Forecasting Solar Power Ramp Events Using Machine Learning Classification Techniques

Mohamed Abuella
Department of Electrical and Computer Engineering
Energy Production and Infrastructure Center
University of North Carolina at Charlotte
Charlotte, USA
mabuella@uncc.edu

Badrul Chowdhury
Department of Electrical and Computer Engineering
Energy Production and Infrastructure Center
University of North Carolina at Charlotte
Charlotte, USA
b.chowdhury@uncc.edu

***Abstract*—The growing integration level of wind and solar energy resources introduces new challenges for the reliable operation of the electric grid. Tasks such as managing high ramp-rates of renewable generation, optimal energy management of energy storage systems, and voltage regulator settings on feeders with distributed generation, may be improved with the availability of solar power forecasts, especially those with accurate ramp event prediction. This paper presents classification techniques to classify and forecast the solar power ramp events. A case study over an entire year is conducted and several evaluation metrics are considered to assess the performance of the classification models of solar power ramp event forecasts.**

***Index Terms*—Classification, evaluation metrics, model selection, ramp events, solar power forecasting**

I. Introduction

The rapid growth of distributed generation (DG) introduces new challenges for reliable operation of the electric power distribution grid. Tasks, such as the optimal management of energy storage systems, voltage regulation of feeders with a significant penetration of solar photovoltaic (PV) power, and managing high ramp-rate conditions of those variable energy resources (VER) due to the slower response of the conventional power plants, depend on the availability of accurate forecasts of the VER output, especially high-quality forecasts of ramp events [1]. The intermittency of DG could adversely impact the operation of power systems that have high shares of VER. Thus, wherever the VER are deployed, it becomes a recommended practice to maintain more than normal operating reserves and energy storage systems to respond to any power mismatches in the system. The capacity of these operating reserves and energy storage systems should be kept to a minimum as much as possible for maximum economic profits of deployment the renewable resources. Therefore, forecasting the wind and solar energy resources becomes a vital tool in the operation of power systems and electricity markets as well [2]. From an operations point of view, the power grid might encounter frequency excursions if the overall response rate is slower than the ramp rate of the solar power generations. Therefore, it is important to provide the forecasts that consider the power ramp events [3], [4].

Most studies in the literature are limited to identification or detection of solar power ramp events in the measured data rather than forecasting. Solar power ramp events forecasting is more challenging as the whole framework depends on the forecasts as inputs to the models. Florita et al. [5] propose a swinging door algorithm to identify the ramp events of wind and solar power, and a new evaluation metric is also proposed to assess the identification performance. A machine learning tool to forecast ramps-up and down of solar irradiance and wind speed is reported in [6]. A study conducted by a team of researchers at Sandia National Laboratory [7] focuses on the variability of solar irradiance by identifying the clear-sky solar irradiance.

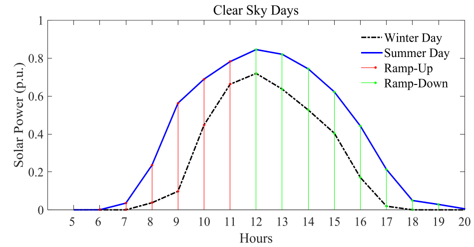
From a forecasting point of view, the majority of the studies attempt to forecast the ramp events of the wind power. Chu et al. [8] track the clouds. They use an artificial neural network to provide very short-term forecasts of the solar irradiance ramps with a horizon of up to ten minutes for two sites in California. They conclude that the performance of the proposed approach outperforms the persistence forecasts. Ferreira et al. conducted a survey [9] about wind power ramp events forecasting studies up to 2010, and they concluded that the topic needs more research in forecasting and evaluation. Gallego-Castillo et al conducted another updated survey [10] which reflects the research of wind power ramp events (WPREs) forecasting has experienced a noticeable increase since the former survey [9] was published. Zheng and Kusiak [11] apply five machine learning regression techniques to forecast the wind power ramp rates from wind power observations in horizons ranging from 10 to 60-minutes. Sevlian and Rajagopal [12] develop an identification technique to detect wind power ramps in historical wind power measurements. Feng et al. [13] propose an elaborate framework of feature selection for short-term wind forecasts, wherein, the outcome of selected features significantly improve the forecasts in 1-hour lead times.

