the performance of conventional PRRC methods through the integration of solar nowcasting information?
- How to ensure the reliability of sensor network-based nowcasting to better cooperate with PRRC operations?
- How the time parameters in a forecast run can affect the forecast applicability on PRRC?
- How to emulate the dynamic performance of PV systems during cloud transitions so as to characterize PV system RRs?

With these research questions, the *objectives* of this thesis can be summarized as follows:

- **Predictive PRRC of PV systems**. As mentioned earlier, solar nowcasting may offer a possibility to implement PRRC more efficiently. In this thesis, the predictive PRRC strategies that integrate solar nowcasts into control process will be developed. The effectiveness of the proposed control strategies will be validated experimentally using nowcasts from a real sensor network.
- **An improved spatio-temporal nowcasting framework for PRRC**. To overcome the limitations that are typically associated with sensor network-based nowcasting, a novel spatio-temporal nowcasting framework will be developed. The applicability of the proposed nowcasting method on predictive PRRC will be evaluated under various control settings and weather conditions.
- **Analysis of operational nowcasting for PRRC**. Since solar nowcasting revolves around operational requirements from grid operators, e.g., time parameters, an in-depth analysis of operational solar nowcasting for predictive PRRC will be carried out. Thad said, an operational nowcasting platform will be set up, and the performance of predictive PRRC will be evaluated in the operational environment, following certain temporal requirements.
- **Emulation of RRs for utility-scale PV systems during cloud transitions**. In order to characterize PV system RRs, emulation tools to mimic the interactions between PV systems and cloud transitions will be developed. Based on emulation, the effect of both endogenous and exogenous factors on PV system RRs will be investigated.

1.4 Thesis Outline

This thesis summarizes the outcomes of the Ph.D. project. The main content of the thesis is based on the collection of papers published during the Ph.D. study. Figure

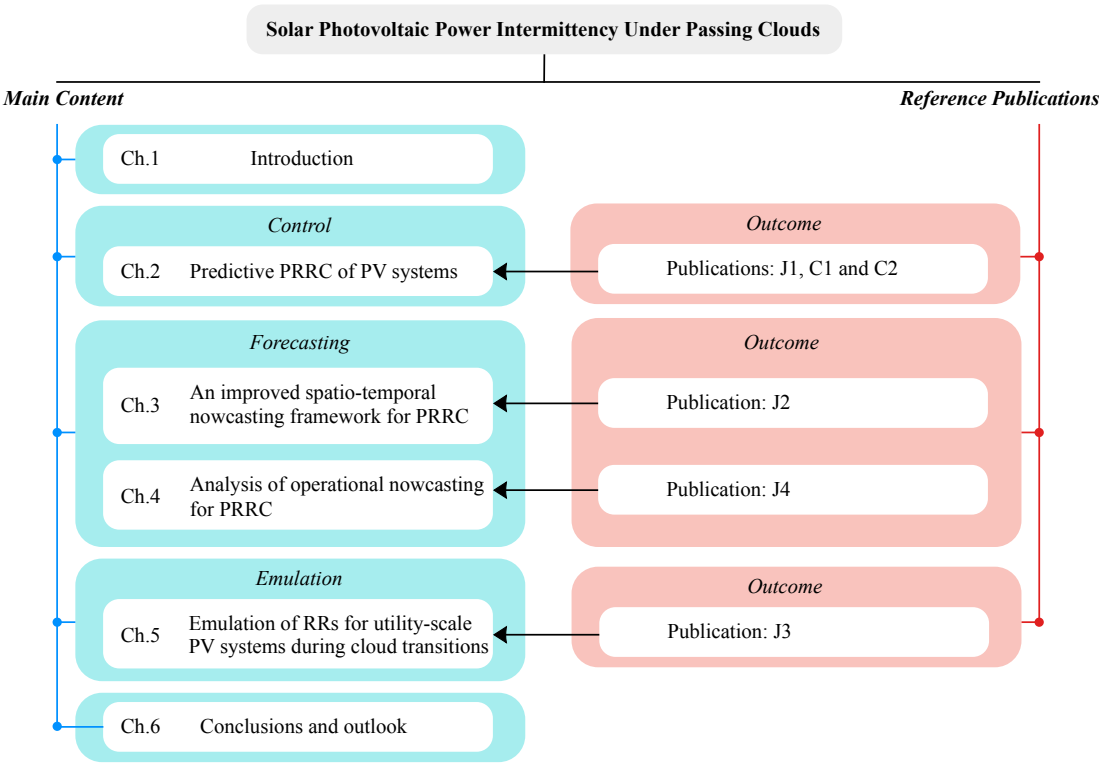


Figure 1.5: Thesis structure and related publications.

1.5 illustrates the structure of the thesis, providing a guideline on how the content is related to the reference publications.

This thesis is organized into six chapters. In **Chapter 1**, the background of the research topic is introduced, alongside with the discussion of motivation and objectives of this thesis. **Chapter 2** focuses on the *control* of PV power RRs, in which two predictive PRRC strategies are proposed based on the conventional ESS and APC methods. Data from an irradiance sensor network is used to produce solar nowcasts and evaluate the control performances. The following two chapters deal with the *forecasting* of PV power RRs for PRRC. In **Chapter 3**, an improved sensor network-based solar nowcasting method is developed, which considers the spatio-temporal statistics separately for stationary and ramp scenarios. The effectiveness of applying the developed nowcasting method to PRRC is also evaluated. **Chapter 4** addresses

the operability of solar nowcasting on PRRC, where the time parameters involved in operational solar nowcasting are identified. To investigate the effect of different forecast time parameters on PRRC, both deterministic and probabilistic solar nowcasting are considered. In **Chapter 5**, the *emulation* of PV power RRs under passing clouds is presented. The characteristics of RRs are analyzed for a range of utility-scale PV systems, for both endogenous and exogenous factors. To mimic the behaviors of an arbitrary partially shaded PV system, a partial shading emulator is developed, and a fully customizable shadow model that can reproduce the natures of a real cloud shadow is introduced. Finally, conclusion remarks of this thesis are summarized in **Chapter 6**, and an outlook into future research is provided.

Chapter 2

Predictive PRRC of PV systems

2.1 Chapter Introduction

The escalating integration level of intermittent solar PV resources into the power grid calls for a critical necessity to implement PRRC for modern PV systems. As discussed earlier, conventional PRRC methods mainly rely on the utilization of ESS and APC. Essentially, these approaches react to the erratic PV power fluctuations, and are termed as *reactive* approaches. Nonetheless, as the high cost of ESS is still hindering its extensive application and APC method cannot deal with ramp-down RRs, the reactive approaches still show limitations to properly address the PRRC requirements. On the other hand, the utilization of solar nowcasting enables a *predictive* approach, which initiates the ramp-down function sufficiently before the passing clouds shade the PV systems and potentially shrinks the required smoothing backup.

In this chapter, two predictive PRRC strategies will be presented. The first strategy does not require any ESS. For ramp-down events, PV power can be curtailed before the actual shading occurs. The second strategy requires only quarter of the energy capacity of conventional ESS control. To provide solar nowcasts, a sensor network-based dynamic spatio-temporal (DST) model is developed. The effectiveness of the nowcasting model and control strategies will be verified on a real irradiance sensor network. The economic analysis will also be provided to validate the feasibility of the proposed strategies.

2.2 Sensor Network-Based Solar Nowcasting

Cloud transients dominate solar irradiance intermittency, while the resultant effects on PV systems are primarily driven by cloud shadows on the ground level. In this section, a sensor network-based cloud shadow tracking algorithm is introduced first. Then, by integrating the derived CSMV information into the classical spatio-temporal model, the DST nowcasting model is formulated, which tunes the regression coefficients in accordance with real-time CSMV measurements.

2.2.1 Correlation-Based CSMV Detection

In this work, a semicircle sensor network configuration introduced in [63] is considered. By sharing a same central sensor, N sensors are grouped into $N-1$ sensor pairs. An exemplified network deployment with $N = 6$ is shown in Figure 2.1, where d_i, $i = 1, 2, \cdots, 6$, denotes the distance between sensor S_i and central sensor S_0. It should be noticed that to ensure the network to effectively sense the cloud shadow movements, the distance between adjacent sensors should lie within the typical cloud dimension range, e.g., 100 - 1000 m [64].

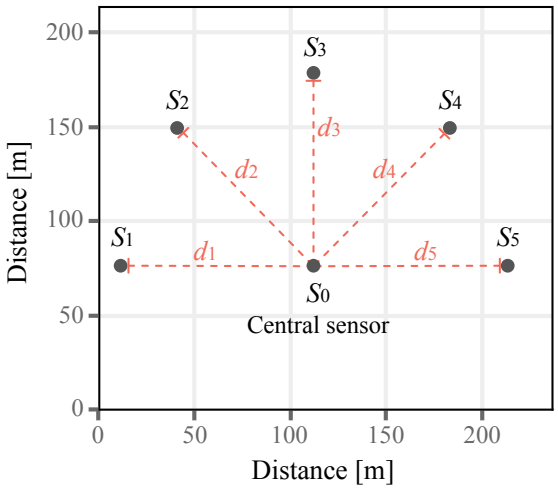


Figure 2.1: An example of semicircle sensor network configuration, shading the same central sensor S_0. The red dashed line indicates the distance between the outer and central sensors.

In order to detect CSMVs, the most-correlated pair (MCP) algorithm previously developed in [65] is performed on the sensor network. The basic principle of MCP is that the outputs of two sensors can be highly correlated with a time lag when they are aligned with the moving direction of CSMVs. To identify the time lag, the cross-correlation coefficient (CCF) is calculated between the central sensor and all other sensors. The sensor pair with the largest CCF is then selected as the MCP, and CSMV can be derived from the time lag and spatial distance between the MCP. For instance, in the context of Figure 2.1, the output time-series of the sensor S_i, $i = 1, 2, \cdots, 6$, and central sensor S_0 can be written as:

$$\boldsymbol{x}_{i,t}^{T_{obs}} = \begin{pmatrix} x_{i,t-T_{obs}} \\ \vdots \\ x_{i,t-2T_s} \\ x_{i,t-T_s} \end{pmatrix}, \ \boldsymbol{y}_{t}^{T_{obs}} = \begin{pmatrix} y_{t-T_{obs}} \\ \vdots \\ y_{t-2T_s} \\ y_{t-T_s} \end{pmatrix}, \tag{2.1}$$

where $\boldsymbol{y}_t^{T_{obs}}$ and $\boldsymbol{x}_{i,t}^{T_{obs}}$ respectively denote the time-series (typically GHI or luminance) of central sensor S_0 and other sensors over an observation window T_{obs}, and T_s is the sensor sampling time. To indicate the MCP, the sensor series $\boldsymbol{x}_{i,t}^{T_{obs}}$ are shifted in time, and the CCF is calculated with an increasing time shift δt, given by:

$$\rho_{i,\delta t} = \frac{Cov(\boldsymbol{x}_{i,t+\delta t}^{T_{obs}}, \boldsymbol{y}_t^{T_{obs}})}{\sqrt{Var(\boldsymbol{x}_{i,t+\delta t}^{T_{obs}})}\sqrt{Var(\boldsymbol{y}_t^{T_{obs}})}}, \tag{2.2}$$

where $\rho_{i,\delta t}$ is the CCF calculated between sensors S_i and S_0, $Cov(\cdot)$ and $Var(\cdot)$ indicate the covariance function and variance function respectively. The sensor pair with the largest $\rho_{i,\delta t}$ is then selected as the MCP, and the corresponding δt indicates the time lag between the MCP. Assuming sensors S_1 and S_0 are identified as the MCP, the CSMV then can be found as:

$$\vec{V} = \frac{\vec{d_1}}{\delta t}. \tag{2.3}$$

2.2.2 A Complete Network Deployment

In practice, various cloud shadows can approach PV systems in different directions. To cover the omnidirectional shadow movements, multiple sub-networks can be deployed around the target PV system to form a complete monitoring network, as shown in Figure 2.2. Once a cloud shadow is approaching the PV system, a large drop can be observed from the sensor readings, then the MCP algorithm will be activated to capture the CSMV.

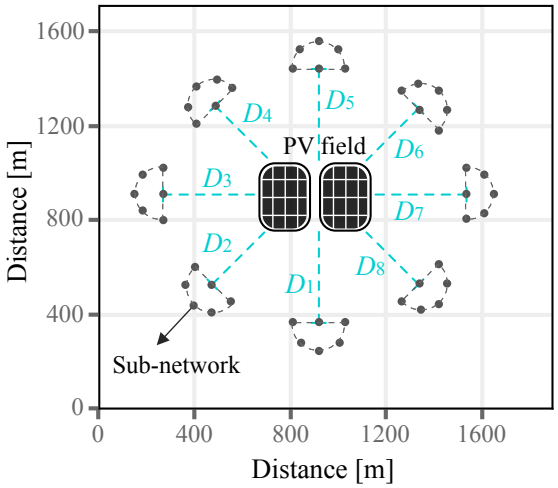


Figure 2.2: An example of a complete monitoring network deployment, consisting of 8 sub-networks. The blue dashed line indicates the distance between the sub-network and the target PV system.

Given the assumption of a persistent cloud shadow movement, if n sub-networks are activated for MCP, the timing for the cloud shadow traveling from the ith sub-network to the focal PV field can be estimated as :

$$\Delta t_i = \frac{D_i}{\vec{V_i}}, \qquad i = 1, 2, \cdots, n, \tag{2.4}$$

where Δt_i is the estimated cloud shadow traveling time from the ith sub-network to the target PV system, $\vec{V_i}$ is the derived CSMV by ith sub-network, and D_i is the distance between the ith sub-network and target PV system. Typically, the distance D can vary from a few hundred meters to several kilometers according to the local cloud velocity

and required RR limit. In [66], the authors have suggested a maximum spatial distance to be ∼10 km to maintain a sufficient correlation between sensors.

2.2.3 DST Nowcasting Model

Based on the derived CSMV information, a spatio-temporal model can be deployed to make nowcasts for the target PV system. Given data from spatial predictors $(\boldsymbol{x}_{1,t}, \boldsymbol{x}_{2,t}, \cdots, \boldsymbol{x}_{n_s,t})$ and corresponding temporal lags $(\tau_1, \tau_2, \cdots, \tau_{n_s})$, where $\boldsymbol{x}_{i,t} = (x_{i,1}, x_{i,2}, \cdots, x_{i,t})^\top$, $i = 1, 2, \cdots, n_s$, is the data from the ith spatial predictor, t is the time instant, and n_s is the number of spatial predictors, a spatio-temporal predictor set can be formulated as:

$$\boldsymbol{X}_t = \begin{pmatrix} {\boldsymbol{x}_{1,t-\tau_1}}^\top \\ {\boldsymbol{x}_{2,t-\tau_2}}^\top \\ \vdots \\ {\boldsymbol{x}_{n_s,t-\tau_{n_s}}}^\top \end{pmatrix} = \begin{pmatrix} x_{1,1-\tau_1} & x_{1,2-\tau_1} & \cdots & x_{1,t-\tau_1} \\ x_{2,1-\tau_2} & x_{2,2-\tau_2} & \cdots & x_{2,t-\tau_2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n_s,1-\tau_{n_s}} & x_{n_s,2-\tau_{n_s}} & \cdots & x_{n_s,t-\tau_{n_s}} \end{pmatrix}, \tag{2.5}$$

and a spatio-temporal model takes the form of linear regression:

$$\boldsymbol{y}_t = \boldsymbol{X}_t \boldsymbol{\beta}, \tag{2.6}$$

where $\boldsymbol{y}_t = (y_1, y_2, \cdots, y_t)^\top$ is the response, $\boldsymbol{\beta} = (\beta_0, \beta_1, \cdots, \beta_p)$ is the regression coefficient, and $p = n_s \cdot t$ is the total number of spatio-temporal predictors. In the case of sensor network-based nowcasting, the model response is the normalized output data of the target PV system, the spatial predictors are the normalized data collected by *all* the sensors in the activated sub-networks, and the corresponding temporal lags are obtained from the cloud shadow traveling time computed by Equation (2.4).

However, one issue with the model in Equation (2.6) is the inclusion of abundant irrelevant predictors, i.e., the unshaded sensors. In most cases, these irrelevant predictors only contribute to model variance instead of accuracy [67]. In order to exclude the irrelevant predictors, we herein introduce a tuning parameter into the model, which is

adjustable according to the real-time data collected by the sensors. At this stage, the DST model is developed, given by:

$$\boldsymbol{y}_t = \boldsymbol{X}_t(\boldsymbol{\beta} + \boldsymbol{\Lambda}), \tag{2.7}$$

where $\boldsymbol{\Lambda} = (\lambda_0, \lambda_1, \cdots, \lambda_p)^\top$ is the tuning parameter, and

$$\lambda_i = \begin{cases} \sum_{j=1}^{p-m} \beta_j^* \cdot \dfrac{\rho_i}{\sum_{i=1}^{m} \rho_i} & \text{if } \rho_i \geq 0.5 \\ -\beta_i & \text{if } \rho_i < 0.5, \end{cases} \qquad i = 0, 1, \cdots, p, \tag{2.8}$$

where λ_i is the tuning parameter for the ith regression coefficient, i.e., β_i, m is the number of relevant predictors,[1] β_j^* is the regression coefficient of the jth irrelevant predictor, and ρ_i denotes the CCF calculated with the corresponding central sensor for the ith relevant predictor. The tuning parameter ensures that the weak correlated sensors (with CCF smaller than 0.5) are excluded from the model, and the more relevant predictors can be assigned with larger weights.

2.3 Predictive PRRC Strategies

In this section, two predictive PRRC strategies are developed. The first strategy improves the conventional APC method to make it practical for power drop conditions, and no ESS is required. The second strategy aims to use the minimum ESS capacity for control.

2.3.1 Strategy 1: Predictive APC Control

As mentioned earlier, the conventional APC method cannot cope with the downward RR conditions since there is no extra energy available to mitigate the sudden power loss. Nonetheless, as the utilization of solar nowcasting can provide additional time buffer,

[1] In this work, if the CCF between a sensor S_i and the central sensor is greater than 0.5 (referred as medium correlated in statistics), the sensor S_i is considered as a relevant spatial predictor. In this regard, m = number of relevant spatial predictors $\times$ corresponding temporal lags.

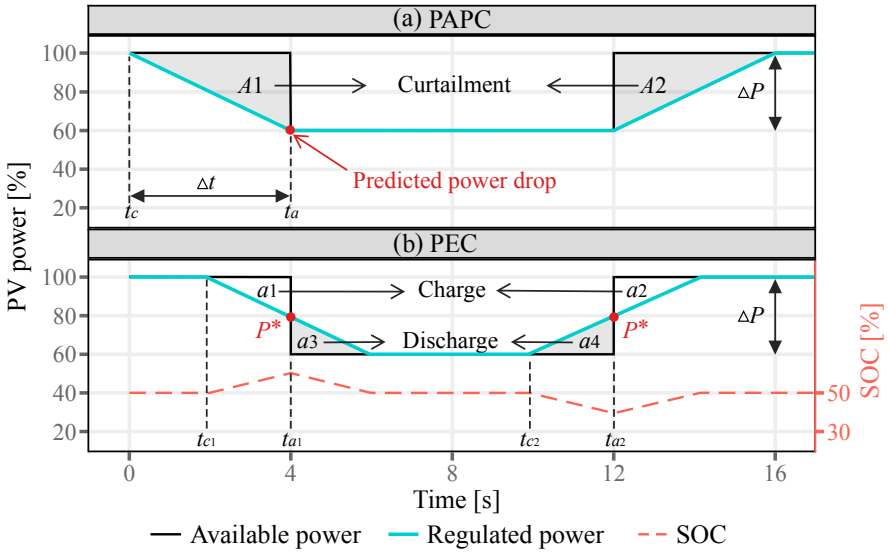


Figure 2.3: Control principle of the proposed (a) PAPC method, and (b) PEC method, respectively.

it may offer an opportunity to proactively regulate the upcoming RRs. Figure 2.3(a) illustrates the control principle of the predictive APC (PAPC) approach.

For the upward RRs, PAPC works similar to the conventional APC method, where the RRs are directly smoothed on the PV inverter level by FPPT (area A_2). For the downward RRs, the integration of solar nowcasting allows to predict the CSMV arrival time t_a and the resultant PV power change ΔP.[2] Given a predefined RR limit of R_s and system rated power of P_{rated}, the proactive control time t_c can be found by:

$$\frac{\Delta P}{(t_a - t_c) \cdot P_{\text{rated}}} = R_s, \tag{2.9}$$

where $t_a - t_c$ defines the estimated control time buffer, i.e., Δt, acquired from solar nowcasting. In this way, the ramp-down curtailment (area A_2) can be implemented prior to the occurrence of the actual power drop. As a result, both upward and downward RR violations can be mitigated solely based on APC, and no ESS is involved.

[2]It should be noticed that the forecast PV power change is a length-h/r time-series, where h and r denote the forecast horizon and forecast resolution respectively. In this case, PAPC outputs its control response at $t = 0$ for a complete timestamps from $t = 1$ to $t = h/r$, which is otherwise unachievable using the "one-step-ahead" control algorithm such as greedy control. More on this in Section 4

2.3.2 Strategy 2: Predictive ESS Control

Recall the conventional ESS control method displayed in Figure 1.3(a), the ESS storage requirement should be determined from the worst case that the PV generation increases or decreases dramatically at the rated operation mode ($\Delta P = P_{\text{rated}}$) within a very short period. In this regard, the required ESS power capacity should equal to the rated power of the PV system. Moreover, in order to handle both positive and negative fluctuations, a doubled ESS energy capacity is typically demanded. Given above, the storage requirement of conventional ESS method can be summarized as:

$$\begin{cases} P_{\text{req}}^{\text{ESS}} = P_{\text{rated}} \\ C_{\text{req}}^{\text{ESS}} = 2\int_t P_{\text{rated}} \end{cases}, \tag{2.10}$$

where $P_{\text{req}}^{\text{ESS}}$ and $C_{\text{req}}^{\text{ESS}}$ denote the ESS power and energy requirement for conventional ESS control, respectively.

To reduce ESS storage requirement, a predictive ESS control (PEC) strategy is herein proposed, whose principle is shown in Figure 2.3(b). In the case of PEC, solar nowcasting is used to predict the CSMV arrival time t_{a_1}, CSMV leaving time t_{a_2}, and corresponding PV power change of ΔP. The PEC is then implemented based on capturing the midpoint P^* between the power changes. For downward RRs, instead of operating in the discharge mode consistently, ESS first charges before PV power decreases to P^* (area a_1). Then, these stored energy is in turn used for compensating the energy loss of area a_3. For upward RRs, before PV power reaches the predicted P^*, ESS first discharges to mitigate the sudden power increase of a_4. In that, ESS earns enough free space to absorb the energy of a_2 after P^*. For PEC, the proactive control time t_c can be found by:

$$\frac{1}{2} \cdot \frac{\Delta P}{\Delta t \cdot P_{\text{rated}}} = R_s, \tag{2.11}$$

where $\Delta t = t_{a_1} - t_{c_1}$ for downward RRs, and $t_{a_2} - t_{c_2}$ for upward RRs. It is worth noting that by seeking the power midpoint P^*, the total energy being absorbed or released

always keeps equilibrium, which indicates the overall energy inside the ESS remains unchanged. In this sense, for the ideal ESS or low self-discharging devices, it avoids the extra SOC control since the SOC value can be reset automatically (see the red dashed line in Figure 2.3(b)). However, as the sign of the first RR is typically unknown, to handle both downward and upward RRs, the initial SOC should be set as 50%.

Referring to Figure 2.3(b), the maximum power through the ESS is half of the power difference ΔP, and the smoothing of each RR event is separated into both charge and discharge procedures. Considering the worst RR scenarios, that is, $\Delta P = P_{\text{rated}}$, the storage requirement of PEC then can be determined as:

$$\begin{cases} P_{\text{req}}^{\text{PEC}} = \frac{1}{2} P_{\text{rated}} \\ C_{\text{req}}^{\text{PEC}} = \frac{1}{2} \int_{\frac{1}{2}\Delta t} P_{\text{rated}} \end{cases} . \tag{2.12}$$

Theoretically, compared with the storage requirement of conventional ESS control in Equation (2.10), the required ESS power capacity of using PEC is halved, and the energy capacity is reduced to the quarter.

2.3.3 Invalid Control Scenarios

In practice, control with perfect nowcasts as seen in Figure 2.3 are generally unavailable. Owing to the inevitable forecast errors, performance of the two predictive PRRC strategies can highly rely on the predictability of upcoming ramps. In that, two quantities, namely, ramp peak occurrence time and ramp peak magnitude, are especially of great importance.[3] To give perspective, Figure 2.4 illustrates the invalid control scenarios of PAPC and PEC under inaccurate predictions of ramp peak occurrence time and ramp peak magnitude, respectively.

In terms of PAPC, it is found that when the predicted ramp peak is earlier or larger than the actual condition, the control can be triggered in advance. As a result, more PV generation is curtailed or wasted (see the red line in the top plot of Figure 2.4(a)

[3]The ramp peak occurrence time is different from cloud arrival time. The former defines the occurrence timestamp of the largest RR, while the latter indicates the instant at which a CSMV is about to affect the PV system.

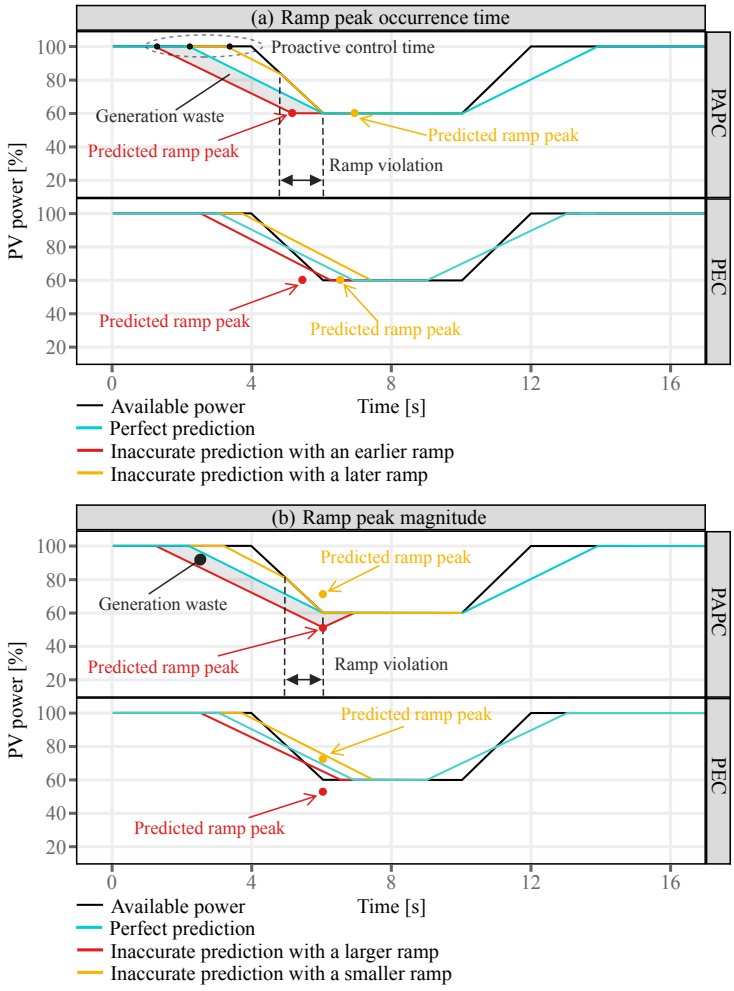


Figure 2.4: Exemplified control of downward RRs using PAPC and PEC under inaccurate predictions of (a) ramp peak occurrence time, and (b) ramp peak magnitude.

and Figure 2.4(b)). On the other hand, if the predicted ramp peak is later or smaller than the fact, the control time is delayed. In that case, an unexpected RR violation can result (see the yellow line in the top plot of Figure 2.4(a) and Figure 2.4(b)). In terms of PEC, the prediction errors in the ramp peak occurrence time and magnitude can lead to a deviation in the identified midpoint P^*. Consequently, the equilibrium between energy absorption and release is broken, and the SOC value will not be set back to 50% automatically. In addition, given the imperfect energy storage efficiency in practice, the consistent non-equilibrium of SOC can eventually lead to an outage of ESS.

2.4 Empirical Results

This section sets out the validation of the developed DST nowcasting and PRRC strategies. Data from a real irradiance sensor network are used to simulate the operation of PAPC and PEC. To evaluate the effectiveness of the proposed control strategies, two comparative case studies are demonstrated, one with scattered clouds and another with mostly clear sky conditions. The applicability of the work is further verified over a one-year period of data. In addition, the economic comparison between different PRRC strategies is provided.

2.4.1 Data

The data used in this work is obtained from Oahu solar measurement grid, an irradiance sensor network installed by National Renewable Energy Laboratory (NREL), located in Hawaii [68]. The network consists of 17 sensors, covering an area of approximately 1 km $\times$ 1.2 km, as shown in Figure 2.5. For each sensor, 1-s GHI data from March 20, 2010 to October 31, 2011 is available.

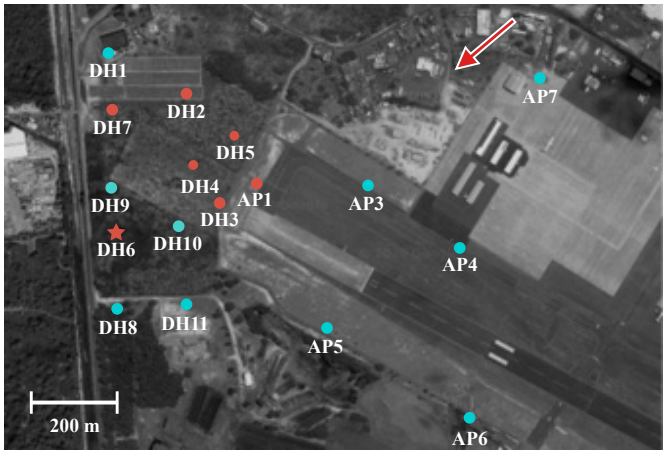


Figure 2.5: Layout of the NREL Oahu sensor network. The 7 sensors for empirical study are marked by red, where the sensor DH6 is selected as the focal system (marked by the red star). Except for sensor DH6, other 6 sensors form a semi-circle sub-network. The arrow in the top shows the wind direction. The scale of the map is shown in the bottom left corner.

Throughout this work, sensor DH6 is selected as the focal system, i.e., the target of nowcasting and control. Since solar-generated power is of interest to PRRC, the GHI measurements at DH6 need to be converted to PV power data. Due to the geographic dispersion in practical PV systems, the diversity in irradiance variabilities can be greatly relieved in PV power profiles [69]. To account for the effects of resource spreading, an irradiance-to-power conversion model introduced in [50] is adopted, which is given by:

$$P(s) = \frac{K}{(\sqrt{S}/(2\pi \cdot 0.02))s + 1} \cdot G(s), \tag{2.13}$$

where $P(s)$ and $G(s)$ respectively denote the PV power and GHI data in the frequency domain, K is the ratio of PV nominal power to the standard irradiance of 1000 W/m^2, and S is PV system area in hectares. Equation (2.13) shows that the system geographic dispersion behaves as a first-order low-pass filter to irradiance variabilities, and the variability becomes inversely proportional to the square root of PV system area. In this work, data from DH6 are used to mimic a 1.5 MW PV system in Cintruénigo, Spain, with $S = 6.4$ hectares [50].

To demonstrate the proposed nowcasting and PRRC strategies, 6 sensors surrounding with DH6 are selected to form a semicircle sub-network, namely, DH7, DH2, DH5, AP1, DH3, and DH4, among which sensor DH4 is used as the central sensor. With this specific network deployment, we herein only consider the days with prevailing trade wind direction from 0° to 90° as PRRC operating days. As a result, a total of 178 days are filtered out from the whole dataset, in which the data of the first 50 days are used for model training, while the left 128 days contribute to the validation set. After several calibration tests, the sensor reading changes for triggering MCP algorithm is set as 100 W/m^2, and the time window for CCF computation is 30 s.

It should be noticed that the cloud shadow traveling time in Equation (2.4) actually defines the forecast horizon of the sensor network. In present context, according to [70], the average distance between the sub-network and DH6 is ∼300 m (329 m between DH5 and DH6, 226 m between DH4 and DH6), and the average wind speed in the island is

10 m/s. On this point, the average forecast horizon of this network deployment can be estimated as 30 s. Given a sensor sampling time of 1 s, the largest detectable CSMV velocity reaches 30 m/s, which should cover most of the weather conditions [71].

2.4.2 Results of DST Nowcasting

To evaluate the performance of DST nowcasting, we consider two example days on July 31, 2010 and January 20, 2011, one with frequent irradiance changes and another with mostly clear sky. Prior to nowcasting, data from the 7 sensors are transformed into clear-sky index (CSI), in order to remove the diurnal trends in time-series. That said, the predictors and response in Equation (2.7) are composed of the converted CSI data. Given a sensor S_i at time t, the CSI is defined as:

$$k^*_{i,t} = \frac{x_{i,t}}{x^{\text{clr}}_{i,t}}, \tag{2.14}$$

where $k^*_{i,t}$ is the CSI for sensor S_i at time t, $x_{i,t}$ denotes the measured GHI (sub-network) or PV power (DH6), and $x^{\text{clr}}_{i,t}$ is the clear-sky expectations. In this work, the Ineichen-Perez clear-sky model is utilized to retrieve $x^{\text{clr}}_{i,t}$. After prediction, the nowcast CSI values are converted back to GHI or PV power using current clear-sky expectations.

Figure 2.6 illustrates the DST nowcasting results for the two example days. It can be seen that the predictions are only produced when large PV power RRs are observed. Moreover, the time interval between two consecutive predictions is inconsistent. Due to the various cloud shadow velocities detected by the sub-network, the forecast horizon naturally differs with different CSMVs. During the two example days, the maximum forecast horizon is found to be 1 min on 17:00 of the cloudless day, while the minimum forecast horizon is 15 s on 12:38 of the overcast day.

The overall performance of DST nowcasting on the two days is summarized in Table 2.1, where a RR limit of $R_s = 10\%/\text{min}$ is considered. It can be observed from Table 2.1 that the PV generation experiences more RR violations on the overcast day. Among the 135 RR violations, 133 RR violations are effectively captured by the DST model.

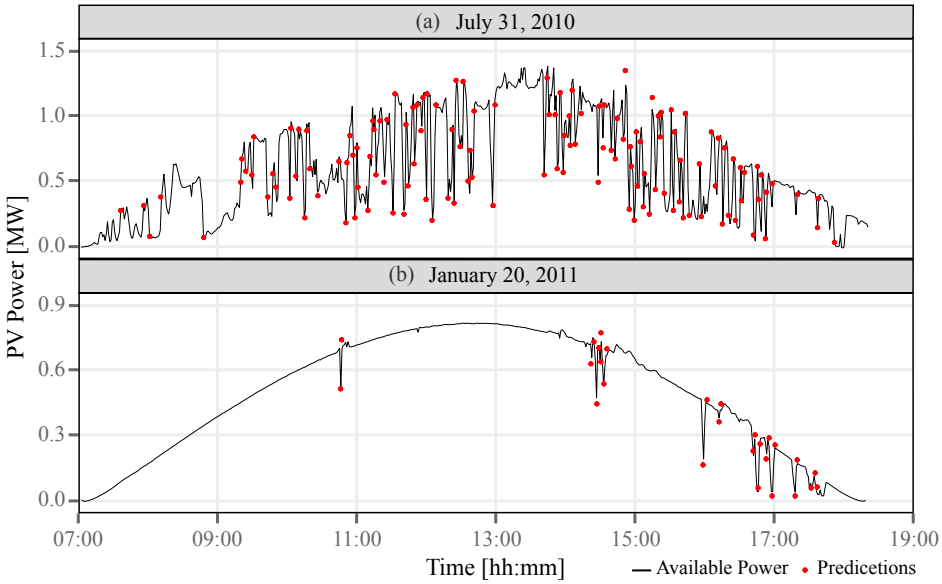


Figure 2.6: DST nowcasting results (red dots) and measured PV generation at DH6 (black line) on (a) July 31, 2010, and (b) January 20, 2011.

Table 2.1: Performance of DST nowcasting on the two example days. A RR limit of 10%/min is applied.

	Overcast day	Cloudless day
Actual RR violations	135	27
Predicted RR violations	133	27
Forecasting horizon [s]	15-30	30-60
nMAE [%]	6.5	5.5

The two missed RR violations have a RR of -10.3%/min and -10.2%/min respectively, which are very close to the RR limit. In that case, the unavoidable measurement and nowcast errors may lead to an omission on these RRs. For the cloudless day, all the 27 RR violations are successfully identified. To quantify the errors in the magnitude of measured and actual RRs, the normalized mean absolute error (nMAE) is computed, which is given by:

$$\text{nMAE} = \frac{\frac{1}{n}\sum_{i=1}^{n}|\hat{P}_i - P_i|}{\frac{1}{n}\sum_{i=1}^{n}P_i} \times 100\%, \tag{2.15}$$

where $\hat{P}_i$ and P_i denote the predicted and measured PV generation at the ith time step, respectively. For the two example days, the nowcasting nMAE on the RR magnitude is found as 6.5% and 5.5%, respectively.

2.4.3 Results of Predictive PRRC Strategies

Figure 2.7 and Figure 2.8 show the PRRC performance using PAPC and PEC strategies respectively, under the RR limit of 10%/s. In addition, a detailed comparison of different PRRC strategies is displayed in Table 2.2, in which the conventional ESS control is used as a benchmark, denoted by ESS. Referring to Equation (2.12), the simulation considers an ESS storage requirement of $P_{\text{req}}^{\text{PEC}} = 0.75$ MW and $C_{\text{req}}^{\text{PEC}} = 62.5$ kW·h for PEC method, while a setup of $C_{\text{req}}^{\text{ESS}} = 250$ kW·h and $P_{\text{req}}^{\text{ESS}} = 1.5$ MW is used for the conventional ESS control.

During the overcast day, it can be seen from Table 2.2 that PEC produces 3 less RR violations than the PAPC method. Besides the two violations caused by RR omission, another 7 and 4 RR violations have been observed for PAPC and PEC respectively, which is resulted from the invalid control scenarios illustrated in Figure 2.4. Theoretically, the

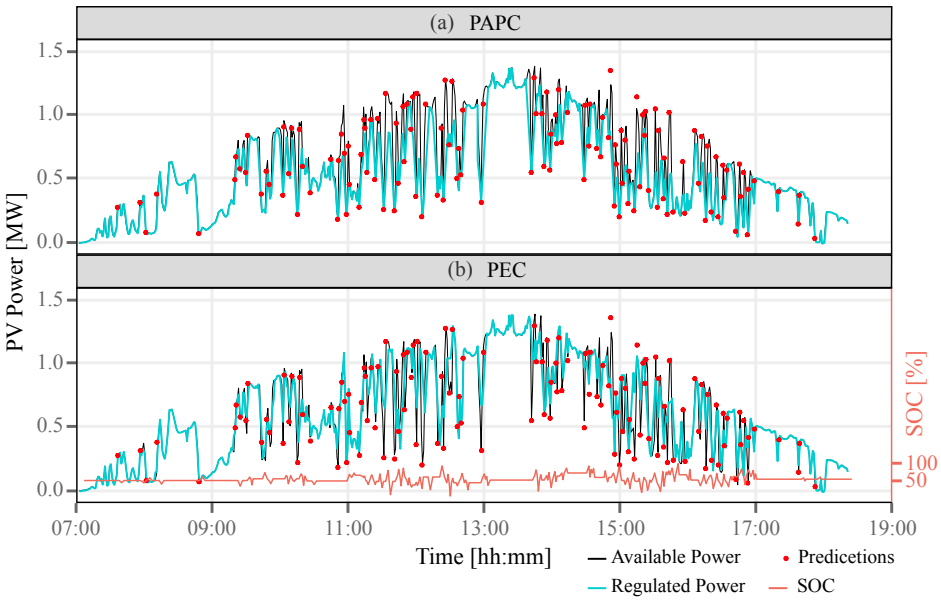


Figure 2.7: PRRC results on July 31, 2010 using (a) PAPC and (b) PEC.

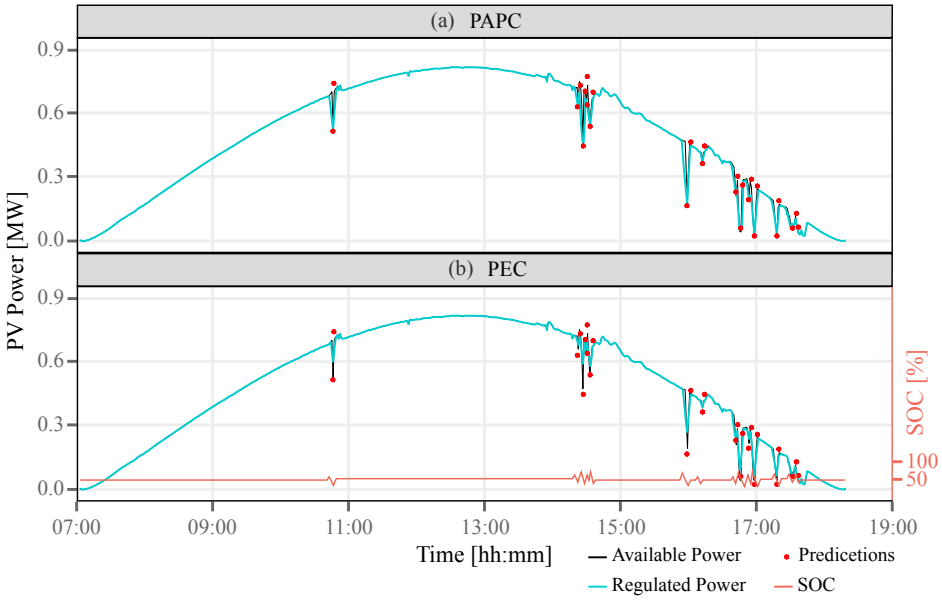


Figure 2.8: PRRC results on January 20, 2011 using (a) PAPC and (b) PEC.

Table 2.2: Comparison of PAPC, PEC, and ESS for the two example days.

