Forecasting Solar Radiation and Photovoltaic Power

Elke Lorenz, Bijan Nouri, Sylvain Cros, Kristian Pagh Nielsen, Rafael Fritz, Garrett Good, Marco Pierro, Guadalupe Sanchez Hernandez, Philippe Lauret, Mathieu David, et al.

9 Forecasting Solar Radiation and Photovoltaic Power

Elke Lorenz,[1] Bijan Nouri,[2] Sylvain Cros,[3] Kristian Pagh Nielsen,[4] Rafael Fritz,[5] Garrett Good,[5] Marco Pierro,[6] Guadalupe Sanchez Hernandez,[7] Philippe Lauret,[8] Mathieu David,[8] Rodrigo Amaro e Silva,[9,10] Carlos Fernandez Peruchena,[11] and Cristina Cornaro[12]

[1] *Fraunhofer Institute for Solar Energy Systems (Fraunhofer ISE), Germany*
[2] *German Aerospace Center (DLR), Institute of Solar Research, Germany*
[3] *Laboratoire de Météorologie Dynamique (LMD)/Institut Pierre-Simon Laplace (IPSL), École Polytechnique, Institut Polytechnique de Paris, Ecole normale supérieure (ENS), Université Paris Sciences & Lettres (PSL), Sorbonne Université, Centre national de la recherche Scientifique (CNRS), France*
[4] *Danish Meteorological Institute, Denmark*
[5] *Fraunhofer Institute for Energy Economics and Energy System Technology (Fraunhofer IEE), Germany*
[6] *Eurac Research, Bolzano, Italy*
[7] *Department of Physics, University of Jaen, Spain*
[8] *University of La Réunion, France*
[9] *Centre Observations, Impacts, Energie (O.I.E.), MINES Paris - PSL Research University, France*
[10] *Instituto Dom Luiz, Faculdade de Ciências, Universidade de Lisboa, Portugal*
[11] *National Renewable Energy Centre of Spain (CENER), Spain*
[12] *University of Rome Tor Vergata, Department of Enterprise Engineering, Italy*

Executive Summary

Solar power forecasting is essential for the reliable and cost-effective system integration of solar energy. It is necessary for a variety of applications that have specific requirements with respect to forecast horizon and spatiotemporal resolution, including the management of electric grids and energy management systems as well as the marketing of solar power.

Different input data and models are suitable for different forecast horizons, generally with a decreasing spatiotemporal resolution with increasing forecast horizon (see Fig. 9-1):

- Short-term irradiance forecasts up to 10–20 minutes ahead resolving irradiance ramps with a temporal resolution of minutes or even less are derived from all-sky imagers (ASIs).
- Irradiance forecasts up to several hours ahead with typical resolutions of 10–15 minutes are derived from satellite images covering large areas.
- Irradiance forecasts from several hours to days ahead essentially rely on numerical weather prediction (NWP) models, which have the capability to describe complex atmospheric dynamics, including advection as well as the formation and dissipation of clouds.

Complementing empirical and physical models, statistical and machine learning (ML) methods are widely used in solar irradiance and power forecasting. To train time-series models, the availability of irradiance and/or photovoltaic (PV) power measurements is crucial, as is proper quality control of the data. Assuming good data quality, these methods can be effectively applied to:

- Improve forecasts with empirical or physical models (postprocessing).

- Combine different input data and forecasts (model blending); here, very short-term forecasting, up to approximately 1 hour ahead, greatly benefits from the use of local online irradiance or PV power measurements as input.
- Derive PV power forecasts from meteorological forecasts.

Besides time-series forecasting, ML algorithms are increasingly used for image prediction using ASI or satellite data, e.g., to compute the optical flow in cloud motion approaches.

State-of-the-art PV power forecasting services do not rely on a single forecasting model but integrate different inputs and models. Prominent examples are intraday forecasting systems up to several hours ahead integrating online measurements, satellite-based forecasts, and NWP model forecasts or day-ahead forecasting systems combining different NWP models, both using statistical and/or ML algorithms for forecast optimization.

Besides forecasting for single PV power plants and portfolios of PV plants, the estimation and forecasting of regionally aggregated PV power is important for grid operators for marketing of PV power and grid management. Here, an additional challenge is that PV power is not measured at a sufficient resolution for most plants in many countries, and information on PV systems is incomplete. Still, because of spatial smoothing effects, forecast errors of regionally aggregated PV power as well as virtual power plants (VPPs) (normalized to their installed power) are much smaller than for single PV plants, depending on the size of the region and the set of PV plants contributing.

Forecast evaluations provide users with necessary information on forecast accuracy, assisting them in choosing between different forecasting services or assessing the risk when a forecast is used as a basis for decisions. Beyond general information on the overall accuracy of deterministic forecasts, probabilistic forecasts provide specific uncertainty information for each forecast value, depending on the weather conditions, and they allow for better risk management.

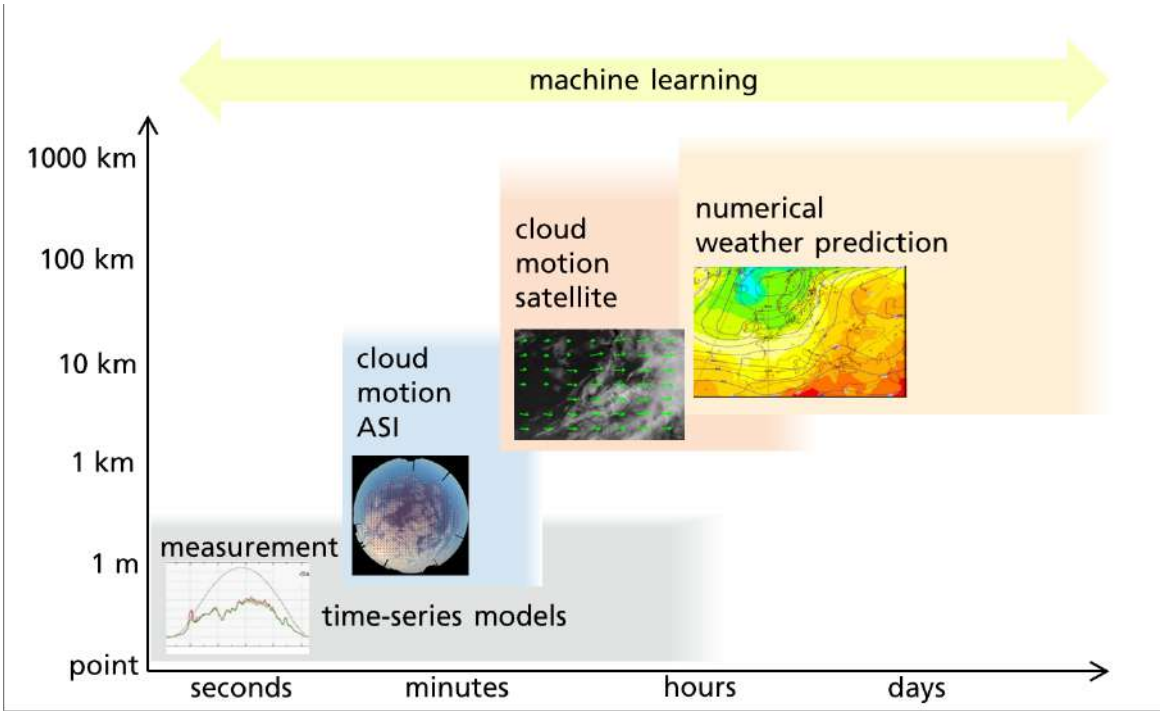

Figure 9-1. Different forecasting methods suitable for various spatial and temporal scales

Empirical and/or physical models are combined with statistical and/or ML models for forecast optimization. The spatial scales of the forecasting methods are defined by spatial resolution and spatial coverage. The temporal scales are defined by temporal resolution, update frequency, and forecast horizon.

Image by Fraunhofer ISE

9.1 Introduction

The variability of solar power generation poses a challenge for electric power systems to balance both generation and demand and to ensure the resilient operation of electric grids.

Besides the deterministic apparent course of the sun, this variability is largely determined by the inherent uncertainty of weather conditions. Therefore, forecasting plays a crucial role for power system operators in managing the electric grid, avoiding congestion, and following protocols (Bessa et al. 2014), and it is also essential for aggregators and energy traders (Pierro et al. 2017). It is one of the most cost-effective solutions to integrate variable renewable energy sources (Notton et al. 2018; Tuohy et al. 2015).

Solar forecasts are used in myriad contexts, with a variety of spatiotemporal scales, and their accuracy can have a great impact on power system performance. Various specific use cases and benefits of solar forecasting have been detailed in the literature, as described in this nonexhaustive list:

- Marketing of solar power by grid operators, plant operators, or direct marketers, reducing the need for balancing power or penalties, depending on national regulations. This is, e.g. described in Antonanzas et al. (2017) for day-ahead markets and in Kaur et al. (2016) for intraday markets. The impact of the update rate of solar forecasts (Cros, Sylvain et al. 2015) and of the presence of storage (David et al. 2021) have also been discussed.

Moreover, concentrating solar power (CSP) plants making use on direct normal irradiance (DNI) forecasts can benefit from bidding in the day-ahead (Kraas et al. 2013) and intraday (Law, Kay, and Taylor 2016) markets, potentially including the scheduling of coupled thermal storage.

- Scheduling storage systems to attenuate power fluctuations in large PV power plants (Marcos et al. 2013), particularly relevant in the presence of ramp-rate restrictions. Cirés et al. (2019) discuss how accurate forecasts reduce storage needs for this effect.
- Supporting generator scheduling of power systems, reducing the need for fossil fuels and costly fast-ramping generators (Brancucci Martinez-Anido et al. 2016), and mitigating operational imbalances (Pierro et al. 2020c). Some studies address the specificities of microgrids as well as insular and hybrid off-grid systems, e.g., (Ramahatana et al. 2022; Simoglou et al. 2014; Jamal et al. 2019).
- For CSP plants, Nouri et al. (2020) show how sky imager forecasts can increase plant efficiency and lifetime.

Solar power forecasting, including both PV and CSP, essentially relies on irradiance forecasting as a first step. Depending on the specific application and requirements regarding forecast horizon and spatiotemporal resolution, different input data and forecasting methods are customarily used. From short to long forecasting horizons (see Figure 9-1), the most important input data and solar forecasting methods are:

- **Time-series models based on local measurements:** They require on-site observations of irradiance and/or PV power, and possibly further meteorological variables that are processed using either statistical methods or artificial intelligence (AI) and ML algorithms, such as neural networks. They might provide meaningful forecasts even up to a few hours ahead under relatively stable sky conditions; however, these methods rarely perform well under variable-sky conditions, given the chaotic behavior of the cloud system and the limited information contained in point-wise observations.
- **Forecasts based on ASIs:** Using information on the local distribution of clouds, collected by one or more ground-based ASIs, the forecast skill can be enhanced relative to time-series models based on local measurements only. This information is key to the generation of solar irradiance forecasts with a temporal resolution on the order of seconds to minutes and a spatial resolution from 10–100 m covering a few square kilometers around the ASIs. The typical forecast horizon of these systems is 10–20 minutes, depending on cloud height and speed.
- **Forecasts based on data from geostationary satellites:** Forecasts up to several hours ahead benefit from wide-area observations of cloud fields. Because of their broad coverage, data from geostationary satellites are an appropriate source for these horizons. Satellite-based forecasts frequently use cloud motion vector (CMV) techniques to extrapolate cloud locations into the future. The typical spatial resolution is from 1–5 km^2 for the current generation of geostationary satellites, with forecast updates every 10–30 minutes, and the typical forecast horizon is 4–6 hours.
- **NWP:** NWP models constitute the main approach for forecast horizons more than several hours and up to several days or weeks ahead. These models predict the evolution of the atmospheric system, including the formation, advection, diffusion, and dissipation of clouds. They are based on a physical description of the dynamic processes occurring in the atmosphere by solving and parameterizing the governing system of equations, and

they depend on an observed set of initial conditions. Current global NWP models cover the Earth with a spatial resolution from approximately 0.1°–0.5° and a temporal resolution from 1–3 hours. Regional models, which are also referred to as limited area models or mesoscale models, typically have an hourly temporal resolution and a spatial resolution of a few kilometers in the covered area.

- **Postprocessing and model blending with statistical and ML models:** When historical or near-real-time on-site solar irradiance or PV yield observations are available, these described methods can be further improved by combining them with ML, resulting in hybrid methods. For NWP forecasts, model output statistic (MOS) techniques are often applied. Here too, satellite-derived irradiance might be used as a reference for model training. Further, state-of-the-art solar irradiance or PV power forecasting services do not rely on a single forecasting model but integrate different input and tools with statistical or ML algorithms, which is referred to as model blending.

Solar power forecasts can be derived based on irradiance forecasts using these different models. Converting irradiance to PV power forecasts requires PV system modeling using parametric PV simulation models and plant data, e.g., nominal power and orientation, and/or ML approaches, learning from PV power measurements.

Generally, solar irradiance and PV power forecasting approaches can be categorized into either physical and empirical models or statistical and ML approaches. Physical models (e.g., radiative transfer models, NWP models, PV module or inverter models) are based on solving basic physical equations. Empirical models are also based on physical considerations, but they do not describe all atmospheric processes in detail. Typical examples are cloud index models, parametric clear-sky models, and cloud motion approaches based on ASI or satellite data.

Statistical and ML algorithms establish the dependence of forecast values (predictands) on input variables (predictors) in a training phase by learning from historical data. Here, it is assumed that patterns in the historical datasets are repeated in the future and thus might be exploited for forecasting. These approaches include classical regression methods, such as autoregressive and autoregressive-integrated moving-average models as well as ML or AI techniques, such as artificial neural networks (ANNs), k-nearest neighbors, or support vector regression. They are widely applied for different purposes in irradiance and PV power forecasting. Coimbra and Pedro (2013) and Diagne et al. (2013) provided an overview of different statistical approaches used for solar irradiance forecasting. Voyant et al. (2017) and Sobri, Koohi-Kamali, and Rahim et al. (2018) reviewed the topic with a focus on the use of ML methods for solar radiation or power forecasting as well as for postprocessing.

For many years, statistical and ML methods were mostly used in time-series forecasting. This includes their application in pure time-series approaches aimed at forecasting solar irradiance or solar power, based solely on local measurements (i.e., time-series approaches with no exogenous input) for forecast horizons from several minutes to several hours ahead. They also play an important role in enhancing the output of physical and empirical forecast models, namely, NWP and CMV forecasts. The community of statistical modeling and AI refers to these models as statistical models with exogenous input. In contrast, meteorologists commonly use the terms statistical postprocessing or, more specifically, MOS in the context of NWP, which is the terminology adopted here.

The availability of ground truth is essential in training ML models. In irradiance and PV power forecasting, mostly irradiance and PV power measurements are used as ground truth data because they are expected to have a relatively small uncertainty compared to other options (see Chapter 10). Still, this uncertainty should not be neglected. The use of satellite-derived irradiance data is also an option, mostly used in postprocessing NWP forecasts.

With the rapid progress of research in AI (e.g., deep learning (DL) and computer vision) during the last few years, the potential offered by these methods in the field of irradiance and PV power forecasting is also growing. ML and AI algorithms are also increasingly used for image-based predictions using ASI or satellite data, e.g., to compute the optical flow for cloud motion approaches but also to directly predict future cloud conditions or irradiance from raw image data.

Besides categorizing solar forecasts by the different forecasting approaches, from a user's point of view, the following general use cases can be distinguished:

- **Forecasts for individual solar systems** are necessary for owners and operators of PV and CSP plants. In particular, day-ahead forecasts of the solar power generated by large PV plants are now mandatory in many countries (Italy, Germany, Spain, Romania, United States, Japan, China, etc.). Further, they can contribute to improved performance supervision and fault detection and to predictive maintenance and operations and maintenance planning as well as the reduction of power ramps. They are also needed by owners, operators, and providers of PV battery systems and energy management systems as an essential input for the predictive scheduling of storage and energy management.
- **Portfolio forecasts,** consisting of an ensemble of PV systems—and potentially also other generators—are frequently used in direct marketing because of their smaller relative forecast errors compared to single-site forecasts, which is due to spatial smoothing. Such portfolios are also referred to as VPPs. It is expected that they will become increasingly important also for the energy management of multilocation companies aiming toward 100% renewable energy supply or for the management of quarters and districts with a high share of solar power.
- **Forecasts of aggregated regional PV power** are needed by grid operators for the grid management and marketing of PV power, depending on national regulations and feed-in tariffs. This comes with the additional challenge that PV power is not measured at a sufficient resolution for most plants in many countries, and information on PV systems is incomplete. Still, as for portfolios, relative forecast errors of regionally aggregated PV power are much smaller than for single PV plants, depending on the size of the region and the set of PV plants contributing.

This chapter provides an overview of the basic concepts of solar irradiance and PV power forecasting by referring to examples and operational models. More complete reviews of the state of the art can be found elsewhere, including in Kleissl (2013), Yang et al. (2018), Visser et al. (2022), and, for PV applications, Antonanzas et al. (2016). The examples presented here have been investigated in the context of the International Energy Agency (IEA) Solar Heating and Cooling Programme (SHC) Task 36 and Task 46 and the Photovoltaic Power Systems Programme (PVPS) Task 16.We illustrate the forecast performance of the different models described in this chapter by showing basic forecasts scores (e.g., root mean square error (RMSE)). A more detailed discussion on forecast evaluation and uncertainty assessment is given

in Chapter 10 (Section 10.2.3). Evaluation results depend on multiple factors, including the climatological and meteorological conditions at the evaluation site and period (season and year), the forecast horizon, the temporal and spatial resolutions of the forecasts, and the forecasting model used. Therefore, both the forecast scores and the differences between the forecasting models might considerably differ depending on these factors. Here, we provide evaluation results for several examples and model benchmarks to illustrate some general findings.

The model descriptions and evaluations of the different irradiance forecasting approaches given here focus on global horizontal irradiance (GHI) and PV applications. Nevertheless, the forecasting methods also apply to DNI to a large extent. A focus on DNI forecasting can be found in Law et al. (2014) and Schroedter-Homscheidt and Wilbert (2017). Other environmental factors—including ambient temperature, air humidity, wind speed, and wind direction—have a nonnegligible impact on the final power yield of solar plants; however, this handbook focuses on the solar resource aspect, and thus the forecasting of these ancillary variables is not discussed further.

Beyond the description of different forecasting models and their performances for different spatiotemporal scales in this chapter, the "IEA Wind Recommended Practices for the Implementation of Renewable Energy Forecasting Solutions" (Möhrlen et al. 2023) focuses on decision support tools for the energy industry. These tools aim to maximize the benefit of renewable energy forecasts, including PV, in operational decision-making. Therefore, detailed guidelines and recommended practices in selecting and evaluating an appropriate forecasting solution for a given application are given.

The following sections first describe the different approaches for irradiance forecasting: time-series forecasting based on local measurements (Section 9.2), irradiance forecasting based on cloud images (ASI or satellites, Section 9.3), NWP (Section 9.4), and postprocessing and model blending with statistical and ML methods (Section 9.5). Then, Section 9.6 addresses PV power prediction for single sites, portfolios, and regionally aggregated power. Next, Section 9.7 introduces probabilistic forecasting. Finally, Section 9.8 provides a summary and recommendations for irradiance forecasting.

9.2 Time-Series Forecasting Based on Measurements

The goal of pure time-series approaches is to derive solar irradiance or power forecasts based solely on local measurements, i.e., without involving any physical modeling. They require real-time access to measurements and are suitable for forecast horizons from several minutes to several hours ahead. Section 9.2.1 introduces persistence as a baseline approach, which constitutes the simplest possible model, using only local irradiance or PV power measurements. Section 9.2.2 provides a short introduction to selected ML models, which are frequently used for applications in time-series forecasting of solar irradiance. Finally, Section 9.2.3 addresses forecasting methods based on pure time-series models and includes a discussion of their advantages and limitations.

9.2.1 Persistence

Persistence is a trivial model that simply assumes that the current situation does not change during a forecast run. Typically, persistence is based on recent on-site measurements. In solar radiation forecasting, persistence is the simplest and most widely used reference model. It is

commonly used to evaluate forecast skill (see Chapter 10, Section 10.5.1.2). It is also used in operational forecasting, e.g., as an input to hybrid forecasting approaches, requiring online measurements as a basis. Alternatively, satellite- or ASI-derived irradiance values can be used as a starting point for persistence, although in that case the forecast uncertainty for very short-term forecast horizons is higher than that for measurement-based persistence.

Several definitions of the persistence of solar irradiance exist, including *simple* persistence; *scaled* persistence, which accounts for solar geometry changes; and more advanced concepts, such as *smart* persistence. The most widely used definitions are presented next.

For day-ahead forecasting, the simplest approach is to assume that irradiance, I (GHI or DNI), persists during a period of 24 hours:

$$I_{\mathrm{per,24h}}(t) = I_{\mathrm{meas}}(t - 24\mathrm{h}). \tag{9-1}$$

A more elaborate option for GHI is to separate the clear-sky and cloudy contributions to solar radiation and to assume that only the cloud conditions persist during a forecast run; this defines the scaled persistence. Clear-sky irradiance is strongly influenced by the deterministic solar geometry and can be described with reasonable accuracy using a clear-sky radiation model (see Chapter 7, Section 7.2.1). In such a modeling approach, the persisting magnitude is the clear-sky index, K_c, calculated as the ratio between the measured GHI and a clear-sky GHI estimate, $\mathrm{GHI_{clear}}$. For forecast horizons of several hours (Δt) ahead, the scaled persistence, $\mathrm{GHI_{per,}}\ K_c$, for time t is then defined as:

$$\mathrm{GHI}_{\mathrm{per\ Kc,\Delta t}}(t) = \mathrm{GHI}_{\mathrm{clear}}(t)\ K_c(t - \Delta t). \tag{9-2}$$

For DNI, a similar approach can be used, now based on the beam clear-sky index or the Linke turbidity factor (Kuhn et al. 2017).

To make the most of persistence in a solar context, the so-called "smart persistence," $\mathrm{GHI_{per\ smart}}$, was proposed in the context of the IEA SHC Task 46. It consists of increasing the integration time that defines the current conditions commensurately to the forecast time horizon. Δt:

$$GHI_{per\ smart,\Delta t}(\mathrm{t}) = GHI_{clear}(\mathrm{t})\frac{1}{\Delta t}\int_{t-2*\Delta t}^{t-\Delta t} K_c(t')dt' \tag{9-3}$$

Or, for measurements available in a discrete time interval, Δt_{meas} :

$$GHI_{per\ smart,\Delta t}(\mathrm{t}) = GHI_{clear}(\mathrm{t})\frac{1}{N}\sum_{i=1}^{N} K_c\left(t - \Delta t\left(1 + \frac{i}{N}\right)\right), \tag{9-4}$$

with $N = \frac{\Delta t}{\Delta t_{meas}}$.

Alternatively, if long-term irradiance measurements are available, combinations of climatology and persistence can be used, as recommended by Yang et al. (2020), as an advanced reference model for forecast evaluation.

Despite its simplicity, persistence can outperform forecasts based on empirical and physical model forecasts for very short-term forecast horizons (see Section 9.3.1.3, Section 9.3.2.2, and Section 9.5.4). A further improvement for these forecast horizons can be achieved by applying statistical and ML models using online measurements as input, as described in the following sections.

9.2.2 Examples of Machine Learning Models Applied for Solar Forecasting

The use of state-of-the-art ML models is popular in both irradiance and PV power forecasting. This section describes several ML approaches that are used frequently, as discussed by Winter et al. (2019), including ANNs, extreme learning machines, gradient-boosted regression trees, and random forests. Further, auto-machine learning (AutoML) has become a hot topic in this field because of its high accuracy, deployment simplicity, and time efficiency. The ML approaches introduced here are used not only for pure time-series forecasting but also for postprocessing and model blending (Section 9.5) and partly also in image prediction (Section 9.3).

9.2.2.1 Artificial Neural Networks

ANNs constitute one of the most versatile ML methods and are known for their use in complex tasks, such as image or speech recognition (LeCun et al. 1989; Sak, Senior, and Beaufays 2014).

As described in Bishop (1995), an ANN consists of a fixed number of nodes, called units, that can take on numerical values and are arranged in several layers. The input layer contains one unit for each feature of the dataset, whereas the output layer, in the case of a single regression problem, is only one unit. The layers between the input and output layers are referred to as hidden layers. The key task is to establish a connection between the nodes by assigning to each unit in one layer the weighted sum of the previous layer's units, and to then apply a nonlinear activation function. In the case of a regression problem, a linear activation function is applied to the weighted sum of the output unit.

By training an ANN on a given set of input and output data, all its weights are adjusted to minimize an error function, typically the mean square error (MSE). This is usually done by back-propagation—an iterative process for calculating the gradient of the error function with respect to each weight (Rumelhart and McClelland 1986). At each step, the weights get updated by using a gradient descent optimization algorithm. An alternative option is the method of adaptive moment estimation, or "Adam," as described by Kingma and Ba (2014). Instead of calculating the gradient of the error function with respect to the full dataset, the weights can be updated at each step only with respect to a subset of the dataset (see Bottou (1998) and Ruder (2017)). The weights can be initialized using a common heuristic, as described by Glorot and Bengio (2010).

To make an ANN able to learn nonlinear relationships between input and output, a nonlinear activation function must be chosen. For example, the leaky rectified linear unit activation function can be used (Maas 2013).

9.2.2.2 Extreme Learning Machines

An extreme learning machine, as proposed by Huang, Zhu, and Siew (2006), is an ANN with a single hidden layer between the input and output layers. Its learning method does not rely on gradient descent. Instead, the weights between the input and hidden layers are chosen randomly. In this way, only the weights between the hidden and output layers need to be determined.

Because this is only a linear regression problem, an analytic solution exists, which can be calculated directly without an iterative optimization algorithm. Hence, training the model is considerably faster while maintaining its good performance.

9.2.2.3 Gradient-Boosted Regression Trees

Gradient-boosted regression trees are an ensemble technique using multiple classification and regression trees (CART), as introduced by Breiman et al. (2017). The CART algorithm creates binary decision trees, which means that at each new node, the data are split into two parts according to a threshold value. Starting with a root node, which, in general, contains all training data, the tree grows until some stop condition is reached. The last nodes form the tree's leaves. Each splitting leads to either another node or a leaf. The leaf contains the class to be predicted. In the case of regression, a leaf returns the mean value of the training samples it contains.

The principle of boosting is described by Friedman (2001). Starting with a single CART tree that is fit to minimize the MSE on the training data, the following trees are trained consecutively so that each new tree predicts the residual error. This residual error is proportional to the gradient of the MSE. By scaling the new tree's prediction with a step size between 0 and 1 and by adding it to the current ensemble, every new tree aims to further reduce the MSE of the ensemble's prediction.

9.2.2.4 Random Forest

A random forest is another technique based on ensembles of CARTs, as presented by Breiman (2001). The ensemble's prediction is the average over all single-tree predictions. Each tree is trained on a bootstrap dataset generated by randomly drawn samples with replacement from the original dataset (Efron 1979). Further, for each node split, only a random subset of features is considered. By randomly omitting data, the resulting trees become less correlated. This reduced correlation of single trees has been observed to reduce the model error.

9.2.2.5 Auto-Machine Learning

AutoML techniques have achieved considerable success in solving and improving any kind of problems in an ever-increasing number of disciplines, further reducing the manual workload. They are based on sequential stages of data processing, from preprocessing the data to hyperparameter tuning and model ensembles, where the use of several base ML models (ANN, random forest, or gradient boosting machines, among others) provide ensemble predictions. A successful ensemble has proven to be the multilayer stack, where the first layer has multiple base models, and whose outputs are concatenated and then fed into the next layer, which itself consists of multiple stacker models. These stackers then act as base models to an additional layer. Note that to take advantage of all available data and also to mitigate overfitting, k-fold ensemble bagging of all models at all layers of the stack is typically used.

9.2.3 Time-Series Forecasting With Statistical and Machine Learning Approaches

Intrahour or hours-ahead solar irradiance and PV power forecasting with pure time-series models use recent measurements of irradiance or PV power as a basic input, possibly complemented by measurements of other variables. Examples are the application of a coupled autoregressive and dynamic system model for forecasting solar radiation on an hourly timescale, as described by Huang et al. (2013), the comparison of ANN and classical time-series models by Reikard (2009),

and the short-term PV power prediction approach of Bacher, Madsen, and Nielsen (2009). Through their review of ML methods, Voyant et al. (2017) concluded that although ANN and autoregression-style methods still dominate statistical forecasting, other methods (e.g., support vector regression, regression tree, random forest, or gradient boosting) are increasingly being used. Although the ranking of such methods is complicated by many factors, it generally holds that a multi-model approach results in an improvement in forecasting performance (Zemouri, Bouzgou, and Gueymard 2019).

For any statistical or ML model, the selection and availability of appropriate input variables, as well as the optimized preprocessing of the data, are of critical importance for good forecast performance. Additionally, the choice of the model and of its configuration (e.g., the ANN architecture or the selection of hyperparameters in ML models) is essential. Finally, the setup of the training sample (e.g., the number of days and sites used for the training) has a noteworthy influence on forecast accuracy. In recent years, automated AutoML tools have been increasingly used for that (see Section 9.2.2.5). They are applied to optimize the ML pipeline, including the identification of the most suitable model and hyperparameters, which can be done using genetic programming.

The advantages and limits of purely statistical approaches are discussed next. High-quality measurements of the actual surface solar irradiance or PV power constitute the best possible starting point for any forecast. In comparison, the assessment of the initial irradiance conditions (i.e., the irradiance analysis) with an empirical or physical forecasting model shows considerably higher uncertainties. Any physics-based forecasting model has an inherent uncertainty, regardless of the forecast horizon, that is caused by limits in the spatial and temporal resolutions, uncertainty in input parameters, and simplifying assumptions within the model. Time-series models exploit the autocorrelation in the time series of solar irradiance, cloud cover and, possibly, other explanatory variables. For very short-term forecast horizons, forecasts based on accurate on-site measurements and statistical methods are affected by forecast errors that are typically smaller than the errors from either NWP analysis or irradiance forecasts derived from satellite or ASI images that correspond to the starting points of the respective forecast runs.

Given the inherent chaotic nature of weather phenomena, any existing autocorrelation decreases as the time lag between time-series instances increases. Hence, the performance of these models is (1) strongly determined by the underlying autocorrelation of each particular weather condition and (2) decreases as the forecast lead time increases. For longer forecast horizons, wide-area observations of clouds or irradiance (e.g., those from satellites images) or NWP model forecasts are necessary to meet forecast skill requirements.

Therefore, pure time-series approaches are typically applied to forecast horizons ranging from several minutes to a few hours ahead. Evidently, their performance compared to other methods strongly depends on the prevailing climate and weather conditions (e.g., the stability of the sky situation), the spatiotemporal resolution of the forecasts, and the models to which they are compared.

Finally, the uncertainty of the irradiance measurements used for both input and as ground truth for model training has a strong impact on the performance of time-series models. This uncertainty strongly depends on the instrument class as well as maintenance of the stations (see

Chapter 3, Chapter 4, and Chapter 10). Any statistical or ML forecasting model will adapt to the irradiance—or PV power—measurements it is trained to. This includes any possible systematic deviations in the measurements, e.g., calibration errors or soiling in case the station under scrutiny is not properly maintained. Therefore, high-quality measurements are crucial for statistical and ML models.

9.2.4 Forecasting With Machine Learning Methods and Data From Sensor Networks

As discussed, irradiance or PV power measurements from a single site can be used as the basis for time-series forecasts, but there are also forecasting methods that seek to explore spatiotemporal correlations from *in situ* solar measurements distributed in space. A review of such methods appears in Benavides Cesar et al. (2022). Bosch and Kleissl (2013) first showed that a pyranometer network was effective in detecting cloud advection. In a follow-up study, Lonij et al. (2013) did one of the earliest forecasting demonstrations, using a network of 80 distributed PV systems (considered as sensors) over an area of 2,500 km^2.Other studies sought to understand some of the nuances involved. In particular, Amaro e Silva and Brito (2017) showed that such an approach surpasses smart persistence from the seconds- to days-ahead timescale as long as the increase in forecasting horizon is met by a larger spatial coverage of the data and a coarser time resolution. Later, Amaro E Silva and Brito (2019) also found that the tilt and orientation of the sensor ensemble impact the added value of such an approach.

The main advantage of this kind of approach is its flexibility regarding the time and spatial scales. Two other benefits are also worth mentioning: The need for imagery and image processing (from satellite or ASI) is eliminated, and useful new data can be generated for use in regional studies regarding variability and upscaling approaches. Conversely, a disadvantage of this approach in solar resource applications is its dependence on potentially expensive and maintenance-intensive sensors. There are also approaches that explore PV generation data from a fleet of PV systems, but the challenge here is that such data sources mostly belong to private companies and are thus not publicly available.

9.3 Irradiance Forecasting Based on Cloud Images

Forecasting the evolution of clouds is essential in irradiance forecasting. Therefore, cloud images providing spatially extended information on clouds are suitable sources to extend forecast horizons compared to pure time-series models based on local measurements only. Two types of cloud images are used in irradiance forecasting: (1) cloud information for small areas with very high spatiotemporal resolution, obtained from ground-based ASIs, and used for very short-term forecast horizons of typically 10–20 minutes ahead; and (2) cloud information derived from geostationary satellite data, covering large areas at high temporal resolution, and suitable to forecast clouds and solar irradiance up to several hours ahead. Though covering different forecast horizons and areas, the basic steps in irradiance forecasting based on ASI and satellite imagery are similar.

At timescales from a few minutes to a few hours, horizontal advection has a strong impact on the temporal evolution of cloud patterns, with the shape of clouds often remaining quite stable. Here, the spatial scale is also extremely important because small-scale cloud structures change faster than larger structures. In these situations, techniques for detecting clouds and their motion

trajectories, referred to as CMV techniques, are used to provide valuable information for irradiance forecasting. Obviously, the performance of these forecasting methods degrades as the importance of local processes of cloud formation and dissipation, such as strong thermally driven convection, increases.

CMV-based techniques consist of the following basic steps:

- Images with cloud information are derived from ASI or satellite data.
- Assuming stable cloud structures and optical properties, the CMVs are determined by identifying matching cloud structures in consecutive cloud images.
- To predict future cloud conditions, the CMVs are applied to the latest available cloud image assuming cloud speed persistence.
- Solar irradiance forecasts are calculated from the predicted cloud structures.

For these steps, physical, empirical, or ML models can be employed, as described in more detail next, for irradiance forecasting based on either ASI (Section 9.3.1) or satellite imagery (Section 9.3.2). In addition to the CMV-based techniques, approaches to directly predict irradiance from cloud images with ML models have been developed during the last few years.