The rest of the paper is organized as follows: Section II introduces the solar power ramp events. Section III explains the model selection procedure and provides an overview of the classification models. Section IV presents numerous evaluation metrics to assess the performance of the classification models. Section V includes the results of the case study. Section VI provides the conclusion.

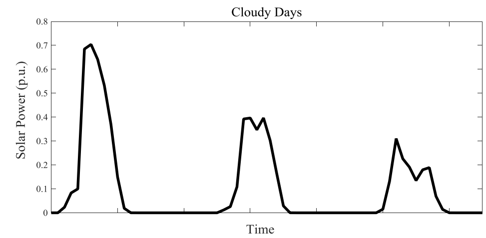
II. Definition of Solar Power Ramp Rates

The application scope of the issues of solar power variability depends on the extent of the variation and the size, or the voltage level of the power system where these issues are taking place. For instance, at the distribution level, the fast ramps affect the charge and discharge mechanism of energy storage devices, as well as voltage regulation equipment on the system, while on the bulk transmission level, the slower ramps have an impact on trading decisions and dispatching of the operating reserve facilities, and their coordination with other generation sources. Therefore, a tool for prediction of solar power ramp events may be needed to mitigate some of these potential issues.

The solar power from PV systems is naturally variable and changes occur in fast and slow ramps due to cloud cover, and the azimuth and zenith angles of the sun during the day and seasons. Thus, the solar power ramps exist not only in the presence of clouds, but also in clear sky situations in the morning and late afternoon, as shown in Fig. 1(a) for clear sky days in the summer and winter.



(a)



(b)

Figure 1. (a) Solar power observations and ramp rates of clear sky summer and winter days; (b) solar power observations of cloudy days

$$Ramp\ Rate, RR_P(t) = \frac{dP(t)}{dt} = \frac{F(t) - P(t-D)}{D} \quad (1)$$

where *F(t)* is the solar power forecast of the t^{th} hour; $P(t-D)$ is the observed solar power of the preceding hour; *D* is the time duration for which the ramp rate is determined (D=1 hour in this study), it is also the moving time window of the rolling forecasts of the ramp events. Therefore, the ramp rate is simply the difference between the solar power forecast at the forecasted hour and the observed solar power of the proceeding hour (i.e, *F(t) – P(t-D)*).

The event could be a ramp-up (positive rate) or a ramp-down (negative rate). It could also be an extreme ramp of a high rate or a normal ramp of a low rate. The variability of the solar power output can be decreased by scattering the solar plant across a larger region, thereby, taking advantage of the geographical smoothing effect.

III. Modeling

A. Classification Models

The classification models that have been implemented for solar power ramp events are reviewed in this section. For more details about the classification models, the interested reader may refer to [14], [15].

1) Naive Bayes: The classification in this model is carried out by the conditional probability of a j^{th} class of the output variable Y at which, the input variable X belongs to a vector X_o. Although it seems to be naive, its concept is the basis of other sophisticated and more powerful classification models. The conditional probabilities of the output classes are estimated by the training set of observations.

2) Linear Discriminant Analysis (LDA): This is a simple parametric model which makes some assumptions about the conditional distribution of its input and output variables. Since the LDA depends on the mean and the variance of the observations of each class, this makes it more sensitive to the observations than other classification models, but in general, it is stable with dispersive observations. It is popular in cases of multiple classes. The coefficients to the linear discriminant are estimated through the training.

3) k-Nearest Neighbors (kNN): A non-parametric model without assumptions which is built based on a simple intuition for determining the observation of the unknown class by measuring the distances to the neighboring observations, and thus the given observation's class belongs to the major class of the nearest observations. k stands for the number of the neighboring observations that are used to identify the class of a given observation, its value is chosen by searching the optimal value of the highest accuracy in the training set.

