A Post-processing Approach for Solar Power Combined Forecasts of Ramp Events

Mohamed Abuella

Supervised by:
Prof. Badrul Chowdhury

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Electrical Engineering

Charlotte
October 1st, 2018

UNC CHARLOTTE
Energy Production and Infrastructure Center (EPIC)

Dissertation Organization

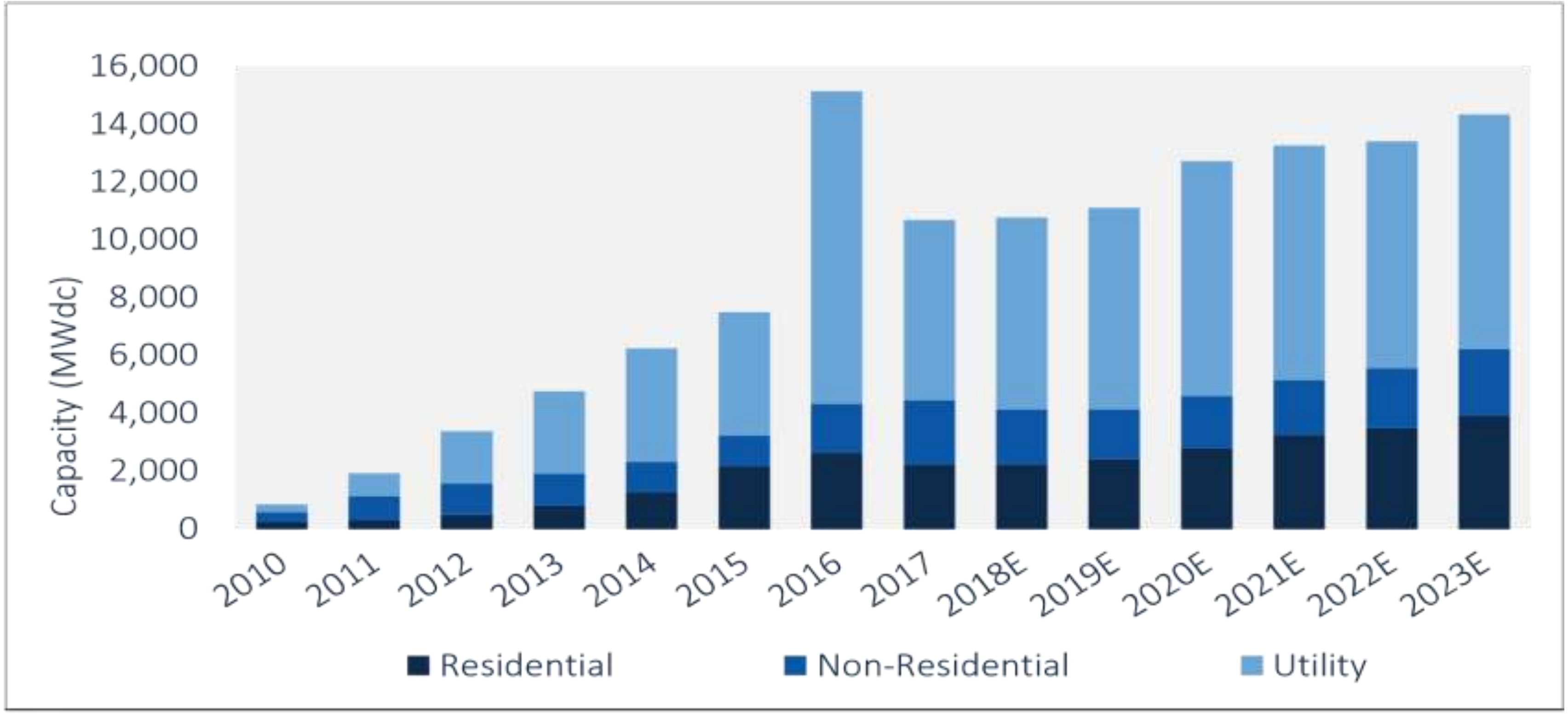
Dissertation Layout		Presentation Outline
Chapter 1 Motivation and Problem Overview	It covers the motivation, problem statement and contribution, as well as the literature review	Highlights, Motivation, Literature Review, Problem Statement & Contribution, Theoretical Background & Methodology
Chapter 2 Theoretical Background	Theoretical background for modeling and forecasting of solar power ramp events	
Chapter 3 Improving Combined Solar Power Forecasts	Applying the adjusting post-processing approach to improve the hourly combined forecasts of solar power	Modeling and Results
Chapter 4 Forecasting of Solar Power Ramp Events	Applying the adjusting post-processing approach to improve the hourly combined forecasts of solar power ramp events	
Chapter 5 Intra-Hour Forecasts of Solar Power and Ramp Events	Applying the adjusting post-processing approach to improve the sub-hourly combined forecasts of solar power and solar power ramp events	
Chapter 6 Conclusions and Future Work	Final conclusions and recommendations of future work	Conclusions and Future Work

https://www.researchgate.net/publication/328757565_A_Post-processing_Approach_for_Solar_Power_Combined_Forecasts_of_Ramp_Events/download

Highlights

- A post-processing approach combines and improves solar power forecasts.
- The approach also adjusts the combined forecasts in terms of ramp events.
- A classification of all possible thresholds and classes of ramp event forecasts.
- A customized cost function for imbalanced classification of ramp events.
- Suitable metrics for the feature selection process and performance evaluation.
- An uncertainty analysis for probabilistic forecasts of solar power ramp events.

Motivation



Source: SEIA/GTM Research

A U.S. PV solar market study * prepared by Solar Energy Industries Association (SEIA) and GTM Research

https://www.seia.org/us-solar-market-insight (June 12, 2018, insight of US Solar Market)

Motivation

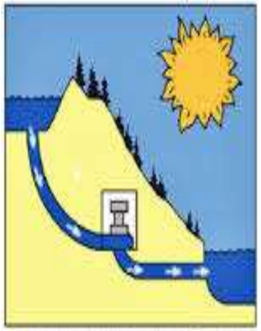
Why Forecast?

$$P_{Supply} = P_{Demand} + P_{Loss}$$

PV Solar Power Generations are Too Variable

Coordination with Operating Reserves and Energy Storage Systems

Reducing Cost and Pollution



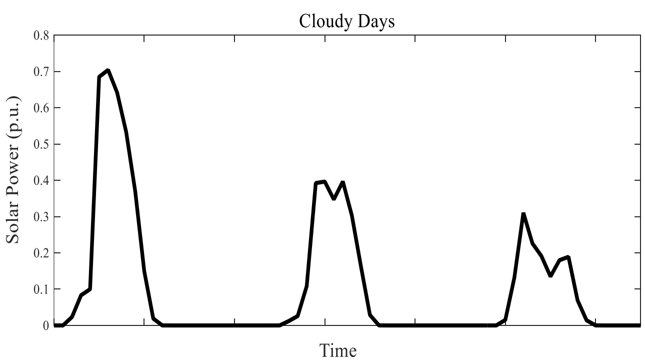
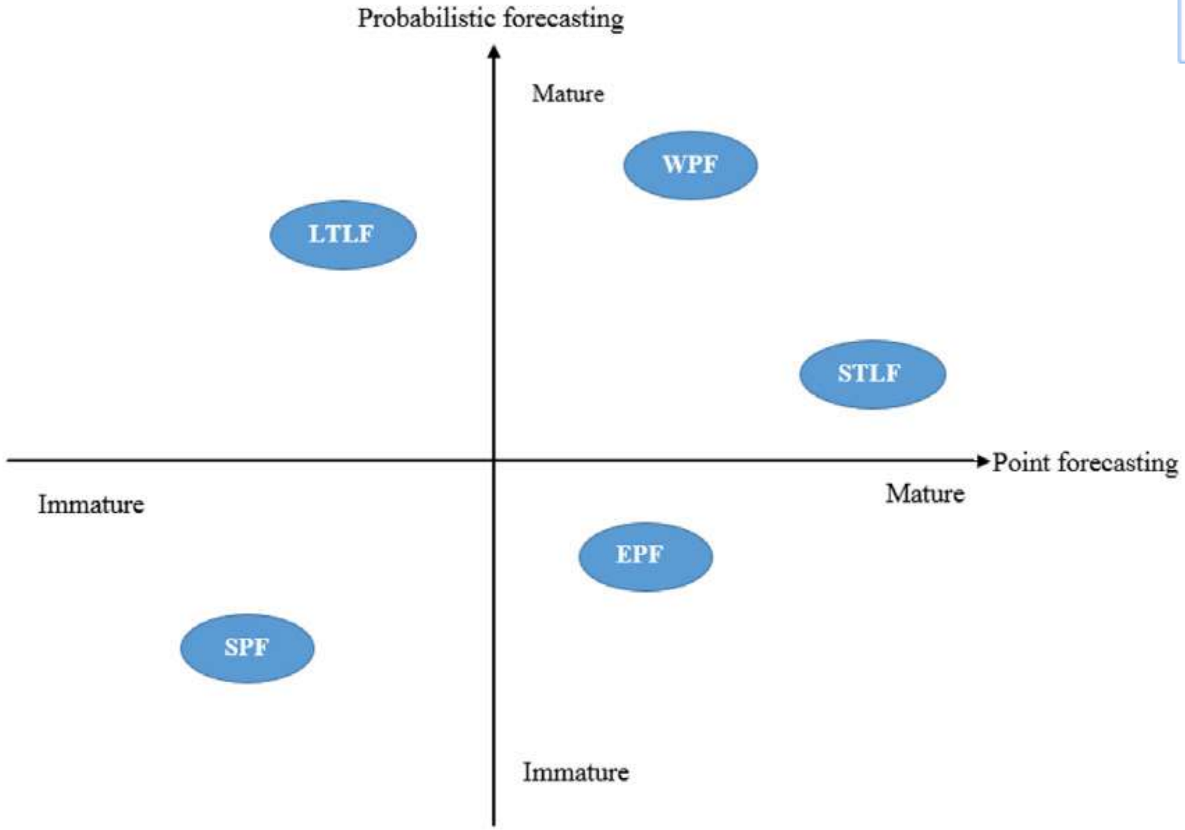


Illustration of the motivation of PV solar power forecasts

Motivation

Challenges



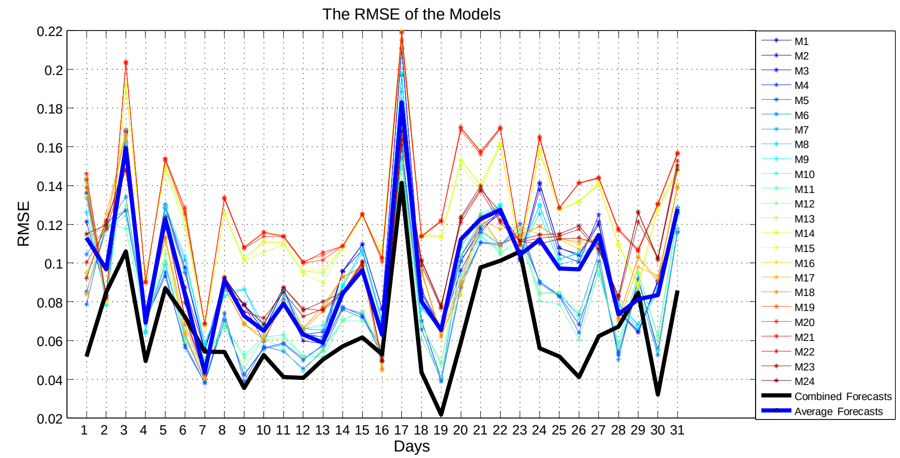
Maturity quadrant of the energy forecasting subdomains (SPF: solar power forecasting; LTLF: long term load forecasting; EPF: electricity price forecasting; WPF: wind power forecasting; STLF: short term load forecasting) .

Hong, Tao, et al. "Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond." *International Journal of Forecasting* 32.3 (2016): 896-913.

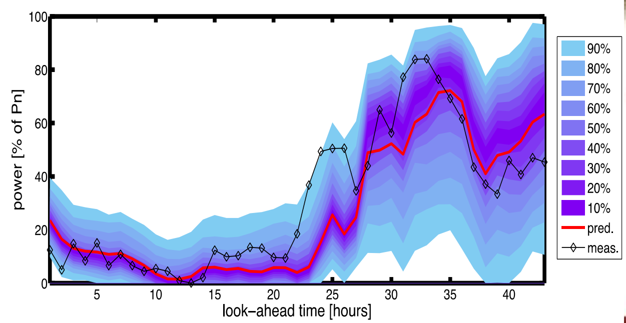
Motivation

Challenges

These are the objectives of the ongoing research on this field that are addressed and recommended for the academy and industry.



- -Accurate V.G. Forecasting;
- -Ensemble Forecasts;
- -Probabilistic Forecasts;
- -Ramp Event Forecasts;
- -Better Evaluation Methods for Forecasts;
- -Forecasts for PV Solar Distributed Generation Systems “Behind-the-Meter Resources”;
- -Study the economic value of V.G. forecasting;
- -Electricity System Operation Adjustments for Optimal Implementation of the Forecast.

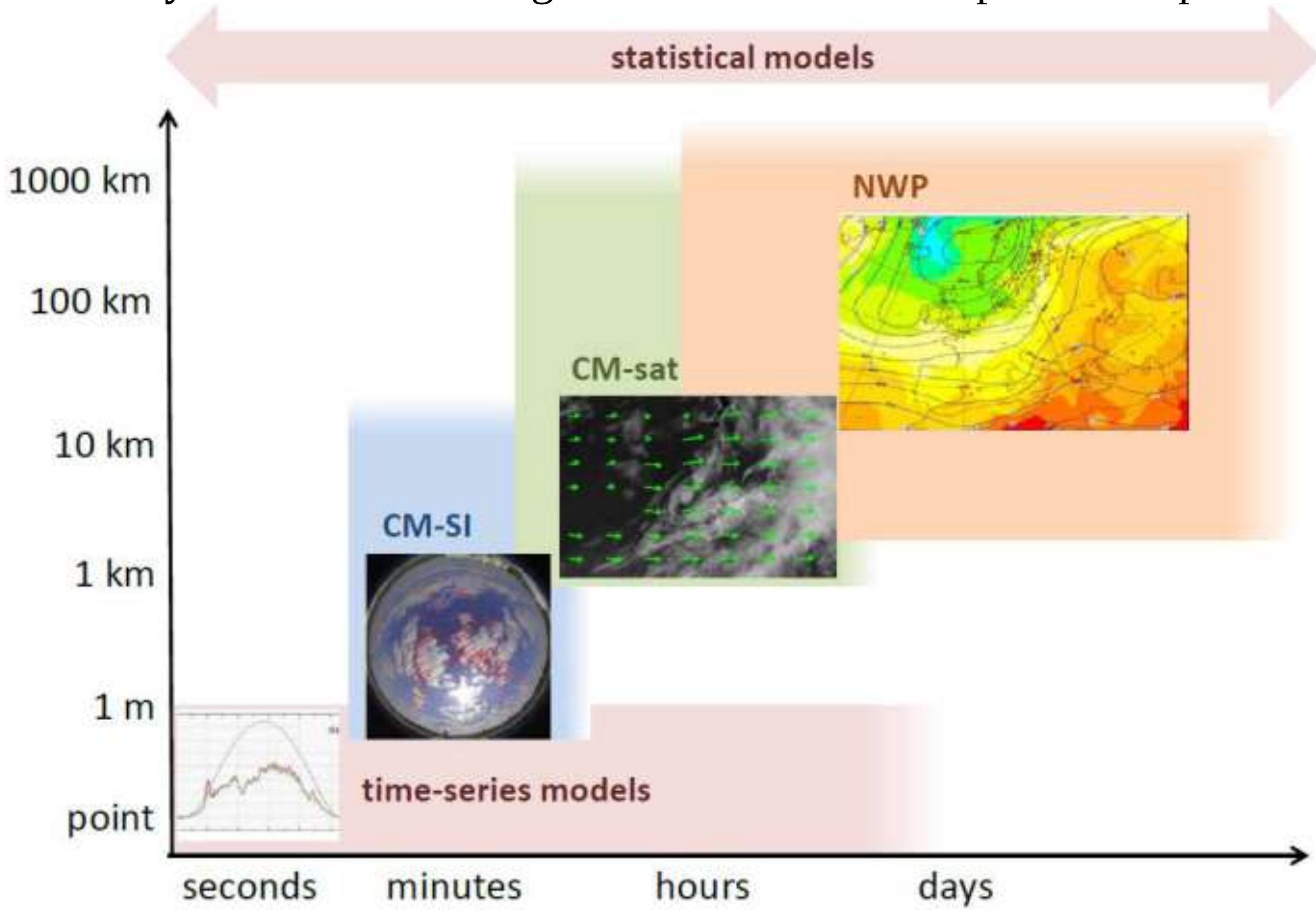


Solar Forecasting: Methods, Challenges, and Performance (IEEE Power & Energy Magazine Nov. 2015)
By Aidan Tuohy, John Zack, Sue Ellen Haupt, Justin Sharp, Mark Ahlstrom, Skip Dise, Eric Grimit, Corinna Möhrlen, Matthias Lange, Mayte Garcia Casado, Jon Black,Melinda Marquis, and Craig Collier.