	Overcast day			Cloudless day		
	PAPC	PEC	ESS	PAPC	PEC	ESS
Energy capacity [kW·h]	-	62.5	250	-	62.5	250
Power capacity [kW)]	-	0.75	1.5	-	0.75	1.5
Charge-discharge times	-	266	135	-	54	27
Final SOC [%]	-	53	50	-	50	50
RR violations	9	6	0	1	0	0
Energy curtailment [%]	12.4	-	-	0.91	-	-

ESS charge-discharge times of PEC should equal to 270, twice than the conventional ESS control. However, due to the unexpected RR violations, the charge-discharge cycle of PEC reduces to 266, which has led to the SOC deviation from 50% to 53% at the end of the operating day. For PAPC, the total energy being curtailed during the day is 886.5 kW·h, which equal to 12.4% of the total generation.

During the cloudless day, all the ramp events are effectively captured. The only RR violation occurs at around 17:00 by using PAPC, where the predicted RR occurrence time is 10% larger and 20 s later than the actual condition, leading to a 30-s RR violation.

Table 2.3: Daily performance of DST model and proposed PRRC strategies over a one-year period.

	PAPC (Dry)			PEC (Dry)			PAPC (Rainy)			PEC (Rainy)		
	Max	Min	Avg	Max	Min	Avg	Max	Min	Avg	Max	Min	Avg
Nowcast nMAE [%]	8.5	0.26	5.8	8.5	0.26	5.8	12.3	2.23	6.6	12.3	2.23	6.6
Ramp omission	6	0	2	6	0	2	10	0	5	10	0	5
Control failures	10	2	6	8	0	4	12	3	8	10	0	5
Energy curtailment [%]	14.2	0.26	6.6	-	-	-	13.5	0.22	5.3	-	-	-
Charge-discharge times	-	-	-	272	30	144	-	-	-	236	24	132
Final SOC [%]	-	-	-	59	44	53	-	-	-	56	48	52

For PEC, the daily charge-discharge times are twice than the conventional ESS method as expected, and the final SOC is set back to 50% successfully. The energy curtailment of PAPC in this day is 66.2 kW·h, accounting for 0.91% of the total generation.

2.4.4 Validation Over One-Year Observations

To validate the applicability of the work, the DST nowcasting model and PRRC strategies are implemented over a full year of data, i.e., 128 days of validation set. The 128 sample days are classified into two groups according to the climate types, namely, 70 days from May to September as the dry season, and 58 days from October to April as the rainy season. The simulation results are shown in Table 2.3.

In terms of DST nowcasting, it can be observed from Table 2.3 that the nowcast accuracy generally decreases in the rainy season, with an average 0.8% increase in nMAE compared with the dry season. As the cloud shadow effect diminishes during the precipitation, the CSMV can become more difficult to detect. In addition, the meteorological variables such as ambient temperature or relative humidity can also change frequently in the rainy season. Under these circumstances, the temporal correlations among the sensors reduce significantly, thus the RRs can become more challenging to predict. Despite the low RR predictability in the rainy season, the overall performance of DST nowcasting is quite promising. Over the one-year observations, 95.8% (9810 of 10,240) RR violations are successfully identified by the DST model, and the annual nMAE comes to 6.2%. These results validate the effectiveness of the developed DST

nowcasting method, which ensures the functionality of the following predictive PRRC strategies.

In terms of PRRC, in the dry season, the higher irradiance intensity and more ramp events have resulted in more energy curtailment for PAPC. The use of PEC also requires the ESS operation mode to switch more frequently, with 12 time more per day on average. According to Table 2.3, the PEC method is found to generally outperform the PAPC in terms of control failures, with two and three fewer daily RR violations on average than PAPC in the dry and rainy seasons respectively. However, as illustrated in Figure 2.4, the invalid control scenarios can break the SOC equilibrium for PEC, which may strain the ESS in the long run. On this point, the average daily final SOC of PEC method is found as 53% after 128 operating days, which can still be acceptable for long-term consistent operations.

2.4.5 Economic Analysis

In this section, the feasibility of the proposed PRRC strategies are evaluated from the economic aspect. In terms of PEC, the integration of an ESS can increase the overall investment of a PV system. On the other hand, even though the ESS is not required for PAPC, the power generation is often suppressed, leading to production waste. Besides, the deployment and maintenance of a nowcasting system also brings extra expenditure for system operators. Hence, the economic value of the proposed PRRC strategies requires further investigation.

Table 2.4 shows the $/W cost of a utility-scale PV system concluded in [72]. Referring to Table 2.4, the overall installation cost of a 1.5 MW PV system should equal to

Table 2.4: Cost of PV installation in $/W.

Component	Cost [$/W]
Module	0.64
Inverter	0.09
Structure and Electrical components	0.26
Other	0.50
Total	1.49

$C_{\text{pv}} = \$2,235,000$. In terms of the nowcasting system, we estimate the cost of a sensor sub-network as \$150. For a 1.5 MW PV system, 20 sub-networks should be enough to enclose the whole system, leading to an installation cost of \$3000. With additional considerations of sensor maintenance and contingencies, the operating cost of a sensor network nowcasting system is estimated as \$0.05/W. On this basis, the total investment of a 1.5 MW nowcasting-integrated PV system can be found as $(1.49 + 0.05)$ \$/W $\times$ 1.5 MW + \$3000 = \$2,313,000. Assuming an annual average energy generation of E kW·h for a 1.5 MW PV system and a PV panel lifetime of Y years, the economic analysis will be conducted under the following PRRC operating modes: 1) PAPC method, 2) PEC method, 3) conventional ESS control, and 4) no reaction.

1) PAPC method: In [73], it has been shown that the RR limit of 10%/min can be frequently violated, i.e., more than 100 times, in a cloudy day. Consequently, significant energy production can be curtailed by using PAPC. Considering an annual average energy curtailment due to PAPC to be E_c kW·h, the PV generation cost under PAPC operating mode can be found as:

$$\$_{\text{papc}} = \frac{(E - E_c) \cdot Y}{C_{\text{pv}} + C_{\text{nowcast}}}, \tag{2.16}$$

where C_{nowcast} denotes the cost of deploying a nowcasting system. In this case, $C_{\text{nowcast}} = 0.05$ \$/W $\times$ 1.5 MW + \$3000 = \$78,000.

2) PEC method: To involve the cost of an energy storage device, the lead-acid battery is herein considered, which is perhaps the most commonly-used ESS type for large-scale industrial application [74]. The overall cost of integrating a lead-acid battery can be calculated as:

$$\$_{\text{la}} = \frac{C_{\text{la}} \cdot I \cdot N}{\eta}, \tag{2.17}$$

where C_{la} is the lead-acid battery cost in \$/kW·h (including capital and operating cost), I is the installed battery capacity in kW·h, N is the total battery replacements during PV lifetime Y years, and η is the battery efficiency. In a general sense, given an

average battery life of y years, $N = Y/y$. However, as PEC method doubles the ESS charge-discharge cycles, the lifespan of a battery is halved. As a result, the battery replacement can be doubled, which equals to $2Y/y$. On the basis of Equation (2.17), the PV generation cost using PEC operating mode is given by:

$$\$_{\text{pec}} = \frac{E \cdot Y}{C_{\text{pv}} + C_{\text{nowcast}} + \$_{\text{la}}}. \tag{2.18}$$

3) Conventional ESS control: The generation cost of conventional ESS method can also be found based on Equation (2.17), but with different battery replacements, i.e., Y/y, and battery capacity, denoted by N' and I' respectively. Given a resultant battery cost of $\$'_{\text{la}}$, the PV generation cost is obtained by:

$$\$ = \frac{E \cdot Y}{C_{\text{pv}} + \$'_{\text{la}}}. \tag{2.19}$$

4) No reaction: In this operating mode, no reaction is performed during ramp events. Although the overall investment is reduced (no need for ESS and nowcasting system), the PV system owners could be penalized by grid operators due to frequent RR violations. In this sense, a weekly PV generation model considering RR violation utilized [75], which is given by:

$$\begin{cases} Q_{\text{week}} = Q_{\text{week}-1} - (RRC_{\text{week}-1} - RRC_{\text{limit}}) \\ P_{\text{week}} = P_{\text{rated}} \cdot Q_{\text{week}} \end{cases}, \tag{2.20}$$

where Q_{week} and $Q_{\text{week}-1}$ are penalty factors which regulate the maximum PV generation of current and previous weeks, $RRC_{\text{week}-1}$ denotes the total number of RR violations in percentage during last week, RRC_{limit} is the predefined RR violation limit in percentage, and P_{week} is the weekly maximum allowable PV generation. Based on Equation (2.20), given a resultant annual energy generation of e kW·h, the PV generation cost is then

Table 2.5: Parameters values for the economic analysis of a 1.5 MW PV system.

E [kW·h]	Y [years]	y [years]	E_c [kW·h]	$\$_{\text{la}}$ [$/kW·h]	I [kW·h]	N	η [%]	I' [kW·h]	N'	e [kW·h]
2,447,781	24	2	198,490	136.5	62.5	24	70	250	12	2,150,000

Table 2.6: Financial comparison among the four PRRC operating modes.

Operating mode	Annual generation [kW·h]	Overall generation (kW·h)	Additional cost ($)	Total PV generation cost ($)	Cost in $/kW·h
1	2,249,291	53,982,984	78,000	2,313,000	0.042
2	2,447,781	58,746,744	370,500	2,605,500	0.044
3	2,447,781	58,746,744	585,000	2,820,000	0.048
4	2,150,000	51,600,000	NA	2,235,000	0.043

calculated as:

$$\$ = \frac{e \cdot Y}{C_{\text{pv}}}. \tag{2.21}$$

Table 2.5 lists the parameter values for the 1.5 MW PV system based on former simulation results and [72]. It should be noticed that different parameter values can significantly influence the final PV generation cost. Besides, PV site location. local market regulations, and selection of RR limit can also vary from case to case. The analysis herein presented is meant to be illustrative and give a broad comparison among the four options.

Table 2.6 displays the PV generation cost of the four PRRC modes. It can be seen that the use of PAPC method shows the least PV generation cost of $0.042/kW·h. Compared with mode 4, the use of PAPC even produces 5% more of total energy generation. In terms of PEC, its generation cost is found $0.042/kW·h more than the PAPC method. However, the use of ESS improves the PV system flexibility, which enables the PV system to participate in grid auxiliary services, thus making extra revenues [76]. Moreover, given the low occurrence of the worst RR scenarios, there are possibilities to shrink the required ESS size further [27]. In this regard, the generation

cost of PEC method could become more competitive. In general, both PAPC and PEC strategies have presented to be more economical for PRRC than conventional ESS control. Although mode 4 shows a similar generation cost to the PEC method, the potential risk of RR violations would prohibit its large-scale application. Again, it is worth noting that the selection between the PAPC and PEC methods is highly dependent on the local conditions. By adapting different parameters into Equations (2.16)-(2.21), the final choice can be made accordingly.

2.5 Chapter Conclusion

The chapter presents two PV PRRC strategies based on the newly developed DST nowcasting method. The proposed PRRC methods take action in advance according to the predicted cloud shadow arrival time and the resultant output power. The PAPC method is approved to be effective to handle the ramp-down RR scenarios, which is difficult for conventional APC methods. The output power can be effectively smoothed without using any ESS. The second strategy PEC enhances the system performance with minimal support from energy storage. The method shows the advantage by neglecting the dedicated SOC control algorithm. In comparison with the conventional solution, the proposed PEC strategy halves the required ESS power capacity, and the energy capacity size is almost reduced to the quarter. The performance of the developed nowcasting model and PRRC strategies are evaluated on a real irradiance sensor network, by simulating the power output of a 1.5 MW solar plant. The results show that the DST nowcasting model can effectively capture the RR events, with an average nMAE smaller than 7%. Furthermore, both of the proposed PRRC methods are proved to be capable of smoothing the PV RRs and outperform the conventional method. The selection criteria for the two proposed PRRC strategies has also been provided.

Chapter 3

An Improved Spatio-Temporal Nowcasting Framework for PRRC

3.1 Chapter Introduction

In previous chapter, a DST nowcasting model has been developed for predictive PRRC implementations. Whereas the model can effectively capture the ramp events, its forecast horizon varies with different CSMV velocities, which is undesirable from a grid operator's perspective. Moreover, when a stricter RR limit is applied, e.g., in timescales of seconds, the MCP algorithm may fail to identify the most relevant predictors since a shorter time window would be used for calculating CCF. In that case, the nowcasting accuracy could deteriorate, leading to more control failures. In view of the above, this chapter aims to develop an improved sensor network-based spatio-temporal nowcasting method for PRRC, which can *preselect* the spatio-temporal predictors more efficiently as well as providing *consistent* solar nowcasts. However, before moving forward, we digress and take a revisit to sensor network-based solar nowcasting, to facilitate the understanding of subsequent materials.

3.1.1 A Revisit to Sensor Network-Based Solar Nowcasting

As mentioned earlier, nowcasting using a sensor network relies on the construction of spatio-temporal predictors. When clouds propagate over a sensor network, one can preselect sufficient lagged time-series data (*temporal predictors*) collected by the neighboring sensors (*spatial predictors*) as predictors for the focal location. Depending on how the cloud information is coupled, the sensor network-based nowcasting can be subdivided into two categories: cloud tracking-based methods and spatio-temporal correlation-based methods.

The cloud tracking-based methods aim to describe the cloud information with CSMVs. Solar nowcasts are then produced by transposing the generated power map in the direction of CSMV. Various CSMV tracking strategies have been introduced in the literature, such as the MCP method [77–79], or peak matching method [80]. The main advantage of using the cloud tracking-based method is that it provides an accurate measure of the time lag between the measured data at the sensors and focal PV system. In another word, the temporal predictors can be adequately preselected. However, these methods often produce nowcasts of variable prediction horizons (such as the DST model), which is limited by the network dimension and CSMV velocity. Although a peer-to-peer method is proposed in [46] to provide consistent solar nowcasts with a fixed prediction horizon, numerous PV references (202 rooftop systems) over a wide geographic dispersion ($\sim$1400 km^2) are required. Furthermore, a simple persistence model or multivariate regression model is typically used in these methods, which often includes insufficient or irrelevant spatial predictors. Consequently, the model becomes inferior to adapt various CSMVs.

Instead of explicitly deriving the CSMV, spatio-temporal correlation-based methods consider the cloud information indirectly. Solar nowcasts are generated by exploiting the spatio-temporal correlations observed among the predictors. However, the number of predictors can become very large as the number of sensors in the network increases. In this situation, several works have implemented the regularized model such as the lasso (least absolute shrinkage and selection operator) for parameter shrinkage [67, 81, 82],

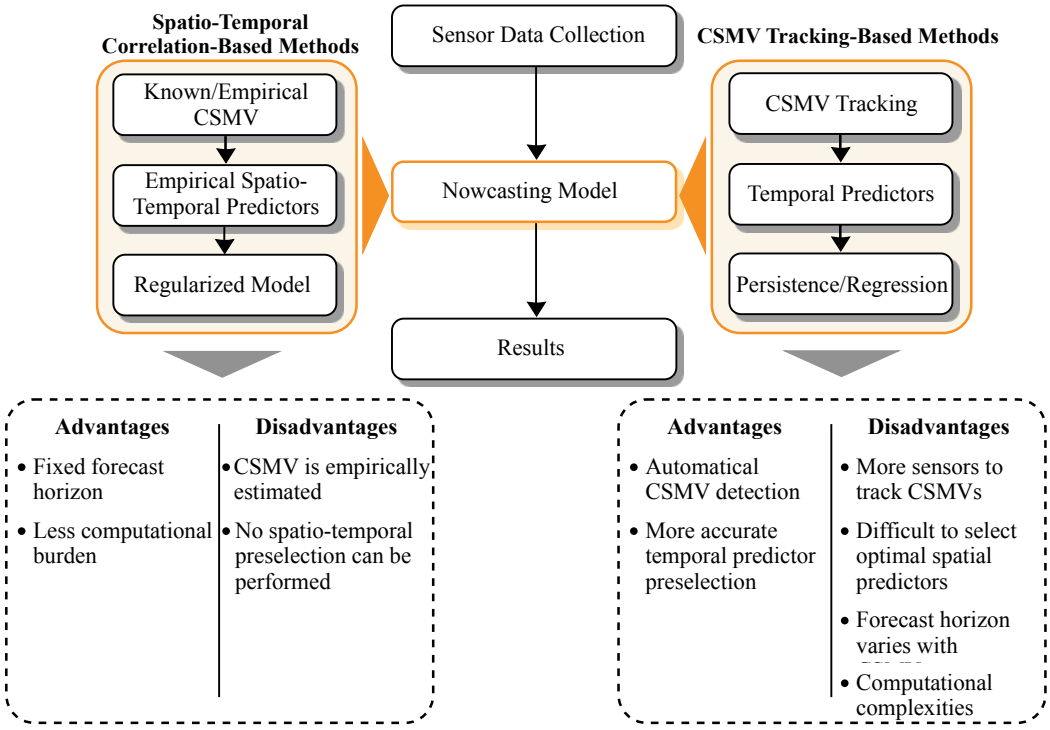


Figure 3.1: Summary of conventional sensor network-based nowcasting methods. Two categories can be identified according to the way of extracting cloud information.

which is able to filter out the highly correlated predictors. The prediction horizon using these methods can be set as a fixed value, which equals to the resolution of training data. However, the methods listed above only consider the homogeneous CSMV impacts through both space and time, and depend on known or empirically estimated CSMV movements. When the CSMV changes frequently, the model may fail to timely adapt the cloud dynamics, leading to the inclusion of both irrelevant spatial and temporal predictors. As a result, the nowcasting performance becomes much worse [83].

Figure 3.1 demonstrates the flowchart of the two conventional sensor network-based nowcasting methods. As previously discussed, these methods have shown some limitations in terms of:

- Lack of a fast and comprehensive preselection mechanism for both spatial and temporal predictors.

- Difficulties to provide consistent solar nowcasts with a fixed forecast horizon while maintaining cloud dynamics.
- Poor model robustness to adapt frequent CSMV changes.

3.1.2 Contributions of Current Work

To address the above issues, a scenario-recognizable preselection (SRP) method is herein developed for sensor network-based solar nowcasting. The proposed SRP method establishes a comprehensive spatio-temporal predictor preselection framework, which enables the most-relevant predictors to be selected in absence or presence of CSMVs. Solar nowcasting with a constant prediction horizon is also achieved using a regularized model, i.e., elastic-net, and the nowcasts can be complemented by the CSMV tracking results when CSMV is detectable. The effectiveness of the SRP-based elastic-net (SRP-Enet) model is validated on a real irradiance sensor network, and a total of 5 case studies are presented for detailed evaluation. Compared with conventional methods, the proposed method significantly improves the nowcasting accuracy, with the feature that the predictions approach closer to the actual ramp peaks. The feasibility of the proposed method on PRRC application is also demonstrated.

3.2 Network Redesign

For a proper design of a sensor network, it is supposed to provide online and high sampling measurements with appropriate spatial resolution to support PV nowcasting applications. The network should also be able to capture the omnidirectional CSMVs. Additionally, it should have a flexible infrastructure and low investment in order to be coupled with various PV systems.

In this work, a concentric sensor network configuration with two cross layers is adopted. The sensor developed herein is made from a mini solar cell. From each sensor, the short-circuit current of the solar cell is measured at 1-s resolution, and mapped to GHI data through a pre-tuned conversion model. The converted GHI data is then

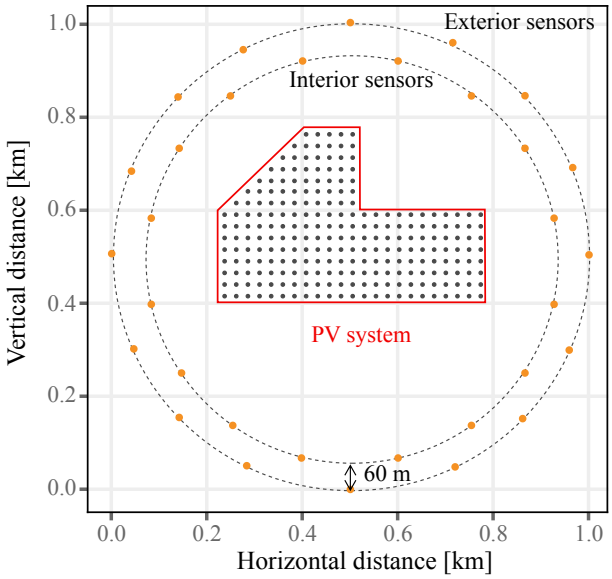


Figure 3.2: Concentric network configuration example of a 5 MW PV system, covering an area of 0.15 km^2. Totally 32 sensor are used to enclose the PV system (16 exterior sensors, and 16 interior sensors). Distance between the exterior and interior layer is around 60 m.

packed and sent to the local server via LoRa wireless communication. More details about the sensor development can be found in Appendix A. Figure 3.2 shows an example layout of the sensor network, where the focal PV system covers an area of 0.15 km^2, and rated at 5 MW. In order to reduce the packet loss rate and data transfer delays, an optimal wireless communication distance is found to be within 500 m. Similar to the network design in Figure 2.1, to prevent a single CSMV from fleeing between the sensor separations, the distance between two adjacent sensors should be less than 200 m. With the above considerations, totally 32 sensors would be deployed to enclose the whole PV system, with 16 exterior and interior sensors respectively, separated by around 60 m.

Compared with other network configurations mentioned in [46, 78, 80], the developed sensor network greatly reduces the required sensor numbers and geographical dispersion. More importantly, it is easy to be generalized and adapted to different PV systems. Last but not least, the network is flexible, and one can conveniently add or remove sensors, even a layer.

3.3 Proposed SRP-Enet Nowcasting

The flowchart of the proposed SRP-Enet nowcasting method is shown in Figure 3.3. As previously discussed, the nowcasting performance can be highly dependent on the ability to capture the spatio-temporal components of the irradiance field. In this regard, the nowcasting method proposed here decomposes the irradiance dynamics into two scenarios, namely, ramp scenarios where there are high possibilities for PV power fluctuations, and stationary scenarios where more smoothed PV generation is likely to be produced. Then, two different spatio-temporal predictor preselection approaches are implemented based on the recognized scenarios. The consistent PV nowcasts with a fixed prediction horizon are provided in stationary scenarios, and complemented by the nowcasts in ramp scenarios when CSMVs are available. A detailed description of the SRP-Enet nowcasting method is presented as following.

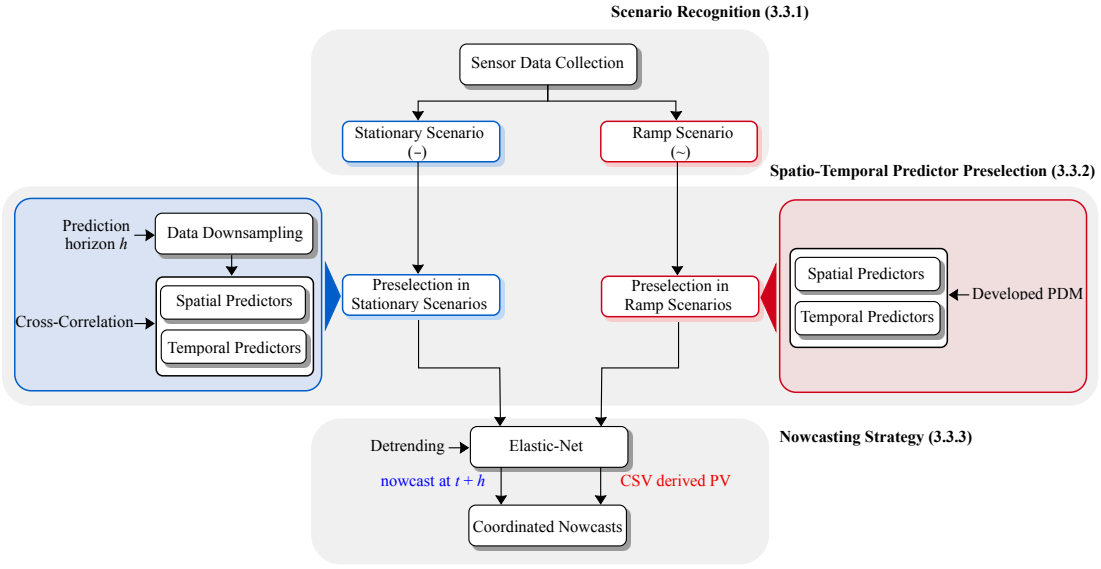


Figure 3.3: A complete flowchart of the proposed SRP-Enet nowcasting method. The SRP-Enet consists of three main blocks, namely scenario recognition, spatio-temporal predictor preselection, and nowcasting strategy. The numbers between parentheses indicate the corresponding sections for detailed description.

3.3.1 Scenario recognition

Suppose $\Phi = \{S_1, S_2, \ldots, S_N\}$ is the complete set of a sensor network, where $N = \texttt{card}(\Phi)$ is the cardinality of the set, indicating the overall sensor numbers, Φ^{ext} and Φ^{int} denote the two subsets of Φ, i.e., the exterior sensor set and interior sensor set respectively, where $N^{ext} = \texttt{card}(\Phi^{ext})$, $N^{int} = \texttt{card}(\Phi^{int})$, and $N^{ext} + N^{int} = N$. Given the instantaneous measurements of two sensors S_i, S_j at time t, i.e., $x_{i,t}$ and $x_{j,t}$, the irradiance time-series can be decomposed into two scenarios:

Stationary scenario: A stationary scenario will be identified if the absolute measurement difference between arbitrary two sensors in the network is less than a predefined threshold value σ:

$$\forall\ S_i,\ S_j \in \Phi,\ |x_{i,t} - x_{j,t}| \leq \sigma. \tag{3.1}$$

Ramp scenario: A ramp scenario will be identified if there exists a pair of exterior and interior sensors whose absolute measurement difference is greater than a predefined threshold value σ:

$$\exists\ S_i \in \Phi^{ext},\ S_j \in \Phi^{int},\ |x_{i,t} - x_{j,t}| > \sigma. \tag{3.2}$$

Based on Equations (3.1) and (3.2), the real-time sensor data stream is processed and labeled as "stationary" $\bar{x}_{i,t}$ or "ramp" $\tilde{x}_{j,t}$, where the hats "$-$" and "$\sim$" denote the stationary scenario and ramp scenario respectively. A scenario recognition signal ζ_t is then generated, given by:

$$\zeta_t = \begin{cases} 0, & \text{stationary scenario} \\ 1, & \text{ramp scenario} \end{cases}. \tag{3.3}$$

It should be noticed that a stationary scenario reveals not only a clear-sky condition, but also a sky with complete cloud coverage, under which a smooth sensor output series can be observed as well. In addition, the recognition of each scenario requires

at least two sensor measurement profiles, which enhances the reliability of recognition (the occurrence of simultaneous measurements error by multiple sensors is significantly decreased).

3.3.2 Spatio-Temporal Predictor Preselection

In ramp scenarios, due to the rapid movements of CSMVs, a fast spatio-temporal predictor preselection algorithm is demanded to cope with cloud dynamics. However, the sensor data in stationary scenarios typically present to be much smoother and less featured in fluctuations, which slows the preselection process. If the ramp scenarios share a similar preselection mechanism to stationary scenarios, the CSMV information may not be reported in time, leading to a significant delay for PRRC. Thus for the proposed SRP method, two different preselection strategies are adopted based on the former scenario recognition.

Preselection in stationary scenarios

For stationary scenarios, the 1-s resolution sensor data need to be first downsampled to h-resolution, since the consistent nowcasts with a fixed forecast horizon of h is required. Given the time instant t, the output time series of an arbitrary sensor $S_i \in \Phi$ and focal PV system can be found by:

$$\overline{X}_{i,t}^{T_{obs}} = \begin{pmatrix} \bar{x}_{i,t-T_{obs}} \\ \vdots \\ \bar{x}_{i,t-2h} \\ \bar{x}_{i,t-h} \end{pmatrix}, \ \overline{Y}_{t}^{T_{obs}} = \begin{pmatrix} \bar{y}_{t-T_{obs}} \\ \vdots \\ \bar{y}_{t-2h} \\ \bar{y}_{t-h} \end{pmatrix}, \tag{3.4}$$

where $\overline{X}_{i,t}^{T_{obs}}$ and $\overline{Y}_{t}^{T_{obs}}$ respectively denote the time series of sensor measurements and focal PV system outputs over an observation window T_{obs}.

To indicate the most-relevant spatio-temporal predictors among the sensor network, similar to Equation (2.2), the correlation coefficient, i.e., $\rho_{i,\delta t}$, is calculated between

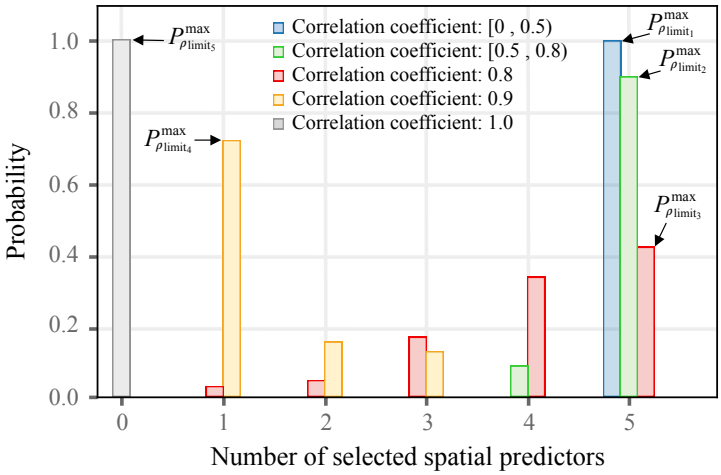


Figure 3.4: An example of determining the threshold correlation coefficient ρ_{limit} for a network with 5 sensors. In this case, sensor data collected over 6 hours are used to generate the distribution, leading to the peak probability set $\{P^{\max}_{\rho_{\text{limit}_1}}, \cdots, P^{\max}_{\rho_{\text{limit}_5}}\}$. In this case, $P^{\max}_{\rho_{\text{limit}_3}}$ is found as the minimum peak, thus $\rho_{\text{limit}} = 0.8$ is selected as the threshold.

each sensor and the target system, with an increasing time shift δt at an interval of h. Then a threshold correlation coefficient ρ_{limit} is applied, and only the sensors with $\texttt{max}(\rho_{i,\delta t}) > \rho_{\text{limit}}$ are selected as spatial predictors. Consequently, the spatial predictor preselection in stationary scenarios can be formulated as:

$$\overline{\Phi} = \{S_i\}, \quad \text{if } \texttt{max}(\rho_{i,\delta t}) > \rho_{\text{limit}}, \tag{3.5}$$

where $\overline{\Phi}$ represents the stationary spatial predictor set, and $\bar{n}_s = \texttt{card}(\overline{\Phi})$ denotes the number of spatial predictors.

Notice that the selection of ρ_{limit} varies case by case, which depends on the network configuration, local meteorological conditions, etc. Figure 3.4 shows an example of finding a proper ρ_{limit} for a network with 5 sensors. For a group of choices $\{\rho_{\text{limit}_1}, \rho_{\text{limit}_2}, \cdots, \rho_{\text{limit}_k}\}$, each element will be fed to Equation (3.5), and the corresponding probability distribution of $\bar{n}_s$ will be calculated. For each probability distribution of $\bar{n}_s$, the maximum probability value is recorded, and appended to the set $\{P^{\max}_{\rho_{\text{limit}_1}}, P^{\max}_{\rho_{\text{limit}_2}}, \cdots, P^{\max}_{\rho_{\text{limit}_k}}\}$. A proper ρ_{limit} is then indicated by the minimum value

in the set, given by:

$$\rho_{\text{limit}} = \underset{\rho_{\text{limit}_j}}{\texttt{argmin}}(\{P^{\max}_{\rho_{\text{limit}_1}}, P^{\max}_{\rho_{\text{limit}_2}}, \cdots, P^{\max}_{\rho_{\text{limit}_k}}\}), \quad j = 1, 2, \cdots, k. \tag{3.6}$$

The reason for seeking the minimum probability peak is that a smaller probability peak always contributes to a smoother distribution (as shown by the red distribution in Figure 3.4). In this regard, the high-relevant spatial predictors are more likely to be selected.

The optimized time shift for each selected spatial predictor is located where the correlation coefficient reaches the maximum, given by:

$$\overline{\gamma}_i = \underset{\delta t}{\texttt{argmax}}(\rho_{i,\delta t}), \quad S_i \in \overline{\Phi}. \tag{3.7}$$

The number of temporal predictors, $\bar{n}_t$ is then indicated by the largest $\overline{\gamma}_i$:

$$\bar{n}_t = \frac{\texttt{max}(\overline{\gamma}_i)}{h}, \quad S_i \in \overline{\Phi}. \tag{3.8}$$

An example of seeking the optimized time shift for an individual sensor is illustrated in Figure 3.5, and the procedures will be conducted for all the sensors simultaneously.

Finally, the preselected stationary spatio-temporal predictors can be written as:

$$\overline{\boldsymbol{X}}_t = \Big[\underbrace{\bar{x}_{i,t-h} \quad \bar{x}_{i,t-2h} \quad \cdots \quad \bar{x}_{i,t-\bar{n}_t h}}_{\bar{n}_t}\Big], \; S_i \in \overline{\Phi}, \tag{3.9}$$

where $\overline{\boldsymbol{X}}_t \in \mathbb{R}^{1\times(\bar{n}_s \times \bar{n}_t)}$.

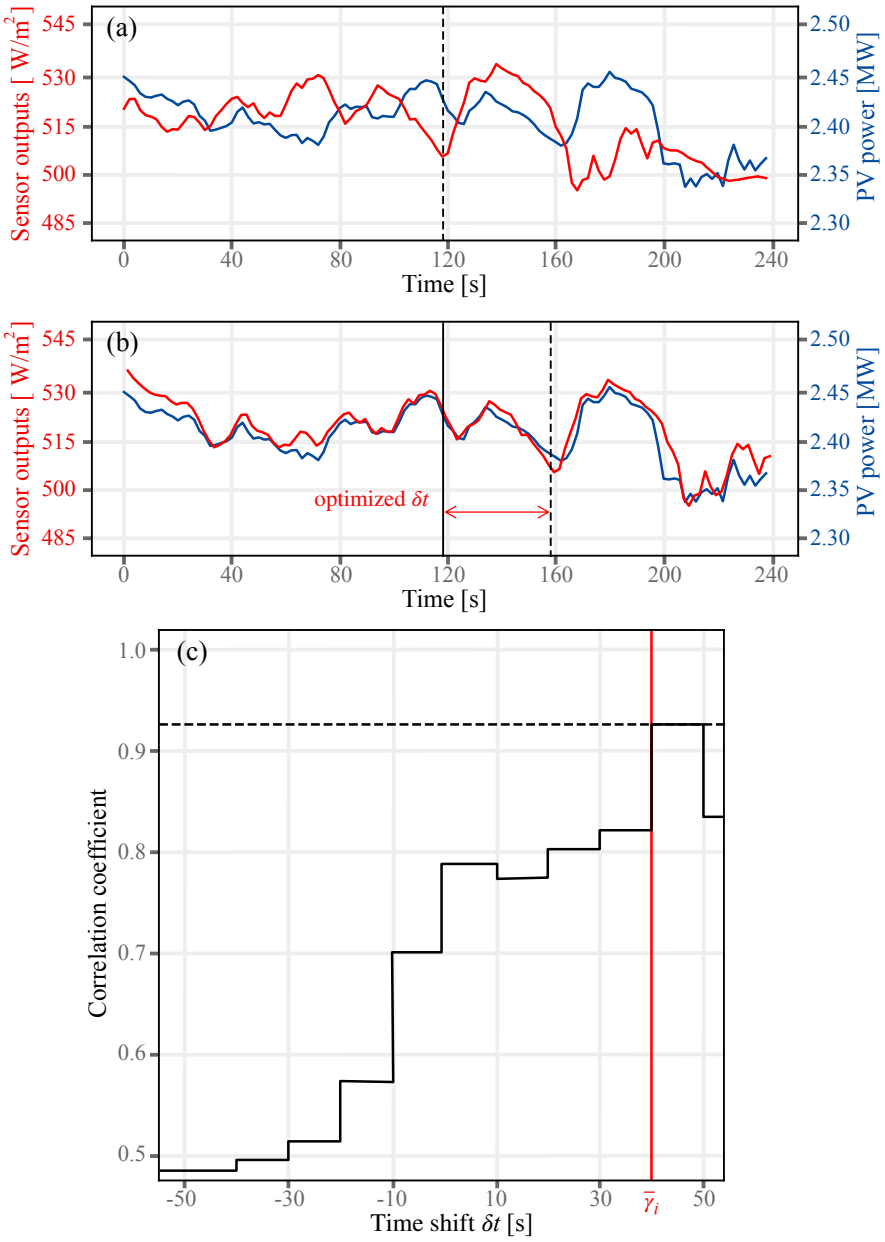


Figure 3.5: An example of seeking the optimized time shift between an individual sensor and focal PV system in a stationary scenario, with ρ_{limit}, $h = 10$ s, over an observation window T_{obs} 240 s. (a) The original time series of sensor measurements (red line) and focal system outputs (blue line). (b) The sensor measurement series is shifted with an optimized δt. (c) The correlation plot in regard to an increasing δt, and the optimized time shift $\bar{\gamma}_i$ is found as 40 s ($\bar{n}_t = 4$), when the correlation coefficient reaches the maximum, 0.927.

Preselection in ramp scenarios

In ramp scenarios, CSMV becomes the vital determinant for spatio-temporal predictor preselection. Two factors, namely cloud shadow size and cloud shadow velocity affect the spatial and temporal correlations among the sensors respectively. On one hand, the unshaded sensors under different CSMV coverage sizes can introduce irrelevant spatial predictors. On the other hand, various cloud shadow velocities can rise changeable time lags between the sensor measurements and focal system. Hence, a faster preselection algorithm capable of capturing CSMV dynamics is demanded. At this stage, the peak difference minimization (PDM) algorithm is developed, which allows an online CSMV tracking, and can update the spatio-temporal predictors adaptively.

The basic principle of PDM is that the cloud shadow transitions in ramp scenarios will show negative and positive peaks in sensor readings, and influence the exterior and interior sensors sequentially. Once the sensors with similar peaks are matched, the corresponding time lag between the sensor readings is available, and the CSMV can be derived. Notice that an assumption has been made that the CSMV remains approximately unchanged during the transition. As shown in last chapter, the traveling time for a CSMV passing through a network may merely last for a few minutes. It is therefore reasonable to think that the assumption can hold within such a short time period.

Algorithm 1 shows the implementation of PDM. A trigger signal is first defined to control the activation of the algorithm, initialized to zero. When an exterior sensor $S_i \in \Phi^{ext}$ is first observed to experience a sudden value change $\Delta\tilde{x}_{i,t_0}$ at time instant t_0, the trigger signal will be set to 1, and PDM will be activated to estimate a time buffer Δt to support PRRC. In this sense, a CSMV has to be computed before $t_0 + \Delta t$ (the estimation of Δt varies with different CSMVs and PRRC regulations). Then PDM will keep tracking other value changes among the rest of sensors. Within the predefined time interval, PDM tries to identify as many sensor value changes as possible, and all the recorded sensors are appended to a new set $\widetilde{\Phi}$, implicating the selected spatial predictors. As a result, only the sensors with sudden value changes (the shaded sensors)

Algorithm 1 Peak Difference Minimization

Input: Sensor set: Φ, Φ^{ext} and Φ^{int}, Sensor sampling time: T_s, Error tolerance: ϵ, Alignments of sensor pairs: $\vec{d}_{i,j}$.

Output: Spatial predictor set: $\widetilde{\Phi}$, Global CSMV: $\vec{V}$.

initial $\widetilde{\Phi} = \{\varnothing\}$, $trigger = 0$, $\delta t = 0$;
for $S_i \in \Phi^{ext}$ **do**
    Record the first sudden sensor reading change, $\Delta\tilde{x}^i_{t_0}$;
    $trigger = 1$;
    Estimate the required time buffer for PRRC, Δt;
    Append sensor S_i to $\widetilde{\Phi}$;
end for
if $trigger == 1$ **then**
    while $t < t_0 + \Delta t$ **do**
        for $S_j \in \{\Phi - \widetilde{\Phi}\}$ **do**
            Record another sudden reading change, $\Delta\tilde{x}^j_t$;
            Append sensor S_j to $\widetilde{\Phi}$;
            for $\delta t = 0; \delta t \leq t_0 + \Delta t - t; \delta t \mathrel{+}= T_s$ **do**
                $\tau_{i,j} = \underset{\delta t}{\texttt{argmin}}(|\Delta\tilde{x}_{i,t_0} - \Delta\tilde{x}_{j,t_0+\delta t}| < \epsilon)$;
            end for
        end for
    end while
end if
if $\widetilde{\Phi} \cap \Phi^{int} \neq \{\varnothing\}$ **then**
    for $S_j \in \{\widetilde{\Phi} - S_i\}$ **do**
        $\vec{V} = \sum \frac{\vec{d}_{i,j}}{\tau_{i,j}}$;
    end for
    return $\widetilde{\Phi}, \vec{V}$;
end if

are considered as spatial predictors, and the unshaded sensors can be excluded.

To determine the peak similarity, instead of comparing the peak values directly, PDM checks how close the value changes are. The biggest advantage of comparing the value changes is that the differencing can eliminate the inherent sensor calibration errors. A small positive value ϵ is then applied as an error tolerance to control the confidence level of peak samilarity. With an increasing time shift δt, two sensors S_i and S_j are said to be correlated at $\tau_{i,j}$ once the difference between their value changes is smaller than ϵ. Together with the known spatial distance $\vec{d}_{i,j}$, the global CSMV, $\vec{V}$ is obtained. Notice that only when the intersection of set $\widetilde{\Phi}$ and set Φ^{int} is not empty would PDM output the results. In another word, the CSMV should cover at least one interior sensor. In

case of a CSMV that only covers exterior sensors, it is deemed to cause no impact on the focal PV system, thus there is no need for preselection.

Subsequently, the time lag between the selected sensors and focal PV system can be estimated by:

$$\tilde{\gamma}_i = \frac{\vec{D}_i}{\vec{V}}, \quad S_i \in \widetilde{\Phi}, \tag{3.10}$$

where $\vec{D}_i$ is the spatial distance between the sensor S_i and the focal PV system. The number of temporal predictors is determined by the largest $\widetilde{\gamma}_i$:

$$\widetilde{n}_t = \texttt{max}(\frac{\widetilde{\gamma}_i}{\Delta t}), \quad S_i \in \widetilde{\Phi}. \tag{3.11}$$

Finally, the preselected ramp spatio-temporal predictors can be formulated as:

$$\widetilde{\boldsymbol{X}}_t = \Big[\underbrace{\tilde{x}_{i,t-T_s} \;\; \tilde{x}_{i,t-2T_s} \;\; \cdots \;\; \tilde{x}_{i,t-\widetilde{n}_t T_s}}_{\widetilde{n}_t}\Big], \; S_i \in \widetilde{\Phi}, \tag{3.12}$$

where $\widetilde{\boldsymbol{X}}_t \in \mathbb{R}^{1\times(\tilde{n}_s \times \tilde{n}_t)}$, and $\tilde{n}_s = \texttt{card}(\widetilde{\Phi})$ indicates the number of selected spatial predictors.