9.3.1 Forecasting Using Ground-Based All-Sky Imagers

Solar irradiance forecasts at subhourly scales with high spatiotemporal resolution can be derived from ground-based ASIs. Such imagers are installed horizontally and sense the whole sky above them (Figure 9-4, upper row). These highly resolved sky images contain information on the cloud cover that can be exploited for forecasting. At times, ASIs are also called whole-sky imagers, sky imagers, or sky cameras. (Note that a sky *imager* is not strictly identical to a sky *camera*; in all publications from IEA PVPS Task 16 participants and within this handbook, the term ASI is normally used.)

ASIs can capture sudden changes in irradiance, which are often referred to as *ramps*, at temporal scales from seconds to minutes (Figure 9-2). Cloud fields sensed from ASIs, or from an assembly of several ASIs, can be resolved with high detail, allowing the partial cloud cover over large PV installations to be modeled and forecasted. The maximum predictable horizon strongly depends on cloud conditions, and it is constrained by the cloud speed and the field of view of the ASIs. This forecast horizon typically ranges from 10–20 minutes, but it can reach more than 30 minutes in favorable cases. Considering the high resolution combined with a very fast update rate, ASI forecasts can play a valuable role in the economic operation of PV plants. Another advantage is that they can significantly reduce the overall societal costs caused by active intermittence mitigation (Wan et al. 2015).

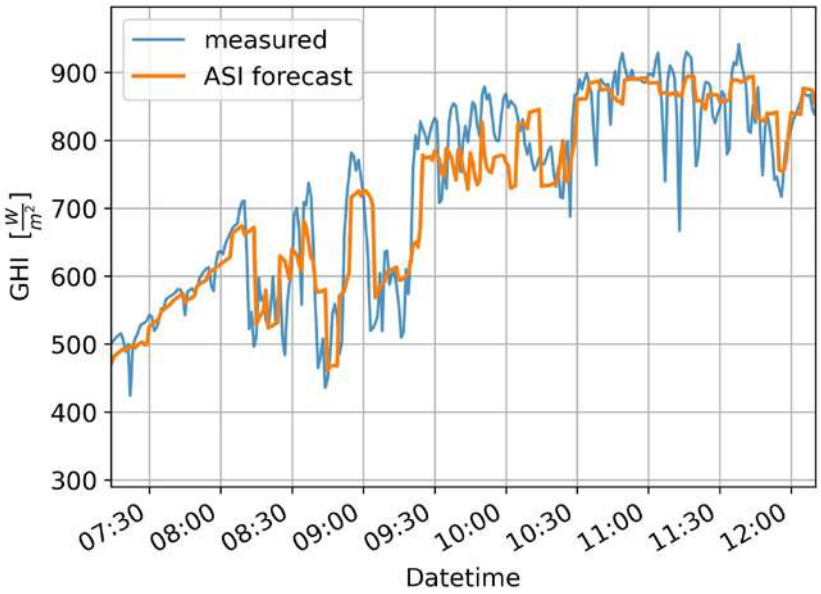

Figure 9-2. Example of minute-resolution GHI forecasts using an ASI

The forecasts with lead times up to 5 minutes ahead are updated every 5 minutes. May 9, 2021. Location: Freiburg in Germany.

Image by Fraunhofer ISE

9.3.1.1 All-Sky Imager Hardware

Currently, there is no defined standard for sky imaging hardware, camera calibration, or image processing techniques. Systems in use include commercially available, low-cost webcams or surveillance cameras, and systems developed specifically for sky imaging ,e.g., (Urquhart et al. 2015). Most systems use digital red-green-blue (RGB) cameras with fish-eye lenses and therefore sense visible radiation only, although some systems work with infrared cameras, which are more expensive. In particular, older RGB systems and some infrared cameras use a downward-looking camera that takes photos of an image of the sky that appears on a roughly spherical upward-looking mirror. This is where the term *imager* comes from. This concept—unlike the smaller lens or dome of fish-eye cameras—has the disadvantage that the *whole* mirror must be cleaned. Moreover, some older systems use sun-tracked "shadowbands" to prevent direct sunlight from reaching the camera. This can reduce lens flare-induced saturated areas in the photos, but the shadowband also covers a noticeable part of the image. Because the required tracking of the shadowband entails higher costs and can lead to system failures, shaded devices have become uncommon in recent years. In addition to the sky imager(s), an ASI forecasting system typically includes a radiometer at the sky imager location. The irradiance measurements are used as additional input to infer irradiance from ASI images and/or as ground truth for model training.

9.3.1.2 All-Sky Imager-Based Forecasting

The classical operation of ASI-based forecasts typically involves a physics-oriented chain of processing steps, e.g., (Marquez and Coimbra 2013). These physics-based or empirical solar forecasting approaches are also referred to as *indirect forecasting* (Lin, Zhang, and Wang 2023).

Nevertheless, because of the great success of ML in computer vision, recent developments show a clear trend to so-called *direct forecasting* approaches (Lin, Zhang, and Wang 2023). These direct approaches use trained models that derive the forecasts directly from the sky images , e.g.,(Chu et al. 2015). The advantages of ML in computer vision have also led to the application of ML methods as a way to improve the individual processing steps of indirect methods (e.g., cloud detection (Hasenbalg et al. 2020)). Figure 9-3 illustrates the general scheme of ASI-based forecast methods. The corresponding processing steps, along with some relevant image preprocessing procedures, are outlined in the following.

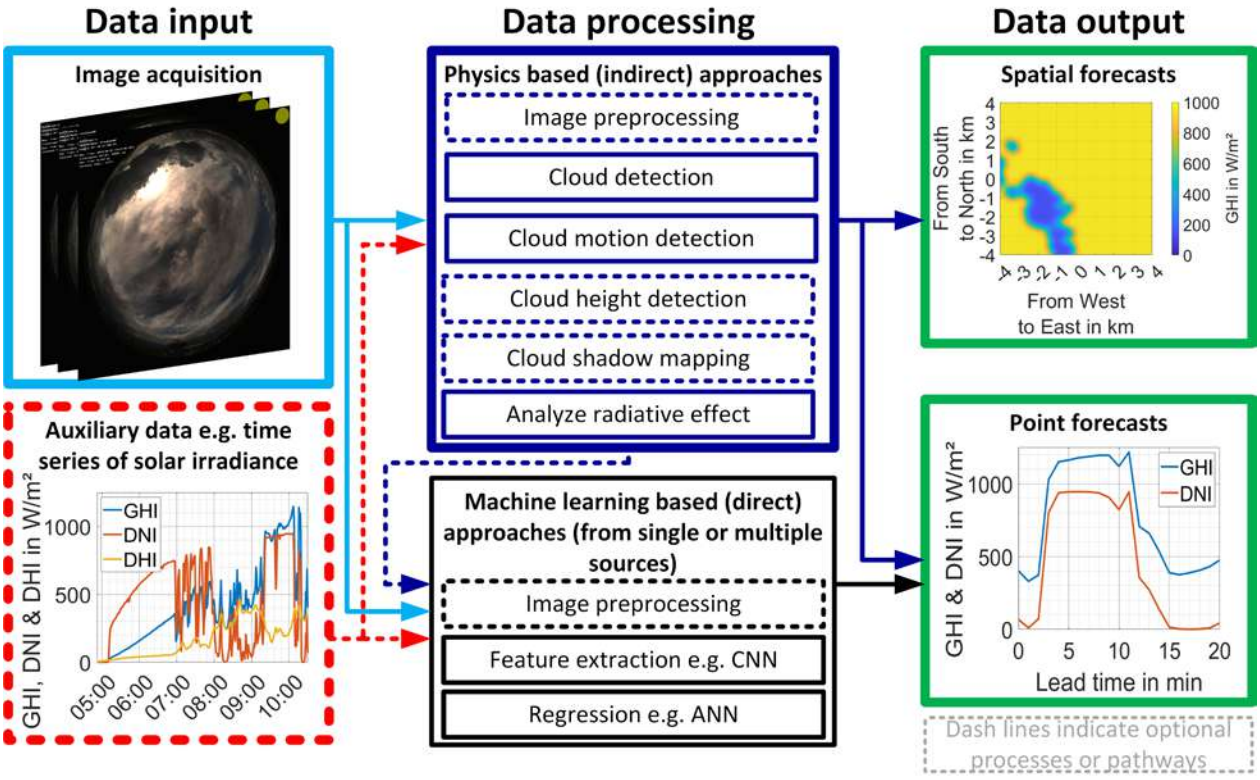

Figure 9-3. General scheme of ASI-based forecasting methods

Image by DLR Institute of Solar Research

Image preprocessing: Both direct and indirect approaches may apply different image preprocessing steps. Following are some frequently used preprocessing steps:

- Fish-eye distortion: Using fish-eye lenses leads to strong radial distortion effects that increase with greater distance from the zenith. Different processing steps might require/benefit from orthogonal projections (e.g., cloud motion detection (Marquez and Coimbra 2013) or cloud height detection (Nouri et al. 2019)). A commonly used approach for fish-eye distortion correction is described in Scaramuzza, Martinelli, and Siegwart (2006).
- Partial masking: Masking disturbing objects like shadowbands (Chow et al. 2011)
- Cropping and downscaling: An ML model's performance might increase if the training datasets are based on higher-resolution imagery; however, this also comes with increased

computational costs (Sun, Szűcs, and Brandt 2018). Cropping and downscaling sky images is therefore common practice.

- High-dynamic-range techniques: The illumination range of sky images can be very large because of the scattering effects of solar radiation (especially under complex cloudiness patterns or in the circumsolar region. High-dynamic-range techniques tackle this challenge by combining multiple images with different exposure times into a single image with dynamic illumination range (Chauvin et al. 2017).

Indirect (empirical) approaches are typically divided into the following main processing steps: (1) cloud detection, (2) cloud motion detection, (3) analysis of the radiative effect, (4) cloud height detection, and (5) cloud shadow mapping. For point forecasts at the sky imager location, information about cloud height is not required because the cloud movement can be parameterized in terms of "pixels per second"; they can be derived using steps (1), (2), and (3) only. The additional processing steps (4) and (5) are required when spatial forecasts are desired (Urquhart et al. 2012).

Cloud detection/segmentation: Three main cloud segmentation techniques are described in the literature and are summarized as follows:

1. Thresholding-based approaches exploit the different spectrally resolved scattering properties of clouds and cloudless sky by analyzing the red-to-blue ratio. Long et al. (2006) proposed a simple empirically defined fixed red-to-blue threshold. A more sophisticated adaptive thresholding, based on a database of image properties during clear-sky conditions (referred to as a "clear-sky library"), was introduced by Chow et al. (2011). An automated way to construct such a database was devised by Shaffery et al. (2020).
2. Superpixel cloud segmentation approaches evaluate multiple pixel features (brightness, color, and texture) and then group pixels into coarse regions (Shi et al. 2017). These compact and perceptually coherent regions are commonly referred to as superpixels.
3. Various ML approaches can be applied to the cloud segmentation tasks. Lately, DL-based approaches using convolutional neural networks (CNNs) have seemed to prevail (Xie et al. 2020). Various benchmarks have shown a clear advantage of DL approaches compared to classical thresholding-based approaches, e.g., ((Hasenbalg et al. 2020; Xie et al. 2020). Some samples of such a comparison are illustrated in Figure 9-4. Critiques mention the requirement of large manually segmented ground truth databases for supervised learning (Lin, Zhang, and Wang 2023); however, this demand can be significantly reduced through the use of transfer learning and pretrained weights for initialization (Fabel et al. 2022). That process involves initially employing a self-supervised approach to establish weights using an extensive unlabeled dataset, followed by a subsequent supervised approach that relies on a small labeled dataset. Ye et al. (2019) combined superpixels with a DL approach (support vector machines) for a semantic segmentation based on nine distinct cloud categories.

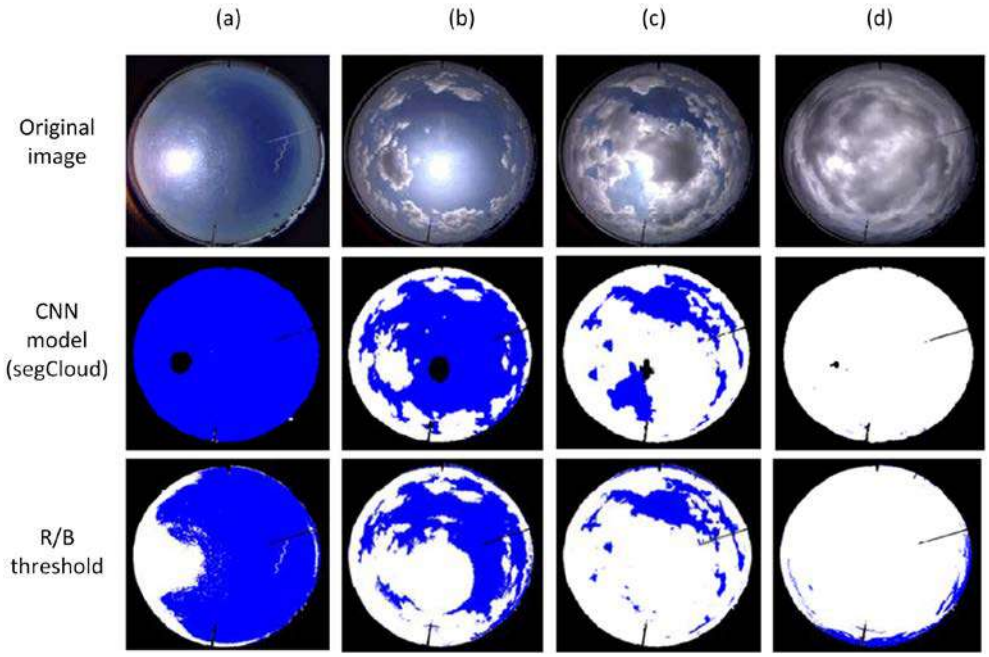

Figure 9-4. Examples of four segmentation results (columns a to d)

Clouds, sky, and the sun are colored white, blue, and black, respectively. The second row illustrates results from a CNN model, and the third row depicts results based on a red-to-blue ratio threshold.

Image by Xie et al. (2020)

Cloud motion detection: Global or multiple CMVs, up to a dense vector field having a single CMV for each pixel, can be derived from sky images. Generally, this is done via similarity maximization, optical flow, or feature tracking. similarity maximization often involves block-matching approaches (see Figure 9-5). Different subparts of consecutive image series are compared for their similarity via cross-correlation (Chauvin et al. 2016). The optical flow approaches assume that cloud motion only leads to a pixel shift at constant pixel brightness (Paragios, Chen, and Faugeras 2006), providing a way to derive dense vector fields. Overall, optical flow typically outperforms similarity maximization approaches; however, this advantage comes at the expense of increased computational costs (Peng et al. 2016). Finally, feature tracking approaches try to find and track unique feature points, such as cloud edges or corners. This method offers notable computational efficiency by prioritizing a few prominent features but is limited by the granularity of the CMVs (Su et al. 2015).

Cloud height detection: Stereoscopic cloud height measurement approaches based on multiple ASIs are described in, e.g., Nguyen and Kleissl (2014). Some of these methods are effective at deriving different cloud heights for the individual clouds seen in the sky image (Peng et al. 2015). That process is exemplified in Figure 9-5. Alternative approaches include the integration of supplementary instruments. In particular, the most accurate determination of the cloud-base height directly above the ASI instrument is currently obtained with ceilometers (Arbizu-Barrena et al. 2015); however, ceilometers are costly and limited to cloud measurements directly above the sensor, and therefore they are not ideally suited for multilayer cloud conditions. Combinations of ASIs with radiometers have also been described, but these approaches do not achieve the performance of a stereoscopic assembly with multiple ASIs (Kuhn et al. 2018).

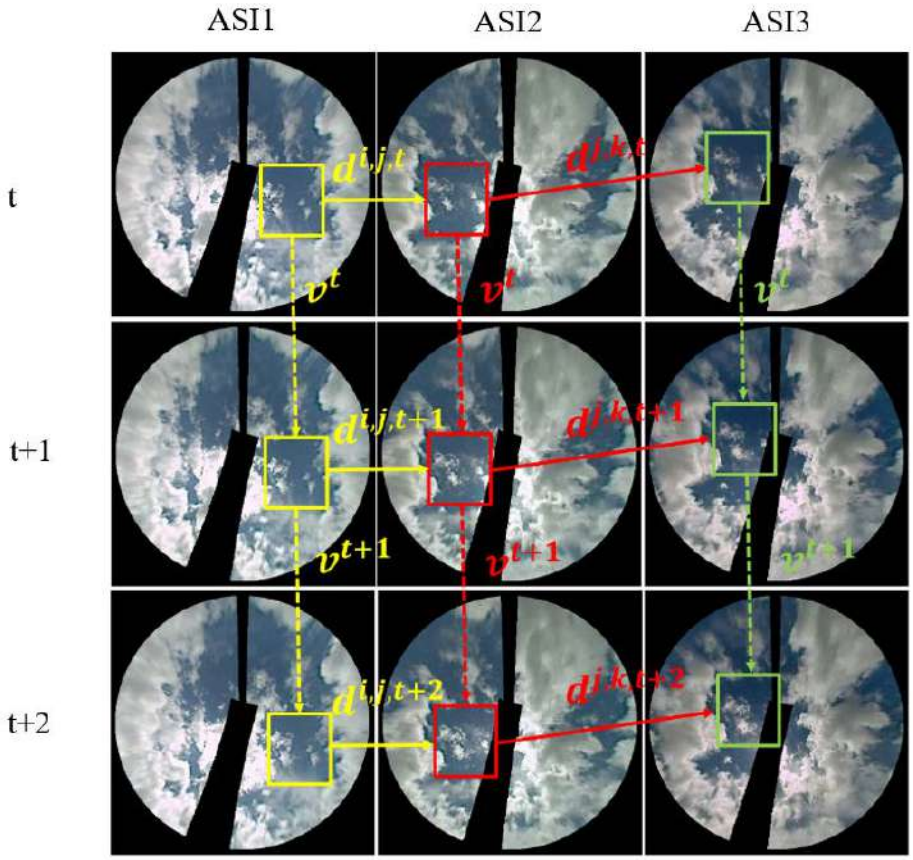

Figure 9-5. Matching cloud blocks for cloud motion and height detection from three consecutive images and three distinct ASIs

In the ASI1, ASI2, and ASI3 images, the cloud block of interest is denoted by the yellow, red, and green boxes, respectively. The detected movement of the cloud block between two consecutive frames is visually represented by a dotted arrow labeled "v." Additionally, the displacement vector between a pair of ASIs captured at the same time stamp is indicated by a solid arrow and labeled "d."

Image by Peng et al. (2015)

Cloud shadow mapping and irradiance forecasting: Cloud shadow maps at the surface are produced by projecting the forecasted cloud scenes with their assigned height using information about the position of the sun and a digital elevation model. Local irradiance or PV power measurements can be used to estimate the cloud effects on irradiance or PV power for either point or spatial forecasts. Urquhart et al. (2013) analyzed the frequency distributions of PV power normalized to clear-sky conditions to determine a clear and a cloudy mode and to assign them to shaded and unshaded cells, respectively. Schmidt et al. (2016) and Dittmann, Holland, and Lorenz (2021) used the clear-sky index derived from recent pyranometer measurements to determine the forecasted all-sky GHI. Similarly, for DNI forecasting, Blanc et al. (2017) used the beam clear-sky index determined from the last 30 minutes of pyrheliometer measurements to derive the cloud transmittance. Ghonima et al. (2012) proposed a method to differentiate thin and thick clouds for various atmospheric conditions using a clear-sky library.

Using multiple ASIs to create networks of ASIs is useful to increase the spatial coverage, the forecast horizon, and the accuracy of observations by providing a more accurate 3D reconstruction of the cloud field (Mejia et al. 2018). Moreover, the combination of several ASI-

derived intermediate results (e.g., segmentation and cloud height) can be used to improve the nowcasts (Blum et al. 2022). Exemplary results of such a network are presented in Figure 9-6.

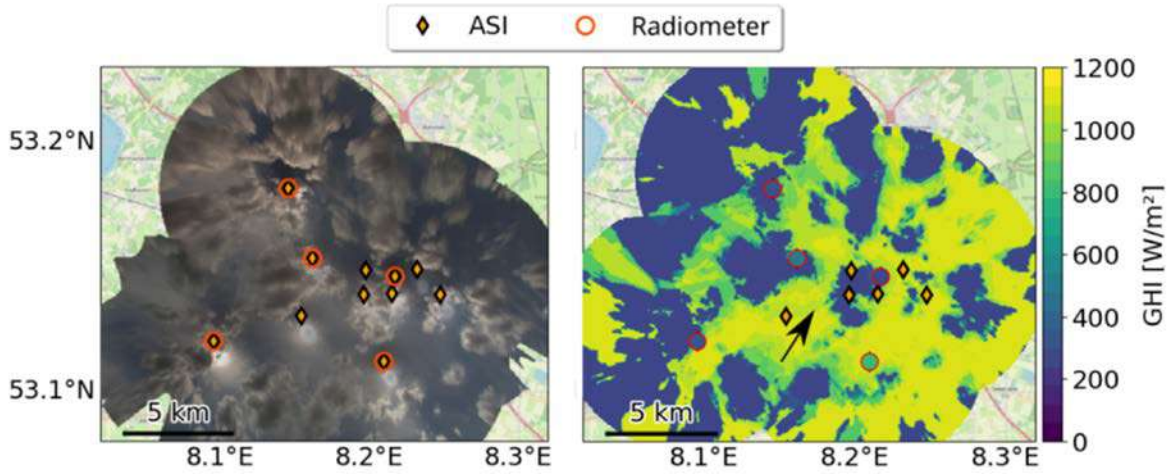

Figure 9-6. Exemplary results obtained by an ASI network

(Left) Stitched raw, undistorted, and georeferenced ASI images and (right) the GHI map derived by the ASI network.

Image by Blum et al. (2022)

Direct approaches use a supervised ML framework, which facilitates the creation of a mapping function that directly associates sky images with their corresponding solar irradiance values. The majority of approaches described in the literature rely on CNNs (Lin et al. 2023). The CNNs automatize the extraction of essential features from the input data. Once the features are extracted, they are flattened to a 1D vector and forwarded to an additional neural network, which proceeds with the regression procedure using solar irradiance as ground truth (Sun, Roth, and Black 2018). Sequential data are a valuable source of information regarding cloud motion, but its processing using traditional 2D CNN kernels is not direct. To overcome this challenge, specialized 3D CNN methods have been developed that are specifically designed to extract cloud features from a sequence of consecutive stacked sky images (Zhao et al. 2019). Another effective strategy involves the fusion of CNNs with recurrent neural networks (RNNs), known for their aptitude in handling subsequent data. Features extracted from the convolutional blocks are flattened and seamlessly passed to the RNN, employing a long short-term memory (LSTM) block for improved memory retention and sequential processing (Zhang et al. 2018). More recently, a variant was introduced called convolutional LSTM (convLSTM) , which directly accepts a sequence of sky images as input, removing the need for intermediate flattening and streamlining the processing of sequential data (Kong et al. 2020). Incorporating auxiliary data, such as solar irradiance time series, into a direct approach can enhance its predictive capabilities. One potential method involves the training of two parallel neural networks, one dedicated to processing images and the other designed for handling time-series data (see Figure 9-7). The extracted features from both networks can then be concatenated and fed into a final neural network, which performs the solar irradiance prediction (Paletta, Arbod, and Lasenby 2021). Recently, attention-based transformer architectures have been employed based on this strategy, and they appear to outperform CNNs and RNNs (Fabel et al. 2023).

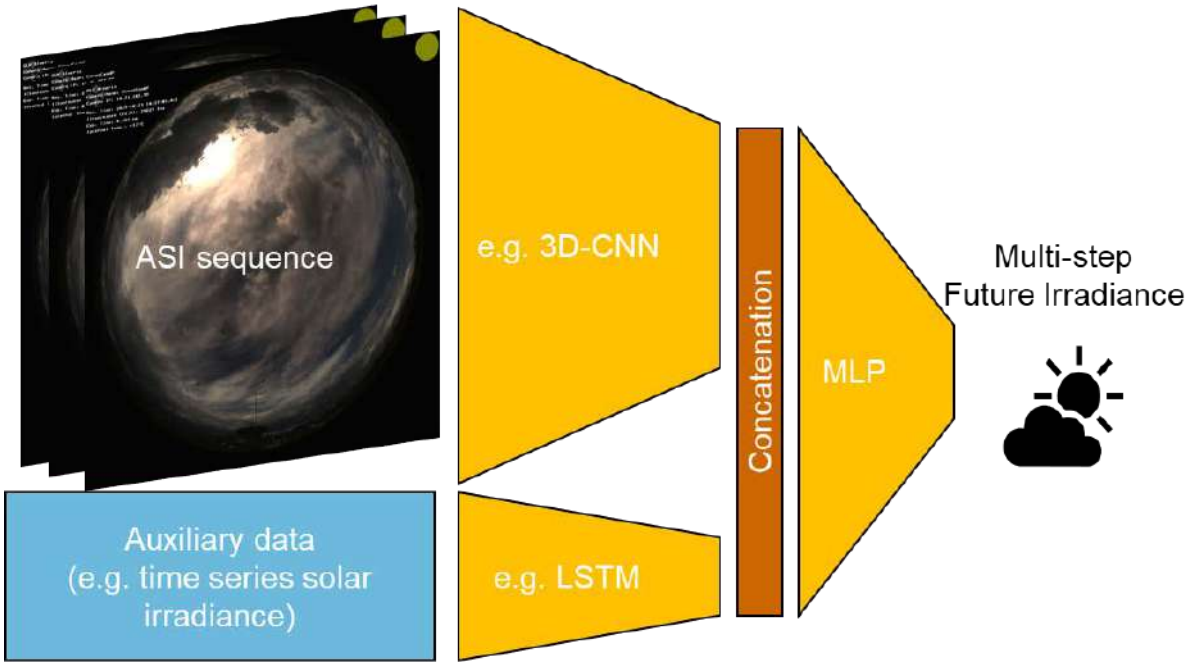

Figure 9-7. Potential general architecture of a multimodal model incorporating ASI sequences and auxiliary data

Image by DLR Institute of Solar Research

There is a noticeable trend toward adopting direct approaches in solar forecasting (Lin et al. 2023). These direct approaches consistently demonstrate impressive performance when assessed using well-established metrics such as RMSE or forecast skill; however, note that these direct approaches often exhibit characteristics similar to a "very smart persistence model," which mitigates the overall deviations. Although this characteristic often leads to lower average errors, it can potentially compromise the ability to accurately represent current conditions or accurately capture ramps (Paletta, Arbod, and Lasenby 2021). As a result, recent approaches have embraced the integration of preliminary results derived from indirect methods, such as CMVs, within the framework of direct approaches (Kamadinata, Ken, and Suwa 2019).

Regardless of direct or indirect methods, almost all ASI-based nowcasting systems provide deterministic forecasts. Recently some new approaches have been designed to provide probabilistic forecasts as well (Nouri et al. 2023; Paletta, Arbod, and Lasenby 2023).

In addition to irradiance nowcasting, ASIs have many other applications that are relevant to meteorology and solar energy. Deriving GHI and/or DNI from sky images is discussed by Dev et al. (2019), Sánchez-Segura et al. (2021), Schmidt et al. (2016), Chauvin et al. (2018), Kurtz and Kleissl (2017), and Gauchet et al. (2012). It is also possible to estimate the sky radiance distribution (Chauvin et al. 2015) and the aerosol optical depth (AOD) (Olmo et al. 2008; Kazantzidis et al. 2017).

9.3.1.3 Performance of All-Sky Imager Irradiance Forecasting: Results of a Benchmarking Exercise

To illustrate the performance of ASI-based forecasts, examples of results from the benchmarking exercise performed within the IEA PVPS Task 16 framework (Logothetis et al. 2022) are

presented next. That exercise aimed at nowcasting GHI with five ASI systems located at the Plataforma Solar de Almeria in southern Spain (Almeria). The experiment lasted 28 days during September–November 2019 and encompassed a large variety of cloud conditions. Six different cloudiness classes, including clear sky, were identified manually from the images. Various forecast lead times, ranging from 1–20 min, were considered. For each cloudiness class, the variation of the RMSE (defined in Chapter 10, Equation 10-4) over increasing forecast lead times (Figure 9-8) indicates that:

- Forecast and persistence errors, as well as differences between forecasts models, strongly depend on the cloud conditions and are smallest for cloud-free conditions.
- Forecast and persistence errors mostly increase with forecast lead time.
- ASI forecasts can outperform persistence in terms of RMSE for all cloud conditions and all lead times (ASI1 and ASI2). All investigated ASI-based models outperform persistence from several minutes onward for most cloud classes.
- Under clear-sky conditions, ASI forecasts show a similar RMSE as persistence, which is expected.

These findings agree with those from the other studies mentioned in Section 9.3.2.1. A comparison of ASI-based methods to satellite-based forecasting is discussed in Section 9.5.3.

In summary, ASI-based forecasting systems have gained popularity in the energy meteorology community because of their exceptional ability to provide very high-resolution intrahour forecasts. In recent years, ML-based direct approaches have gained ground; however, pure ML approaches have been found to have certain weaknesses (e.g., strong smoothing of predictions to minimize deviations). As a result, some of the most recent works advocate for integrating indirect physics-based approaches with direct data-driven approaches.

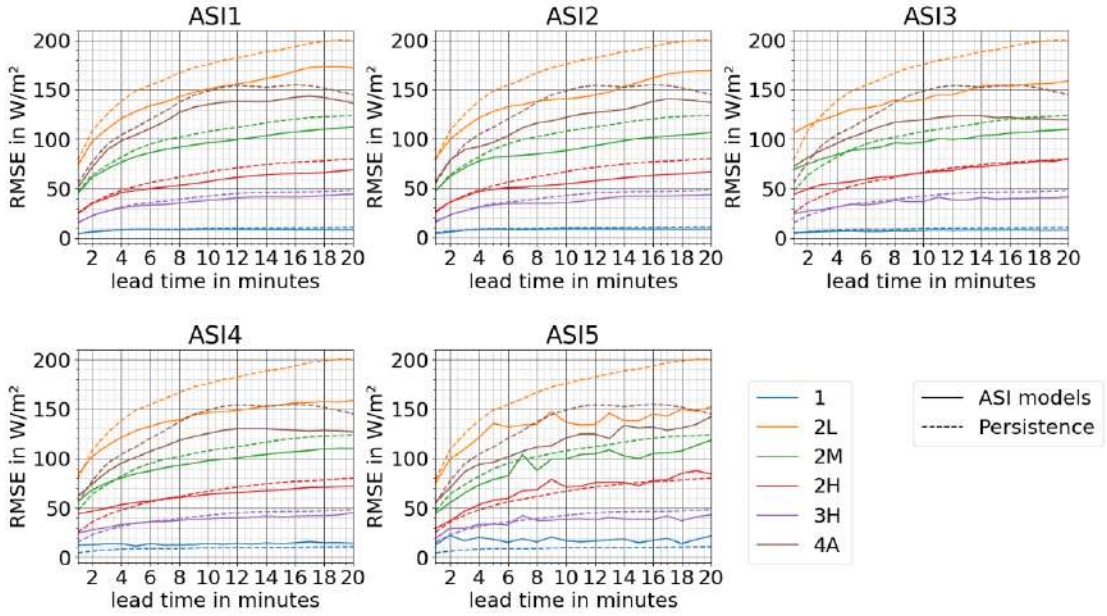

Figure 9-8. RMSE of five different ASI forecasting approaches (model, solid lines) compared to scaled persistence (PERS, dashed lines) over forecast lead times up to 20 minutes ahead for six different cloud classes as presented by Logothetis et al. (2022)

Dataset: Minute values: 28 days from September–November 2019 in southern Spain (Almeria). Cloud classes: 1: cloud-free (or almost cloud-free); 2L: scattered low clouds; 2 M: scattered multiple clouds; 2H scattered high/middle clouds; 3H: scattered high/middle clouds during half the day, cloud-free during the other half; 4A: overcast cloud conditions during half the day, scattered clouds during the other half.

Image by DLR with data from University of Patras and DLR

9.3.2 Satellite-Based Forecasts

Geostationary meteorological satellites have been operational since the late 1970s. The different geostationary meteorological satellites are now able to observe the complete Earth every 15 or 10 minutes with a spatial resolution of 1–3 km (see Chapter 7, Section 7.4.1). Each satellite observes the Earth from a different location—see Figure 9-9 for the 0º degree service (main mission) of Meteosat Second Generation (MSG) satellites. Each satellite has an onboard multichannel sensor to observe the radiance reflected by the atmosphere and clouds in an appropriate spectral range. This information can be used to assess the attenuating effects of clouds and ultimately estimate the incident radiation reaching a solar panel at any location, applying the models discussed in Chapter 7. In solar forecasting, this information is used to obtain forecasts of solar irradiance up to a few hours ahead and over large areas with a high-revisit frequency without requiring any specialized hardware and thus at low cost.

Figure 9-9. Image of the full Earth disc by MSG satellite located at 0º latitude and longitude

Image by EUMETSAT, contains EUMETSAT Meteosat data, 2023

9.3.2.1 Satellite-Based Forecasting Approaches

Satellite-based forecasting approaches include cloud motion analysis and direct ML methods, such as those employed in ASI forecasting, as discussed.

9.3.2.1.1 Satellite-Based Forecasts Using Cloud Motion Analysis

Since the late 1970s, when the first geostationary meteorological satellites became operational, cloud motion analysis has become a prominent technique to retrieve large-scale wind information for weather forecast models. Originally, this analysis was made by an operator who observed similar cloud patterns between consecutive images and manually drew the most plausible wind vectors between these patterns (Wolfgang Benesch 2007). This rather time-consuming and error-prone approach was later automated by detecting the minimal error between blocks of subsequent images and is still called "block-matching." Its first well-known application for solar energy forecasts was set by Lorenz, Hammer, and Heinemann (2004). The block-matching approach produces a CMV field that is applied on a current image to extrapolate it and to infer the future spatial distribution of clouds (see Figure 9-10).