7

Literature Review

Taxonomy of solar forecasting methods based on temporal and spatial resolution#



CM-sat: cloud motion by satellite images;

CM-SI: cloud motion by sky-imagers

NWP: Numerical weather prediction Systems

Statistical models: blending and post-processing all of the methodologies.

M. Sengupta, A. Habte, C. Gueymard, S. Wilbert, and D. Renne, Best practices handbook for the collection and use of solar resource data for solar energy applications," Tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States), 2017

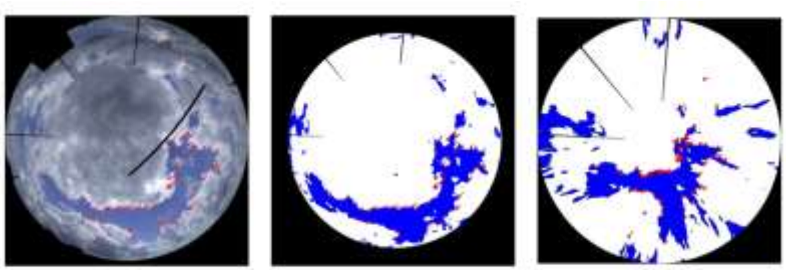
Yang, D., Kleissl, J., Gueymard, C. A., Pedro, H. T., & Coimbra, C. F. (2018). History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining. *Solar Energy*. (**This review paper published in **2018,** and reviewing **1000** solar forecast studies).

Literature Review

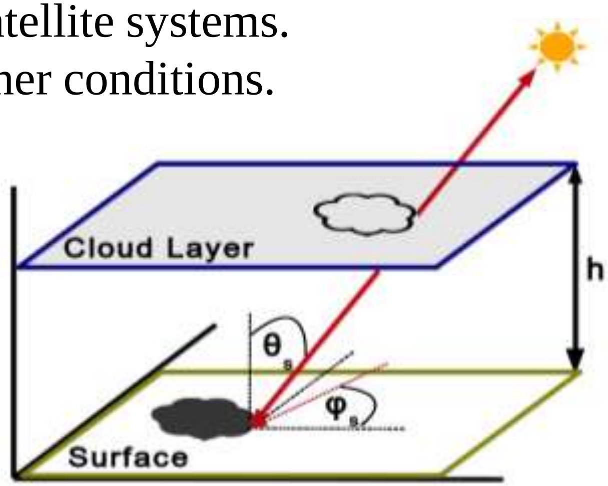
Main methods for solar irradiance / power ramp events:

1) *Forecast Approaches:*

a) Physical models to track the cloud motion: sky imagers and satellite systems.
b) Statistical models that are selected based on forecasts of weather conditions.



Raw and projected cloud map form sky imaging devices [R.1]



Shadow projection by sky imagine devices [R.1]

2) *Identification/Detection Approaches:*

a) Anomaly detection methods, such as swinging door algorithm [R.4].
b) Classification by using machine learning techniques.

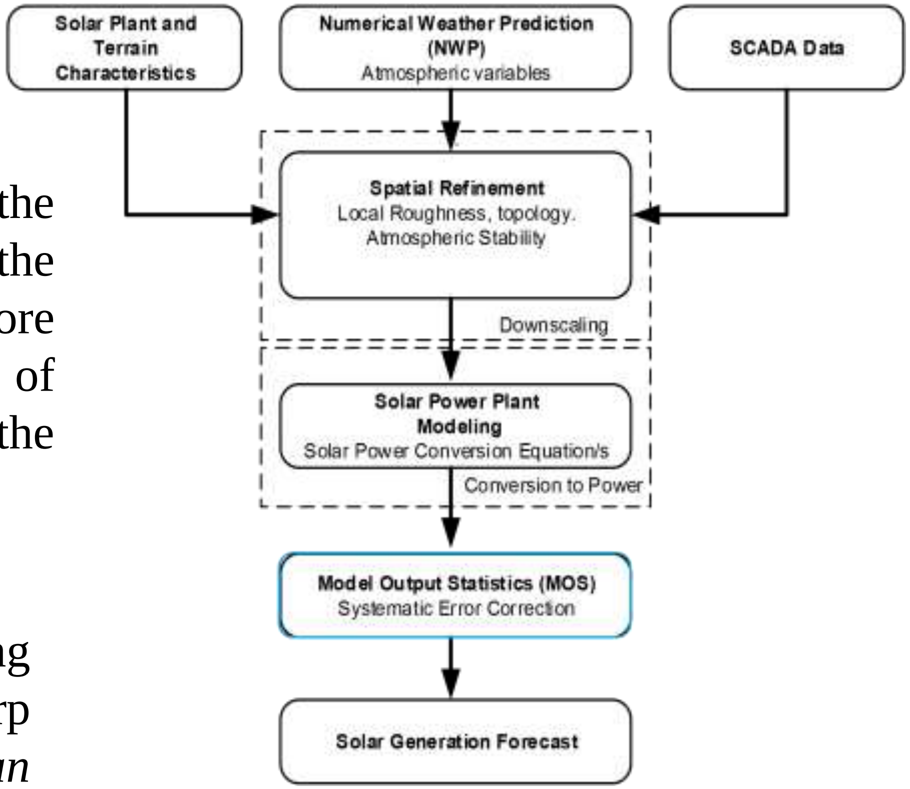
[R.1] Schmidt, T., Calais, M., Roy, E., Burton, A., Heinemann, D., Kilper, T., & Carter, C. (2017). Short-term solar forecasting based on sky images to enable higher PV generation in remote electricity networks. *Renewable Energy and Environmental Sustainability.*
[R.2] Palmer, D., Koubli, E., Cole, I., Betts, T., & Gottschalg, R. (2018). Satellite or ground-based measurements for production of site specific hourly irradiance data: Which is most accurate and where?. *Solar Energy*, *165*, 240-255.
[R.3] Bessa, R. J., Trindade, A., & Miranda, V. (2015). Spatial-temporal solar power forecasting for smart grids. *IEEE Transactions on Industrial Informatics*, *11*(1), 232-241.
[R.4] Cui, M., Zhang, J., Florita, A., Hodge, B. M., Ke, D., & Sun, Y. (2015, August). Solar power ramp events detection using an optimized swinging door algorithm. In *ASME 2015 International Design Engineering and Computers and Information in Engineering Conference*.

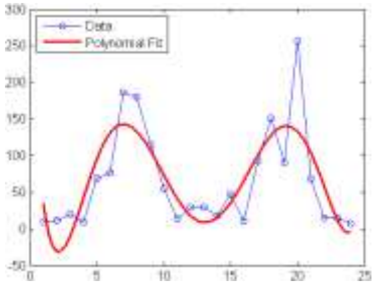
Problem Statement and Contribution

The Problem Statement

Combining of different forecasts can reduce the systemic bias of the individual models, boost the overall accuracy, and make the performance more robust, but it also smooths out the sharp changes of the forecasts, which leads to reduced accuracy of the combined forecasts for ramp events.

The post-processing methods (MOS) are affecting the ramp event forecasts by smoothing the sharp changes in the raw forecasts. (*Kleissl 2013; Inman et al. 2013*)





Kleissl, J. (2013). Solar energy forecasting and resource assessment. Academic Press.

Inman, R. H., Pedro, H. T., & Coimbra, C. F. (2013). Solar forecasting methods for renewable energy integration. Progress in energy and combustion science, 39(6), 535-576.

Problem Statement and Contribution

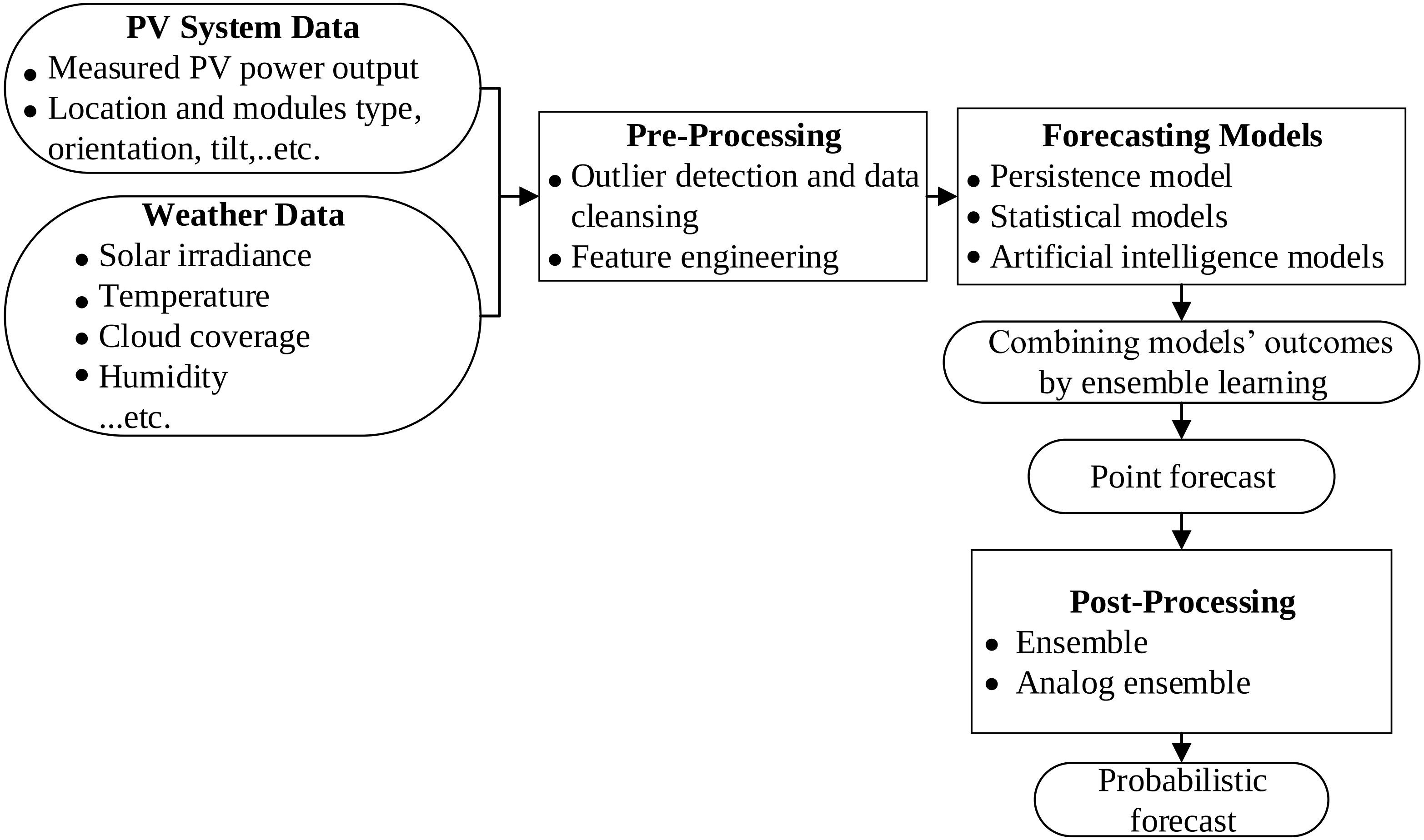
The Contribution

Due to the issue that was highlighted (*Kleissl 2013; Inman et al. 2013*), it is therefore, there is room for improvement by applying the proposed approach for adjusting the combined forecasts of solar power in terms of ramp events.

To the best of our knowledge, this is the first attempt to tackle this issue of the combined forecasts for solar power ramp events.

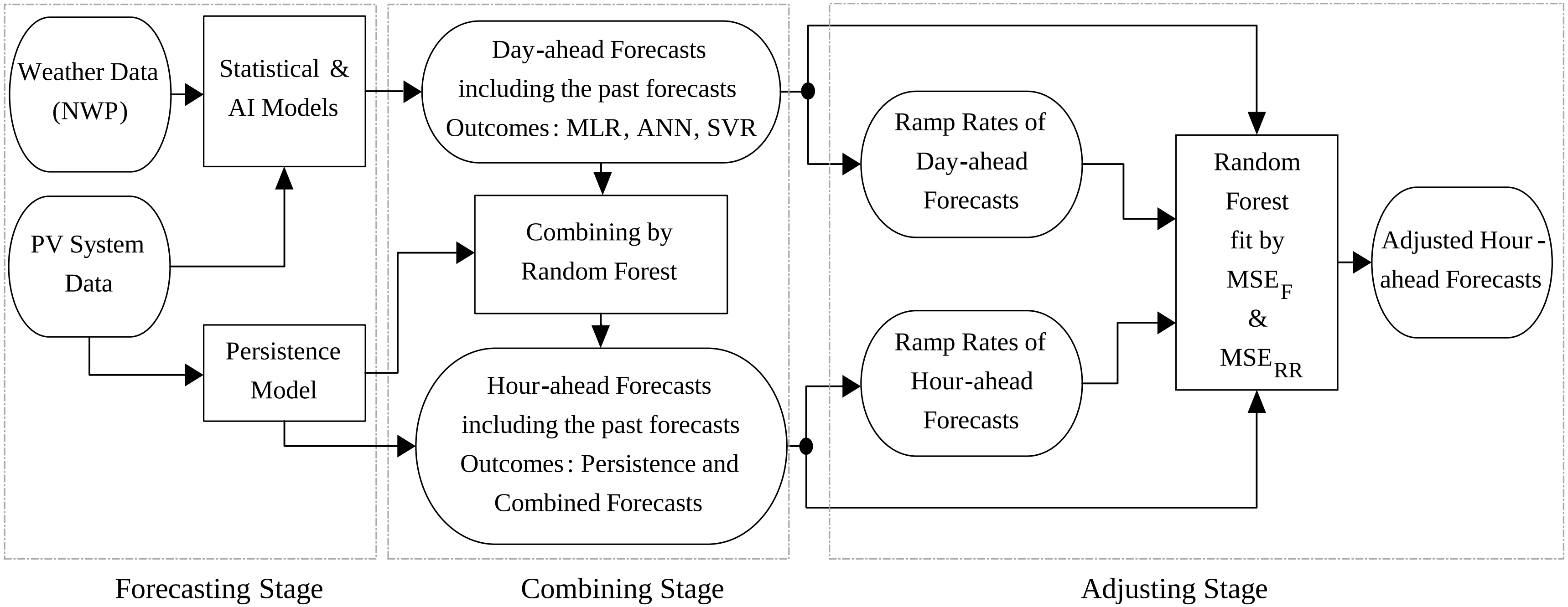
- A post-processing approach combines and improves solar power forecasts.
- The approach also adjusts the combined forecasts in terms of ramp events.
- A classification of all possible thresholds and classes of ramp event forecasts.
- A customized cost function for imbalanced classification of ramp events.
- Suitable metrics for the feature selection process and performance evaluation.
- An uncertainty analysis for probabilistic forecasts of solar power ramp events.