4) Decision Tree: The decision trees can be implemented efficiently for problems of binary decisions since they are recursively splitting the data into two sets. The binary splitting procedure is conducted by applying conditional tests on samples of each variable, and it continues until the minimum number of the samples is reached at each node. The tree splitting should be terminated (i.e., tree pruning) before the tree overfits with the data during the training. The maximum number of splits and the minimum number of samples (leaf size) are set for more efficient and less complexity of the decision tree model. One of the advantages of the decision tree model is that it can deal with both continuous and categorical variables.

5) Logistic Regression: In a nutshell the logistic regression is a special case of the generalized linear regression model. Like the LDA model, the logistic regression is also a parametric model, since it has some assumptions. Its output is categorical, while its input variables can be either continuous or categorial. The maximum likelihood method is used during the training to estimate the coefficients of the models and then, those fitted coefficients are applied for the test set.

6) Random Forests (RF): Random forests have been proposed to tackle the correlated classification and regression trees (CART). The trees of the RF are more various and uncorrelated as they are grown by a random number of features and observation samples. Two main parameters are required to be set in the RF: the number of trees B (forest size), and the minimum number n_{min} of observations per node (leaf size). It is worth mentioning that the performance of the RF is not overly sensitive to the values of these parameters. In addition, the RF does not rely on cross-validation to estimate the parameters because it has a built-in out-of-bag (OOB) estimation algorithm, which validates the performance of the trees with samples that are not used in the training. Thus, the robustness and flexibility are the main advantages of the RF model.

7) Support Vector Machines (SVM): This is a machine learning technique that is widely implemented for data classification applications. Essentially, the classification by the SVM is carried out as a binary classification. For cases of more than two classes, the one-vs.-rest strategy is adopted. The SVM model utilizes kernel tricks to make the classes more separable, since they are mapped into a higher dimensional space. The common kernels are linear, polynomial, and radial basis functions (RBF). The hyperparameters of the SVM model are C as an optimization penalty parameter, and γ, which is a kernel parameter. The optimal hyperparameters are found by a grid search. A complete grid search could be time-consuming; so heuristic search algorithms can be used instead. The SVM model is less prone to overfitting and local optima issues.

8) Artificial Neural Network (ANN): The artificial neural network has also been implemented for pattern recognition problems. The basic ANN has three layers, input, output, and hidden layers. The neurons are in the hidden layer, and they connect the input layer with the output layer. If the number of neurons is too large, this can cause an overfitting issue, whence the ANN model will perform with high accuracy in the training stage, but will perform poorly in the testing stage. Thus, for proper regularization, a suitable number of neurons is designated for the ANN model. The back-propagation algorithm for training an ANN is based on the gradient descent, and can sometimes converge to a local optima, instead of the global optimum.

B. Model Selection

The available historical data contains the solar power observations and predictions of various weather conditions, such as solar irradiance, cloud cover information, air temperature, relative humidity, wind speed, etc. In addition to several solar power forecasts, including NWP-driven day-ahead forecasts generated by multiple linear regression (MLR), support vector regression (SVR) and ANN models. Hour-ahead time-series forecasts from the persistence and nonlinear autoregressive exogenous (NARX) models. Moreover, the hour-ahead combined forecasts are also included. Further details about these solar power forecasts can be found in [16]. The ramp rates of the solar power forecasts are calculated as in (1), and linked with the features of the classification models.

Using the irrelevant features (input variables) in the classification model leads to more complexity of the model. Therefore, the subset of features are selected by applying the greedy search approach in the training set of the available data, through a cross-validation strategy, to find the most effective subset of features for each classification model. The wrapper technique is adopted for the feature search, which considers the interaction of the input variables with each other, and their correlation to the solar power ramp events (the output) of the classification models. The wrapper search is carried out for all possible combinations of input variables, retaining the most effective variables, and removing the less effective ones from the selected input variables [17]. Since the objective of the classification is to forecast the classes of the ramp events of solar power, the difference between the true and false forecasted events of high-rate ramp events of solar power (Diff. Index), as shown in (2), is chosen as a score for selecting the best subset of input variables by the wrapper search approach.