A complete preselection function

At this stage, a complete spatio-temporal preselection mechanism can be formed, where the following functions should apply:

$$\begin{cases} n_s = \texttt{sel}(\texttt{card}(\overline{\Phi}), \texttt{card}(\widetilde{\Phi}), \zeta_t) \\ n_t = \texttt{max}(\texttt{sel}(\frac{\overline{\gamma}_i}{\Delta t}, \frac{\widetilde{\gamma}_i}{\Delta t}, \zeta_t)) \end{cases}, \tag{3.13}$$

and

$$\boldsymbol{X}_t = \Big[x_{i,t-t_s} \; x_{i,t-2t_s} \cdots x_{i,t-n_t t_s}\Big], \quad S_i \in \texttt{sel}(\overline{\Phi}, \widetilde{\Phi}, \zeta_t), \tag{3.14}$$

where $\texttt{sel}(a, b, \zeta_t)$ is the selection function, which outputs a when $\zeta_t = 0$ (stationary scenarios), and outputs b when $\zeta_t = 1$ (ramp scenarios). In this way, different prediction preselection strategies can be automatically applied based on the scenario recognition signal.

Notice that although the use of the largest time lag $\widetilde{\gamma}_i$ in Equations (3.8) and (3.11) may result in irrelevant temporal predictors included in the model, the computational complexity is greatly reduced. This problem will be further addressed in the following section, where a penalized model is introduced for additional predictor filtering.

3.3.3 Nowcasting Strategy

Nowcasting model

As above mentioned, the preselection in both scenarios may introduce irrelevant temporal predictors. On this point, a penalized regression model becomes useful to provide further predictor filtering. The ridge regression and lasso are two frequently used penalized regression models. The former penalizes the residual sum of squares using an ℓ_2-penalty, while the latter takes ℓ_1-penalty. The biggest advantage of ridge regression is its stability and strong tolerance to small changes in model inputs. However, the ridge regression estimates always retain a whole set of predictors. In contrast, lasso can completely exclude the unexpected predictors from the model by shrinking their parameters to zero. Nonetheless, lasso may lose effectiveness when strong collinearity or aggregation effect is observed among the predictors, which is just the case for the sensors under a same cloud coverage. Consequently, only few or even one predictor is selected, and the model become less interpretable. In this work, we introduce the elastic-net, whose regularization term is a convex combination of ℓ_1-penalty and ℓ_2-penalty. The integration of both ℓ_1 and ℓ_2 penalties allows elastic-net to learn a sparse model with fewer zero weights than lasso, while maintaining the stability as ridge regression.

Given a dataset with n predictors and m samples $X \in \mathbb{R}^{m \times n}$, and responses

$Y \in \mathbb{R}^{m\times 1}$, the elastic-net estimator is given by:

$$\boldsymbol{\beta}^{\text{Enet}} = \underset{\boldsymbol{\beta}}{\texttt{argmin}}\Big\{\|X\boldsymbol{\beta} - Y\|_2^2 + \lambda\alpha\|\boldsymbol{\beta}\|_1 + \lambda(1-\alpha)\|\boldsymbol{\beta}\|_2^2\Big\}, \tag{3.15}$$

where $\boldsymbol{\beta}^{\text{Enet}} \in \mathbb{R}^{n\times 1}$ is the regression parameter, and $\lambda > 0$ is the regularization parameter which regulates the strength of the penalty. The larger the value of regularization parameter λ, the greater the amount of shrinkage and thus the parameters become more robust to collinearity. Specially, the elastic net transforms into ridge regression when $\alpha = 0$, and becomes lasso when $\alpha = 1$. In this work, all the parameters are selected using the k-fold cross validation.

Coordinated nowcasts

In stationary scenarios, a fixed forecast horizon of h is available. However, the forecast horizon in ramp scenarios, i.e., $\widetilde{n}_t T_s$, is limited by the size of the network and CSMV velocities. The larger the CSMV velocity and the smaller the network dimensions, the shorter prediction is provided. To aggregate the predictions in two scenarios, a coordinated framework is developed, as shown in Figure 3.6.

In the stationary scenarios, the spatio-temporal correlations among the sensors present to be more stable due to the steady irradiance resources. In this sense, the stationary predictor set $\overline{\boldsymbol{X}}_t$ is only updated for each time interval ΔT to adapt the gradual attributes change, e.g. ambient temperature, pressure and humidity, etc. In the ramp scenarios, beside updating the ramp predictor set $\widetilde{\boldsymbol{X}}_t$, the latest preselected stationary predictors is also used for consistent PV nowcasting at $t + h$. Then the $t + h$ nowcast with resolution of h is upsampled to the 1-s sequence, with backward interpolation. When the CSMV is detected, the interpolated value at $t + \tilde{n}_t T_s$ will be replaced by the nowcasting result. The value sequence is then downsampled back to h by calculating the mean of the sequence. In this way, consistent PV nowcasts with a fixed forecast horizon of h becomes available in both of the scenarios, while the cloud dynamics is contained.

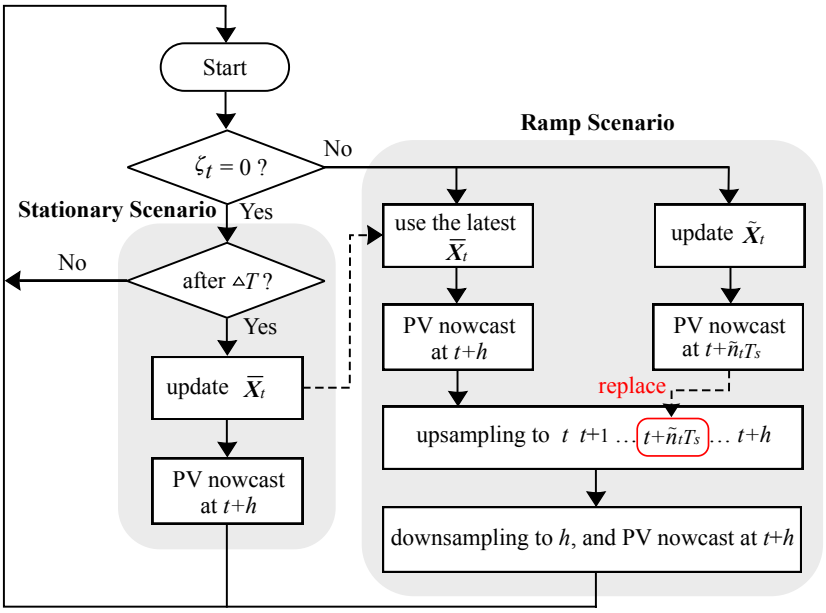


Figure 3.6: Flowchart of the developed coordinated nowcasting. The consistent PV nowcasts with a fixed forecast horizon at $t+h$ is available in stationary scenarios. The forecasts are further complemented when the CSMV is detected at $t+\widetilde{n}_t T_s$ in ramp scenarios.

3.4 Results

In this section, data from a concentric network prototype deployed in XJTLU, Suzhou, China, are considered, whose layout is shown in Figure 3.7. The network consists of 3 exterior sensors $\Phi^{ext} = \{S_2, S_3, S_4\}$ and 2 interior sensors $\Phi^{int} = \{S_1, S_5\}$. Another sensor is used to mimic the generation of a 5 MW PV system based on the irradiance-to-power conversion model in Equation (2.13). The specific network deployment is displayed in Table 3.1. The central server is equipped with Intel Core i7 2.9-GHz CPU, and all the processing is carried out using Python. The system is operated continuously from November 1, 2018 to April 1, 2019. Due to the limited experimental configuration, only the data measured in the days with dominated wind directions from 0° to 60° north are used. After several calibration tests, the correlation coefficient threshold is found as $\rho_{\text{limit}} = 0.8$, the threshold value for scenario recognition is chosen to be $\sigma = 2.5$, and the error tolerance of PDM is set as $\epsilon = 5$.

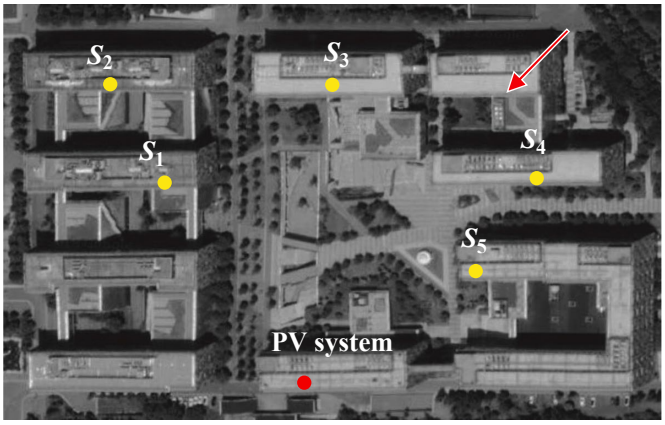


Figure 3.7: Layout of the deployed XJTLU sensor network prototype, with 3 exterior sensors and 2 interior sensors. Arrow in the top shows the dominate wind direction.

Table 3.1: Spatial distances in meters [m] between the sensors. The maximum distance is found as 224 m between S_2 and S_4. The minimum distance is found as 55 m between S_4 and S_5.

	S_1	S_2	S_3	S_4	S_5	PV
S_1	-	58	100	193	168	130
S_2	58	-	116	**224**	200	185
S_3	100	116	-	116	122	152
S_4	193	**224**	116	-	**55**	158
S_5	168	200	122	**55**	-	100

3.4.1 Error Metrics

To evaluate the proposed nowcasting strategy, three error metrics are used in this work, namely, the normalized root mean square error (nRMSE), forecast skill (FS), and normalized peak mean absolute error (nPMAE).

Given the measured PV generation P_t, and the predicted power $\hat{P}_t$ at time instant t, nRMSE is given by:

$$\text{nRMSE} = \frac{\sqrt{\frac{1}{m}\sum_{t=1}^{m}(P_t - \hat{P}_t)^2}}{\frac{1}{m}\sum_{t=1}^{m} P_t} \times 100\%. \tag{3.16}$$

The FS proposed in [84] is given by:

$$\mathrm{FS} = 1 - \frac{\mathrm{nRMSE}}{\mathrm{nRMSE}_p} \tag{3.17}$$

where nRMSE_p is the nRMSE produced by the persistence model. The metric FS equals to 1 for a perfect forecast, and takes a negative value when the proposed model is inferior to the persistence model.

In the case of PRRC, it has been shown in Figure 2.4 that the forecast accuracy on RR peak values (both positive and negative) is critical for control implementations. In this regard, the nPMAE is herein proposed, given by:

$$\mathrm{nPMAE} = \frac{\frac{1}{m_p}\sum_{t=1}^{m_p}|P_t^p - \hat{P}_t^p|}{\frac{1}{m_p}\sum_{i=t}^{m_p} P_t^p} \times 100\%, \tag{3.18}$$

where m_p is the number of peaks, P_t^p and $\hat{P}_t^p$ denote the measured and predicted peak values respectively.

3.4.2 Results of Predictor Preselection using SRP

To demonstrate the effect of SRP preselection, we consider a preselection example on a day with moderate variability, on 2018 December 24. In this example, consistent PV nowcasts with a prediction horizon $h = 10$ s are generated. Thus the data processed in stationary scenarios are averaged into 10-s intervals. In addition, a fixed PRRC time buffer for PDM is chosen as $\Delta t = 5$ s, and the stationary predictors is set to be updated for each $\Delta T = 10$ min.

Before implementing SRP, the scenario recognition signal is generated to describe the real-time PV dynamics, as shown in Figure 3.8(a). In this case, 13% data points are identified as "ramp". Figure. 3.8(b) and Fig. 3.8(c) show the spatial and temporal preselection results respectively. It can be seen that 5 spatial predictors, that is, a whole set of sensors are mostly selected in stationary scenarios, and at least 3 spatial predictors are included. This indicates that strong spatial correlations are observed

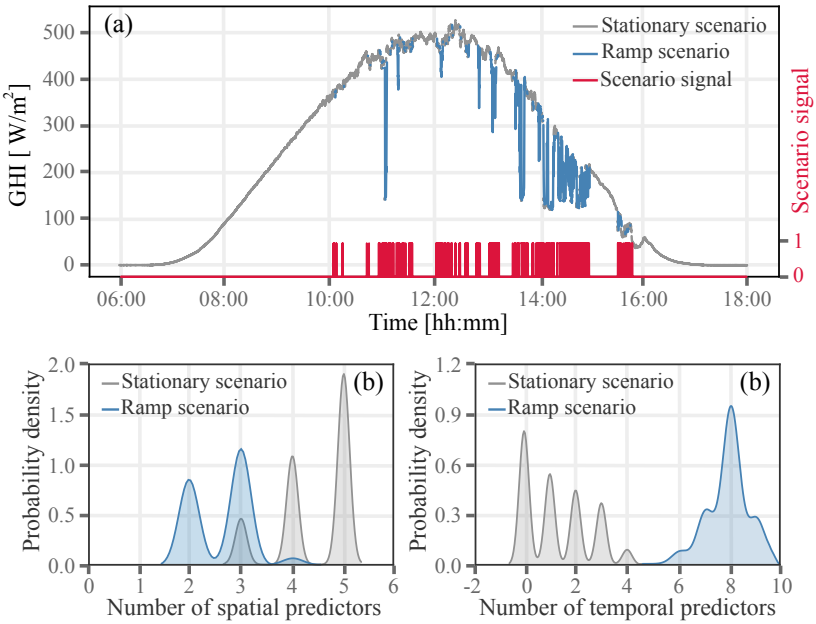


Figure 3.8: The preselection example for a moderate day on 2018 December 24. (a) PV generation profiles and scenario recognition results, where the grey line and blue line represent the recognized stationary and ramp scenarios respectively, and the red line shows the scenario signal. (b) Results of spatial predictor preselection using SRP. (c) Results of temporal predictor preselection using SRP.

among the sensors during these periods. In ramp scenarios, fewer spatial predictors are selected. Except for the two necessary along-wind sensors, PDM typically preselects one, and maximum two (in very few cases) additional spatial predictors. The selection is mainly limited by the predefined PRRC time buffer, which is fixed as 5 s in this case. On this point, the effect of varying Δt will be further studied in Section 3.4.4.

For the temporal predictor preselection, it can be seen that only few or even no temporal predictor is selected in stationary scenarios (0 in most cases). This is mainly due to the steady irradiance received during these periods, which greatly relieves the solar temporal dynamics. In ramp scenarios, 8 predictors are mostly selected, implicating a time lag of 8 s. Considering an average distance of 56.5 m between the focal system and interior sensors (S_4 and S_5), the average detected CSMV velocity is estimated as 7 m/s. Notice that the actual wind speed during the day has been reported as 6.5 m/s, which proves the proposed SRP method to be effective to track CSMV.

3.4.3 Results of Coordinated Nowcasting

Once the spatio-temporal predictors are preselected, the 10-s ahead nowcasting is generated by coordinating the results in the two scenarios. For each prediction, only the adjacent 720 data points (approximately 20% data of the whole day) are used for training, implicating the most recent 120 and 12 minutes data respectively for stationary and ramp scenarios. The workflow to produce nowcasts is similar to the previous chapter, namely the GHI data from sensors are transformed to CSI, then the nowcast CSI are converted back to GHI or PV power data for evaluation. For a single nowcast, the processing time can be controlled within 10 ms, which makes the proposed method feasible for real-time control.

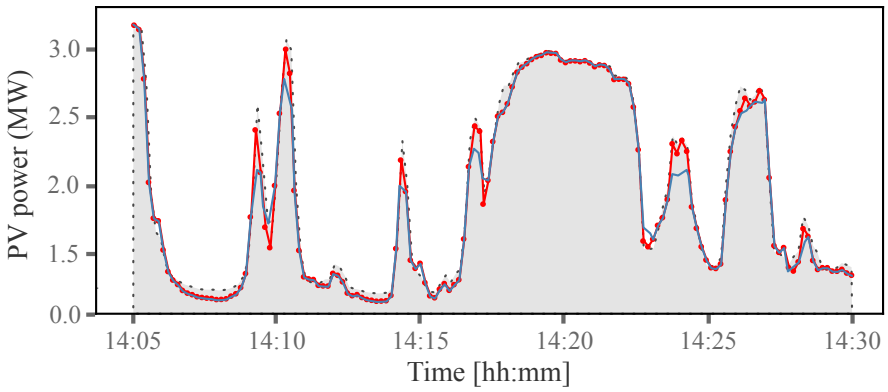


Figure 3.9: Nowcasting using SRP-Enet (red line) and measured PV generation (black dotted line) for a period on 2018 December 24. Nowcasting results before coordination (using stationary predictors only) is represented by the blue solid line. After coordination, the predictions approach closer to the actual peaks.

To visualize the coordination effects, a snapshot from 14:05 to 14:30 is plotted in Figure 3.9. It can be seen that the coordination improves the nowcasts significantly. Before coordination, only stationary predictors are fed into the model, thus it performs better when less fluctuations are observed. For the nowcasts during large ramps, the performance deteriorates as the predicted values always fail to reach the ramp peaks. After coordination, the ramp predictors begin to enter the model, making it more adaptive to fast moving CSMVs. As a result the nowcasts generated by stationary

predictors are further complemented, with the predicted values approaching closer to the ramp peaks. During this specific time period, the nPMAE before coordination is found as 15.80%, and reduces to 8.47% (43.4% improvements) after coordination. For the nowcasts of the entire day, the nRMSE, FS, and nPMAE are 4.71%, 0.40, 7.98%, respectively.

3.4.4 Case Study 1: Nowcasting under Various PRRC Time Buffers

The previous example assumes a fixed PRRC time buffer. In practice, however, the required time buffer can vary from case to case according to different PRRC regulations. Therefore, in this case study, we investigate the effect of various PRRC time buffers on nowcasting accuracy. Recall the coordination flowchart in Figure 3.6, the ramp predictor set ought to be updated before time $t + h$, implicating $\Delta t < h$. Otherwise the prediction becomes useless. In this sense, we herein consider the 30-s ahead nowcasting to provide more choices of Δt. The results are depicted in Table 3.2.

Table 3.2: Nowcasting performance for various PRRC time buffers. 30-s ahead nowcasting is generated on 2018 December 24.

Δt [s]	nRMSE [%]	nPMAE [%]	FS
2	5.05	10.56	0.36
5	4.71	7.98	0.40
10	**4.64**	**7.07**	**0.41**
20	5.35	12.05	0.32
30	5.68	16.87	0.27
60	5.68	16.87	0.27

It can be seen from Table 3.2 that the accuracy reduction with different time buffers is marginal. The use of 5-s and 10-s time buffers generate comparable results, and outperform the other cases. This result aligns with the discussions in Section 4.2. With an average wind velocity of 6.5 m/s, it generally takes 8 s for a CSMV to arrive the focal PV system. Hence, to effectively preselect the relevant predictors in a ramp scenario, a searching time near 8 s is preferred. However, for PRRC applications, the CSMV has to be identified before it reaches the focal system. In this regard, the 5-s choice seems

to be superior to the 10-s case although the 10-s time buffer offers slight improvements.

It is also found that when $\Delta t > 10$ s, the nowcasting performance begins to deteriorate. This is due to the slow preselection process that may not be able to timely update the ramp predictors. For the case of $\Delta t = 20$ s, there is only a 10-s interval before generating nowcasts (since the ramp predictors should be updated before $t + 30$ s). Consequently, only the CSMVs occurred between t and $t + 20$ can be identified. As Δt further increases, eventually there would be no CSMV being reported (such as the 30-s and 60-s cases). As a result, only stationary predictors are fed into the model, and the model becomes less robust in ramp scenarios with larger nPMAE errors. On the other hand, when the time buffer is insufficient (such as the 2 s case), fewer predictors are preselected, which also hinders the nowcasting performance.

3.4.5 Case Study 2: Nowcasting using Various Training Data Lengths

In this case study, we evaluate the effect of training data length on the nowcasting performance. The 10-s ahead nowcasts are generated as the previous example. In addition, another 9 cloudy days with similar wind directions are selected for validation (totally 10 days). To benchmark the proposed SRP-Enet method, the persistence (Pers), linear regression (LR), and long short-term memory recurrent neural networks (LSTM) models are used.

The Pers is the most commonly-used and simplest type of solar nowcasting model. It assumes the conditions (irradiance, temperature, cloud coverage, etc.) remain the same between t and $t + h$. Thus, Pers model only shows promising results at very short time horizons, making it a standard for benchmarking solar nowcasting [31]. On the other hand, LR is a universal benchmark model for spatio-temporal forecasting. When a short prediction horizon is desired, the strong linearity between the measurements of sensors and focal PV system makes the regression-based models a priority for predictions [67]. Furthermore, in order to compare with the recent advances in AI-enhanced technology, the LSTM model, which is typically applied to time-series forecasting, is included for comparison as well [85]. The LSTM model used here is constructed with 50 neurons in

the first hidden layer and 1 neuron in the output layer. The model is fit for 50 epochs with a batch size of 10. The LR and LSTM models use the same training data length as the proposed method. Moreover, the CSMV information is unknown to the two models, thus sufficient predictors are assumed, i.e. a whole set of 5 spatial predictors and 10 temporal predictors. The Pers model directly transforms the measurements of sensor S_5 to PV generation. The results are displayed in Table 3.3.

Table 3.3: 10-s ahead nowcasting performance of the proposed SRP-Enet method and three benchmarking models with various training data points. The results are averaged over the 10 days.

Training length	Metrics	Pers	LR	LSTM	SRP-Enet
360	nRMSE (%)	8.24	8.82	9.92	6.57
	nPMAE (%)	18.20	19.96	14.70	12.67
	FS	-	-0.07	-0.20	0.20
720	nRMSE (%)	8.24	6.74	8.34	**5.83**
	nPMAE (%)	18.20	15.92	14.61	**8.57**
	FS	-	0.18	-0.01	**0.29**
1080	nRMSE (%)	8.24	6.70	8.19	6.18
	nPMAE (%)	18.20	15.21	14.37	10.42
	FS	-	0.19	0.01	0.25
1440	nRMSE (%)	8.24	6.63	7.91	6.29
	nPMAE (%)	18.20	14.90	14.08	11.24
	FS	-	0.20	0.04	0.24
1800	nRMSE (%)	8.24	**6.58**	**7.64**	6.37
	nPMAE (%)	18.20	**14.62**	**13.97**	11.78
	FS	-	**0.20**	**0.07**	0.23

It is an interesting finding that the proposed SRP-Enet method generates the best nowcasts with 720 data points (20% training length). As the training data length further increases, its performance deteriorates instead. This is mainly due to the effects of SRP, where the predictors are constantly updated. As a result, data collected several hours ago may not be suitable to fit the latest predictors. For LR and LSTM models, their accuracies decrease when the training data become fewer. Even when the data is sufficient such as using 1800 points (data collected within 5 hours, accounting for

40% of the whole data set), due to the large number of irrelevant predictors, their performances are still inferior to that of of SRP-Enet. Furthermore, it can be seen that the more advanced LSTM model is beaten even by a simple LR model in the tests. This observation is consistent with the discussions in [67] that the regression-based methods could be more suitable in solar nowcasting.

We note that the "curse" of training data length may be relaxed in practical operations by using historical data with similar meteorological conditions. In this regard, the performance of LSTM-type models may be further improved. However, this case study verifies that the proposed SRP-Enet method can use much fewer training data while achieving similar or better results than the other models, which makes it more advantageous when historical data are unavailable.

3.4.6 Case Study 3: Nowcasting with Various Forecast Resolutions

In previous studies, performance of the SRP-Enet along with several benchmarking models is evaluated at a forecast resolution of 10 s. In this case study, nowcasts with various forecast resolutions are presented. Data from the selected 10 days are averaged into 10, 20, 30, 60, and 300 s intervals, respectively. Based on the results from case study 2, the most recent 20% data are used for training. Table 3.4 shows the nowcasting results.

It can be concluded from Table 3.4 that the proposed SRP-Enet method shows better performance than the benchmarking models for all the forecast resolutions. For the nowcasts within 60 s, cloud dynamics still dominate the accuracy, thus the inclusion of temporal predictors in the model can significantly improve the nowcasts. For $r =$ 300 s, the impacts of CSMV are greatly relieved thanks to the well-known temporal smoothing effect [70]. In this case, the inclusion of unnecessary temporal predictors is likely to deteriorate the nowcasts, which accounts for the unacceptable results of the LR model. Although the redundant temporal predictors are also included in the SRP-Enet, the inherent regularized term of elastic-net model provides a supplementary predictor filtering. Thus the SRP-Enet is still comparable to the Pers model. It is also observed

Table 3.4: Nowcasting performance of the proposed SRP-Enet and benchmarking models at various forecast resolutions. The results are averaged over the 10 days.

r [s]	Metrics	Pers	LR	LSTM	SRP-Enet
10	nRMSE (%)	8.24	6.74	8.34	5.83
	nPMAE (%)	18.20	15.92	14.61	8.57
	FS	-	0.18	-0.01	0.29
20	nRMSE (%)	7.84	6.00	7.33	4.86
	nPMAE (%)	15.34	12.58	13.22	7.26
	FS	-	0.23	0.06	0.35
30	nRMSE (%)	6.91	5.67	6.84	4.59
	nPMAE (%)	10.59	9.53	10.07	6.88
	FS	-	0.18	0.01	0.34
60	nRMSE (%)	5.53	4.92	5.37	4.24
	nPMAE (%)	8.28	7.82	8.15	6.32
	FS	-	0.11	0.03	0.23
300	nRMSE (%)	2.35	4.02	2.59	2.30
	nPMAE (%)	4.50	7.08	5.11	3.95
	FS	-	-0.7	-0.1	0.02

that as the forecast resolution expands, the nPMAE differences between the models decreases rapidly. This also verifies the temporal smoothing effect at larger forecast resolutions.

3.4.7 Case study 4: Nowcasting in Different Weathers

So far the studies have shown that the proposed SRP-Enet model is able to produce better results in cloudy days. In this case study, we investigate the nowcasting performance in different weathers, namely, the sunny, cloudy, and rainy days. For each weather type, 10 days are selected during the experimental days. Based on the previous studies, 20% data are used for training to generate 10-s ahead PV nowcasts. Figure 5.4 shows the examples of nowcasting in the 3 typical weather conditions.

It can be seen from Figure 5.4 that the SRP-Enet nowcasts approach to the ramp peaks closely in all weather conditions, which is otherwise unachievable using other methods. This property should become especially important for PRRC operations.

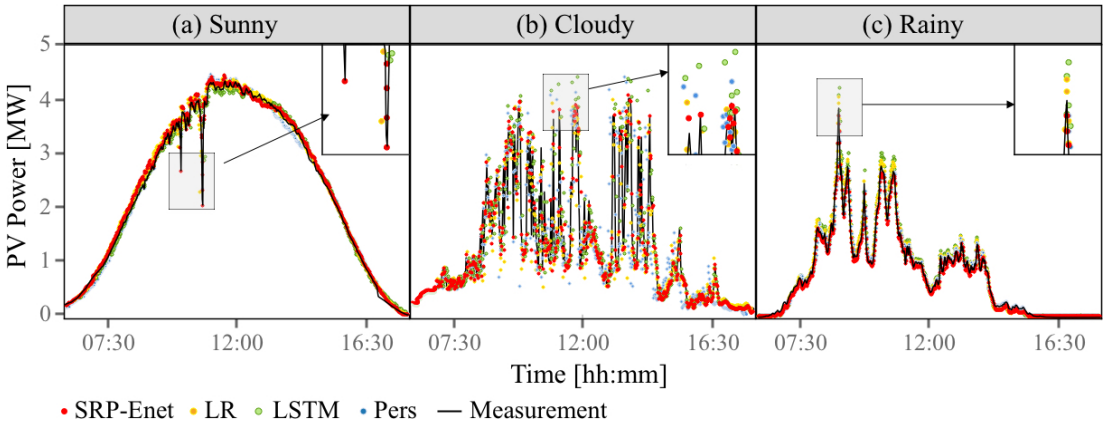


Figure 3.10: Nowcasting examples under typical weather conditions (a) sunny, (b) cloudy, (c) rainy. The proposed SRP-Enet method approaches the closest to the ramp peaks compared with other methods.

Table 3.5 depicts the nowcasting performance during the respective 10 days. It is evident from Table 3.5 that the SRP-Enet method generally outperforms the benchmarking methods in all weathers, with an apparent reduction in terms of nPMAE. In the cloudy days, the SRP-Enet nowcasts show an average nRMSE and nPMAE improvements over 13.5% and 41.3% compared with the benchmarking methods. At first glance, the SRP-Enet does not show as much nRMSE improvements in sunny days as it does in cloudy days. Due to the infrequent cloud coverages in the sunny days, even a large error in the predicted power fluctuations may not significantly influence the nRMSE. In this regard, the nRMSE is unfaithful to tell the reliability of the nowcasts. On the other hand, the nPMAE, which aims to describe the ability of predicting PV ramps, becomes useful. In sunny days, the nPMAE improvements by the SRP-Enet is more than 20% over the other methods. Another observation can be made is that the SRP-Enet nowcasts in the rainy days present to be inferior to Pers in terms of nRMSE. As the CSMV diminishes in precipitation, the raining periods can be often recognized as stationary scenarios. However, the meteorological features such as ambient temperature or humidity can change rapidly during these periods. Thus, an interval of $\Delta T = 10$ min for updating stationary predictors may not response to these sudden changes in time.

Table 3.5: Nowcasting performance in different weather conditions. 10 typical days are selected for each weather type. The results are averaged over the respective days.

Weather	Metrics	Pers	LR	LSTM	SRP-Enet
Sunny	nRMSE (%)	2.71	2.37	3.19	**2.35**
	nPMAE (%)	6.64	5.98	7.02	**4.81**
	FS	-	0.13	-0.18	**0.13**
Cloudy	nRMSE (%)	8.24	6.74	8.34	**5.83**
	nPMAE (%)	18.20	15.92	14.61	**8.57**
	FS	-	0.18	-0.01	**0.29**
Rainy	nRMSE (%)	**4.59**	5.01	5.65	4.92
	nPMAE (%)	7.86	9.26	10.99	**7.17**
	FS	-	-0.09	-0.23	**-0.07**

We have noticed that in some areas, additional weather types such as snowstorm and sandstorm may also be observed. In these cases, however, the proposed SRP-Enet forecasting can still be utilizable, since the sensors can feel the snow drop or sand cover as similar as the PV system is experiencing (the sensor material, tilt angle etc. are similar to the focal PV system).

3.4.8 Case study 5: applying SRP to all the models

The last case study in this chapter evaluates the effects of SRP. Nowcasts with and without uing SRP are studied for all the models. When SRP is unavailable, sufficient spatio-temporal predictors, i.e., a full set of spatial predictors $n_s = 5$ with temporal lag $n_t = 10$, are considered. The configurations of this case study are identical to case study 4, namely, generating 10-s ahead nowcasts with a training length of 20%. The results are generalized on the 10 cloudy days to exclude the non-cloud factors (such as the sudden change of temperature and humidity in rainy days). Figure 3.11 shows the snapshots of nowcasts using different models before and after applying SRP.

Before SRP, a large number of irrelevant predictors are fed to the models. Consequently, the LR nowcasts typically become much more fluctuated than expected, and the LSTM nowcasts present to be lagging to the measurements. Benefiting from its

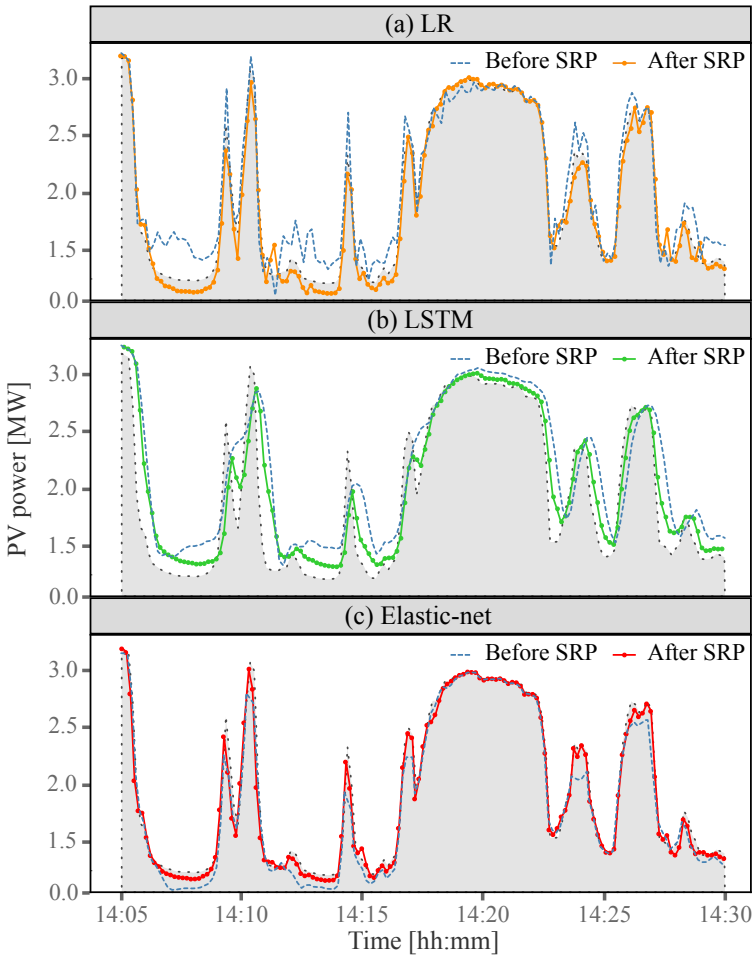


Figure 3.11: Nowcasts before and after applying SRP to (a) LR, (b) LSTM, and (c) Elastic-net models for a period on 2018 December 24. The measured PV generation is represented by the black dotted line.

penalty term, the elastic-net produces much smoother nowcasts than LR, and there is no lag between the predicted and measured time series. However, it is still far beyond satisfactory since it always fails to reach the ramp peaks. After SRP, it can be seen that the nowcasts for all the models are significantly improved. With proper predictors preselected, the LR nowcasts become more stable. The lagging issue of LSTM is also relieved. More importantly, all the nowcasts approach closer to the ramp peaks.

The error metrics before and after applying SRP for the models are shown in Table 3.6. It is observed that although the benchmarking models use the similar predictors,

Table 3.6: Nowcasting performance before and after applying SRP. The results are averaged over the 10 days.

	Metrics	LR	LSTM	Elastic-Net
Before	nRMSE (%)	6.74	8.34	6.18
	nPMAE (%)	15.92	14.61	10.36
	FS	0.18	-0.01	0.25
After	nRMSE (%)	**6.12**	**6.58**	**5.83**
	nPMAE (%)	**10.02**	**10.27**	**8.57**
	FS	**0.26**	**0.20**	**0.29**

their performances are still inferior to the elastic-net based nowcasts. This observation validates the effectiveness of the penalty term of the elastic-net model, which is able to provide additional predictor reduction. In this case, the nRMSE improvements after applying SRP are found as 9.19%, 21.1%, and 5.66%, respectively for the LR, LSTM, and elastic-net models. The nPMAE improvements are found as 37.06%, 29.71%, and 17.28%, respectively.

3.5 Application to Predictive PRRC

In this section, we demonstrate the practical application of the proposed SRP-Enet nowcasting on predictive PRRC. Especially, we draw the attention on PAPC since it does not involve any external auxiliary device such as ESS, thus the value of nowcasts can be more properly quantified.[1]

To evaluate the performance of PAPC, two metrics are used, namely, the ramp smoothing rate (RSR) and energy curtailment ratio (ECR). The RSR assesses the reduction in RR violations after PAPC, which is defined as:

$$\text{RSR} = 1 - \frac{\sum_{t=1}^{n'} \mathbf{I}(|R_t^*|, R_s)}{\sum_{t=1}^{n'} \mathbf{I}(|R_t|, R_s)} \times 100\%, \tag{3.19}$$

[1]Recall Figure 2.4, for the PEC method, even though the nowcasts could contain large errors, the use of an ESS can still help to smooth out the RRs. In this case, it would become unclear how the more accurate nowcasts can benefit control process. Hence, to show the value of the proposed nowcasting method, we herein only demonstrate PAPC but not PEC.

where R_t^* and R_t are the RR at time t after and before PAPC respectively, and R_s is the required RR limit. $\mathbf{I}(|R_t^*|, R_s)$ is 1 if $|R_t^*| > R_s$, and equals to 0 otherwise. Generally, a larger RSR indicates less RR violations, and the regulated output power becomes smoother. The ECR quantifies the amount of curtailed energy due to PAPC, given by:

$$\text{ECR} = 1 - \frac{\sum_{t=1}^{n'} \tilde{P}_t}{\sum_{t=1}^{n'} P_t} \times 100\%. \tag{3.20}$$

In most cases, one wishes ECR to be small so that more production can be extracted from PV systems. Typically, ECR needs to be interpreted together with RSR since less curtailment may introduce more RR violations.

Before implementing PAPC, a RR limit R_s should be first defined. In this study, we follow the regulation in [15], where the maximum allowed RR is suggested to be 100 kW/s for the PV systems above 11 kW. It should be noticed that in case of PRRC, the time buffer Δt in PDM is no longer a fixed value. Given the first sensor reading change at t_0, the control time buffer equals to

$$\Delta t = t_c - t_0 \tag{3.21}$$

where t_c is the proactive control time defined in Equation (2.9), which changes with different CSMVs.

Table 3.7 shows the PAPC performance over the 10 cloudy days. According to the previous case studies, the 10-s ahead PV nowcasts are generated with a training data length of 20%. To benchmark the proposed SRP-Enet method, the DST model developed in previous chapter is considered. Since the DST nowcasting works only when ramp violations are observed, the metrics except for nPMAE are discarded. It can be concluded from Table 3.7 that the proposed SRP-Enet method generally outperforms the DST model. In both moderate and highly variable days, the SRP-Enet method achieves lower nPMAE and higher RSR. It is also observed that in highly variable days, the DST method seems to sacrifice less energy. However, this is mainly due to its more control failures. In another word, the energy has not been fully curtailed to smooth out

Table 3.7: PAPC performance using SRP-Enet and DST over the 10 cloudy days.

	Moderate		Highly variable	
Metrics	DST	SRP-Enet	DST	SRP-Enet
nPMAE (%)	8.09	**7.25**	10.08	**9.96**
RSR (%)	95.99	**96.05**	95.47	**96.42**
ECR (%)	2.23	**1.66**	**2.89**	3.05

the ramp-down fluctuations due to the nowcasting errors. We note that the effective use of DST model should require a larger and denser sensor network. However, the feasibility and simplicity are also important considerations for practical application. According to the results presented above, the proposed SRP-Enet method is proved to be more reliable than DST model when fewer sensors are considered.

3.6 Chapter Conclusion

This chapter presents a novel sensor network-based PV nowcasting method with a newly developed spatio-temporal predictor preselection, which can be used for PAPC. Compared with the conventional PV nowcasting methods, the main improvements of the proposed method can be summarized as:

- The proposed SRP preselection establishes a comprehensive predictor preselection mechanism for both spatial and temporal predictors in absence or presence of CSMVs.
- The proposed SRP preselection decomposes the PV dynamics into two scenarios, and the results in two scenarios can be coordinated to provide consistent PV nowcasts at a fixed forecast horizon, with CSMV information well contained. Specifically, the use of the developed SRP preselection on different models shows an average nRMSE and nPMAE improvements over 11% and 25% respectively.
- The proposed SRP-Enet nowcasting reveals strong adaptability under various weather conditions, especially in the highly cloudy days, with an average nRMSE

and nPMAE improvements over 13.5% and 41.3% respectively. In addition, it is able to use fewer training data while achieving similar or better results than the benchmarking models which use a larger training set.

The feasibility of integrating SRP-Enet nowcasts to PAPC operation is also evaluated through practical experiments. The results show that the use of SRP-Enet nowcasts on PAPC outperforms the previously developed DST model, with less control failures and energy curtailment.

Chapter 4

Analysis of Operational Nowcasting for PRRC

4.1 Chapter Introduction

In previous two chapters, the feasibility of integrating solar nowcasting into PRRC operations have been demonstrated. While the practices therein presented could potentially bring in more nowcasting applications in the field of PRRC, an important issue—*operability*—is left out.

In a general sense, the operability denotes the operational nature of solar forecasting [86]. Since the main purpose of generating solar forecasts is to assist in power system operations, the forecasts ought to follow certain operational requirements in terms of *temporal issues*, such as forecast horizon, resolution, or update rate [87]. As grid integration involves a variety of timescales that relate to different system operations, the operational requirements naturally differ. To give perspective, Table 4.1 summarizes the operational forecasting requirements used by California Independent System Operators (CAISO) [88], where three different forecast requirements for intra-hour, intra-day, and day-ahead solar forecasting are enacted, corresponding to power balance, real-time market, and day-ahead market operations, respectively. While the operational

Table 4.1: Overview of operational requirements for intra-hour, intra-day, and day-ahead solar forecasting in CAISO.

Forecast type	Target horizon	System operation	Operational forecasting requirement
Intra-hour	15 min-1 h	Power balance	The real-time economic dispatching of CAISO requires thirteen 5-min forecasts to be submitted 7.5 min prior to the operating hour. The process repeats every 5 min.
Intra-day	1 h-1 day	Real-time market	The short-term unit commitment of CAISO requires 5-h-ahead forecasts to be submitted 75 min before the operating hour, with a 15-min forecasting resolution. The process repeats every hour.
Day-ahead	$>$ 1 day	Day-ahead market	CAISO requires the hourly day-ahead forecasts to be submitted at 5:30 A.M. on the day before the operating day, corresponding to a 18.5-42.5-h-ahead forecasting horizon.

forecasting has been carefully addressed in intra-hour, intra-day, and day-ahead solar forecasting studies [87, 89, 90], such temporal issues are often overlooked in solar nowcasting practices.

In light of the above, this chapter aims to address the operability of solar nowcasting by demonstrating an operational solar nowcasting practice on PAPC.[1] Specifically, various temporal issues related to operational solar nowcasting will be clarified, and their impacts on operational nowcasting and PAPC will be investigated. To perform PAPC, both deterministic and probabilistic nowcasts will be considered.