Graphical Abstract of the Proposed Adjusting Approach



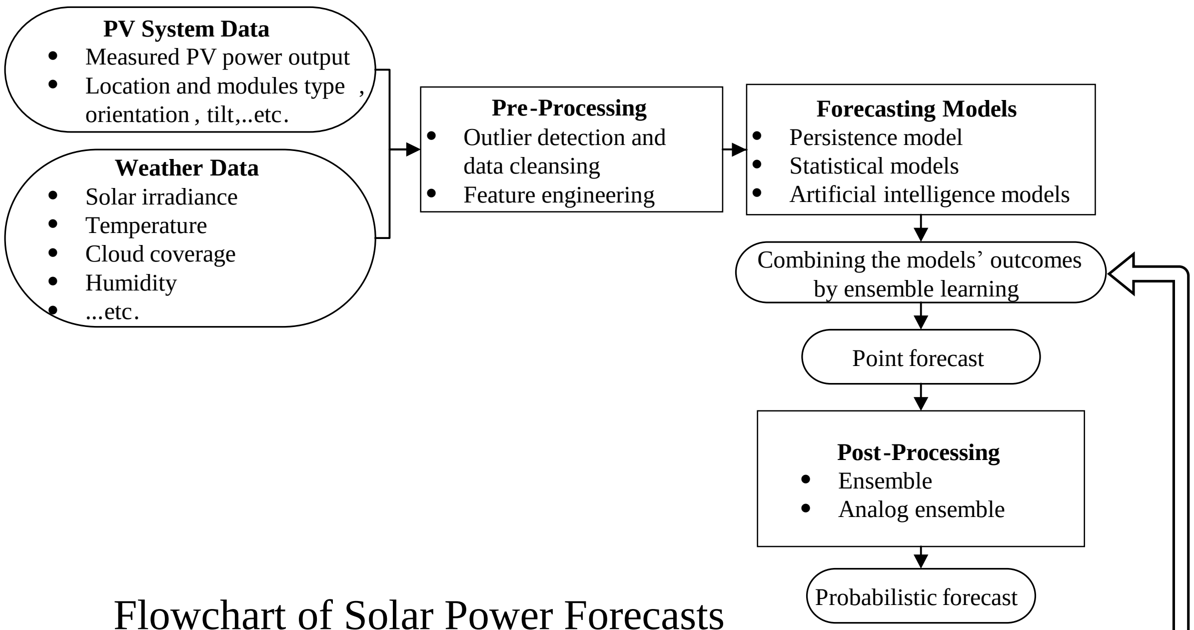
Flowchart of Solar Power Forecasts

Graphical Abstract of the Proposed Adjusting Approach

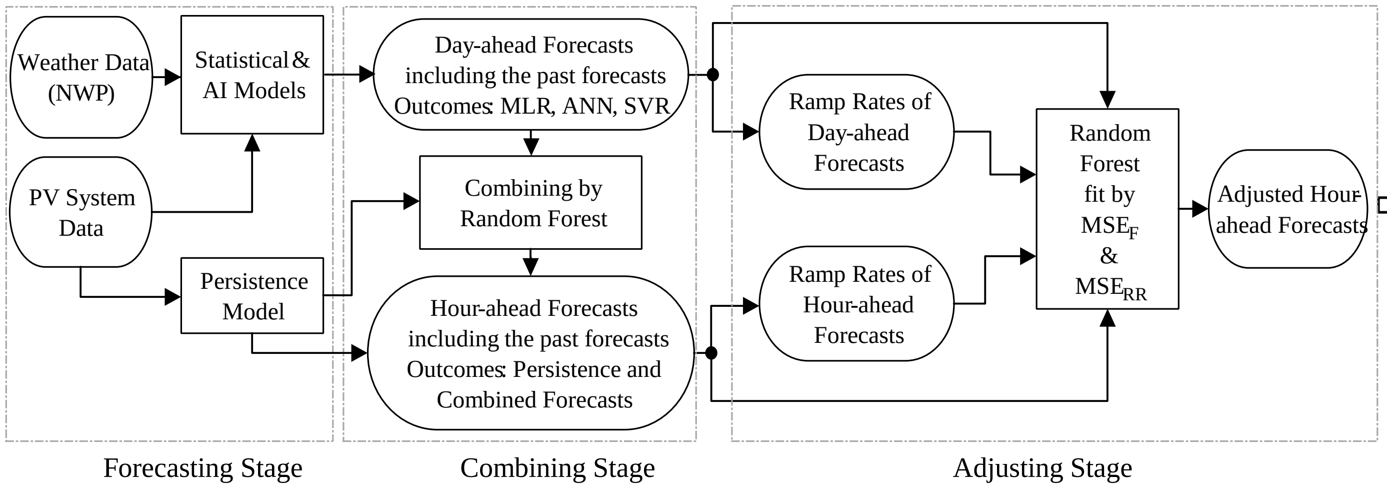


Block diagram of the adjusting approach

Graphical Abstract of the Proposed Adjusting Approach



Flowchart of Solar Power Forecasts

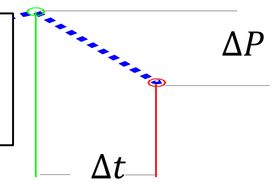


Block diagram of the adjusting approach

Solar Power Ramp Rates

Solar power ramp rate (RR) is *the change of solar power during a certain time interval***.**

$$Ramp\ Rate, RR(t) = \frac{dP(t)}{dt} = \frac{P(t+D) - P(t)}{D}$$



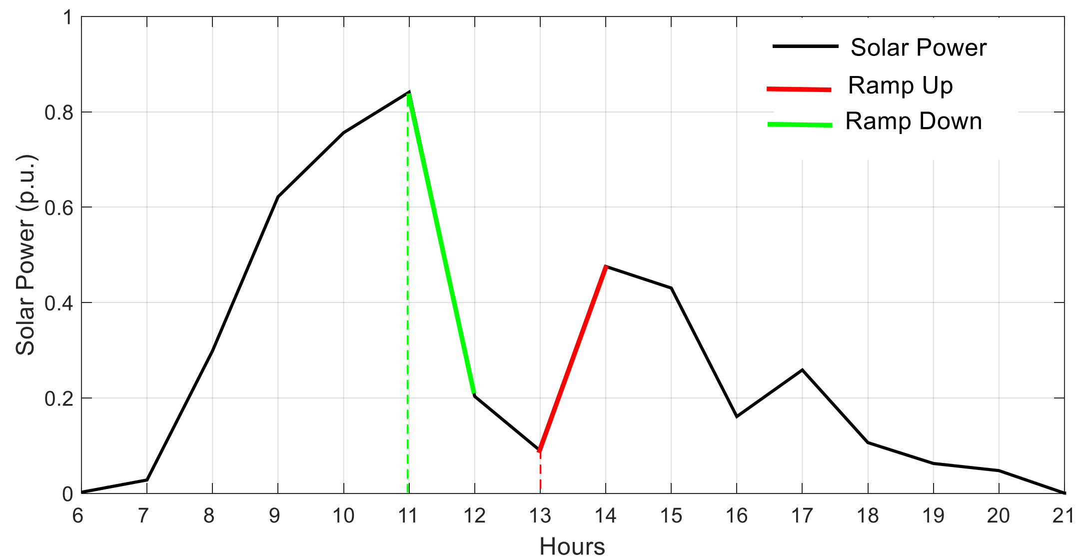
where P(t) is the solar power of the target hour, it can also be its forecast F(t); D is the time duration for which the ramp rate is determined.

For the illustrated cloudy day below:

Ramp rate, $\frac{\Delta P}{\Delta t} = \frac{0.2 - 0.85}{12{:}00 - 11{:}00} = -0.65\ (-65\%)\ ramp\ down\ of\ its\ normal\ capacity,\ (pu/hr)$

Ramp rate, $\frac{\Delta P}{\Delta t} = \frac{0.48 - 0.1}{14{:}00 - 13{:}00} = +0.38\ (+38\%)\ ramp\ up\ of\ its\ normal\ capacity,\ (pu/hr)$

Some ramps are with low rates, while others with high rates.



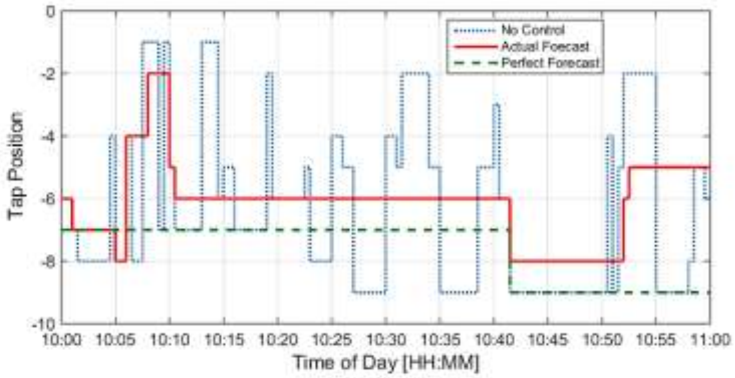
Ramp Events During a Cloudy Day

Potential Applications

There are several applications of power systems that rely on solar power ramp event forecasts

Distribution level:

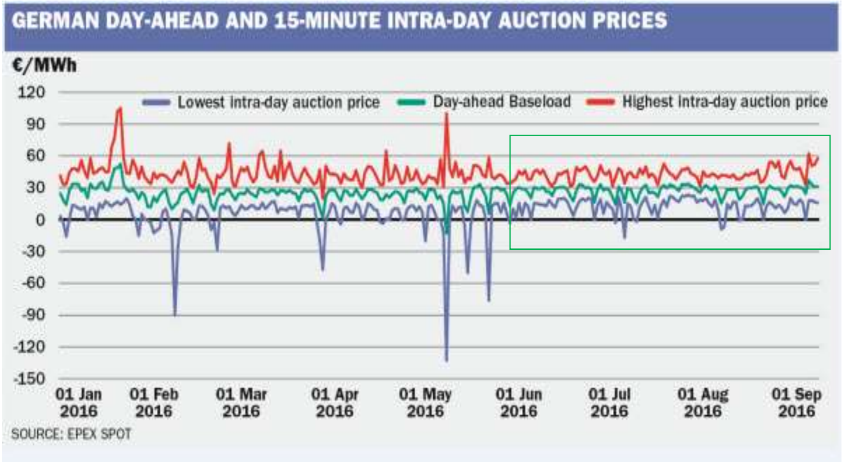
- Optimizing the voltage regulation equipment.
- Control schemes of energy storage systems.



Optimizing the Transformer's Tap Changer position sequences using the solar forecast

Transmission / bulk level:

- Trading & dispatching the operating reserve.
- Managing the ramp capability / system flexibility with high-level of renewable energy integration.



EPEX: European power exchange spot trading

Data Preprocessing

Data Description:

Dataset	Golden, CO	Cocoa, FL	Eugene, OR	Canberra
Country	USA	USA	USA	Australia
Climate type	Semi-arid	Subtropical	Marine coast	Oceanic
Latitude (°, -S)	39.74	28.39	44.05	-35.16
Longitude (°, -W)	-105.18	-80.46	-123.07	149.06
Elevation above sea (m)	1798	12	145	595
Number of panels	11	11	11	8
Panel tilt (°) from horizontal	40	28.5	44	36
Panel orientation (°) clockwise from North	180	180	180	38
Total capacity (W)	1252	1272	1290	1560
Time period of observations	Aug. 2012 to Sep. 2013	Jan. 2011 to March 2012	Dec. 2012 to Jan. 2014	April 2012 to May 2014
Data resolution	15min	5min	5min	1hr
Missing (% of observations)	18%	17%	10%	0%
Variability (data resolution) Std.Div.	(15min) 0.256 (1hr) 0.119	(5min) 0.251 (1hr) 0.164	(5min) 0.250 (1hr) 0.161	(1hr) 0.259

B. Marion, A. Anderberg, C. Deline, J. del Cueto, M. Muller, G. Perrin, J. Rodriguez, S. Rummel, T. J. Silverman, F. Vignola, et al., New data set forvalidating pv module performance models," in Photovoltaic Specialist Conference (PVSC), 2014 IEEE 40th, pp. 1362{1366, IEEE, 2014.

https://crowdanalytix.com/contests/global-energy-forecasting-competition-2014-probabilistic-solar-power-forecasting

Data Preprocessing

Data Description:

PV solar system is near Canberra, Australia, consisting of 8 panels, its nominal power of (1560W), and panel orientation 38° clockwise from the north, with panel tilt (of 36°). The historical observed solar power data are normalized to the rated capacity (i.e., 1560W).

Weather predictions are produced by a global numerical weather prediction system, European Centre for Medium-Range Weather Forecasts (ECMWF).