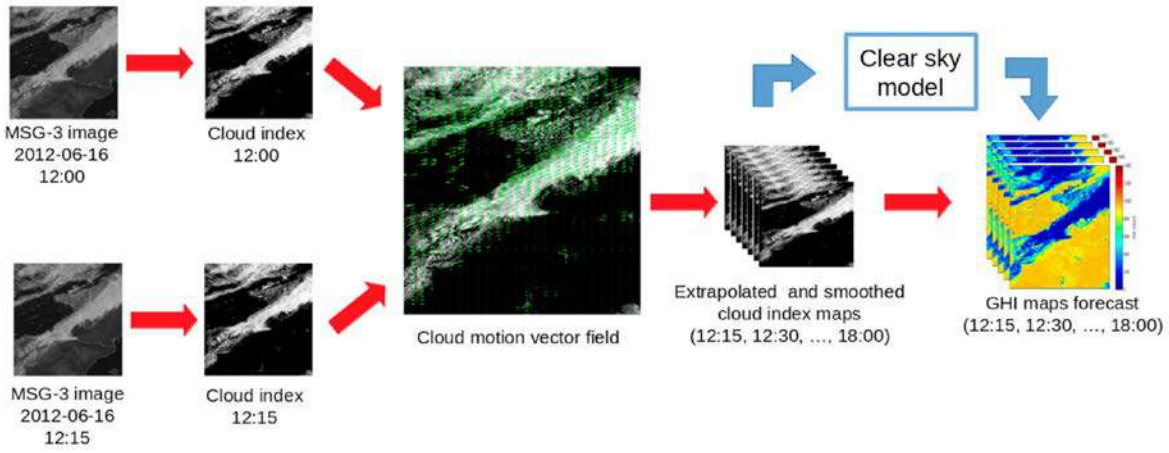

Figure 9-10. Generic process of a satellite-based forecast using CMV analysis

Image by Cros et al. (2020)

In recent decades, several methods have been designed based on these general principles; however, their differences can be established according to the method selected to resolve each main step, namely:

- The CMV field computation
- The extrapolation technique
- The satellite-derived variable to extrapolate, including the method to derive it.

These steps are further detailed in the following.

The block-matching technique has been widely used in cloud motion analysis thanks to its simplicity to ease of use (Lorenz, Hammer, and Heinemann 2004; Perez and Hoff 2013, 233–265; Kühnert, Lorenz, and Heinemann 2013, 267–297; Alonso-Montesinos and Batlles 2015; Cros et al. 2020). Its implementation requires tuning several parameters, such as vector grid resolution and sizes of the blocks to match.

9.3.2.1.1.1 Cloud Motion Vector Field Computation

The development of cloud motion analysis for ASI irradiance forecasting as well as advances in ML-based modeling have inspired researchers to also consider optical flow analysis techniques. These methods have the advantage of producing a dense CMV field (one vector per pixel) and thus propose more realistic extrapolated images than those issued from block-matching. Originating from the computer vision research area, most methods are conveniently available in precoded and well-documented libraries (Lucas and Kanade 1981; Farnebäck 2003; (Sun, Roth, and 2010); however, Bresky and Daniels (2006) noted that optical flow analysis does not provide a noticeable difference for cloud cover forecasts compared to block-matching mainly because the optical flow analysis was designed for fluid animated video (e.g., 24 images/s), whereas satellites provide 1 image every 5–10 minutes in the best case. Nevertheless, the convenient implementation of the optical flow analysis has attracted many solar energy forecasters (Cros et al. 2014; Nonnenmacher and Coimbra 2014; Sirch et al. 2017; Urbich, Bendix, and Müller 2019, Kallio-Myers et al. 2020; Kosmopoulos et al. 2020; Prasad and Kay 2021). A recent comparison of four optical flow methods and block-matching applied for CMV forecasting using

Geostationary Operational Environmental Satellite (GOES)-East satellite images for six sites in South America (Aicardi, Musé, and Alonso-Suárez 2022) shows better forecasting scores of the optical flow methods than the block-matching (see Section 9.3.2.2).

Another technique inspired from rainfall radar data processing is referred to as cloud-shape matching using contour and centroids (Wang et al. 2019). In parallel, sectoral cloud tracking (Schroedter-Homscheidt and Pulvermüller 2011) divides an image part into several parts around the sun position for direct irradiance forecast. Finally, wind vectors derived from NWP models are also used despite the associated operational constraints to obtain these additional data. This can explain why the main references on this approach in an operational context (Müller and Remund 2014; Miller et al. 2018) originate from providers of NWP forecasts.

9.3.2.1.1.2 The Extrapolation Technique

The extrapolation techniques applied to infer future images are not clearly documented in most of the aforementioned publications. Some details can be found in Cros et al. (2020), Gallucci et al. (2018), and Aicardi, Musé, and Alonso-Suárez (2022). The impact of different extrapolation approaches (called "push" and "pull," or their combination) on forecast quality is found to be small. The important aspect is to extrapolate each forecasted image in a step-by-step approach, with consideration for the possibly curved motions defined by the CMV. As a final step in the prediction of future images, smoothing filters are applied to eliminate randomly varying small-scale structures that are hardly predictable; this step can considerably reduce the forecast RMSE (Aicardi, Musé, and Alonso-Suárez 2022; Kühnert 2015; Lorenz, Hammer, and Heinemann 2004); see Chapter 10, Section 10.6.2.3.

9.3.2.1.1.3 Variable Used to Describe Cloudiness

Existing methods are also differentiated by the choice of the variable used to describe cloudiness. Many methods use a dimensionless cloud index or cloud albedo derived from the broadband visible channel of satellite sensors, e.g., using the semiempirical Heliosat method (Hammer et al. 2003; Rigollier, Lefèvre, and Wald 2004; Mueller et al. 2012). Carrière et al. (2021) rather used the clear-sky index, K_c, which is fully correlated with the cloud index. The advantage here is that K_c can be also defined as an objective physical variable. Kosmopoulos et al. (2020) experimented with the use of cloud optical depth because this variable is provided in standard physical satellite-derived products, such as the Nowcasting Satellite Application Facility (SAF-NWC)[56] or the Advanced Very-High-Resolution Radiometer (AVHRR) Processing scheme Over cLouds, Land, and Ocean (APOLLO) (Kriebel et al. 2003). Similarly, Wang et al. (2019) used cloud microphysical data from cloud physical property (CPP)-Surface Insolation under Clear and Cloudy Skies (SICCS).

9.3.2.1.2 Bidimensional Machine Learning Approach

CMV-based methods are proven to provide satisfactory results for nowcasting and have been operationally implemented around the world by different service providers using various geostationary satellites; however, their performance is known to reach a limit when cloud cover does not follow a net advection motion. Cloud appearance or disappearance caused by convective situations, cold or warm fronts passing, or coastal effects are not considered and can induce large forecast errors. To overcome this, some researchers have attempted to compensate

[56] See https://www.nwcsaf.org/.

for this lack of information by using ML approaches. Different types of ML algorithms can be applied following this principle. For instance, Licciardi et al. (2015) used neural networks (autoencoder combined with nonlinear principal component analysis). Dambreville et al. (2014) proposed an autoregression model using external variables (AR-X), where the surrounding GHI is the main input and the statistical CMV field provides external information. André et al. (2019) underlined that the autoregressive approach can be penalized by the lack of cloud motion information. Cros, Deroubaix, and Schmutz (2015) extended the AR-X approach with a dynamic CMV.

According to the aforementioned authors, such approaches can significantly improve the forecast accuracy on very short time horizons (up to approximately 2 hours). For longer time horizons, CMV seems more robust.

9.3.2.1.3 Probabilistic Forecasting

Information on the uncertainty associated with any kind of forecasting is essential for many applications. In particular, probabilistic forecasting (see Section 9.7) constitutes a concrete answer for electricity trading or microgrid management. Probabilistic forecasting is a common feature when using NWP models because they can be operated to produce ensemble forecasts. The synthesis of several forecast members is, by definition, a probabilistic result. Conversely, forecasts based on CMV are, by design, deterministic. The associated uncertainty is typically derived from metrics such as RMSE. Carrière et al. (2021) proposed adding a Gaussian distribution of the errors in the direction of each CMV to obtain a probabilistic speed and direction of cloud motion and then obtain probabilistic forecasts. Figure 9-11 shows the CMV field (top left), a probabilistic representation of the area where the clouds come from (top right), and an example of a probabilistic GHI forecast run (bottom).

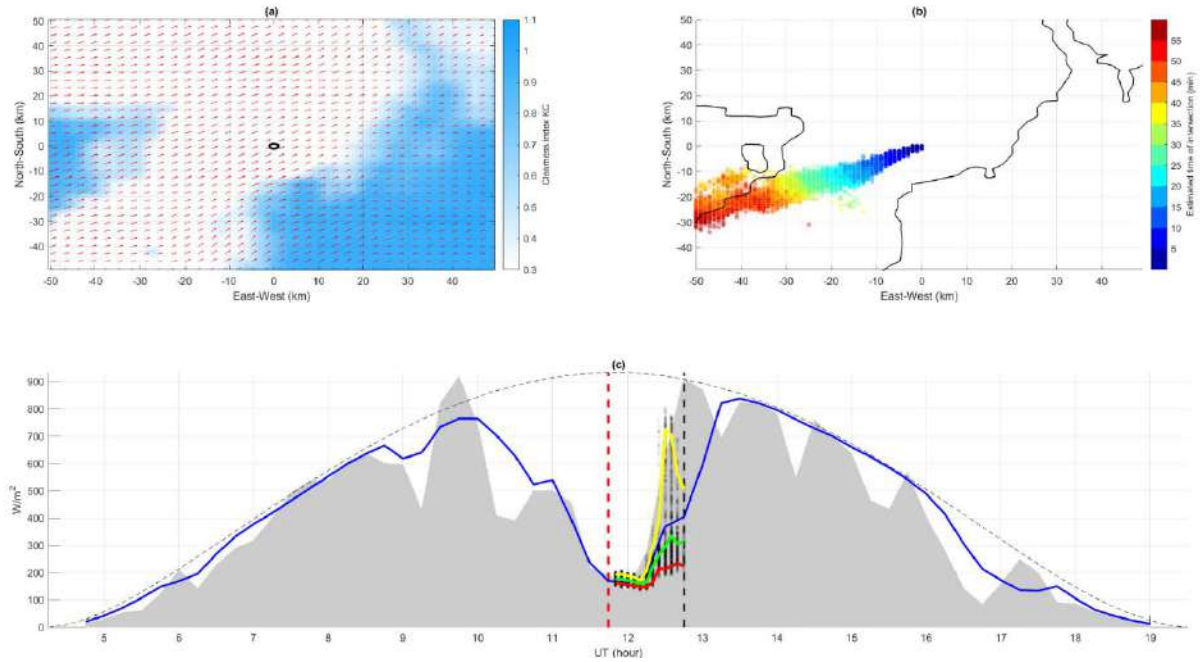

Figure 9-11. Snapshot of CMV-based probabilistic forecast

Example of the CMV-based probabilistic forecast on Carpentras, July 11, 2016, at 12:00 UT, for 1 hour ahead every 5 minutes. (a) Clear-sky index map at 12:00 UT, along with CMV (small red arrows), with a radius of 50 km around the location of interest. (b) Contour map of the clear-sky index map with 0.5 as a threshold, along with dots identifying pixels converging to the monitored perimeter of interest. The color of the dots represents the estimated time of intersection. (c) The gray area represents the reference measured GHI at a 15-minute time step. The blue line represents the corresponding estimation from satellites. The black dashed line represents the corresponding clear-sky GHI. The probabilistic forecasting issued at 12:00 UT (vertical red dashed line) up to 2 hours ahead (vertical black dashed line) is represented by the ensemble of potential converging clear-sky index, K_c, in black dots. For illustration purposes, the 10th, 50th, and 90th percentiles of this ensemble are represented by the red, green, and yellow lines, respectively.

Image by Carrière et al. (2021)

9.3.2.1.4 Satellite Data as Input to Deep Learning Forecast Techniques

Using satellite data as input to DL approaches is becoming increasingly popular in solar forecasting. Using satellite images provides several advantages:

- Data are regularly spatialized with long homogenous archives.
- GHI and/or cloud index are bounded values, limiting the risk of divergence in the training process.
- The available variables are linked by known physical relationships (PV power, GHI, cloud index, …), and thus an elaborate indexation process is not required.
- Small-scale stochastic cloud motion cannot be accurately modeled by thermodynamic equations. DL approaches might overcome this limit to some extent.

Several recent investigations have developed forecast models by applying CNNs on various cloud or GHI datasets derived from geostationary satellites.

Several recent works developed forecast models by applying various CNNs on various cloud or GHI dataset derived from geostationary satellite. Berthomier, Pradel, and Perez (2020) used the U-Net model to forecast the evolution of SAF-NWC cloud masks over the next 2 hours. Nielsen,

Iosifidis, and Karstoft (2021) developed a convLSTM architecture applied to the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Satellite Application Facility on Climate Monitoring (CM SAF) Surface Solar Radiation Data Set - 2 (SARAH-2) satellite-based maps to forecast the surface solar irradiance. Gallo et al. (2022) also trained a convLSTM model with ground and MSG data. Kellerhals, Leeuw, and Rodriguez Rivero (2022) applied a convGRU model on MSG-CPP data for cloud nowcasting. Similar approaches have been applied to the imagery from the Communication Ocean and Meteorological Satellite (COMS) (Ahn, Yu, and Yeom 2022) and Himawari-9 (Jiang et al. 2019) satellites.

9.3.2.2 Performance of Satellite-Based Irradiance Forecasting: Example Comparing Different Cloud Motion Vector Techniques

This section illustrates the performance of satellite-based irradiance forecasting using the CMV technique in the study by Aicardi, Musé, and Alonso-Suárez (2022) as an example. The study compares four optical flow methods and block-matching to obtain hourly GHI forecasts up to 5 hours ahead using GOES-East satellite images for six sites in Uruguay. All methods are operated with individual optimization of the parameter settings and smoothing.

For the different approaches, an analysis of the variation of RMSE with forecast lead time is illustrated in Figure 9-12 and provides results that agree with many of the other studies discussed in Section 9.3.2.1:

- Satellite-based predictions outperform persistence (even in combination with climatology) for forecast horizons from 1–5 hours ahead, with a clear advantage beyond 2 hours ahead.
- The RMSE of GHI forecasts strongly increases with the forecast horizon, as expected.

Further, the comparison shows that different optical flow forecasting methods perform very similarly and can outperform simple block-matching. A comparison of satellite-based methods with both ASI and NWP forecasting is given in Section 9.5.4.

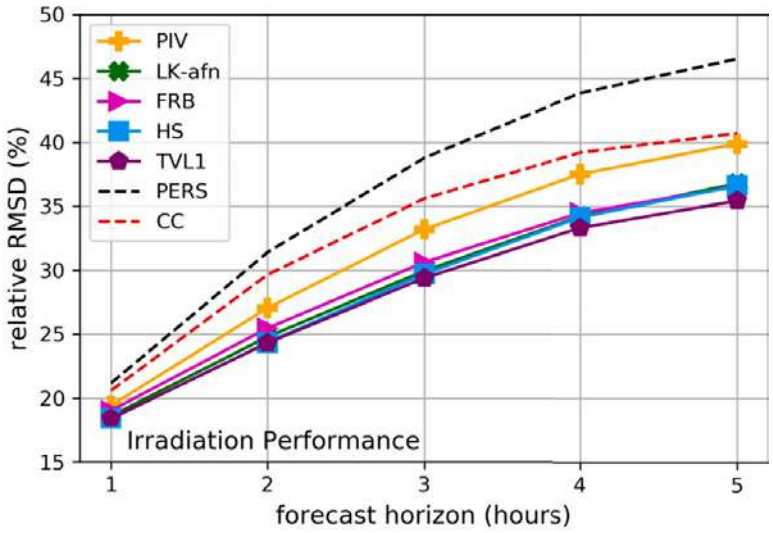

Figure 9-12. Relative RMSE of five different satellite-based CMV forecasting approaches (solid lines, PIV: block-matching; LK-afn: optical flow, Lukas Kanade; FRB: optical flow, Farnebäck; HS: optical flow, Horn and Schnuck; TVL1: optical flow with L1 norm; compared to persistence (PERS) and a model combining persistence and climatology, CC (dashed lines), over forecast lead times up to 5 hours. Dataset: Hourly values, years: 2015 and 2016, six sites in Uruguay.

Image by Aicardi et al. (2022)

In summary, satellite-based forecasting plays an important role in solar irradiance and PV power forecasting up to several hours ahead, typically with hourly or 15-minute resolutions, which is particularly important for intraday power markets. Thanks to the broad coverage of geostationary satellites, forecasts can be provided over large areas with a high-revisit frequency and without any additional hardware requirements. Whereas CMV techniques are still very popular in satellite-based irradiance forecasting, direct ML-based approaches have been increasingly used during the last few years and show promising results, similar to ASI forecasting.

9.4 Numerical Weather Prediction

NWP models are routinely operated by weather services to forecast the state of the atmosphere. Starting from initial conditions that are derived from routine Earth observations from worldwide networks of meteorological sensors, the temporal evolution of the atmosphere is simulated by solving the equations that describe the physical processes occurring in the atmosphere (Figure 9-13). Such physical modeling is the main forecasting approach for time horizons longer than approximately 5 hours ahead. A comprehensive overview of NWP modeling was given by Bauer, Thorpe, and Brunet (2015).

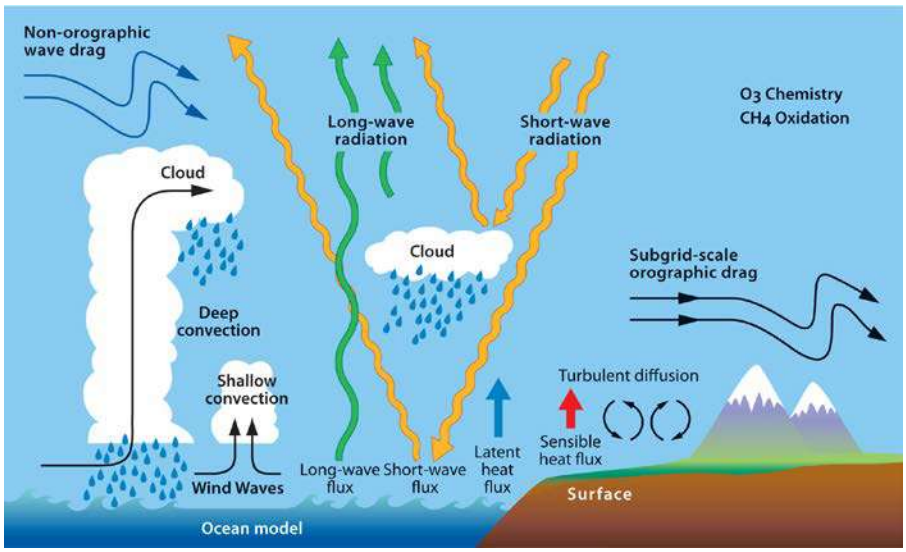

Figure 9-13. Illustration of atmospheric processes modeled in NWP

Image by ECMWF (https://www.ecmwf.int/en/research/modelling-and-prediction/atmospheric-physics), accessed on 28.04.2024; Creative Commons Attribution 4.0 International Public License

9.4.1 Basic Principles of Numerical Weather Prediction Forecasting and Challenges in Irradiance Forecasting

Global NWP models predict the future state of the atmosphere worldwide. To determine the initial state from which an NWP model is run, data assimilation techniques are applied to make efficient use of worldwide meteorological observations (Jones and Fletcher 2013). These include observations from ground-based weather stations, buoys, radiosondes, airplanes, and spaceborne sensors (i.e., satellites), see Figure 9-14.

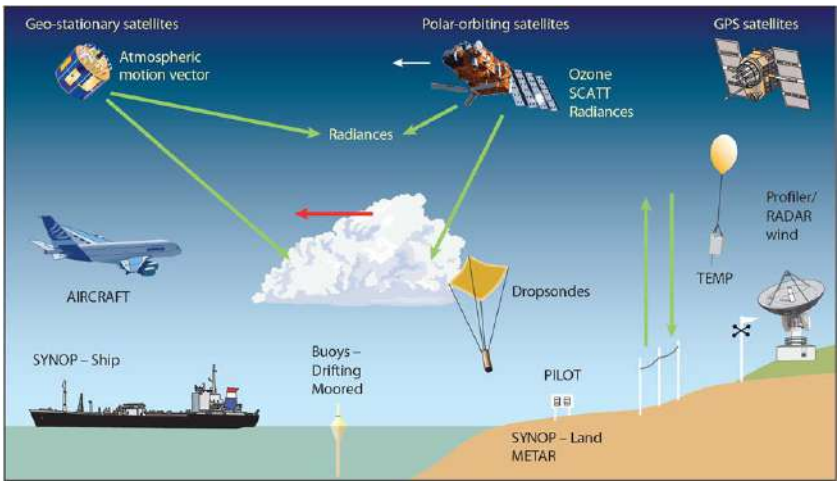

Figure 9-14. Sensors collecting meteorological observations

Image by ECMWF (https://www.ecmwf.int/en/research/data-assimilation/observations), accessed on 28.04.2024; Creative Commons Attribution 4.0 International Public License

NWP models operate by solving conservation equations for the atmosphere and the surface layers immediately below using spatial and temporal discretization. The spatiotemporal resolution of this discretization determines the computational cost of the simulation. Primarily,

the equations of momentum and energy are solved, but equations for hydrometeors, water vapor, radiative transfer, soil moisture, etc., are also needed.

In addition, many physical processes occur on spatial scales smaller than the grid size, including, most importantly, condensation, convection, turbulence, as well as scattering and absorption of shortwave and longwave radiation (see Figure 9-13). This depends on the model resolution, though. Models with a spatial resolution approaching 1 km can resolve deep convection, whereas models with a finer resolution of approximately 100 m can resolve shallow convection as well. The effect of unresolved processes on the mean flow at the model's grid size is evaluated with the so-called parameterizations of atmospheric physics. They include interactions of the land and ocean with the atmosphere, vertical and temporal development of the planetary boundary layer, cumulus triggering and cloud microphysics, as well as shortwave and longwave radiation. These physical parameterizations are key components of obtaining accurate predictions with NWP models. They bridge the small-scale and large-scale processes, and they prompt the convergence of the numerical routines that solve the physical equations.

Currently, global NWP models are run by approximately 15 national and international weather services, and their spatial resolution ranges from approximately 10 km–50 km. The temporal resolution of the global model outputs is typically 1 or 3 hours, their forecasts are normally updated every 6 or 12 hours, and they are run 10 days or longer into the future.

Regional models—also called limited area models or mesoscale models—cover only a limited area of the Earth. They take the lateral boundary conditions from a concurrent or previous global NWP model run and refine the spatial and temporal grid of the global NWP model. A typical example of successively finer grids is shown in Figure 9-15. They use initial conditions from a previous run with the same model and need several months to get slowly varying variables, such as soil water or temperatures, correct. Weather services typically operate mesoscale models with a spatial resolution ranging from 1–10 km, and they provide hourly forecasts, though higher resolutions are feasible. Compared to global models, the higher spatial resolution of mesoscale models allows for explicit modeling of small-scale atmospheric phenomena.

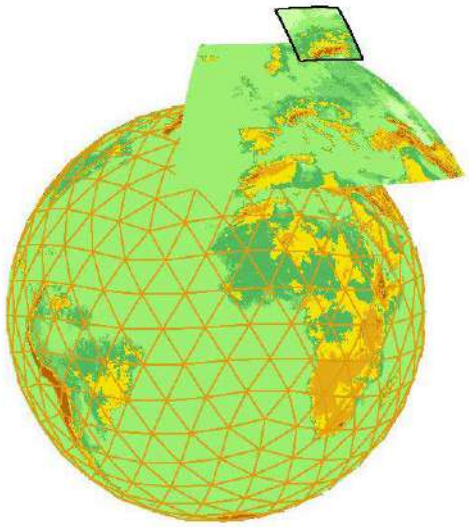

Figure 9-15. Downscaling from global to regional NWP with a higher spatial and temporal resolution

Image by Deutscher Wetterdienst DWD
(https://www.dwd.de/EN/research/weatherforecasting/num_modelling/06_nwp_emergency_response_system/num_weather_prediction_emergency_system_node.html) accessed on 28.04.2024

Clouds and their optical properties are of primary importance for irradiance forecasting. They are strongly affected by the model's initial conditions and by several processes in and around the atmospheric system. For instance, the deep soil's water concentration affects the available cloud water via evapotranspiration from plants and trees. In parallel, clouds related to weather systems at approximately 1,000 km (synoptic) scale are most accurately forecasted. This includes clouds along a warm front with consistent rising motion over a large region.

Thunderstorm clouds (cumulonimbus) appear dark gray and can block most of the solar irradiance. They can form in conditionally strongly unstable atmospheric conditions and are the most difficult to predict. They can form due to cascading outflows from other cumulonimbus clouds that are in essence of a chaotic nature. This process is inherently impossible to accurately predict in day-ahead forecasts. In practice, a region with the likely formation of such clouds is forecasted based on an ensemble of NWP models (see Section 9.7). Here, the members of the ensemble can have perturbed initial conditions, surface conditions, and physics parametrizations. The output from several NWP models can also be combined into multi-model ensembles to improve accuracy.

Fog is also often difficult to forecast. Recent work in several weather centers has focused on better representations of aerosols and cloud-aerosol microphysical interactions to improve these forecasts. Cloud-aerosol interactions have also been shown to be important for cloud ice nucleation, which can create high-level cirrus clouds. These developments are likely to improve the cloud forecasts in the coming years. Additionally, explicit aerosol forecasts in regional NWP models will improve the clear-sky irradiance forecasts compared to the previous models that mostly use fixed aerosol climatologies based on average historical conditions.

Most current NWP models offer the surface GHI as a normal output. Some models also offer direct and/or diffuse irradiance on a horizontal surface as a normal output, or even DNI forecasts. Note that if direct horizontal irradiance, noted DIR in Chapter 8, is accumulated over 1 hour by the model, its conversion to DNI can only be approximate; thus, whenever possible, it is best to obtain DNI directly. Additionally, some models provide clear-sky irradiance, which includes the effect of aerosols and other atmospheric constituents. This clear-sky irradiance is necessary to obtain the clear-sky index. Although in principle the surface irradiance output can be used directly in solar energy applications, in practice additional postprocessing is customarily applied to improve forecast accuracy (see Section 9.5).

9.4.2 Examples of Operational Numerical Weather Prediction Models

Some examples of NWP models are given and specific references are provided with respect to the application and evaluation of irradiance forecasts in the context of solar energy applications. Note that the sample of operational models and applications given here is nonexhaustive; it simply summarizes the research experience and lessons learned from research completed within the frameworks of the IEA SHC Task 36 and Task 46 (now completed) as well as the current IEA PVPS Task 16.

9.4.2.1 Global Numerical Weather Prediction Models

Many of today's operational global NWP models are listed in Table 9-1, including their current spatial resolution. Of these, this section introduces the Integrated Forecasting System (IFS) by the European Centre for Medium-Range Weather Forecasts (ECMWF), and the Global Forecast Systems (GFS) by the U.S. National Centers for Environmental Prediction (NCEP).

Table 9-1. Nonexhaustive List of Global NWP Models and the National Bureaus That Run Them

Meteorological Bureau	Country	Model (Resolution)[a]
BOM	Australia	ACCESS-G (12–17 km)
CMA	China	GRAPES-GFS (26 km)
CMC	Canada	GEM (22 km)
DWD	Germany	ICON (13 km)
JMA	Japan	GSM (20 km)
KMA	South Korea	KIM (12 km)
ECMWF	Europe[b]	IFS (9 km)
MF	France	ARPEGE (7.5-37 km)
NCEP	United States	GFS (23 km)
NCMRWF	India	NEPS-G (12 km)
RUMS	Russia	SL-AV (25 km)
UKMO	United Kingdom	UM (10 km)

[a] Some models are run with varying resolutions for different areas. For these, the range of resolutions is given.
[b] In addition to European member states and cooperating states, Morocco and Israel are cooperating states.

The ECMWF's IFS is a global model currently being operated on 137 vertical levels for high-resolution deterministic forecasts. Since the Summer of 2023, all 51 ensemble members of the IFS have been run with a horizontal grid resolution of approximately 9 km. During each model run, many resources are spent on analyzing the initial model state to be as accurate as possible and balanced. If the model state is not balanced, the model will quickly go from the analyzed state to one that significantly differs. For instance, if more water vapor is present than can be withheld in the atmosphere, this will quickly precipitate and be removed. To avoid this, the temperature needs to be adequate to contain the amount of water vapor to be assimilated in the initial state. To efficiently ingest observations and initialize each model run, the method used in the IFS is called 4-Dimensional Variational (4D-Var) data assimilation (Bonavita and Lean 2021).

IFS irradiance forecasts have been extensively evaluated for PV power forecasting (Lorenz et al. 2009; Lorenz et al. 2011). For example, their excellent performance has been shown over the United States, Canada, and Europe (Perez and Hoff 2013) and over Europe (Lorenz et al. 2016) in the benchmarking studies performed under the auspices of IEA SHC Task 36 and Task 46, respectively; see Section 9.4.3. For research purposes, IFS forecasts are available from the archive free of charge in full resolution. Additionally, a subset of IFS real-time forecasts is made available to the public free of charge.[57]

Similarly, NCEP's GFS is frequently used in PV power forecasting, in part because its forecasts are provided free of charge to any user. It is currently being operated at a spatial resolution of approximately 13 km over 64 vertical levels; however, the outputs are provided using a regular latitude/longitude grid with a relaxed resolution of 0.25º and 46 levels, with an hourly resolution up to 120 hours ahead and a 3-hour resolution up to 240 hours ahead. The model is cycled every 6 hours. Comparisons of intraday GHI forecasts of the GFS and IFS forecasts are discussed in Mathiesen and Kleissl (2011) and Perez et al. (2018); see Figure 9-20.

9.4.2.2 Regional Numerical Weather Prediction Models

Meteorological observations are shared internationally among almost 200 member countries and territories around the world. Many of these countries also run their own national regional weather models, including their own initial state data analysis, which gives an advantage to these models. For instance, the U.K. Met Office develops and runs the Unified Model (UM), which is also used in the UM core partnering countries: Australia, India, Singapore, New Zealand, and South Korea. The German Weather Service (DWD) develops and runs forecasts with the Consortium for Small-scale MOdeling (COSMO) model in collaboration with institutes from Greece, Israel, Italy, Poland, Romania, and Switzerland. In addition, they develop and run the ICOsahedral Nonhydrostatic (ICON) model in a broader consortium including the Max Planck Institute of Meteorology, COSMO, and other computing and meteorological institutes in Germany and Switzerland. Météo-France leads ACCORD (A Consortium for Convection-scale modeling Research and Development), which includes 26 national meteorological services across Europe and the Mediterranean region.

Beyond the models developed and run by national weather centers and consortia of these, the mesoscale (or regional) Weather Research and Forecasting (WRF) model (Skamarock et al.

[57] See https://www.ecmwf.int/en/forecasts/datasets/open-data.

2005) is an open-source NWP model that is widely used in energy meteorology. It was developed in the framework of a long-term collaborative effort of several institutes led by the National Center for Atmospheric Research (NCAR) in the United States. It is now a "community model," meaning that it is publicly and freely available, and it can receive contributions from all participants. The WRF model is nonhydrostatic, has multiple nesting capabilities, and offers several schemes for each different parameterization of the atmospheric physical processes. This makes the WRF model adaptable to widely different climate conditions and different applications over virtually any region of interest.

The WRF model has been extensively evaluated in the context of solar energy applications, and it was also part of the IEA SHC Task 36 NWP benchmark (Perez et al. 2013). Other studies from the last few years have evaluated the model over widely different regions, including Lara-Fanego et al. (2012) in Southern Spain; Isvoranu and Badescu (2013) in Romania; Zempila et al. (2016) in Greece; Aryaputera, Yang, and Walsh (2015) in Singapore; He, Yuan, and Yang (2016) in China; Lima et al. (2016) in Brazil; Gueymard and Jiménez (2018) in Kuwait; and Sosa-Tinoco et al. (2016) in Mexico.

An important milestone in the use of the WRF model for solar radiation applications has been the recent development of WRF-Solar, a dedicated suite of WRF model parameterizations for solar radiation forecasting (Deng et al. 2014; Ruiz-Arias, Dudhia, and Gueymard 2014; Thompson and Eidhammer 2014) within the U.S. Department of Energy's Sun4Cast project (Haupt et al. 2016). Some of these improvements, and others, have been summarized by Jiménez et al. (2016). Moreover, the Sun4Cast project has contributed to the development of the Multisensor Advection Diffusion nowCast (MADCast) system (Descombes et al. 2014), which is a particular configuration of the WRF model for the fast assimilation of satellite reflectance images. That configuration can be used to obtain a proxy field to cloud fraction that can be subsequently advected in WRF and used to compute solar radiation nowcasts. Lee et al. (2017) presented a comparative evaluation of WRF-Solar, MADCast, and satellite-based forecasts and found that WRF-Solar performed generally well at predicting GHI under challenging situations in California. Beyond MADCast, MAD-WRF (Jiménez et al. 2022) blends satellite information directly into WRF.

To extend the WRF-Solar capabilities beyond deterministic forecasts, the WRF-Solar Ensemble Prediction System (WRF-Solar EPS) has been developed. WRF-Solar EPS introduces stochastic perturbations in the most relevant variables for solar irradiance forecasts that have been identified with tangent linear models of selected parameterizations (Yang et al. 2021). The model provides a user-friendly configuration to set the characteristics of the perturbations for each variable and to select the variables to perturb. A detailed description and evaluation of WRF-Solar EPS for the United States is given in Sengupta et al. (2022); see Chapter 10, Section 10.5.2.1. The WRF model is run operationally for solar irradiance forecasting at several public and private entities.

9.4.3 Performance of NWP GHI Forecasts: Results of a Benchmarking Exercise

Results of the IEA SHC Task 36 benchmark of NWP GHI forecasts (Perez et al. 2013) are shown here to illustrate the forecast performance of NWP models up to several days ahead. Although the performance of NWP irradiance forecasts has been clearly improved since then, the general findings outlined here are still valid, as confirmed in many other studies, e.g., those given

in Section 9.4.2. A new worldwide benchmark of NWP irradiance forecasts is planned for the next phase of the IEA PVPS Task 16.

Perez et al. (2013) evaluated different NWP models for a variety of climates in the United States, Canada, and Europe. The evaluations are performed for seven sites in the United States from May 2009–April 2010; three sites in Canada from June 2009–May 2010; and 24 sites in Europe, namely, in Germany, Austria, Switzerland, and Spain, from July 2007–June 2008. The evaluated models include IFS and different implementations of the WRF model for all sites, mostly with no postprocessing.