4.2 Design of PAPC Operating Environment

This section elaborates the PAPC operation design. Several implementational issues, i.e., time parameters, data, PAPC algorithm, and operating timeline are clarified.

[1] As discussed in Chapter 3, to quantify the effect of nowcasting settings more properly, we herein only consider the PAPC implementation to exclude the support from ESS.

4.2.1 Temporal Issues in Operational Solar Nowcasting

Previously in [87], four time parameters are introduced to fully characterize the temporal issues in operational forecasting, namely, forecast horizon, forecast resolution, forecast lead time, and forecast rolling update rate. Nonetheless, as the above time parameters are specially defined in intra-day and day-ahead forecasting scenarios, the situation slightly changes when it comes to solar nowcasting.

Firstly, given that intra-day and day-ahead solar forecasting are commonly applied in power systems research, forecast lead time is typically required by system operators to schedule grid reserve capacities, unit commitment, and economic dispatch [91]. However, such prior scheduling is no longer needed in terms of PAPC, since the high reaction speed of FPPT (typically < 1 s) enables nearly real-time PV power control [92]. Besides the forecast lead time, the forecast rolling update rate needs to be reconsidered for solar nowcasting as well. For intra-day and day-ahead solar forecasting, the forecasts are often run much further into the future than the required lead time. In this case, it is natural to consider the forecast rolling update rate to refresh that forecast, thus providing more accurate and valuable information for system operators. However, involving rolling forecasts may become infeasible for solar nowcasting in practice, as it could be too computationally intensive to update the nowcasts every several seconds [93].

Based on the above viewpoints, three time parameters for operational solar nowcasting are herein defined, namely, forecast horizon ($\mathcal{H}$), forecast resolution ($\mathcal{R}$), and forecast model update rate ($\mathcal{M}$).[2] Analogous to the intra-hour, intra-day, and day-ahead forecasting cases, the *forecast horizon* denotes the time span from the first to the last required nowcasts in each submission, and the *forecast resolution* is the temporal resolution of the required nowcasts. The *forecast model update rate*, on the other hand, is a newly defined time parameter for solar nowcasting. As noted in previous chapters, the performance of solar nowcasting can be highly characterized by intra-day meteorological

[2]In principle, these time parameters should be termed as nowcast horizon, nowcast resolution, and nowcast model update rate. However, to keep consistency with the the intra-hour, intra-day, and day-ahead forecasting cases, the word "forecast" is used. Moreover, as solar nowcasting is also known as intra-minute solar forecasting, the naming generally applies.

variations, such as wind speed, cloud motions, or temperature. In this regard, the nowcasting models or the predictors need to be constantly updated to adapt the frequent weather changes. With above definitions, a triplet can be used to denote the temporal requirements in operational solar nowcasting, i.e., $(\mathcal{H}, \mathcal{R}, \mathcal{M})$. For example, if the 5-s nowcasts out to 30 s are required, and the nowcast model is updated every 10 min, the triplet ($\mathcal{H}^{30s}$, $\mathcal{R}^{5s}$, $\mathcal{M}^{10min}$) fully describes such temporal requirements.

4.2.2 Data and Implementation of PAPC

The empirical part of this chapter considers the same dataset as Chapter 2, i.e., the Oahu solar measurement grid, as shown in Figure 4.1. To demonstrate PAPC, sensor DH4 is selected as the focal system, i.e., the target of nowcasting and control. Since solar-generated power is of interest as to PAPC operations, the GHI measurements at DH4 are converted to PV power data using the similar irradiance-to-power conversion model in Equation (2.13). Moreover, in order to produce isotropic nowcasts, we herein only consider 8 sensors with approximate proximities to DH4 as predictors (marked by red circles in Figure 4.1). The sensor pairwise distance varies from 103 m (DH4 and DH5) to 226 m (DH4 and DH6).

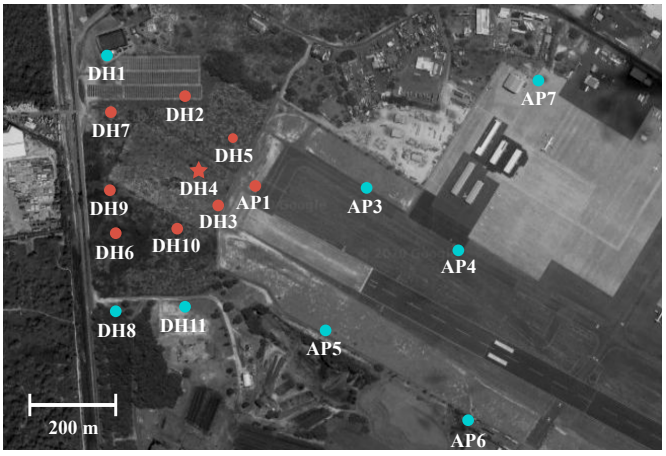


Figure 4.1: Layout of the NREL Oahu sensor network. The 9 sensors for empirical study are marked by red, where the sensor DH4 is selected as the PAPC target (marked by the red star). Surrounding the target sensor DH4, other 8 sensors form a closed circular deployment. The scale of the map is shown in the bottom left corner.

Given the above operation setup, the PAPC is implemented as follows. Assuming the nowcasts are produced with the forecast horizon h and resolution r, and the control sampling time of PAPC is t_s. For an arbitrary nowcast run at time t, a nowcast series with length h/r is available:

$$\hat{\boldsymbol{P}}_t = \{\hat{P}_{t+r}, \hat{P}_{t+2r}, \cdots, \hat{P}_{t+h}\}, \tag{4.1}$$

where $\hat{P}_t$ denotes the nowcast for DH4 at time t. However, to integrate the nowcasts into PAPC design, the forecast resolution should match the control sampling time. In this case, the nowcasts with resolution r needs to be downscaled[3] to t_s, leading to a length-h/t_s nowcast series:

$$\hat{\boldsymbol{p}}_t = \{\hat{p}_{t+t_s}, \hat{p}_{t+2t_s}, \cdots, \hat{p}_{t+h}\}, \tag{4.2}$$

where $\hat{p}_t$ denotes the downscaled nowcast for $\hat{P}_t$. Based on Equation. (4.2) and the instantaneous measurement at DH4, PAPC firstly constructs a time series $\boldsymbol{S}_t$,

$$\boldsymbol{S}_t = \{P_t, \hat{p}_{t+t_s}, \hat{p}_{t+2t_s}, \cdots, \hat{p}_{t+h}\}, \tag{4.3}$$

where P_t is the measurement of DH4 at time t, and the length of $\boldsymbol{S}_t$ is $1 + h/t_s$. Given Equation. (4.3), a resultant RR series can be calculated as:

$$\boldsymbol{R}_t = \{\hat{R}_t, \hat{R}_{t+t_s}, \hat{R}_{t+2t_s}, \cdots, \hat{R}_{t+h-1}\}, \tag{4.4}$$

where

$$\hat{R}_{t+i\cdot t_s} = \frac{\boldsymbol{S}_t[i+2] - \boldsymbol{S}_t[i+1]}{t_s}, \tag{4.5}$$

is the RR calculated between the $(i+2)^{\text{th}}$ and $(i+1)^{\text{th}}$ elements of time series $\boldsymbol{S}_t$, and $i = 0, 1, \cdots, h/t_s - 1$. Finally, by identifying the location and value of the minimum RR

[3]In this work, the downscaling is performed by linear interpolation.

in $\boldsymbol{R}_t$, denoted by k and $\hat{R}_{\min}$ respectively, PAPC takes the following operations:

$$\tilde{p}_{t+i} = \begin{cases} \hat{p}_{t+i} & \text{if } \hat{R}_{\min} \geq -R_s \\ \hat{p}_{t+i+1} + R_s \cdot t_s & \text{if } \hat{R}_{\min} < -R_s, \end{cases} \tag{4.6}$$

where $i = k-1, k-2, \cdots, 1$, R_s is the required RR limit, and $\tilde{p}_t$ is the power signal for FPPT. It should be noticed that the power signal may be invalid as it becomes greater than the maximum available power. In that case, $\tilde{p}_t = P_t$.

4.2.3 Operating Timeline

According to Section 4.2.1, it is now clear that in order to integrate operational solar nowcasting into PAPC, three time parameters should be involved, i.e., $\mathcal{H}$, $\mathcal{R}$, and $\mathcal{M}$. At this point, a PAPC operating timeline is designed, which will be used throughout this chapter. Figure 4.2 depicts an example timeline, exemplified under the nowcasting setting ($\mathcal{H}^{15\text{s}}$, $\mathcal{R}^{1\text{s}}$, $\mathcal{U}^{5\text{min}}$). As PAPC only works during the daytime, the operation is assumed to start at 7:30:00 and lasts for 10 hours. In contrast to the hourly or day-ahead solar forecasting which looks several days into the history, solar nowcasting relies mostly on data collected on the actual operation day to make inference. It has been shown in Chapter 3 that when sensor network-based nowcasting is considered, training with recent data (e.g., several hours) often leads to a higher nowcast accuracy than using longer training length, as the spatio-temporal correlations within the network could change frequently. In this regard, this work uses the most recent 1-h data for training at each model updating stage. Notice that for 1-s measurements, the 1-h training data leads to a length-3600 training set, which is large enough for the nowcasting models considered in this work.

Based on Figure 4.2, the process of constructing $\boldsymbol{S}_t$ for PAPC is displayed in Table 4.2. Firstly, the nowcast submission starts at 7:30:00, and fifteen 1-s nowcasts are produced for the period of 7:30:01-7:30:15 (the second column in Table 4.2). Referring to Equation (4.3), $\boldsymbol{S}_1$ should cover the period of 7:30:00-7:30:15 (the last column in

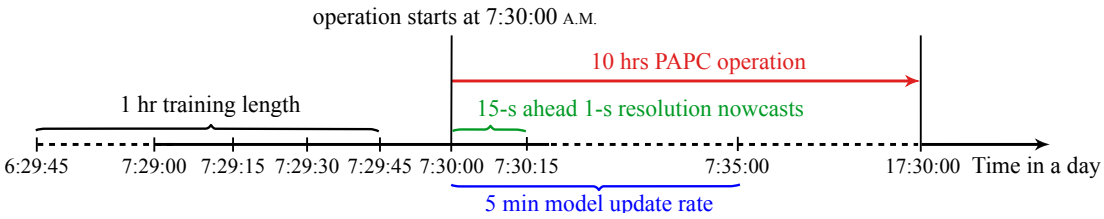


Figure 4.2: An example of PAPC operating timeline exemplified under ($\mathcal{H}^{15s}$, $\mathcal{R}^{1s}$, $\mathcal{U}^{5min}$).

Table 4.2: An illustration of PAPC operation under the nowcasting setting ($\mathcal{H}^{15s}$, $\mathcal{R}^{1s}$, $\mathcal{M}^{5min}$). The timestamps marked by bold indicate the model updating time.

Submission time	Forecast timestamps	$\boldsymbol{S}_t$ timestamps
07:30:00	07:30:01-07:30:15	07:30:00-07:30:15
07:30:01	07:30:16	07:30:01-07:30:16
07:30:02	07:30:17	07:30:02-07:30:17
⋮		
07:35:00	07:35:15	07:35:00-07:35:15
⋮		
17:30:00	17:30:15	17:30:00-17:30:15

Table 4.2). At 7:30:01, the second nowcast is submitted, which contains, however, only a single nowcast for 7:30:16. To construct $\boldsymbol{S}_2$, the same nowcasts as $\boldsymbol{S}_1$ for the period of 7:30:02-7:30:15 will be used. Then together with the instantaneous measurement at 7:30:01 and a new nowcast for 7:20:16, $\boldsymbol{S}_2$ can be formed. The process continues until the end of the operation day. It should be noticed that for each nowcast submission, a complete nowcast time series, that is, fifteen 1-s nowcasts are actually available. However, as mentioned earlier, such rolling operation could bring more confusions for PAPC.[4] Hence, except for the first nowcast submission, a single 15-s ahead nowcast is produced for the rest.

[4]Let us consider an example here. Suppose the operation follows the rolling manner and the control sampling time is 1 s, then at $t = 1$, we have $\boldsymbol{S}_1 = \{P_1, \hat{p}_2, \hat{p}_3, \cdots, \hat{p}_{16}\}$. At $t = 2$, PAPC is tuning P_2, and $\boldsymbol{S}_2 = \{P_2, p_3^*, p_4^*, \cdots, p_{17}^*\}$, where p^* denotes the downscaled nowcast produced at $t = 2$. In this case, there are multiple nowcasts produced at the same timestamp, i.e., $\hat{p}_3$ and p_3^*, $\hat{p}_4$ and p_4^*, $\cdots$, $\hat{p}_{16}$ and p_{16}^*. The confusion comes as which nowcast should be used.

4.3 Nowcasting Methods

In this section, the nowcasting models applied to PAPC are introduced. A total of eight models, i.e., four deterministic and four probabilistic models, are considered for GHI nowcasting, most of which come from two recent solar forecasting benchmark articles [94, 95]. In addition, the error metrics for both deterministic and probabilistic nowcasts are explained.

4.3.1 Nowcasting Setup

At each model updating stage, the nowcasting models are built based on a size n-by-p training matrix, consisting of n samples and p predictors,

$$\boldsymbol{X}_{\text{train}} = \begin{pmatrix} \boldsymbol{x}_1^\top \\ \boldsymbol{x}_2^\top \\ \vdots \\ \boldsymbol{x}_n^\top \end{pmatrix} = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,p} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,p} \end{pmatrix}, \tag{4.7}$$

and a length-n training response vector containing CSI values at DH4,

$$\boldsymbol{y}_{\text{train}} = \begin{pmatrix} y_1 & y_2 & \cdots & y_n \end{pmatrix}^\top. \tag{4.8}$$

Provided one has n' nowcasts to be made, whose index resumes from n, a test matrix can be written as:

$$\boldsymbol{X}_{\text{test}} = \begin{pmatrix} \boldsymbol{x}_{n+1}^\top \\ \boldsymbol{x}_{n+2}^\top \\ \vdots \\ \boldsymbol{x}_{n+n'}^\top \end{pmatrix} = \begin{pmatrix} x_{n+1,1} & x_{n+1,2} & \cdots & x_{n+1,p} \\ x_{n+2,1} & x_{n+2,2} & \cdots & x_{n+2,p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n+n',1} & x_{n+n',2} & \cdots & x_{n+n',p} \end{pmatrix}. \tag{4.9}$$

By looping through the rows of $\boldsymbol{X}_{\text{test}}$, a vector of nowcasts,

$$\hat{\boldsymbol{y}}_{\text{test}} = \begin{pmatrix} \hat{y}_{n+1} & \hat{y}_{n+2} & \cdots & \hat{y}_{n+n'} \end{pmatrix}^{\top}, \tag{4.10}$$

can be obtained.

Recall the operating timeline, the training process is based on the data collected during the last 1 hour. In this regard, $\boldsymbol{X}_{\text{train}}$ should contain $n = 3600/r$ samples, where r is the forecast resolution. Moreover, to identify the spatio-temporal predictors that to be fed into $\boldsymbol{X}_{\text{train}}$, the SRP method developed in Chapter 3 is used. On the other hand, the value of n' is determined by the forecast model update rate. For instance, given $\mathcal{M}^{5\text{min}}$, $n' = 300/r$.

4.3.2 Models for Deterministic Nowcasting

In this work, four models are considered for deterministic nowcasting: 1) smart persistence (SP), 2) ordinary least squares (OLS) regression, 3) lasso regression, and 4) k-nearest neighbors (kNN) regression. While the first three methods are often used as reference models in deterministic solar forecasting [94], kNN regression is selected to typify the pattern-matching-based methods. Notice that the more advanced artificial intelligence methods are excluded in this work. This is because that most of AI-based methods require longer training period that is far below the solar nowcasting requirements [93]. Moreover, it has been reported in Chapter 3 that the regression-based methods are generally more suitable for solar nowcasting, due to the strong correlations between predictors and response.

Smart persistence

The SP model refers to a persistence model on CSI. It assumes the nowcast CSI time-series is equal to the most recent available CSI measurements. Given the time

parameters ($\mathcal{H}^h$, $\mathcal{R}^r$), the SP model can be written as:

$$\begin{pmatrix} \hat{y}_{t+r} & \hat{y}_{t+2r} & \cdots & \hat{y}_{t+h} \end{pmatrix} = \begin{pmatrix} y_{t+r-h} & y_{t+2r-h} & \cdots & y_t \end{pmatrix}. \tag{4.11}$$

For instance, given the exemplified operating timeline in Figure 4.2, the measured CSI values at DH4 from 7:29:46 to 7:30:00 will be directly used as nowcasts for the period of 7:30:01-7:30:15.

OLS regression

Regression is a statistical process for estimating the relationships among variables. Mathematically, a linear regression model is given by:

$$y_i = \boldsymbol{x}_i^\top \boldsymbol{\beta} + \varepsilon_i, \tag{4.12}$$

where $i = 1, \cdots, n$, $\boldsymbol{\beta}$ are the regression coefficients, $\boldsymbol{x}$ are the predictors, y is the response, and ε_i is a zero-mean, homogeneous error. Given training data $\boldsymbol{X}_{\text{train}}$ and $\boldsymbol{y}_{\text{train}}$, the OLS estimator for $\boldsymbol{\beta}$ is

$$\boldsymbol{\beta}^{\text{OLS}} = \underset{\boldsymbol{\beta}}{\texttt{argmin}} \; ||\boldsymbol{X}_{\text{train}}\boldsymbol{\beta} - \boldsymbol{y}_{\text{train}}||_2^2. \tag{4.13}$$

Lasso regression

Instead of minimizing the sum of squared errors as in OLS, the lasso regression penalizes the residual sum of squares using the ℓ_1-penalty, which often shrinks the regression coefficients of uncorrelated predictors to exactly zero. The lasso estimator is given as

$$\boldsymbol{\beta}^{\text{lasso}} = \underset{\boldsymbol{\beta}}{\texttt{argmin}} \; ||\boldsymbol{X}_{\text{train}}\boldsymbol{\beta} - \boldsymbol{y}_{\text{train}}||_2^2 + \lambda||\boldsymbol{\beta}||_1, \tag{4.14}$$

where λ is a tuning parameter that regulates the strength of the penalty. In this work, λ is selected using k-fold cross validation.

kNN regression

As one of the simplest pattern matching methods, a kNN model uses its training set as a feature space. By specifying a query vector of current feature state, the k closest samples will be identified based on some distance measure, e.g., Euclidean distance. In a regression task, the k neighbors are unified to produce a collective prediction.

In the present context, the row vectors in $\boldsymbol{X}_{\text{test}}$, i.e., the preselected p predictors are used as queries during pattern matching. For a row vector in the test set, $\boldsymbol{x}_*^\top$, the Euclidean distances between $\boldsymbol{x}_*^\top$ and each row of $\boldsymbol{X}_{\text{train}}$ are computed, and the CSI observations that correspond to the k smallest Euclidean distances are selected. Subsequently, based on the individual responses from the k neighbors, $\tilde{y}_1, \tilde{y}_2, \cdots, \tilde{y}_k$, a collective prediction can be formed by:

$$\hat{y} = \frac{\sum_{i=1}^{k} \beta_i \tilde{y}_i}{\sum_{i=1}^{k} \beta_i}, \tag{4.15}$$

where

$$\beta_i = \frac{1 - d_i}{d_{\max} - d_{\min}}, \tag{4.16}$$

is the weight for the i^{th} neighbor, d_i is the Euclidean distance between the i^{th} neighbor and $\boldsymbol{x}_*^\top$, and $d_{\max}$ and $d_{\min}$ are the maximum and minimum values among the Euclidean distance set respectively. In this work, the value of k is set to be 30, which follows the choice in [96].

4.3.3 Models for Probabilistic Nowcasting

Since the aforementioned SP, OLS, and kNN methods can also be extended to probabilistic models, we herein adopt the same three methods for probabilistic nowcasting, namely, persistence ensemble (PeEn), OLS with normal predictive distribution, and analog ensemble (AnEn). In addition, as a representative of nonparametric approaches, the quantile regression (QR) is considered as well. In this work, the results of probabilis-

tic nowcasting are recorded in terms of quantiles. For each model, a set of 19 quantiles is considered:

$$Q = \{0.05, 0.1, 0.15, \cdots, 0.9, 0.95\}. \tag{4.17}$$

Persistence ensemble

Whereas the SP model considers the most recent available CSI values as nowcasts, PeEn takes the CSI values recorded at N most recent available timestamps to create an ensemble, and thus an empirical predictive distribution. Following the operating timeline, PeEn in this work uses the last 1-h measurements at DH4 to form an ensemble, i.e., $N = 3600/r$, where r is the required nowcast resolution. For example, given the submission process depicted in Table 4.2, instead of submitting a single nowcast at 7:30:01, which equals to the CSI of DH4 at 7:29:46, 1-h of CSI values at DH4 from 6:29:46 to 7:29:46 ($N = 3600$) are used to produce an empirical distribution. Based on the predictive distribution, the 19 quantiles are then estimated.

OLS with normal predictive distribution

Recall the OLS estimator in Equation (4.13), with a new sample of predictors, $\boldsymbol{x}_*$, the predicted response is:

$$\hat{y}_* = \boldsymbol{x}_*^\top \boldsymbol{\beta}^{\mathrm{OLS}}. \tag{4.18}$$

For OLS, the variance of this prediction can be formulated by:

$$\mathbb{V}(\hat{y}_*) = s^2 \cdot \boldsymbol{x}_*^\top (\boldsymbol{X}_{\mathrm{train}}^\top \boldsymbol{X}_{\mathrm{train}})^{-1} \boldsymbol{x}_*, \tag{4.19}$$

where s^2 is an unbiased estimate of sample error variance $\mathbb{V}(\varepsilon_*)$. Subsequently, since the prediction error $\hat{\varepsilon}_* = y_* - \hat{y}_*$, the variance of $\hat{\varepsilon}_*$ can be written as:

$$\begin{aligned}\mathbb{V}(\hat{\varepsilon}_*) &= \mathbb{V}(\varepsilon_*) + \mathbb{V}(\hat{y}_*) \\ &= s^2 + s^2 \cdot \boldsymbol{x}_*^\top (\boldsymbol{X}_{\text{train}}^\top \boldsymbol{X}_{\text{train}})^{-1} \boldsymbol{x}_*.\end{aligned} \tag{4.20}$$

The probabilistic OLS nowcasts are obtained by assuming the errors follow a normal distribution. In that, the q^{th} quantile of the probabilistic nowcast can be found by $\hat{y}_* + z_q\sqrt{\mathbb{V}(\hat{\varepsilon}_*)}$, where z_q is the q^{th} quantile of a standard normal distribution. For instance, when $q = 0.5$, $z_q = 0$; and when $q = 0.95$, $z_q = 1.645$.

Analog ensemble

The term "analog" was firstly used in [97] to describe the weather patterns that resemble each other. In the solar forecasting literature, AnEn is widely designed to cooperate with NWP forecasts, due to its fast computation speed [98]. From a machine learning context, AnEn can be thought as a probabilistic version of kNN method. Given a feature query, AnEn searches the k most relevant analogs in the database. Then these analogs are used jointly to form a predictive distribution. In this work, the similar 30 neighbors identified by kNN are also used for AnEn, and the 19 quantiles are generated based on the estimated distribution.

Quantile regression

Unlike OLS regression that provides a framework for estimating conditional mean models, QR defines a regression model based on conditional quantile functions. Given the q^{th} conditional quantile, QR solves:

$$\boldsymbol{\beta}^{\text{QR}} = \underset{\boldsymbol{\beta}}{\texttt{argmin}} \sum_{i=1}^{n} \rho_q(y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}), \tag{4.21}$$

where $\rho_q(u) = u(q - I_{u<0})$ is a piecewise linear function for some $q \in (0, 1)$. When $u > 0$, $\rho_q(u)$ has a gradient q, whereas when $u < 0$, $\rho_q(u)$ has a gradient $q - 1$. To produce the

19 quantiles, each element in the quantile set Q is fed to the QR model. Hence, a total of 19 QR models are built at each model updating stage.

4.3.4 Evaluation Metrics

Metrics for deterministic nowcasting

Three metrics are used in this work to evaluate the deterministic nowcasts, namely, the normalized mean bias error (nMBE), nRMSE, and FS. Whereas nMBE is used to assess the systematic bias in the nowcasts, nRMSE is used to evaluate whether the nowcasts contain large errors. Finally, FS is used to determine the improvement of each model over the reference model, in this case, the SP model. Besides the nRMSE and FS defined in Equations (3.16) and (3.17), the nMBE is give by:

$$\text{nMBE} = \frac{\frac{1}{n'}\sum_{t=1}^{n'}(\hat{P}_t - P_t)}{\frac{1}{n'}\sum_{t=1}^{n'} P_t} \times 100\%. \tag{4.22}$$

where n' is the number of nowcasts, and $\hat{P}_t$ and P_t are the predicted and measured PV power at time t respectively.

Metrics for probabilistic nowcasting

To evaluate the probabilistic nowcasts, the prediction interval coverage probability (PICP), prediction interval normalized average width (PINAW), and normalized continuous ranked probability score (nCRPS) are used.

The PICP is given by:

$$\text{PICP} = \frac{1}{n'}\sum_{t=1}^{n'} \epsilon_t \times 100\%, \tag{4.23}$$

where ϵ_t equals to 1 if the observation at time t falls inside the lower and upper bounds of the prediction interval (PI); otherwise ϵ_t is 0. PICP quantifies the reliability of nowcasts, and is evaluated at a nominal coverage probability, i.e., $(1-\alpha)\times 100\%$. Theoretically, PICP should be greater than $(1-\alpha)\times 100\%$ so that the probabilistic nowcasts can

be considered as reliable. In this work, $\alpha = 0.1$ is considered, which corresponds to a nominal coverage probability of 90%, that is, the 90% PI. Given the quantiles in Eq. (4.17), the 90% PI is obtained based on the probabilistic nowcasts at $q = 0.05$ and $q = 0.95$.

The PINAW quantifies the average width of PI, which is defined as:

$$\text{PINAW} = \frac{1}{n'} \frac{\sum_{t=1}^{n'} (U_t - L_t)}{P_{\max} - P_{\min}} \times 100\%, \tag{4.24}$$

where U_t and L_t denote the upper and lower bounds of the PI respectively, and $P_{\max}$ and $P_{\min}$ are the maximum and minimum values of observations. Generally, a smaller PINAW corresponds to a narrower predictive distribution, which is desirable as the nowcasts becomes sharper. Nonetheless, since a nowcast can be sharp but completely exclude the target, solely using PINAW could be unfaithful to indicate the nowcasting quality. On this point, PINAW should be interpreted together with PICP.

Lastly, nCRPS compares the cumulative distribution function of the probabilistic nowcasts and observations, which is given by:

$$\text{nCRPS} = \frac{\frac{1}{n'} \sum_{t=1}^{n'} \int_0^{\infty} (F^{\hat{P}_t}(x) - \mathbb{1}(x - P_t))^2 dx}{\frac{1}{n'} \sum_{t=1}^{n'} P_t}, \tag{4.25}$$

where $F^{\hat{P}_t}$ is the cumulative distribution function of the predicted $\hat{P}_t$, and $\mathbb{1}(x - y_t)$ is the Heaviside step function shifted to P_t. From the view of deterministic nowcasting, nCRPS reduces to mean absolute percentage error. In this sense, a lower nCRPS indicates a more accurate probabilistic nowcast.

4.4 Empirical Study

In this section, following the operation design described in Section 4.2, a total of four case studies are presented. The first three case studies aim to evaluate the effect of time parameters on operational solar nowcasting and PAPC performance. Both deterministic and probabilistic nowcasts are applied to various PAPC operating scenarios. In the

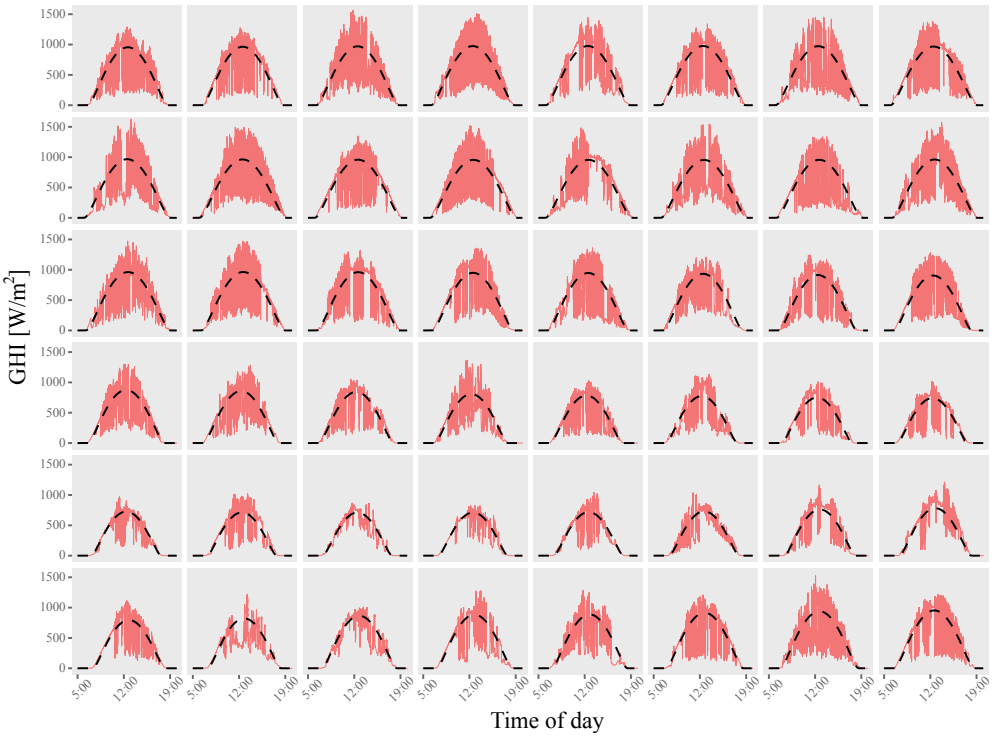


Figure 4.3: Visualization of the 48 validation days. GHI is plotted using red solid lines. The Ineichen-Perez clear sky irradiance is plotted in black dashed lines.

last study, PAPC operation is extended to three different PV systems, to quantify the impacts of system capacities. Throughout the four case studies, 48 cloudy days are identified from the whole dataset (see Figure 4.3), which contributes to over 1.7 million data points for validation.[5] To perform PAPC, the empirical validation considers a control sampling time of 1 s, and a RR limit of 25 kW/s. Furthermore, since PAPC performance can be highly dependent on nowcasting accuracy, isolating the nowcast error is of interest. On this point, PAPC with perfect nowcasts is also demonstrated, denoted as Perf.

[5]Since PAPC mainly works for cloudy days, we herein only opt the days with rapid irradiance variations. However, as rainfall can also cause rapid irradiance changes, days with high humidity are excluded. The overall data is 10-h operation × 3600 nowcasts per hour × 48 days = 1,728,000.

4.4.1 Case Study 1: The Effect of Various Forecast Horizons

In this case study, we evaluate the effect of forecast horizons on operational solar nowcasting and PAPC. The 10, 15, 30, and 60-s ahead nowcasts are produced under the nowcasting setting ($\mathcal{R}^{1s}, \mathcal{M}^{5min}$). Based on Equation (2.13), data from DH4 are used to mimic a practical PV system in Sesma, Spain, with a capacity of 0.99 MW and an area of 4.2 Ha [69]. In order to assess the performance of probabilistic nowcasts on PAPC, the metrics RSR and ECR are calculated using the lower bounds of 10%, 50%, and 90% PIs, respectively (See Appendix B for more details). The results are displayed in Table 4.3.

In terms of deterministic nowcasting, it can be seen from Table 4.3 that in general, all models show increasing nMBE and nRMSE errors as the forecast horizon expands. Nonetheless, the OLS, lasso, and kNN models show the highest FS for $\mathcal{H}^{15s}$. This may be due to the fact that during the 48 validation days, the network is dominated by trade winds with an angle of approximately 60° from north, and an average wind speed from 5 m/s to 11 m/s. As a result, the most relevant predictors would become DH5 with lags from 10 to 20.[6] On this point, once the forecast horizon is greater than 30 s, abundant irrelevant predictors enter the models, and it becomes more difficult to approach the optimized predictors, thus the model accuracy deteriorates rapidly. In this case, lasso outperforms the other methods thanks to its ℓ_1-penalty (see the $\mathcal{H}^{60s}$ case), from which additional predictor filtering is available.

In terms of probabilistic nowcasting, all models show increasing PINAW and nCRPS as the forecast horizon increases. In the other words, the sharpness and accuracy of the probabilistic nowcast is decreasing. As to calibration, PICPs of OLS and QR show a similar tendency as FS in deterministic nowcasts, i.e., the PICP increases before forecast horizon reaches 30 s, and decreases for the 30-s and 60-s ahead nowcasts, which is again, owing to the poor model predictability when $\mathcal{H} \geq 30$ s. On the other hand, PeEn and AnEn have relatively flat PICPs across all the horizons. This is due to the fact that

[6] Recall the spatial distance between DH4 and DH5, i.e., 103 m, it yields a temporal lag from 10.3 s to 20.6 s.

Table 4.3: Performance of operational nowcasting and PAPC under $(\mathcal{R}^{1s}, \mathcal{M}^{5min})$, but with various forecast horizons. The metrics are presented as daily "mean ± standard deviation". The column-wise best results are in bold with gray shade.

		Deterministic nowcasting metrics			PAPC metrics			
					RSR [%]		ECR [%]	
$\mathcal{H}$ [s]	Method	nMBE [%]	nRMSE [%]	FS [%]	PRED	PERF	PRED	PERF
10	SP	**0.00 ± 0.01**	9.47 ± 2.17	0.00 ± 0.00	86.99 ± 4.74	87.88 ± 4.59	2.49 ± 0.78	0.08 ± 0.07
	OLS	0.29 ± 0.37	**7.22 ± 3.12**	**22.85 ± 28.51**	89.99 ± 4.05	87.88 ± 4.59	2.07 ± 0.96	0.08 ± 0.07
	lasso	-0.23 ± 0.42	7.72 ± 3.20	17.03 ± 30.71	90.82 ± 3.85	87.88 ± 4.59	**2.05 ± 0.81**	0.08 ± 0.07
	kNN	-0.19 ± 0.45	8.46 ± 3.12	8.94 ± 30.22	**90.94 ± 3.71**	87.88 ± 4.59	2.21 ± 0.82	0.08 ± 0.07
15	SP	**0.01 ± 0.02**	13.38 ± 3.06	0.00 ± 0.00	96.45 ± 1.56	95.64 ± 1.78	3.54 ± 1.09	0.08 ± 0.07
	OLS	0.78 ± 0.52	**8.74 ± 2.79**	**34.38 ± 15.01**	96.59 ± 1.72	95.64 ± 1.78	2.80 ± 1.00	0.08 ± 0.07
	lasso	-0.30 ± 0.48	9.54 ± 2.79	27.95 ± 16.25	97.09 ± 1.88	95.64 ± 1.78	**2.73 ± 0.88**	0.08 ± 0.07
	kNN	-0.15 ± 0.53	10.32 ± 2.79	21.72 ± 17.25	**97.29 ± 1.70**	95.64 ± 1.78	2.91 ± 0.91	0.08 ± 0.07
30	SP	**-0.01 ± 0.04**	22.52 ± 5.19	0.00 ± 0.00	98.26 ± 1.17	97.67 ± 1.39	6.14 ± 1.85	0.09 ± 0.07
	OLS	3.14 ± 1.61	**17.85 ± 4.18**	**20.57 ± 6.84**	98.52 ± 1.15	97.67 ± 1.39	6.82 ± 2.33	0.09 ± 0.07
	lasso	-0.74 ± 0.84	18.94 ± 4.19	15.42 ± 7.08	**99.17 ± 0.72**	97.67 ± 1.39	**6.10 ± 1.98**	0.09 ± 0.07
	kNN	-0.24 ± 1.11	20.35 ± 4.40	9.03 ± 8.51	99.11 ± 0.71	97.67 ± 1.39	6.56 ± 1.91	0.09 ± 0.07
60	SP	**-0.02 ± 0.07**	33.24 ± 7.90	0.00 ± 0.00	98.24 ± 1.18	98.20 ± 1.53	9.60 ± 2.89	0.09 ± 0.07
	OLS	6.44 ± 2.51	30.87 ± 7.08	6.88 ± 4.08	98.47 ± 1.05	98.20 ± 1.53	12.84 ± 3.86	0.09 ± 0.07
	lasso	-1.05 ± 1.18	**29.81 ± 6.90**	**10.08 ± 4.44**	98.90 ± 1.26	98.20 ± 1.53	**10.23 ± 2.99**	0.09 ± 0.07
	kNN	-0.38 ± 1.56	31.18 ± 7.04	5.84 ± 5.92	**98.98 ± 0.81**	98.20 ± 1.53	10.71 ± 2.80	0.09 ± 0.07

		Probabilistic nowcasting metrics ($\alpha = 0.1$)			PAPC metrics					
					RSR [%]			ECR [%]		
$\mathcal{H}$ [s]	Method	PICP [%]	PINAW [%]	nCRPS [%]	10% PI	50% PI	90% PI	10% PI	50% PI	90% PI
10	PeEn	88.59 ± 4.32	56.42 ± 5.77	20.55 ± 6.08	**95.39 ± 4.96**	**97.43 ± 3.47**	**99.76 ± 0.62**	16.93 ± 10.03	30.79 ± 10.35	55.48 ± 6.30
	OLS	**95.65 ± 1.78**	19.54 ± 4.41	4.02 ± 1.42	90.61 ± 3.98	93.13 ± 3.62	93.04 ± 3.98	3.06 ± 1.24	8.60 ± 2.73	19.26 ± 5.48
	AnEn	84.48 ± 5.79	**11.45 ± 2.87**	3.35 ± 1.18	89.54 ± 4.73	90.83 ± 5.11	93.29 ± 5.05	2.53 ± 0.92	4.98 ± 1.69	**11.45 ± 3.61**
	QR	90.57 ± 3.64	15.23 ± 4.32	**2.93 ± 1.20**	89.81 ± 4.53	90.69 ± 4.50	97.51 ± 2.98	**1.93 ± 0.84**	**2.66 ± 1.07**	15.71 ± 5.25
15	PeEn	88.65 ± 4.16	56.43 ± 5.77	20.60 ± 6.10	**98.43 ± 1.55**	**99.06 ± 1.35**	**99.91 ± 0.22**	16.96 ± 10.04	30.81 ± 10.35	55.50 ± 6.30
	OLS	**95.95 ± 1.70**	25.04 ± 6.30	5.02 ± 1.42	97.18 ± 1.52	97.89 ± 1.24	97.10 ± 1.32	4.15 ± 1.36	11.29 ± 3.29	24.91 ± 6.91
	AnEn	84.26 ± 6.20	**15.32 ± 4.84**	4.06 ± 1.19	96.10 ± 2.01	97.03 ± 1.70	98.12 ± 1.39	2.95 ± 1.02	6.03 ± 2.03	**15.50 ± 5.12**
	QR	90.64 ± 3.51	20.08 ± 6.70	**3.66 ± 1.15**	96.22 ± 2.14	96.59 ± 2.41	99.21 ± 1.06	**2.40 ± 0.85**	**3.46 ± 1.32**	20.90 ± 7.14
30	PeEn	88.92 ± 3.62	56.44 ± 5.77	20.72 ± 6.14	**98.87 ± 1.26**	**99.21 ± 1.30**	**99.92 ± 0.18**	17.03 ± 10.07	30.88 ± 10.35	55.54 ± 6.31
	OLS	**94.32 ± 1.23**	42.35 ± 9.08	9.67 ± 2.44	98.29 ± 1.22	98.29 ± 0.85	98.18 ± 1.01	9.31 ± 2.95	21.65 ± 6.10	42.42 ± 9.87
	AnEn	83.24 ± 6.33	**29.44 ± 7.88**	8.25 ± 2.27	98.43 ± 1.19	98.71 ± 0.94	99.00 ± 0.85	6.14 ± 2.79	12.49 ± 4.58	**29.34 ± 6.96**
	QR	89.18 ± 3.38	38.80 ± 11.70	**7.94 ± 2.19**	98.20 ± 0.81	98.60 ± 0.87	99.87 ± 0.22	**5.14 ± 1.69**	**8.23 ± 3.84**	39.35 ± 10.11
60	PeEn	89.11 ± 3.25	56.47 ± 5.77	20.93 ± 6.22	**98.85 ± 1.27**	**99.19 ± 1.30**	**99.92 ± 0.18**	17.15 ± 10.13	30.99 ± 10.36	55.61 ± 6.30
	OLS	**90.73 ± 1.46**	55.90 ± 9.36	15.87 ± 3.82	98.55 ± 1.04	98.12 ± 1.03	98.32 ± 1.50	16.22 ± 4.61	32.49 ± 8.24	56.17 ± 9.77
	AnEN	82.65 ± 5.89	**38.13 ± 7.19**	**14.32 ± 3.92**	98.21 ± 1.39	98.28 ± 1.37	98.38 ± 1.94	11.32 ± 4.89	20.13 ± 6.22	**38.13 ± 6.21**
	QR	87.00 ± 3.27	58.02 ± 13.55	14.73 ± 3.83	98.39 ± 0.87	98.97 ± 0.77	99.84 ± 0.41	**10.07 ± 3.40**	**17.27 ± 7.60**	53.61 ± 8.48

PeEn and AnEn only generate nowcasts based on the past measurements, and no training refinement, e.g., predictor preselection, is involved. Among these models, OLS shows the best coverage, while AnEn is found to have the highest sharpness (with the smallest PINAW). The predictive distributions from PeEn, however, are not sharp. Since the goal of probabilistic forecasting is to maximize the sharpness of the predictive distribution subject to calibration, OLS seems to be the best option to generate decent probabilistic nowcasts. Notice that it is possible to decrease/increase the number of ensemble members to make PeEn/AnEn more reliable, at a cost of decreasing calibration/sharpness. However, such trade-off is beyond the scope of current paper.