Available data

No.	Input Variable, (X)
1	Cloud Water Content
2	Cloud Ice Content
3	Surface Pressure
4	Relative Humidity
5	Cloud Cover
6	10m - U Wind
7	10m - V Wind
8	2-m Temperature
9	Surface solar radiation down

No.	Input Variable, (X)
10	Surface thermal radiation down
11	Top net solar radiation
12	Total precipitation
13	Heat Index
14	Wind Speed
15	Hours
16	Months
17	Days of Month
18	Days of Year

Data partition into training and testing sets

	Month	Year	Partition
From	April	2012	Training Set
To	May	2013	
From	June	2013	Testing Set
To	May	2014	

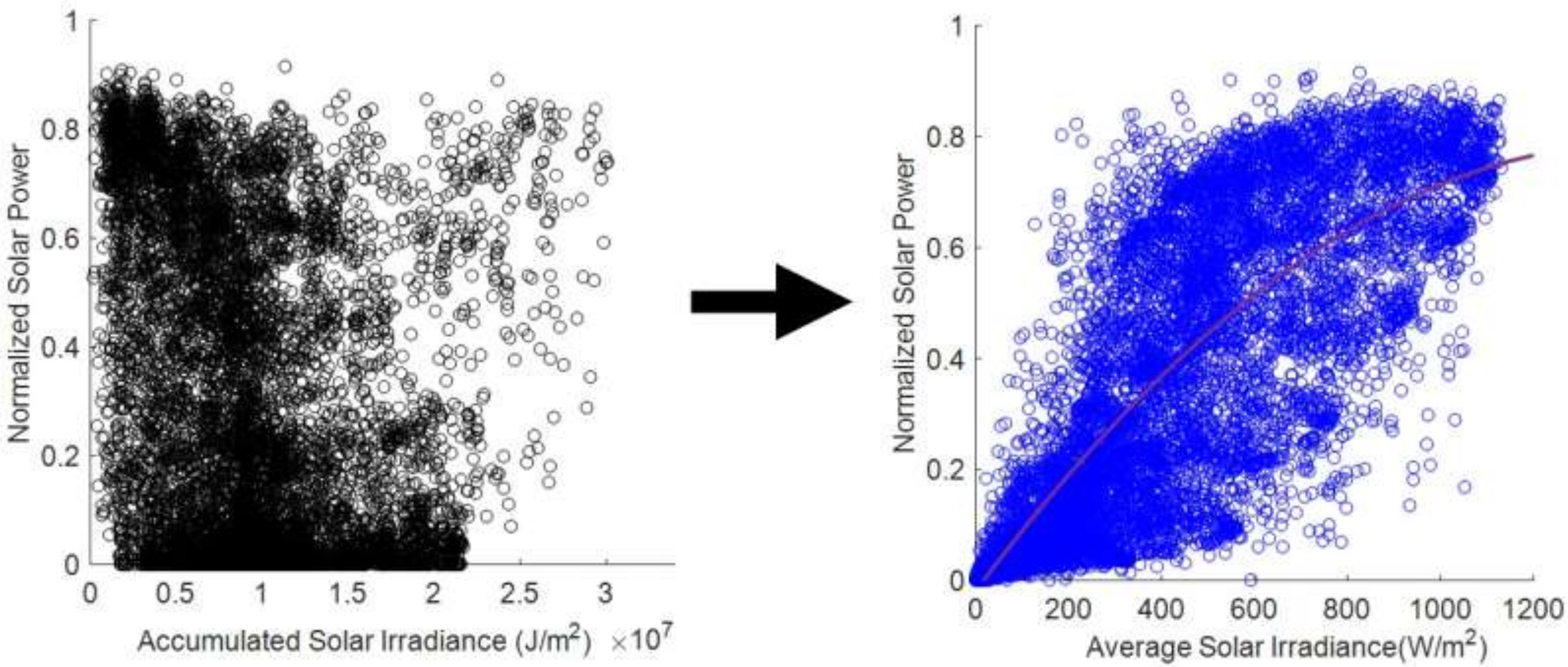
https://crowdanalytix.com/contests/global-energy-forecasting-competition-2014-probabilistic-solar-power-forecasting

Data Preprocessing

Flowchart of Data Preparation

$$Avg(t) = \frac{Accum(t+1) - Accum(t)}{3600}$$

Solar Power vs. Surface Solar Irradiance Down (SSRD) provide by NWP

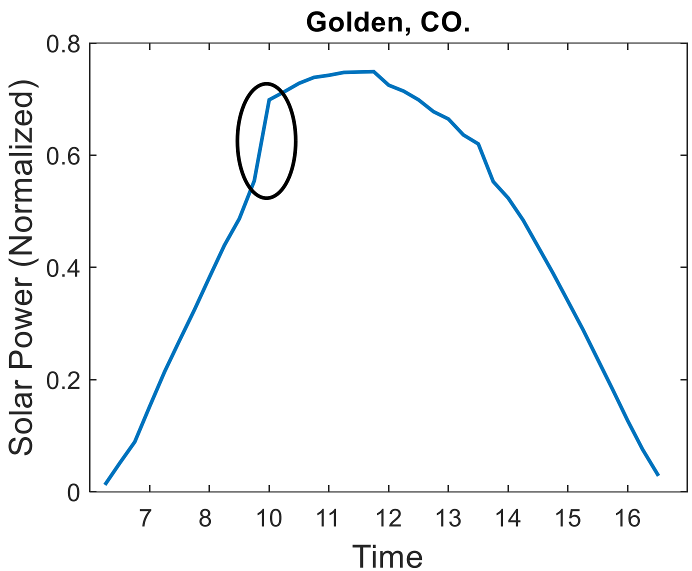


Scatter plot of the observed solar power vs. Solar Irradiance

Data Preprocessing

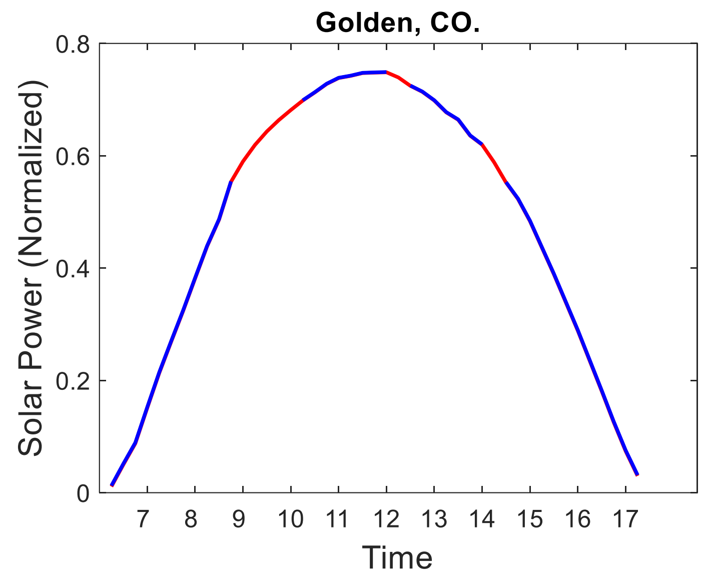
Flowchart of Data Preparation

September 24th, 2013



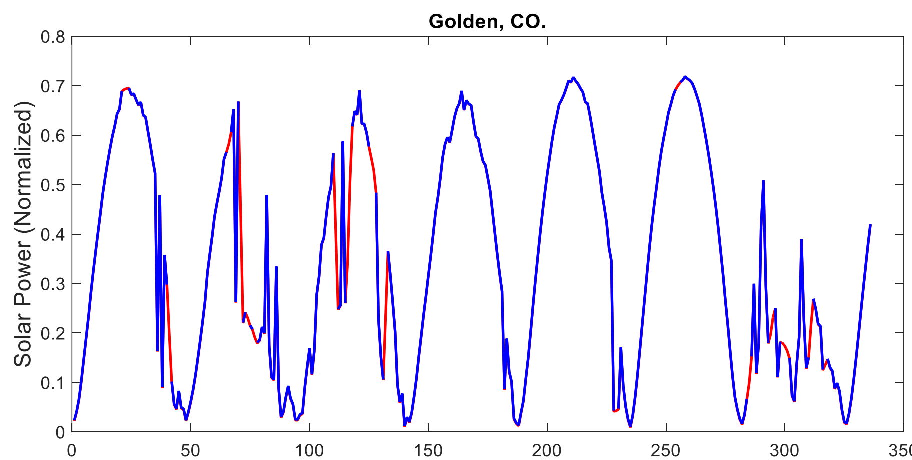
There are missing data: All minutes at hour=9 and some minutes at hour=10, 12, and 14.

Filling of missing data by interpolation



One week:

August 14th to 20th ,2012



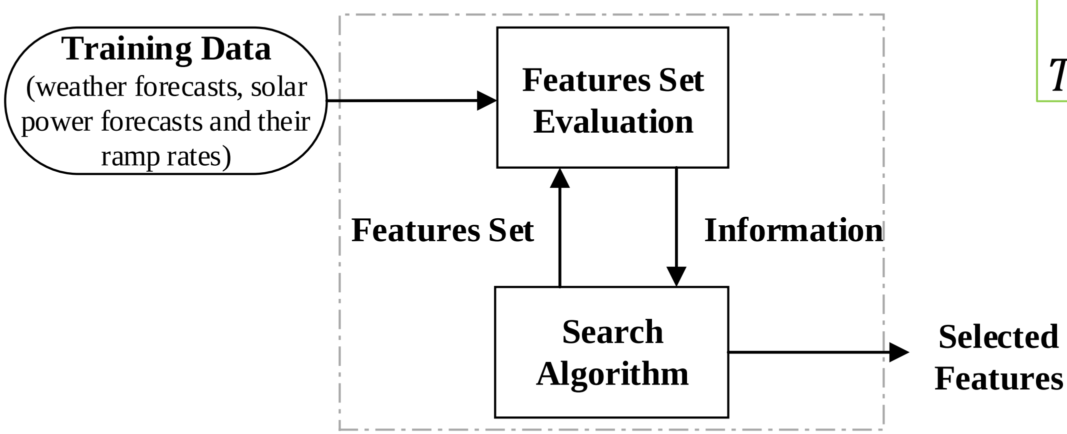
Data Preprocessing

Flowchart of Data Preparation

1. Pick up a feature from the available features set;
2. Run the model with this feature;
3. Score the model, by using the following score: Max(Diff. Index), where ***Diff. Index*** is the difference between true and false ramp events;
4. Add another feature to the selected features;
5. Repeat steps 2 and 3;
6. Choose subset of features with the best score, remove the selected from the available features;
7. Repeat steps 1 to 6;
8. If there is no longer any feature to select, Stop.

Objective: Increase the true events,
Decrease the false events.

True Events ↑ & *False Events* ↓



Wrapper Approach

Data Preprocessing

Available data

No.	Input Variable, (X)	No.	Input Variable, (X)
1	Cloud Water Content	10	Surface thermal radiation down
2	Cloud Ice Content	11	Top net solar radiation
3	Surface Pressure	12	Total precipitation
4	Relative Humidity	13	Heat Index
5	Cloud Cover	14	Wind Speed
6	10m - U Wind	15	Hours
7	10m - V Wind	16	Months
8	2-m Temperature	17	Days of Month
9	Surface solar radiation down	18	Days of Year

Models' parameters and their selected input variables

Model	Model Parameters	Selected Features (Input Variables)
MLR	The regression coefficients (βs) of the MLR model are found using OLS with the training set	The candidate MLR model: $\beta_0+\beta_1X_9+\beta_2X_8+\beta_3X_{10}+\beta_4X_{12}+\beta_5X_2+\beta_6X_4+\beta_7X_{16}+\beta_8X_{15}+\beta_9X_9^2+\beta_{10}X_9^3+\beta_{11}X_9*X_{15}+\beta_{12}X_9*X_{16}+\beta_{13}X_9*X_{17}+\beta_{14}X_9^2*X_{15}+\beta_{15}X_9^2*X_{16}+\beta_{16}X_9*X_8*X_{15}+\beta_{17}X_9*X_{10}*X_{15}+\beta_{18}X_9*X_4*X_{15}+\beta_{19}X_9*X_{12}*X_{15}+\beta_{20}X_9*X_2*X_{15}*X_{17}+\beta_{21}X_9^2*X_{17}+\beta_{22}X_5*X_{15}+\beta_{23}X_8*X_{15}+\beta_{24}X_1*X_{15}+\beta_{25}X_2*X_{15}+\beta_{26}X_{12}*X_{15}+\beta_{27}X_4*X_{15}+\beta_{28}X_{10}*X_{15}+\beta_{29}X_{11}*X_{16}+\beta_{30}X_{11}*X_{17}$
ANN	Hidden layers=1 Neurons=20	X_1, X_2, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}, X_{11}, X_{12}, X_{15}
SVR	Kernel type= RBF C=50 and Gamma=1	X_4, X_5, X_8, X_9, X_{10}, X_{11}, X_{12}, X_{14}, X_{15}
RF	RF Size, B=100 Trees Leaf size, n_{min}=5 Input samples, m=6	The models' outcomes and their ramp rates: -Day-ahead forecasts: MLR, ANN, SVR -Hour-ahead forecasts: Persistence and combined forecasts Using two loss functions: MSE_F & MSE_{RR}

Evaluation Metrics

These are negatively oriented metrics except the skill score, which is the higher is the better.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - F_i)^2}$$

$$RMSE_{RR} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(RR_{Pi} - RR_{Fi})^2}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|P_i - F_i|$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(P_i - F_i)$$

$$Pb_{q,i}(F_q, P_i) = \begin{cases} (1 - \frac{q}{100})(F_q - P_i), & \text{if } P_i < F_q \\ \frac{q}{100}(P_i - F_q), & \text{if } P_i \geq F_q \end{cases}$$

Quantiles, $q \in [1- 99]$

$$\text{Pinball}, PB = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{99}\sum_{q=1}^{q=99} PB_{q,i}\right)$$

$$Skill\ Score\ (\%) = \left(1 - \frac{Metric_{method}}{Metric_{reference}}\right) * 100$$

Evaluation Metrics

The evaluation metrics to assess the classification for solar ramp events:

Objective: Increase the true events,
Decrease the false events.

$$True\ Events \uparrow \quad \& \quad False\ Events \downarrow$$

$$Total\ Accuracy = \frac{True\ Events}{Total\ Events}$$

$$Precision = \frac{True\ High}{(True\ High + False\ High)}$$

$$Recall = \frac{True\ High}{(True\ High + False\ Low)}$$

$$Balance\ Precision = \frac{1}{4}\sum_{class=1}^{4}\frac{True\ class}{(True\ class + False\ class)}$$

$$F1\ score = \frac{2*(Precision*Recall)}{(Precision + Recall)}$$

$$Diff.\ Index = (True - False)\ of\ High\ Rate\ Events$$

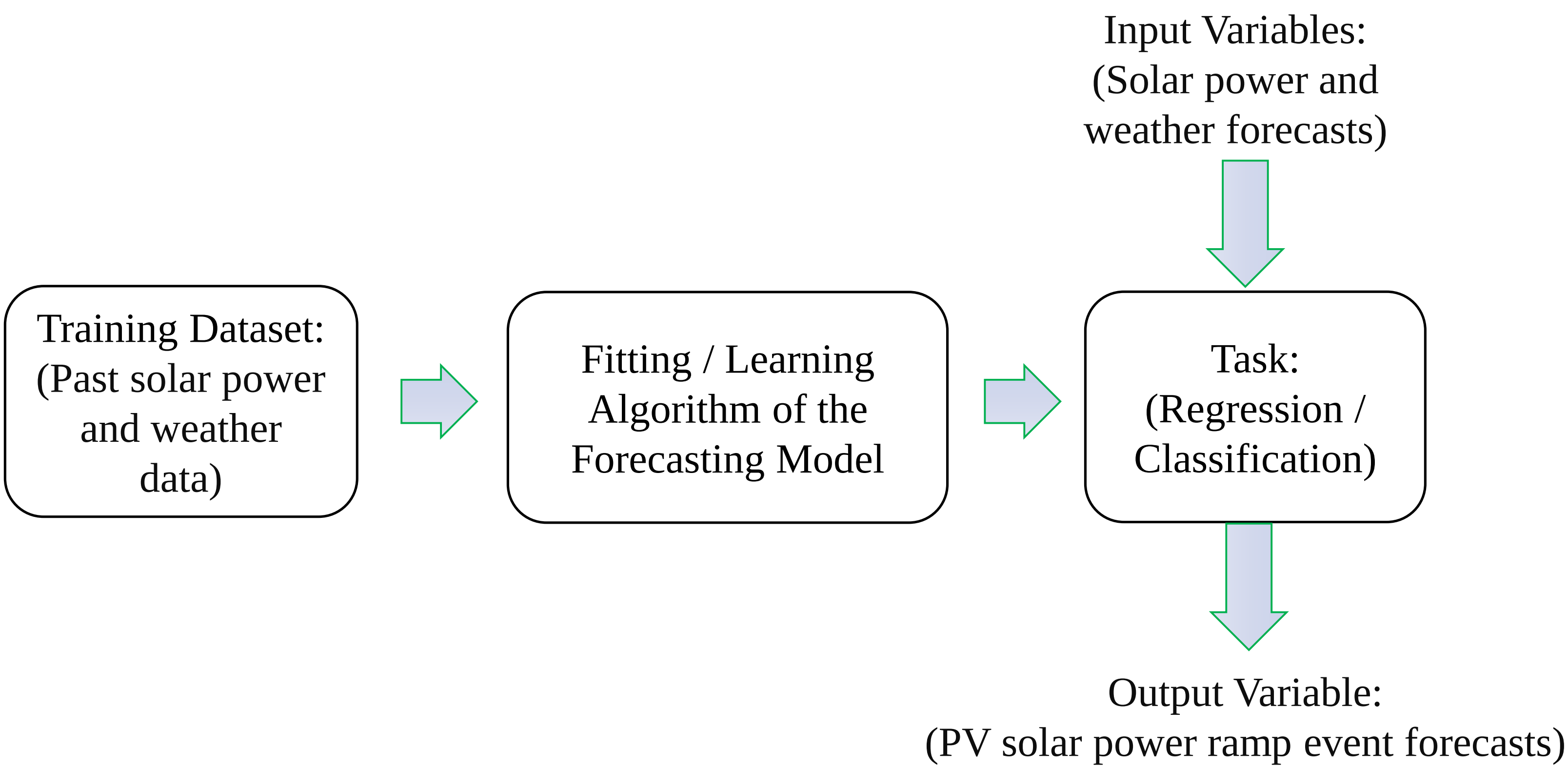
Predicted Events			
High-Rate	**True** High-Rate	**False** High-Rate	
Low-Rate	**False** Low-Rate	**True** Low-Rate	
	High-Rate	Low-Rate	**Observed Events**

Confusion matrix of possible cases of solar power ramp events

The most suitable metrics for our application are the Diff. Index and the F1 score.