An analysis of the RMSE variation with the forecast lead time, here grouped to entire days up to several days ahead (Figure 9-16), indicates the following:

- NWP models clearly outperform persistence from the first day onward.
- The RMSE of NWP forecasts slightly increases with increasing lead time.
- A big difference exists in the RMSE of different NWP models, which can be partly attributed to differences in their spatiotemporal resolutions (see Chapter 10, Section 10.3). The IFS forecasts show comparatively low RMSE values for all investigated areas.
- Both the absolute and relative RMSE values strongly depend on climatic conditions. The relative RMSE is typically smaller in sunny areas, as shown in the bottom plot of Figure 9-16 for Spain.

A comparison of the performance of NWP forecasts to satellite-based forecasts and persistence for intraday forecasting is given in Section 9.5.3.

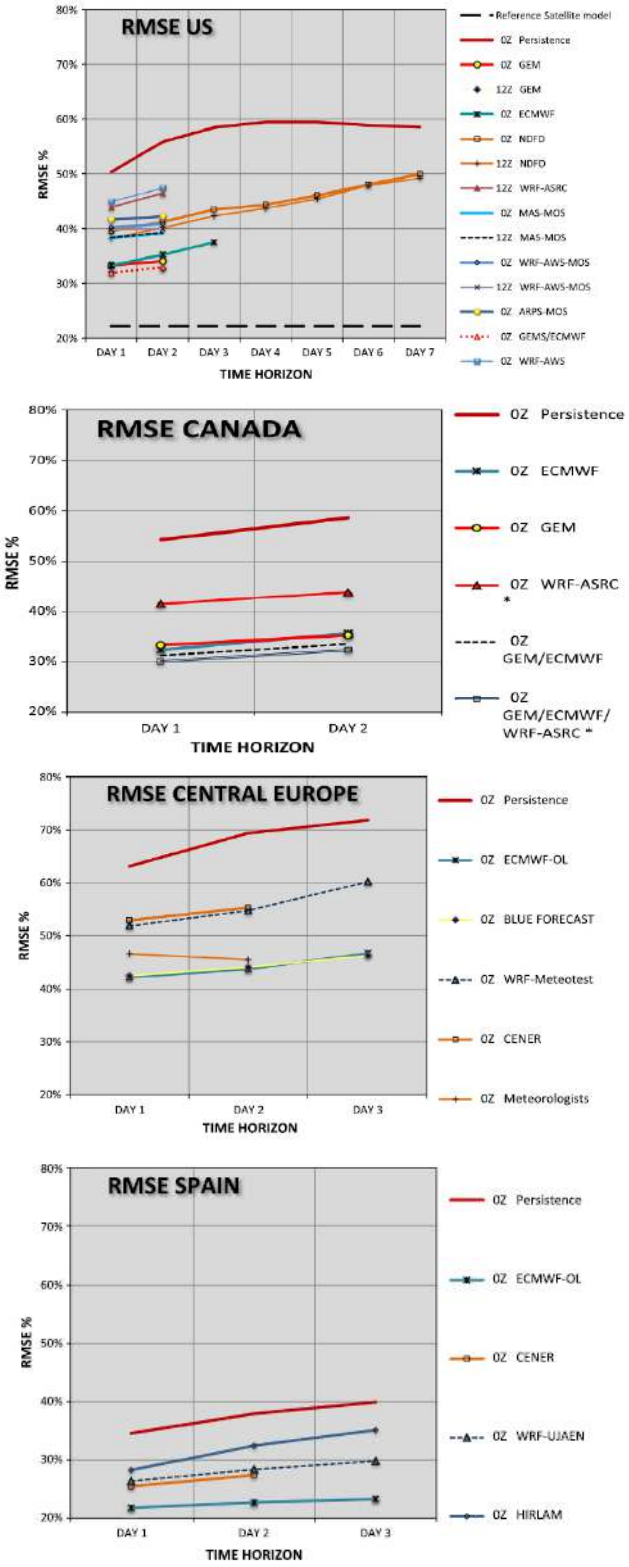

Figure 9-16. Relative RMSE of GHI forecasts for different NWP models and model combinations over forecast horizons up to 3 days

United States: composite of 7 sites (first row); Canada: composite of 3 sites (second row); Central Europe (composite of 21 sites (third row); and Spain: composite of 3 sites (last row). *Image by Perez et al. (2013)*

9.5 Postprocessing and Model Blending With Statistical and Machine Learning Methods

Postprocessing empirical or physical model outputs plays an important role in irradiance and PV power forecasting.

Particularly, various postprocessing methods are applied to:

- Reduce model errors of physical or empirical models by considering unaccounted or partially accounted local and regional effects (e.g., topography or aerosols)
- Combine the outputs of different models (model blending)
- Derive quantities that are not included in the normal model outputs
- Calibrate NWP ensemble forecasts
- Derive probabilistic forecasts from deterministic forecasts.

Currently, postprocessing is mostly performed by statistical or ML methods. These are referred to as "ML models with exogenous input" in the ML community. Nevertheless, empirical and physical models are also employed for postprocessing, especially to derive quantities that are not included in the normal model outputs. Further, a traditional method to obtain improved local forecasts from NWP model outputs is to involve the human knowledge of forecast experts. Especially in difficult forecast situations, such as fog, this alternate approach offers potential for improved irradiance forecasts.

To train ML algorithms, the availability of irradiance and/or PV power measurements is crucial (see Section 9.2). These are used as ground truth for model training. Whereas time-series models require near-real-time data as inputs, postprocessing algorithms can be trained on historic (or "offline") data. Whenever near-real-time (or "online") measurements are available, online training (which consists of regularly retraining the model with new measurements as they become available) is an option to better adapt ML models to current conditions, e.g., seasonal changes. Moreover, satellite-derived irradiance data with their high spatial resolution are a suitable reference for postprocessing NWP model forecasts.

Again, note that uncertainty and especially systematic deviations of ground truth data have a large impact on ML model performance (see Section 9.2.3). Therefore, thorough quality control is essential before using measurements as ground truth for model training. This is also the case for satellite-based irradiance estimates if they are used as ground truth.

The following sections summarize various postprocessing methods for the deterministic forecasting applications enumerated here. The application of statistical and ML models in the context of probabilistic forecasting is addressed in Section 9.7.

9.5.1 Model Output Statistics

MOS are widely used to refine the output of NWP models, primarily to account for local variations in weather and surface conditions. Already more than 50 years ago, Glahn and Lowry (1972) used measurements and/or climatology for specific locations as a basis to adapt the forecasts. Overall, MOS techniques constitute a powerful tool to adapt the results from NWP or satellite-based models to site-specific conditions (e.g., (Gueymard et al. 2012). The set of

predictors consists of various NWP outputs and might be extended by including any relevant information—for example, prior observations or climatological values.

Originally, the term *model output statistics* was associated with the use of regression equations; however, a generalization of this concept now involves other statistical or ML approaches. A bias correction of ECMWF irradiance forecasts in dependence of solar elevation and clear-sky index has been applied by Lorenz et al. (2009). Kalman filters have been proposed by Pelland, Galanis, and Kallos (2013) to improve the irradiance forecasts of the Canadian Global Environmental Multiscale (GEM) model and by Diagne et al. (2014) for WRF solar irradiance forecasts. The application of ANNs to predicted variables from a weather forecasting database has been investigated by, e.g., Marquez and Coimbra (2011). Gastón et al. (2009) used an ML algorithm to enhance the solar irradiance forecasts of the SKIRON model. Pierro et al. (2015) proposed an MOS technique to correct WRF-based GHI forecasts by coupling two intermediate MOS corrections, consisting of correlations with relative humidity and ANNs, respectively. Other powerful postprocessing approaches have been thoroughly reviewed by Yang and Van Der Meer (2021).

9.5.2 Model Blending

Combining —or "blending"—the output of different models can considerably increase the forecast accuracy. First, simple averaging is beneficial for models with similar accuracy, exploiting the fact that forecast errors of different models are usually not perfectly correlated (Lorenz et al. 2016; Perez et al. 2013). Blending methods using more advanced techniques can also account for strengths and weaknesses of the different models for certain situations, for example, by adapting the contribution of each model depending on the weather situation.

A common customer request consists of obtaining short-term forecasts as a single continuous product, even if different observation technologies are used as input. Model blending with statistical and ML approaches is applied to produce such seamless irradiance forecasts covering horizons from several minutes to several days ahead. They integrate different inputs suitable for the different forecast horizons with an optimized weighting. These inputs might include measurements, ASI- and satellite-based forecasts, as wells as NWP forecasts.

Various approaches to this aim have been proposed, mostly based on measurements and/or satellite-based predictions in combination with NWP models. For instance, a weighted average of satellite-based and NWP forecasts was investigated in Lorenz and Heinemann (2012), with the weights optimized for each forecast horizon using linear regression. Kühnert (2015) additionally integrated PV power measurements using the same approach. Bacher, Madsen, and Nielsen (2009) applied an autoregressive model for hourly solar power forecasting combining measurements and NWP forecasts. Sanfilippo et al. (2016) applied a multi-model approach to solar forecasting using supervised classification to select the best predictions from support vector regression and diverse stochastic models. Wolff et al. (2016) and Aguiar et al. (2016) combined forecasts based on support vector regression and ANNs, respectively. Yang et al. (2017) used a hierarchical scheme and minimization of the trace of the forecast error covariance matrix. Within the context of the Sun4Cast project, NCAR's DICast system (Haupt et al. 2018) has been applied to blend multiple solar radiation forecasts. This system—which has already been applied in other forecasting areas, such as transportation, agriculture, and wind energy—consists of a two-step

process: (1) a statistical bias correction process using a dynamic MOS and (2) optimization of the model blending weights for each lead time (Haupt et al. 2016).

Further, the integration of ASI-based forecasting methods to blending models were demonstrated to be valuable for short-term high-resolution forecasting. Pedro et al. (2018) and Huang et al. (2019) assessed intrahour hybrid forecasting models that combine statistical or ML methods with measurements and information extracted from sky imagery and found substantial improvements compared to simple time-series models. Recently, combinations of measurements with ASI- and satellite-based forecasts have also been investigated in López-Cuesta et al. (2023) and Straub et al. (2024, 2023).

Example evaluations of the performance of blending models are given in Section 9.5.4 and demonstrate the benefit of combining different input data over single modes performance.

9.5.3 Postprocessing to Derive Additional Quantities

Not all variables of interest in the context of solar energy forecasting (i.e., global tilted irradiance [GTI], DNI, or PV power) are always available as direct NWP outputs or as a result of CMV forecasts. Postprocessing can be applied to derive these quantities. To that aim, empirical methods are typically employed, but statistical or ML methods as well as physical models are also frequently used to derive the desired quantity from the direct output of the forecasting model.

Although GHI has now become a standard output of NWP models, this was not the case when the field of solar forecasting started to emerge. For example, Perez et al. (2007) proposed an empirical solar radiation forecast model relating sky-cover predictions from the National Digital Forecast Database to the clear-sky index to derive GHI forecasts.

The irradiance components (diffuse horizontal irradiance [DHI] and DNI) are still not provided as normal outputs by all irradiance forecasting systems. To derive them from GHI forecasts, several empirical diffuse or direct fraction models can be used, which were originally developed for application to measurements and later applied to satellite data (see Chapter 7, Section 7.3.1). These models are also being used in DNI forecasting systems that are based on a GHI forecast, as discussed e.g., in Schroedter-Homscheidt, Benedetti, and Killius (2017). For DNI forecasts, several physical postprocessing approaches have also been proposed, specifically for better consideration of aerosols. Breitkreuz et al. (2009) proposed a forecasting approach for direct and diffuse irradiance based on the combination of a chemistry transport model and an NWP model in which forecasts of AOD are directly collected from the chemistry transport model outputs. Similarly, Gueymard and Jiménez (2018) used WRF-Solar with hourly inputs of aerosol forecasts from the National Aeronautics and Space Administration's (NASA's) Goddard Earth Observing System Model 5 (GEOS-5) atmospheric analysis model. Such aerosol forecasts, together with other remote sensing data (ground albedo and ozone) and NWP parameters (water vapor and clouds) are used as input to radiation transfer calculations to derive the irradiance forecasts. A similar approach was used by Lara-Fanego et al. (2012) to derive DNI from WRF output using aerosol observations from the Moderate Resolution Imaging Spectroradiometer (MODIS) onboard the Terra satellite.

In the context of PV applications, forecasting GTI (or plane-of-array [POA] irradiance) or directly PV power is also of interest, as discussed in Section 9.6.1.

9.5.4 Performance of Blending Models: Examples for Intrahour and Intraday Forecasting

The performance of different blending models is illustrated along with a comparison of the different single-model irradiance forecasts used as input to them. The examples shown here cover high-resolution *intrahour* forecasts that integrate measurements and/or ASI and satellite-based forecasts as well as *intraday* forecasts integrating measurements, satellite-based, and NWP forecasts.

9.5.4.1 High-Resolution Intrahour Forecasting

The examples for high-resolution intrahour model blending presented here include:

- A comparison of minute-resolution ASI forecasts, satellite-based CMV forecasts, and a combination of both with lead-time-dependent weights for the DLR-operated Eye2Sky network in Northern Germany, which integrates several ASIs and irradiance measurement stations
- A comparison of ground-based persistence, ASI-based and satellite-based forecasts, as well as a combination of them using linear regression, again with lead-time-dependent weights, but for a radiometric network of eight stations in Freiburg (Southern Germany) (Straub et al. 2024).

For these comparisons, the satellite-based forecasts with their native resolution of 15 minutes are up-sampled to minute-resolution forecasts.

Analyses of the RMSE of the different minute-resolution forecast models over lead times of up to 30 minutes for the two examples (Figure 9-17 and Figure 9-18, respectively) illustrate the following:

- ASI forecasts outperform ground-based persistence from lead times of about 2 minutes onward (Figure 9-18), which agrees with the results in Section 9.3.1.3.
- Up-sampled satellite-based forecasts outperform scaled persistence from about 5 minutes onward (Figure 9-18).
- ASI forecasts outperform satellite-based forecasts up to lead times ranging from 10–20 minutes ahead, depending on whether one is using a network of ASIs or a single ASI, the quality of the models applied to derived forecasts from ASI and the satellite images as well as the climatic conditions for the evaluation site and period.
- Blending models clearly outperform single-model forecasts for all lead times.

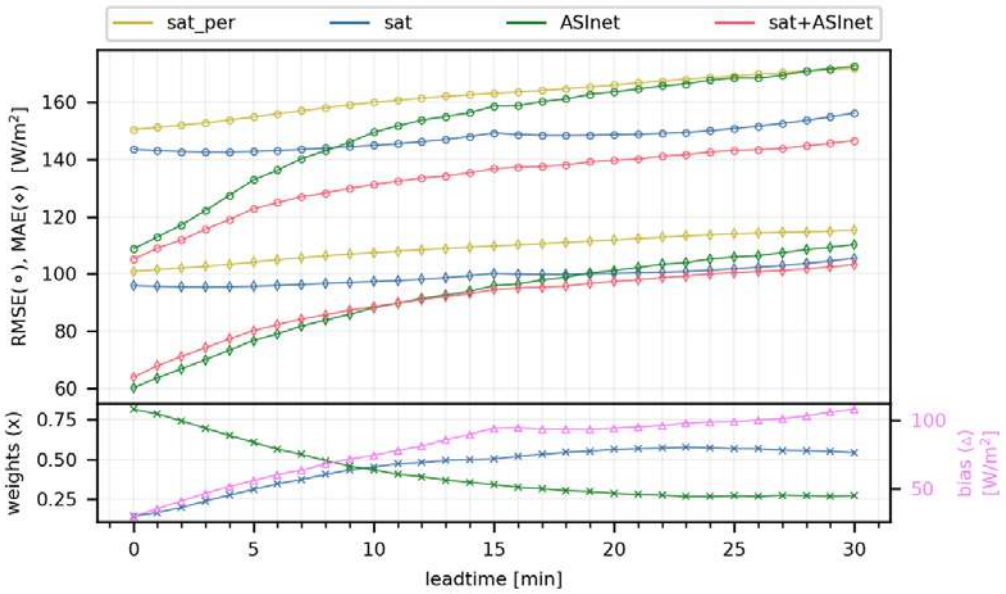

Figure 9-17. RMSE, MAE, and bias (MBE), as well as weight of the blending model of minute-resolution GHI forecasts over forecast lead times up to 30 minutes ahead for satellite-based persistence (sat_per), satellite-based forecasts (sat), forecasts derived from the Eye2Sky ASI network (ASInet), and a combination of the latter two (sat + ASInet)

Dataset: Two validation sites within the DLR Eye2Sky network, August 2020. *Image by DLR Institute of Networked Energy Systems*

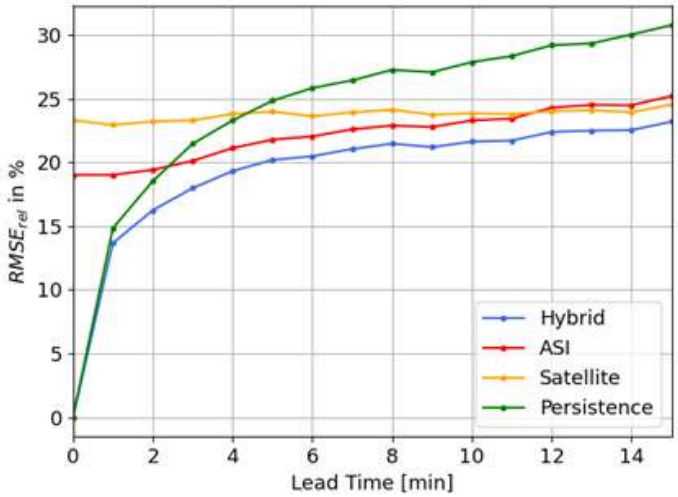

Figure 9-18. Relative RMSE of minute-resolution GHI forecasts over lead times up to 15 minutes ahead for ground-based persistence (persistence), ASI-based forecasts (ASI), satellite-based CMV forecasts (satellite), and a combination of the three (hybrid).

Dataset: Eight sites in Freiburg, May 2021–April 2022; test data: every fourth day, including only complete forecast runs for all models.

Image by Straub et al. (2024)

9.5.4.2 Intraday Forecasting

The examples for intraday forecasting with model blending presented here include:

- A comparison of smart persistence, satellite-based CMV forecasts, different NWP models, and various combination approaches, all at an hourly resolution for intraday and day-ahead forecasting, conducted at seven U.S. sites (Perez et al. 2018).
- A comparison of ground-based scaled persistence, satellite-based CMV forecasts, and IFS forecasts, and a combination those three methods using linear regression with lead-time-dependent weights for sites in Germany, all at a 15-minute resolution.

Analyses of the RMSE over forecast lead times for these examples (Figure 9-19, Figure 9-20) illustrate the following:

- Ground-based-scaled or smart persistence performs better than NWP model forecasts in the first hour or even up to 3 hours ahead, depending on the NWP model (Figure 9-19).
- The performance of the satellite-based forecasts is similar to, or slightly better than, persistence in the first hour (Figure 9-19). For more than 1 hour ahead, satellite-based forecasts clearly outperform persistence. For forecasts at a 15-minute resolution, an advantage of the satellite-based CMV forecasts over scaled persistence is found from 30 minutes onward (Figure 9-20).
- Satellite-based forecasts outperform NWP models up to lead times ranging from 2–5 hours ahead, depending on the NWP model and also on the satellite model and evaluation site and period.
- The RMSE of persistence and satellite-based CMV forecasts increases much faster with the lead time than the RMSE of NWP models.
- The performance of NWP models considerably differs for intraday as well as day-ahead forecasts (Figure 9-19; compare also Section 9.4.3).
- The blending models clearly outperform single-model forecasts for all lead times. Whereas for intraday forecasting, a combination of persistence, satellite-based forecasts, and NWP forecasts is beneficial (Figure 9-20), integrating several NWP models considerably reduces the forecast RMSE in day-ahead forecasting (Figure 9-19).
- The performance of blending models differs with training data, e.g., using site-independent or site-specific ground truth, or using either ground-measured or satellite-derived irradiance values (Figure 9-19). The evaluations show that satellite-derived irradiance data are a suitable alternative to ground measurements for model training if the latter are not available.

These findings agree with other studies, e.g., those given in Section 9.5.2.

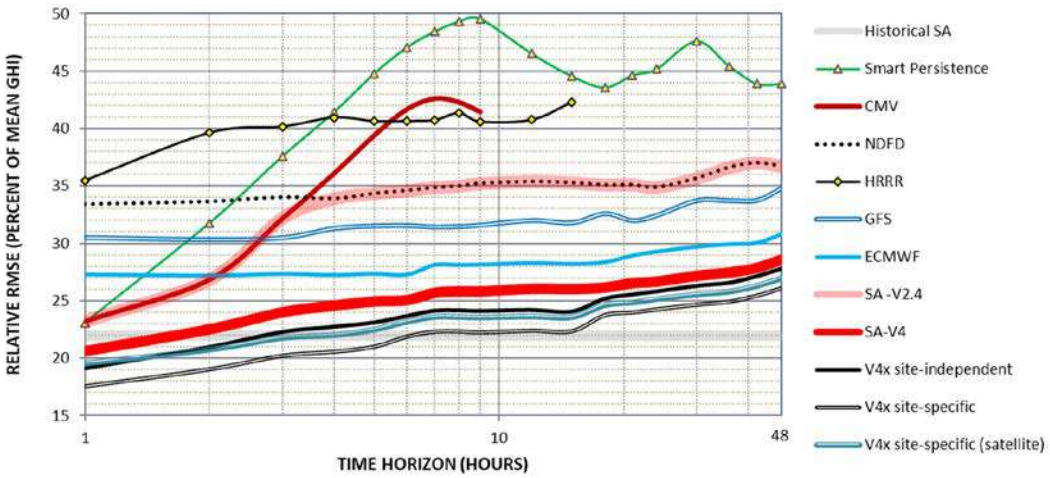

Figure 9-19. Relative RMSE of hourly GHI forecasts up to 48 hours ahead, with different approaches: ground-based smart persistence, satellite-based CMV forecasts (CMV), different NWP forecasts: NDFD, HRRR, GFD by NCEP, ECMWF IFS, and different blending models: SolarAnywhere (SA) V2.4, SA V4, V4 site independent; V4 site specific; and V4 site specific (satellite)

The RMSE of satellite-derived irradiance data (historical SolarAnywhere) is given for comparison. Data: Seven Surface Radiation Budget Network (SURFRAD) sites in the United States from July 2015–April 2016.

Image by Perez et al. (2018)

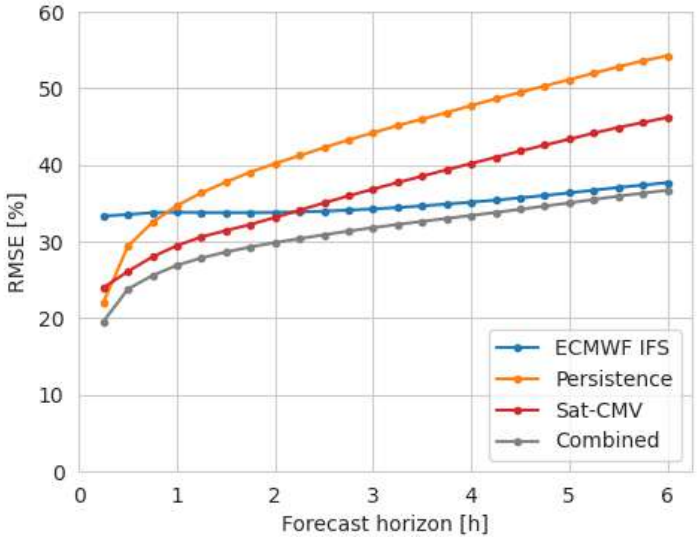

Figure 9-20. Relative RMSE (normalized to the average GHI) of 15-minute-resolution GHI forecasts over lead times up to 6 hours ahead for ground-based scaled persistence (persistence), satellite-based CMV forecasts (Sat-CMV), ECMWF IFS irradiance, and a combination of the three (combined)

Data: Eighteen sites in Germany operated by the German Weather Service during 2018.

Image by Fraunhofer ISE

9.6 PV Power Forecasting

PV power forecasts—pertaining to either a given PV plant, a portfolio of plants, or aggregated regional PV power—are important for plant operators, grid operators, and the marketing of the produced energy. They are based on irradiance predictions with the different models described in sections 9.2 through 9.5. To convert irradiance forecasts into PV power forecasts, physics-based or statistical and ML methods can be applied (Figure 9-21). Both approaches can also be combined. Complementing irradiance forecasts, near-real-time PV power measurements can be included as inputs.

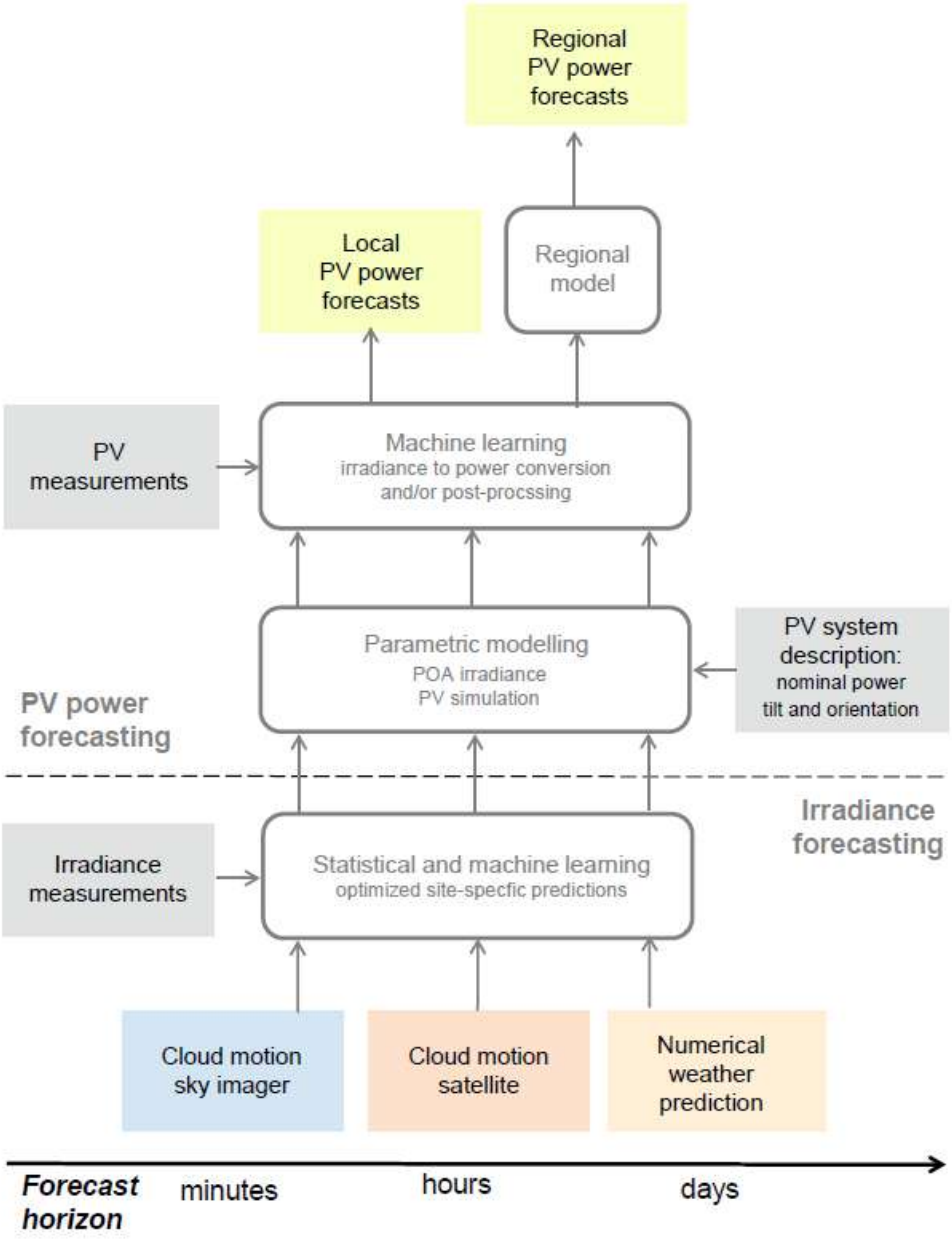

Figure 9-21. Overview of basic modeling steps in PV power prediction

Irradiance prediction: Different forecasting models for different forecast horizons (cloud motion from ASI and satellite data, NWP) and combinations with statistical and ML approaches for optimized site-specific predictions. PV power prediction: Conversion of irradiance to PV power with parametric PV simulation models and/or statistical ML approaches; regional PV power predictions require a regional model (e.g., upscaling) as a last step.

Image reproduced from Lorenz (2018)

Physics-based parametric modeling involves transposing GHI to POA irradiance and then applying a PV simulation model. For this, information on the characteristics of the PV system configuration is required in addition to the meteorological input data. The most important plant data include information on the installed capacity and module tilt and orientation. Further information on the module efficiency and inverter performance can also help to model the PV power output as a function of irradiance and temperature. For more detailed simulations,

additional information can also be used, for example, to model the local shading conditions. Alternatively, the relationship between the PV power output and irradiance forecasts and other input variables can be reverse-engineered with ML methods based on historical datasets containing measured and predicted PV power. In practice, these approaches are often combined; for instance, statistical postprocessing using measured PV power data can be applied to improve predictions with parametric simulation models.

PV power forecasting for plant operation, e.g., storage management, and direct marketing requires forecasts for single plants (Section 9.6.1). In contrast, PV power prediction for grid operators requires forecasts of the aggregated PV power generation for their grid areas or for grid nodes (i.e., regional forecasts are needed instead of single-site forecasts). These regional predictions are typically obtained by regional models, such as upscaling (Section 9.6.2). Portfolio forecasts are based on the same modeling approaches as single PV plants because they are typically generated as the sum of the forecast of the plants contributing to the portfolio. With respect to forecast performance, they benefit from regional smoothing effects, such as forecasts of regionally aggregated PV.

9.6.1 PV Power Forecasting for Single Plants

One way to forecast the production of a PV power plant is to apply a PV power simulation model to the forecast of the relevant predicting variables, primarily irradiance and ambient temperature, but possibly also other meteorological variables (e.g., (Pelland, Galanis, and Kallos 2013; Kühnert 2015)).

Here, the transposition of GHI into GTI to obtain the POA irradiance constitutes the first modeling step. Unless DNI and DHI are explicitly provided by the forecast model, this requires splitting GHI into its direct and diffuse irradiance components. For that purpose, many empirical diffuse or direct fraction models are available (see Section 9.5.3 and Chapter 7, Section 7.3.1 for discussion). Next, the direct and diffuse components are projected or "transposed" to the POA irradiance. The transposition of the direct irradiance is only geometric and thus straightforward. The transposition of the diffuse irradiance to POA requires models that ideally also account for the directional distribution of radiance over the sky, describing anisotropic effects, such as horizon brightening and circumsolar irradiance. Again, empirical models that were developed for the transposition of measured and satellite-derived irradiance data can be directly applied here. An overview of such models is given Chapter 7, Section 7.3.1.

In the next step, the POA irradiance is converted into PV power output. The available PV simulation models and tools, such as those described in more details in Chapter 11, Section 11.7, have been developed mostly in the context of long-term-yield predictions. They can also be directly applied to PV power forecasting. Deeper insight into the modeling of PV power and corresponding variables can be achieved with the tools provided by pvlib,[58] a software package for modeling PV systems (Andrews et al. 2014). Most simple PV simulation models use only the GTI on the POA as input. State-of-the-art PV simulation models consider additional influencing factors. The DC module efficiency depends on the POA irradiance and decreases with increasing temperature, which is typically considered in PV power forecasting systems. Additionally, it is secondarily affected by wind speed and direction, e.g., (Beyer et al. 2004). The angular and

[58] See https://pvlib-python.readthedocs.io/en/stable.

spectral distributions of irradiance are other influencing factors; however, these comparatively small effects are usually not included in detail in current PV power forecasting systems.

The availability of information on PV system parameters can be a problem when using such parametric models. A natural approach is to use the metadata available for the PV system—most importantly, peak power and orientation; however, this information is frequently missing or erroneous, especially for smaller PV systems, which can significantly decrease the PV power forecast quality. Information on the module and inverter specifications is less critical, hence standard values are typically used.

When historical power measurements are available for a PV plant, data-driven models tend to deliver the best results (e.g., Inman, Pedro, and Coimbra 2013; Gensler et al. 2016). Such models can reverse-engineer a lot of effects that depart from the idealized scenario by a physical model or that would require too detailed information, such as how different plant materials perform, how soiled the plant usually is, how the incident radiation is shaded or reflected by local surroundings, or how the plant has aged. Depending on how impactful these factors are at a particular site, generalized physical models might perform quite badly.

However, it is emphasized that measurement issues and plant outages can impact the performance of ML models. Failures of technical components are likely to have large impacts, but the same is true with curtailment caused by grid operation or electricity market price constraints. Proper quality control of PV power measurements and/or adapted training procedures (e.g., Saint-Drenan 2015) are therefore essential when applying ML models in PV power forecasting.

Data-driven models are used in different ways for PV power forecasts. They can be applied for the postprocessing of forecasts with PV simulation tools (e.g., Kühnert 2015). They can also be used to learn PV system parameters from historical data (e.g., Saint-Drenan 2015), or applied to directly transform predicted irradiance and other meteorological parameters into PV power forecasts.

Note that the ultimate objective of such regression or ML models is to minimize an error metric, typically the RMSE between the predictions and ground truth. These methods can only learn from the patterns that exist in the training data, which must be representative of the use case. The result tends to be smoothed, or they do not properly represent very high or low values to avoid large errors. ML can better learn specific factors about a power plant, such as when exactly it starts production at sunrise. But it might also focus the model on best predicting afternoons or summers when production values—and thus absolute errors—are larger. In any case, ML-generated results should always be visually inspected. Nevertheless, ML methods can learn both nonlinear and nonstationary relationships specific to any particular plant (Das et al. 2018; Ulbricht et al. 2013) and are prevalent when historical data is available.

It is not uncommon to train multiple models to the same plant for different weather inputs, time horizons, or seasons. Whereas a physical model uses an established relationship to derive power from irradiance, an ML method implicitly contains some kind of postprocessing to correct bias and optimize the performance between particular input and target data. Hence, different models

might be trained to calculate power from measurements, satellite-derived irradiance data, and intraday or day-ahead forecasts to eventually deliver the lowest errors possible in each use case.

9.6.2 Regional PV Estimation and Forecasting

Regional PV power estimates or forecasts aim to represent the aggregated PV output that is fed into the grid over regions that can range from a few streets feeding into a single, low-voltage substation, to a municipality, to a control zone of a distribution system operator or a transmission system operator, or even an entire country. A region could be geometrically defined, for which the allocation of which plants are in which region is clear, or it could be defined by a shared grid node, in which case detailed grid data might be needed or the selection of plants in the region might be uncertain or dynamic. A regional forecast represents an ensemble of contiguous plants in an area of interest, either geographically or within a branch of the electric grid. Thus, it tends to differ from a portfolio forecast that is used (e.g., in direct marketing).