Regarding PAPC, the results of Perf reveal that for ideal nowcasts, the control performance improves as forecast horizon increases, with increasing RSR and nearly unchanged ECR. The similar trend can be observed for both deterministic and probabilistic models across the $\mathcal{H}^{5s}$, $\mathcal{H}^{10s}$, and $\mathcal{H}^{15s}$ cases. However, when the forecast horizon is beyond model predictability, i.e., the $\mathcal{H}^{30s}$ and $\mathcal{H}^{60s}$ cases, the increasing nowcasting errors have led to a rapid growing in ECR. Furthermore, it is an interesting finding that all models produce even better RSR results than Perf. For deterministic models, the models with negative nMBE, i.e., lasso and kNN, are found most likely to have a larger RSR. This is because when the nowcasts underestimate the targets, it is identical to provide additional "reserves" for PAPC operation. In some cases, these reserved power could compensate the deficiency in a shorter forecast horizon, thus more ramps can be regulated.[7] In terms of probabilistic models, PINAW is found to be highly related to RSR. More specifically, a larger PINAW typically leads to a higher RSR. Moreover, it is observed that PAPC using larger PIs can significantly improve the RSR. This is now foreseeable since expanding PIs is analogous to increasing the amount of reversed power.

[7]Let us consider a 5-s ahead nowcasting example here. Given measurements {1000, 1000, 1000, 1000, 800} in kW, a sudden power drop of 200 kW/s occurs at $t = 5$ s. Under a RR limit of 25 kW/s, Perf outputs {1000, 975, 950, 925, 800}, which still leads to a 125 kW/s violation. To fully smooth the ramp, the forecast horizon needs to be expanded to 8 s. However, when a negatively biased nowcast is considered, e.g., 900 kW at $t = 1$ s, the output becomes {900, 875, 850, 825, 800}. In this case, the ramp can be successfully regulated.

4.4.2 Case Study 2: The Effect of Different Forecast Resolutions

In this case study, operational nowcasts and PAPC with various forecast resolutions are presented. The PAPC setting is identical to the previous study, namely, a RR limit of 25 kW/s is applied to the Sesma 0.99 MW system. Moreover, based on the results from case study 1, the nowcasting setting ($\mathcal{H}^{15s}, \mathcal{M}^{5min}$) is considered to ensure a proper model predictability. On this basis, data from DH4 are aggregated (summed up) into 1, 3, 5, and 15 s resolutions, from which fifteen-, five-, three-, and one-step-ahead nowcasts are generated at each nowcast run, respectively. Nonetheless, since PAPC requires a similar nowcasting interval as control sampling time, i.e., 1 s, the aggregated nowcasts need to be downscaled before PAPC use. To this end, this case study uses the stepwise-averaging method to disaggregate nowcasts. Simply speaking, each aggregated nowcast is downscaled by its mean value. It should be noticed that for each forecast resolution, the nowcasting metrics are calculated based on the corresponding aggregated dataset. The PAPC metrics, instead, are computed using the original DH4 data.[8] The results are depicted in Table 4.4.

It can be observed from Table 4.4 that both deterministic and probabilistic models show increasing nowcasting accuracy, i.e., smaller nRMSE and nCRPS, as the forecast resolution expands from 1 s to 5 s. In general, when a lower forecast resolution is considered (a larger value of $\mathcal{R}$), the temporal variability in solar time-series is greatly relieved. Hence, it is much easier to produce accurate nowcasts for a highly-aggregated dataset. However, it is also found that for $\mathcal{R}^{15s}$, the accuracy of all models, except for SP and PeEn, decreases rapidly. The poor performance of the 15-s resolution nowcasts may be due to the shrinkage of training data length. Recall the operating timeline in Figure 4.2, the $\mathcal{R}^{15s}$ case solely uses $3600/15 = 240$ data points to make inference, thus the model training could be divergent. On the other hand, SP and PeEn are found to benefit more from the temporal smoothing in large forecast resolutions, with increasing

[8] For example, given an original dataset $\mathcal{D} = \{1000\text{ W}, 900\text{ W}, 800\text{ W}\}$ in a resolution of 1 s, if the 3-s resolution nowcasts are of interest, the dataset needs to be aggregated into $\mathcal{D}^* = \{2700\text{ W}\}$. Provided a nowcast $\mathcal{P}^* = \{2400\text{ W}\}$, it should be further downscaled into $\mathcal{P} = \{800\text{ W}, 800\text{ W}, 800\text{ W}\}$ for PAPC use. While the nowcasting metrics are calculated using $\mathcal{D}^*$ and $\mathcal{P}^*$, the PAPC metrics are based on $\mathcal{D}$ and $\mathcal{P}$.

Table 4.4: Same as Table 4.3, but the results are computed for different forecast resolutions with $(\mathcal{H}^{15s}, \mathcal{M}^{5min})$. A larger value of $\mathcal{R}$ indicates a lower forecast resolution.

		Deterministic forecasting metrics			PAPC metrics			
					RSR [%]		ECR [%]	
$\mathcal{R}$ [s]	Method	nMBE [%]	nRMSE [%]	FS [%]	PRED	PERF	PRED	PERF
1	SP	**0.01 ± 0.02**	13.38 ± 3.06	0.00 ± 0.00	96.45 ± 1.56	95.64 ± 1.78	3.54 ± 1.09	0.08 ± 0.07
	OLS	0.78 ± 0.52	**8.74 ± 2.79**	**34.38 ± 15.01**	96.59 ± 1.72	95.64 ± 1.78	2.80 ± 1.00	0.08 ± 0.07
	lasso	-0.30 ± 0.48	9.54 ± 2.79	27.95 ± 16.25	97.09 ± 1.88	95.64 ± 1.78	**2.73 ± 0.88**	0.08 ± 0.07
	kNN	-0.15 ± 0.53	10.32 ± 2.79	21.72 ± 17.25	**97.29 ± 1.70**	95.64 ± 1.78	2.91 ± 0.91	0.08 ± 0.07
3	SP	**-0.01 ± 0.02**	13.26 ± 3.03	0.00 ± 0.00	96.81 ± 2.05	95.97 ± 1.81	3.12 ± 0.96	0.60 ± 0.21
	OLS	0.18 ± 0.37	**6.94 ± 3.12**	**47.17 ± 19.92**	97.00 ± 1.49	95.97 ± 1.81	2.01 ± 0.95	0.60 ± 0.21
	lasso	-0.09 ± 0.34	7.06 ± 3.19	46.24 ± 20.38	96.88 ± 1.60	95.97 ± 1.81	**1.98 ± 0.82**	0.60 ± 0.21
	kNN	-0.20 ± 0.43	9.16 ± 3.21	33.20 ± 22.15	**97.21 ± 1.49**	95.97 ± 1.81	2.29 ± 0.89	0.60 ± 0.21
5	SP	**-0.01 ± 0.02**	13.05 ± 2.98	0.00 ± 0.00	96.93 ± 1.85	96.24 ± 1.74	2.68 ± 0.82	1.08 ± 0.35
	OLS	0.17 ± 0.37	**6.88 ± 3.08**	**46.80 ± 19.79**	97.24 ± 1.57	96.24 ± 1.74	2.18 ± 0.94	1.08 ± 0.35
	lasso	-0.05 ± 0.32	6.97 ± 3.14	46.08 ± 20.12	97.29 ± 1.43	96.24 ± 1.74	**2.15 ± 0.81**	1.08 ± 0.35
	kNN	-0.21 ± 0.54	8.69 ± 3.18	29.41 ± 21.85	**97.67 ± 1.40**	96.24 ± 1.74	2.54 ± 0.92	1.08 ± 0.35
15	SP	**-0.01 ± 0.02**	11.23 ± 2.58	0.00 ± 0.00	97.21 ± 1.58	97.86 ± 1.27	**1.60 ± 0.20**	3.18 ± 0.97
	OLS	0.27 ± 0.26	**7.14 ± 2.74**	**36.14 ± 18.28**	98.00 ± 1.38	97.86 ± 1.27	3.30 ± 1.13	3.18 ± 0.97
	lasso	-0.03 ± 0.29	7.28 ± 2.76	34.93 ± 18.27	98.01 ± 1.37	97.86 ± 1.27	3.21 ± 1.07	3.18 ± 0.97
	kNN	-0.29 ± 1.43	11.77 ± 3.06	-7.06 ± 23.85	**98.29 ± 1.31**	97.86 ± 1.27	3.69 ± 1.59	3.18 ± 0.97

		Probabilistic forecasting metrics ($\alpha = 0.1$)			PAPC metrics					
					RSR [%]			ECR [%]		
$\mathcal{R}$ [s]	Method	PICP [%]	PINAW [%]	nCRPS [%]	10% PI	50% PI	90% PI	10% PI	50% PI	90% PI
1	PeEn	88.65 ± 4.16	56.43 ± 5.77	20.60 ± 6.10	**98.43 ± 1.55**	**99.06 ± 1.35**	**99.91 ± 0.22**	16.96 ± 10.04	30.81 ± 10.35	55.50 ± 6.30
	OLS	**95.95 ± 1.70**	25.04 ± 6.30	5.02 ± 1.42	97.18 ± 1.52	97.89 ± 1.24	97.10 ± 1.32	4.15 ± 1.36	11.29 ± 3.29	24.91 ± 6.91
	AnEn	84.24 ± 6.20	**15.32 ± 4.84**	4.06 ± 1.19	96.10 ± 2.01	97.03 ± 1.70	98.12 ± 1.39	2.95 ± 1.02	6.03 ± 2.03	**15.50 ± 5.12**
	QR	90.64 ± 3.51	20.08 ± 6.70	**3.66 ± 1.15**	96.22 ± 2.14	96.59 ± 2.41	99.21 ± 1.06	**2.40 ± 0.85**	**3.46 ± 1.32**	20.90 ± 7.14
3	PeEn	88.57 ± 4.09	56.09 ± 5.77	20.57 ± 6.10	**98.43 ± 1.54**	**99.06 ± 1.35**	**99.91 ± 0.22**	16.91 ± 10.00	30.71 ± 10.36	55.35 ± 6.28
	OLS	**95.15 ± 1.65**	16.58 ± 3.86	3.75 ± 1.45	97.16 ± 1.44	97.44 ± 1.79	97.86 ± 1.35	2.88 ± 1.25	7.81 ± 2.75	17.36 ± 5.47
	AnEn	85.76 ± 4.61	**10.90 ± 2.47**	3.74 ± 1.21	96.68 ± 1.80	96.95 ± 1.97	97.06 ± 1.90	2.97 ± 1.03	5.44 ± 1.85	**11.23 ± 3.79**
	QR	89.84 ± 3.44	12.92 ± 3.63	**2.79 ± 1.21**	96.86 ± 1.67	97.01 ± 1.68	98.68 ± 1.39	**1.93 ± 0.85**	**2.60 ± 1.08**	14.18 ± 5.01
5	PeEn	88.42 ± 4.05	56.03 ± 5.83	20.53 ± 6.09	**98.43 ± 1.54**	**99.05 ± 1.36**	**99.91 ± 0.22**	16.85 ± 9.94	30.59 ± 10.34	55.18 ± 6.29
	OLS	**94.89 ± 1.59**	16.10 ± 3.96	3.66 ± 1.44	97.31 ± 1.49	97.56 ± 1.37	97.75 ± 1.50	3.00 ± 1.23	7.64 ± 2.72	16.67 ± 5.42
	AnEn	85.23 ± 4.68	**11.70 ± 2.41**	3.72 ± 1.26	97.16 ± 1.47	97.22 ± 1.71	96.84 ± 2.25	3.42 ± 1.24	5.99 ± 2.09	**11.75 ± 3.97**
	QR	89.34 ± 3.58	12.65 ± 3.62	**2.77 ± 1.20**	97.29 ± 1.58	97.36 ± 1.50	98.56 ± 1.39	**2.12 ± 0.86**	**2.79 ± 1.10**	13.76 ± 4.85
15	PeEn	87.84 ± 3.89	55.49 ± 5.88	20.22 ± 6.06	**98.50 ± 1.52**	**99.07 ± 1.37**	**99.93 ± 0.18**	16.53 ± 9.60	29.89 ± 10.29	54.11 ± 6.18
	OLS	**94.32 ± 1.94**	16.49 ± 3.87	3.73 ± 1.25	98.06 ± 1.39	98.08 ± 1.39	97.98 ± 1.40	4.08 ± 1.33	8.37 ± 2.50	16.84 ± 4.81
	AnEn	84.17 ± 5.09	17.61 ± 3.44	6.89 ± 1.63	97.99 ± 1.33	97.96 ± 1.54	98.20 ± 1.52	5.41 ± 2.81	9.03 ± 4.35	**16.86 ± 6.72**
	QR	88.73 ± 3.67	**13.94 ± 4.18**	**3.02 ± 1.08**	97.98 ± 1.46	98.12 ± 1.34	98.50 ± 1.28	**3.22 ± 1.13**	**4.14 ± 1.46**	14.68 ± 4.51

nowcasting accuracy across all scenarios.

In terms of PAPC, it can be concluded from the performance of PERF that both RSR and ECR increase for a lower forecast resolution. Similar to the previous finding, deterministic models with negative nMBE and probabilistic models with larger PINAW are found to produce higher RSR. Among these models, SP and PeEn show consistent improvements in PAPC across all resolutions. With little surprise, SP even outperforms PERF in the case of $\mathcal{R}^{15s}$, with a comparable RSR but a much smaller ECR. The reason could be traced to the downscaling step. As the forecast resolution increases, the downscaling error also increases. Since the downscaling method herein used is more analogous to a persistence model on the mean values, the error variance caused by downscaling can be smaller for the SP nowcasts.

4.4.3 Case Study 3: The Effect of Various Forecast Model Updating Rates

This case study evaluates the last time parameter in operational solar nowcasting, that is, the forecast model update rate. For each operation day, four model update rates are considered, i.e., 5, 15, 30, and 60 min. Moreover, to isolate the downscaling errors, the 15-s ahead 1-s resolution nowcasts are generated. Similar to the previous case studies, PAPC is implemented for the Sesma PV system with a RR limit of 25 kW/s. The results of nowcasting and PAPC are displayed in Table 4.5.

Compared with the forecast horizon and forecast resolution, the effect of forecast model update rate seems to be more straightforward—except for SP and PeEn, all models show monotonically decreasing nowcasting performance as $\mathcal{M}$ increases (a slower model updating rate). This is no surprise since the nowcast time series often carries considerably large amount of variabilities, in that, refreshing the nowcasting models in a higher frequency clearly produces more accurate nowcasts. On the other hand, as the SP and PeEn models do not involve any training process, varying the forecast model update rate reveals no effect on their performance.

As to PAPC, it is noted from Table 4.5 that the differences arising from various

Table 4.5: Same as Table 4.3, but the results are computed for different forecast model update rates with $(\mathcal{H}^{15s}, \mathcal{R}^{1s})$. Since the model update rate only affects training accuracy, the performance of PERF remains the same. In general, a larger value in $\mathcal{M}$ denotes a slower forecast model update rate.

		Deterministic forecasting metrics			PAPC metrics			
					RSR [%]		ECR [%]	
$\mathcal{M}$ [min]	Method	nMBE [%]	nRMSE [%]	FS [%]	PRED	PERF	PRED	PERF
5	SP	**0.01 ± 0.02**	13.38 ± 3.06	0.00 ± 0.00	96.45 ± 1.56	95.64 ± 1.78	3.54 ± 1.09	0.08 ± 0.07
	OLS	0.78 ± 0.52	**8.74 ± 2.79**	**34.38 ± 15.01**	96.59 ± 1.72	95.64 ± 1.78	2.80 ± 1.00	0.08 ± 0.07
	lasso	-0.30 ± 0.48	9.54 ± 2.79	27.95 ± 16.25	97.09 ± 1.88	95.64 ± 1.78	**2.73 ± 0.88**	0.08 ± 0.07
	kNN	-0.15 ± 0.53	10.32 ± 2.79	21.72 ± 17.25	**97.29 ± 1.70**	95.64 ± 1.78	2.91 ± 0.91	0.08 ± 0.07
15	SP	**0.01 ± 0.02**	13.38 ± 3.06	0.00 ± 0.00	96.45 ± 1.56	95.64 ± 1.78	3.54 ± 1.09	0.08 ± 0.07
	OLS	0.75 ± 0.51	**8.83 ± 2.71**	**33.57 ± 14.98**	96.47 ± 1.85	95.64 ± 1.78	2.80 ± 0.95	0.08 ± 0.07
	lasso	-0.41 ± 0.66	9.83 ± 2.84	25.61 ± 17.00	97.14 ± 1.90	95.64 ± 1.78	**2.75 ± 0.88**	0.08 ± 0.07
	kNN	-0.40 ± 0.74	10.95 ± 2.90	16.67 ± 19.20	**97.67 ± 1.39**	95.64 ± 1.78	3.00 ± 0.90	0.08 ± 0.07
30	SP	**0.01 ± 0.02**	13.38 ± 3.06	0.00 ± 0.00	96.45 ± 1.56	95.64 ± 1.78	3.54 ± 1.09	0.08 ± 0.07
	OLS	0.71 ± 0.48	**8.90 ± 2.58**	**32.83 ± 14.49**	96.60 ± 1.88	95.64 ± 1.78	2.80 ± 0.90	0.08 ± 0.07
	lasso	-0.57 ± 0.80	10.09 ± 2.98	23.00 ± 22.00	97.25 ± 1.88	95.64 ± 1.78	**2.73 ± 0.82**	0.08 ± 0.07
	kNN	-0.49 ± 0.82	11.31 ± 2.85	13.31 ± 21.94	**97.29 ± 1.63**	95.64 ± 1.78	3.13 ± 0.85	0.08 ± 0.07
60	SP	**0.01 ± 0.02**	13.38 ± 3.06	0.00 ± 0.00	96.45 ± 1.56	95.64 ± 1.78	3.54 ± 1.09	0.08 ± 0.07
	OLS	0.69 ± 0.58	**9.31 ± 2.65**	**29.16 ± 17.42**	96.51 ± 1.97	95.64 ± 1.78	2.90 ± 0.94	0.08 ± 0.07
	lasso	-0.69 ± 0.81	10.19 ± 2.93	22.08 ± 20.84	97.04 ± 1.83	95.64 ± 1.78	**2.80 ± 0.94**	0.08 ± 0.07
	kNN	-1.05 ± 1.26	12.12 ± 3.43	4.02 ± 26.30	**97.17 ± 1.77**	95.64 ± 1.78	3.26 ± 1.10	0.08 ± 0.07

		Probabilistic forecasting metrics ($\alpha = 0.1$)			PAPC metrics					
					RSR [%]			ECR [%]		
$\mathcal{M}$ [min]	Method	PICP [%]	PINAW [%]	nCRPS [%]	10% PI	50% PI	90% PI	10% PI	50% PI	90% PI
5	PeEn	88.65 ± 4.16	56.43 ± 5.77	20.60 ± 6.10	**98.43 ± 1.55**	**99.06 ± 1.35**	**99.91 ± 0.22**	16.96 ± 10.04	30.81 ± 10.35	55.50 ± 6.30
	OLS	**95.95 ± 1.70**	25.04 ± 6.30	5.02 ± 1.42	97.18 ± 1.52	97.89 ± 1.24	97.10 ± 1.32	4.15 ± 1.36	11.29 ± 3.29	24.91 ± 6.91
	AnEn	84.24 ± 6.20	**15.32 ± 4.84**	4.06 ± 1.19	96.10 ± 2.01	97.03 ± 1.70	98.12 ± 1.39	2.95 ± 1.02	6.03 ± 2.03	**15.50 ± 5.12**
	QR	90.64 ± 3.51	20.08 ± 6.70	**3.66 ± 1.15**	96.22 ± 2.14	96.59 ± 2.41	99.21 ± 1.06	**2.40 ± 0.85**	**3.46 ± 1.32**	20.90 ± 7.14
15	PeEn	88.65 ± 4.16	56.43 ± 5.77	20.60 ± 6.10	**98.43 ± 1.55**	**99.06 ± 1.35**	**99.91 ± 0.22**	16.96 ± 10.04	30.81 ± 10.35	55.50 ± 6.30
	OLS	**95.51 ± 2.00**	24.95 ± 6.25	5.05 ± 1.36	96.88 ± 2.11	97.70 ± 1.28	97.45 ± 1.07	4.14 ± 1.30	11.25 ± 3.20	24.80 ± 6.78
	AnEn	82.72 ± 7.01	**16.02 ± 5.04**	4.46 ± 1.25	95.99 ± 2.26	97.01 ± 1.87	98.08 ± 1.45	3.06 ± 1.07	6.20 ± 2.10	**16.08 ± 5.17**
	QR	89.79 ± 4.10	20.01 ± 6.68	**3.71 ± 1.13**	96.08 ± 2.02	96.60 ± 2.41	99.27 ± 1.01	**2.41 ± 0.83**	**3.46 ± 1.30**	20.82 ± 7.00
30	PeEn	88.65 ± 4.16	56.43 ± 5.77	20.60 ± 6.10	**98.43 ± 1.55**	**99.06 ± 1.35**	**99.91 ± 0.22**	16.96 ± 10.04	30.81 ± 10.35	55.50 ± 6.30
	OLS	**95.03 ± 2.22**	24.77 ± 6.19	5.07 ± 1.31	96.84 ± 2.00	97.87 ± 1.19	97.41 ± 1.81	4.12 ± 1.25	11.17 ± 3.14	24.59 ± 6.65
	AnEn	81.00 ± 7.54	**16.51 ± 4.97**	4.73 ± 1.19	95.99 ± 2.25	97.22 ± 1.74	98.02 ± 1.74	3.17 ± 1.08	6.34 ± 2.08	**16.53 ± 5.03**
	QR	88.54 ± 4.62	19.91 ± 6.63	**3.73 ± 1.08**	96.17 ± 2.20	96.53 ± 2.26	99.19 ± 1.05	**2.41 ± 0.81**	**3.47 ± 1.28**	20.63 ± 6.84
60	PeEn	88.65 ± 4.16	56.43 ± 5.77	20.60 ± 6.10	**98.43 ± 1.55**	**99.06 ± 1.35**	**99.91 ± 0.22**	16.96 ± 10.04	30.81 ± 10.35	55.50 ± 6.30
	OLS	**93.50 ± 3.13**	25.20 ± 5.72	5.36 ± 1.25	96.94 ± 1.70	97.83 ± 1.33	97.53 ± 1.43	4.25 ± 1.24	11.44 ± 2.91	25.00 ± 6.17
	AnEn	78.63 ± 8.02	**17.52 ± 5.22**	5.40 ± 1.51	96.06 ± 2.14	97.26 ± 1.68	97.99 ± 1.44	3.38 ± 1.43	6.64 ± 2.34	**17.10 ± 5.36**
	QR	84.54 ± 6.20	19.47 ± 6.44	**3.88 ± 1.11**	96.20 ± 1.91	96.63 ± 1.83	99.29 ± 0.94	**2.48 ± 0.84**	**3.51 ± 1.29**	20.32 ± 6.69

forecast model updating rates are small. For the $\mathcal{M}^{5\text{min}}$ and $\mathcal{M}^{60\text{min}}$ cases, the PAPC metrics for both deterministic and probabilistic models typically show less than 1% variations. From the system operators' point of view, this is encouraging, since a slower model updating rate means a smaller memory storage, thus the better applicability.

4.4.4 Case Study 4: The Impact of PV Capacities on Operational Solar Nowcasting and PAPC

So far the studies have been dealing with various nowcasting settings. From a system operator's perspective, it is natural to wonder how the nowcasting and PAPC can behave on different PV systems. With this viewpoint, this case study moves the focus from the "extrinsic" nowcasting settings to an inherent property of PV systems, that is, system capacity. Besides the aforementioned 0.99 MW system in Sesma, we herein consider another two PV systems located in Milagro, Spain—a 0.143 MW system with an area of 0.63 Ha, and a 9.5 MW system with an area of 52 Ha [69]. In this case study, the nowcasts are generated under the nowcasting setting $(\mathcal{H}^{15\text{s}}, \mathcal{R}^{1\text{s}}, \mathcal{M}^{5\text{min}})$.

Figure 4.4 shows the deterministic nowcasting results for the three PV systems. A direct column-wise comparison leads to the conclusion that the nowcasts become more convergent to the measurements as system capacity increases. Such reduction in the spread of scatter plots indicates a decreasing nRMSE in the nowcasts. In the case of 0.143 MW system, the mean nRMSEs for SP, OLS, lasso, and kNN models are found as 20.65%, 12.38%, 12.86%, 13.70%. While for the 9.5 MW system, the mean nRMSE reduces to 7.27%, 5.81%, 6.76%, 7.37%, respectively. Nonetheless, the skills of all models (except for SP) are observed to decrease in larger PV systems. As the performance of solar nowcasting is dominated by the movements of cloud shadows, a larger PV system often suffers less from cloud transitions thanks to the well-known geographic smoothing effect [99]. In this regard, the system predicability improves, thus even a naive model, e.g., persistence model, is able to produce quite reliable results.

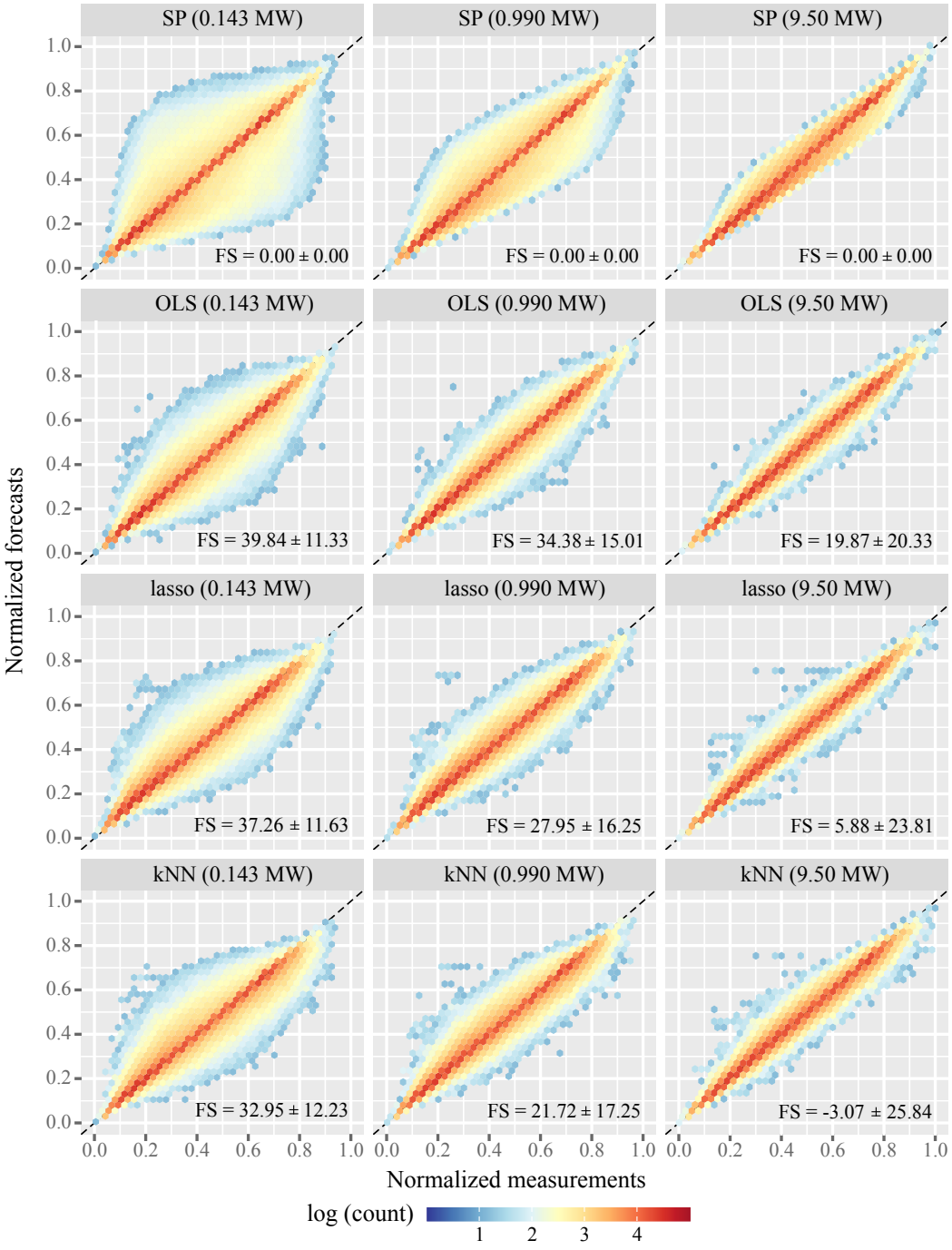


Figure 4.4: Scatter plots of deterministic nowcasts versus measurements for PV systems with different capacities. Hexagon binning algorithm is used for visualization; the color varies from red to lightblue while the number of scatter points per bin varies from high to low. Each plot is drawn based on 48 validation days. The daily means and standard deviations of FS are indicated on the plots.

In terms of probabilistic nowcasting, Figure 4.5 depicts the distributions of nCRPSs of the three PV systems. It is evident from Figure 4.5 that the probabilistic nowcasts of a larger PV system typically have a smaller mean and standard deviation of nCRPS. In addition, all probabilistic nowcasts show increasing sharpness as system capacity increases, with higher PICP but slightly wider PINAW. Among these methods, QR is found to have the best accuracy in nCRPS, whereas OLS again reveals the optimal balance between PICP and PINAW. It is also observed that the performance of AnEn improves rapidly when a larger system is considered. Compare the results of 9.5 MW case to the 0.143 MW case, the PICP of AnEn increases over 10%. Given AnEn often produces the sharpest nowcasts, it can be inferred that AnEn could be more preferable for utility-scale PV systems.

Figure 4.6 presents the resulting PAPC performance. Notice that owing to the rare cases of RR $>$ 25kW/s for the 0.143 MW system (less than 10 violations during the validation days), a RR limit of 10 kW/s is applied here instead. To allow an 'apples to apples' comparison between deterministic and probabilistic nowcasts, the 10% PI of probabilistic forecasts is used for PAPC. As Figure 4.6 shows, both RSR and ECR performance deteriorate substantially for larger PV systems, for all models. While a mean RSR $\sim$95% can be observed for all models in the 0.143 MW case, only PeEn reveals a mean RSR higher than 80% for the 9.5 MW system, at a cost of mean ECR greater than 20%. As system capacity increases, a larger PV system typically shows larger power variations in magnitude during irradiance changes, thus it becomes more difficult to follow the 10 kW/s RR limit. In addition, since it is much easier for larger systems to violate the RR limit, PAPC can be triggered more frequently. As a consequence, more power generation is curtailed. It is also observed that SP has generated more reliable PAPC results for larger PV systems. In the 9.5 MW case, SP becomes the best option among all methods, with the second highest RSR and low ECR. This again indicates that SP is able to benefit more from the system geographic smoothing. As a probabilistic version of SP, PeEn also reveals superiority in larger systems, where it shows the highest RSR as usual, but increasingly comparable ECR.

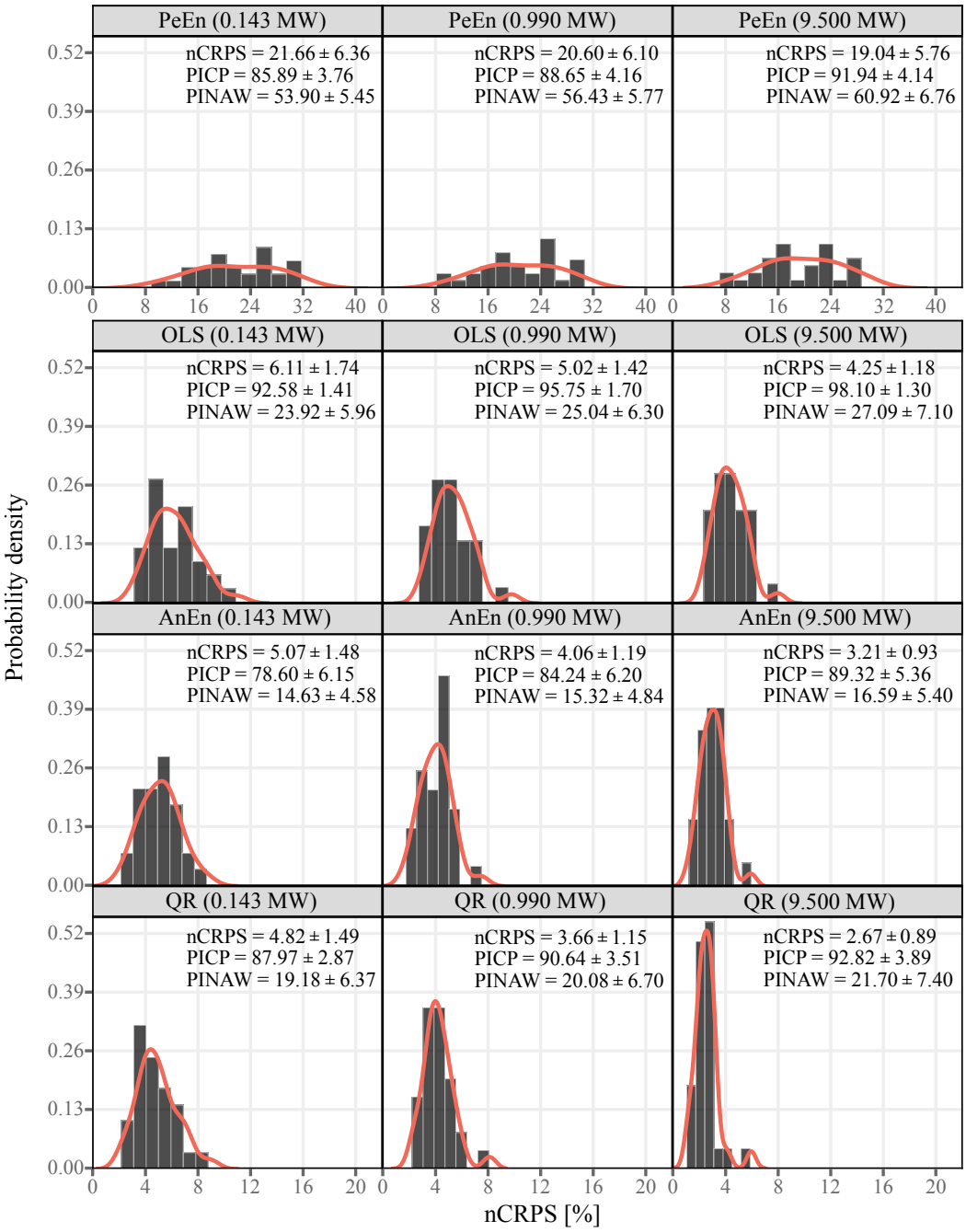


Figure 4.5: nCRPS distributions of the probabilistic nowcasts for PV systems with different capacities. Each distribution plot is based on the whole 48 validation days. The daily means and standard deviations of nCRPS, PICP, and PINAW are indicated on the plots.

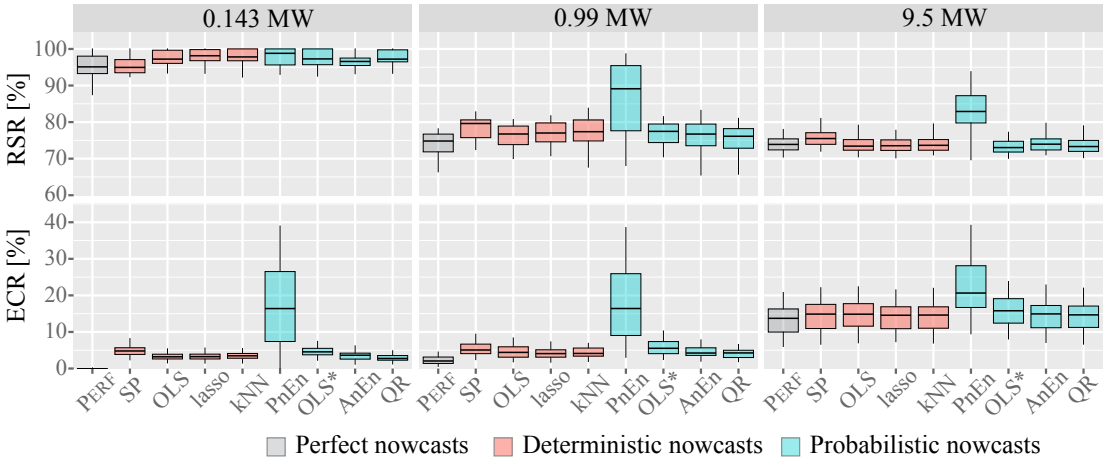


Figure 4.6: Boxplots for RSR and ECR of all models at the 3 PV systems. For probabilistic nowcasts, the lower bound of 10% PI is applied to PAPC. The probabilistic OLS method is annotated by OLS*.

4.5 Chapter Conclusion

This chapter presents a study on the operability of solar nowcasting, which demonstrates an operational solar nowcasting practice for PAPC. Three time parameters related to operational solar nowcasting are proposed, namely, forecast horizon, forecast resolution, and forecast model update rate. The effect of the three time parameters on operational solar nowcasting and PAPC are evaluated through four case studies. The main findings are summarized as follows:

- In general, the accuracy of both deterministic and probabilistic nowcasting decreases for a longer forecast horizon, a higher forecast resolution, and a slower forecast model updating rate. However, such tendency could disappear for the forecast horizon and resolution, as the skill of the nowcasting—the measure of goodness of the nowcasts against a reference model, is of interest. In that case, the effect of forecast horizon is found subject to the local wind speed, where all models show the highest skill as the forecast horizon matches the temporal lags among the predictors. In addition, due to the shrinkage of training data length, nowcasting produced at a lower resolution can show inferior skill to that at a

higher resolution.

- In the ideal case, i.e., assuming the nowcasts are perfect, the performance of PAPC improves as the forecast horizon increases, with a boost in RSR and nearly unchanged ECR. Moreover, both RSR and ECR are found to increase for a lower forecast resolution. On the other hand, the practical PAPC performance is observed highly dominated by the nowcasting results, especially in terms of nMAE (deterministic) and PINAW (probabilistic). More specifically, under the similar nowcasting skills, the negatively-biased (negative nMAE) deterministic nowcasts and the wider (a larger PINAW) probabilistic nowcasts are more likely to generate better PAPC results. Among the three time parameters, the forecast model updating rate is found to have the minimal impact on PAPC, with less than 1% variations between the 5-min and 60-min model updating rates.

- Besides the time parameters, the capacity of a PV system also reveals notable effects on the performance of operational nowcasting and PAPC. Generally, both deterministic and probabilistic models show increasing nowcasting accuracy as the system capacity increases, whereas the skills of all models are found to decrease. Due to the geographic smoothing effect, the nowcast time series of a larger PV system typically contains less variabilities, thus all models are able to produce comparable nowcasting results. In contrast, the performance of PAPC is found to deteriorate as system capacity increases. As the magnitude of the power variations increases in larger PV systems, it becomes more difficult to follow the required RR limit (in unit of power). However, it should be noticed that the situation will reverse when a RR limit is defined in percentage of capacity, in which a smaller PV system can suffer more from the RR violations.

Chapter 5

Emulation of RRs for Utility-Scale PV Systems During Cloud Transitions

5.1 Chapter Introduction

In Chapter 4, it has been shown that the nowcasting and PRRC performance can vary greatly with different PV systems. Underestimating PV system RRs may trigger failures of PRRC, thus violating RR regulations. Overestimating RRs may result in unnecessary deployment of ramping sources, leading to the expenditure of operational costs. In this sense, having the knowledge of system[1] RR characteristics becomes crucial for PV system operators.

As discussed in Section 1.1, current studies still show some deficiencies to investigate PV system RRs. The main limitations are:

- Exhaustive sensing of cloud shadow transitions both spatially and in time are unavailable for most existing utility-scale PV systems. Thus the studies using data from practical PV systems are typically limited to the analyses of endogenous

[1] Throughout this section, the term "system" refers to the PV system.

factors.

- Due to heavy computational burdens, most studies that are based on the computer-aided simulators are constrained to evaluate a single or several PV arrays. While a utility-scale system can be simulated by downscaling the simulation resolution, the simulation fidelity decreases. Moreover, difficulty also arises for integrating the natures of cloud shadows into simulations.

- The analytical modeling of PV power fluctuations, though more efficient, often contains certain simplifications of cloud shadings (e.g., assumes a statistically invariant and isotropic cloud field), which makes it inferior for studying the RR characteristics with respect to cloud shadow properties.

In this context, the goal of this work is to comprehensively investigate the RRs of utility-scale PV systems during cloud shadow transitions, for both endogenous and exogenous factors, and in resolution of module-level. For this purpose, this work first sets forth a partial shading emulator, which can emulate the module-level responses of utility-scale systems under passing clouds. In order to reproduce the cloud shadow natures, a fractal-based cloud shadow model is introduced. The cloud shadow model considers both the characteristics of irradiance transitions and the spatial diversities of cloud coverages. Based on the developed emulation tools, the effects of two endogenous factors, namely PV array arrangement and system orientation, and three exogenous factors, namely shadow intensity, shadow velocity, and shadow size on RRs are explored for a series of utility-scale PV systems ranging from 1 MW to 60 MW. The cloud shadow transitions identified from real measurements are also exploited to assess the characteristics of RRs in reality.

5.2 Methodology

This section elaborates the modeling process of utility-scale PV systems and cloud shadows. Based on the mathematical model of a single PV module, a partial shading

emulator is developed to emulate the behaviors of PV strings under partial shading, and a PV matrix is constructed to model the aggregation of PV arrays. The cloud shadow model is based on a modified fractal approach, which allows user-defined shadow properties to be synthesized. The synthesized shadow pattern is stored in a matrix form as well. A detailed description of the emulation process is presented as following.

5.2.1 Modeling of PV systems

Mathematical model of PV module

In this study, an experimentally verified mathematical model of PV module presented in [100] is adopted. The model is based on the well-known one-diode PV cell model that provides the following relationship between cell current and voltage:

$$I = I_{ph} - I_o(e^{\frac{V+R_sI}{AV_T}} - 1) - \frac{V + R_sI}{R_p}, \tag{5.1}$$

where I is the cell current, I_{ph} is the light-generated current that is affected by instantaneous irradiance value, I_o is the dark saturation current, V is the cell voltage, R_s and R_p are the series and parallel resistances, A is ideality factor, and V_T is the thermal voltage. The thermal voltage can be further represented by $V_T = kT/q$, where k is the Boltzmann constant, T is the cell temperature, and q is the electron charge. The model for a PV module is then obtained by scaling the parameters of Equation (5.1) by the number of series-connected cells. The PV module model takes two inputs, namely the irradiance value and module temperature, and outputs the corresponding I-V characteristic curve.