Methodology

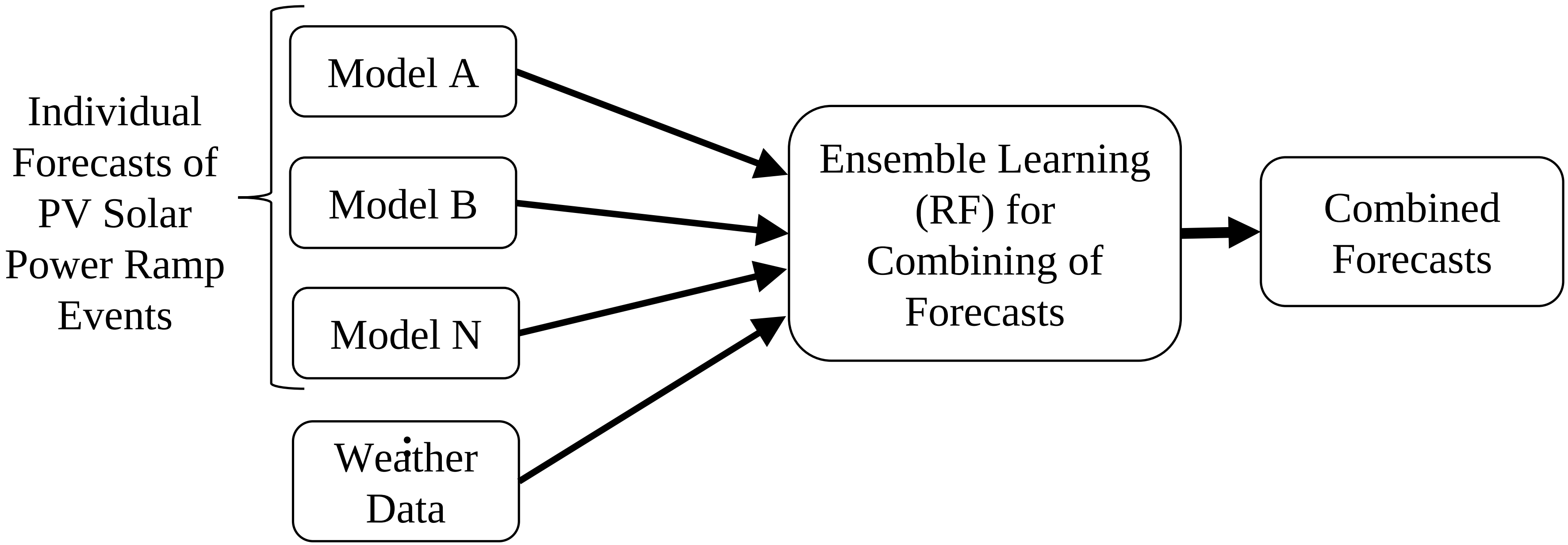
Regression or Classification Models



Block diagram of PV solar power ramp event forecasting models

Methodology

Ensemble Forecasts: Combining Various Models



General diagram of combining different models

$$F_{comb} = W_A * M_A + W_B * M_B + W_C * M_C \ldots . + W_N * M_N$$

Method of Combining The Models

Random forest (RF) is chosen to be the ***ensemble learning*** method for combining the various models' outcomes.

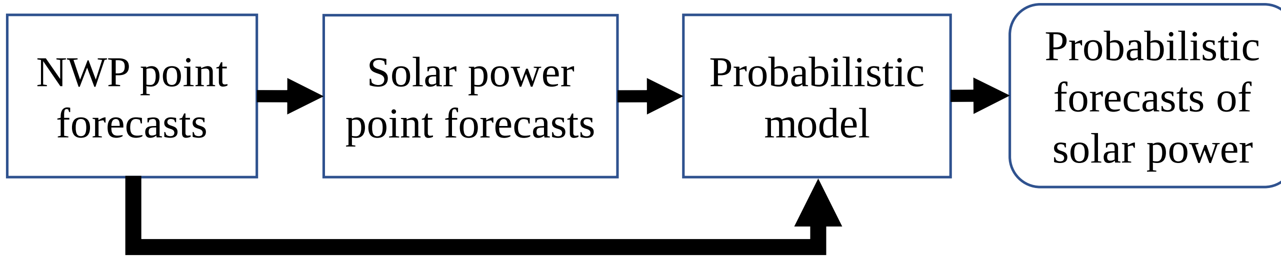
T. Hastie, R. Tibshirani, J. Friedman, and others, *The elements of statistical learning*, 2nd Edition. Springer-Verlag New York, 2009.

Methodology

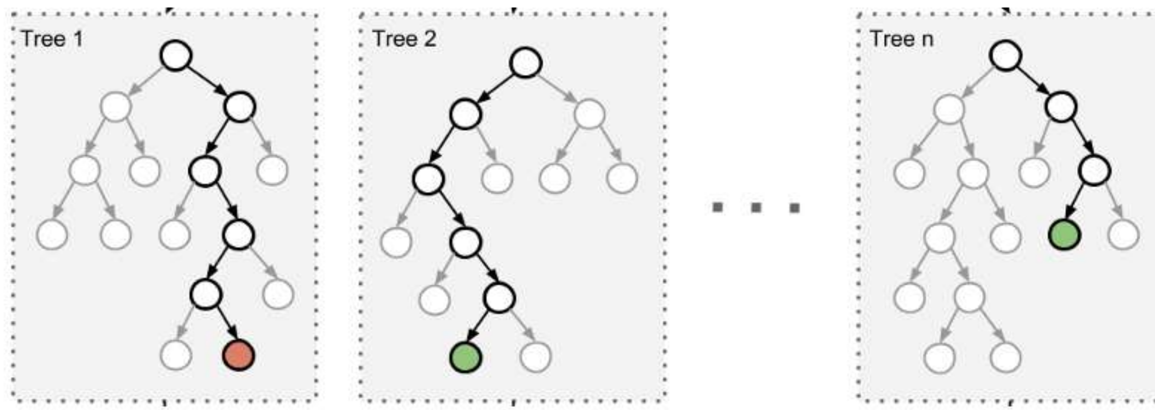
Probabilistic Forecasts

Ensemble-based probabilistic forecasts

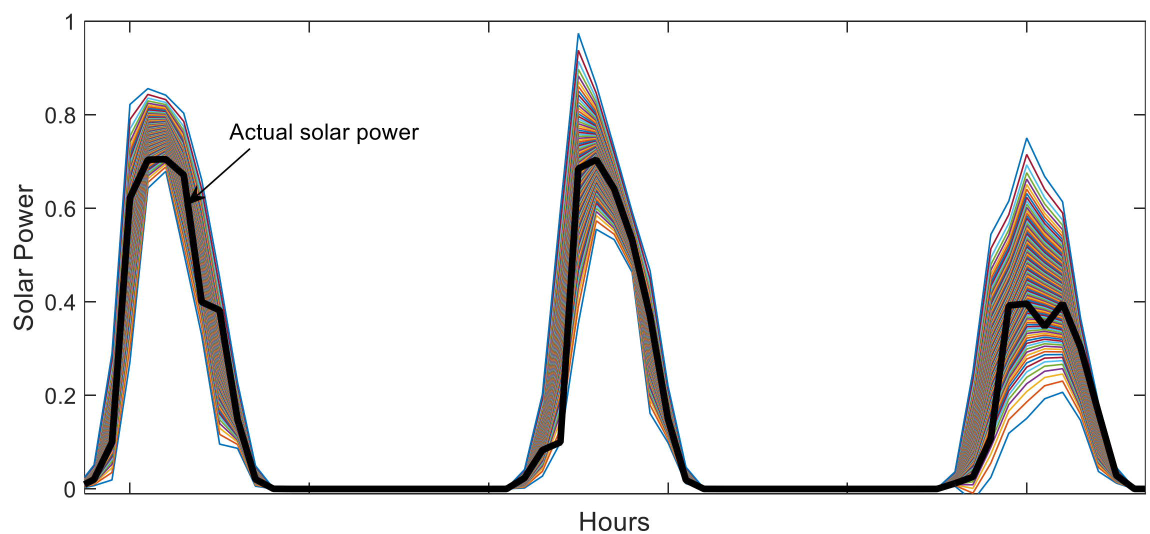
(a)



(b)



(c)



a) Diagram of ensemble-based probabilistic forecasts,

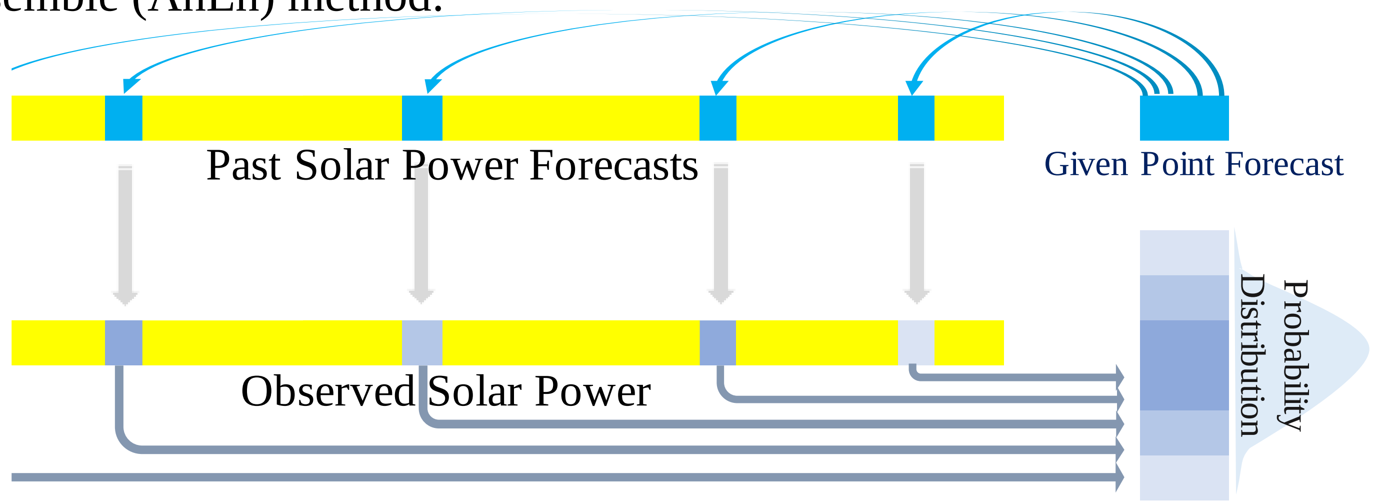
b) Splitting mechanism of trees in random forest,

c) Sample of ensemble-based probabilistic forecasts of solar power of 3 days

Methodology

Probabilistic Forecasts

Analog Ensemble (AnEn) method:



Schematic diagram of analog ensemble method

$$\left|F_{\text{Given}}^{Hr} - F_{\text{Past}}^{Hr}\right| \leq \varepsilon, \qquad \varepsilon = 0.1$$

where F_{Given}^{Hr} denotes the given point forecast at an hour *Hr*, for which the prediction interval will be estimated, F_{Past}^{Hr} the point forecasts at the same hour of the day.

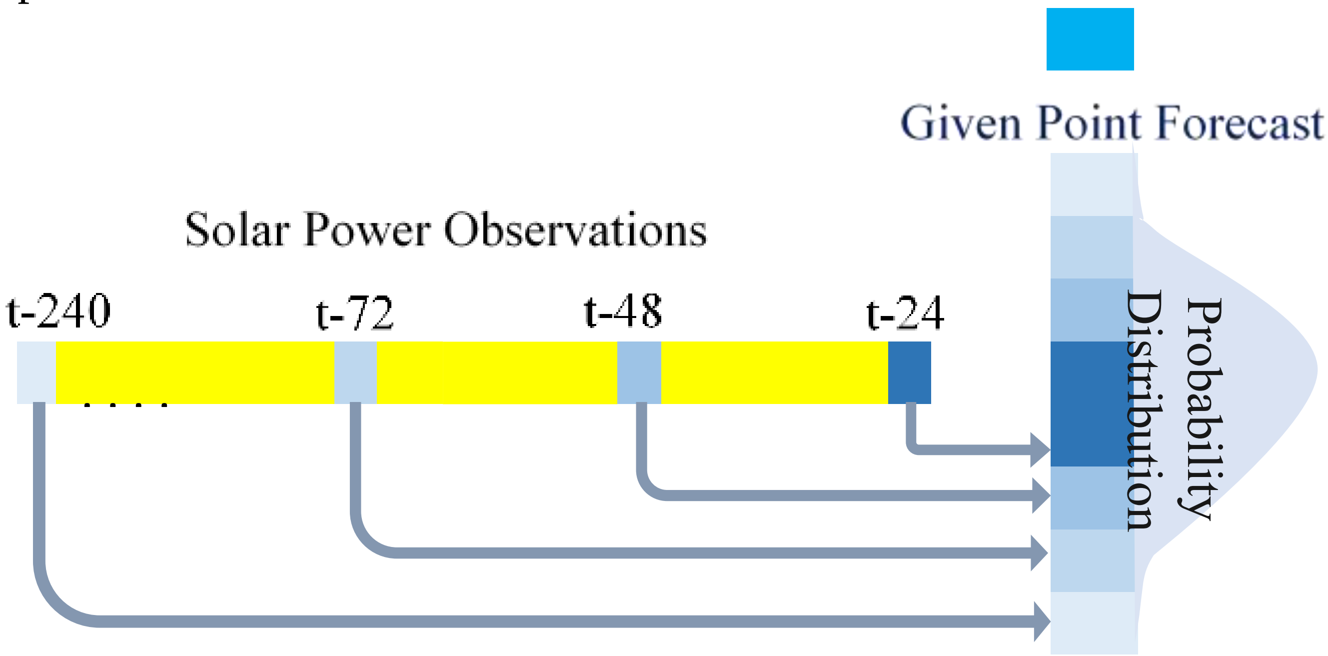
Delle Monache, L., Eckel, F. A., Rife, D. L., Nagarajan, B., & Searight, K. (2013). Probabilistic weather prediction with an analog ensemble. *Monthly Weather Review*, *141*(10), 3498-3516.