Different approaches can estimate and forecast regional power, depending on what kind of data are available. Reference plant power measurements can be useful to estimate PV power output or to train forecasts. Weather-based models can use satellite-derived irradiance for real-time power estimation or NWP forecasts to predict future power generation. If metadata on the individual plants is available, the variation of the solar resource and plant characteristics within the region can be considered.

Data availability, however, poses a major challenge. Ideally, detailed registries of all installed plants would exist, and measurements would be available for all of them. Yet, in many countries, many solar plants (especially small ones) have been installed without any record. Existing data often belong to private companies (particularly measurements) or government institutions and often cannot be shared because of legal or privacy restrictions. Some private companies sell commercial datasets, but they are not necessarily representative, and sometimes they are aggregated to anonymize private data. Even among distribution system operators and transmission system operators, data quality and sharing can be a challenge.

Some national and international efforts are underway to create better and even public databases. At the very least, it is necessary to know how much PV capacity is installed in a region up to a given date. Organizations like the European Network of Transmission System Operators for Electricity (ENTSO-E) or the International Renewable Energy Agency (IRENA) aim to provide annual estimates of national PV generation and installed capacity. Detailed registries of data on individual plants are still needed, however. National efforts here include SOWISP in Spain (Jiménez-Garrote et al. 2023) and the market master data register (MaStR) in Germany. Methods for the automated recognition of plants based on satellite or aerial imagery are also improving, aiming to find missing installations and/or to validate information on their characteristics, e.g. (Kleebauer, Horst, and Reudenbach 2021). Nevertheless, this approach is still mostly experimental, so obtaining complete and accurate registries is still difficult in most countries.

Despite these challenges, an important aspect of regional forecasts is that the larger the region and the number of PV plants, the less the uncertainty because positive errors partly balance negative errors. Moreover, including more plants in the mix improves balancing the uncertainty of their characteristics, just as a larger region improves balancing the meteorological uncertainty of the solar resource.

9.6.2.1 Regional PV Power Without Plant Data and Measurements

In some countries, access to PV plant data and measurements is very restricted. When a provider of regional PV estimates and forecasts has no access to plant data or measurements in a region, some general approaches are still possible to evaluate the regional PV production directly from openly available meteorological information, such as satellite-derived irradiance data or NWP analyses or forecasts. If historical power estimates are available from a third party having access to more detailed data (such as from a grid operator), the strictly weather-based model can be optimized to best match the estimates with a method such as upscaling while avoiding the intermediate step of obtaining the details of the installed plants.

Perhaps the simplest and most economic approach is to treat the entire region as a VPP, whereby the solar resource in the region is averaged to a single value and the power is calculated as if originating from a single plant whose characteristics are somewhat representative of the aggregate. This method is known as *model inputs average*, and the power conversion can be made using the same methods as described in Section 9.6.1 for single PV plants (da Silva Fonseca Junior et al. 2014; Zamo et al. 2014; Pierro et al. 2020a). It is also possible to use multiple VPPs for different subregions or characteristics and to combine these into a single model (Pierro et al. 2017). It is not possible, however, to optimize a model for all the existing individual plants because there are too many degrees of freedom to train into a single regional target.

New areas of research include larger-scale ML methods, such as random forest or CNNs, as well as DL. These can be trained to a target estimate using the raw input of gridded weather data, attempting to learn the intraregional dependencies of the regional PV fleet; however, such approaches are not yet as widely researched as single-plant models. They run the risk of overtraining and are computationally expensive for large areas.

9.6.2.2 PV Plant Data

PV plant data for the power plants contributing to regional PV power feed-in is valuable information for regional estimates and forecasting. In its most basic form, plant data include the size, location, and installation date of the plants. This is used by more advanced regional methods that attempt to account for the regional variations of the installed plants as well as for the spatial variability of clouds. Plant location data normally consist of exact coordinates, but sometimes only a city or postal code is available. Depending on the spatial resolution of the reference plants or weather data, the latter might also suffice.

Additional installation information that is necessary or useful to calculate PV production but that is not always available includes, in roughly descending order of importance: the azimuthal direction and tilt of the arrays, the size of the inverter, whether the operator consumes any fraction of the power before feeding it in (with an estimate of how much they do consume on average), and whether there is storage or feed-in limits. The usefulness of some of this information depends on whether the value of interest is the overall production, the actual power feed-in, or the load to the grid.

Plant data can be collected from different sources, such as national registries, commercial datasets, and grid operators. It is more common to find data about large solar plants than small, household solar plants. For some regional PV estimation and forecasting approaches, statistical

modeling is used to account for missing or underrepresented data by using the available data from PV plants to extrapolate how common such plants are overall. Typical PV system characteristics, e.g., tilt and azimuth for different PV system sizes, were investigated in various studies, e.g., (Killinger et al. 2018). For example, the typical orientation of PV plants correlates with their size and the kind of surface on which they are installed, e.g., ground-mounted or rooftop systems. And whether a plant has storage or is designed for self-consumption can depend on its size and age, whether it is part of a household, the kind of incentives that were available at the time of installation, etc.

For regional forecasting, the differences between thousands of plants can quickly even out, although statistical differences might exist when comparing e.g., urban and rural areas. Figure 9-22 shows the distribution of array orientations in the German MaStR registry as a function of the number of plants as well as their capacity; the difference demonstrates how larger solar plans tend to be optimized, with a larger share of the systems oriented toward the equator with close to optimum tilt, whereas a larger variation in tilt as well as azimuth angles is found with smaller plants.

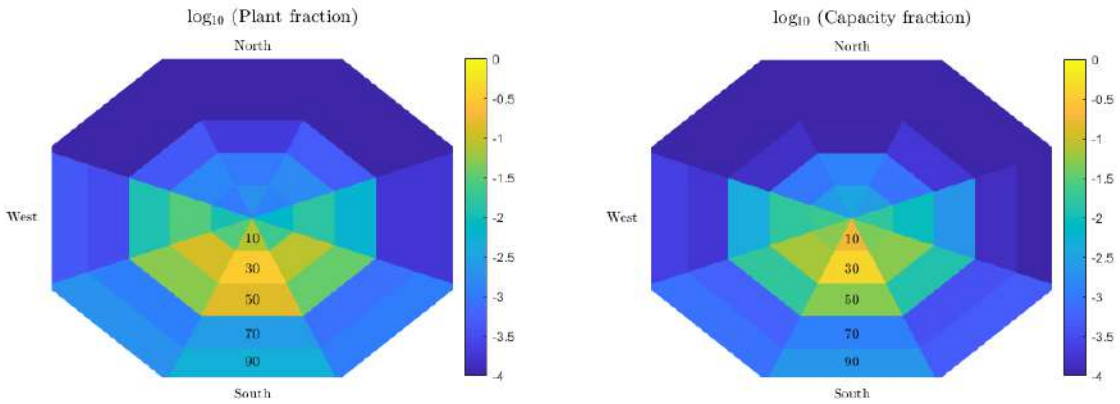

Figure 9-22. Fraction (on a log scale) of the number of plants (left) and overall capacity (right) installed for different plant geometries in Germany according to the MaStR registry as of June 2023

The angles are discretized in eight azimuth and five tilt categories. A linear scale would show essentially all plants to be somewhat southward facing, whereas the log scale better illustrates the frequency of all orientations.

Image by Fraunhofer IEE

9.6.2.3 Regional PV Power Upscaling

If the measurements of all PV power plants in a region are available, the regional production or feed-in would simply be their sum. Measurement availability, however, remains the exception. A region can contain millions of plants with only tens or at most thousands of them (typically large solar plants rather than household rooftop installations) being monitored at high temporal resolutions; however, it is possible to upscale the available power measurements to obtain a value for the entire region without any additional inputs, such as meteorological data, or physical modeling.

Note that there might be two different upscaled values: (1) a real-time estimate for markets and grid operation that is based on the available live measurements and (2) a more accurate estimate that can be calculated days or even weeks later, when all possible data have been gathered. The quality of the upscaling depends on the number and representativeness of the measurements.

The current popularity of the upscaling method stems, at least to some extent, from its large adoption in wind energy, where it is more suitable. Relatively many turbines are monitored, and they behave much more similarly to one another than do PV plants with diverse orientations. As PV installations increase, however, so should the popularity of more complex upscaling methods that account for different PV characteristics (Killinger et al. 2016).

If the reference plants in a region are both well distributed and representative, the first upscaling option is to simply scale the sum of their power measurements to the installed capacity of the entire region according to the plant data. Operationally, such an approach can also first be performed on smaller subregions or plant categories and then aggregated to different regions of interest, which considerably improves accuracy (Kühnert 2015; Lorenz et al. 2011; Saint-Drenan et al. 2016).

Another approach that uses data on PV system location relies on the spatial interpolation or extrapolation of the measurements to derive power values for the plants that are located between them (or simply to points on an installed capacity grid), as illustrated in Figure 9-23. There are many ways of doing this; see Li and Heap (2014) for a review. Inverse distance weighting is simple and robust and thus most common (Bright et al. 2018; Saint-Drenan et al. 2011), though kriging is also popular with irradiation data (Jamaly and Kleissl 2017; Yang et al. 2013).

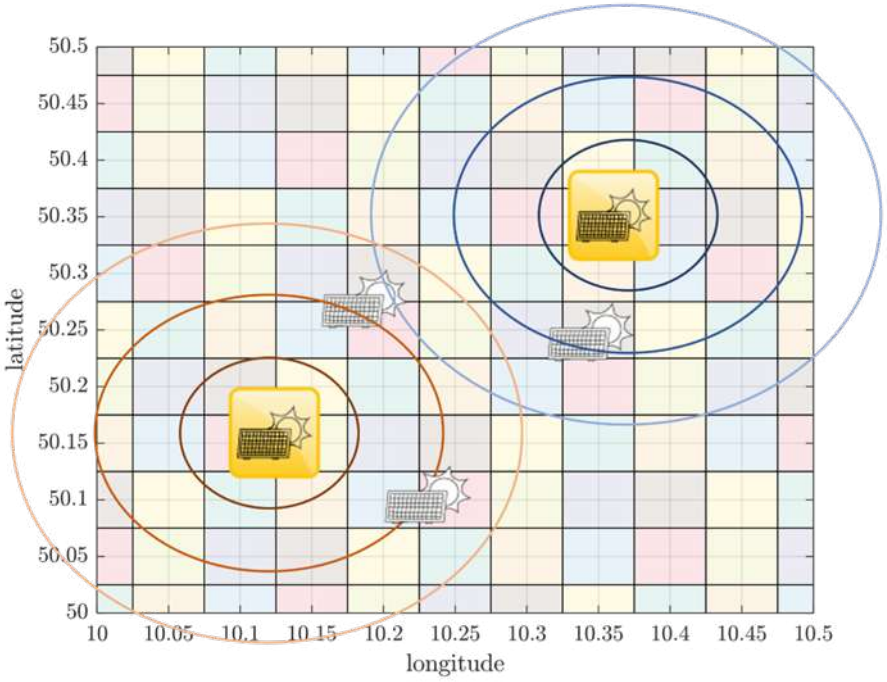

Figure 9-23. Illustration of the upscaling method based on two monitored PV plants (yellow) to also consider nonmonitored plants (black and white)

The weights for each reference plant decrease with distance. *Image by Fraunhofer IEE*

The upscaling approach used with reference measurements for regional power estimation is also popular for regional power forecasting, whereby the problem of regional forecasting is transformed into forecasting the production from every reference plant. As the reference plants have historical measurements to train on, this facilitates a data-driven approach of using, e.g., ML methods to forecast plant power from weather prediction data.

9.6.2.4 Physical Modeling of All PV Plants

Finally, all PV plants in a region can be simulated without any power measurements from meteorological data in combination with generalized physical models. Physical models can describe all known characteristics of each plant in a region and simulate their power output from appropriate local meteorological data, usually at higher spatiotemporal resolution than offered by reference plants. For computational efficiency, nearby plants with the same primary characteristics can also be grouped together.

A generalized physical model is fairly accurate and has the advantage of being able to provide results anywhere, even if a plant is not monitored. Note, however, that for any given plant, a generalized model cannot be as accurate as a model trained to actual measurements. Registry data do not include the efficiency curve of a particular inverter, how the plant might be shaded over different angles, or any other small details that a trained model might consider. In regional forecasting, it is generally sufficient to assume typical values for most plant characteristics and to statistically model or optimize the most important ones.

If the plant geometries are not known, a single POA can be used, or a distribution of tilts and orientations can be simulated at all locations and statistically weighted, e.g., (Saint-Drenan, Good, and Braun 2017). Similar statistical modeling is used by the Copernicus Climate Change Service to generate regional PV power for each region in Europe, with values for each country updated monthly and made available through the Copernicus Data Store.[59] If no local statistics on plant geometry are available, it is possible to modify the known geometry distributions from other countries according to the locally optimized tilt angle (Saint-Drenan et al. 2018). An alternative is to optimize the distribution of geometries to a power estimate. The relationship between plant tilt angles and latitude also somewhat depends on local architectural constraints. For instance, the geometry of installations made on flat roofs, which are typical in some arid climates, might be latitude-optimized as freestanding modules, whereas plants on tilted roofs are typically adapted to the roof orientation and therefore show a larger variation of geometries. In the future, moreover, diverse orientations might also be purposefully selected to better distribute the PV production throughout the day, even in the case of large installations.

Figure 9-24 shows a snapshot of a spatially resolved forecast from such a physical model based on a German plant registry using statistical modeling of the plants' orientations. The power production and feed-in are calculated everywhere according to the local cloud features and can be aggregated according to the local installed capacity. Although the figure's normalized values highlight the solar resource rather than the installed PV capacity, the self-consumption values demonstrate how the plants can have different characteristics in populated areas.

[59] See https://cds.climate.copernicus.eu.

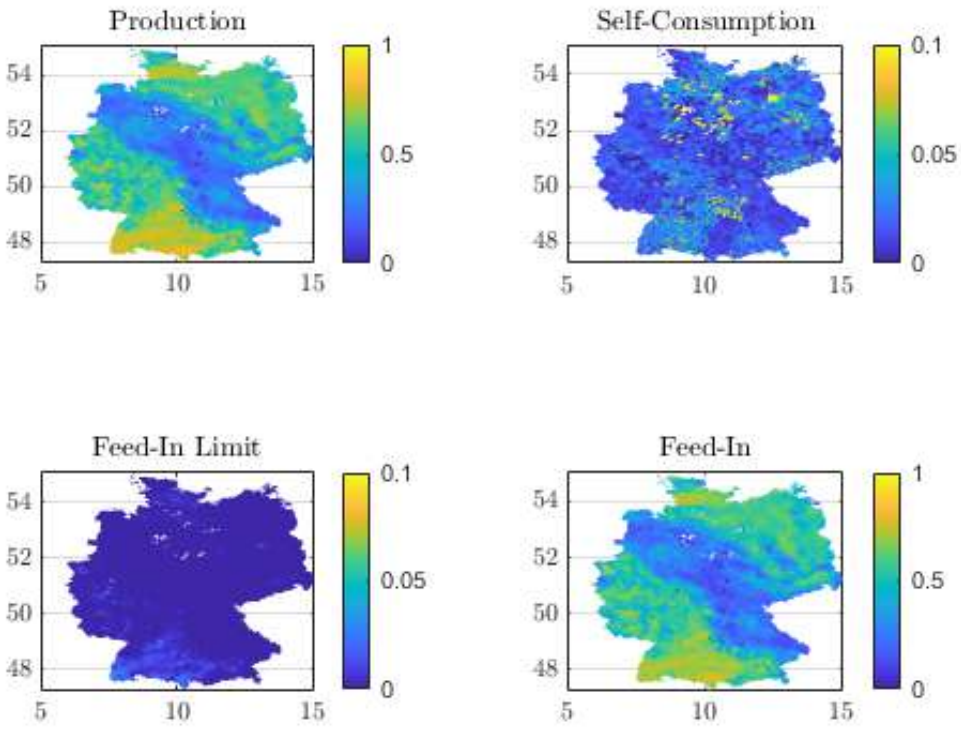

Figure 9-24. Snapshot of estimated local German PV values, normalized by installed capacity, according to real-time satellite data and a physical power model with probabilistic plant information from Fraunhofer IEE

"Feed-in limit" indicates power lost due to plants whose feed-in power is capped at, e.g., 70% of capacity.

Image by Fraunhofer IEE

9.6.2.5 PV Self-Consumption and Regional Power Feed-In

All of the energy produced by a solar power plant is not necessarily fed into the grid, which complicates the estimation and forecasting of PV power feed-in that are needed for smooth grid operation. Many small- or medium-size plants are installed on residences or commercial buildings, and such plants are now increasingly designed to allow behind-the-meter usage (i.e., self-consumption by the user), which was not necessarily the case in the past.

At any given moment, a user typically first self-consumes as much produced energy as possible, stores another fraction if battery storage exists, and feeds in the rest. The self-consumption and feed-in power might also be limited by the inverter size and any feed-in capping by the grid. The energy, S, that can be directly self-consumed at time t is:

$$S(t) = min\,(P_{AC}(t), L(t)), \qquad (9\text{-}5)$$

where P_{AC} is the current PV power generation, and L is the current system load. If the nominal power of a PV plant is small compared to typical loads, it will generally never feed in. If it is very large, then the load becomes less important; however, the load uncertainty of a particular household is typically much greater than that of the solar resource, making it very difficult to accurately forecast the feed-in power of any given plant with self-consumption. Loads are unique to every building and user because they depend on human behavior, have weekly and seasonal

cycles, and exhibit spikes when, e.g., compressors for refrigeration turn on or off. This is an interesting challenge to data-driven and upscaling forecasting strategies because most measurements are made only at the meter, making the measurements nontransferable, even between plants with otherwise identical characteristics. An advantage of physical models is that they do simulate PV generation and thus can be used to distinguish between power generation and feed-in.

If available, registry data on plant self-consumption tend to be either Boolean, i.e., whether or not a plant self-consumes at all, or (more rarely) a scalar value of how much of the production is consumed on average. The time dependency of the load or self-consumption is typically not recorded. Simple strategies to account for the general self-consumption on a plant-by-plant basis do exist in the form of either multiplication by a constant factor or subtraction of a constant load ("band method"). Standard load profiles can be used to introduce a time dependency, although these can greatly differ from any given real load because profiles are generally smooth and designed to represent an average of many users for grid resource planning. The aggregated demand over large regions is much more predictable and should make self-consumption easier to model regionally. Though this is technically a plant-by-plant problem, as self-consumption is determined by local load and generation (Eq. 9-5) and additionally local feed-in limits, battery size or maximum loading speed might also come into play. On a single-plant level, stochastic bottom-up models can be used for local consumption and PV generation, as proposed by Karalus et al. (2023).

The share of self-consumption depends on national regulations and strongly varies between countries and even between regions within a country. For areas with a low self-consumption rate, simple approaches can also be expected to yield reasonable results, on average. Figure 9-25 shows a week of satellite-based estimates of self-consumption compared to production and feed-in power in a control area in Germany Annually, this statistical model estimates that self-consumption currently amounts to approximately 3% of total production in the control zone, which is similar to the value obtained by aggregating the self-consumption averages in the plant metadata. The same model estimates that feed-in limits were rarely active and only reduced the feed-in power by approximately 0.3% annually.

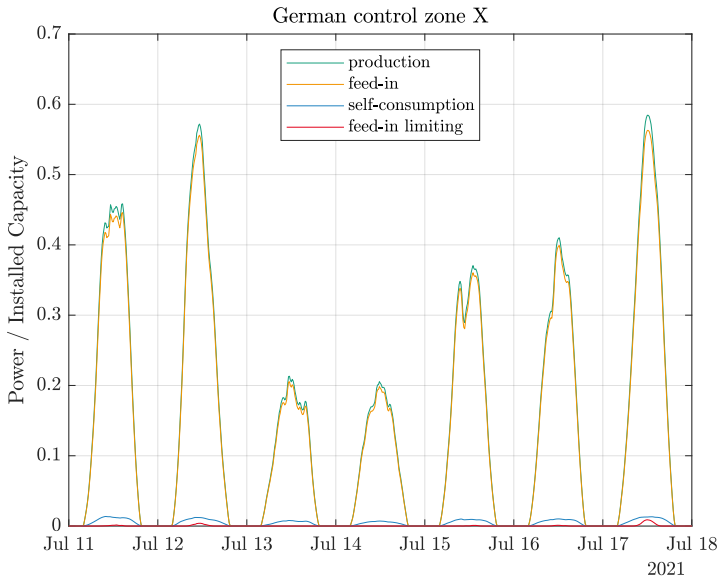

Figure 9-25. One-week time series obtained with the same physical power model and probabilistic plant information from Fraunhofer IEE shown in Figure 9-24 but now aggregated to a control zone

Image by Fraunhofer IEE

The seasonal dependence of PV self-consumption is illustrated in Figure 9-26 for another control area in Germany with an average self-consumption share of 9.5% in 2018 (Karalus et al. 2023). During summer with higher PV generation, self-consumption rates are smaller than during winter with low PV generation. The figure also shows different shares of self-consumption for different portfolios of PV-load systems, depending on PV system size and residential or commercial loads, among other factors.

Figure 9-26. Monthly self-consumption derived with a stochastic bottom-up model for a German control area (TransnetBW) during 2018

Top: Monthly average of daily specific PV generation (normalized to the installed PV power) split into feed-in power (dark colors) and self-consumption (light colors) for a portfolio of PV systems with feed-in monitoring (SOL, green) and for all PV systems in the control area (All, orange). Bottom: Corresponding self-consumption rates per month.

Image by Karalus et al. (2023)

It must be emphasized that the share of PV self-consumption is quickly growing in many countries. In Germany, for example, the majority of new PV installations are designed for self-consumption. At the current rate of PV growth, self-consumption will soon have a first-order effect on regional PV power feed-in. It is difficult to find measurements of both production and feed-in power for plants throughout a region to validate self-consumption models because usually only feed-in power is measured, the measurements are not communicated, or the data are protected by privacy laws. In the future, behind-the-meter forecasting will play a crucial role (Erdener et al. 2022).

9.6.2.6 Regional PV Best Practices

Estimates and forecasts of regional PV production and feed-in can be made by the different methods described here, depending on which plant data, reference measurements, and/or regional estimates are available. Complementary to these basic approaches, preprocessing techniques such as calibrating and blending weather predictions to measurements are common, as are postprocessing techniques such as model output statistics or multi-models trained to good estimates. Figure 9-27 illustrates such a forecasting chain.

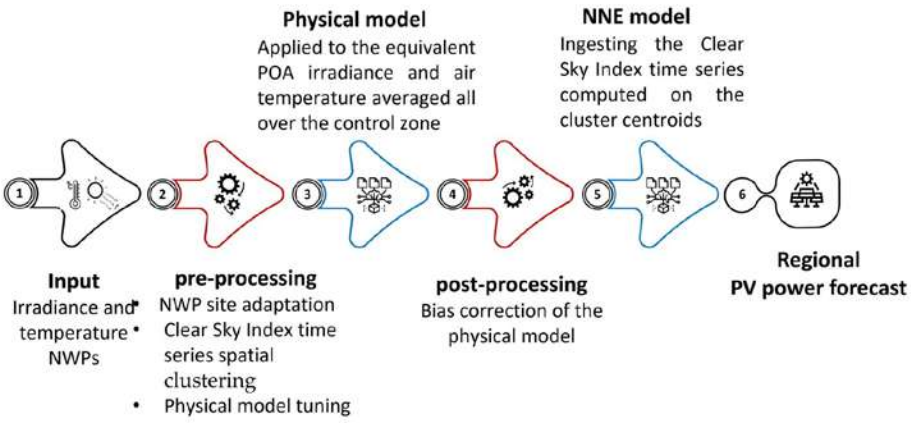

Figure 9-27. Diagram of a hybrid method used to predict regional solar generation

NNE model denotes Neural Network Ensemble model. Image by Eurac Research

Different approaches can also be combined to form hybrid models: A generalized ML model can replace a physical model for a nonmonitored plant or a VPP; physically simulated plants can complement an upscaling approach; and statistical modeling can be restricted to represent only nonmonitored plants, to be combined with data-driven plant forecasts for dynamic aggregation later. Some examples can be found in the literature, e.g., (Gigoni et al. 2018; Pierro et al. 2020a, 2017, 2015), whereas others are proprietary techniques used by commercial forecast providers.

Postprocessing steps are commonly used to improve performance and correct model bias, particularly when there is a target estimate. Different products can finally be combined into a model mix using ML, as is typically performed by a grid operator when receiving data from several forecast providers. Note that such calibrations are not always a better representation of reality because they might only minimize the error to a target estimate that can itself be biased. For instance, this occurs because of the upscaling of unrepresentative plants or the effects of self-consumption and storage.

9.6.2.7 Performance of Regional PV Power Forecasting: Example Evaluation for Italy and Germany

Regional forecasts show much lower uncertainties than single-site forecasts. This also holds for portfolio forecasts for distributed PV systems. By enlarging the footprint of the forecast region of interest, forecast errors are reduced, e.g., (da Silva Fonseca Junior et al. 2014; David et al. 2016a; Hoff and Perez 2012; Kühnert 2015; Lorenz et al. 2009, 2011; Pierro et al. 2020a; Saint-Drenan et al. 2016). This phenomenon, also called regional averaging or smoothing effect, is related to the correlation between the forecast errors at different locations. The larger the region, the less correlated the irradiance conditions are between different sites, and thus also solar forecast errors. This subsequently leads to a higher accuracy of the regional PV power forecasts.

An example is shown in Figure 9-28, which depicts the RMSE of hourly day-ahead forecasts in Italy, obtained by predicting the PV generation of different control areas using averaged model inputs and directly forecasting the power generation at market zone level (Pierro et al. 2020a).

In addition, a measure of PV power variability is displayed in Figure 9-28. With $P(t)$ denoting the PV power output at time *t,* the change in PV power for a given time step, Δt, is defined as:

$$\Delta P_{\Delta t} = P(t) - P(t - \Delta t). \quad (9\text{-}6)$$

Hourly values and a time step, Δt, of 24 hours are specifically considered in Figure 9-28.

The PV power variability in each zone is defined as the standard deviation, $\sigma(\Delta P_{\Delta t})$, as proposed by Perez et al. (2016), which is equivalent to the RMSE of the persistence of PV power:

$$RMSE_{per} = \frac{1}{\sqrt{N}} \sqrt{\sum_{i=1}^{N} (\Delta P_{\Delta t})^2} = \sigma(\Delta P_{\Delta t}). \quad (9\text{-}7)$$

Here, it is commonly assumed that the temporal average of $\Delta P_{\Delta t}$ should be zero.

Both the variability and the forecast errors decrease with an increase in the size of the region and the number of PV systems considered. These quantities can be well fitted either by a hyperbolic function, similar to the one proposed in Perez et al. (2016) or by an exponential function, as proposed by Lorenz et al. (2009). As shown in Figure 9-28, by enlarging the footprint of the forecast region from the prediction of the PV generation in each market zone in Italy to the prediction of the PV generation over all of Italy, the RMSE can decrease from 5.5% (market zones average) to 3.6% (countrywide).

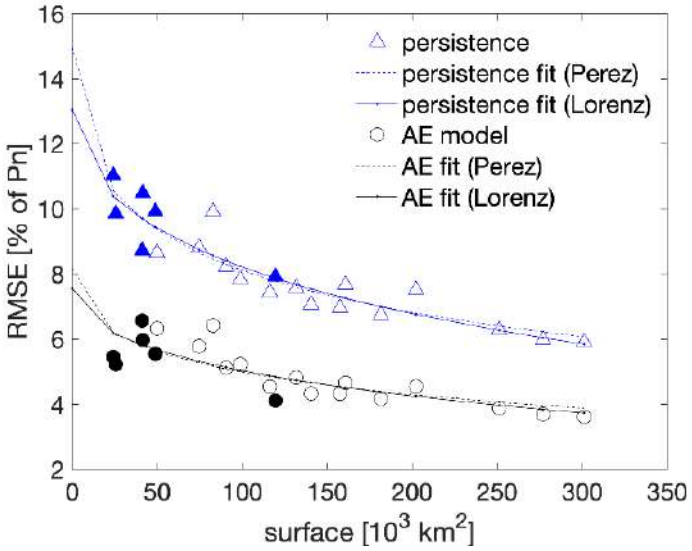

Figure 9-28. Smoothing effect over Italy: Relative RMSE (normalized to the nominal power, Pn) of regional forecasts with an analog ensemble (AE, circles) and persistence (triangles) as a function of the area size of the market zones in Italy (full circles/triangles) and for areas merging several adjacent market zones (empty circles/triangles)

Dashed lines correspond to a fit using the function proposed by Perez et al. (2016), and solid lines correspond to a fit using the function proposed by Lorenz et al. (2009).

Image by Pierro et al. (2020a)

Another important aspect about regional forecasting is that, depending on forecast horizon, performance differences between models and model ranking can change compared to single-site forecasting. An example is illustrated here for the German PV power forecasting system described in Kühnert (2015); see Figure 9-29. Whereas for single sites NWP-based forecasts outperform PV power measurement-based persistence within lead times of less than 1 hour

ahead, measurement-based persistence shows a considerably smaller German-average RMSE than NWP-based forecasts for up to 3 hours ahead. Also, the improvement in the first hours of satellite-based forecasts over NWP-based forecasts is found to be much larger for regional forecasts than for single-site forecasts.

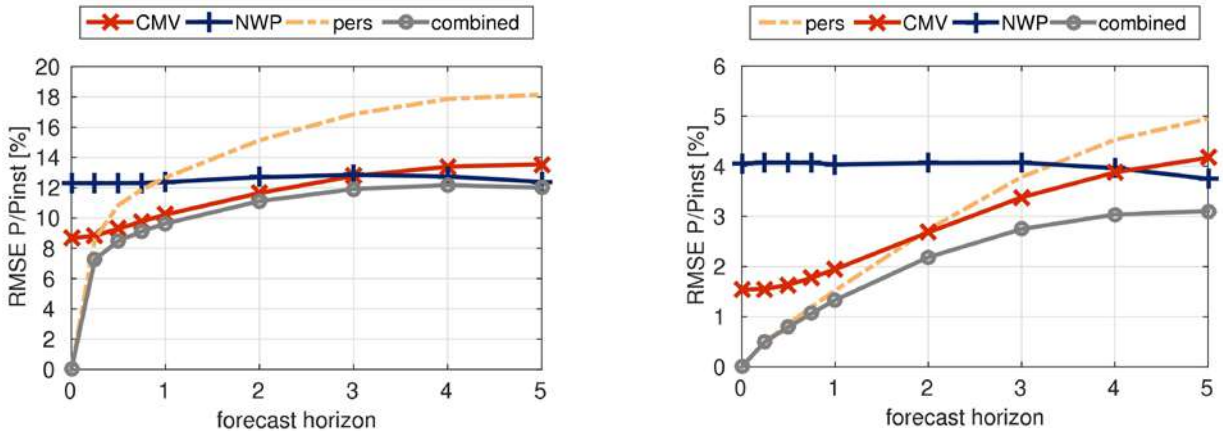

Figure 9-29. Relative RMSE (normalized to the installed PV power, P_{inst}) of 15-minute resolutions over PV power forecasts for lead times up to 5 hours ahead. The results are based on the persistence of PV power measurements (pers), satellite-based CMV (cmv), NWP (ECMWF IFS and DWD COSMO-EU), and a combination of the three (stepwise linear model).

Left: Average single site RMSEs; right: RMSE of aggregated regional PV power considering all sites. Data: 921 PV stations in Germany, May 2013–November 2013, solar zenith angles below 80°.

Image by Kühnert (2015)

Finally, although regional PV power forecasting benefits from spatial averaging because forecast errors decrease especially under variable cloud conditions, challenging weather conditions still remain in regional forecasting. For instance, the large-scale formation and dissipation of fog is difficult to predict and can lead to large regional forecast errors (Köhler et al. 2017). Similarly, snow adhering to PV systems can have a large impact on regional forecast accuracy (Lorenz, Heinemann, and Kurz 2012).

To summarize, regional PV generation exhibits much lower variability and increased forecast accuracy than single PV systems. Exploiting these averaging effects requires a well-developed grid infrastructure and a structure of the energy market adapted to decentralized and variable renewable energy systems.

The accuracy of irradiance forecasts has greatly improved over time, and it has also contributed to reduce uncertainties in the mean irradiance conditions relevant to regional PV feed-in power. The current quality of regional estimates and forecasts thus greatly depends on the quality of PV modeling and the availability of plant data and/or measurements. The procurement of such data is necessary to facilitate a successful energy transition and will require a concerted effort in most countries in the coming years.

9.7 Probabilistic Solar Forecasts

Any forecast is inherently uncertain, and the proper assessment of its associated uncertainty offers grid and plant operators a more informed decision-making framework. For example, a forecast that includes prediction intervals is of genuine added value and, if appropriately incorporated into grid operations, might increase the value of solar power generation (Morales et al. 2014).

This section is restricted to the univariate[60] context that corresponds to those probabilistic forecasts that do not consider the spatiotemporal dependencies generated by stochastic processes such as solar power generation. Two types of solar probabilistic forecasts are considered here: *quantile forecasts* and *ensemble forecasts* (i.e., those using the Ensemble Prediction System (EPS)). Quantile forecasts are quite versatile probabilistic models and, as such, might address a wide range of forecasting time horizons, whereas NWP-based EPS forecasts generally provide probabilistic forecasts for one or several days ahead. Probabilistic forecasting requires a rather complex verification framework, which is introduced in Chapter 10, Section 10.5.1.5. The evaluation framework is based on visual diagnostic tools and a set of scores that mostly originate from the weather forecast verification community (Wilks 2019). What follows constitutes an overview of the basic concepts related to solar probabilistic forecasting methods. Comprehensive overviews regarding forecasting methods and the verification of solar probabilistic forecasts metrics can be found in Antonanzas et al. (2016); Lauret, David, and Pinson (2019); and Van Der Meer, Widén, and Munkhammar (2018).

9.7.1 Nature of Probabilistic Forecasts of Continuous Variables

In contrast to deterministic forecasts, probabilistic forecasts provide additional information about the inherent uncertainty embodied in all forecasting models. The probabilistic forecast of a continuous variable, such as solar power generation or solar irradiance, takes the form of either a cumulative distribution function (CDF), $F(Y)$, or a probability distribution function (PDF), $f(Y)$, of the random variable of interest, Y (e.g., GHI). In particular, the CDF of a random variable, Y, is given as:

$$F(y) = P(Y \leq y) \tag{9-8}$$

where $P(Y \leq y)$ represents the probability that Y is less or equal to y.