In the case of module-level emulation, each cell in a module is assumed to receive the uniform irradiance, and the voltage losses across internal bypass diodes are considered using a single diode (multiplied by the number of bypass diodes). Notice that although one PV cell should be the smallest unit that can be used for emulation, the differences arising from using a module as the basic unit have been found moderately small especially for partial shading analyses, since the scale of a cloud is typically much larger than a

Table 5.1: Specification of the JAM72S09 395/PR PV module at STC.

Parameter	Specification
Number of cells, N_s	72
Short-circuit current, I_{sc}	10.27 A
Open-circuit voltage, V_{oc}	49.64 V
Current at MPP, I_{mpp}	9.76 A
Voltage at MPP, V_{mpp}	40.48 V
Power at MPP, P_{mpp}	395.08 W
Temperature coefficient on PV current, K_i	0.06 %/°C
Temperature coefficient on PV voltage, K_v	-0.3 %/°C
Number of bypass diodes, N_{bypass}	3

module [55]. In this regard, PV module-level emulation should be accurate enough for the analysis presented in this work. The parameters of the PV module model are fitted to the characteristics of the JAM72S09 395/PR PV module, which is specially designed for utility-scale system integration. The specification of the PV module at standard test conditions (STC) is displayed in Table 5.1.

PV array topology

The internal array topology is critical for the performance of utility-scale PV systems. The series-parallel array topology with central inverters has been preferred by a majority of existing systems for its simplicity of installation and low cost. However, the main drawback comes as the use of single MPPT typically causes significant power mismatch losses in the partially shaded PV modules [101]. To counteract partial shading effects, a range of array topologies have been developed, such as multi-string, total-cross-tied, bridge-link, or honey-comb [102]. Among these, the multi-string topology has been suggested to be a big potential for future large-scale integration due to its flexibility [103] and commercial availability [104]. To this end, this study will mainly focus on the multi-string topology. Notice it has been reported in [105] that different array topologies can only contribute to small power fluctuation differences during cloud transitions. In this context, the results presented in this work can also be meaningful for evaluating

other array topologies.

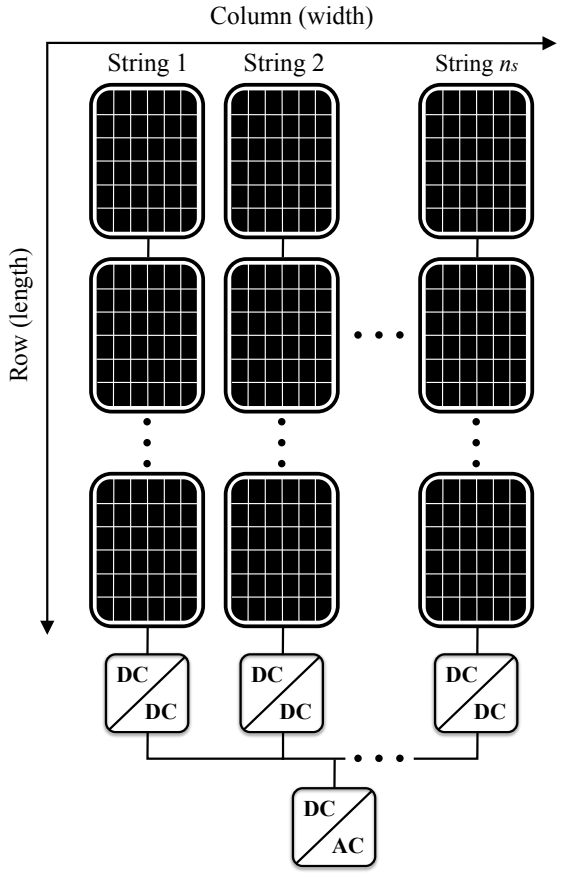


Figure 5.1: Electrical connections of an array of multi-string PV configuration.

Figure 5.1 shows the layout of multi-string array topology. Multiple modules are connected in series to form a PV string. The output of each string is modulated by an independent DC/DC converter, which is controlled for MPPT. The common DC bus is linked to the grid through a centralized DC/AC inverter. In this way, the partial shading effects are minimized at string level since each PV string is individually extracted and processed. Given a PV array composed of n_s strings in columns, with each string series-connected by n_m modules in rows, a multi-string PV system with $a \times b$ arrays

can be constructed in matrix form, given by:

$$\mathbf{\Pi} = \begin{pmatrix} \mathbf{\Gamma}_{11} & \mathbf{\Gamma}_{12} & \dots & \mathbf{\Gamma}_{1b} \\ \mathbf{\Gamma}_{21} & \mathbf{\Gamma}_{22} & \dots & \mathbf{\Gamma}_{2b} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{\Gamma}_{a1} & \mathbf{\Gamma}_{a2} & \dots & \mathbf{\Gamma}_{ab} \end{pmatrix}, \tag{5.2}$$

and

$$\mathbf{\Gamma}_{pq} = \begin{pmatrix} (\gamma_{11}^{pq}, \chi_{11}^{pq}) & (\gamma_{12}^{pq}, \chi_{12}^{pq}) & \dots & (\gamma_{1n_s}^{pq}, \chi_{1n_s}^{pq}) \\ (\gamma_{21}^{pq}, \chi_{21}^{pq}) & (\gamma_{22}^{pq}, \chi_{22}^{pq}) & \dots & (\gamma_{2n_s}^{pq}, \chi_{2n_s}^{pq}) \\ \vdots & \vdots & \ddots & \vdots \\ (\gamma_{n_m1}^{pq}, \chi_{n_m1}^{pq}) & (\gamma_{n_m2}^{pq}, \chi_{n_m2}^{pq}) & \dots & (\gamma_{n_mn_s}^{pq}, \chi_{n_mn_s}^{pq}) \end{pmatrix}, \tag{5.3}$$

where $\mathbf{\Pi} \in \mathbb{R}^{an_m \times bn_s}$ and $\mathbf{\Gamma}_{pq} \in \mathbb{R}^{n_m \times n_s}$, $p = 1, \cdots, a$, $q = 1, \cdots, b$, are the matrix representations of the entire PV system and an individual PV array respectively. Each element in the array matrix is a pair of variables, where γ_{ij}^{pq} and χ_{ij}^{pq} respectively denote the irradiance and temperature of the module in the ith row and jth column of the array $\mathbf{\Gamma}_{pq}$. The *I-V* characteristic of each module is then obtained by assigning corresponding γ_{ij}^{pq} and χ_{ij}^{pq} to the PV module model.

Emulating partial shading effects

The PV model presented in [100] only considers the characteristics of a single module. When it comes to an aggregation of modules such as a PV string or array, however, the aggregated system output can be limited by the shaded modules under passing clouds. In this case, the module-integrated bypass diodes provide an alternative current path, and both current and power curves of a PV string or array can exhibit multiple peaks [106]. For the studied multi-string configuration, distributed MPP trackers are integrated at string level, thus the characteristic of each string should be considered

individually.

To account for the partial shading effects, most simulator-based studies are dedicated to deriving the characteristics of bypass diodes using physical models or complex intelligent schemes. Although accurate, the complicated modeling process and heavy computations have limited these studies to small-scale PV systems (typically several PV arrays) [107]. At this stage, a partial shading emulator is developed, which does not require the explicit modeling of bypass diodes, and the behaviors of a partially shaded PV string can be emulated based on the I-V curves of series-connected modules. The principle of the emulator is based on two verified analytical statements presented in [108]: For a PV string consisting of n_m modules, with each module sharing a similar open-circuit voltage V_{oc} at STC, then

1) the PV string contains n_m regions, and separated by amount equal to V_{oc}, that is $((j-1)V_{oc},\ jV_{oc})$, where $j = 2, \cdots, n_m$;

2) the bypass diodes of shaded modules are activated near $V_{oc},\ 2V_{oc}, \cdots, (n_m - 1)V_{oc}$ of the string.

Algorithm 1 demonstrates the implementation of the proposed emulator, where the computer functions are written in `verbatim`. The emulator takes the I-V characteristics of n_m series-connected PV modules as inputs, denoted by:

$$\mathbf{I} = (\hat{I}_1 \cdots \hat{I}_{n_m})^\top \quad \mathbf{V} = (\hat{V}_1 \cdots \hat{V}_{n_m})^\top, \quad \mathbf{I}, \mathbf{V} \in \mathbb{R}^{n_m \times m}, \tag{5.4}$$

where vectors $\hat{I}_j = (i_{j1} \cdots i_{jm})$ and $\hat{V}_j = (v_{j1} \cdots v_{jm})$, $j = 1, \cdots, n_m$, represent the I-V characteristic of the jth PV module in the string. m is a user-defined value that declares the number of points needed in Equation (5.1) to find the I-V curve. In another word, each pair of (v_{jk}, i_{jk}), $k = 1, \cdots, m$, denotes a single point on the I-V curve of the jth module. We herein use $m = 100$ as suggested in [100]. Since a greater irradiance value always contributes to a higher peak on the I-V curve, the emulator first sorts $\mathbf{I}$ and $\mathbf{V}$ in descending order with respect to $\hat{I}_j[1]$, that is i_{j1}, leading to the sorted

Algorithm 1 Partial Shading Emulator

Input: *I-V* characteristics of n_m PV modules: $\mathbf{I} = (\hat{I}_1 \cdots \hat{I}_{n_m})^\top$ and $\mathbf{V} = (\hat{V}_1 \cdots \hat{V}_{n_m})^\top$, Module open-circuit voltage: V_{oc}.

Output: String *I-V* characteristic with n_m series-connected PV modules: $\widetilde{I}$ and $\widetilde{V}$.

$\widetilde{I} \leftarrow (\varnothing)$, $\widetilde{V} \leftarrow (\varnothing)$, $\hat{X} \leftarrow (\varnothing)$ ▷ Initialization
$\mathbf{I}^* \leftarrow \texttt{sort}(\mathbf{I})$ ▷ Sort $\mathbf{I}$ with respect to $I_j[1]$
$\mathbf{V}^* \leftarrow \texttt{sort}(\mathbf{V}, \mathbf{I}^*)$ ▷ Sort $\mathbf{V}$ with respect to $\mathbf{I}^*$
for $j = 1 : n_m$ **do** ▷ Loop to create voltage regions
    $\mathbf{V}^*[j] \leftarrow \mathbf{V}^*[j] + (j-1) \cdot V_{oc} + V_{\text{bypass}}$ ▷ Statement 1
    if $j \geq 2$ **then**
        $x \leftarrow \texttt{argmin}_x(\hat{I}_{j-1}[x] \leq \hat{I}_j[1])$
        $\hat{X}[-1] \leftarrow x$ ▷ Indexes of cross-points
    end if
end for
for $j = 2 : n_m - 1$ **do** ▷ Loop to combine the curves
    $\widetilde{V} \leftarrow \hat{V}_j[\hat{X}[j-1] : \hat{X}[j]]$ ▷ Statement 2
    $\widetilde{I} \leftarrow \hat{I}_j[\hat{X}[j-1] : \hat{X}[j]]$
end for
$\widetilde{V}[1 : \hat{X}[1]] \leftarrow \hat{V}_1[1 : \hat{X}[1]]$. ▷ The first voltage region
$\widetilde{I}[1 : \hat{X}[1]] \leftarrow \hat{I}_1[1 : \hat{X}[1]]$
$\widetilde{V}[\hat{X}[-1] : -1] \leftarrow \hat{V}_{n_m}[\hat{X}[-1] : -1]$ ▷ The last region
$\widetilde{I}[\hat{X}[-1] : -1] \leftarrow \hat{I}_{n_m}[\hat{X}[-1] : -1]$

vectors $\mathbf{I}^*$ and $\mathbf{V}^*$. Then refer to statement 1, all the elements in $\mathbf{V}^*$ are increased by $(j-1) \cdot V_{oc} + V_{\text{bypass}}$ to create n_m voltage regions, where V_{bypass} is the voltage drops across the bypass diodes in a module. In order to determine the activation points of bypass diodes, the emulator seeks the cross-points between each $\hat{I}_j$-$\hat{V}_j$ and $\hat{I}_{j-1}$-$\hat{V}_{j-1}$ curves, which gives the vector $\hat{X} = (x_1 \cdots x_{n_m-1})$. Each element x_j denotes the index or position of the bypass diode activation between the jth and $(j-1)$th module. Then according to statement 2, the slices $\hat{I}_j[x_{j-1} : x_j]$ and $\hat{V}_j[x_{j-1} : x_j]$, that is $(i_{jx_{j-1}} \cdots i_{jx_j})$ and $(v_{jx_{j-1}} \cdots v_{jx_j})$, are stripped and appended to vectors $\widetilde{I}$ and $\widetilde{V}$ respectively. Joined by the first and last voltage regions, the vectors $\widetilde{I}$ and $\widetilde{V}$ finally contribute to the emulated *I-V* characteristic for the entire PV string under partial shading.

An example of emulating a PV string with three modules under irradiance 1000 W/m^2, 800 W/m^2, 200 W/m^2, and temperature 25°C is illustrated in Figure 5.2. To

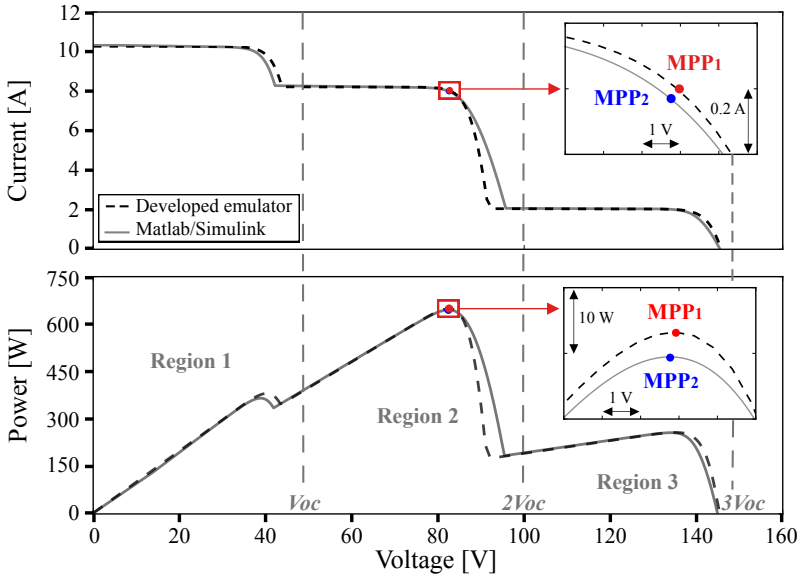


Figure 5.2: Emulated PV string characteristics under partial shading using Matlab/Simulink and developed emulator, (a) I-V curve. (b) P-V curve.

benchmark the emulations, a verified Matlab/Simulink model introduced in [109] is used. It can be seen that the emulated curves are separated by $V_{oc}, 2V_{oc}$ and $3V_{oc}$ respectively, forming 3 regions, which aligns with the previous two statements. Compared with Matlab simulations, the results obtained from the emulator match well both in current and power curves. The most notable difference can be observed at the voltage at about 90 V in region 2. The main reason for this variation is that the emulator simplifies the effects of bypass diodes without fully considering its series and parallel resistance, and a constant forward voltage drop of $V_{\text{bypass}} = 2.1$ V is assumed for each module (3 bypass diodes with 0.7 V for each). In this case, the maximum available string power is found as 663.2 W and 659.5 W respectively using the developed emulator and Matlab/Simulink, leading to 0.56% MPP error.

Table 5.2 further compares the two methods with different array sizes ($n_m \times n_s$) and number of shadings in terms of computational time and MPP error (benchmarked by Simulink results). It can be seen from Table 5.2 that only small MPP error exists between the developed emulator and Simulink model. Generally, the error increases as

Table 5.2: Performance comparison between the developed emulator and Matlab/Simulink with different PV array sizes ($n_m \times n_s$) under partial shading. The simulations are carried on PV module level.

Array size	Number of shadings	Computational time [s]		MPP error [%]
		Emulator	Simulink	
5×1	1	0.008	1.07	0.08
	2	0.008	1.26	0.52
	5	0.019	1.31	0.58
10×1	1	0.008	2.68	0.05
	5	0.022	2.86	0.45
	10	0.035	3.10	0.62
10×2	1	0.008	6.57	0.04
	10	0.034	6.84	0.37
	20	0.065	6.95	0.44
10×6	1	0.008	10.39	0.02
	30	0.093	26.26	0.05
	60	0.163	34.11	0.09
10×10	1	0.008	162.17	0.01
	50	0.127	271.59	0.02
	100	0.241	443.25	0.04

more shaded modules are involved. However, it decreases rapidly as array size expands. In the case of modeling a 10 × 10 PV array (∼ 40 kW), there are only 0.03% MPP differences between emulating a single and full modules of shading, with the latter producing 0.04% MPP error. Thus we can conclude that for modeling a utility-scale PV system that contains multiple arrays, the error can further decrease. With regard to the computational time, it only takes 0.0019 s and 0.241 s respectively for the emulator to emulate the fully shaded 5×1 and 10×10 arrays, and the time barely changes when the number of shadings remains the same. On the contrary, the Simulink model is found to be more sensitive to both array size and number of shadings. For simulating a fully shaded 10×10 array, it takes 443.25 s for one simulation. Although one can perform advanced intelligent algorithms to further optimize the Simulink model, e.g., use curve fitting to derive the effects of bypass diodes, the simulation of one single

shadow transition over a PV array with 168 modules can still reach up to 1038 s [105]. The situation becomes even worse when it comes to a utility-scale system. The above results reveal that the developed emulator can achieve a good compromise between accuracy and efficiency for the module-level emulations of large-scale PV systems.

It should be noticed that the developed emulator is not only applicable for PV systems with multi-string topology. Once PV strings are emulated by the emulator, the I-V characteristic of paralleled strings can be easily obtained by cumulating their $\widetilde{I}$, but selecting the minimum $\widetilde{V}$ since the voltage across the paralleled strings should be equal and limited by the smallest one. In this way, a PV array with arbitrary module and string connections can be emulated.

5.2.2 Cloud shadow modeling

Shadow pattern generation using fractals

In this section, we deal with the modeling of cloud shadow patterns. To reproduce the irregular shapes of cloud shadows, a fractal approach introduced in [56] is adopted, where fractals are created using the diamond-square algorithm. Given the fractal pixels $N \times N$, fractal dimension D, and a constant σ, the algorithm is recursive and takes $\log_2(N-1)$ stages to complete. At each stage, the center points are calculated as the mean of four corner points in the same square plus a random variable as a function of D and σ. The edge points are obtained in a similar way but based on their three neighbors. The output of the algorithm is a three-dimensional fractal surface. We refer the readers to the original publication for more details.

In this work, the algorithm is further modified, in particular by initializing all pixels along four fractal edges to be zero. In addition, a scaling factor $\lambda = D/2$ is introduced to penalize σ to σ/λ at each stage. In this way, discontinuities at fractal boundaries are resolved, and peaks can only appear near the fractal center, which more closely resembles an actual cloud. The parameters D and σ herein used are adapted to generate low-altitude cumulus clouds as suggested in [56], which gives two fractal dimension values $D_1 = 1.9$ and $D_2 = 1.35$, and $\sigma = 20$. In this case, λ is calculated as $(D_1 + D_2)/2$.

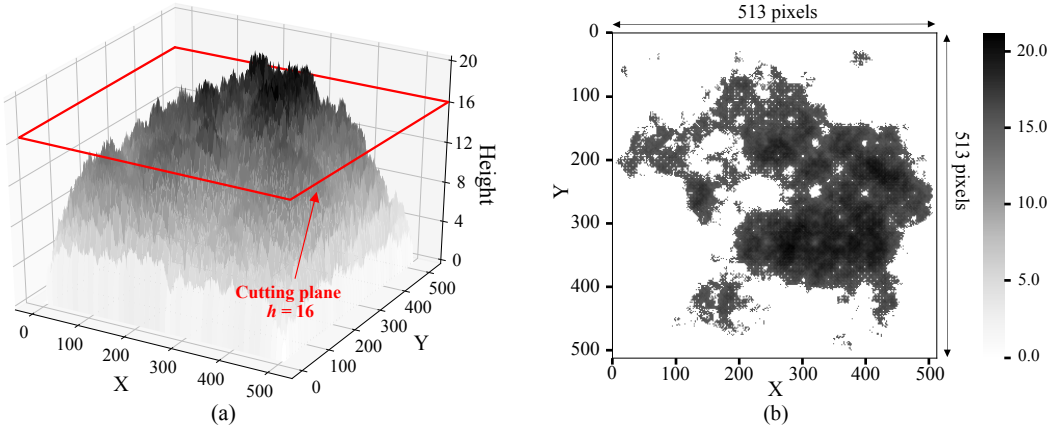


Figure 5.3: An example of generating fractals using the modified diamond-square algorithm. (a) Generated 513×513 fractal surface. By introducing the scaling factor $\lambda = 1.6$, the peaks tend to appear near the fractal center. (b) Cloud shadow pattern obtained by inserting a cutting plane to the fractal surface at $h = 16$, leading to a relative pattern size S equal to 50% of the total pixels. By initializing the fractal edge points to be zero, the generated shadow pattern becomes marginal continuous without fragments on the edges.

The cumulus clouds are considered since these clouds are the primary concern for PV system operations [42].

An example of creating a fractal surface with $N = 513$ is shown in Figure 5.3(a). It can be seen that all edge points are set to zero and the peaks of the fractal tend to appear near the center as expected. The fractal surface is an intermediate that allows a cloud shadow pattern to be generated by inserting a cutting plane of height h. The points with heights greater than h are regarded as the shadow pattern while the other points below h are discarded. The relative pattern size S is defined as the number of non-zero pixels in the shadow pattern. Figure 5.3(b) shows a shadow pattern instance generated with $h = 16$. We can see that the generated shadow pattern is marginally continuous without fragments on the four edges, which is otherwise unachievable using conventional fractal approaches. In this case, the relative shadow size S equals to 131,585 pixels, accounting for 50% of the total fractal pixels.

For practical applications, one can customize an arbitrary shadow size by assigning

a pixel dimension to the relative pattern size. A pixel dimension of d means that each pixel accounts for a geographic dimension of d^2 m^2. Since the emulation resolution herein considered is a PV module, the pixel dimension is set equal to the shorter dimension of the module. Given a user-defined shadow size of s m^2, and the module geographic dimension of $d_1 \times d_2$, where d_1 and d_2 represent the module length and width respectively, a proper h then can be found by:

$$h = \underset{\mu}{\texttt{argmin}}(|S_\mu \cdot \texttt{min}(d_1^2, d_2^2) - s| < \epsilon), \quad h_{min} \leq \mu \leq h_{max}, \tag{5.5}$$

where S_μ denotes the relative pattern size obtained at height μ, $\texttt{min}(a, b)$ is a selection function that returns the smaller value between a and b, h_{min} and h_{max} are the minimum and maximum heights of the fractal surface respectively, and ϵ is the error tolerance. After several calibration tests, the value of ϵ is set as 2.

Synthesis of shadow thickness

The pattern shown in Figure 5.3(b) is still far from mimicking a real shadow since the shadow thickness is not well-interpolated. To this end, the shadow intensity α is defined, given by:

$$\alpha = 1 - \frac{G_t}{G_{\text{clear}}}, \tag{5.6}$$

where G_t is the irradiance under cloud coverage, and G_{clear} is the clear-sky irradiance. The shadow intensity describes the attenuation of irradiance during cloud shadow transitions, and a larger shadow intensity indicates a lower irradiance value.

To synthesize the shadow thickness with shadow intensity of α, the shadow pattern is decomposed into K layers by linearly interpolating $K - 2$ new cutting surfaces between h_{max} and h_s, where h_s is the h value determined by Equation (5.5), and

$K = h_{max} - h_s + 1$. The pixel values of the kth layer, that is ω_k can be found by:

$$\begin{cases} \omega_k = 1, & k = 1 \\ \omega_k = \dfrac{1-\alpha}{K} \cdot (K-k), & 1 < k < K \\ \omega_k = 1 - \alpha, & k = K \end{cases} . \tag{5.7}$$

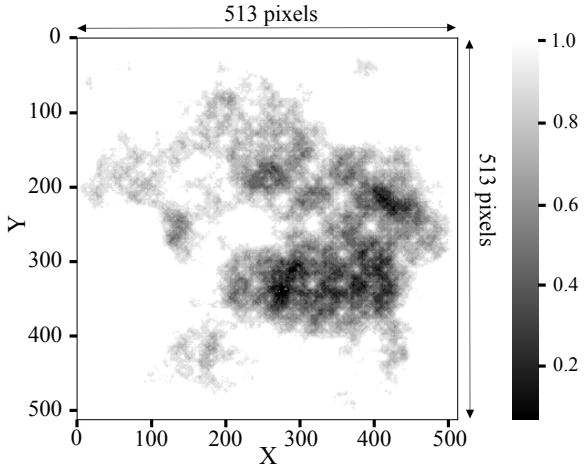


Figure 5.4: Synthesized shadow pattern with shadow intensity of 0.9. The darkest pixel corresponds to an irradiance attenuation to 0.1 (10% of G_{clear}).

Figure 5.4 shows an example of synthesizing shadow thickness for the pattern in Figure 5.3(b), with $\alpha = 0.9$. It can be seen that the thickness of the synthesized shadow pattern becomes clearer, which makes it resemble an actual shadow more. For this pattern, the darkest pixel value is 0.1, corresponding to a full irradiance attenuation of 90%. Since the shadow pattern is an image with $N \times N$ pixels, it can also be represented in a matrix form. The shadow matrix $\mathbf{\Omega}$ is given by:

$$\mathbf{\Omega} = \begin{pmatrix} \omega_{11} & \omega_{12} & \dots & \omega_{1N} \\ \omega_{21} & \omega_{22} & \dots & \omega_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \omega_{N1} & \omega_{N2} & \dots & \omega_{NN} \end{pmatrix}, \quad \in \mathbb{R}^{N \times N}, \tag{5.8}$$

where the element ω_{ij}, $i = 1, \cdots, N$, $j = 1, \cdots, N$ describes the shadow intensity of one pixel, covering a geographic dimension of d m $\times$ d m.

5.3 Case Study: A Complete Emulation Process

In this section, we demonstrate a case study of applying the proposed modeling methods to emulate PV generation during shadow transitions. The studied PV system is assumed to consist of 7 $\times$ 20 arrays, and each array is composed of 5 strings, series-connected by 40 modules. It has also been assumed that both adjacent PV strings and arrays are separated by 2 m, and there are no gaps between series-connected modules. According to the data sheet of JAM72-S09 395/PR PV module, the module dimension is 1.98 m in length and 1 m in width. With this configuration, the geographic dimension of an individual PV array is 79.2 m $\times$ 13 m, and rated at 79.016 kW, which is close to a typical PV array of utility-scale PV systems [103]. The entire system covers an area of 566.4 m $\times$ 298 m, and rated at 11.06 MW. The system is initialized to be operated under uniform irradiance of 1000 W/m^2, and a constant module temperature of 25 °C is considered.

Figure 5.5 illustrates the flowchart of the complete emulation process. Following the numbers between brackets in blocks, the emulation is conducted as follows:

Step 1: Determine the parameters of the PV system configuration, and construct a PV matrix $\mathbf{\Pi}$. As aforementioned, we can specify the module geographic dimension d_1 = 1.98 m and d_2 = 1 m, array configuration $n_m = 40$, $n_s = 5$, and array arrangement a = 7, b = 20. Then the PV matrix is obtained in the form of matrices (5.2) and (5.3), where the elements are initialized as (1000, 25). Notice that to account for the gaps, additional rows and columns of zeros have been interpolated into the PV matrix (See Appendix C).

Step 2: Generate a fractal surface by specifying the parameters of the fractal model. The suggested parameter values are $D_1 = 1.9$, $D_2 = 1.3$, $\sigma = 20$, and $\lambda = 1.6$.

Step 3: Specify the shadow size and pixel dimension d ($d = 1$ m in this case), then

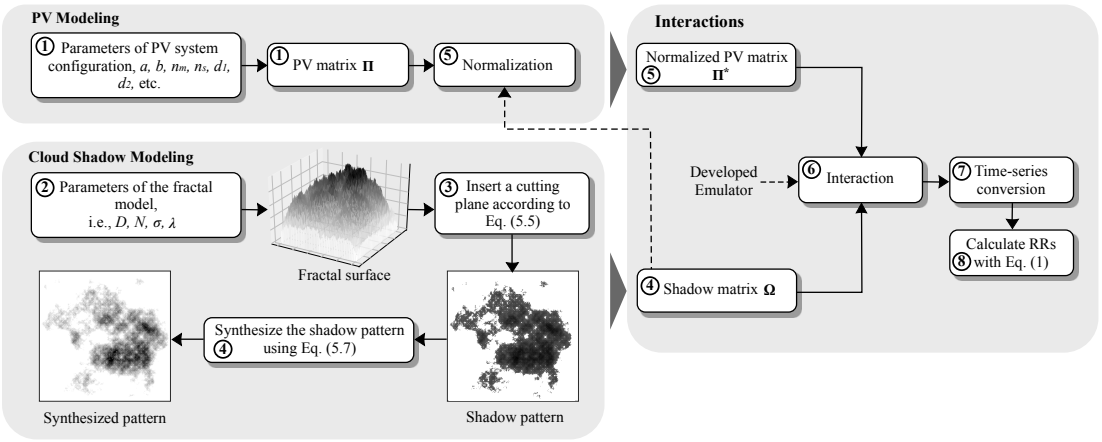


Figure 5.5: Flowchart of the proposed emulation methods. The numbers between brackets in blocks indicate the emulation sequence.

insert a cutting plane to the fractal surface according to Equation (5.5).

Step 4: Specify the shadow intensity, and synthesize shadow thickness using Equation (5.7). For demonstrative purpose, we herein adopt a similar shadow pattern shown in Figure 5.4, leading to a shadow intensity of 0.9 and shadow size of 513 m $\times$ 513 m. The shadow matrix $\mathbf{\Omega}$ is then obtained.

Step 5: Normalization. So far we have developed the PV matrix $\mathbf{\Pi}$ and shadow matrix $\mathbf{\Omega}$. However, the elements of the two matrices show different geographic dimensions as the element in PV matrix denotes a 1.98 m $\times$ 1 m area, while an element in shadow matrix indicates a geographic dimension of 1 m $\times$ 1 m. In order to normalize the two matrices, the PV matrix are split into $(d_1/d) \times (d_2/d)$ elements (in this case 1.98 $\times$ 1 elements). The normalized PV matrix $\mathbf{\Pi}^*$ shows an element geographic dimension of 1 m $\times$ 1 m. Details about the matrix normalization can be found in Appendix C.

Step 6: Interaction of the PV and shadow matrices. To emulate the dynamic shadow transitions, one can consistently move the shadow matrix over the normalized PV matrix. In this case study, we consider a horizontal cloud shadow movement, thus the shadow matrix should affect the PV matrix column by column. The emulation starts as the shadow matrix is about to enter the PV matrix, and ends as the shadow

matrix completely leaves the PV matrix. During interactions, the intersection of the two matrices indicates the shading on the PV system. Recalling Equation (5.6), the module irradiance under cloud coverage then can be calculated as $(1-\omega_{ij})\cdot 1000$, where $\omega_{ij} \in \mathbf{\Pi}^* \bigcap \mathbf{\Omega}$. At each moving step, the PV matrix is updated and fed to the developed emulator.

Step 7: Time-series conversion. The output of step 6 is a time-independent power sequence. In order to assign temporal attributes to the sequence, the shadow velocity should be specified. In this context, a constant shadow velocity of 10 m/s is considered. Since the shadow matrix travels at 1 column per step, the distance of which equals to the module width (1 m), it should take the shadow 0.1 s for a single moving step. Then the output power sequence can be converted to time-series.

Step 8: Calculate PV system RRs.

Figure 5.6(a) and Figure 5.6(b) display the emulated system output power and resultant RRs calculated with $\Delta t = 1$ s respectively. The emulation takes 811 steps to complete, indicating a shadow transition of 81.1 s. To make the emulation more intuitive, Figure 5.6(c) further visualizes the model interactions, where the PV field is highlighted by dashed lines, and the shadow moves horizontally from right to left. It can be observed from Figure 5.6(a) and Figure 5.6(b) that the PV output power and resultant RRs are not exactly symmetrical as the shadow is entering and leaving the system. Referring to Figure 5.4, the right part of this particular shadow pattern presents to be thicker and sharper than the left part. As a consequence, the RR changes faster as the right part of the shadow is affecting the PV field. This asymmetry is similar to the variabilities observed at realistic utility-scale PV systems [110], thus yielding the necessity of reproducing the shadow natures. During the shadow transition, the system mean RR is found to be 0.60%/s, and the maximum instantaneous RR reaches 2.23%/s at 67 s when the shadow center is leaving the PV field (see the middle plot of Figure 5.6(c)).

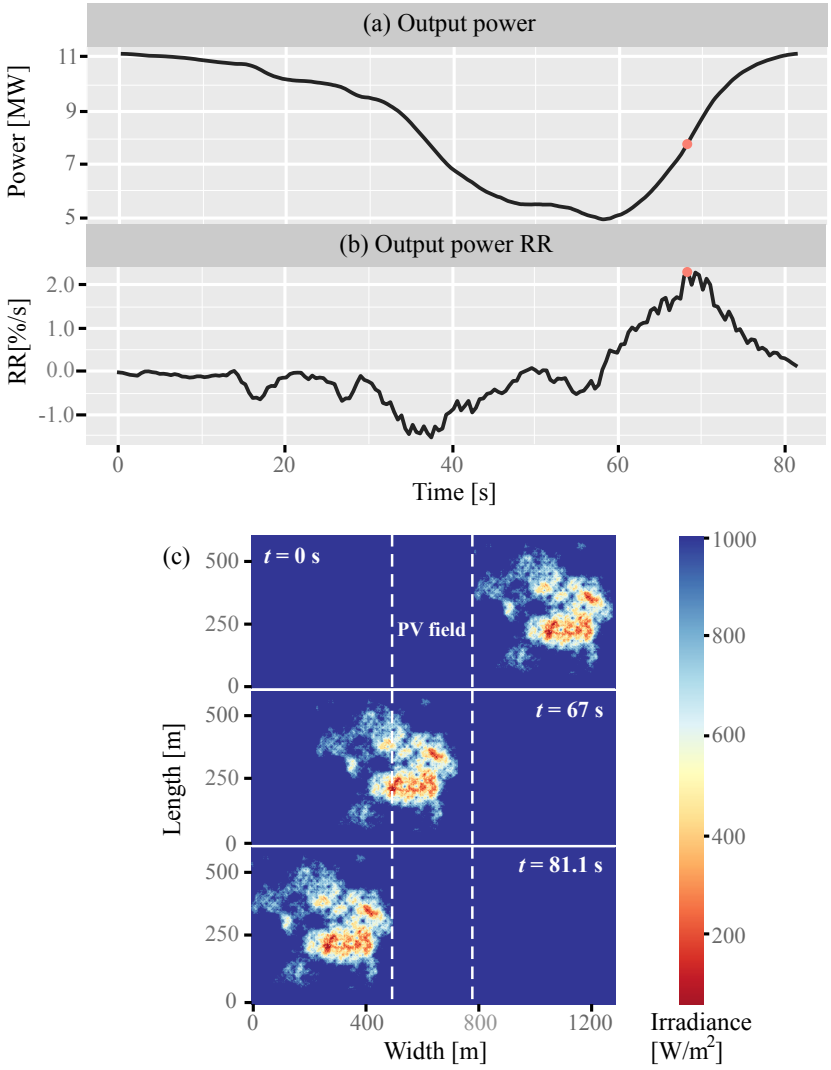


Figure 5.6: Emulated (a) output power time-series and (b) resultant RRs calculated with $\Delta t = 1$ s for a 11.06 MW PV system under a horizontal shadow movement at 10 m/s. The emulation takes 811 steps to complete, indicating a shadow transition of 81.1 s. The maximum instantaneous RR reaches 2.23% at the instant 67 s, which is labeled in red. (c) Visualization of model interactions at emulation step $t = 0$ s (top plot), 67 s (middle plot), and 81.1 s (bottom plot). The PV field covers an area of 566.4 m $\times$ 298 m, highlighted by dashed lines.

5.4 Results

In this section, PV system RRs due to both endogenous and exogenous factors during cloud shadow transitions are investigated. In addition, a total of 3747 cloud shadow transitions identified from real measurements are utilized for assessing the RR characteristics in reality. Six PV systems ranging from 1 MW to 60 MW are emualted. The modeled array configuration is identical to the previous case study, that is, consisting of 40×5 modules and covering an area of 79.2 m $\times$ 13 m. Throughout the emulations, the system is initialized under uniform irradiance of 1000 W/m^2. Moreover, a constant module temperature of 25 °C is considered for simplicity. During fast shadow transitions, the module temperature differences within a large-scale PV system are typically small, which can only lead to slight impacts on system operations [111].

5.4.1 Effects of Endogenous Factors: Array Arrangement and System Orientation

Two endogenous factors are investigated herein, namely array arrangement and system orientation. For each studied PV system, four different array arrangements are evaluated, with increased number of arrays in row while decreased in column. To account for system orientations, the PV matrix is rotated accordingly. A system orientation of 90° denotes an array configuration same as Figure 5.1, where the PV strings are aligned in columns, and the orientation angle decreases as the system rotates in clockwise. Nine different cloud shadow patterns have been produced for generalization, with $N = 513$, $S = 50\%$ total pixels, and $\alpha = 0.9$ (see Appendix D). The shadows are assumed to move horizontally at a fixed velocity of 10 m/s. Notice that the parameter values selected here, i.e. shadow intensity, shadow velocity, and shadow size, are only for demonstrative purpose, a detailed evaluation on the effects of these so-called exogenous factors will be further addressed in Section 5.4.2. The emulation results are shown in Figure 5.7 and Table 5.3.

It can be seen from Figure 5.7 and Table 5.3 that the systems with longer dimension

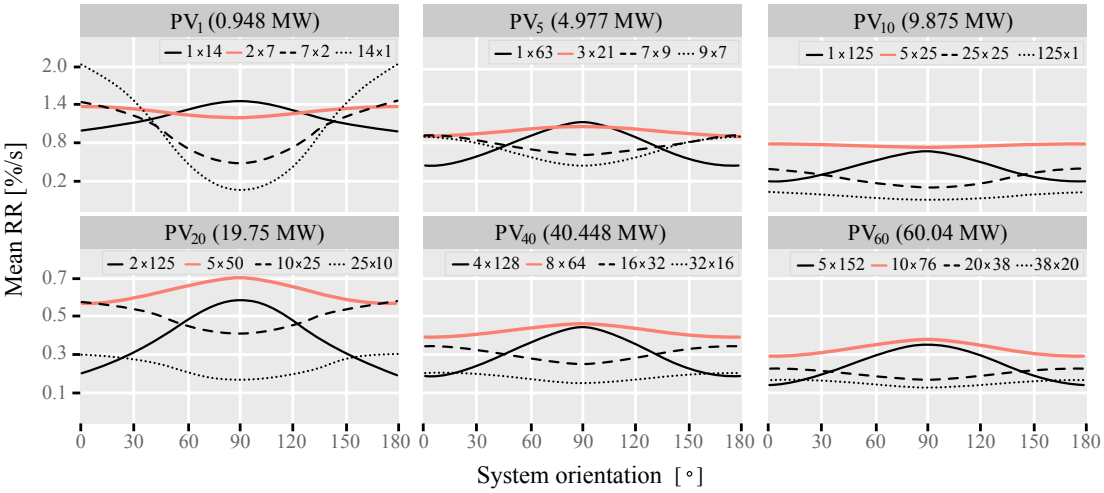


Figure 5.7: Mean RR for the six studied PV systems with respect to different array arrangements and system orientations. The performance of the most-squared systems are plotted in red. The results are averaged onver 9 different cloud shadow patterns.

in width typically experience the largest and smallest RR at orientation around 90° and 0° respectively, while systems with longer dimension in length show an opposite result. This indicates that the system RR becomes the largest as the shadow moves along the system longer dimension, and becomes the smallest as the shadow moves along the shorter dimension.

It is also observed that for systems whose dimensions are smaller or comparable to shadow diameters (513 m × 513 m), such as PV_1 and PV_5 systems, the arrangements that are closer to a shape of square reveal less sensitivity to system orientations (see the cases of 2 by 7 and 3 by 21 arrangements). As the system becomes more elongated, the effect of system orientation presents to be stronger. Regarding systems larger than PV_5, though, the system dimensions become much larger than shadow diameters, thus the spatial dispersion within the PV plant can greatly reduce the ramp magnitudes. In these cases, variations of RR due to system orientations become the slightest for the longest arrangement (more arrays aligned in rows), and largest for the widest arrangement (more arrays aligned in columns). Referring to the array configuration in Figure 5.1, it implies that the in-plant spatial smoothing increases as the system elongates in the

Table 5.3: Emulation results for the six studied PV systems in terms of different array arrangements and system orientations. The performance of the most-squared systems are marked in bold. The results are averaged over 9 different cloud shadow patterns.

System	Capacity [MW]	Array arrangement [row×column]	Dimension (length × width) [m × m]	Maximum (Maximum variation) of mean RR for different system orientations [%/s]	Maximum (Maximum variation) of instantaneous RR for different system orientations [%/s]
PV_1	0.948	1×14	79.2×208	1.51 (0.34)	11.41 (8.06)
		2×7	**160.4×103**	**1.49 (0.17)**	**6.06 (1.96)**
		7×2	566.4×28	1.52 (0.78)	5.71 (3.62)
		14×1	1134.8×13	2.05 (1.65)	10.15 (8.83)
PV_5	4.977	1×63	79.2×943	1.08 (0.60)	8.73 (7.59)
		3×21	**241.6×313**	**1.03 (0.16)**	**3.11 (0.54)**
		7×9	566.4×133	0.88 (0.28)	2.80 (1.23)
		9×7	728.8×103	0.87 (0.30)	2.61 (1.47)
PV_{10}	9.875	1×125	79.2×1873	0.68 (0.48)	5.98 (5.48)
		5×25	**404×373**	**0.80 (0.06)**	**2.19 (0.23)**
		25×5	2028×73	0.38 (0.22)	1.12 (0.71)
		125×1	10148×13	0.11 (0.06)	0.65 (0.50)
PV_{20}	19.75	2×125	160.4×1873	0.58 (0.39)	2.99 (2.50)
		5×50	**404×748**	**0.71 (0.18)**	**1.95 (0.92)**
		10×25	810×373	0.56 (0.12)	1.67 (0.66)
		25×10	2028×154	0.30 (0.12)	0.93 (0.48)
PV_{40}	40.448	4×128	322.6×1918	0.45 (0.27)	1.46 (1.02)
		8×64	**647.6×958**	**0.47 (0.09)**	**1.07 (0.27)**
		16×32	1287.2×478	0.33 (0.08)	0.92 (0.31)
		32×16	2596.4×238	0.20 (0.07)	0.65 (0.24)
PV_{60}	60.04	5×152	404×2278	0.34 (0.21)	0.98 (0.64)
		10×76	**810×1138**	**0.37 (0.08)**	**0.86 (0.33)**
		20×38	1622×518	0.24 (0.06)	0.63 (0.15)
		38×20	3083.6×284	0.16 (0.04)	0.55 (0.13)

same direction of string alignment. This is foreseeable since a system with more strings aligned in columns would suffer more from partial shading effects. However, it should be noticed that although the most square-like alignments produce larger RRs than the elongated alignments in these cases, there are only slight differences in terms of RR variations, which means the square-like alignments are still advantageous to counteract the orientation effects.