Alessandrini, S., Delle Monache, L., Sperati, S., & Cervone, G. (2015). An analog ensemble for short-term probabilistic solar power forecast. *Applied energy*, *157*, 95-110.

Methodology

Probabilistic Forecasts

Persistence probabilistic method:



Schematic diagram of persistence probabilistic method

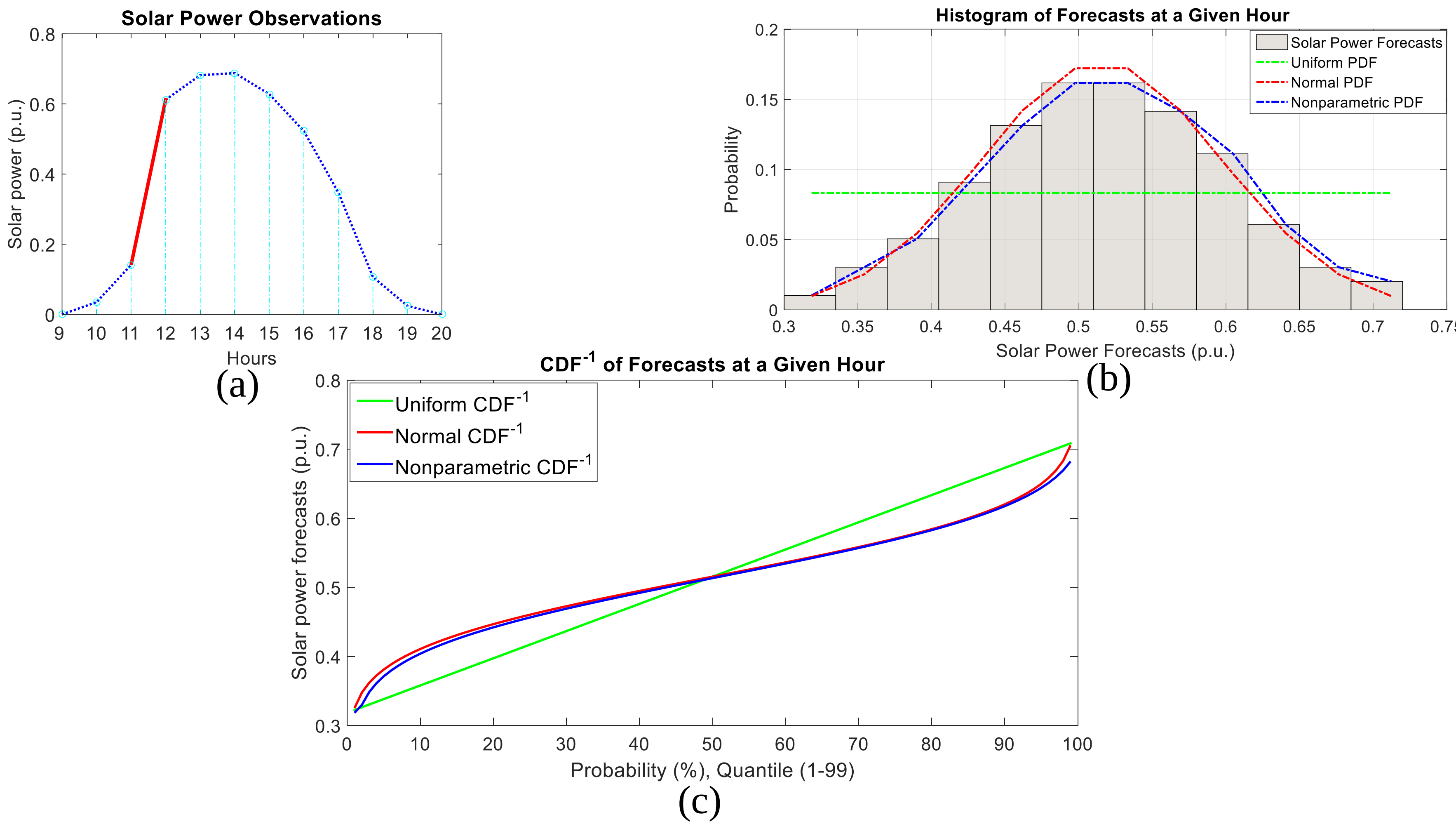
The 10, 20 and 30 recent observed powers are carried out.

It is found that the recent 10 observed solar powers at the given hour with CDF distribution achieve more accurate persistence probabilistic forecasts.

Methodology

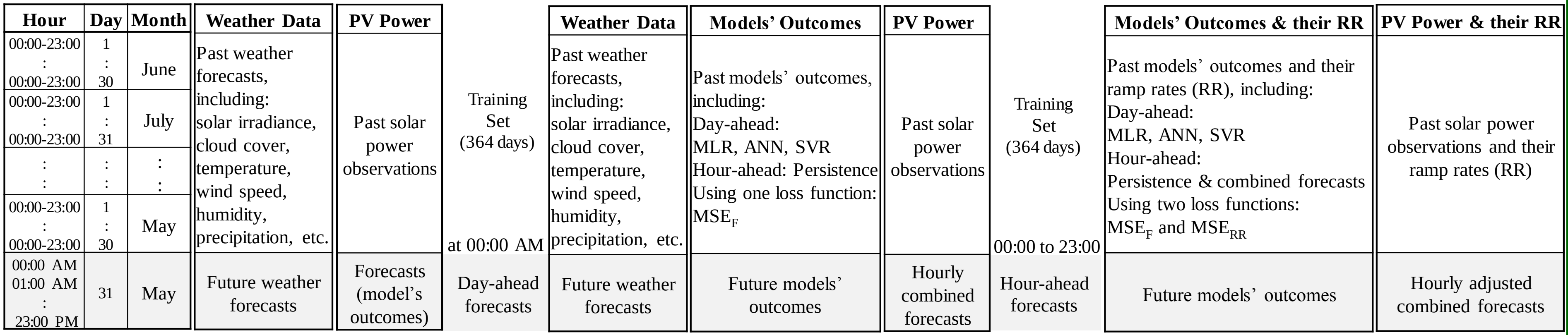
Probabilistic Forecasts

Probability distributions of the random forest outcomes at 12:00 pm on May 29th



(a) Solar power observations of the given day, (b) histograms of the random forest outcomes of the forecast at 12:00pm,(c) estimated CDFs for the probabilistic forecasts at 12:00pm

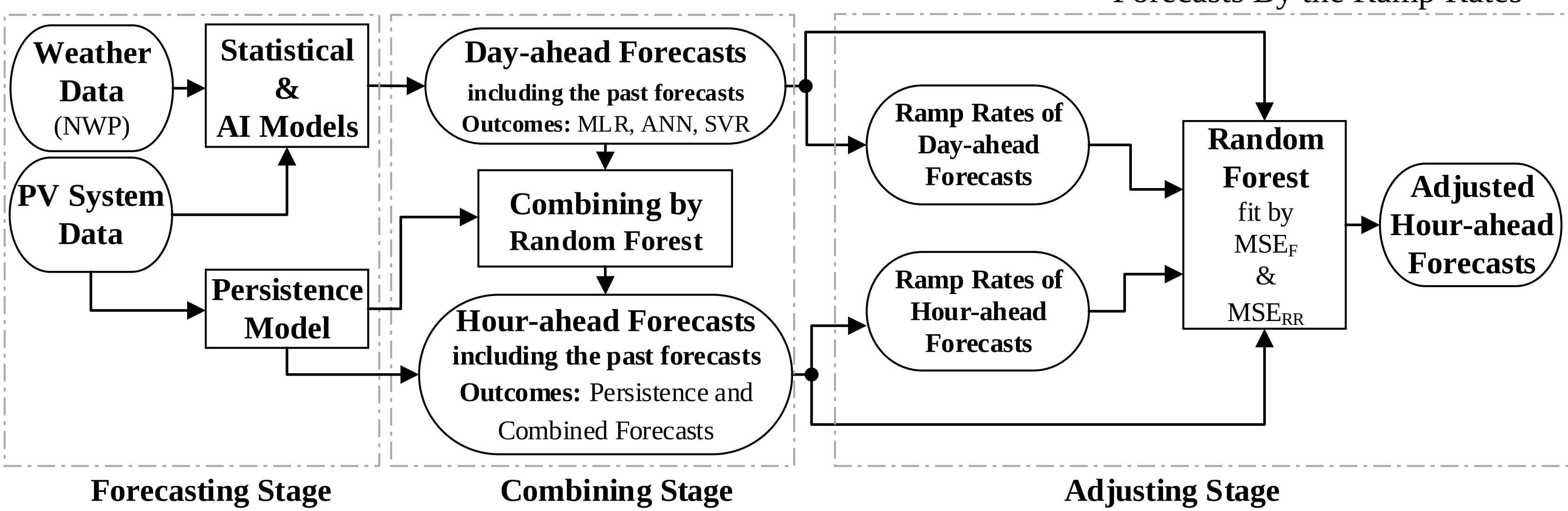
Methodology



(a) Producing Different Models' Outcomes

(b) Combining by Ensemble Learning

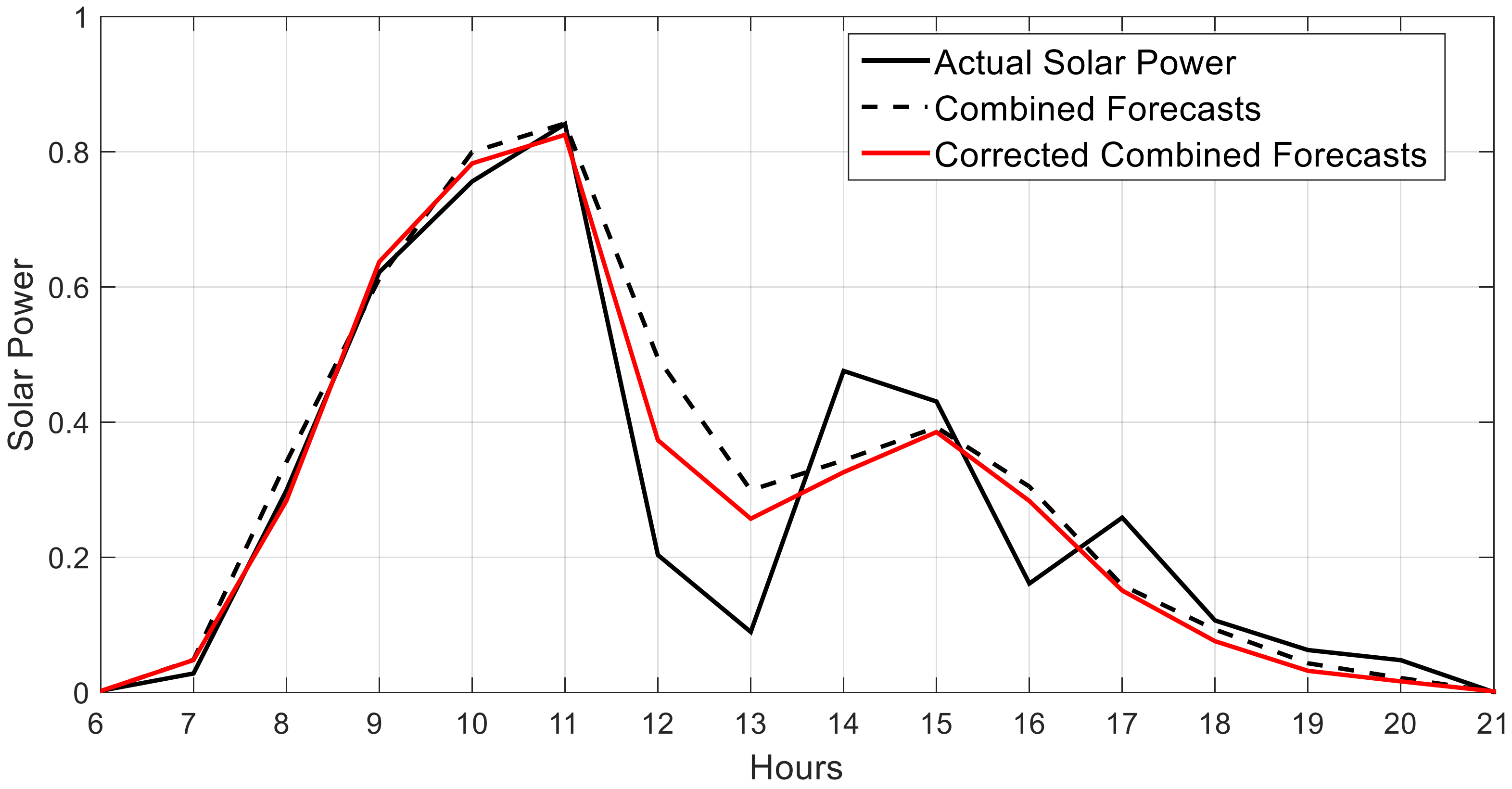
(c) Adjusting and Correcting the Combined Forecasts By the Ramp Rates



Schematic and block diagrams of **the adjusting post-processing approach**

Modeling and Results

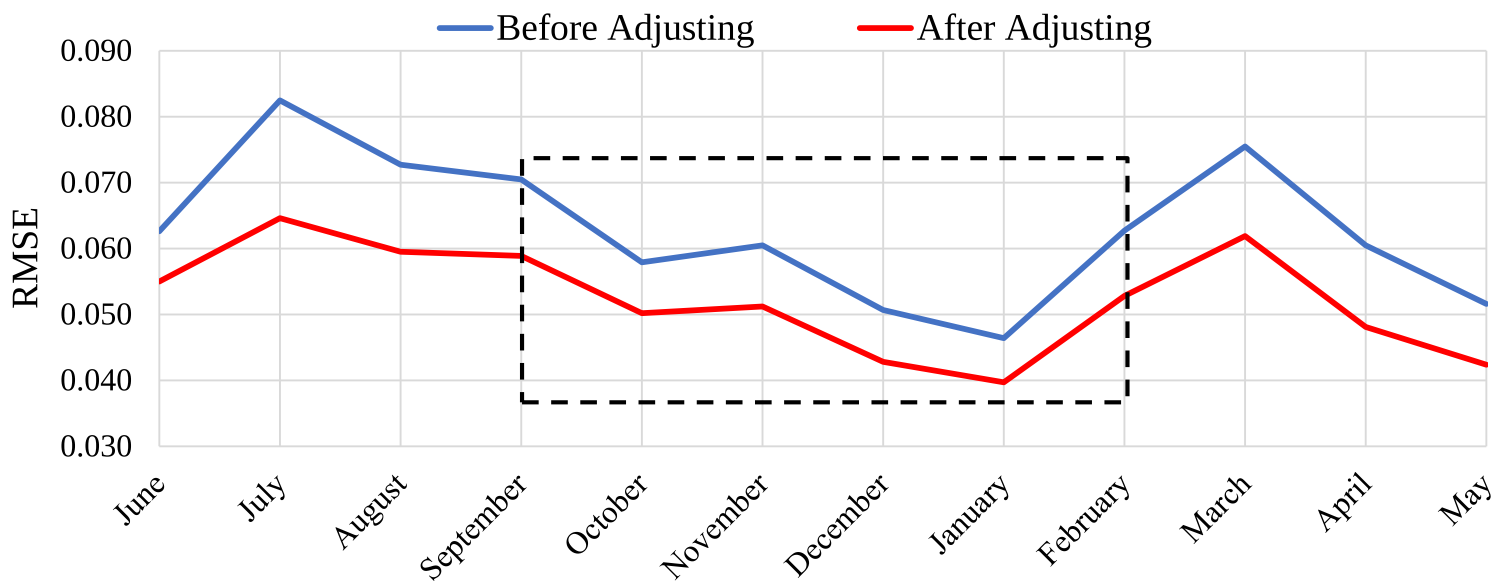
Implementing the ***adjusting approach*** for Improving Combined Solar Power Forecasts