The predictive distribution can be summarized by a set of discrete quantiles. The quantile, q_τ, at probability level $\tau \in [0,1]$ is defined as follows:

$$q_\tau = F^{-1}(\tau), \tag{9-9}$$

where F^{-1} is the so-called *quantile function*. A quantile, q_τ, corresponds to the threshold value below which an event, y, materializes with a probability level, τ.

[60] Future work will be devoted to multivariate probabilistic models capable of capturing the spatiotemporal correlations present in irradiance and PV forecasts.

Further, prediction intervals (also called interval forecasts) can be inferred from a set of quantiles. Prediction intervals define the range of values within which the observation is expected to be with a certain probability (i.e., its nominal coverage rate) (Pinson et al. 2007). For example, a central prediction interval with a coverage rate of 95% is estimated by using the quantile $q_{\tau=0.025}$ as the lower bound and $q_{\tau=0.975}$ as the upper bound. Figure 9-30 shows an example of probabilistic forecasts of solar irradiance where prediction intervals have been computed for nominal coverage rates ranging from 20%–80%.

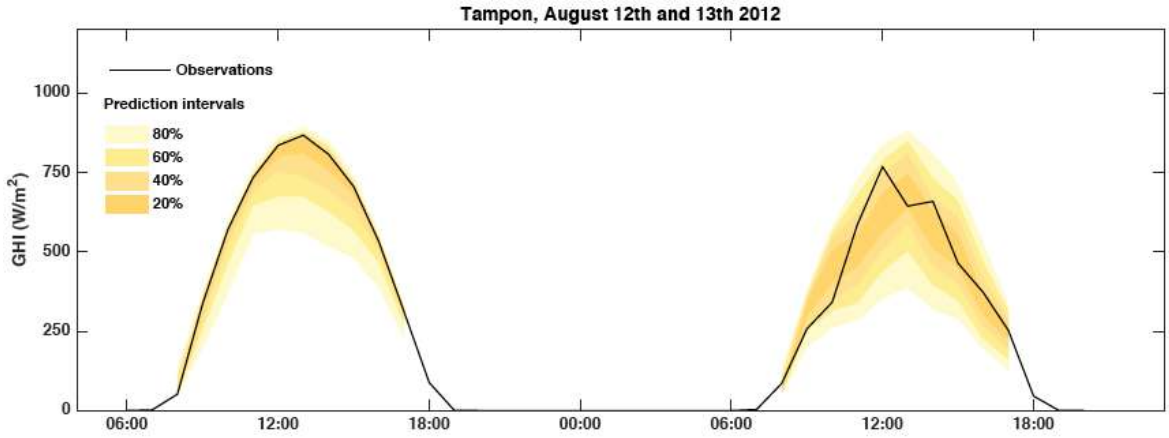

Figure 9-30. Example of probabilistic solar irradiance forecasts: 2 days of measured GHI at Le Tampon, France, and associated 1-hour-ahead forecasts with prediction intervals (yellow) generated with the quantile regression forest model

Image by PIMENT, University of La Reunion

9.7.2 Quantile Forecasts

Two approaches are commonly used in the community to generate quantile forecasts (see Figure 9-31) for different forecast horizons. As input, they use either online ground observations and satellite images for intraday forecasting or NWP deterministic forecasts, which are more effective for day-ahead forecasting. The former approach (e.g., Bacher, Madsen, and Nielsen 2009; Pedro et al. 2018) consists of directly generating the quantiles of the predictive distribution of the variable of interest (e.g., GHI, DNI, or PV power). The latter approach (e.g., David et al. 2016; Grantham, Gel, and Boland 2016; Lorenz et al. 2009; Pierro et al. 2020b) seeks to produce the interval forecasts from the combination of a deterministic (point) forecast and quantiles of the prediction error. In both approaches, the quantiles can be estimated either by assuming a parametric law for the predictive distribution or by nonparametric methods, which make no assumptions about the shape of the predictive distribution.

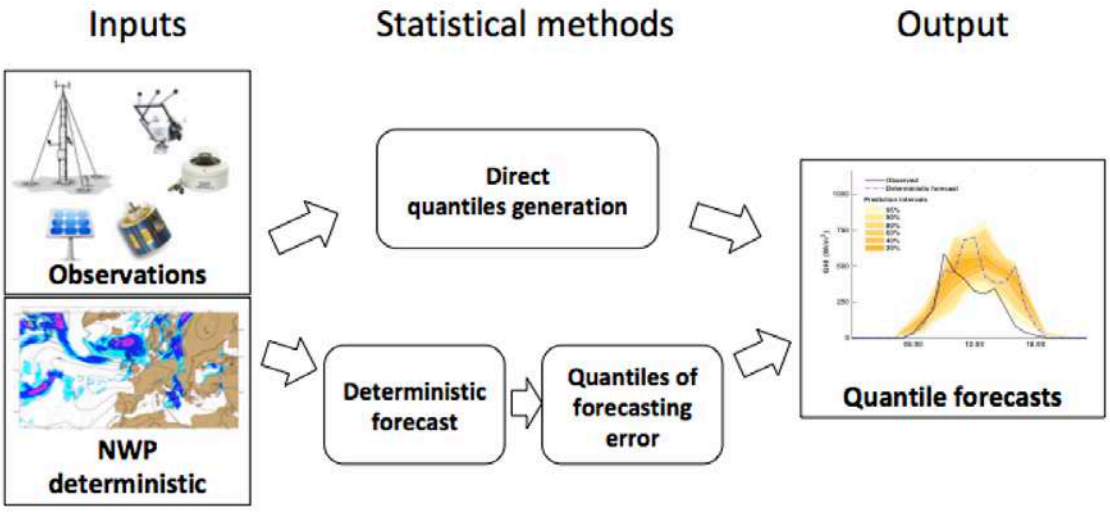

Figure 9-31. Two typical workflows used to generate quantile forecasts from recent past observations and/or deterministic NWP forecasts

Image by PIMENT, University of La Reunion

9.7.2.1 Parametric Methods

Parametric models assume that the variable of interest or the prediction error follows a known law of distribution (e.g., a doubly truncated Gaussian for GHI or a Gaussian for the error distribution). Only a few quantities (e.g., mean and variance) are needed to fully characterize the predictive distribution. Consequently, this approach is particularly interesting in an operational context because it requires low computational effort.

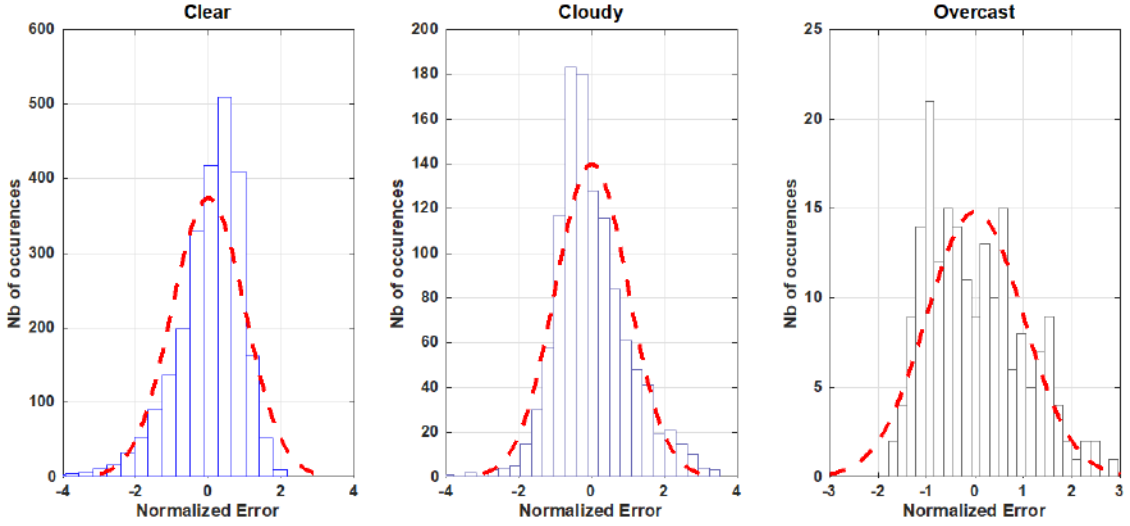

Figure 9-32. PDF of the normalized error (zero mean and unit variance) of the hourly profile of day-ahead forecasts of the clear-sky index provided by ECMWF for three different sky conditions and for the site of Saint-Pierre (21.34°S, 55.49°E), Reunion, France, in 2012

The red dashed line represents the fitted standard normal PDF.

Image from David and Lauret (2018)

In the solar forecasting community, it is very common to fit a Gaussian distribution to the errors even though errors derived from deterministic forecasts of solar irradiance or of the clear-sky index do not usually follow a Gaussian distribution (see Figure 9-32). For instance, Lorenz et al. (2009) developed a probabilistic irradiance forecasting model by assuming a Gaussian distribution of the error of the deterministic GHI forecasts generated by the IFS. More precisely, the predictive CDF was a Gaussian distribution with a mean corresponding to the point forecast and a standard deviation derived from a fourth-degree polynomial function for different classes of cloud index and solar elevation. For intrahour and intraday solar irradiance probabilistic forecasts, David et al. (2016) assumed a Gaussian error distribution of the deterministic forecast to generate a predictive CDF with a Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) model. Instead of fitting a parametric PDF to the error distribution, Fatemi, Kuh, and Fripp (2018) proposed a framework for parametric probabilistic forecasts of solar irradiance using the beta distribution and standard two-sided power distribution.

9.7.2.2 Nonparametric Methods

To circumvent the necessity of making assumptions about the shape of the predictive distribution, numerous nonparametric methods have been proposed in the literature, e.g., Van Der Meer, Widén, and Munkhammar 2018. Examples of techniques include bootstrapping (Efron 1979; Grantham, Gel, and Boland 2016), kernel density estimation (Parzen 1962), and k-nearest neighbors (Pedro et al. 2018). Here, two prominent and simple nonparametric methods are briefly discussed: the quantile regression and the analog ensemble (AnEn) technique.

Quantile regression models relate quantiles of the variable of interest (predictand) to a set of explanatory variables (predictors). Statistical or ML techniques—such as linear quantile regression, quantile regression forest, or gradient boosting (David and Lauret 2018; Van Der Meer, Widén, and Munkhammar 2018)—are commonly used to produce the set of discrete quantiles with probability levels spanning the unit interval (see Figure 9-33).

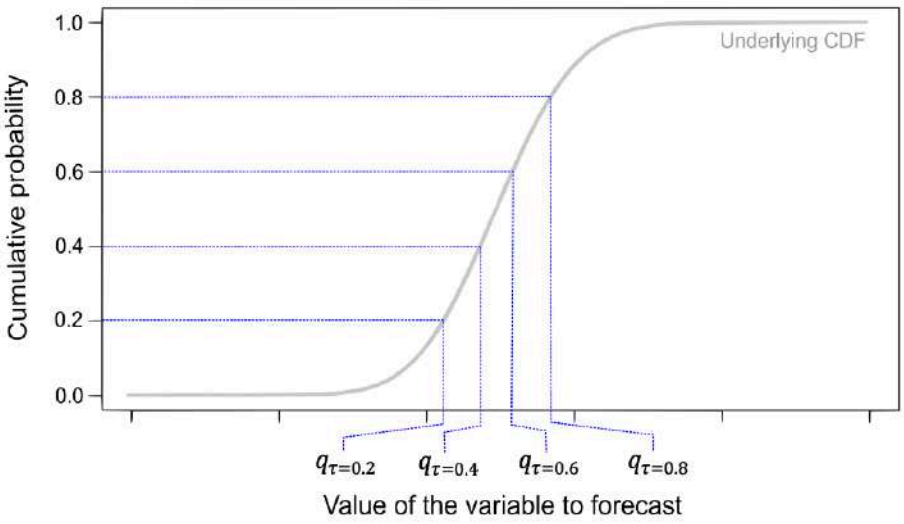

Figure 9-33. Illustration of a set of four discrete quantiles with probabilities ranging from 0.2–0.8

Image by PIMENT, University of La Reunion

The following summarizes the linear quantile regression method first proposed by Koenker and Bassett (1978); see David, Luis, and Lauret (2018) for details about the implementation of other regression methods, including other variants of the linear quantile regression, quantile regression forest, quantile regression neural network, and boosting.

The linear quantile regression technique estimates a set of quantiles of the CDF, F, of some response variable, Y (the predictand), by assuming a linear relationship between the quantiles of Y (q_τ) and a set of explanatory variables, X (the predictors):

$$q_\tau = \beta_\tau X + \epsilon, \tag{9-10}$$

where β_τ is a vector of the parameters to be optimized at each probability level, τ, and ϵ represents a random error term (Koenker and Bassett 1978).

Numerous implementations of the linear quantile regression technique (and of its related variants) have been proposed in the literature to generate quantile forecasts for different forecast horizons and using different types of predictors, X; see, e.g., Bacher, Madsen, and Nielsen (2009), Bakker et al. (2019), and Zamo et al. (2014) for NWP-based forecasts; Lauret, David, and Pedro (2017) for time-series forecasting; Nouri et al. (2023) for ASI forecasting; and Van Der Meer, Widén, and Munkhammar (2018) for a wider review.

The AnEn method (Delle Monache et al. 2013) is a simple nonparametric technique used to build the predictive distributions. The aim is to search for similar forecasted conditions in the historical data and to create a probability distribution with the corresponding observations. Alessandrini et al. (2015) applied an AnEn approach to a set of predicted meteorological variables (e.g., GHI, cloud cover, and air temperature) generated by the Regional Atmospheric Modeling System (RAMS). Note that the AnEn technique is mostly employed for day-ahead forecasting and generates the predictive distribution using NWP deterministic forecasts.

9.7.3 Ensemble Prediction System

9.7.3.1 Definition

The EPS corresponds to a perturbed set of forecasts generated by slightly changing the initial conditions of the control run and of the modeling of unresolved phenomena (Leutbecher and Palmer 2008). Figure 9-34 shows a schematic representation of an ensemble forecast generated by an NWP model. The trajectories of the perturbed forecasts (blue lines) can strongly differ from the control run (red line). The spread of the resulting members (blue-shaded area) represents the forecast uncertainty. For example, the ECMWF provides an ensemble forecast from the IFS model. It consists of 1 control run and 50 “perturbed” members.

Though members of the ensemble are not directly linked to the notion of quantiles, they can be seen as discrete estimates of a CDF when they are sorted in ascending order. Lauret, David, and Pinson (2019) proposed different ways to associate these sorted members to a CDF.

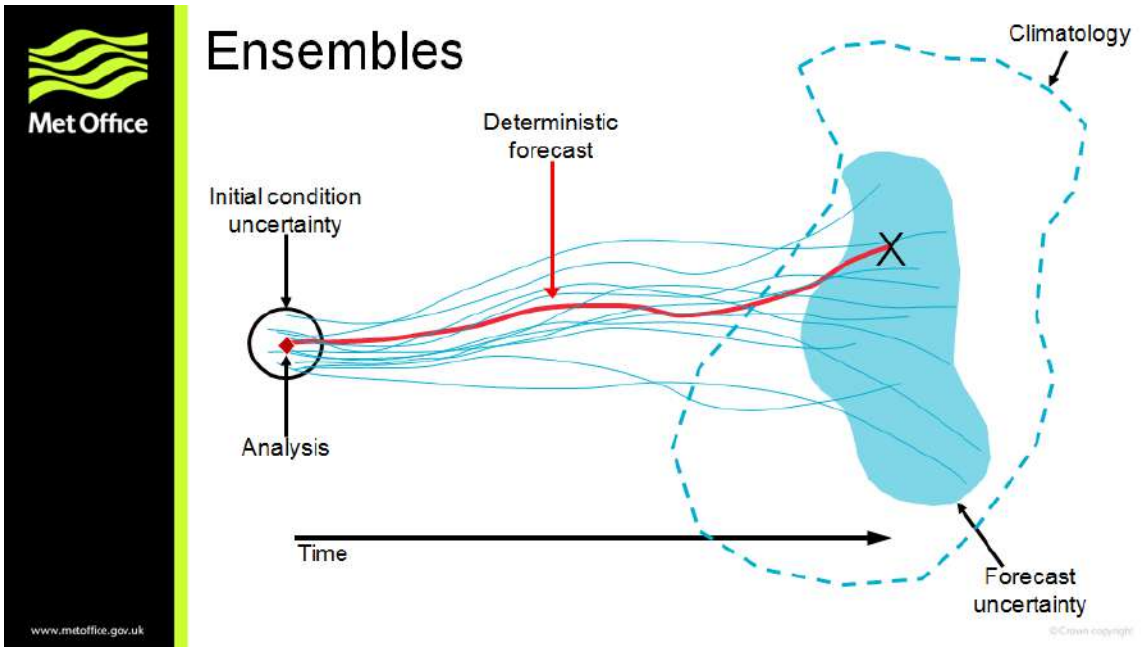

Figure 9-34. A schematic illustration of an ensemble forecast generated with an NWP model

Image from Met Office, © British Crown copyright (2021)

9.7.3.2 Postprocessing of the Ensemble Prediction System

Global and regional NWPs are designed to forecast a large variety of meteorological variables (precipitation and temperature being of utmost importance for society) and have not previously focused on the accurate generation of the different components of solar radiation. Consequently, raw ensembles provided by meteorological centers suffer from a lack of accuracy, a lack of calibration, or both (Leutbecher and Palmer 2008). See, e.g., Yang (2020) for definitions and discussions about the specific meaning of accuracy, calibration, and other specialized terms in the field of forecasting, some of which are further discussed in Chapter 10, Section 10.5.1.5. Overall, raw ensemble forecasts are systematically refined by postprocessing techniques (also called calibration techniques) to further improve their quality.

The aim of postprocessing is to apply a statistical calibration to the PDF drawn by the raw initial ensemble forecasts to optimize a specific metric used to assess the quality of probabilistic forecasts (e.g., the continuous ranked probability score [CRPS] described in Chapter 10, Section 10.5.1.5) In addition to having a coarse spatial resolution, the ensemble forecasts from NWPs are known to be underdispersive. i.e., they exhibit a lack of spread (Leutbecher and Palmer 2008). To address this, Sperati, Alessandrini, and Delle Monache (2016) proposed two different correction methods already used in the realm of wind forecasting: the variance deficit method, designed by Buizza, Richardson, and Palmer (2003) and the ensemble MOS method proposed by Gneiting et al. (2005). Even if these methods cannot be considered parametric, they are based on the characteristics of a normal distribution. Indeed, such a distribution is appealing because it can be assessed with only two parameters: the mean and the standard deviation, which are related to the average bias and the spread of the ensemble, respectively.

Another method of calibration is based on the rank histogram (see Chapter 10, Section 10.6.3.2), which was initially proposed by Hamill and Colucci (1997) for precipitation forecasts. Zamo et

al. (2014) applied this method to the Météo-France EPS, called PEARP, to generate probabilistic solar forecasts. The aim of this method is to build a calibrated CDF from the rank histogram derived from past forecasts and observations. Other techniques of EPS calibration exist in meteorology. For example, Pinson (2012) and Pinson and Madsen (2009) suggested a framework for the calibration of wind ensemble forecasts. Junk, Delle Monache, and Alessandrini (2015) proposed an original calibration model, based on the combination of nonhomogeneous Gaussian regression and AnEn models, for wind speed forecasting applied to ECMWF-EPS predictions. Likewise, Hamill and Whitaker (2006) suggested an adaptation of the AnEn technique for the calibration of ensemble precipitation forecasts using the statistical moments of the distribution, such as the mean and spread of the members as predictors. See Wilks (2019) for a thorough review of univariate ensemble postprocessing methods.

9.7.4 Benchmark Probabilistic Models

This section describes benchmark or baseline probabilistic models used to gauge the performance of new proposed probabilistic methods using skill scores, such as the continuous rank probability skill score. By analogy with the deterministic approach, persistence ensemble (PeEn) models based on GHI (Alessandrini et al. 2015) and on the clear-sky index (David et al. 2016) have been proposed. The empirical CDF of a PeEn forecast is simply built with the most recent k past measurements of solar irradiance. Considering an infinite number of past measurements, the PeEn simply becomes the climatology. In numerous other fields of meteorology, climatology is often considered to be a reference that can be used to test the performance of probabilistic models (Wilks 2019). That is because the climatology is perfectly reliable; however, it has no resolution.

Climatological reference models for probabilistic solar irradiance forecasting should account for the deterministic course of solar irradiance. The complete-history persistence ensemble proposed by Yang (2019) corresponds to a conditional climatology where the time of day is used as a predictor. The so-called clear-sky-dependent climatology (CSD-Clim) (Le Gal La Salle, David, and Lauret 2021)) is based on a similar approach but using the clear-sky irradiance as predictor instead of the time of day. Another simple approach consists of deriving the distribution of the clear-sky index from a long-term dataset and deriving the irradiance distribution by multiplication with clear-sky irradiances (Nouri et al. 2023). A comparison of these baseline models based on clear-sky index distributions derived from a long-term dataset of GHI and DNI in Almeria (Spain) by Nouri et al. (2023) shows only minor differences between the different approaches with a slight advantage for CSD-Clim. Because all three benchmark models consider the current conditions of the sun's position and atmospheric turbidity, the influence of further discretization over the time of day or clear-sky irradiance is small. Finally, for ensemble forecasts, the CRPS of the raw ensemble can serve as a benchmark.

9.8 Summary and Recommendations for Irradiance Forecasting

Solar power forecasting is essential for the reliable and cost-effective system integration of solar energy. It is used for a variety of applications with specific requirements with respect to forecast horizon and spatiotemporal resolution. To meet these needs, different solar irradiance and power forecasting methods have been developed, including physical and empirical models, as well as statistical and ML approaches. Based on these developments, forecasting services of good quality are now available for users.

An overview of the basic characteristics of different forecasting approaches is given in Table 9-2: (1) time-series models based on local measurements, (2) ASI and satellite-based forecasting, and (3) regional and global NWP models. These characteristics include forecast horizon, update frequency, temporal resolution, spatial resolution, and coverage, complemented with some practical information. Beyond using one of these forecasting models, blending of different forecasts using statistical and ML models (Section 9.5) is frequently applied to enhance forecast accuracy.

The different forecasting approaches with their corresponding input data and their applicability for different requirements are summarized in the following, with the spatiotemporal resolution of irradiance forecasts generally decreasing with increasing forecast horizon:

- Short-term forecasting up to approximately 1 hour ahead greatly benefit from the use of local online irradiance or PV power measurements as input; however, pure time-series approaches, based on local measurements only, are outperformed by approaches integrating empirical and/or physical model forecasts from a few minutes to hours onward, depending on the spatiotemporal scale of the forecasts and the climatic conditions of the forecast location.
- Short-term irradiance forecasts up to 20 minutes ahead that resolve irradiance ramps with a temporal resolution of minutes or even less can be derived from ASIs using cloud motion and/or ML-based methodologies. Using information on the local cloudiness around the site under scrutiny, state-of-the-art ASI models outperform persistence based on single-site measurements for high-resolution intrahour forecasting. Local hardware typically consists of a radiometric station, one ASI for point forecasting, and at least two ASIs for spatially resolved irradiance fields.
- Irradiance forecasts up to several hours ahead with typical resolutions from 10–15 minutes are derived from geostationary satellite data covering large areas without requiring local hardware. Satellite-based forecasting models are typically based on cloud motion approaches and increasingly also involve ML techniques. State-of-the-art satellite-based forecasts outperform persistence from approximately 30 minutes onward, and NWP forecasts up to several hours ahead.
- Irradiance forecasts from several hours to days ahead essentially rely on NWP models, with their capability to describe complex atmospheric dynamics, including advection as well as the formation and dissipation of clouds. Typically, NWP forecasts are provided with hourly resolution in the first days; such forecasts cover countries or continents with regional models and the entire Earth with global models.

Table 9-2. Basic Characteristics of Different Forecasting Approaches

	Time-Series Using Local Measurements	All-Sky Imager Based	Satellite Based	NWP Global	NWP Regional
Forecast horizon	Intrahour and intraday: 15 minutes–2 hours	Intrahour: 10–20 minutes	intraday: 4–8 hours	Days ahead: 10–15 days	Days ahead: 2–3 days
Update frequency	1 second to 1 hour	~10 seconds to 1 minute	10–15 minutes	6–24 hours	3–6 hours
Temporal resolution	1 second to 1 hour	~10 seconds to 1 minute	10–15 minutes	1–12 hours	1 hour
Spatial resolution	Point	10–100 m	0.5–3 km at subsatellite point	9–20 km	1–10 km
Coverage	Point	Point to 12 km	Satellite field of view (full Earth disc, continents)	Global	Countries, continents
Local hardware	Radiometric station or meters for PV power measurements	ASI(s) and radiometric station	None	None	None
Forecast providers	Private companies	Private companies	Private companies, some national weather services	National and international weather services	National and private weather services
Comments	Forecast horizon is linked to temporal resolution	Spatially extended forecasts require cloud height measure-ments or multiple ASIs; coverage and forecast horizon depend on cloud conditions	Spatial/ temporal resolutions depend on satellite/ spectral channels; spatial resolution decreases with distance from subsatellite point	Update frequency and temporal resolution might decrease with increasing forecast horizon	

NWP model forecasts are typically provided by international and national weather services, with a list of global NWP models and their providers given in Table 9-1. Regional NWP model

forecasts are provided by numerous national and private weather services. In addition, Table 9-3 gives information on some companies providing ASI-based and/or satellite-based forecasts. It is emphasized that this list is based on the experience of the IEA PVPS Task 16 participants and does not pretend to be exhaustive.

Table 9-3. Examples of Companies Providing Irradiance Forecasts based on ASI or satellite data

Company	Website	ASI	Satellite
CSPServices	https://www.cspservices.de/meteorological-services/	x	
Flucrum3D	https://www.fulcrum3d.com/cloudcam/	x	
meteo for energy	https://www.meteoforenergy.com/en/	x	x
Meteotest	https://solarwebservices.ch/		x
Reuniwatt	https://reuniwatt.com/en/247-all-sky-observation-sky-insight/	x	x
SoDa	https://www.soda-pro.com/		x
SolarAnywhere	https://www.solaranywhere.com/products/solaranywhere-forecast/		x
Solargis	https://solargis.com/products/solar-power-forecast/overview		x
Solcast	https://solcast.com/forecast-accuracy		x
Steadysun	https://www.steady-sun.com/solar-energy-forecasting/	x	x

As a complement to empirical and physical models, statistical and ML methods are widely used in solar irradiance and power forecasting. They also exploit the rapid development in AI techniques along with ever-increasing computational resources:

- Statistical and ML approaches are effectively applied to improve forecasts obtained with empirical or physical models (postprocessing). Through training against high-quality irradiance measurements (ground truth), they can reduce systematic meteorological forecast errors.
- Training of statistical and ML approaches to PV power measurements additionally provides a way to derive plant-specific models that account for the characteristics of a given PV plant or even to replace PV simulation models.
- With the fast advances in computer vision, ML techniques are now successfully applied to the prediction of cloud images (from ASIs or satellites), including algorithms to compute the optical flow in cloud motion approaches. From the predicted cloud image irradiance, forecasts can be derived in a subsequent step.
- Direct ML forecasting approaches can combine all kinds of possible observations (e.g., images from different sources or meteorological measurements of sensor networks) to predict solar irradiance or PV power.

For the training of time-series models, the availability of irradiance and/or PV power measurements is crucial. Great care is essential in the selection and quality control of this ground truth data, depending on the intended model usage. ML models adapt to ground truth as is, including, e.g., potential measurement faults, degradation, and soiling. Depending on the application, this might constitute either a problem (e.g., the irradiance forecasts might be biased if the ground truth originates from soiled sensors) or an advantage (e.g., when predicting the

current status of a PV plant for marketing the produced power, in the presence of soiling and/or degradation). Concerning the postprocessing of NWP model forecasts, satellite-derived data also provide a suitable reference for training.

Overall, the best possible accuracy of irradiance or PV power forecasts can be achieved by integrating different input data and methodologies. Prominent examples include:

- High-resolution intrahour forecasting systems combine local measurements and ASI data with empirical and ML approaches, and possibly also integrate satellite-based forecasts.
- Forecasting systems for the intraday energy market up to several hours ahead integrate online measurements and/or satellite-based forecasts as well as NWP -based forecasts with statistical and/or ML approaches.
- Forecasting systems from several hours to several days ahead use different NWP models as input in combination with statistical and/or ML approaches.

The performance of the different forecast models depends on multiple factors that have different impacts depending on forecast horizon and meteorological conditions:

- The capability of the models to predict changes in clouds and irradiance—for instance, persistence cannot predict approaching clouds.
- The performance of the models for irradiance retrieval/analysis for a forecast lead time of zero—persistence is then error-less, and the satellite-derived irradiance has a lower uncertainty than NWP analysis.
- The model's input data and parameters as well as the area covered by the input data—for instance, the larger the monitored cloud scene, the larger the forecast horizon can be.
- The computer time to execute a model run—the faster a model run, the less time that has passed since the observations fed into the model have been taken at the time of forecast delivery.
- The spatiotemporal resolution of the forecasts
- The capability of the model to correctly predict the AOD, especially for DNI forecasting in arid regions.

Besides forecasting for single PV power plants, the estimation and forecasting of regionally aggregated PV power is important for grid operators. Here, an additional challenge is that the information on all the PV power plants contributing to the overall regional feed-in power is often incomplete. Moreover, for most plants in many countries, the PV power is not monitored at a sufficient temporal resolution. Therefore, regional models, such as upscaling, have been developed and are effectively applied to derive and forecast regionally aggregated PV power. Because of spatial smoothing effects, the forecast errors of regionally aggregated PV power (normalized to the installed power) are much smaller than for single PV plants, depending on the size of the region and the set of contributing PV plants.

Forecast evaluations provide users with the necessary information on forecast accuracy and assist them in selecting between different forecasting services or assessing the risk when a forecast is used as a basis for decisions. In this chapter, different forecasts are compared using RMSE values as a basic score. A more detailed discussion and recommendations for the evaluation and uncertainty assessment of irradiance forecasting are given in Chapter 10.

Compared to deterministic forecasts, probabilistic forecasts have the great advantage to also add specific uncertainty information for each forecast value, depending on weather conditions. Probabilistic forecasts take the form of CDFs or PDFs. They are summarized by quantiles from which prediction intervals can be inferred. Quantiles can be estimated using either a parametric or a nonparametric approach. In the latter case, statistical or ML techniques can be used to estimate the quantiles. Although NWP ensemble members are not directly linked to the notion of quantiles, different propositions exist to infer a CDF from an ensemble.

Finally, forecasting solar power should be evaluated in the context of the system integration of solar power, where elaborate strategies are needed to provide the necessary power to meet the demand at any instant. These strategies include spatial smoothing for grid-integrated PV complemented by wind power and increasingly also the use of storage (batteries), curtailment, and shifting of loads to times with abundant PV generation. Applying these strategies reduces the variability of solar power as well as forecast errors.

References

Aguiar, L. M., B. Pereira, P. Lauret, F. Díaz, and M. David. 2016. "Combining Solar Irradiance Measurements, Satellite-Derived Data and a Numerical Weather Prediction Model to Improve Intra-Day Solar Forecasting." *Renewable Energy* 97: 599–610. https://doi.org/10.1016/j.renene.2016.06.018.

Ahn, H., J. Yu, and J.-M. Yeom. 2022. "Deep Learning-Based Prediction of Solar Surface Irradiance With Geostationary Satellite Images." Presented at the 2022 IEEE 17th Annual System of Systems Engineering Conference (SOSE), Rochester, NY, 311–315. https://doi.org/10.1109/SOSE55472.2022.9812657.

Aicardi, D., P. Musé, and R. Alonso-Suárez. 2022. "A Comparison of Satellite Cloud Motion Vectors Techniques to Forecast Intra-Day Hourly Solar Global Horizontal Irradiation." *Solar Energy* 233: 46–60. https://doi.org/10.1016/j.solener.2021.12.066.

Alessandrini, S., L. Delle Monache, S. Sperati, and G. Cervone. 2015. "An Analog Ensemble for Short-Term Probabilistic Solar Power Forecast." *Applied Energy* 157: 95–110. https://doi.org/10.1016/j.apenergy.2015.08.011.

Alonso-Montesinos, J., and F. J. Batlles. 2015. "Solar Radiation Forecasting in the Short- and Medium-Term Under All Sky Conditions. *Energy* 83: 387–393. https://doi.org/10.1016/j.energy.2015.02.036.

Amaro e Silva, R., and M. C. Brito. 2019. "Spatio-Temporal PV Forecasting Sensitivity to Modules' Tilt and Orientation." *Applied Energy* 255, 113807. https://doi.org/10.1016/j.apenergy.2019.113807.

Amaro e Silva, R., and M. C. Brito. 2017. "Understanding Spatio-Temporal Solar Forecasting."

André, M., R. Perez, T. Soubdhan, J. Schlemmer, R. Calif, and S. Monjoly. 2019. "Preliminary Assessment of Two Spatio-Temporal Forecasting Technics for Hourly Satellite-Derived Irradiance in a Complex Meteorological Context." *Solar Energy* 177: 703–712. https://doi.org/10.1016/j.solener.2018.11.010.

Andrews, R. W., J. H. Stein, C. Hansen, and D. Riley. 2014. "Introduction to the Open Source PV LIB for Python Photovoltaic System Modelling Package." Presented at the 2014 IEEE 4^{0t}h Photovoltaic Specialists Conference (PVSC), Denver, CO, USA, 0170–0174. https://doi.org/10.1109/PVSC.2014.6925501.

Antonanzas, J., D. Pozo-Vázquez, L. A. Fernandez-Jimenez, and F. J. Martinez-de-Pison. 2017. "The Value of Day-Ahead Forecasting for Photovoltaics in the Spanish Electricity Market." *Solar Energy* 158: 140–146. https://doi.org/10.1016/j.solener.2017.09.043.

Antonanzas, J., N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de-Pison, and F. Antonanzas-Torres. 2016. "Review of Photovoltaic Power Forecasting." *Solar Energy* 136: 78–111. https://doi.org/10.1016/j.solener.2016.06.069.

Arbizu-Barrena, C., D. Pozo-Vázquez, J. A. Ruiz-Arias, and J. Tovar-Pescador. 2015. "Macroscopic Cloud Properties in the WRF NWP Model: An Assessment Using Sky Camera and Ceilometer Data." *JGR Atmospheres* 120. https://doi.org/10.1002/2015JD023502.