These findings suggest that for the areas with a regular cloud shadow moving direction, one should place PV arrays in an elongated alignment, with system longer dimension in the same direction of string alignment, and in perpendicular to the shadow movements. While for the areas with unknown or frequently changing shadow movements, a squared system alignment is preferred to minimize the effects of system orientations.

5.4.2 Effects of Exogenous Factors: Shadow Intensity, Shadow Velocity, and Shadow Size

In this subsection, we evaluate the effects of exogenous factors that come from three shadow properties, i.e. shadow intensity, shadow velocity, and shadow size on PV system RRs. Each factor is studied by means of control variates, where the other two factors are fixed at their featured values. The featured values of shadow intensity, velocity, and size are selected as 0.9, 30 m/s, and 100% system size, which approximate the maximum observations from real measurements (see Figure 5.10 in Section 5.4.3). The maximum observations are considered since these worst RR scenarios are the main concerns for system operations e.g., storage sizing [27], PV plant planning and control [112]. Furthermore, based on the results from previous subsection, we herein only consider the most square-like arrangements in Table 5.3, so that the system orientation effects are minimized. Similarly, the results are generalized with the nine cloud shadow patterns. It should be stipulated that although shadow moving direction is a notable factor that affects system performance as well, the effects are identical to changing system orientations. In this regard, the effects of shadow moving directions are not

reiterated here, and a constant horizontal shadow movement will be considered in the remainder of this paper.

The mean and maximum instantaneous RRs of the six studied PV systems due to different shadow properties are shown in Figure 5.8. To visualize the effects more intuitively, the values of shadow intensity, shadow size, and shadow velocity are normalized to their respective featured values. It can be seen that both mean and maximum RRs increase linearly with the increasing shadow velocity. This is reasonable since the shadow velocity presents to be linear to the cloud transition interval. However, this linearity disappears for the case of shadow intensity. Instead, a logarithmical growth is observed. This nonlinearity between the shadow intensity and RRs should be largely caused by the geographic dispersion of a PV system, as a result of which the system power fluctuations can be significantly smoother than the irradiance fluctuations. Similar observations have also been reported in [69] for practical large-scale PV systems, where the system geographic dispersion is found as a low-pass filter to irradiance fluctuations. Notice that when a smaller system, e.g., a single PV array is considered, the effects of geographic dispersion are greatly reduced. In this case, the shadow intensity can become linear to both mean and maximum instantaneous RRs [113]. In terms of shadow size, it shows the strongest effects on system maximum instantaneous RR. When the shadow size increases to 20% of system size, the maximum instantaneous RR increases to 3.88%/s, 3.39%/s, 3.23%/s, 1.86%/s, 1.48%/s, and 1.08%/s respectively for the PV_1, PV_5, PV_{10}, PV_{20}, PV_{40}, and PV_{60} systems, corresponding to 36.78 kW/s, 168.72 kW/s, 318.98 kW/s, 367.35 kW/s, 598.63 kW/s, and 648.43 kW/s. However, it is beaten by shadow intensity when comes to mean RR. This is mainly due to a longer cloud transition interval required in case of an expanded shadow size, where the mean ramp-rate, to some extent, are alleviated in temporal scale. As the shadow intensity increases to 0.2, it leads to mean RR of 2.16%/s, 1.26%.s, 1.05%/s, 0.82%/s, 0.42%/s, and 0.29%/s for the systems respectively, corresponding to 20.48 kW/s, 62.71 kW/s, 103.69 kW/s, 161.95 kW/s, 169.88 kW/s, and 174.11 kW/s.

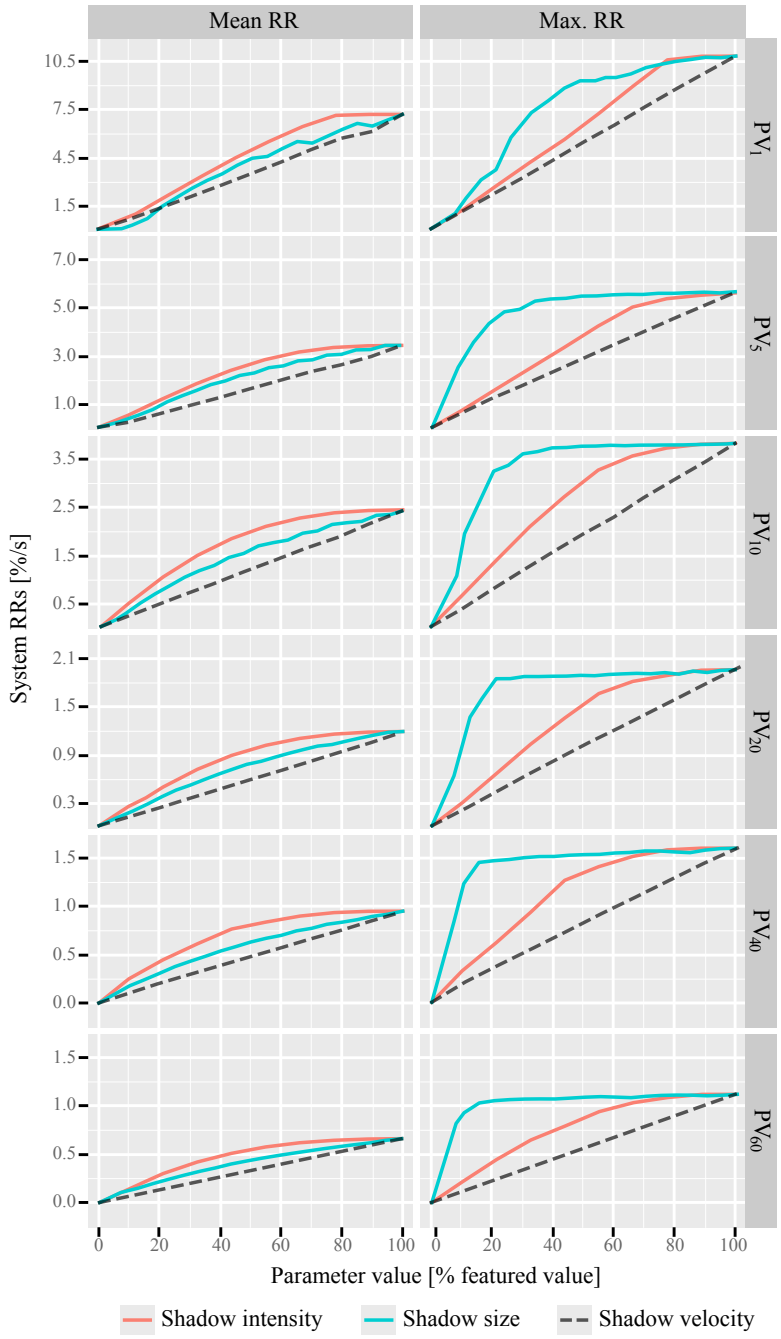


Figure 5.8: Mean and maximum instantaneous RRs of the 6 studied PV systems under different shadow characteristics. The values of shadow intensity, shadow size, and shadow velocity are normalized to their respective featured values: 0.9, 100% system size, and 30 m/s.

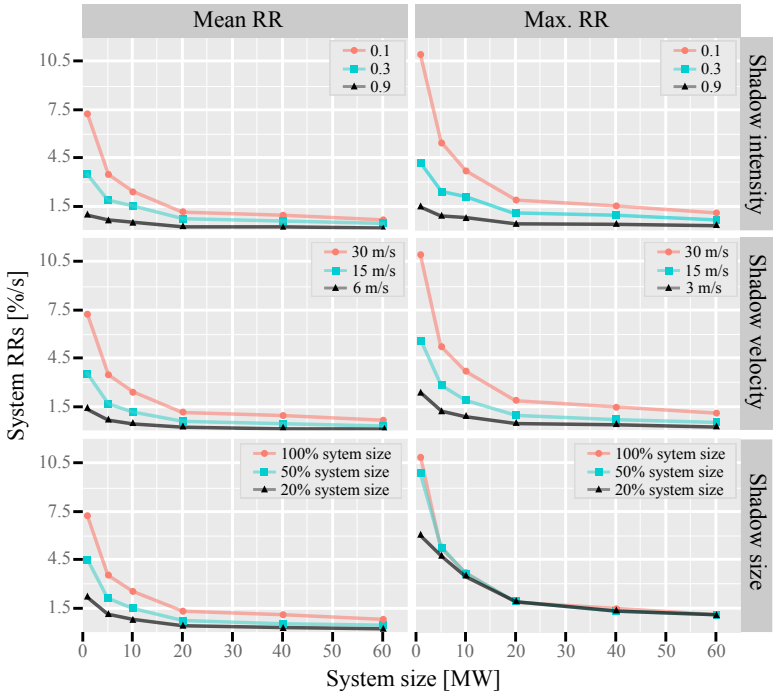


Figure 5.9: Mean and maximum instantaneous RRs under different shadow properties in terms of different system sizes. Exponential decay is observed for all the shadow properties as the size of system expands.

Figure 5.9 further illustrates the effects of exogenous factors on different system sizes. It can be seen that for all the studied shadow properties, both mean and maximum instantaneous RRs decrease exponentially as system size increases. This exponential decay is also in line with the observations of practical large-scale PV systems presented in [99].

5.4.3 PV System RRs under Identified Shadow Transitions

In the previous subsection, PV system RRs along with several exogenous factors are evaluated with empirically selected values. In this subsection, a total of 3747 shadow transitions identified from real measurements are applied to assess and quantify the PV system RRs in reality. Similarly, only the most square-like systems in Table 5.3 are considered in order to exclude the system orientation effects. The emulation period ranges from 6 s to 396 s according to different PV system capacities and shadow

transitions.

The cloud shadow transitions are exploited using solar irradiance data collected over 6 months from XJTLU sensor network (see Chapter 3). For each sensor, GHI data with a sampling time of 1 s is available. In order to exclude potential measurement errors, the 10% shadow intensity limit is applied. The shadow velocity is calculated using the most correlated pair algorithm previously introduced in Section 2.2.1, which is based on the time lags between a central sensor and its neighbors. Notice that to precisely measure the shadow size, a dense sensor network dispersed over a wide area is required. Unfortunately, such measurements are currently unavailable. In this regard, we measure the shadow length instead, which is a production of shadow velocity and sensor shading period. To obtain the shadow pattern, Equation (5.5) is modified so as to determine a proper h with a specified shadow length L, and the fractal surface are generated with $N = L$. Other parameters of the cloud shadow model are set equal to the previous example, where $D_1 = 1.9$, $D_2 = 1.35$, $\sigma = 20$, and $\lambda = 1.6$. (More information about the sensor network and identification of shadow transitions can be found in Appendix E).

The scatter plots of the three measured shadow properties are displayed in Figure 5.10. It can be seen that both medians and 90th percentiles remain approximately constant, indicating that there are no clear correlations between each two of the properties. In the identified shadow transitions, the medians of the shadow intensity, shadow velocity, and shadow length are 0.40, 9.09 m/s, and 161.32 m. The 90th percentiles are 0.65, 13.34 m/s, and 418.80 m respectively.

As previously mentioned, various RR regulations have been imposed by different utilities, where the RR limit differs in terms of time step Δt and measuring unit (unit of power or unit of percentage). In this work, we follow the state-of-art RR regulation stipulated in Danish grid code, where the maximum RR is suggested to be 100 kW/s for systems above 11 kW [15]. The Danish RR limit is studied since it is enacted specially for large-scale PV systems, and (probably) is the most rigorous regulation among existing standards (RR limit at second level). Notice that for other choices of

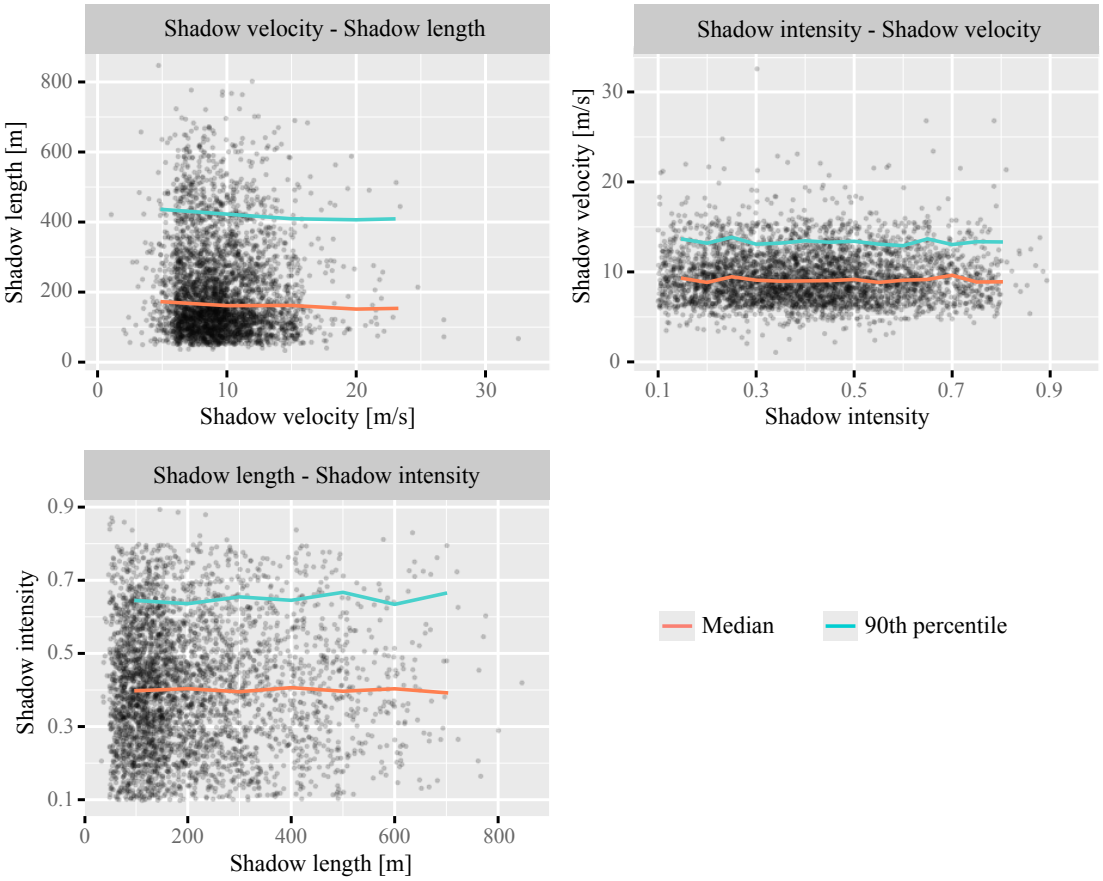


Figure 5.10: Scatter plots between the measured shadow intensity, shadow velocity and shadow length. The medians and 90th percentiles are plotted in red and blue lines respectively, showing that there are no clear correlations between each two of the properties.

RR limit (e.g., 10%/min required as Germany [14] and Puerto Rico [16]), one can follow the same process described in Figure 5.5, but calculate RRs with different time steps. The comparison of different RR limits can become another huge topic, which may be possible in a future study.

Figure 5.11 shows the cumulative distributions of RRs during the identified shadow transitions for the six studied systems, where the portions that lie on the right side of RR limit (red solid line) indicate the percentages of RR violations. The system performance in terms of mean and maximum instantaneous RRs, and RR violations are

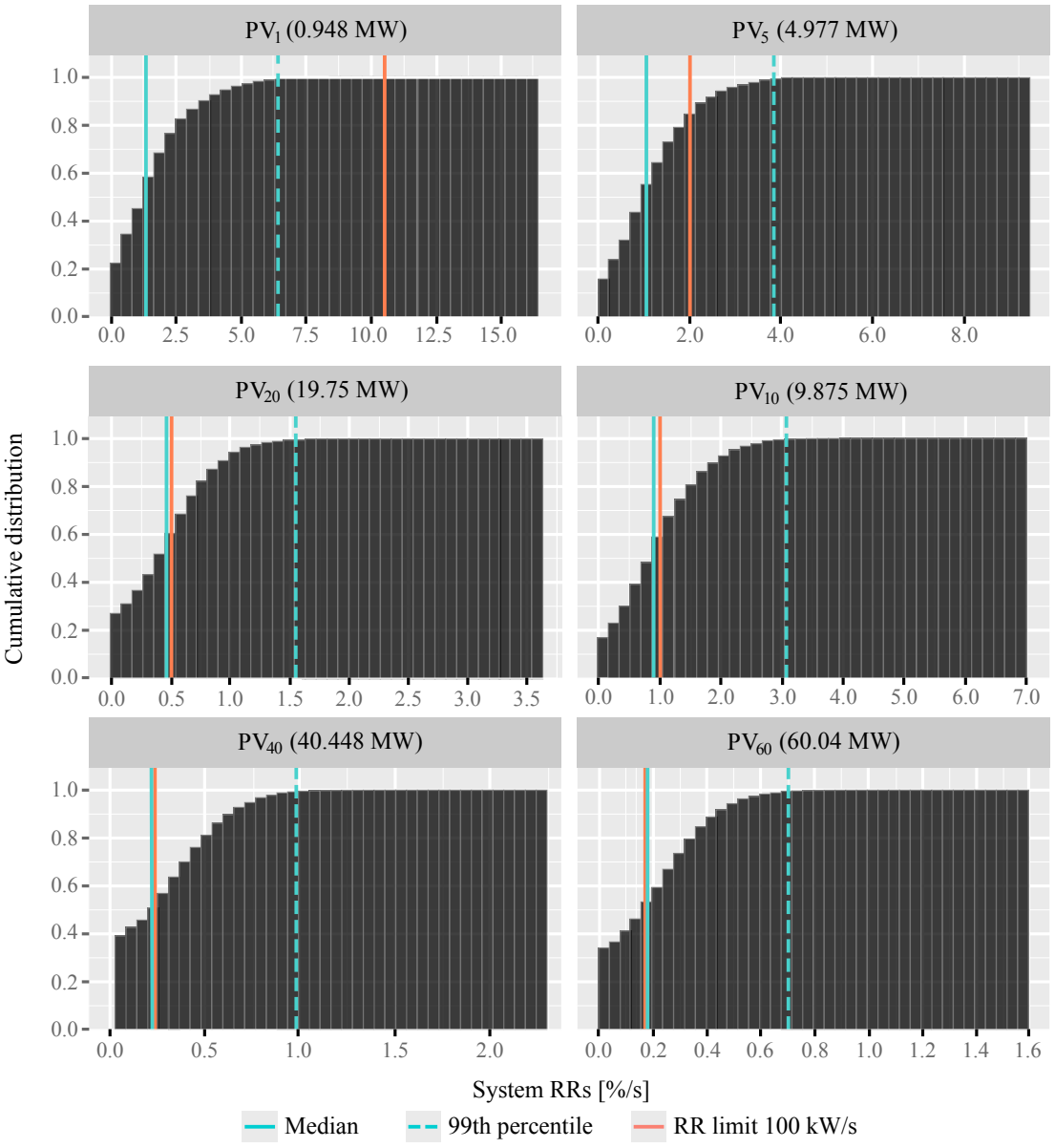


Figure 5.11: Cumulative distributions of RRs for the six studied systems. The medians and 99th percentiles are plotted in blue solid and dashed lines respectively. The portions that lie on the right side of RR limit (red solid line) indicate the RR violations.

Table 5.4: System mean and maximum instantaneous RRs during all the identified shadow transitions. The RR violations are calculated as the percentage of RRs greater than 100 kW/s to the total RRs.

System	Mean RR [kW/s]	Max. RR [kW/s]	RR violations [%]
PV_1	15.93	156.58	0.02
PV_5	59.15	468.61	17.11
PV_{10}	92.95	695.33	44.03
PV_{20}	95.35	714.82	44.19
PV_{40}	97.10	916.56	48.18
PV_{60}	109.47	960.49	51.57

further summarized in Table 5.4. It can be observed that the RR limit is seldom violated for the PV_1 system, with over 99% of the shadow transitions complying with the regulation. Referring to Figure 5.8, it almost requires a shadow with all featured property values to cause a RR of 100 kW/s (10.55%/s) on PV_1 system, which is the rare cases among the measurements (48 of 3747 shadows). As the system capacity increases, the RR limit approaches closer to the medians (blue solid line), implicating that the system has experienced more RR violations. During the identified shadow transitions, the maximum RRs of the 6 systems are found as 156.58 kW/s, 468.61 kW/s, 695.33 kW/s , 714.82 kW/s, 916.56 kW/s, and 960.49 kW/s respectively, which have all exceeded the RR limit. For the PV_{60} system, RR violations should account for 51.57% of total RRs, that is, over half of time the system would be of concerns for grid operators.

To validate the emulations, Table 5.5 further compares the characteristics of the emulated PV_1 and PV_{10} systems with two practical PV plants of similar capacities presented in [50], namely Sesma (0.99 MW) and Milagro (9.5 MW) plants. The characteristics of Sesma and Milagro plants are obtained based on one year practical power measurements with 1 second resolution. As displayed in Table 5.5, the 90th percentiles of RR distributions for the Sesma and Milagro plants are reported as 6%/s and 3%/s respectively, indicating that 90% of the observed RRs should be less than

Table 5.5: Validation of the emulated RR characteristics of PV_1 and PV_{10} systems using one year observations of two practical PV plants, Sesma and Milagro plants. The maximum RRs of the two practical plants are estimated from the distribution plots.

System	Capacity [MW]	90th percentile [%/s]	Max. RR [%/s]
PV_1	0.948	5.18	16.52
Sesma	0.99	6.00	$\approx$ 13.00
PV_{10}	9.875	2.12	7.04
Milagro	9.5	3.00	$\approx$ 6.00

6%/s and 3%/s respectively for the 0.99 MW and 9.5 MW systems. For the emulated RR distributions, the 90th percentiles of the PV_1 (0.948 MW) and PV_{10} (9.875 MW) systems are found as 5.18%/s and 2.12%/s respectively, which generally match the Sesma and Milagro plants. To bring more insights on the RR distribution, e.g., the outlier, we also compare the maximum RR between the emulations and measurements (notice that the value of maximum RR is not directly provided in [50], thus we estimate it from the distribution plots). It can be seen from Table 5 that the differences in maximum RRs between the emulated PV_1, PV_{10} systems and Sesma, Milagro plants are practically small. Considering the errors brought by different system configurations (approximate 1 % according to Figure 5.7), the emulated systems can be promising to describe the practical PV systems.

5.5 Chapter Conclusion

This chapter presents a comprehensive study on the RRs of utility-scale PV systems under passing clouds. Compared with previous studies, this work is carried out based on the emulations of utility-scale PV systems at a higher resolution i.e., module-level. The irradiance characteristics and spatial diversities of real cloud coverages are also considered in more detail. To emulate the module-level behaviors of a partially shaded PV system, a partial shading emulator is proposed. Compared with conventional computer-aided simulators, the proposed emulator significantly improves the computational efficiency

while retaining promising accuracy. In order to reproduce the cloud shadow natures, a modified fractal approach is developed, which allows user-defined shadow properties to be synthesized. Based on the developed emulation tools, the PV system RRs due to both endogenous factors (i.e., PV array arrangement and system orientation) and exogenous factors (i.e., shadow intensity, shadow velocity) have been investigated.

The characteristics of the RRs are evaluated for a series of utility-scale PV systems ranging from 1 MW to 60 MW. The main findings can be summarized as:

- The most square-like array arrangement reveals the least sensitivity to system orientations. As the system expands, however, the in-plant spatial smoothing can also significantly dampen the orientation effects, and the smoothing becomes stronger as the system elongates in the direction of string alignment.

- Regarding the exogenous factors, the mean and maximum instantaneous RRs appear to be linear to the shadow velocity, while a logarithmical growth is observed for both shadow intensity and shadow size. Moreover, the shadow intensity and shadow size are found to dominate system performance in terms of mean and maximum instantaneous RRs respectively. As system capacity increases, the effects of all the studied exogenous factors decrease in an exponential decay manner.

In order to characterize the RRs in reality, a total of 3747 cloud shadow transitions identified from real measurements are also explored. For the Danish 100 kW/s RR limit, it is found that a utility-scale PV system can frequently violate the regulation even if the multi-string array configuration is considered, which inherently counteracts the partial shading effects. On the whole, a system with larger capacity presents to suffer more from ramp violations. Specially in the case of a 60 MW system, more than 50% of the identified cloud shadow transitions have triggered ramp violations.

Chapter 6

Conclusions and Outlook

6.1 Summary of Contents

The short-term PV power intermittency caused by passing clouds is becoming a major concern for grid operators. As the penetration of PV systems boosts, the rapid power fluctuations pose more challenges to maintain grid transient stability. In this context, the main focus of this thesis is on facilitating the mitigation of solar PV intermittency. Specifically, this thesis revolves around *control*, *forecasting*, and *emulation* of cloud-induced PV power RRs. A brief summary of this thesis is presented as follows.

In **Chapter 1**, the demands and challenges for mitigating solar PV intermittency is discussed. To address the adverse impacts from the PV systems, PRRC has been widely imposed by grid operators. As current PRRC strategies still show limitations to efficiently comply with RR regulations, solar forecasting, especially intra-minute forecasting or nowcasting, provides a remedy to enhance the controllability of PV systems. On the other hand, emulation of PV system RRs during cloud transitions can offer a better understanding on the RR characteristics for system operators. With the prior knowledge of system RRs, PV intermittency may be managed more effectively.

In **Chapter 2**, two novel PRRC strategies are presented, which utilize the sensor network-based solar nowcasting. The first strategy, termed as PAPC, does not require any external storage systems. During ramp-down events, the integration of a nowcasting

system allows to suppress the PV generation before the actual cloud shading occurs. The second strategy, FEC, reduces the required ESS capacity to quarter of the conventional method. Moreover, it shows an advantage by neglecting the dedicated SOC control operations. Economic analysis is also demonstrated to validate the feasibility of the proposed strategies.

As the two predictive PRRC strategies highly rely on the accurate predictions of upcoming RRs, **Chapter 3** develops an improved sensor network-based solar nowcasting method, which can optimally preselect the spatio-temporal predictors for nowcasting models. The proposed SRP preselection method enables a fast and precise predictor preselection in different scenarios, and provides consistent PV nowcasts with cloud information interpolated. The experimental results reveal that the proposed method has strong robustness in various weather weather condition, with fewer training data needed. Moreover, compared with conventional methods, the proposed method significantly improves the nowcasting accuracy, with predicted ramps approaching closer to actual ramp peaks. The effectiveness of the developed SRP-Enet model for PAPC is also evaluated. The results show that the proposed method can be promising to improve the performance of PRRC.

Chapter 4 brings forth the operability of solar nowcasting on PRRC implementations. To that end, three time parameters involved in operational solar nowcasting are identified, i.e., forecast horizon, forecast resolution, and forecast model updating rate. Based on an actual irradiance sensor network, both deterministic and probabilistic nowcasting are generated, and integrated into PAPC operations. The empirical studies reveal that the performance of both operational solar nowcasting and PAPC are highly characterized by the choice of time parameters. In addition, the PV system capacity also shows notable impacts on the results.

In **Chapter 5**, the emulation of cloud-induced system RRs is addressed. With the objective of producing high-resolution and high-accuracy emulations of utility-scale PV systems under passing clouds, a partial shading emulator is proposed, which can efficiently mimic the behaviors of a partially shaded PV system, in accuracy of PV

module level. Then a fully customizable shadow model that can reproduce the natures of a real cloud shadow is introduced. Based on the developed emulation tools, system RRs rising from two endogenous factors i.e., PV array arrangement and system orientation, and three exogenous factors i.e., shadow intensity, shadow velocity, and shadow size are studied. Furthermore, in order to assess the RRs in reality, a total of 3747 cloud shadow transitions exploited from real measurements have been applied for emulations. The results reveal that the RRs caused by passing clouds are critical problems for system operations, and a larger system can suffer more from ramp violations, indicating that PRRC strategies should be essential for contemporary utility-scale PV systems.

6.2 Outlook

At this stage, it is clear that the mitigation of solar PV intermittency relies on the joint efforts by solar forecasters and system operators. On one hand, solar forecasters quantify the solar resource dynamics, and provide decent solar forecasts for system operators. On the other hand, system operators make the best use of the forecasts to enhance system stability. Clearly, to advance the integration of PV systems further, it is necessary to unify the knowledge from the two fields. Hence, in what follows, an outlook into future studies is presented. From both a solar forecaster's and a system operator's point of view, various research topics in demand are discussed.

6.2.1 From A Solar Forecaster's Perspective

From solar irradiance to PV power forecasting

To bridge solar forecasting and power system operations, a natural requirement is to forecast the solar-generated power. However, as most modern solar forecasting works dig heavily on exploring new input features for solar irradiance forecasting, PV power forecasting is less addressed [114].

Generally, the approaches for converting irradiance to PV power can be classified into parametric and and nonparametric ones [30]. Nonparametric approaches conceive

the PV system as a black box, which aim to derive a one-stop model for direct irradiance-power mapping. The irradiance-to-power conversion model used in this thesis comes to this category. On the other hand, parametric approaches usually consist of three modeling process—separating diffuse horizontal irradiance from the GHI forecasts using a separation model [115]; converting the horizontal irradiance components to the POA (plane of array irradiance) irradiance with a transposition model [116]; and a PV system model to convert the POA irradiance to output power [117].

Since the conversion involves several procedures, it is therefore of interest to study how the errors can propagate from solar irradiance forecasts to PV power forecasts. Besides, as shown in Chapter 5, PV power generation can be highly characterized by both endogenous and exogenous factors. On this point, the effects of different system and meteorological attributions on PV power forecasting need to be further clarified.

Probabilistic solar forecasting

A weather forecast is intrinsically five-dimensional, spanning space, time, and probability [118]. In that, solar forecasts ought to be essentially *probabilistic*. According to Chapter 4, it is evident that probabilistic solar nowcasting can favor PAPC by offering more implementational flexibilities—one can apply different PIs or quantiles of a probabilistic nowcast to the control, under a trade-off between RSR and ECR. Besides the PAPC application, probabilistic solar forecasting also reveals monetary and reliability benefits over deterministic equivalents, in terms of market bidding [119], battery management [120], power system auxiliary services [121], and economic scheduling [122], etc.

Given the above evidence, clearly, it is more advantageous for system operators to integrate probabilistic solar forecasting into their decision-making process. In fact, the probabilistic forecasts can be flexibly post-processed into deterministic forecasts, if only a single best guess is of interest. In that case, one may either summarize the predictive distribution through a statistical functional, or combine the component forecasts [123]. As the value of probabilistic forecasting being exploited, it is believed that probabilistic solar forecasting will become a trending topic.

Forecast reconciliation

In case study 2 of Chapter 4, it is observed that the low-resolution nowcasts typically improve the performance of PAPC. In this regard, it reveals a necessity to produce more accurate nowcasts at a larger $\mathcal{R}$.

Given the high-resolution measurements, if the forecasts at a lower resolution are expected, one can simply aggregate the data before forecasting. However, as aggregating the dataset can significantly reduce the length of training data, the forecasting models could be underfitting. Alternatively, one can aggregate several high-resolution forecasts into a low-resolution forecast, which is known as *forecast aggregation*. However, due to the modeling errors, the forecasts are in general not aggregated consistent (e.g., the sum of three 1-s resolution forecasts is not equal to the single forecast made using 3-s resolution data). This problem is not properly addressed until a recent work in [124], where a regression-based temporal reconciliation framework is proposed to unify all forecasts produced at different resolutions.

In [125], such temporal reconciliation is firstly applied to solar forecasting. In general, it is found that reconciliation methods not only provide consistent forecasts, but also improve the forecasting accuracy at different hierarchy levels. As the forecast reconciliation is still at an early stage in solar forecasting community, this aspect ought to be studied in depth to provide system operators with new insights.

6.2.2 From A System Operator's Perspective

Coordination with PV power reserve control

In Chapter 4, it is found that PAPC performance presents a high correlation with metrics nMAE (deterministic) and PINAW (probabilistic). In general, deterministic nowcasts with a smaller negative nMAE, and probabilistic nowcasts with a larger PINAW are more likely to produce better PAPC results, especially in terms of RSR. This finding implies that operating PV power below its maximum available power could bring additional benefits to PAPC. From a system operator's point of view, such operation is referred as

power reserve control (PRC) [126].

Given the maximum available power P_{avai}, and the required amount of reserved power ΔP, PRC is described as $P_{\text{pv}} = P_{\text{avai}} - \Delta P$, where P_{pv} is the regulated PV power. While ΔP is a user-defined value, P_{avai} often needs to be estimated [127]. In this regard, PRC can be directly linked to PAPC, as the ith nowcast made at timestamp t in PAPC can be also used as P_{avai} for $t + i$. In fact, in most cases, PRC and PRRC are simultaneously required by system operators [15]. Based on this viewpoint, it is necessary to seek a way to coordinate both control operations.

Towards a unified PRRC standard

As the penetration of PV systems keeps growing, PRRC is becoming a universal requirement to facilitate grid operations. As of now, a range of PRRC regulations have been imposed by different grid operators (see Table 1.1). Besides the different timescales at which these standards are enacted, one major concern is the unit of RR limit. For example, given the Danish 100 kW/s RR limit, it can be seen in Chapter 5 that a 60 MW system can frequently exceed the limitation. On the other hand, a 1 MW system is found seldom violating the regulation. In this regard, one may conclude that a larger PV system could suffer more from RR violations. However, the situation reverses when a RR limit is set up in percentage of capacity. In that case, it leads to a conclusion that a smaller system is more likely to violate the regulation. For PV system operators, such conflicted conclusion apparently can mislead the PRRC options.

Another confusion comes from the calculation of RR. Provided with the Puerto Rico 10%/min RR limit, there are actually several ramp calculation methods available—one may use the difference between two endpoints of any 60-s interval, or use the minimum and maximum values between two endpoints, or use the whole differencing sum of the 60-s time series. Besides, the "10%/min" can also be expressed in different schemes, e.g., 5%/30 s, 1%/6 s, 0.16%/1 s, etc. For each scheme, different ramp calculation methods also apply. Therefore, when enacting a PRRC regulation, it is suggested that the RR limit should be more carefully defined with specific averaging time.

Lastly, when it comes to a regulation, the penalty of violating the rule naturally needs to be declared. Due to the extra cost for system operators to comply with RR regulations, e.g., forecasting system deployment, power curtailment, and ESS installation, it is worthwhile considering penalties on RR violations, so that the monetary value of PRRC can be quantified, and the ESS can be more properly sized. Unfortunately, until the writing of this thesis, such penalty is not yet being imposed in a standard regulation.

In light of the above, there is an urgent need to promote a unified PRRC standard. However, unifying does not mean that all regulations should necessarily share a similar RR limit. Instead, they ought to be formulated under a unified framework, including unit, the way of calculating RR, and revolving around penalty.

Appendix A

Sensor Deployment

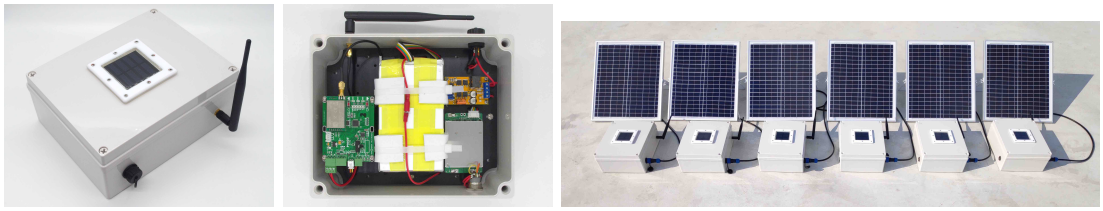
The sensor developed herein is based on a 5 cm $\times$ 5 cm mini solar cell. Sensing directly using a solar cell shows advantages of presenting more similar characteristics to the target PV systems, e.g. the sensitivity to ambient temperature, humidity, etc. Hence, the sensor becomes more reliable on detecting CSMV impacts, with less measurement variances caused by "non-cloud" factors.

The basic theory behind is that the short-circuit current of a solar cell can be quasi-linear to the GHI it receives [128]. In this sense, once we can measure the instantaneous short-circuit current of the solar cell, and establish an appropriate conversion model, the corresponding GHI can be derived. Given a tuning parameter β and intercept α, a linear conversion model can be written as:

$$G = \beta \cdot \frac{G_{STC}}{I_{SCS}} \cdot I_{SC} + \alpha, \tag{A.1}$$

where G and I_{SC} denote the measured GHI and short-circuit current respectively, and G_{STC} and I_{SCS} correspond to the GHI and short-circuit current at standard test conditions ($G_{STC} = 1000$ W/m^2 with cell temperature of 25°C or 298 K).

In order to obtain the short-circuit current of the solar cell, a small precise resistor (0.1 Ω) is connected in shunt. In this way, the current through the resistor can be estimated as the short-circuit current. For each sensor, the short-circuit current is



(a) (b) (c)

Figure A.1: Solar sensor prototype, (a) water-proof shell, (b) inside configuration, (c) PV panel for self-charging.

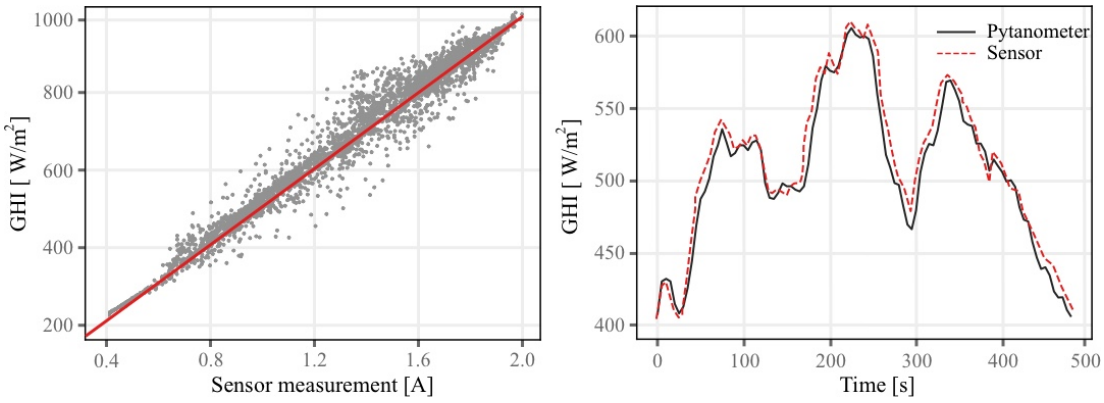


Figure A.2: **Left**. Example of solar sensor calibration with a pyranometer ($\beta = 1.31$, $\alpha = 78.26$). **Right.** Comparison of calibrated sensor measurements and pyranometer readings during cloud transitions (temporal resolution of 1 s).

recorded at 1-s resolution using a STM32 micro-controller, powered by a Li-Po battery. After converting the short-circuit current to GHI using Equation (A.1), the GHI data are packed and transmitted to the local server via a LoRa wireless communication module. To assure a consistent and autonomous outdoor operation, the sensor is equipped with a water-proof shell and a PV panel for self-charging, see Figure A.1. The cable-less configuration also provides more flexibilities to the network design.

Regarding the calibration, a commercial pyranometer is placed closely to each sensor, and data from a total of 12 hours are applied. Figure A.2 shows a calibration example for a single sensor. In this case, the regression parameters β and α are found as 1.31 and 78.26, respectively. The calibrated sensor measurements are compared

with the pyranometer readings over several cloud transition periods, and the mean absolute error is found to be 1.25%. It should be noticed that for a more accurate measurement, the individual cell temperature information is demanded. However, the additional temperature data will occupy more communication channels, which can halve the number of sensors in the network.

Appendix B

Probabilistic Nowcasts for PAPC

Figure B.1 shows an example of applying probabilistic solar nowcasts to PAPC operation. Three different PIs, i.e., 10%, 50%, and 90% PIs are extracted from the predictive distributions generated by OLS.

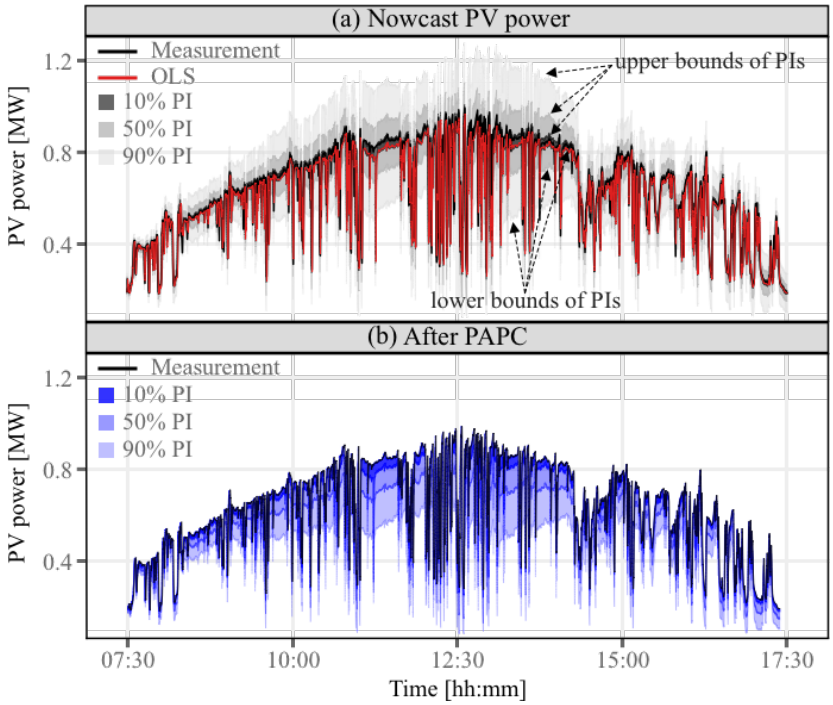


Figure B.1: The results of (a) OLS probabilistic nowcasts, and (b) PAPC on an example operating day, following the operating timeline in Figure 4.2.