To obtain an adequate regularization and avoid the overfitting and underfitting issues of the machine learning based-classification models, a grid search is implemented through cross-validation in the dataset, to tune the parameters of each model with the corresponding selected features.

IV. Evaluation Metrics

The following evaluation metrics are used to assess the performance of the classification techniques:

$$\text{Diff. Index} = (\text{True} - \text{False}) \text{ of High-Rate Events} \tag{2}$$

$$\text{Total Accuracy} = \frac{\text{True Events}}{\text{Total Events}} \tag{3}$$

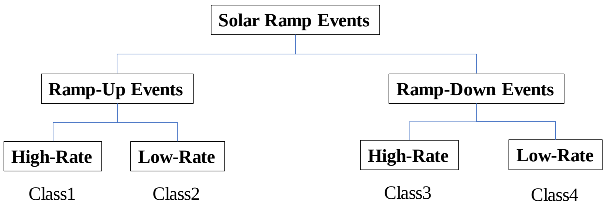
$$\text{Precision} = \frac{\text{True High}}{\text{True High + False High}} \tag{4}$$

$$\text{Recall (Sensitivity)} = \frac{\text{True High}}{\text{True High + False Low}} \tag{5}$$

$$\text{Balanced Precision} = \frac{1}{4} \sum_{Class=1}^{n} \frac{\text{True Class}}{\text{True Class + False Class}} \tag{6}$$

$$\text{F}_1 \text{ score} = \frac{2.(\text{Precision} \times \text{Recall})}{(\text{Precision + Recall})} \tag{7}$$

Where High denotes high-rate ramp events and Low refers to low rate ramps; True events are when the events are predicted to belong to the same classes as found in the actual observations, while False is indicated when the events are predicted to be in classes other than those found in the actual observations. Some of these metrics are



Class	Class1	Class2	Class3	Class 4	Total
Events	131	1290	31	2376	3828

(a)

Predicted Events			
High-Rate	True High-Rate	False High-Rate	
Low-Rate	False Low-Rate	True Low-Rate	
	High-Rate	Low-Rate	Observed Events

(b)

Figure 2. (a) Classes of solar power ramp events; (b) confusion matrix of possible cases of solar power ramp events

used in other applications [9], [18]. The Diff. Index is the difference between the True and False events of high-rate ramps. The Total Accuracy metric gives equal weighting to all classes, and therefore, is affected by the low-rate events. Precision is the ratio of the true events of forecasted class to the total events of the same class in the forecasts. Recall is the ratio of the true events of forecasted class to total events of the same class in the observations; it does not consider the false events of the observed class as does Precision. The Balance Precision metric is the average of the precision of each class; it is also slightly impacted by the low-rate ramps. The F_1 score mitigates the fact that the Recall and the Precision metrics share an inverse relationship to each other. The most suitable metrics for our application are the Diff. (True-False) and the F_1 score.

V. Case Study

A. Data Description

The solar PV power system is located in Australia. The weather forecasts data and the measured solar power data were recorded from April 2012 to May 2014. The weather forecasts are from European Center for Medium-Range Weather Forecasts (ECMWF), which is a global numerical weather prediction (NWP) model. The data are publicly available; refer to [19].

This study presents classification techniques to classify and forecast the solar power ramp events. The objective of implementing the classification techniques for the solar power ramp events forecasting is to increase the true events and decrease the false events of high-rate ramp events. The moving time window of the rolling forecasts of solar power ramp events is 1 hour (i.e., D, duration=1 hr). The forecasts of solar power ramp events over an entire year are generated and several evaluation metrics are used to assess the forecasts of the ramp events of solar power. The test part of the data contains 12 months - from May 2013 to June 2014. After filtering the night hours and very low ramps at sunrise and sunset, the total ramp events of the test data include 3828 occasions. The classes are defined by the direction and rate of the ramps while the ramp duration is 1-hour, which is also the forecast horizon. As shown in Fig. 2(a), the solar power ramp events are classified into four classes and the distribution of the ramp events for each class based on a threshold ramp rate of 0.4 per unit/hr of the rating capacity. The threshold (0.4 pu/hr) is chosen as a mid-level value of ramp rates since the maximum rate is 0.76 pu/hr, and it is also above the ramp rates of the daily normal ramp events which occur in the mornings and late afternoons. Setting the threshold to the medium value of ramp rates is also common in wind power ramp events forecasting [10]. The confusion matrix of the possible cases of the classification is shown in Fig. 2(b).