Aryaputera, A. W., D. Yang, and W. M. Walsh. 2015. "Day-Ahead Solar Irradiance Forecasting in a Tropical Environment." *Journal of Solar Energy Engineering* 137, 051009. https://doi.org/10.1115/1.4030231.

Bacher, P., H. Madsen, and H. A. Nielsen. 2009. "Online Short-Term Solar Power Forecasting." *Solar Energy* 83: 1772–1783. https://doi.org/10.1016/j.solener.2009.05.016.

Bakker, K., K. Whan, W. Knap, and M. Schmeits. 2019. "Comparison of Statistical Post-Processing Methods for Probabilistic NWP Forecasts of Solar Radiation." *Solar Energy* 191: 138–150. https://doi.org/10.1016/j.solener.2019.08.044.

Bauer, P., A. Thorpe, and G. Brunet. 2015. "The Quiet Revolution of Numerical Weather Prediction." *Nature* 525: 47–55. https://doi.org/10.1038/nature14956.

Benavides Cesar, L., R. Amaro E Silva, M. Á. Manso Callejo, and C.-I. Cira. 2022. "Review on Spatio-Temporal Solar Forecasting Methods Driven by In Situ Measurements or Their Combination With Satellite and Numerical Weather Prediction (NWP) Estimates." *Energies* 15, 4341. https://doi.org/10.3390/en15124341.

Berthomier, L., B. Pradel, and L. Perez. 2020. "Cloud Cover Nowcasting With Deep Learning." Presented at the 2020 IEEE 10th International Conference on Image Processing Theory, Tools, and Applications (IPTA), Paris, France, 1–6. https://doi.org/10.1109/IPTA50016.2020.9286606.

Bessa, R., C. Moreira, B. Silva, and M. Matos. 2014. "Handling Renewable Energy Variability and Uncertainty in Power Systems Operation." *WIREs Energy & Environment* 3: 156–178. https://doi.org/10.1002/wene.76.

Beyer, H.G., A. Drews Betcke Jethro, D. Heinemann, E. Lorenz, G. Heilscher, and S. Bofinger. 2004. "Identification of a General Model for the MPP Performance of PV-Modules for the Application in a Procedure for the Performance Check of Grid Connected Systems."

Bishop, C. M. 1995. *Neural Networks for Pattern Recognition*. Oxford: Clarendon Press, Oxford University Press.

Blanc, P., P. Massip, A. Kazantzidis, P. Tzoumanikas, P. Kuhn, S. Wilbert, D. Schüler, and C. Prahl. 2017. "Short-Term Forecasting of High Resolution Local DNI Maps With Multiple Fish-Eye Cameras in Stereoscopic Mode." Presented at the SOLARPACES 2016: International Conference on Concentrating Solar Power and Chemical Energy Systems, Abu Dhabi, United Arab Emirates, 140004. https://doi.org/10.1063/1.4984512.

Blum, N. B., S. Wilbert, B. Nouri, J. Stührenberg, J. E. Lezaca Galeano, T. Schmidt, D. Heinemann, T. Vogt, A. Kazantzidis, and R. Pitz-Paal. 2022. "Analyzing Spatial Variations of Cloud Attenuation by a Network of All-Sky Imagers." *Remote Sensing* 14, 5685. https://doi.org/10.3390/rs14225685.

Bonavita, M., and P. Lean. 2021. "4D-VAR for Numerical Weather Prediction." *Weather* 76: 65–66. https://doi.org/10.1002/wea.3862.

Bosch, J. L., and J. Kleissl. 2013. "Cloud Motion Vectors From a Network of Ground Sensors in a Solar Power Plant. *Solar Energy* 95, 13–20. https://doi.org/10.1016/j.solener.2013.05.027.

Bottou, L. 1998. "Online Learning and Stochastic Approximations."

Brancucci Martinez-Anido, C., B. Botor, A. R. Florita, C. Draxl, S. Lu, H. F. Hamann, and B.-M. Hodge. 2016. "The Value of Day-Ahead Solar Power Forecasting Improvement." *Solar Energy* 129: 192–203. https://doi.org/10.1016/j.solener.2016.01.049.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45: 5–32. https://doi.org/10.1023/A:1010933404324.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 2017. *Classification and Regression Trees*, 1st ed. New York: Routledge. https://doi.org/10.1201/9781315139470.

Breitkreuz, H., M. Schroedter-Homscheidt, T. Holzer-Popp, and S. Dech. 2009. "Short-Range Direct and Diffuse Irradiance Forecasts for Solar Energy Applications Based on Aerosol Chemical Transport and Numerical Weather Modeling." *Journal of Applied Meteorology and Climatology* 48: 1766–1779. https://doi.org/10.1175/2009JAMC2090.1.

Bresky, W., and J. Daniels. 2006. "The Feasibility of an Optical Power Flow Algorithm for Estimating Atmospheric Motion."

Bright, J. M., S. Killinger, D. Lingfors, and N. A. Engerer. 2018. "Improved Satellite-Derived PV Power Nowcasting Using Real-Time Power Data From Reference PV Systems." *Solar Energy* 168: 118–139. https://doi.org/10.1016/j.solener.2017.10.091.

Buizza, R., D. S. Richardson, and T. N. Palmer. 2003. "Benefits of Increased Resolution in the ECMWF Ensemble System and Comparison With Poor-Man's Ensembles." *Quarterly Journal of the Royal Meteorological Society* 129: 1269–1288. https://doi.org/10.1256/qj.02.92.

Carrière, T., R. Amaro E. Silva, F. Zhuang, Y.-M. Saint-Drenan, and P. Blanc. 2021. "A New Approach for Satellite-Based Probabilistic Solar Forecasting With Cloud Motion Vectors." *Energies* 14, 4951. https://doi.org/10.3390/en14164951.

Chauvin, R., J. Nou, J. Eynard, S. Thil, and S. Grieu. 2018. "A New Approach to the Real-Time Assessment and Intraday Forecasting of Clear-Sky Direct Normal Irradiance." *Solar Energy* 167: 35–51. https://doi.org/10.1016/j.solener.2018.02.027.

Chauvin, R., J. Nou, S. Thil, and S. Grieu. 2015. "Modelling the Clear-Sky Intensity Distribution Using a Sky Imager." *Solar Energy* 119: 1–17. https://doi.org/10.1016/j.solener.2015.06.026.

Chauvin, R., J. Nou, S. Thil, and S. Grieu. 2016. "Cloud Motion Estimation Using a Sky Imager." Presented at the SOLARPACES 2015: International Conference on Concentrating

Solar Power and Chemical Energy Systems, Cape Town, South Africa, 150003. https://doi.org/10.1063/1.4949235.

Chauvin, R., J. Nou, S. Thil, and S. Grieu. 2017. "Generating High Dynamic Range Images Using a Sky Imager." *IFAC-PapersOnLine* 50: 219–224. https://doi.org/10.1016/j.ifacol.2017.08.037.

Chow, C. W., B. Urquhart, M. Lave, A. Dominguez, J. Kleissl, J. Shields, and B. Washom. 2011. "Intra-Hour Forecasting With a Total Sky Imager at the UC San Diego Solar Energy Testbed." *Solar Energy* 85: 2,881–2,893. https://doi.org/10.1016/j.solener.2011.08.025.

Chu, Y., M. Li, H. T. C. Pedro, and C. F. M. Coimbra. 2015. "Real-Time Prediction Intervals for Intra-Hour DNI Forecasts." *Renewable Energy* 83: 234–244. https://doi.org/10.1016/j.renene.2015.04.022.

Cirés, E., J. Marcos, I. De La Parra, M. García, and L. Marroyo. 2019. "The Potential of Forecasting in Reducing the LCOE in PV Plants Under Ramp-Rate Restrictions." *Energy* 188, 116053. https://doi.org/10.1016/j.energy.2019.116053.

Coimbra, C. F. M., and H. T. C. Pedro. 2013. "Stochastic-Learning Methods." *Solar Energy Forecasting and Resource Assessment*. Elsevier. https://doi.org/10.1016/B978-0-12-397177-7.00015-2.

Cros, S., E. Buessler, L. Huet, N. Sébastien, and N. Schmutz. 2015. "The Benefits of Intraday Solar Irradiance Forecasting to Adjust the Day-Ahead Scheduled PV Power."

Cros, S., J. Badosa, A. Szantaï, and M. Haeffelin. 2020. "Reliability Predictors for Solar Irradiance Satellite-Based Forecast." *Energies* 13, 5566. https://doi.org/10.3390/en13215566.

Cros, S., M. Deroubaix, and N. Schmutz. 2015. "Method and Device for Forecasting Cloudiness by Statistical Processing of Data Selected by Spatial Analysis."

Cros, S., N. Sébastien, O. Liandrat, and N. Schmutz. 2014. "Cloud Pattern Prediction From Geostationary Meteorological Satellite Images for Solar Energy Forecasting." Edited by A. Comerón, E. I. Kassianov, K. Schäfer, R. H. Picard, K. Stein, and J. D. Gonglewski. Presented at the SPIE Remote Sensing, Amsterdam, Netherlands, 924202. https://doi.org/10.1117/12.2066853.

da Silva Fonseca Junior, J. G., T. Oozeki, H. Ohtake, K. Shimose, T. Takashima, and K. Ogimoto. 2014. "Regional Forecasts and Smoothing Effect of Photovoltaic Power Generation in Japan: An Approach With Principal Component Analysis." *Renewable Energy* 68: 403–413. https://doi.org/10.1016/j.renene.2014.02.

Dambreville, R., P. Blanc, J. Chanussot, and D. Boldo. 2014. "Very Short-Term Forecasting of the Global Horizontal Irradiance Using a Spatio-Temporal Autoregressive Model." *Renewable Energy* 72: 291–300. https://doi.org/10.1016/j.renene.2014.07.012.

Das, U. K., K. S. Tey, M. Seyedmahmoudian, S. Mekhilef, M. Y. I. Idris, W. Van Deventer, B. Horan, and A. Stojcevski. 2018. "Forecasting of Photovoltaic Power Generation and Model Optimization: A Review." *Renewable and Sustainable Energy Reviews* 81: 912–928. https://doi.org/10.1016/j.rser.2017.08.017.

David, M., and P. Lauret. 2018. "Solar Radiation Probabilistic Forecasting." In *Wind Field and Solar Radiation Characterization and Forecasting*. Edited by R. Perez. Springer International Publishing. https://doi.org/10.1007/978-3-319-76876-2_9.

David, M., F. Ramahatana, P. J. Trombe, and P. Lauret. 2016. "Probabilistic Forecasting of the Solar Irradiance With Recursive ARMA and GARCH Models." *Solar Energy* 133: 55–72. https://doi.org/10.1016/j.solener.2016.03.064.

David, M., J. Boland, L. Cirocco, P. Lauret, and C. Voyant. 2021. "Value of Deterministic Day-Ahead Forecasts of PV Generation in PV + Storage Operation for the Australian Electricity Market." *Solar Energy* 224: 672–684. https://doi.org/10.1016/j.solener.2021.06.011.

David, M., M. A. Luis, and P. Lauret. 2018. "Comparison of Intraday Probabilistic Forecasting of Solar Irradiance Using Only Endogenous Data." *International Journal of Forecasting* 34: 529–547. https://doi.org/10.1016/j.ijforecast.2018.02.003.

Delle Monache, L., F. A. Eckel, D. L. Rife, B. Nagarajan, and K. Searight. 2013. "Probabilistic Weather Prediction With an Analog Ensemble." *Monthly Weather Review* 141: 3,498–3,516. https://doi.org/10.1175/MWR-D-12-00281.1.

Deng, A., B. Gaudet, J. Dudhia, and K. Alapaty. 2014. "Implementation and Evaluation of a New Shallow Convection Scheme."

Descombes, G., D. Auligne, H.-C. Lin, D. Xu, S. Schwartz, and F. Vandenberghe. 2014. "Multi-Sensor Advection Diffusion nowCast (MADCast) for Cloud Analysis and Short-Term Prediction." https://doi.org/10.5065/D62V2D37.

Dev, S., F. M. Savoy, Y. H. Lee, and S. Winkler. 2019. "Estimating Solar Irradiance Using Sky Imagers." *Atmospheric Measurement Techniques* 12: 5,417–5,429. https://doi.org/10.5194/amt-12-5417-2019.

Diagne, M., M. David, J. Boland, N. Schmutz, and P. Lauret. 2014. "Post-Processing of Solar Irradiance Forecasts From WRF Model at Reunion Island." *Solar Energy* 105: 99–108. https://doi.org/10.1016/j.solener.2014.03.016.

Diagne, M., M. David, P. Lauret, J. Boland, and N. Schmutz. 2013. "Review of Solar Irradiance Forecasting Methods and a Proposition for Small-Scale Insular Grids." *Renewable and Sustainable Energy Reviews* 27: 65–76. https://doi.org/10.1016/j.rser.2013.06.042.

Dittmann, A., N. Holland, and E. Lorenz. 2021. "A New Sky Imager Based Global Irradiance Forecasting Model With Analyses of Cirrus Situations." *metz* 30: 101–113. https://doi.org/10.1127/metz/2020/1024.

Efron, B. 1979. "Bootstrap Methods: Another Look at the Jackknife." *Annals of Statistics* 7. https://doi.org/10.1214/aos/1176344552.

Erdener, B. C., C. Feng, K. Doubleday, A. Florita, and B.-M. Hodge. 2022. "A Review of Behind-the-Meter Solar Forecasting." *Renewable and Sustainable Energy Reviews* 160: 112224. https://doi.org/10.1016/j.rser.2022.112224.

Fabel, Y., B. Nouri, S. Wilbert, N. Blum, D. Schnaus, R. Triebel, L. F. Zarzalejo, E. Ugedo, J. Kowalski, and R. Pitz-Paal. 2023. "Combining Deep Learning and Physical Models: A Benchmark Study on All-Sky Imager-Based Solar Nowcasting Systems." *Solar RRL* solr.202300808. https://doi.org/10.1002/solr.202300808.

Fabel, Y., B. Nouri, S. Wilbert, N. Blum, R. Triebel, M. Hasenbalg, P. Kuhn, L. F. Zarzalejo, and R. Pitz-Paal. 2022. "Applying Self-Supervised Learning for Semantic Cloud Segmentation of All-Sky Images." Atmospheric Measurement Techniques 15: 797–809. https://doi.org/10.5194/amt-15-797-2022.

Farnebäck, G. 2003. "Two-Frame Motion Estimation Based on Polynomial Expansion." In *Image Analysis, Lecture Notes in Computer Science*. Edited by J. Bigun and T. Gustavsson. Berlin, Heidelberg: Springer. https://doi.org/10.1007/3-540-45103-X_50.

Fatemi, S. A., A. Kuh, and M. Fripp. 2018. "Parametric Methods for Probabilistic Forecasting of Solar Irradiance." *Renewable Energy* 129: 666–676. https://doi.org/10.1016/j.renene.2018.06.022.

Friedman, J. H. 2001. "Greedy Function Approximation: A Gradient Boosting Machine." Annals of Statistics 29. https://doi.org/10.1214/aos/1013203451.

Gallo, R., M. Castangia, A. Macii, E. Macii, E. Patti, and A. Aliberti. 2022. "Solar Radiation Forecasting With Deep Learning Techniques Integrating Geostationary Satellite Images." *Engineering Applications of Artificial Intelligence* 116, 105493. https://doi.org/10.1016/j.engappai.2022.105493.

Gallucci, D., F. Romano, A. Cersosimo, D. Cimini, F. Di Paola, S. Gentile, E. Geraldi, S. Larosa, S. Nilo, E. Ricciardelli, and M. Viggiano. 2018. "Nowcasting Surface Solar Irradiance With AMESIS via Motion Vector Fields of MSG-SEVIRI Data." *Remote Sensing* 10, 845. https://doi.org/10.3390/rs10060845.

Gastón, M., E. Lorenz, S. Lozano, D. Heinemann, M. Blanco, and L. Santigosa. 2009. "Comparison of Global Irradiance Forecasting Approaches."

Gauchet, C., Blanc, P., Espinar, B., Charbonnier, B., Demengel, D. 2012. Surface solar irradiance estimation With low-cost fish-eye camera. Presented at the Workshop on "Remote Sensing Measurements for Renewable Energy."

Gensler, A., J. Henze, B. Sick, and N. Raabe. 2016. "Deep Learning for Solar Power Forecasting—An Approach Using AutoEncoder and LSTM Neural Networks." Presented at the

2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Budapest, Hungary, 002858–002865. https://doi.org/10.1109/SMC.2016.7844673.

Ghonima, M. S., B. Urquhart, C. W. Chow, J. E. Shields, A. Cazorla, and J. Kleissl. 2012. "A Method for Cloud Detection and Opacity Classification Based on Ground-Based Sky Imagery." *Atmospheric Measurement Techniques* 5: 2,881–2,892. https://doi.org/10.5194/amt-5-2881-2012.

Gigoni, L., A. Betti, E. Crisostomi, A. Franco, M. Tucci, F. Bizzarri, and D. Mucci. 2018. "Day-Ahead Hourly Forecasting of Power Generation From Photovoltaic Plants." *IEEE Transactions on Sustainable Energy* 9: 831–842. https://doi.org/10.1109/TSTE.2017.2762435.

Glahn, H. R., and D. A. Lowry. 1972. "The Use of Model Output Statistics (MOS) in Objective Weather Forecasting." Journal of Applied Meteorology and Climatology 11: 1,203–1,211. https://doi.org/10.1175/1520-0450(1972)011<1203:TUOMOS>2.0.CO;2.

Glorot, X., and Y. Bengio. 2010. "Understanding the Difficulty of Training Deep Feedforward Neural Networks." Presented at the International Conference on Artificial Intelligence and Statistics.

Gneiting, T., A. E. Raftery, A. H. Westveld, and T. Goldman. 2005. "Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation." *Monthly Weather Review* 133: 1098–1118. https://doi.org/10.1175/MWR2904.1.

Grantham, A., Y. R. Gel, and J. Boland. 2016. "Nonparametric Short-Term Probabilistic Forecasting for Solar Radiation." *Solar Energy* 133: 465–475. https://doi.org/10.1016/j.solener.2016.04.011.

Gueymard, C., and P. Jiménez. 2018. "Validation of Real-Time Solar Irradiance Simulations Over Kuwait Using WRF-Solar." Presented at the ISES EuroSun 2018 Conference—12th International Conference on Solar Energy for Buildings and Industry, International Solar Energy Society, Rapperswil, 1–11. https://doi.org/10.18086/eurosun2018.09.14.

Gueymard, C., W. Gustafson, A. Etringer, and P. Storck. 2012. "Evaluation of Procedures to Improve Solar Resource Assessments: Optimum Use of Short-Term Data From a Local Weather Station to Correct Bias in Long-Term Satellite-Derived Solar Radiation Time Series."

Hamill, T. M., and J. S. Whitaker. 2006. "Probabilistic Quantitative Precipitation Forecasts Based on Reforecast Analogs: Theory and Application." *Monthly Weather Review* 134: 3,209–3,229. https://doi.org/10.1175/MWR3237.1.

Hamill, T. M., and S. J. Colucci. 1997. "Verification of Eta–RSM Short-Range Ensemble Forecasts." *Monthly Weather Review* 125: 1312–1327. https://doi.org/10.1175/1520-0493(1997)125<1312:VOERSR>2.0.CO;2.

Hammer, A., D. Heinemann, C. Hoyer, R. Kuhlemann, E. Lorenz, R. Müller, and H. G. Beyer. 2003. "Solar Energy Assessment Using Remote Sensing Technologies." *Remote Sensing of Environment* 86: 423–432. https://doi.org/10.1016/S0034-4257(03)00083-X.

Hasenbalg, M., P. Kuhn, S. Wilbert, B. Nouri, and A. Kazantzidis. 2020. "Benchmarking of Six Cloud Segmentation Algorithms for Ground-Based All-Sky Imagers." *Solar Energy* 201, 596–614. https://doi.org/10.1016/j.solener.2020.02.042.

Haupt, S. E., B. Kosović, T. Jensen, J. K. Lazo, J. A. Lee, P. A. Jiménez, J. Cowie, G. Wiener, T. C. McCandless, M. Rogers, S. Miller, M. Sengupta, Y. Xie, L. Hinkelman, P. Kalb, and J. Heiser. 2018. "Building the Sun4Cast System: Improvements in Solar Power Forecasting." *Bulletin of the American Meteorological Society* 99: 121–136. https://doi.org/10.1175/BAMS-D-16-0221.1.

Haupt, S. E., B. Kosovic, T. L. Jensen, J. Lee, P. Jiménez Munoz, K. Lazo, R. Cowie, T. McCandless, M. Pearson, M. Wiener, S. Alessandrini, L. Delle Monache, D. Yu, Z. Peng, D. Huang, J. Heiser, S. Yoo, P. Kalb, S. Miller, M. Rogers, and L. Hinkleman. 2016. *The Sun4Cast® Solar Power Forecasting System: The Result of the Public-Private-Academic Partnership to Advance Solar Power Forecasting*. https://doi.org/10.5065/D6N58JR2.

He, X., C. Yuan, and Z. Yang. 2016. "Performance Evaluation of Chinese Solar Radiation Forecast Based on Three Global Forecast Back Ground Fields."

Hoff, T. E., and R. Perez. 2012. "Modeling PV Fleet Output Variability." *Solar Energy: Progress in Solar Energy* 3 86: 2,177–2,189. https://doi.org/10.1016/j.solener.2011.11.005.

Huang, G.-B., Q.-Y. Zhu, and C.-K. Siew. 2006. "Extreme Learning Machine: Theory and Applications." *Neurocomputing* 70: 489–501. https://doi.org/10.1016/j.neucom.2005.12.126.

Huang, J., M. Korolkiewicz, M. Agrawal, and J. Boland. 2013. "Forecasting Solar Radiation on an Hourly Timescale Using a Coupled AutoRegressive and Dynamical System (CARDS) Model." *Solar Energy* 87: 136–149. https://doi.org/10.1016/j.solener.2012.10.012.

Huang, J., M. M. Khan, Y. Qin, and S. West. 2019. "Hybrid Intra-hour Solar PV Power Forecasting Using Statistical and Skycam-Based Methods." Presented at the 2019 IEEE 4^{6t}h Photovoltaic Specialists Conference (PVSC), Chicago, IL, 2,434–2,439. https://doi.org/10.1109/PVSC40753.2019.8980732.

Inman, R. H., H. T. C. Pedro, and C. F. M. Coimbra. 2013. "Solar Forecasting Methods for Renewable Energy Integration." *Progress in Energy and Combustion Science* 39: 535–576. https://doi.org/10.1016/j.pecs.2013.06.002.

Isvoranu, D., and V. Badescu. 2013. "Comparison Between Measurements and WRF Numerical Simulation of Global Solar Irradiation in Romania." *Annals of West University of Timisoara - Physics* 57: 24–33. https://doi.org/10.1515/awutp-2015-0103.

Jamal, T., C. Carter, T. Schmidt, G. M. Shafiullah, M. Calais, and T. Urmee. 2019. "An Energy Flow Simulation Tool for Incorporating Short-Term PV Forecasting in a Diesel-PV-Battery Off-Grid Power Supply System." *Applied Energy* 254, 113718. https://doi.org/10.1016/j.apenergy.2019.113718.

Jamaly, M., and J. Kleissl. 2017. "Spatiotemporal Interpolation and Forecast of Irradiance Data Using Kriging." *Solar Energy* 158: 407–423. https://doi.org/10.1016/j.solener.2017.09.057.

Jiang, H., N. Lu, J. Qin, W. Tang, and L. Yao. 2019. "A Deep Learning Algorithm to Estimate Hourly Global Solar Radiation From Geostationary Satellite Data." *Renewable and Sustainable Energy Reviews* 114, 109327. https://doi.org/10.1016/j.rser.2019.109327.

Jiménez, P. A., J. Dudhia, G. Thompson, J. A. Lee, and T. Brummet. 2022. "Improving the Cloud Initialization in WRF-Solar With Enhanced Short-Range Forecasting Functionality: The MAD-WRF Model." *Solar Energy* 239: 221–233. https://doi.org/10.1016/j.solener.2022.04.055.

Jiménez, P. A., J. P. Hacker, J. Dudhia, S. E. Haupt, J. A. Ruiz-Arias, C. A. Gueymard, G. Thompson, T. Eidhammer, and A. Deng. 2016. "WRF-Solar: Description and Clear-Sky Assessment of an Augmented NWP Model for Solar Power Prediction." *Bulletin of the American Meteorological Society* 97: 1,249–1,264. https://doi.org/10.1175/BAMS-D-14-00279.1.

Jiménez-Garrote, A., G. Sánchez-Hernández, M. López-Cuesta, and D. Pozo-Vázquez. 2023. "SOWISP—A Retrospective High Spatial and Temporal Resolution Database of the Installed Wind and Solar PV Power in Spain." *Solar Energy* 256: 44–54. https://doi.org/10.1016/j.solener.2023.03.009.

Jones, A. S., and S. J. Fletcher. 2013. "Data Assimilation in Numerical Weather Prediction and Sample Applications." In *Solar Energy Forecasting and Resource Assessment*. Elsevier. https://doi.org/10.1016/B978-0-12-397177-7.00013-9.

Junk, C., L. Delle Monache, and S. Alessandrini. 2015. "Analog-Based Ensemble Model Output Statistics." *Monthly Weather Review* 143: 2,909–2,917. https://doi.org/10.1175/MWR-D-15-0095.1.

Kallio-Myers, V., A. Riihelä, P. Lahtinen, and A. Lindfors. 2020. "Global Horizontal Irradiance Forecast for Finland Based on Geostationary Weather Satellite Data. *Solar Energy* 198: 68–80. https://doi.org/10.1016/j.solener.2020.01.008.

Kamadinata, J. O., T. L. Ken, and T. Suwa. 2019. "Sky Image-Based Solar Irradiance Prediction Methodologies Using Artificial Neural Networks." *Renewable Energy* 134: 837–845. https://doi.org/10.1016/j.renene.2018.11.056.

Karalus, S., B. Köpfer, P. Guthke, S. Killinger, and E. Lorenz. 2023. "Analysing Grid-Level Effects of Photovoltaic Self-Consumption Using a Stochastic Bottom-Up Model of Prosumer Systems." *Energies* 16, 3059. https://doi.org/10.3390/en16073059.

Kaur, A., L. Nonnenmacher, H. T. C. Pedro, and C. F. M. Coimbra. 2016. "Benefits of Solar Forecasting for Energy Imbalance Markets." *Renewable Energy* 86: 819–830. https://doi.org/10.1016/j.renene.2015.09.011.

Kazantzidis, A., P. Tzoumanikas, E. Nikitidou, V. Salamalikis, S. Wilbert, C. Prahl. 2017. "Application of Simple All-Sky Imagers for the Estimation of Aerosol Optical Depth." Presented at the SOLARPACES 2016: International Conference on Concentrating Solar Power and

Chemical Energy Systems, Abu Dhabi, United Arab Emirates, 140012. https://doi.org/10.1063/1.4984520.

Kellerhals, S. A., F. De Leeuw, and C. Rodriguez Rivero. 2022. "Cloud Nowcasting With Structure-Preserving Convolutional Gated Recurrent Units." *Atmosphere* 13, 1632. https://doi.org/10.3390/atmos13101632.

Killinger, S., F. Braam, B. Müller, B. Wille-Haussmann, and R. McKenna. 2016. "Projection of Power Generation Between Differently-Oriented PV Systems." *Solar Energy* 136: 153–165. https://doi.org/10.1016/j.solener.2016.06.075.

Killinger, S., D. Lingfors, Y.-M. Saint-Drenan, P. Moraitis, W. Van Sark, J. Taylor, N. A. Engerer, and J. M. Bright. 2018. "On the Search for Representative Characteristics of PV Systems: Data Collection and Analysis of PV System Azimuth, Tilt, Capacity, Yield, and Shading." *Solar Energy* 173: 1087–1106. https://doi.org/10.1016/j.solener.2018.08.051.

Kingma, D. P., and J. Ba. 2014. "Adam: A Method for Stochastic Optimization." https://doi.org/10.48550/ARXIV.1412.6980.

Kleebauer, M., D. Horst, and C. Reudenbach. 2021. "Semi-Automatic Generation of Training Samples for Detecting Renewable Energy Plants in High-Resolution Aerial Images." *Remote Sensing* 13, 4793. https://doi.org/10.3390/rs13234793.

Kleissl, J. 2013. *Solar Energy Forecasting and resource Assessment*. Academic Press.

Koenker, R., and G. Bassett. 1978. "Regression Quantiles." *Econometrica* 46: 33. https://doi.org/10.2307/1913643.

Köhler, C., A. Steiner, Y.-M. Saint-Drenan, D. Ernst, A. Bergmann-Dick, M. Zirkelbach, Z. Ben Bouallègue, I. Metzinger, and B. Ritter. 2017. "Critical Weather Situations for Renewable Energies—Part B: Low Stratus Risk for Solar Power." *Renewable Energy* 101: 794–803. https://doi.org/10.1016/j.renene.2016.09.002.

Kong, W., Y. Jia, Z. Y. Dong, K. Meng, and S. Chai. 2020. "Hybrid Approaches Based on Deep Whole-Sky-Image Learning to Photovoltaic Generation Forecasting." *Applied Energy* 280, 115875. https://doi.org/10.1016/j.apenergy.2020.115875.

Kosmopoulos, P., D. Kouroutsidis, K. Papachristopoulou, P. I. Raptis, A. Masoom, Y.-M. Saint-Drenan, P. Blanc, C. Kontoes, and S. Kazadzis. 2020. "Short-Term Forecasting of Large-Scale Clouds Impact on Downwelling Surface Solar Irradiation." *Energies* 13, 6555. https://doi.org/10.3390/en13246555.

Kraas, B., M. Schroedter-Homscheidt, and R. Madlener. 2013. "Economic Merits of a State-of-the-Art Concentrating Solar Power Forecasting System for Participation in the Spanish Electricity Market." *Solar Energy* 93: 244–255. https://doi.org/10.1016/j.solener.2013.04.012.

Kriebel, K. T., G. Gesell, M. Ka¨stner, and H. Mannstein. 2003. “The Cloud Analysis Tool APOLLO: Improvements and Validations.” *International Journal of Remote Sensing* 24: 2,389–2,408. https://doi.org/10.1080/01431160210163065.

Kuhn, P., M. Wirtz, N. Killius, S. Wilbert, J. L. Bosch, N. Hanrieder, B. Nouri, J. Kleissl, L. Ramirez, M. Schroedter-Homscheidt, D. Heinemann, A. Kazantzidis, P. Blanc, and R. Pitz-Paal. 2018. “Benchmarking Three Low-Cost, Low-Maintenance Cloud Height Measurement Systems and ECMWF Cloud Heights Against a Ceilometer.” *Solar Energy* 168: 140–152. https://doi.org/10.1016/j.solener.2018.02.050.

Kuhn, P., S. Wilbert, D. Schüler, C. Prahl, T. Haase, L. Ramirez, L. Zarzalejo, A. Meyer, L. Vuilleumier, P. Blanc, J. Dubrana, A. Kazantzidis, M. Schroedter-Homscheidt, T. Hirsch, and R. Pitz-Paal. 2017. “Validation of Spatially Resolved All Sky Imager Derived DNI Nowcasts.” Presented at the SOLARPACES 2016: International Conference on Concentrating Solar Power and Chemical Energy Systems, Abu Dhabi, United Arab Emirates, 140014. https://doi.org/10.1063/1.4984522.

Kühnert, J. 2015. “Development of a Photovoltaic Power Prediction System for Forecast Horizons of Several Hours.”

Kühnert, J., E. Lorenz, and D. Heinemann. 2013. “Satellite-Based Irradiance and Power Forecasting for the German Energy Market.” In *Solar Energy Forecasting and Resource Assessment*. Elsevier. https://doi.org/10.1016/B978-0-12-397177-7.00011-5.

Kurtz, B., and J. Kleissl. 2017. “Measuring Diffuse, Direct, and Global Irradiance Using a Sky Imager.” *Solar Energy* 141: 311–322. https://doi.org/10.1016/j.solener.2016.11.032.

Lara-Fanego, V., J. A. Ruiz-Arias, D. Pozo-Vázquez, F. J. Santos-Alamillos, and J. Tovar-Pescador. 2012. “Evaluation of the WRF Model Solar Irradiance Forecasts in Andalusia (Southern Spain).” *Solar Energy* 86: 2,200–2,217. https://doi.org/10.1016/j.solener.2011.02.014.

Lauret, P., M. David, and H. Pedro. 2017. “Probabilistic Solar Forecasting Using Quantile Regression Models.” *Energies* 10, 1591. https://doi.org/10.3390/en10101591.

Lauret, P., M. David, and P. Pinson. 2019. “Verification of Solar Irradiance Probabilistic Forecasts.” *Solar Energy* 194: 254–271. https://doi.org/10.1016/j.solener.2019.10.041.

Law, E. W., A. A. Prasad, M. Kay, and R. A. Taylor. 2014. “Direct Normal Irradiance Forecasting and its Application to Concentrated Solar Thermal Output Forecasting—A Review. *Solar Energy* 108: 287–307. https://doi.org/10.1016/j.solener.2014.07.008.

Law, E. W., M. Kay, and R. A. Taylor. 2016. “Evaluating the Benefits of Using Short-Term Direct Normal Irradiance Forecasts to Operate a Concentrated Solar Thermal Plant.” *Solar Energy* 140: 93–108. https://doi.org/10.1016/j.solener.2016.10.037.

Le Gal La Salle, J., M. David, and P. Lauret. 2021. “A New Climatology Reference Model to Benchmark Probabilistic Solar Forecasts.” *Solar Energy* 223: 398–414. https://doi.org/10.1016/j.solener.2021.05.037.

LeCun, Y., B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. 1989. "Backpropagation Applied to Handwritten Zip Code Recognition." *Neural Computation* 1: 541–551. https://doi.org/10.1162/neco.1989.1.4.541.

Lee, J. A., S. E. Haupt, P. A. Jiménez, M. A. Rogers, S. D. Miller, and T. C. McCandless. 2017. "Solar Irradiance Nowcasting Case Studies Near Sacramento." *Journal of Applied Meteorology and Climatology* 56: 85–108. https://doi.org/10.1175/JAMC-D-16-0183.1.

Leutbecher, M., and T. N. Palmer. 2008. "Ensemble Forecasting." *Journal of Computational Physics* 227: 3,515–3,539. https://doi.org/10.1016/j.jcp.2007.02.014.

Li, J., and A. D. Heap. 2014. "Spatial Interpolation Methods Applied in the Environmental Sciences: A Review." *Environmental Modelling & Software* 53: 173–189. https://doi.org/10.1016/j.envsoft.2013.12.008.