To implement PAPC, the lower bounds of the PIs are fed into the control process described in Equations (4.1)-(4.6). Referring to Figure B.1(b), PAPC with wider PIs will sacrifice more energy for smoothing. In this example day, the RSRs using 10%, 50%, and 90% PIs are 95.83%, 97.22%, and 98.09, while the ECRs are found as 6.26%, 16.28, and 35.20, respectively.

Appendix C

Reconstruction of PV Matrix

Consider a PV array consisting of 2 strings, with each string series-connected by 2 modules ($n_m = 2$, $n_s = 2$), a PV system composed of 2×2 ($a = 2$, $b = 2$) arrays then can be modeled as:

$$\mathbf{\Pi} = \begin{pmatrix} (\gamma_{11}^{11}, \chi_{11}^{11}) & (\gamma_{12}^{11}, \chi_{12}^{11}) & (\gamma_{11}^{12}, \chi_{11}^{12}) & (\gamma_{12}^{12}, \chi_{12}^{12}) \\ (\gamma_{21}^{11}, \chi_{21}^{11}) & (\gamma_{22}^{11}, \chi_{22}^{11}) & (\gamma_{21}^{12}, \chi_{21}^{12}) & (\gamma_{22}^{12}, \chi_{22}^{12}) \\ (\gamma_{11}^{21}, \chi_{11}^{21}) & (\gamma_{12}^{21}, \chi_{12}^{21}) & (\gamma_{11}^{22}, \chi_{11}^{22}) & (\gamma_{12}^{22}, \chi_{12}^{22}) \\ (\gamma_{21}^{21}, \chi_{21}^{21}) & (\gamma_{22}^{21}, \chi_{22}^{21}) & (\gamma_{21}^{22}, \chi_{21}^{22}) & (\gamma_{22}^{22}, \chi_{22}^{22}) \end{pmatrix}, \tag{C.1}$$

where $\mathbf{\Pi} \in \mathbb{R}^{4\times4}$, and the PV arrays are framed in grey. Each element in matrix (C.1) is composed of a pair of variables, indicating the received irradiance and temperature of a module respectively, and covering a geographic dimension of 1.98 m $\times$ 1 m. Notice that the module temperature will be omitted in the following matrices for simplicity.

To transform the PV matrix into the same geographic dimension as the shadow matrix, that is 1 m $\times$ 1 m, each element in matrix (C.1) is split into $(1.98/1) \times (1/1)$ elements. Since the number of elements should be an integer, the 1.98 elements are rounded up to 2, which gives the normalized PV matrix:

$$\mathbf{\Pi}^n = \begin{pmatrix} \begin{Bmatrix} \gamma'^{11}_{11} \\ \gamma''^{11}_{11} \\ \gamma'^{11}_{21} \\ \gamma''^{11}_{21} \end{Bmatrix} \begin{Bmatrix} \gamma'^{11}_{12} \\ \gamma''^{11}_{12} \\ \gamma'^{11}_{22} \\ \gamma''^{11}_{22} \end{Bmatrix} & \begin{Bmatrix} \gamma'^{12}_{11} \\ \gamma''^{12}_{11} \\ \gamma'^{12}_{21} \\ \gamma''^{12}_{21} \end{Bmatrix} \begin{Bmatrix} \gamma'^{12}_{12} \\ \gamma''^{12}_{12} \\ \gamma'^{12}_{22} \\ \gamma''^{12}_{22} \end{Bmatrix} \\ \begin{Bmatrix} \gamma'^{21}_{11} \\ \gamma''^{21}_{11} \\ \gamma'^{21}_{21} \\ \gamma''^{21}_{21} \end{Bmatrix} \begin{Bmatrix} \gamma'^{21}_{12} \\ \gamma''^{21}_{12} \\ \gamma'^{21}_{22} \\ \gamma''^{21}_{22} \end{Bmatrix} & \begin{Bmatrix} \gamma'^{22}_{11} \\ \gamma''^{22}_{11} \\ \gamma'^{22}_{21} \\ \gamma''^{22}_{21} \end{Bmatrix} \begin{Bmatrix} \gamma'^{22}_{12} \\ \gamma''^{22}_{12} \\ \gamma'^{22}_{22} \\ \gamma''^{22}_{22} \end{Bmatrix} \end{pmatrix}, \tag{C.2}$$

where $\mathbf{\Pi}^n \in \mathbb{R}^{8\times4}$, and a PV module is denoted by two elements grouped in braces, with each element showing a geographic dimension of 1 m $\times$ 1 m (this gives a module dimension of 2 m $\times$ 1 m). When interacted with the shadow matrix, we take the mean value of the two braced elements as the irradiance value received by a PV module.

To account for the gaps between adjacent PV strings and arrays, additional rows and columns of zeros can be interpolated. Given a geographic dimension of 1 m $\times$ 1 m, a gap of 2 m in rows and columns can be modeled by inserting two rows and columns of zeros respectively. A complete PV matrix with string and array gaps interpolated is represented by:

$$\mathbf{\Pi}^* = \begin{pmatrix} \begin{Bmatrix} \gamma'^{11}_{11} \\ \gamma''^{11}_{11} \\ \gamma'^{11}_{21} \\ \gamma''^{11}_{21} \end{Bmatrix} \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix} \begin{Bmatrix} \gamma'^{11}_{12} \\ \gamma''^{11}_{12} \\ \gamma'^{11}_{22} \\ \gamma''^{11}_{22} \end{Bmatrix} & \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix} & \begin{Bmatrix} \gamma'^{12}_{11} \\ \gamma''^{12}_{11} \\ \gamma'^{12}_{21} \\ \gamma''^{12}_{21} \end{Bmatrix} \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix} \begin{Bmatrix} \gamma'^{12}_{12} \\ \gamma''^{12}_{12} \\ \gamma'^{12}_{22} \\ \gamma''^{12}_{22} \end{Bmatrix} \\ \begin{matrix} 0 \\ 0 \end{matrix} \quad \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \quad \begin{matrix} 0 \\ 0 \end{matrix} & \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} & \begin{matrix} 0 \\ 0 \end{matrix} \quad \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix} \quad \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{Bmatrix} \gamma'^{21}_{11} \\ \gamma''^{21}_{11} \\ \gamma'^{22}_{21} \\ \gamma''^{22}_{21} \end{Bmatrix} \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix} \begin{Bmatrix} \gamma'^{21}_{12} \\ \gamma''^{21}_{12} \\ \gamma'^{21}_{22} \\ \gamma''^{21}_{22} \end{Bmatrix} & \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix} & \begin{Bmatrix} \gamma'^{22}_{11} \\ \gamma''^{22}_{11} \\ \gamma'^{22}_{21} \\ \gamma''^{22}_{21} \end{Bmatrix} \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix} \begin{Bmatrix} \gamma'^{22}_{12} \\ \gamma''^{22}_{12} \\ \gamma'^{22}_{22} \\ \gamma''^{22}_{22} \end{Bmatrix} \end{pmatrix}, \tag{C.3}$$

where $\mathbf{\Pi}^* \in \mathbb{R}^{10\times10}$.

Appendix D

Nine Shadow Patterns

Figure D.1 shows the nine cloud shadow patterns used in Section 4.1 and Section 4.2 for generalization. The shadow patterns are produced with $N = 513$, $S = 50\%$ total pixels, and $\alpha = 0.9$.

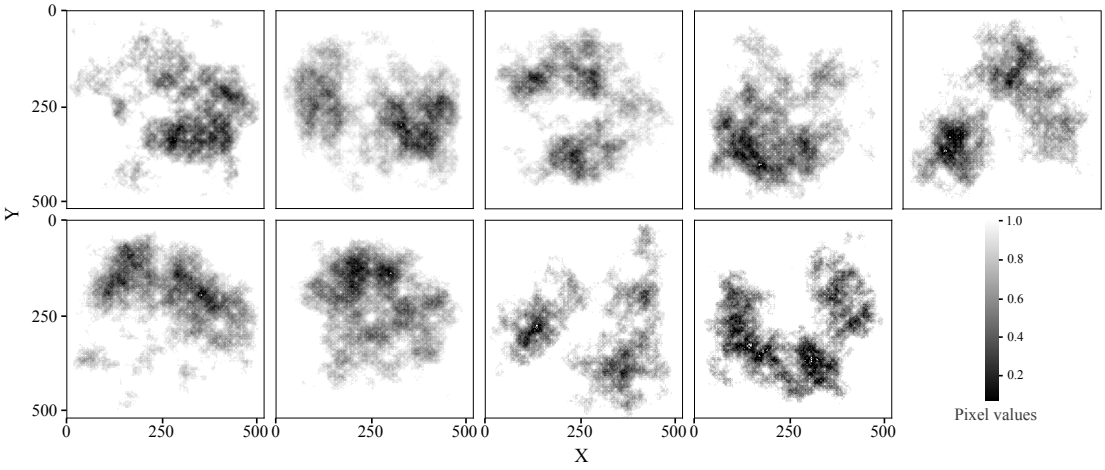


Figure D.1: Nine cloud shadow patterns used for generalization.

Appendix E

Identification of Shadow Characteristics

This appendix introduces the identification of shadow characteristics using XJTLU sensor network. The network is composed of 6 sensors, and measures GHI data with sampling time of 1 s. Notice that although it can be feasible to measure the shadow intensity and shadow velocity based on a sensor network, the shadow size still remains problematic since the measurements can be limited by network dimensions. In this regard, we measure the shadow length instead. As an alternative, advanced imaging systems (e.g., sky imagers and shadow cameras) can also be promising for cloud shadow detection. However, difficulties can rise during the conversion from cloud conditions to ground level irradiance due to the inaccurate measurement of cloud base height. For practical applications, one may use a combination of multiple measuring systems to obtain more precise shadow properties. The method presented herein is only meant to provide a possible solution when sensor network measurements are available.

E.1 Shadow velocity

To identify the shadow velocity, the MCP algorithm previously introduced in Section 2.2.1 is performed on the sensor network. In this case, sensor S_3 is selected as the

central sensor, and the 6 sensors can form 5 sensor pairs.

E.2 Shadow intensity

Figure E.1 shows an example of an identified shadow transition, which is based on the measurements of central sensor S_3. As suggested in [64], a shadow transition is identified when an irradiance drop is followed by steady shading and then by an irradiance rise. During the shadow transition, irradiance drops from 802.2 W to 200.8 W (the minimum value), thus the shadow intensity is found as 0.75.

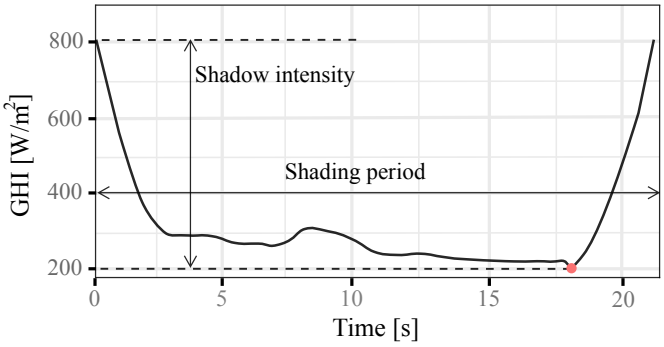


Figure E.1: Example of an identified shadow transition, with shadow intensity 0.75 and shading period 21.5 s.

E.3 Shadow length

The shadow length is calculated as the product of shadow velocity and shading period. For the shadow transition shown in Figure E.1, the shading period is 21.5 s. Consider a shadow velocity of 10 m/s, it then leads to a shadow length of 215 m.

In order to generate a shadow pattern, however, Equation (5.5) requires a shadow size as input. Since we only measure the shadow length, the equation should be modified, which is given by:

$$h = \underset{\mu}{\texttt{argmin}}(|L_\mu \cdot d - L| < \epsilon), \quad h_{min} \leq \mu \leq h_{max}, \tag{E.1}$$

where $L_\mu \leq N$ denotes the number of non-zero pixels in a single row of shadow matrix obtained at height μ, d is the pixel dimension, L is a user-defined shadow length, and ϵ is the error tolerance (we use $\epsilon = 2$ after several calibration tests). The equation states that given a cutting plane of height μ and the resultant shadow matrix $\mathbf{\Omega}_\mu$, if there exists a row in the matrix $\mathbf{\Omega}_\mu$ whose non-zero pixels satisfy the relationship in Equation E.1, then the height μ will be used for shadow generation.

References

[1] Q. Wang, M. Su, R. Li, and P. Ponce, "The effects of energy prices, urbanization and economic growth on energy consumption per capita in 186 countries," *J. Clean. Prod*, vol. 225, pp. 1017–1032, 2019.

[2] M. Hu, R. Li, W. You, Y. Liu, and C. Lee, "Spatiotemporal evolution of decoupling and driving forces of CO2 emissions on economic growth along the Belt and Road," *J. Clean. Prod*, vol. 277, p. 123272, 2020.

[3] M. P. Pablo-Romero, A. Sánchez-Braza, and A. Galyan, "Renewable energy use for electricity generation in transition economies: Evolution, targets and promotion policies," *Renew. Sustan. Energy Rev.*, vol. 138, p. 110481, 2021.

[4] A. Zervos, C. Lins, and J. Muth, "RE-thinking 2050-A 100Vision for the European Union," in *EREC European Renewable Energy Council*, p. 76, 2010.

[5] J. Yan, Y. Yang, P. Campana, and J. He, "City-level analysis of subsidy-free solar photovoltaic electricity price, profits and grid parity in China," *Nat. Energy*, vol. 4, pp. 709–717, 2019.

[6] IEA, "IEA PVPS Annual Report," 2019.

[7] Q. Hou, N. Zhang, E. Du, M. Miao, F. Peng, and C. Kang, "Probabilistic duck curve in high PV penetration power system: Concept, modeling, and empirical analysis in China," *Appl. Energy*, vol. 242, pp. 205–215, 2019.

[8] X. Sun, J. M. Bright, C. A. Gueymard, B. Acord, P. Wang, and N. A. Engerer, "Worldwide performance assessment of 75 global clear-sky irradiance models using principal component analysis," *Renew. Sustan. Energy Rev.*, vol. 111, pp. 550–570, 2019.

[9] P. Ineichen and R. Perez, "A new airmass independent formulation for the Linke turbidity coefficient," *Sol. Energy*, vol. 73, pp. 151–157, 2002.

[10] S. Shivashankar, S. Mekhilef, H. Mokhlis, and M. Karimi, "Mitigating methods of power fluctuation of photovoltaic (PV) sources – A review," *Renew. Sustan. Energy Rev.*, vol. 59, pp. 1170–1184, 2016.

[11] R. Rajan, F. M.Fernandez, and Y. Yang, "Primary frequency control techniques for large-scale PV-integrated power systems: A review," *Renew. Sustan. Energy Rev.*, vol. 144, p. 110998, 2021.

[12] S. Marharjan, D. S. Kumar, and A. M. Khambadkone, "Enhancing the voltage stability of distribution network during PV ramping conditions with variable speed drive loads," *Appl. Energy*, vol. 264, p. 114733, 2020.

[13] M. Lave, J. Kleissl, A. Ellis, and F. Mejia, "Simulated PV power plant variability: Impact of utility-imposed ramp limitations in Puerto Rico," *2013 IEEE 39th Photovoltaic Specialists Conference (PVSC)*, pp. 1817–1821, 2013.

[14] E. Troester, "New german grid codes for connecting pv systems to the medium voltage power grid," in *Proc. of 2nd Int. Workshop Concentrating Photovoltaic Power plants: Opt. Design, Prod., Grid Connection*, 2009.

[15] Energinet.dk, "Technical regulation 3.2.2 for pv power plants with a power output above 11 kw," tech. rep. doc. 14/17997-39, 2016.

[16] V. Gevorgian and S. Booth, "Review of PREPA technical requirements for interconnecting wind and solar generation," Tech. Rep. NREL/TP-5D00-57089, National Renewable Energy Laboratory (NREL), 2013.

[17] B. Crăciun, T. Kerekes, D. Séra, R. Teodorescu, and U. D. Annakage, "Power ramp limitation capabilities of large PV power plants with active power reserves," *IEEE Trans. Sustain. Energy*, vol. 8, pp. 573–581, 2017.

[18] R. van Haaren, M. Morjaria, and V. Fthenakis, "An energy storage algorithm for ramp rate control of utility scale PV (photovoltaics) plants," *Energy*, vol. 91, pp. 894–902, 2015.

[19] S. A. Pourmousavi and T. K. Saha, "Evaluation of the battery operation in ramp-rate control mode within a PV plant: A case study," *Sol. Energy*, vol. 166, pp. 242–254, 2018.

[20] I. de la Parra, J. Marcos, and L. Marroyo, "Dealing with the implementation of ramp-rate control strategies – Challenges and solutions to enable PV plants with energy storage systems to operate correctly," *Sol. Energy*, vol. 169, pp. 242–248, 2018.

[21] A. Sangwongwanich, Y. Yang, and F. Blaabjerg, "A cost-effective power ramp-rate control strategy for single-phase two-stage grid-connected photovoltaic systems," *Proceedings of the 8th Annual IEEE Energy Conversion Congress and Exposition, ECCE 2016*, pp. 1–7, 2016.

[22] I. de la Parra, J. Marcos, M. García, and L. Marroyo, "Improvement of a control strategy for PV power ramp-rate limitation using the inverters: Reduction of the associated energy losses," *Sol. Energy*, vol. 127, pp. 262–268, 2016.

[23] E. Bullich-Massagué, M. Aragüés-Peñalba, A. Sumper, and O. Boix-Aragones, "Active power control in a hybrid pv-storage power plant for frequency support," *Sol. Energy*, vol. 144, pp. 49–62, 2017.

[24] Á. D. J. do Nascimento and R. Rüther, "Evaluating distributed photovoltaic (PV) generation to foster the adoption of energy storage systems (ESS) in time-of-use frameworks," *Sol. Energy*, vol. 208, pp. 917–929, 2020.

[25] J. Marcos, O. Storkel, L. Marroyo, M. Garcia, and E. Lorenz, "Storage requirements for PV power ramp-rate control," *Sol. Energy*, vol. 99, pp. 28–35, 2014.

[26] H. D. Tafti, A. Sangwongwanich, Y. Yang, J. Pou, G. Konstantinou, and F. Blaabjerg, "An adaptive control scheme for flexible power point tracking in photovoltaic systems," *IEEE Trans. Ind. Electron.*, vol. 34, pp. 5451–5463, 2018.

[27] I. de la Parra, J. Marcos, M. Garcia, and L.Marroyo, "Control strategies to use the minimum energy storage requirement for PV power ramp-rate control," *Sol. Energy*, vol. 111, no. 332-343, 2015.

[28] A. Ahmed and M. Khalid, "A review on the selected applications of forecasting models in renewable power systems," *Renew. Sustan. Energy Rev.*, vol. 100, pp. 9–21, 2019.

[29] D. Yang, J. Kleissl, C. Gueymard, H. Pedro, and C. Coimbra, "History and trends in solar irradiance and pv power forecasting: A preliminary assessment and review using text mining," *Solar Energy*, vol. 168, pp. 60–101, 2017.

[30] M. P. Almeida, M. Muñoz, I. de la Parra, and O. Perpiñán, "Comparative study of PV power forecast using parametric and nonparametric PV models," *Sol. Energy*, vol. 155, pp. 854–866, 2017.

[31] R. H. Inman, H. T. C. Pedro, and C. F. M. Coimbra, "Solar forecasting methods for renewable energy integration," *Prog. Energ. Combust.*, vol. 39, pp. 535–576, 2013.

[32] D. Larson, L. Nonnenmacher, and C. F. M. Coimbra, "Day-ahead forecasting of solar power output from photovoltaic plants in the American Southwest," *Renew. Energy*, vol. 91, pp. 11–20, 2016.

[33] K. Bakker, K. Whan, W. Knap, and M. Schmeits, "Comparison of statistical post-processing methods for probabilistic NWP forecasts of solar radiation," *Sol. Energy*, vol. 191, pp. 138–150, 2019.

[34] M. Diagne, M. David, P. Lauret, J. Boland, and N. Schmutz, "Review of solar irradiance forecasting methods and a proposition for small-scale insular grids," *Renew. Sustan. Energy Rev.*, vol. 27, pp. 65–76, 2013.

[35] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. M. de Pison, and F. Antonanzas-Torres, "Review of Photovoltaic Power Forecasting," *Sol. Energy*, vol. 136, pp. 78–111, 2016.

[36] D. Yang, W. M. Walsh, Z. Dong, P. Jirutitijaroen, and T. G. Reindl, "Blocking matching algorithms: Their applications and limitations in solar irradiance forecasting," *Energy Procedia*, vol. 33, pp. 335–342, 2013.

[37] I. Urbich, J. Bendix, and R. Müller, "A novel approach for the short-term forecast of the effective cloud albedo," *Remote Sens-Basel*, vol. 10, p. 955, 2018.

[38] G. M. Yagli, D. Yang, O. Gandhi, and D. Srinivasan, "Can we justify producing univariate machine-learning forecasts with satellite-derived solar irradiance?," *Appl. Energy*, vol. 259, p. 114122, 2020.

[39] J. Polo, S. Wilbert, J. Ruiz-Arias, R. Meyer, C. Gueymard, M. Súri, L. Martín, T. Mieslinger, P. Blanc, I. Grant, J. Boland, P. Ineichen, J. Remund, R. Escobar, A. Troccoli, M. Sengupta, K. Nielsen, D. Renne, N. Geuder, and T. Cebecauer, "Preliminary survey on site-adaptation techniques for satellite-derived and reanalysis solar radiation datasets," *Sol. Energy*, vol. 132, pp. 25–37, 2016.

[40] B. Li and J. Zhang, "A review on the integration of probabilistic solar forecasting in power systems," *Sol. Energy*, vol. 210, pp. 68–86, 2020.

[41] C. W. Chow, S. Belongie, and J. Kleissl, "Cloud motion and stability estimation for intra-hour solar forecasting," *Sol. Energy*, vol. 115, pp. 645–655, 2015.

[42] F. Barbieri, S. Rajakaruna, and A. Ghosh, "Very short-term photovoltaic power forecasting with cloud modeling: A review," *Renew. Sustan. Energy Rev.*, vol. 75, pp. 242–263, 2017.

[43] P. Kuhn, S. Wilbert, C. Prahl, D. Schüler, T. Haase, T. Hirsch, M. Wittmann, L. Ramirez, L. Zarzalejo, A. Meyer, L. Vuilleumier, P. Blanc, and R. Pitz-Paal, "Shadow camera system for the generation of solar irradiance maps," *Sol. Energy*, vol. 157, pp. 157–170, 2017.

[44] A. Lorenzo, W. Holmgren, and A. Cronin, "Irradiance forecasts based on an irradiance monitoring network, cloud motion, and spatial averaging," *Sol. Energys*, vol. 122, pp. 1158–1169, 2015.

[45] P. Kuhn, M. Wirtz, S. Wilbert, J. L. Bosch, G. Wang, L. Ramirez, D. Heinemann, and R. Pitz-Paal, "Field validation and benchmarking of a cloud shadow speed sensor," *Sol. Energy*, vol. 173, pp. 229–245, 2018.

[46] B. Elsinga and W. Sark, "Short-term peer-to-peer solar forecasting in a network of photovoltaic systems," *Appl. Energy*, vol. 206, pp. 1464–1483, 2017.

[47] D. Yang and N. Chen, "Expanding existing solar irradiance monitoring network using entropy," *IEEE Trans. Sustain. Energy*, vol. 6, pp. 1208–1215, 2015.

[48] E. Wiemken, H. G. Beyer, W. Heydenreich, and K. Kiefer, "Power characteristics of PV esembles: experiences from the combined power production of 100 grid connected PV systems distributed over the area of Germany," *Sol. Energy*, vol. 70, pp. 513–518, 2001.

[49] E. Curtright and J. Apt, "The character of power output from utility-scale photovoltaic systems," *Prog. Photovolt. Res. Appl.*, vol. 16, pp. 241–247, 2008.

[50] J. Marcos, L. Marroyo, E. Lorenzo, D. Alvira, and E. lzco, "Power output fluctuations in large scale PV plants: one year observations with one second resolution and a derived analytic model," *Prog. Photovolt. Res. Appl.*, vol. 19, pp. 218–227, 2011.

[51] M. Dhimish, V. Holmes, B. Mehrdadi, M. Dales, B. Chong, and L. Zhang, "Seven

indicators variations for multiple PV array configurations under partial shading and faulty PV conditions," *Sol. Energy*, vol. 113, pp. 438–460, 2017.

[52] N. Belhaouas, M. S. A. Cheikh, P. Agathoklis, M. R. Oularbi, B. Amrouche, K. Sedraoui, and N. Djilali, "PV array power output maximization under partial shading using new shifted PV array arrangements," *Appl. Energy*, vol. 187, pp. 326–337, 2017.

[53] D. P. Winston, S. Kumaravel, B. P. Kumar, and S. Devakirubakaran, "Performance improvement of solar PV array topologies during various partial shading conditions," *Sol. Energy*, vol. 196, pp. 228–242, 2020.

[54] S. R. Pendem and S. Mikkili, "Modelling and performance assessment of PV array topologies under partial shading conditions to mitigate the mismatching power losses," *Sol. Energy*, vol. 160, pp. 303–321, 2018.

[55] K. Lappalainen and S. Valkealahti, "Size of a basic simulation unit in PV system partial shading studies," in *35th European Photovoltaic Solar Energy Conference and Exhibition*, pp. 1647–1651, 2018.

[56] C. Cai and D. C. Aliprantis, "Cumulus cloud shadow model for analysis of power system with photovoltaics," *IEEE Trans. Power Syst.*, vol. 28, pp. 4496–4506, 2013.

[57] M. Jazayeri, K. Jazayeri, and S. Uysal, "Generation of spatially dispersed irradiance time-series based on real cloud patterns," *Sol. Energy*, vol. 158, pp. 977–994, 2017.

[58] M. Lave, M. J. Reno, and R. J. Broderick, "Creation and value of synthetic high-frequency solar inputs for distribution system qsts simulations," in *2017 IEEE 44th Photovoltaic Specialist Conference (PVSC)*, pp. 3031–3033, 2017.

[59] T. E. Hoff and R. Perez, "Quantifying pv power output variability," *Sol. Energy*, vol. 84, pp. 1782–1793, 2010.

[60] T. E. Hoff and R. Perez, "Modeling PV fleet output variability," *Sol. Energy*, vol. 86, pp. 2177–2189, 2012.

[61] M. Lave, J. Kleissl, and J. S. Stein, "A Wavelet-based variability model (WVM) for solar PV power plants," *IEEE Trans. Sustain. Energy*, vol. 4, pp. 501–509, 2013.

[62] A. R. Dyreson, E. R. Morgan, S. H. Monger, and T. L. Acker, "Modeling solar irradiance smoothing for large PV power plants using a 45-sensor network and the wavelet variability model," *Sol. Energy*, vol. 110, pp. 482–495, 2014.

[63] J. Bosch, Y. Zheng, and J. Kleissl, "Deriving cloud velocity from an array of solar radiation measurements," *Sol. Energy*, vol. 87, pp. 196–203, 2013.

[64] K. Lappalainen and S. Valkealahti, "Analysis of shading periods caused by moving clouds," *Sol. Energy*, vol. 135, pp. 188–196, 2016.

[65] J. Bosch and J. Kleissl, "Cloud motion vectors from a network of ground sensors in a solar power plant," *Sol. Energy*, vol. 95, pp. 13–20, 2013.

[66] D. Yang, Z. Dong, T. Reindl, P. Jirutitijaroen, and W. M. Walsh, "Solar irradiance forecasting using spatio-temporal empirical kriging and vector autogression models with parameter shrinkage," *Sol. Energy*, vol. 103, pp. 550–562, 2014.

[67] D. Yang, Z. Ye, L. Lim, and Z. Dong, "Very short term irradiance forecasting using the lasso," *Sol. Energy*, vol. 114, pp. 314–326, 2015.

[68] M. Sengupta and A. Andreas, "Oahu Solar Measurement Grid (1-Year Archive): 1-Second Solar Irradiance," Tech. Rep. DA-5500-56506, National Renewable Energy Laboratory (NREL), 2010.

[69] J. Marcos, L. Marroyo, E. Lorenzo, D. Alvira, and E. lzco, "From irradiance to output power fluctuations: the pv plant as a low pass filter," *Prog. Photovolt. Res. Appl.*, vol. 19, pp. 505–510, 2011.

[70] L. M. Hinkelman, "Differences between along-wind and cross-wind solar irradiance variability on small spatial scales," *Sol. Energy*, vol. 88, pp. 192–203, 2013.

[71] K. Lappalainen and S. Valkealahti, "Apparent velocity of shadow edges caused by moving clouds," *Sol. Energy*, vol. 138, pp. 47–52, 2016.

[72] R. Fu, D. Chung, T. Lowder, D. Feldman, K. Ardani, and R. Margolis, "U.S. solar photovoltaic system cost benchmark: Q1 2016," tech. rep., National Renewable Energy Laboratory (NREL), 2016.

[73] M. Jamaly, J. L. Bosch, and J. Kleissl, "Aggregate ramp rates of distributed photovoltaic systems in San Diego county," *IEEE Trans. Sustain. Energy*, vol. 4, pp. 519–526, 2013.

[74] W. Omran, M. Kazerani, and M. Salama, "Investigation of methods for reduction of power fluctuations generated from large grid-connected photovoltaic systems," *IEEE Trans. Energy Convers.*, vol. 26, pp. 318–327, 2011.

[75] A. Makibar, L. Narvarte, and E. Lorenzo, "On the relation between battery size and pv power ramp rate limitation," *Sol. Energy*, vol. 142, pp. 182–193, 2017.

[76] B. Xu, Y. Dvorkin, D. Kirschen, C. A. Silva-Monroy, and J. Watson, "A comparison of policies on the participation of storage in u.s. frequency regulation markets," in *Power and Energy Society General Meeting (PESGM)*, pp. 1–5, 2016.

[77] M. Lipperheide, J. Bosch, and J. Kleissl, "Embedded nowcasting method using cloud speed persistence for a photovoltaic power plant," *Sol. Energy*, vol. 112, pp. 232–238, 2015.

[78] X. Chen, Y. Du, and H. Wen, "Forecasting based power ramp-rate control for PV systems without energy storage," in *IEEE 3rd International Future Energy Electronics Conference and ECCE Asia (IFEEC 2017 - ECCE Asia)*, pp. 733–738, 2017.

[79] X. Chen, Y. Du, W. Xiao, and S. Lu, "Power ramp-rate control based on power Forecasting For PV Grid-tied Systems with Minimum Energy Storage," in *The 43rd Annual Conference of the IEEE Industrial Electronics Society (IECON)*, (Beijing), 2017.

[80] S. Achleitner, A. Kamthe, and A. E. Cerpa, "Sips: Solar irradiance prediction system," *ACM/IEEE IPSN*, 2014.

[81] A. W. Aryaputera, D. Yang, L. Zhao, and W. M. Walsh, "Very short-term irradiance forecasting at unobserved locations using spatio-temporal kriging," *Sol. Energy*, vol. 122, pp. 1266–1278, 2015.

[82] X. G. Agoua, R. Girard, and G. Kariniotakis, "Short-term spatio-temporal forecasting of photovoltaic power production," *IEEE Trans. Sustain. Energy*, vol. 9, pp. 538–546, 2018.

[83] D. Yang, "On adding and removing sensors in a solar irradiance monitoring network for areal forecasting and pv performance evaluation," *Sol. Energy*, vol. 155, pp. 1417–1430, 2017.

[84] Y. Chu, B. Urquhart, S. Gohari, H.T J. Kleissl, and C. Coimbra, "Short-term reforecasting of power output from a 48 mw solar pv plant," *Sol. Energy*, vol. 112, pp. 68–77, 2015.

[85] H. Zhou, Y. Zhang, L. Yang, Q. Liu, K. Yan, and Y. Du, "Short-term photovoltaic power forecasting based on long short term memory neural network and attention mechanism," *IEEE Acess*, vol. 7, pp. 78063–78074, 2019.

[86] D. Yang, "A guideline to solar forecasting research practice: Reproducible, operational, probabilistic or physically-based, ensemble, and skill (ROPES)," *J. Renew. Sustain. Energy*, vol. 11, p. 022701, 2019.

[87] D. Yang, E. Wu, and J. Kleissl, "Operational solar forecasting for the real-time market," *Int. J. Forecasting*, vol. 35, pp. 1499–1519, 2019.

[88] Y. V. Makarov, P. V. Etingov, J. Ma, Z. Huang, and K. Subbarao, "Incorporating uncertainty of wind power generation forecast into power system operation, dispatch, and unit commitment procedures," *IEEE Trans. Sustain. Energy*, vol. 2, pp. 433–442, 2011.

[89] D. Yang and Z. Dong, "Operational photovoltaics power forecasting using seasonal time series ensemble," *Sol. Energy*, vol. 166, pp. 529–541, 2018.

[90] D. Yang, "Standard of reference in operational day-ahead deterministic solar forecasting," *J. Renew. Sustain. Energy*, vol. 11, p. 053702, 2019.

[91] M. Cui and J. Zhang, "Estimating ramping requirements with solar-friendly flexible ramping product in multi-timescale power system operations," *Appl. Energy*, vol. 225, pp. 27–41, 2018.

[92] H. D. Tafti, G. Konstantinou, C. D. Townsend, G. G. Farivar, S. Ceballos, J. Pou, and J. E. Fletcher, "Comparative analysis of flexible power point tracking algorithms in photovoltaic systems," in *2020 IEEE Energy Conversion Congress and Exposition (ECCE)*, pp. 110–115, 2020.

[93] A. Mashlakov, T. Kuronen, L. Lensu, A. Kaarna, and S. Honkapuro, "Assessing the performance of deep learning models for multivariate probabilistic energy forecasting," *Appl. Energy*, vol. 285, p. 116405, 2021.

[94] H. Pedro, D. Larson, and C. Coimbra, "A comprehensive dataset for the accelerated development and benchmarking of solar forecasting methods," *J. Renew. Sustain. Energy*, vol. 11, p. 036102, 2019.

[95] D. Yang, D. Meer, and J. Munkhammar, "Probabilistic solar forecasting benchmarks on a standardized dataset at Folsom, California," *Sol. Energy*, vol. 206, pp. 628–639, 2020.

[96] H. Pedro and C. Coimbra, "Nearest-neighbor methodology for prediction of intra-

hour global horizontal and direct normal irradiances," *Renew. Energy*, vol. 80, pp. 770–782, 2015.

[97] E. N. Lorenz, "Atmospheric perdictability as revealed by naturally occuring analogues," *J. Atmos. Sci.*, vol. 26, pp. 636–646, 1969.

[98] S. Alessandrini, L. Monache, S. Sperati, and G. Cervone, "An analog ensemble for short-term probabiistic solar power forecast," *Appl. Energy*, vol. 157, pp. 95–110, 2015.

[99] J. Marcos, L. Marroyo, E. Lorenzo, and M. García, "Smoothing of PV power fluctuations by geographical dispersion," *Prog Photovolt Res Appl*, vol. 20, pp. 226–237, 2012.

[100] M. Villalva, J. Gazoli, and E. Filho, "Comprehensive approach to modeling and simulation of photovoltaic arrays," *IEEE Trans. Power Electron.*, vol. 24, pp. 1198–1208, 2009.

[101] A. Sinha, K. C. Jana, and M. K. Das, "An inclusive review on different multi-level inverter topologies, their modulation and control strategies for a grid connected photo-voltaic system," *Sol. Energy*, vol. 170, pp. 633–657, 2018.

[102] O. Bingöl and B. Özkaya, "Analysis and comparison of different PV array configurations under partial shading conditions," *Sol. Energy*, vol. 160, pp. 336–343, 2018.

[103] A. Cabrera-Tobar, E. Bullich-Massagué, M. Aragüés-Peñalba, and O. Gomis-Bellmunt, "Topologies for large scale photovoltaic power plants," *Renew. Sustan. Energy Rev.*, vol. 59, pp. 309–319, 2016.

[104] S. Kouro, J. I. Leon, D. Vinnikov, and L. G. Franquelo, "Grid-Connected Photovoltaic Systems: An overview of Recent Research and Emerging PV Converter Technology," *IEEE Trans. Ind. Electron. M.*, vol. 9, pp. 47–61, 2015.

[105] K. Lappalainen and S. Valkealahti, "Output power variation of different PV array configurations during irradiance transitions caused by moving clouds," *Appl. Energy*, vol. 190, pp. 902–910, 2017.

[106] C. R. S. Reinoso, D. H. Milone, and R. H. Buitrago, "Simulation of photovoltaic centrals with dynamic shading," *Appl. Energy*, vol. 103, pp. 278–289, 2013.

[107] I. de la Parra, M. Muñoz, E. Lorenzo, M. García, J. Marcos, and F. Martínez-Moreno, "PV performance modelling: A review in the light of quality assurance for large PV plants," *Sol. Energy*, vol. 78, pp. 780–797, 2017.

[108] R. Ahmad, A. Murtaza, H. Sher, U. Shami, and S. Olalekan, "An analytical approach to study partial shading effects on pv array supported by literature," *Renew and Sustan Energy Rev*, vol. 74, pp. 721–732, 2017.

[109] N. Blair, A. Dobos, J. Freeman, T. Neises, and M. Wagner, "System Advisor Model, SAM 2014.1.14: General description," Tech. Rep. NREL/TP-6A20-61019, National Renewable Energy Laboratory (NREL), 2014.

[110] J. Marcos, I. de la Parra, M. García, and L. Marroyo, "Simulating the variability of dispersed large PV plants," *Prog. Photovolt. Res. Appl.*, vol. 25, pp. 680–691, 2016.

[111] M. Muñoz, M. García, I. de la Parra, J. Marcos, L. Marroyo, and E. Lorenzo, "Module temperature dispersion within a large PV array: Observations at the amareleja PV plant," *IEEE J. Photovolt.*, vol. 8, pp. 1725–1731, 2018.

[112] K. Lappalainen, G. Wang, and J. Kleissl, "Estimation of the largest expected photovoltaic power ramp rates," *Appl. Energy*, vol. 278, p. 115636, 2020.

[113] K. Lappalainen and S. Valkealahti, "Effects of irradiance transitions on the ouput power fluctuations of different pv array configurations," in *Proceedings of IEEE innovative smart grid technologies-Asia conference*, pp. 705–711, 2016.

[114] D. Yang, S. Alessandrini, J. Antonanzas, F. Antonanzas-Torres, V. Badescu, H. G. Beyer, R. Blaga, J. Boland, J. M. Bright, C. F. M. Coimbra, M. David, Â. Frimane, C. A. Gueymard, T. Hong, M. J. Kay, S. Killinger, J. Kleissl, P. Lauret, E. Lorenz, D. van der Meer, M. Paulescu, R. Perez, O. Perpiñán-Lamigueiro, I. M. Peters, G. Reikard, D. Renné, Y. Saint-Drenan, Y. Shuai, R. Urraca, H. Verbois, F. Vignola, C. Voyant, and J. Zhang, "Verification of deterministic solar forecasts," *Sol. Energy*, vol. 210, pp. 20–37, 2020.

[115] D. Yang and C. A. Gueymard, "Ensemble model output statistics for the separation of direct and diffuse components from 1-min global irradiance," *Sol. Energy*, vol. 208, pp. 591–603, 2020.

[116] H. Quan and D. Yang, "Probabilistic solar irradiance transposition models," *Renew. Sustan. Energy Rev.*, vol. 125, p. 109814, 2020.

[117] M. J. Mayer and G. Gróf, "Extensive comparison of physical models for photovoltaic power forecasting," *Appl. Energy*, vol. 283, p. 116239, 2021.

[118] M. R. Allen and D. A. Stainforth, "Towards objective probabilistic climate forecasting," *Nature*, vol. 419, p. 228, 2002.

[119] A. Brusaferri, M. Matteucci, and P. P. A. Vitali, "Bayesian deep learning based method for probabilistic forecast of day-ahead electricity prices," *Appl. Energy*, vol. 250, pp. 1158–1175, 2019.

[120] D. van der Meer, G. Wang, and J. Munkhammar, "An alternative optimal strategy for stochastic model predictive control of a residential battery energy management system with solar photovoltaic," *Appl. Energy*, vol. 283, p. 116289, 2021.

[121] F. Fahiman, S. Disano, S. Erfani, P. Mancarella, and C. Leckie, "Data-driven dynamic probabilistic reserve sizing based on dynamic bayesian belief networks," *IEEE Trans. Power Syst.*, vol. 34, pp. 2281–2291.

[122] R. Appino, J. Á. G. Ordiano, R. Mikut, T. Faulwasser, and V. Hagenmeyer, "On the use of probabilistic forecasts in scheduling of renewable energy sources coupled to storages," *Appl. Energy*, vol. 210, pp. 1207–1218, 2018.

[123] D. Yang and D. van der Meer, "Post-processing in solar forecasting: Ten overarching thinking tools," *Renew. Sustan. Energy Rev.*, vol. 140, p. 110735, 2021.

[124] G. Athanasopoulos, R. J. Hyndman, N. Kourentzes, and F. Petropoulos, "Forecasting with temporal hierarchies," *Eur. J. Open Res.*, vol. 262, pp. 60–74, 2017.

[125] D. Yang, H. Quan, V. Disfani, and C. Rodríguez-Gallegos, "Reconciling solar forecasts: Temporal hierarchy," *Sol. Energy*, vol. 158, pp. 332–346, 2017.

[126] A. Sangwongwanich, Y. Yang, and F. Blaabjerg, "A sensorless power reserve control strategy for two-stage grid-connected pv systems," *IEEE Trans. Power Electron.*, vol. 32, no. 11, pp. 8559–8569, 2017.

[127] X. Li, H. Wen, Y. Zhu, L. Jiang, Y. Hu, and W. Xiao, "A novel sensorless photovoltaic power reserve control with simple real-time mpp estimation," *IEEE Trans. Power Electron.*, vol. 34, pp. 7521 – 7531, 2018.

[128] M. Muñoz-García, A. Melado-Herreros, J. Balenzategui, and P. Barrerio, "Low-cost irradiance sensors for irradiation assessments inside tree canopies," *Sol. Energy*, vol. 103, pp. 143–153, 2014.