Table I
(A) The Most important features; (B) the selected features and the parameters for each model

(A)

No.	Most Important Features
1	Cloud water content, NWP output
2	Cloud cover, NWP output
3	Top net solar radiation, NWP output
4	Hour-ahead combined forecasts of solar power
5	Ramp rates of NWP-driven day-ahead solar power forecasts by ANN
6	Ramp rates of NWP-driven day-ahead solar power forecasts by SVR
7	Ramp rates of hour-ahead combined forecasts of solar power
8	Ramp rates of time-series hour-ahead forecasts of solar power by NARX
9	Ramp classes of persistence hour-ahead forecasts of solar power
10	Ramp classes of NWP-drive day-ahead solar power forecasts by ANN
11	Ramp classes of NWP-driven day-ahead solar power forecasts by SVR
12	Ramp classes of hour-ahead combined forecasts solar power

(B)

Model	Parameters	Selected Features
Naïve Bayes	Distribution=Normal; distribution parameters are estimated in the training.	1, 5, 11
LDA	Its coefficients (μ) are fitted in the training.	1, 2, 3, 6, 9, 10, 12
Decision Tree	Max of splits=15; Min leaf size=1	1, 12
kNN	Euclidean distance; k=15 (nearest 15 neighbors)	1, 4, 6, 7, 8
Logistic Regression	Its coefficients (β) are fitted in the training.	1, 3, 11, 12
Random Forests	Forest size=100 trees; Min. leaf size=1	1, 3, 11, 12
SVM	Kernel= Radial basis function; C=184; gamma=5	1, 3, 11, 12
ANN	Hidden layer=1; Neurons=10	1, 5, 12

B. Results and Evaluation

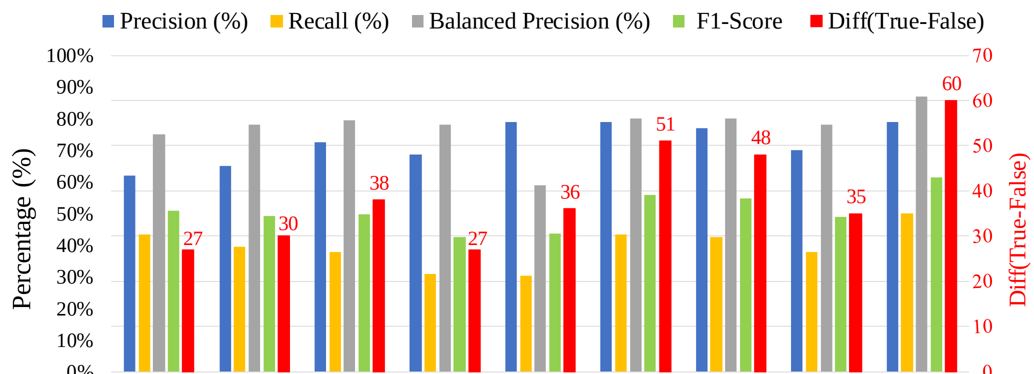
Based on the wrapper approach to search the best features for the classification models, the most effective features were found, and these are shown in Table I(A). The selected features and the parameters of each classification model are given in Table I(B), the features are represented by numbers associated with them as in Table I(A).

Table II shows the results of implementing the classification techniques to forecasts the solar power ramp events. Fig. 3 displays a graph of the results of the classification techniques in terms of different evaluation metrics for the 162 high-rate ramp events. There are 3828 total ramp events identified, 3666 of them are low-rate and 162 are high-rate ramp events.