Combined forecasts of solar power for a cloudy day before and after applying the adjusting

Modeling and Results

Improving Combined Solar Power Forecasts



Monthly RMSEs of combined forecasts before and after applying the adjusting approach

Comparison of hour-ahead forecasts over a complete year

Method	Persistence	MLR	ANN	SVR	Simple Average	Ensemble (Before Adjusting)	Ensemble (After Adjusting)
MBE (Bias)	0.0756	-0.1498	-1.852	-4.291	-1.554	0.0469	0.0747
RMSE	0.1209	0.0763	0.0681	0.0700	0.0667	0.0628	**0.0523**
$RMSE_{RR}$	0.1383	0.0771	0.0722	0.0747	0.0796	0.0750	0.0698
RMSE Improve (%)	**57%**	**31%**	**23%**	**25%**	**22%**	**17%**	---

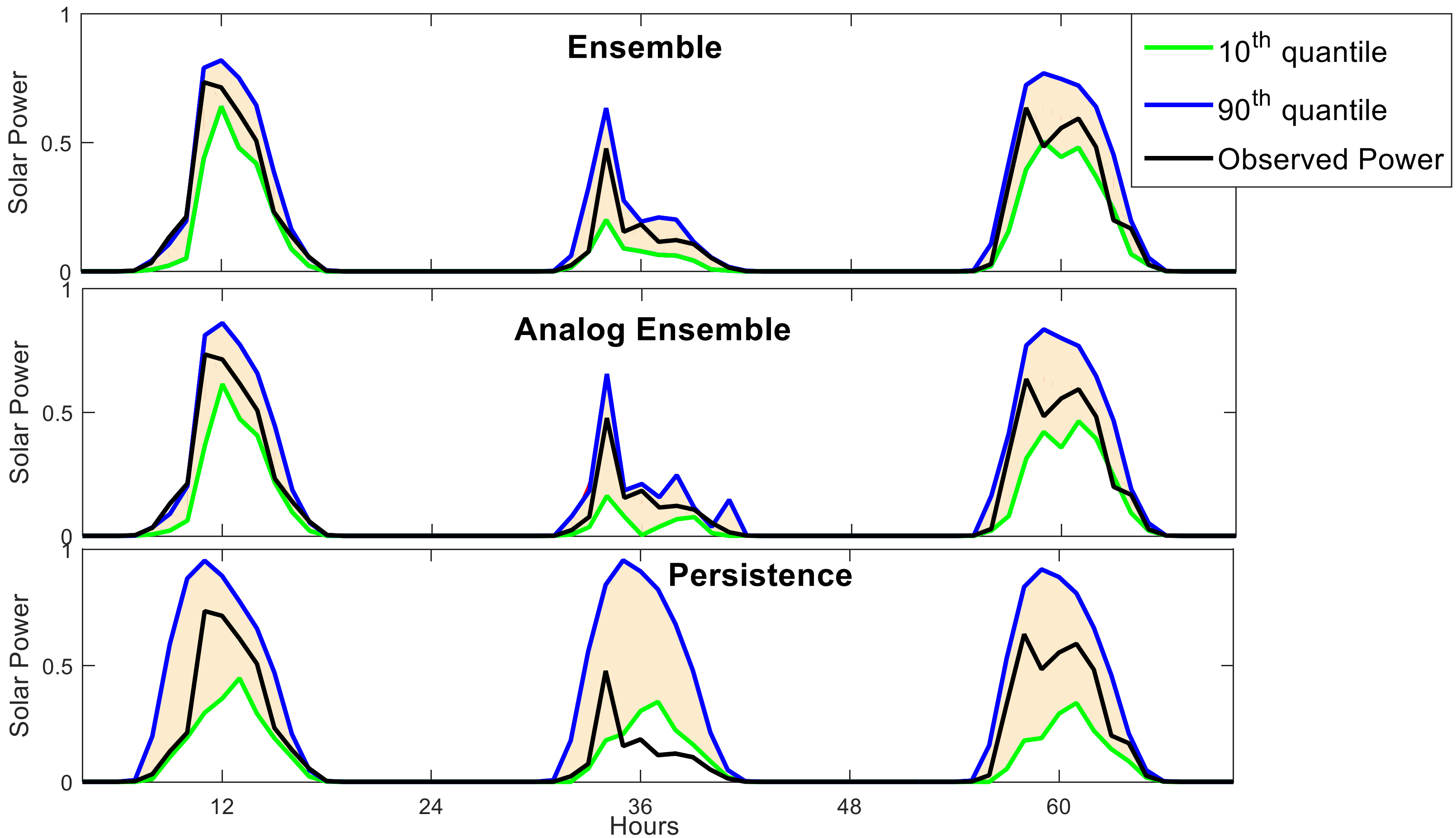
Modeling and Results

Probabilistic Forecasts

$$Pb_{q,i}(F_q, P_i) = \begin{cases} \left(1 - \frac{q}{100}\right)(F_q - P_i), & \text{if } P_i < F_q \\ \frac{q}{100}\,(P_i - F_q), & \text{if } P_i \geq F_q \end{cases}$$

Quantiles, $q \in [1\text{-}99]$

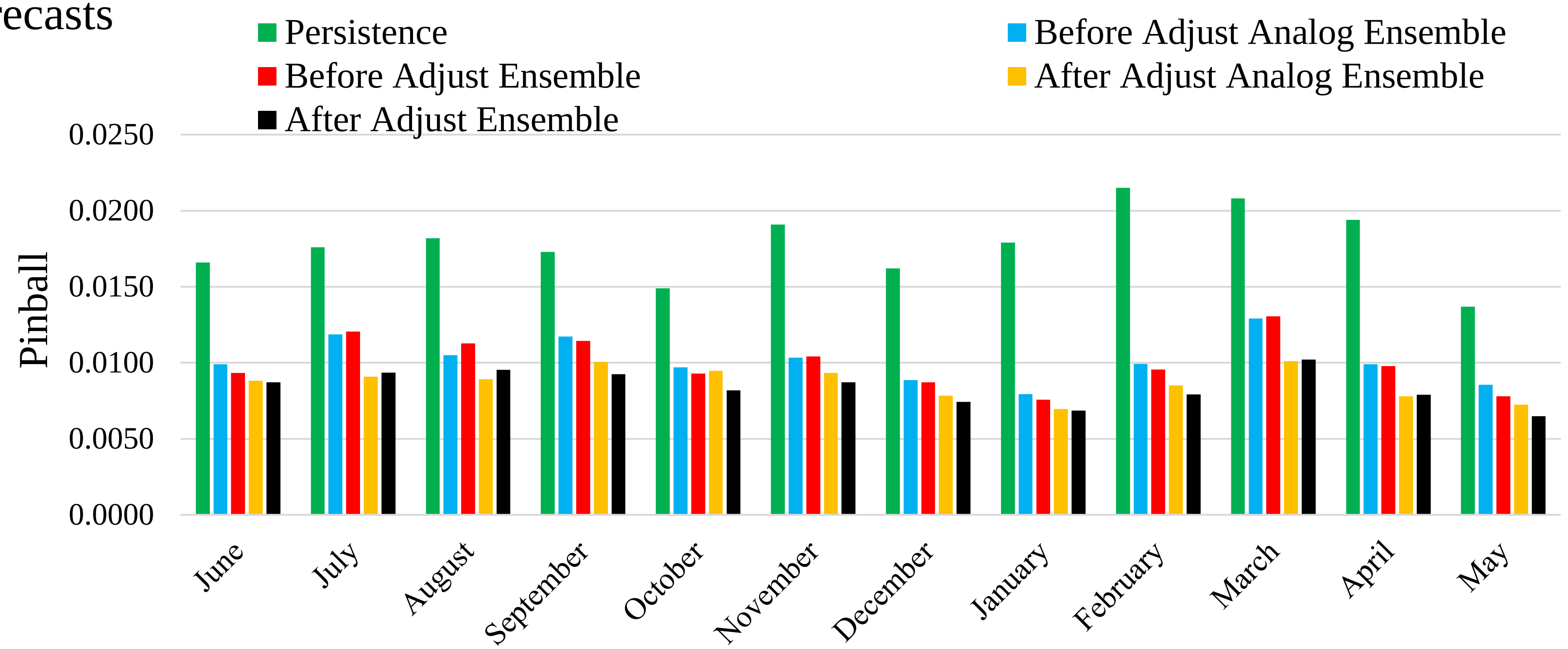
$$\text{Pinball}, PB = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{99}\sum_{q=1}^{q=99} PB_{q,i}\right)$$



Graphs of the probabilistic forecasts of the three methods for three days

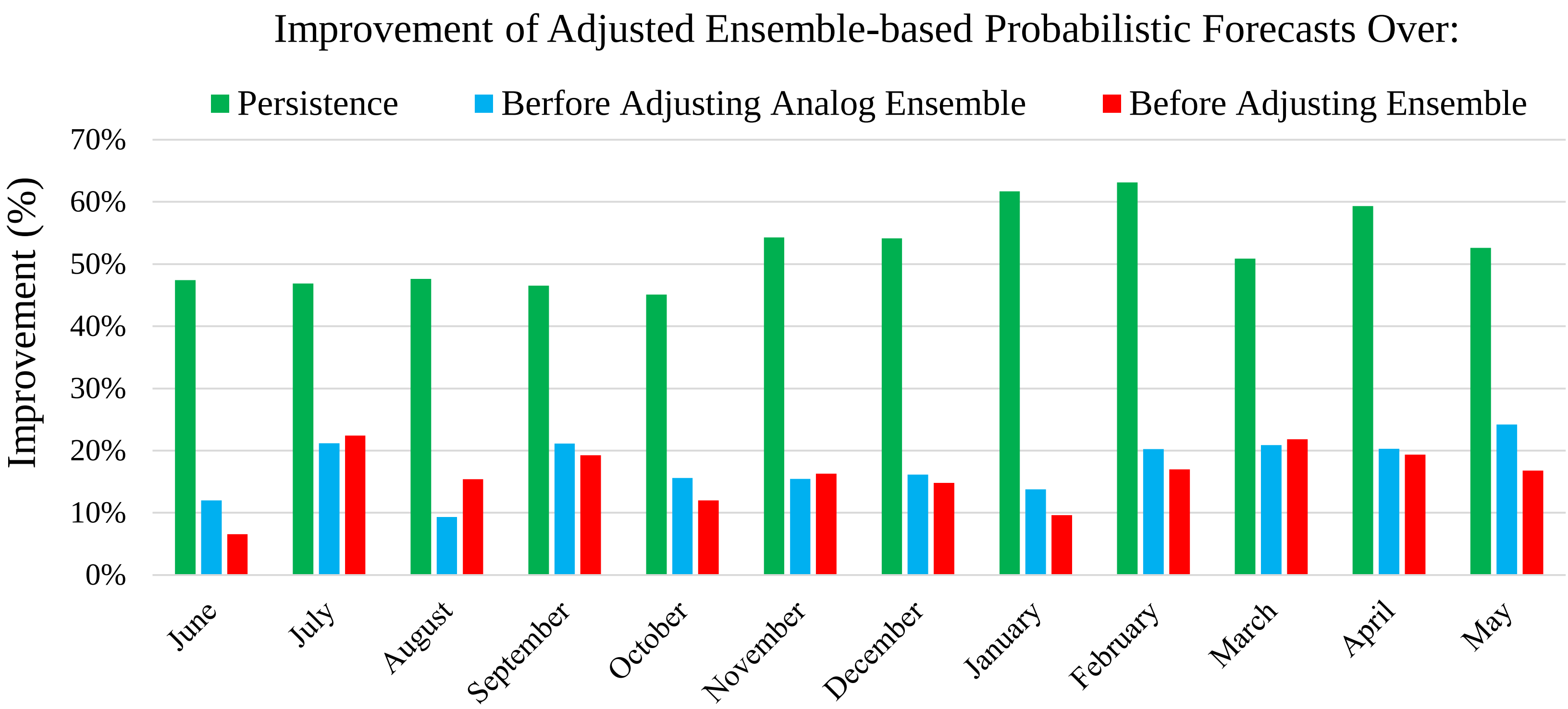
Probabilistic Forecasts

Modeling and Results



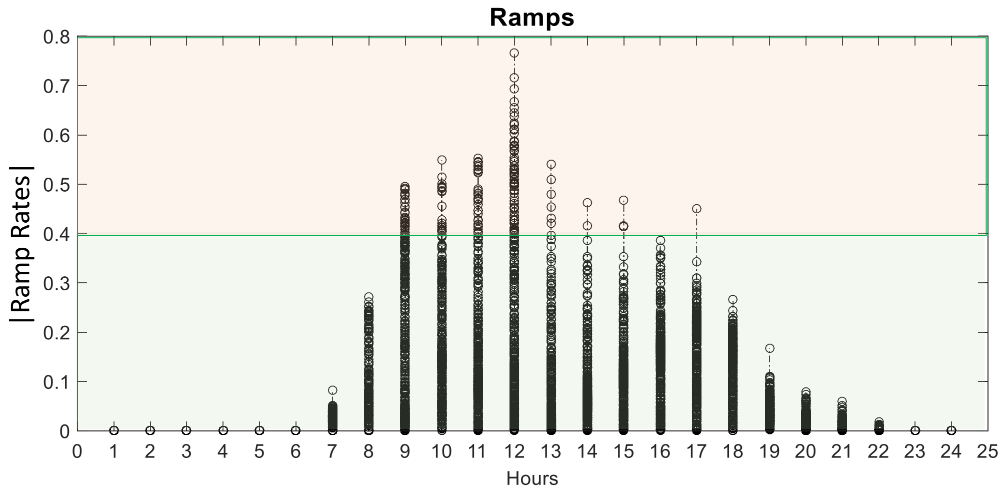
Month	Pinball					Improvement of Adjusted Ensemble Over:		
	Persistence	Before Adjusting		After Adjusting				
		Analog Ensemble	Ensemble	Analog Ensemble	Ensemble	Persistence	Before adjust AnEn	Before adjust Ensemble
June	0.0166	0.0099	0.0093	0.0088	0.0087	47%	12%	7%
July	0.0176	0.0119	0.0121	0.0091	0.0094	47%	21%	22%
August	0.0182	0.0105	0.0113	0.0089	0.0095	48%	9%	15%
September	0.0173	0.0117	0.0114	0.0101	0.0092	47%	21%	19%
October	0.0149	0.0097	0.0093	0.0095	0.0082	45%	16%	12%
November	0.0191	0.0103	0.0104	0.0093	0.0087	54%	15%	16%
December	0.0162	0.0089	0.0087	0.0078	0.0074	54%	16%	15%
January	0.0179	0.0080	0.0076	0.0070	0.0069	62%	14%	10%
February	0.0215	0.0099	0.0095	0.0085	0.0079	63%	20%	17%
March	0.0208	0.0129	0.0131	0.0101	0.0102	51%	21%	22%
April	0.0194	0.0099	0.0098	0.0078	0.0079	59%	20%	19%
May	0.0137	0.0086	0.0078	0.0073	0.0065	53%	24%	17%
Average	**0.0178**	**0.0102**	**0.0100**	**0.0087**	**0.0084**	**52%**	**18%**	**16%**