Licciardi, G. A., R. Dambreville, J. Chanussot, and S. Dubost. 2015. :Spatiotemporal Pattern Recognition and Nonlinear PCA for Global Horizontal Irradiance Forecasting." *IEEE Geoscience and Remote Sensing Letters* 12: 284–288. https://doi.org/10.1109/LGRS.2014.2335817.

Lima, F. J. L., F. R. Martins, E. B. Pereira, E. Lorenz, and D. Heinemann. 2016. "Forecast for Surface Solar Irradiance at the Brazilian Northeastern Region Using NWP Model and Artificial Neural Networks." *Renewable Energy* 87: 807–818. https://doi.org/10.1016/j.renene.2015.11.005.

Lin, F., Y. Zhang, and J. Wang. 2023. "Recent Advances in Intra-Hour Solar Forecasting: A Review of Ground-Based Sky Image Methods." *International Journal of Forecasting* 39: 244–265. https://doi.org/10.1016/j.ijforecast.2021.11.002.

Logothetis, S.-A., V. Salamalikis, S. Wilbert, J. Remund, L. F. Zarzalejo, Y. Xie, B. Nouri, E. Ntavelis, J. Nou, N. Hendrikx, L. Visser, M. Sengupta, M. Pó, R. Chauvin, S. Grieu, N. Blum, W. Van Sark, and A. Kazantzidis. 2022. "Benchmarking of Solar Irradiance Nowcast Performance Derived From All-Sky Imagers." *Renewable Energy* 199: 246–261. https://doi.org/10.1016/j.renene.2022.08.127.

Long, C. N., J. M. Sabburg, J. Calbó, and D. Pagès. 2006. "Retrieving Cloud Characteristics From Ground-Based Daytime Color All-Sky Images." *Journal of Atmospheric and Oceanic Technology* 23: 633–652. https://doi.org/10.1175/JTECH1875.1.

Lonij, V. P. A., A. E. Brooks, A. D. Cronin, M. Leuthold, and K. Koch. 2013. "Intra-Hour Forecasts of Solar Power Production Using Measurements From a Network of Irradiance Sensors." *Solar Energy* 97: 58–66. https://doi.org/10.1016/j.solener.2013.08.002.

López-Cuesta, M., R. Aler-Mur, I. M. Galván-León, F. J. Rodríguez-Benítez, and A. D. Pozo-Vázquez. 2023. "Improving Solar Radiation Nowcasts by Blending Data-Driven, Satellite-Images-Based, and All-Sky-Imagers-Based Models Using Machine Learning Techniques." *Remote Sensing* 15, 2328. https://doi.org/10.3390/rs15092328.

Lorenz, E. 2018. “Solar Irradiance Forecasting for System Integration of Solar Energy.”

Lorenz, E., A. Hammer, and D. Heinemann. 2004. “Short-Term Forecasting of Solar Radiation Based on Satellite Data.”

Lorenz, E., and D. Heinemann. 2012. “Prediction of Solar Irradiance and Photovoltaic Power.” In *Comprehensive Renewable Energy*. Elsevier. https://doi.org/10.1016/B978-0-08-087872-0.00114-1.

Lorenz, E., D. Heinemann, and C. Kurz. 2012. “Local and Regional Photovoltaic Power Prediction for Large-Scale Grid Integration: Assessment of a New Algorithm for Snow Detection: Assessment of a New Algorithm for Snow Detection. Progress in Photovoltaics: Research and Applications 20: 760–769. https://doi.org/10.1002/pip.1224.

Lorenz, E., J. Hurka, D. Heinemann, and H. G. Beyer. 2009. “Irradiance Forecasting for the Power Prediction of Grid-Connected Photovoltaic Systems.” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 2: 2–10. https://doi.org/10.1109/JSTARS.2009.2020300.

Lorenz, E., J. Kühnert, D. Heinemann, K. P. Nielsen, J. Remund, and S. C. Müller. 2016. “Comparison of Global Horizontal Irradiance Forecasts Based on Numerical Weather Prediction Models With Different Spatio-Temporal Resolutions.” *Progress in Photovoltaics* 24: 1626–1640. https://doi.org/10.1002/pip.2799.

Lorenz, E., T. Scheidsteger, J. Hurka, D. Heinemann, and C. Kurz. 2011. “Regional PV Power Prediction for Improved Grid Integration.” *Progress in Photovoltaics* 19: 757–771. https://doi.org/10.1002/pip.1033.

Lucas, B., and K. Kanade. 1981. “An Iterative Image Registration Technique With an Application to Stereo Vision (IJCAI).”

Maas, A. L. 2013. “Rectifier Nonlinearities Improve Neural Network Acoustic Models.”

Marcos, J., L. Narvarte, I. Berazaluce, R. González, and J. Samuel. 2013. “Attenuation of Power Fluctuations in Large PV Power Plants: The Use of Prediction to Optimize the Storage System.” Presented at the 28th European Photovoltaic Solar Energy Conference and Exhibition, 4,252–4,255. https://doi.org/10.4229/28THEUPVSEC2013-5BV.7.56.

Marquez, R., and C. F. M. Coimbra. 2011. “Forecasting of Global and Direct Solar Irradiance Using Stochastic Learning Methods, Ground Experiments, and the NWS Database.” *Solar Energy* 85: 746–756. https://doi.org/10.1016/j.solener.2011.01.007.

Marquez, R., and C. F. M. Coimbra. 2013. “Intra-Hour DNI Forecasting Based on Cloud Tracking Image Analysis.” *Solar Energy* 91: 327–336. https://doi.org/10.1016/j.solener.2012.09.018.

Mathiesen, P., and J. Kleissl. 2011. "Evaluation of Numerical Weather Prediction for Intra-Day Solar Forecasting in the Continental United States." *Solar Energy* 85: 967–977. https://doi.org/10.1016/j.solener.2011.02.013.

Mejia, F. A., B. Kurtz, A. Levis, Í De La Parra, and J. Kleissl. 2018. "Cloud Tomography Applied to Sky Images: A Virtual Testbed." *Solar Energy* 176: 287–300. https://doi.org/10.1016/j.solener.2018.10.023.

Miller, S. D., M. A. Rogers, J. M. Haynes, M. Sengupta, and A. K. Heidinger. 2018. "Short-Term Solar Irradiance Forecasting via Satellite/Model Coupling." *Solar Energy* 168: 102–117. https://doi.org/10.1016/j.solener.2017.11.049.

Möhrlen, C., G. Giebel, and J.W. Zack. 2023. *IEA Wind Recommended Practice for the Implementation of Renewable Energy Forecasting Solutions*. Elsevier. https://doi.org/10.1016/C2021-0-03549-5.

Morales, J., M. Zugno, S. Pindea, and P. Pinson. 2014. "Electricity Market Clearing With Improved Scheduling of Stochastic Production." *European Journal of Operational Research* 235: 765–774.

Mueller, R., T. Behrendt, A. Hammer, and A. Kemper. 2012. "A New Algorithm for the Satellite-Based Retrieval of Solar Surface Irradiance in Spectral Bands." *Remote Sensing* 4: 622–647. https://doi.org/10.3390/rs4030622.

Müller, S. C., and J. Remund. 2014. "Satellite-Based Shortest Term Solar Energy Forecast System for Entire Europe for the Next Hours." Presented at the 29th European Photovoltaic Solar Energy Conference and Exhibition, 2,589–2,590. https://doi.org/10.4229/EUPVSEC20142014-5BV.1.6.

Nguyen, D. (Andu), and J. Kleissl. 2014. "Stereographic Methods for Cloud Base Height Determination Using Two Sky Imagers." *Solar Energy* 107: 495–509. https://doi.org/10.1016/j.solener.2014.05.005.

Nielsen, A. H., A. Iosifidis, and H. Karstoft. 2021. "IrradianceNet: Spatiotemporal Deep Learning Model for Satellite-Derived Solar Irradiance Short-Term Forecasting." *Solar Energy* 228: 659–669. https://doi.org/10.1016/j.solener.2021.09.073.

Nonnenmacher, L., and C. F. M. Coimbra. 2014. "Streamline-Based Method for Intra-Day Solar Forecasting Through Remote Sensing." *Solar Energy* 108: 447–459. https://doi.org/10.1016/j.solener.2014.07.026.

Notton, G., M.-L. Nivet, C. Voyant, C. Paoli, C. Darras, F. Motte, and A. Fouilloy. 2018. "Intermittent and Stochastic Character of Renewable Energy Sources: Consequences, Cost of Intermittence, and Benefit of Forecasting." *Renewable and Sustainable Energy Reviews* 87: 96–105. https://doi.org/10.1016/j.rser.2018.02.007.

Nouri, B., K. Noureldin, T. Schlichting, S. Wilbert, T. Hirsch, M. Schroedter-Homscheidt, P. Kuhn, A. Kazantzidis, L. F. Zarzalejo, P. Blanc, Z. Yasser, J. Fernández, and R. Pitz-Paal. 2020. "Optimization of Parabolic Trough Power Plant Operations in Variable Irradiance Conditions

Using All-Sky Imagers." *Solar Energy* 198: 434–453. https://doi.org/10.1016/j.solener.2020.01.045.

Nouri, B., P. Kuhn, S. Wilbert, N. Hanrieder, C. Prahl, L. Zarzalejo, A. Kazantzidis, P. Blanc, and R. Pitz-Paal. 2019. "Cloud Height and Tracking Accuracy of Three All-Sky Imager Systems for Individual Clouds." *Solar Energy* 177: 213–228. https://doi.org/10.1016/j.solener.2018.10.079.

Nouri, B., S. Wilbert, N. Blum, Y. Fabel, E. Lorenz, A. Hammer, T. Schmidt, L. F. Zarzalejo, and R. Pitz-Paal. 2023. "Probabilistic Solar Nowcasting Based on All-Sky Imagers." *Solar Energy* 253: 285–307. https://doi.org/10.1016/j.solener.2023.01.060.

Olmo, F. J., A. Cazorla, L. Alados-Arboledas, M. A. López-Álvarez, J. Hernández-Andrés, and J. Romero. 2008. "Retrieval of the Optical Depth Using an All-Sky CCD Camera." *Applied Optics* 47, H182. https://doi.org/10.1364/AO.47.00H182.

Paletta, Q., G. Arbod, and J. Lasenby. 2021. "Benchmarking of Deep Learning Irradiance Forecasting Models From Sky Images—An In-Depth Analysis." https://doi.org/10.48550/ARXIV.2102.00721.

Paletta, Q., G. Arbod, and J. Lasenby. 2023. "Omnivision Forecasting: Combining Satellite and Sky Images for Improved Deterministic and Probabilistic Intra-Hour Solar Energy Predictions." *Applied Energy* 336, 120818. https://doi.org/10.1016/j.apenergy.2023.120818.

Paragios, N., Y. Chen, and O. Faugeras. 2006. *Handbook of Mathematical Models in Computer Vision*. Boston, MA: Springer. https://doi.org/10.1007/0-387-28831-7.

Parzen, E. 1962. "On Estimation of a Probability Density Function and Mode." *The Annals of Mathematical Statistics* 33: 1,065–1,076. https://doi.org/10.1214/aoms/1177704472.

Pedro, H. T. C., C. F. M. Coimbra, M. David, and P. Lauret. 2018. "Assessment of Machine Learning Techniques for Deterministic and Probabilistic Intra-Hour Solar Forecasts." *Renewable Energy* 123: 191–203. https://doi.org/10.1016/j.renene.2018.02.006.

Pelland, S., G. Galanis, and G. Kallos. 2013. "Solar and Photovoltaic Forecasting Through Post-Processing of the Global Environmental Multiscale Numerical Weather Prediction Model." *Progress in Photovoltaics* 21: 284–296. https://doi.org/10.1002/pip.1180.

Peng, Z., D. Yu, D. Huang, J. Heiser, and P. Kalb. 2016. "A Hybrid Approach to Estimate the Complex Motions of Clouds in Sky Images." *Solar Energy* 138: 10–25. https://doi.org/10.1016/j.solener.2016.09.002.

Peng, Z., D. Yu, D. Huang, J. Heiser, S. Yoo, and P. Kalb. 2015. "3D Cloud Detection and Tracking System for Solar Forecast Using Multiple Sky Imagers." *Solar Energy* 118: 496–519. https://doi.org/10.1016/j.solener.2015.05.037.

Perez, R., and T. E. Hoff. 2013. "Chapter 10: SolarAnywhere Forecasting." In *Solar Energy Forecasting and Resource Assessment*. Edited by J. Kleissl. Boston, MA: Academic Press, Boston. https://doi.org/10.1016/B978-0-12-397177-7.00010-3.

Perez, R., E. Lorenz, S. Pelland, M. Beauharnois, G. Van Knowe, K. Hemker, D. Heinemann, J. Remund, S. C. Müller, W. Traunmüller, G. Steinmauer, D. Pozo, J. A. Ruiz-Arias, V. Lara-Fanego, L. Ramirez-Santigosa, M. Gaston-Romero, and L. M. Pomares. 2013. Comparison of Numerical Weather Prediction Solar Irradiance Forecasts in the U.S., Canada, and Europe." *Solar Energy* 94: 305–326. https://doi.org/10.1016/j.solener.2013.05.005.

Perez, R., J. Schlemmer, S. Kivalov, J. Dise, P. Keelin, M. Grammatico, T. Hoff, and A. Tuohy. 2018. "A New Version of the SUNY Solar Forecast Model: A Scalable Approach to Site-Specific Model Training."

Perez, R., K. Moore, S. Wilcox, D. Renné, and A. Zelenka. 2007. "Forecasting Solar Radiation—Preliminary Evaluation of an Approach Based Upon the National Forecast Database." *Solar Energy* 81: 809–812. https://doi.org/10.1016/j.solener.2006.09.009.

Perez, R., M. David, T. E. Hoff, M. Jamaly, S. Kivalov, J. Kleissl, P. Lauret, and M. Perez. 2016. "Spatial and Temporal Variability of Solar Energy." *Foundations and Trends in Renewable Energy* 1: 1–44. https://doi.org/10.1561/2700000006.

Pierro, M., C. Cornaro, D. Moser, A. Betti, M. Moschella, E. Collino, D. Ronzio, D. Van der Meer, J. Widen, L. Visser, T. A. AlSkaif, and W. Van Sark. 2020a. *Regional Solar Power Forecasting*. No. Report IEA-PVPS T16-01: 2020.

Pierro, M., F. Bucci, C. Cornaro, E. Maggioni, A. Perotto, M. Pravettoni, and F. Spada. 2015. "Model Output Statistics Cascade to Improve Day Ahead Solar Irradiance Forecast." *Solar Energy* 117: 99–113. https://doi.org/10.1016/j.solener.2015.04.033.

Pierro, M., M. De Felice, E. Maggioni, D. Moser, A. Perotto, F. Spada, and C. Cornaro. 2017. "Data-Driven Upscaling Methods for Regional Photovoltaic Power Estimation and Forecast Using Satellite and Numerical Weather Prediction Data." *Solar Energy* 158: 1,026–1,038. https://doi.org/10.1016/j.solener.2017.09.068.

Pierro, M., M. De Felice, F. Maggioni, D. Moser, A. Perotto, F. Spada, and C. Cornaro. 2020b. "Residual Load Probabilistic Forecast for Reserve Assessment: A Real Case Study." *Renewable Energy* 149: 508–522. https://doi.org/10.1016/j.renene.2019.12.056.

Pierro, M., R. Perez, M. Perez, D. Moser, and C. Cornaro. 2020c. "Italian Protocol for Massive Solar Integration: Imbalance Mitigation Strategies." *Renewable Energy* 153: 725–739. https://doi.org/10.1016/j.renene.2020.01.145.

Pinson, P. 2012. "Adaptive Calibration of (u,v)-Wind Ensemble Forecasts." *Quarterly Journal of the Royal Meteorological Society* 138: 1273–1284. https://doi.org/10.1002/qj.1873.

Pinson, P., and H. Madsen. 2009. "Ensemble-Based Probabilistic Forecasting at Horns Rev." *Wind Energy* 12: 137–155. https://doi.org/10.1002/we.309.

Pinson, P., H. Aa. Nielsen, J. K. Møller, H. Madsen, and G. N. Kariniotakis. 2007. "Non-Parametric Probabilistic Forecasts of Wind Power: Required Properties and Evaluation." *Wind Energy* 10: 497–516. https://doi.org/10.1002/we.230.

Prasad, A. A., and M. Kay. 2021. "Prediction of Solar Power Using Near-Real Time Satellite Data." *Energies* 14, 5865. https://doi.org/10.3390/en14185865.

Ramahatana, F., J. La Salle, P. Lauret, and M. David. 2022. "A more efficient microgrid operation through the integration of probabilistic solar forecasts." Sustainable Energy, Grids and Networks 31. https://doi.org/10.1016/j.segan.2022.100783.

Reikard, G. 2009. "Predicting Solar Radiation at High Resolutions: A Comparison of Time Series Forecasts." *Solar Energy* 83: 342–349. https://doi.org/10.1016/j.solener.2008.08.007.

Rigollier, C., M. Lefèvre, and L. Wald. 2004. "The Method Heliosat-2 for Deriving Shortwave Solar Radiation From Satellite Images." *Solar Energy* 77: 159–169. https://doi.org/10.1016/j.solener.2004.04.017.

Ruder, S. 2017. "An Overview of Gradient Descent Optimization Algorithms."

Ruiz-Arias, J. A., J. Dudhia, and C. A. Gueymard. 2014. "A Simple Parameterization of the Short-Wave Aerosol Optical Properties for Surface Direct and Diffuse Irradiances Assessment in a Numerical Weather Model." *Geoscientific Model Development* 7: 1,159–1,174. https://doi.org/10.5194/gmd-7-1159-2014.

Rumelhart, D. E., and J. L. McClelland. 1986. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 1, Foundations, Computational Models of Cognition and Perception*. Cambridge, MA: MIT Press.

Saint-Drenan, Y. M., G. H. Good, and M. Braun. 2017. "A Probabilistic Approach to the Estimation of Regional Photovoltaic Power Production." *Solar Energy* 147: 257–276. https://doi.org/10.1016/j.solener.2017.03.007.

Saint-Drenan, Y. M., G. H. Good, M. Braun, and T. Freisinger. 2016. "Analysis of the Uncertainty in the Estimates of Regional PV Power Generation Evaluated With the Upscaling Method." *Solar Energy* 135: 536–550. https://doi.org/10.1016/j.solener.2016.05.052.

Saint-Drenan, Y.-M. 2015. "A Probabilistic Approach to the Estimation of Regional Photovoltaic Power Generation Using Meteorological Data." Kassel, Germany.

Saint-Drenan, Y.-M., L. Wald, T. Ranchin, L. Dubus, and A. Troccoli. 2018. "An Approach for the Estimation of the Aggregated Photovoltaic Power Generated in Several European Countries From Meteorological Data." *Advances in Science and Research* 15: 51–62. https://doi.org/10.5194/asr-15-51-2018.

Saint-Drenan, Y.-M., S. Bofinger, K. Rohrig, and B. Ernst. 2011. "Regional Nowcasting of the Solar Power Production With PV-Plant Measurements and Satellite Images." Presented at the

ISES Solar World Congress 2011, International Solar Energy Society, Kassel, Germany, 1–11. https://doi.org/10.18086/swc.2011.11.09.

Sak, H., A. Senior, and F. Beaufays. 2014. "Long Short-Term Memory Based Recurrent Neural Network Architectures for Large Vocabulary Speech Recognition." https://doi.org/10.48550/ARXIV.1402.1128.

Sánchez-Segura, C. D., L. Valentín-Coronado, M. I. Peña-Cruz, A. Díaz-Ponce, D. Moctezuma, G. Flores, and D. Riveros-Rosas. 2021. "Solar Irradiance Components Estimation Based on a Low-Cost Sky-Imager." *Solar Energy* 220: 269–281. https://doi.org/10.1016/j.solener.2021.02.037.

Sanfilippo, A., L. Martin-Pomares, N. Mohandes, D. Perez-Astudillo, and D. Bachou. 2016. "An Adaptive Multi-Modeling Approach to Solar Nowcasting." *Solar Energy* 125: 77–85. https://doi.org/10.1016/j.solener.2015.11.041.

Scaramuzza, D., A. Martinelli, and R. Siegwart. 2006. "A Toolbox for Easily Calibrating Omnidirectional Cameras." Presented at the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, 5,695–5,701. https://doi.org/10.1109/IROS.2006.282372.

Schmidt, T., J. Kalisch, E. Lorenz, and D. Heinemann. 2016. "Evaluating the Spatio-Temporal Performance of Sky-Imager-Based Solar Irradiance Analysis and Forecasts." *Atmospheric Chemistry and Physics* 16: 3,399–3,412. https://doi.org/10.5194/acp-16-3399-2016.

Schroedter-Homscheidt, M., A. Benedetti, and N. Killius. 2017. "Verification of ECMWF and ECMWF/MACC's Global and Direct Irradiance Forecasts With Respect to Solar Electricity Production Forecasts." *Meteorologische Zeitschrift*: 1–19. https://doi.org/10.1127/metz/2016/0676.

Schroedter-Homscheidt, M., and B. Pulvermüller. 2011. "Verification of Direct Normal Irradiance Forecasts for the Concentrating Solar Thermal Power Plant Andasol-3 Location." Presented at the SolarPaces 2011, Granada, Spain.

Schroedter-Homscheidt, M., and S. Wilbert. 2017. "Methods to Provide Meteorological Forecasts for Optimum CSP System Operations." In *The Performance of Concentrated Solar Power (CSP) Systems*. Elsevier. https://doi.org/10.1016/B978-0-08-100447-0.00008-0.

Sengupta, M., P. A. Jiménez, J.-H. Kim, J. Yang, and Y. Xie. 2022. *Probabilistic Cloud Optimized Day-Ahead Forecasting System Based on WRF-Solar (Final Report)*. https://doi.org/10.2172/1855782.

Shaffery, P., A. Habte, M. Netto, A. Andreas, and V. Krishnan. 2020. "Automated Construction of Clear-Sky Dictionary From All-Sky Imager Data." *Solar Energy* 212: 73–83. https://doi.org/10.1016/j.solener.2020.10.052.

Shi, C., Y. Wang, C. Wang, and B. Xiao. 2017. "Ground-Based Cloud Detection Using Graph Model Built Upon Superpixels." *IEEE Geoscience and Remote Sensing Letters* 14: 719–723. https://doi.org/10.1109/LGRS.2017.2676007.

Simoglou, C. K., E. G. Kardakos, E. A. Bakirtzis, D. I. Chatzigiannis, S. I. Vagropoulos, A. V. Ntomaris, P. N. Biskas, A. Gigantidou, E. J. Thalassinakis, A. G. Bakirtzis, and J. P. S. Catalão. 2014. "An Advanced Model for the Efficient and Reliable Short-Term Operation of Insular Electricity Networks With High Renewable Energy Sources Penetration." *Renewable and Sustainable Energy Reviews* 38: 415–427. https://doi.org/10.1016/j.rser.2014.06.015.

Sirch, T., L. Bugliaro, T. Zinner, M. Möhrlein, and M. Vazquez-Navarro. 2017. "Cloud and DNI Nowcasting With MSG/SEVIRI for the Optimized Operation of Concentrating Solar Power Plants." *Atmospheric Measurement Techniques* 10: 409–429. https://doi.org/10.5194/amt-10-409-2017.

Skamarock, C., B. Klemp, J. Dudhia, O. Gill, M. Barker, W. Wang, and G. Powers. 2005. "A Description of the Advanced Research WRF Version 2." https://doi.org/10.5065/D6DZ069T.

Sobri, S., S. Koohi-Kamali, and N. A. Rahim. 2018. "Solar Photovoltaic Generation Forecasting Methods: A Review." *Energy Conversion and Management* 156: 459–497. https://doi.org/10.1016/j.enconman.2017.11.019.

Sosa-Tinoco, I., J. Peralta-Jaramillo, C. Otero-Casal, A. López- Agüera, G. Miguez-Macho, and I. Rodríguez-Cabo. 2016. "Validation of a Global Horizontal Irradiation Assessment From a Numerical Weather Prediction Model in the South of Sonora–Mexico." *Renewable Energy* 90: 105–113. https://doi.org/10.1016/j.renene.2015.12.055.

Sperati, S., S. Alessandrini, and L. Delle Monache. 2016. "An Application of the ECMWF Ensemble Prediction System for Short-Term Solar Power Forecasting." *Solar Energy* 133: 437–450. https://doi.org/10.1016/j.solener.2016.04.016.

Straub, N., W. Herzberg, A. Dittmann, and E. Lorenz. 2023. "Combination of a Novel All Sky Imager Based Approach for High-Resolution Solar Irradiance Nowcasting With Persistence and Satellite Nowcasts for Increased Accuracy." Presented at the EU PVSEC 2023. https://doi.org/10.4229/EUPVSEC2023/4CO.8.3.

Straub, N., W. Herzberg, A. Dittmann, and E. Lorenz. 2024. "Blending of a Novel All Sky Imager Model With Persistence and a Satellite Based Model for High-Resolution Irradiance Nowcasting." *Solar Energy* 269, 112319. https://doi.org/10.1016/j.solener.2024.112319.

Su, F., W. Jiang, J. Zhang, H. Wang, and M. Zhang. 2015. "A Local Features-Based Approach to All-Sky Image Prediction." *IBM Journal of Research and Development* 59 (6): 1–6:10. https://doi.org/10.1147/JRD.2015.2397772.

Sun, D., S. Roth, and M. J. Black. 2010. "Secrets of Optical Flow Estimation and Their Principles."

Sun, Y., G. Szűcs, and A. R. Brandt. 2018. "Solar PV Output Prediction From Video Streams Using Convolutional Neural Networks." *Energy and Environmental Science* 11: 1,811–1,818. https://doi.org/10.1039/C7EE03420B.

Thompson, G., and T. Eidhammer. 2014. "A Study of Aerosol Impacts on Clouds and Precipitation Development in a Large Winter Cyclone." *Journal of the Atmospheric Sciences* 71: 3,636–3,658. https://doi.org/10.1175/JAS-D-13-0305.1.

Tuohy, A., J. Zack, S. E. Haupt, J. Sharp, M. Ahlstrom, S. Dise, E. Grimit, C. Mohrlen, M. Lange, M. G. Casado, J. Black, M. Marquis, and C. Collier. 2015. "Solar Forecasting: Methods, Challenges, and Performance." *IEEE Power and Energy Magazine* 13: 50–59. https://doi.org/10.1109/MPE.2015.2461351.

Ulbricht, R., U. Fischer, W. Lehner, and H. Donker. 2013. "First Steps Towards a Systematical Optimized Strategy for Solar Energy Supply Forecasting."

Urbich, I., J. Bendix, and R. Müller. 2019. "The Seamless Solar Radiation (SESORA) Forecast for Solar Surface Irradiance—Method and Validation." *Remote Sensing* 11, 2576. https://doi.org/10.3390/rs11212576.

Urquhart, B., B. Kurtz, E. Dahlin, M. Ghonima, J. E. Shields, and J. Kleissl. 2015. "Development of a Sky Imaging System for Short-Term Solar Power Forecasting." *Atmospheric Measurement Techniques* 8: 875–890. https://doi.org/10.5194/amt-8-875-2015.

Urquhart, B., C. W. Chow, A. Nguyen, J. Kleissl, M. Sengupta, J. Blatchford, and D. Jeon. 2012. "Towards Intra-Hour Solar Forecasting Using Two Sky Imagers at a Large Solar Power Plant."

Urquhart, B., M. Ghonima, D. (A.) Nguyen, B. Kurtz, C. W. Chow, and J. Kleissl. 2013. "Sky-Imaging Systems for Short-Term Forecasting." In *Solar Energy Forecasting and Resource Assessment*. Elsevier. https://doi.org/10.1016/B978-0-12-397177-7.00009-7.

Van Der Meer, D. W., J. Widén, and J. Munkhammar. 2018. "Review on Probabilistic Forecasting of Photovoltaic Power Production and Electricity Consumption." *Renewable and Sustainable Energy Reviews* 81: 1,484–1,512. https://doi.org/10.1016/j.rser.2017.05.212.

Visser, L., E. Lorenz, D. Heinemann, W. G. J. H. H. Van Sark. 2022. "Solar Power Forecasts." In *Comprehensive Renewable Energy*. Elsevier. https://doi.org/10.1016/B978-0-12-819727-1.00135-7.

Voyant, C., G. Notton, S. Kalogirou, M.-L. Nivet, C. Paoli, F. Motte, and A. Fouilloy. 2017. "Machine Learning Methods for Solar Radiation Forecasting: A Review." *Renewable Energy* 105: 569–582. https://doi.org/10.1016/j.renene.2016.12.095.

Wan, C., J. Zhao, Y. Song, Z. Xu, J. Lin, and Z. Hu. 2015. "Photovoltaic and Solar Power Forecasting for Smart Grid Energy Management." *CSEE Journal of Power and Energy Systems* 1: 38–46. https://doi.org/10.17775/CSEEJPES.2015.00046.

Wang, P., R. Van Westrhenen, J. F. Meirink, S. Van Der Veen, and W. Knap. 2019. "Surface Solar Radiation Forecasts by Advecting Cloud Physical Properties Derived From Meteosat Second Generation Observations." *Solar Energy* 177: 47–58. https://doi.org/10.1016/j.solener.2018.10.073.

Wilks, D. S. 2019. *Statistical Methods in the Atmospheric Sciences*. Elsevier. https://doi.org/10.1016/C2017-0-03921-6.

Winter, K., D. Beinert, S. Vogt, and R. Fritz. 2019. "PV Power Forecast Comparison of Physical and Machine Learning Models."

Wolff, B., J. Kühnert, E. Lorenz, O. Kramer, and D. Heinemann. 2016. "Comparing Support Vector Regression for PV Power Forecasting to a Physical Modeling Approach Using Measurement, Numerical Weather Prediction, and Cloud Motion Data." *Solar Energy* 135: 197–208. https://doi.org/10.1016/j.solener.2016.05.051.

Wolfgang B. 2007. "60 Years Operational Satellites: An Overview From the Beginning of Operational Satellite Applications at an NMS up to 2030."

Xie, W., D. Liu, M. Yang, S. Chen, B. Wang, Z. Wang, Y. Xia, Y. Liu, Y. Wang, and C. Zhang. 2020. "SegCloud: A Novel Cloud Image Segmentation Model Using a Deep Convolutional Neural Network for Ground-Based All-Sky-View Camera Observation." Atmospheric Measurement Techniques 13: 1,953–1,961. https://doi.org/10.5194/amt-13-1953-2020.

Yang, D. 2019. "A Universal Benchmarking Method for Probabilistic Solar Irradiance Forecasting." *Solar Energy* 184: 410–416. https://doi.org/10.1016/j.solener.2019.04.018.

Yang, D. 2020. "Ensemble Model Output Statistics as a Probabilistic Site-Adaptation Tool for Solar Irradiance: A Revisit." *Journal of Renewable and Sustainable Energy* 12, 036101. https://doi.org/10.1063/5.0010003.

Yang, D., and D. Van Der Meer. 2021. "Post-Processing in Solar Forecasting: Ten Overarching Thinking Tools." *Renewable and Sustainable Energy Reviews* 140, 110735. https://doi.org/10.1016/j.rser.2021.110735.

Yang, D., C. Gu, Z. Dong, P. Jirutitijaroen, N. Chen, and W. M. Walsh. 2013. "Solar Irradiance Forecasting Using Spatial-Temporal Covariance Structures and Time-Forward Kriging." *Renewable Energy* 60: 235–245. https://doi.org/10.1016/j.renene.2013.05.030.

Yang, D., H. Quan, V. R. Disfani, and L. Liu. 2017. "Reconciling Solar Forecasts: Geographical Hierarchy." *Solar Energy* 146: 276–286. https://doi.org/10.1016/j.solener.2017.02.010.

Yang, D., J. Kleissl, C. A. Gueymard, H. T. C. Pedro, and C. F. M. Coimbra. 2018. "History and Trends in Solar Irradiance and PV Power Forecasting: A Preliminary Assessment and Review Using Text Mining." *Solar Energy* 168: 60–101. https://doi.org/10.1016/j.solener.2017.11.023.

Yang, D., 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.-M. Saint-Drenan, Y. Shuai, R. Urraca, H. Verbois, F. Vignola, C. Voyant, and J. Zhang. 2020. "Verification of Deterministic Solar Forecasts." *Solar Energy* 210: 20–37. https://doi.org/10.1016/j.solener.2020.04.019.

Yang, J., J.-H. Kim, P. A. Jiménez, M. Sengupta, J. Dudhia, Y. Xie, A. Golnas, and R. Giering. 2021. "An Efficient Method to Identify Uncertainties of WRF-Solar Variables in Forecasting Solar Irradiance Using a Tangent Linear Sensitivity Analysis." *Solar Energy* 220: 509–522. https://doi.org/10.1016/j.solener.2021.03.044.

Ye, L., Z. Cao, Y. Xiao, and Z. Yang. 2019. "Supervised Fine-Grained Cloud Detection and Recognition in Whole-Sky Images." *IEEE Transactions on Geoscience and Remote Sensing* 57: 7,972–7,985. https://doi.org/10.1109/TGRS.2019.2917612.

Zamo, M., O. Mestre, P. Arbogast, and O. Pannekoucke. 2014. "A Benchmark of Statistical Regression Methods for Short-Term Forecasting of Photovoltaic Electricity Production. Part II: Probabilistic Forecast of Daily Production." *Solar Energy* 105: 804–816. https://doi.org/10.1016/j.solener.2014.03.026.

Zemouri, N., H. Bouzgou, and C. A. Gueymard. 2019. "Multimodel Ensemble Approach for Hourly Global Solar Irradiation Forecasting." *European Physical Journal Plus* 134: 594. https://doi.org/10.1140/epjp/i2019-12966-5.

Zempila, M.-M., T. M. Giannaros, A. Bais, D. Melas, and A. Kazantzidis. 2016. "Evaluation of WRF Shortwave Radiation Parameterizations in Predicting Global Horizontal Irradiance in Greece." *Renewable Energy* 86: 831–840. https://doi.org/10.1016/j.renene.2015.08.057.

Zhang, J., R. Verschae, S. Nobuhara, and J.-F. Lalonde. 2018. "Deep Photovoltaic Nowcasting." *Solar Energy* 176: 267–276. https://doi.org/10.1016/j.solener.2018.10.024.

Zhao, X., H. Wei, H. Wang, T. Zhu, and K. Zhang. 2019. "3D-CNN-Based Feature Extraction of Ground-Based Cloud Images for Direct Normal Irradiance Prediction." *Solar Energy* 181: 510–518. https://doi.org/10.1016/j.solener.2019.01.096.