M. Abuella and B. Chowdhury, "Forecasting Solar Power Ramp Events Using Machine Learning Classification Techniques," in 2018 IEEE 8th International Symposium on Power Electronics for Distributed Generation Systems (PEDG), 2018, pp. 16.

Table II
DETAILED RESULTS OF THE SOLAR POWER RAMP EVENTS FORECASTS BY THE CLASSIFICATION TECHNIQUES

Classification Technique	All True Events	High-Rate True Events	All False Events	High-Rate False Events	Total Accuracy (%)	Precision (%)	Recall (%)	Balanced Precision (%)	F1 score (%)	Diff. index (True-False)
Naïve Bayes	3165	70	663	43	83%	62%	43%	75%	51%	27
LDA	3288	64	540	34	86%	65%	40%	78%	49%	30
Decision Trees	3077	61	751	23	80%	73%	38%	80%	50%	38
kNN	3200	50	628	23	84%	68%	31%	78%	43%	27
Logistic Regression	3102	49	726	13	81%	79%	30%	59%	44%	36
Random Forest	3118	70	710	19	81%	79%	43%	80%	56%	51
SVM	3125	69	703	21	82%	77%	43%	80%	55%	48
ANN	3259	61	569	26	85%	70%	38%	78%	49%	35
Combined Classifiers	3309	81	519	21	86%	79%	50%	87%	61%	60
Out of	3828	162	3828	3666						



Method	Naïve Bayes	LDA	Decision Trees	kNN	Logistic Regression	Random Forest	SVM	ANN	Combined Classifiers
Precision (%)	62%	65%	73%	68%	79%	79%	77%	70%	79%
Recall (%)	43%	40%	38%	31%	30%	43%	43%	38%	50%
Balanced Precision (%)	75%	78%	80%	78%	59%	80%	80%	78%	87%
F1-Score (%)	51%	49%	50%	43%	44%	56%	55%	49%	61%
Diff(True-False)	27	30	38	27	36	51	48	35	60

Figure 3. Solar power ramp events forecasts by the classification techniques of the high-rate ramp events (162 events)

Since the direction of ramps are easily forecasted with high accuracy, they are combined for the evaluation metrics for a binary classification problem with two main classes, as follows: high-rate events include both up and down ramp events. Low-rate event also include both up and down ramp events. The Diff. (true-false) is represented by integers, not percentages, so that its scale is shown on the right hand side y-axis on the graph.

For the individual calcification methods, the best forecasts are obtained by random forest (RF) and SVM, where the Diff Index = 48 and F_1 score = 55% for SVM, and Diff Index = 51 and F_1 score = 56% for RF. The combined forecasts of the different classification methods yields the best performance with Diff Index = 60 and F_1 score = 64%.

As shown in Table II, in terms of the precision, the logistic regression model gives a competitive precision when compared to the RF model, since it has a lower number of false events with respect to true events of high-rate classes. On the other hand, in terms of the recall, the poor performance of the logistic regression is identified, as it captures a lower number of true events of high-rate classes compared to the RF. Thus, for an overall assessment of the classification model performance, evaluation metrics such as the Diff. Index and F_1 score are more useful.

VI. CONCLUSION

Several classification techniques are implemented to forecast the solar power ramp events by using features including solar power forecasts and weather predictions. This study presents the challenging aspect of ramp forecasting, and it is not comparable to studies that detect the ramp events by using historical solar power observations and meteorological measurements. For a general assessment of the classification model performance for solar power ramp events forecasting, the evaluation metrics that consider the precision and the recall together, such as Diff. Index and F_1 score, should be used, in order to properly weigh both the true and the false events of high-rate ramp events.

In the individual classification models, the RF and SVM models yield the most accurate forecasts of solar power ramp events. In addition, combining the outcomes of the models improves the accuracy and leads to a more robust performance. The classification techniques (i.e., RF, SVM, and combined classifiers) outperform the solar power forecasts that are used as features to these classification techniques, and hence, this is one of the advantages of using the classification techniques with several solar power forecasts as inputs.

References

[1] J. Kleissl, *Solar Energy Forecasting and Resource Assessment*. Elsevier, 2013.

[2] J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno, *Integrating renewables in electricity markets - Operational problems*, vol. 205 of *International Series in Operations Research & Management Science*. Boston, MA: Springer US, 2014.

[3] B.-I. Craciun, T. Kerekes, D. Sera, R. Teodorescu, and U. D. Annakkage, "Power Ramp Limitation Capabilities of Large PV Power Plants With Active Power Reserves," *IEEE Transactions on Sustainable Energy*, vol. 8, pp. 573–581, apr 2017.

[4] A. Tuohy, 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, "Solar Forecasting: Methods, Challenges, and Performance," *IEEE Power and Energy Magazine*, vol. 13, no. 6, pp. 50–59, 2015.

[5] A. Florita, B.-M. Hodge, and K. Orwig, "Identifying wind and solar ramping events," in *Green Technologies Conference, 2013 IEEE*, pp. 147–152, IEEE, 2013.

[6] Y. Hirata and K. Aihara, "Predicting ramps by integrating different sorts of information," *The European Physical Journal Special Topics*, vol. 225, no. 3, pp. 513–525, 2016.

[7] M. J. Reno and C. W. Hansen, "Identification of periods of clear sky irradiance in time series of GHI measurements," *Renewable Energy*, vol. 90, pp. 520–531, 2016.

[8] Y. Chu, H. T. C. Pedro, M. Li, and C. F. M. Coimbra, "Real-time forecasting of solar irradiance ramps with smart image processing," *Solar Energy*, vol. 114, pp. 91–104, 2015.

[9] C. Ferreira, J. Gama, L. Matias, A. Botterud, and J. Wang, "A survey on wind power ramp forecasting.," tech. rep., Argonne National Laboratory (ANL), 2011.

[10] C. Gallego-Castillo, A. Cuerva-Tejero, and O. Lopez-Garcia, "A review on the recent history of wind power ramp forecasting," *Renewable and Sustainable Energy Reviews*, vol. 52, pp. 1148–1157, dec 2015.

[11] H. Zheng and A. Kusiak, "Prediction of Wind Farm Power Ramp Rates: A Data-Mining Approach," *Journal of Solar Energy Engineering*, vol. 131, no. 3, p. 031011, 2009.

[12] R. Sevlian and R. Rajagopal, "Detection and statistics of wind power ramps," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3610–3620, 2013.

[13] C. Feng, M. Cui, B.-m. Hodge, and J. Zhang, "A data-driven multi-model methodology with deep feature selection for short-term wind forecasting," *Applied Energy*, vol. 190, pp. 1245–1257, 2017.

[14] G. James, D. Witten, T. Hastie, and R. Tibishirani, *An Introduction to Statistical Learning*, vol. 112. 2013.

[15] T. Hastie, R. Tibshirani, J. Friedman, and Others, *The elements of statistical learning*. Springer-Verlag New York, 2 edition ed., 2009.

[16] M. Abuella and B. Chowdhury, "Improving Combined Solar Power Forecasts Using Estimated Ramp Rates: Data-driven Post-processing Approach," *IET Renewable Power Generation*, 2018.

[17] M. R. Hossain, A. M. T. Oo, and A. Ali, "The effectiveness of feature selection method in solar power prediction," *Journal of Renewable Energy*, vol. 2013, 2013.

[18] A. Dal Pozzolo, O. Caelen, R. A. Johnson, and G. Bontempi, "Calibrating probability with undersampling for unbalanced classification," in *Computational Intelligence, 2015 IEEE Symposium Series on*, pp. 159–166, IEEE, 2015.

[19] "Global Energy Forecasting Competition 2014, Probabilistic solar power forecasting." [Online]. Available: http://www.crowdanalytix.com/contests/global-energy-forecasting-competition-2014.