Modeling and Results



Modeling and Results

Implementing several classification models to forecasts solar power ramp events



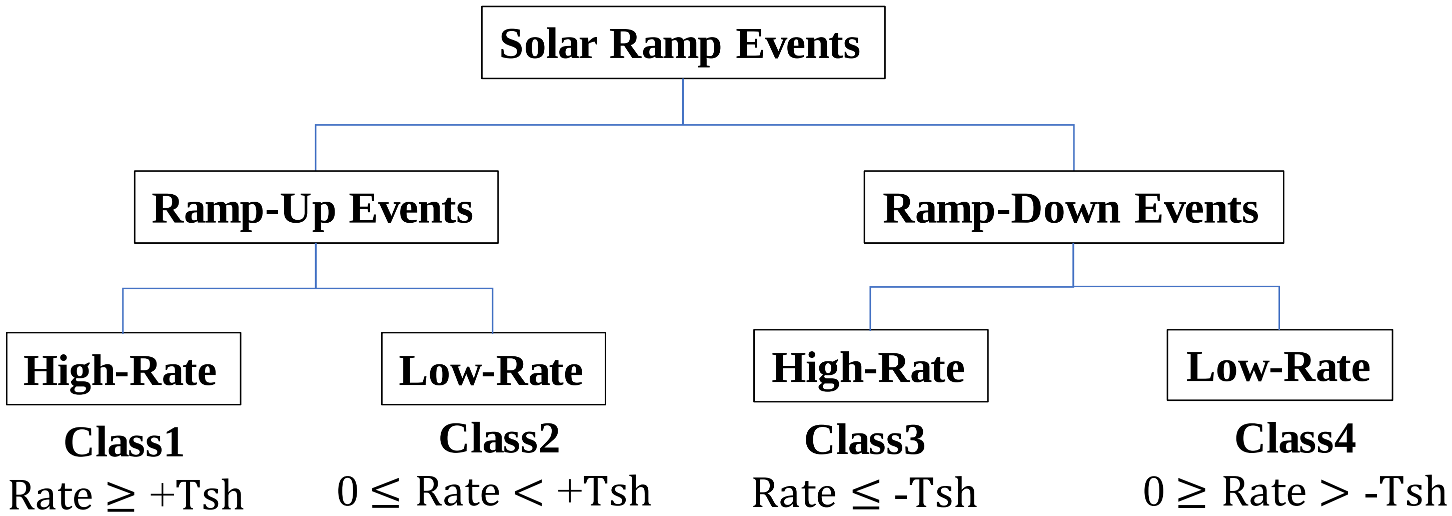
Classes of solar power ramp events in the case study

Objective: Increase the true events,
Decrease the false events.

True Events ↑ & *False Events* ↓

Ramp Classes as following:

Class1: Ramp up of high rate, |rate| ≥ Tsh
Class2: Ramp up of low rate, |rate| < Tsh
Class3: Ramp down of high rate, |rate| ≥Tsh
Class4: Ramp down of low rate, |rate| < Tsh



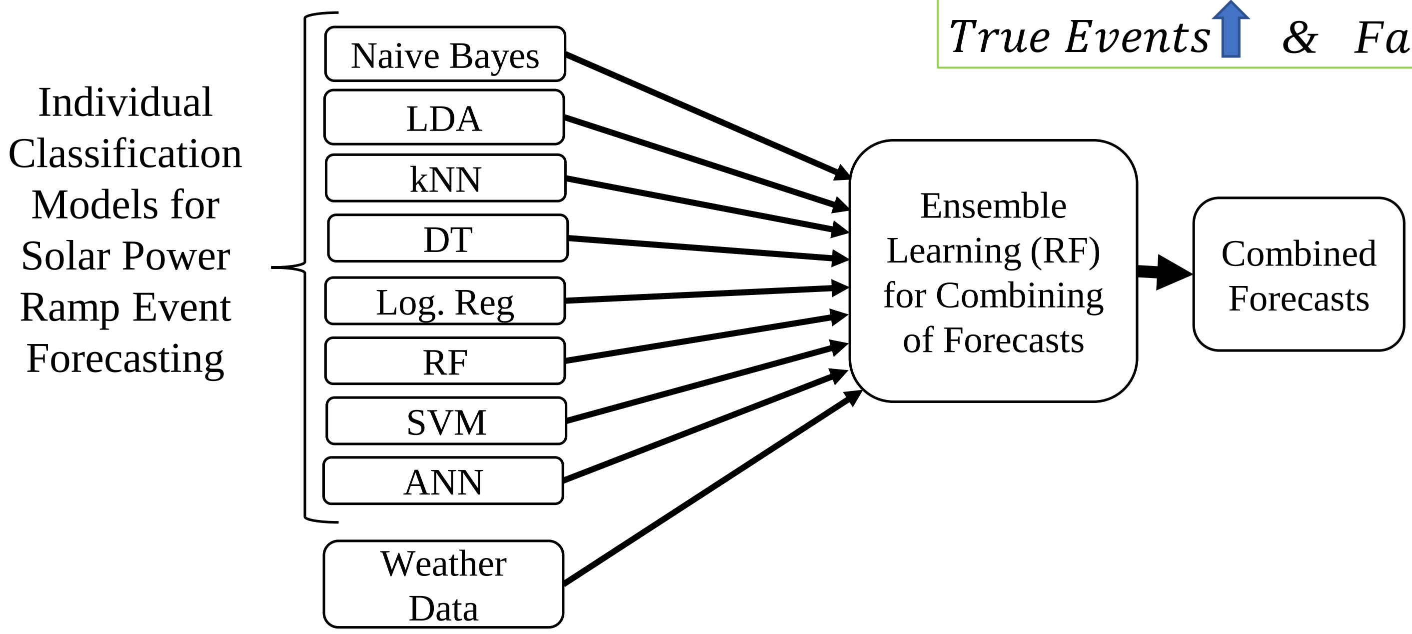
Distribution of the classes of solar power ramp events at threshold (Tsh) =0.4pu/hr.

Modeling and Results

Implementing several classification models to forecasts solar power ramp events

Objective: Increase the true events, Decrease the false events.

True Events ↑ & *False Events* ↓



Combining classification models for solar power ramp event forecasting

$$F_{comb}=W_A*M_A+ W_B*M_B + W_C*M_C \text{ ....} + W_N*M_N$$

where W_A is an assigned weight for an outcome of model (A)

T. Hastie, R. Tibshirani, J. Friedman, and others, *The elements of statistical learning*, 2nd Edition. Springer-Verlag New York, 2009.

Modeling and Results

Available Features for classification of solar power ramp events

Additional Features:

- (yellow) Solar power forecasts & their ramp rates
- (green) Classes (Labels) of Ramp Event

Weather Data

No.	Variable Name
1	Cloud Water Content
2	Cloud Ice Content
3	Surface Pressure
4	Relative Humidity
5	Cloud Cover
6	10m- U Wind
7	10m- V Wind
8	2-m Temperature
9	Surface solar radiation down
10	Surface thermal radiation down
11	Top net solar radiation
12	Total precipitation
13	Heat Index
14	Wind Speed

Day and Hour-ahead Forecasts

No.	Variable Name	RMSE	Horizon
15	Persistence Farm 1	0.1209	hour-ahead
16	Persistence Farm 2	0.2108	hour-ahead
17	Persistence Farm 3	0.1393	hour-ahead
18	MLR	0.0763	day-ahead
19	ANN	0.0681	day-ahead
20	SVR	0.0715	day-ahead
21	Combined 0.0628	0.0628	hour-ahead
22	Combined 0.0554	0.0554	hour-ahead
23	Combined 0.0523	0.0523	hour-ahead
24	Combined 0.0579	0.0579	hour-ahead
25	ARIMA 0.0928	0.0928	hour-ahead
26	ARIMAX 0.0915	0.0915	hour-ahead
27	NAR Hr1	0.0890	hour-ahead
28	NAR Hr2	0.1384	hour-ahead
29	NARX Hr1	0.0760	hour-ahead
30	NARX Hr2	0.1419	hour-ahead

Forward Ramp Rates

No.	Variable Name
31	MLR
32	ANN
33	SVR

No.	Variable Name
50	std. dev (all ramp rates)

Classes of the Forecasts

51	Persistence Farm 1
52	Persistence Farm 2
53	Persistence Farm 3
54	MLR
55	ANN
56	SVR
57	Combined 0.0628
58	Combined 0.0554
59	Combined 0.0523
60	Combined 0.0579
61	ARIMA 0.0928
62	ARIMAX 0.0915
63	NAR Hr1
64	NAR Hr2
65	NARX Hr1
66	NARX Hr2

Backward Ramp Rates

No.	Variable Name
34	Persistence Farm 1
35	Persistence Farm 2
36	Persistence Farm 3
37	MLR
38	ANN
39	SVR
40	Combined 0.0628
41	Combined 0.0554
42	Combined 0.0523
43	Combined 0.0579
44	ARIMA 0.0928
45	ARIMAX 0.0915
46	NAR Hr1
47	NAR Hr2
48	NARX Hr1
49	NARX Hr2

Modeling and Results

Combined Forecasts of Ramps

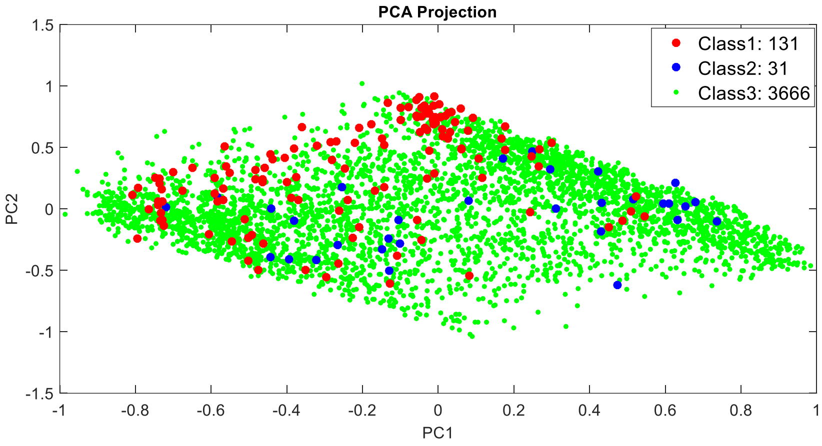
Data projection onto PC1 and PC2:

For clarity, low-rate classes (2&4) become as a one class, class3: ➔

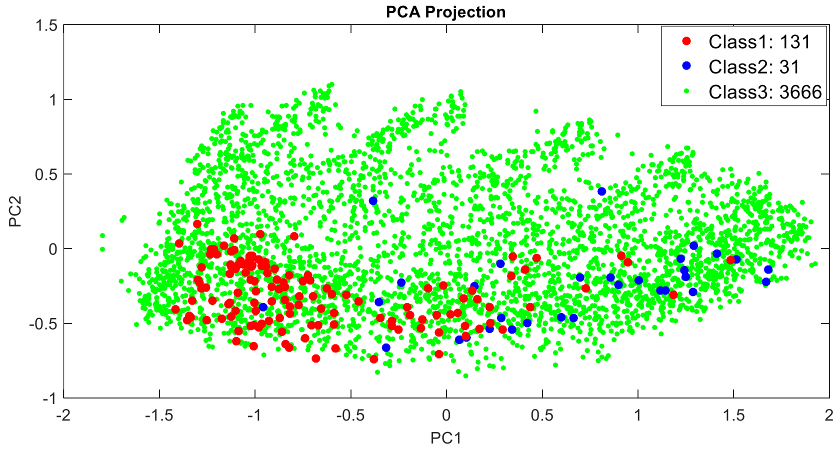
Threshold=0.4pu/hr

Class	Class1	Class2	Class3	Total
Samples	131	31	3666	3828

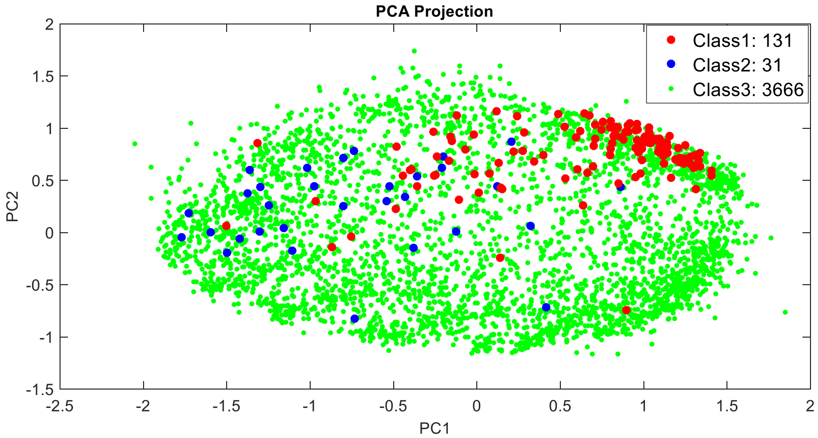
1) All 14 weather variables



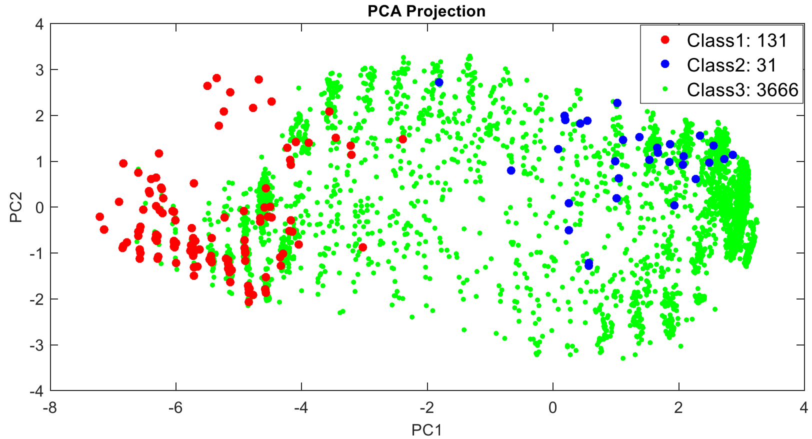
2) All 30 Features (*add solar power forecasts*)



3) All 50 Features (*add ramp rates of forecasts*)



4) All 66 Features (*add class labels of forecasts*)



Modeling and Results

(a) The most 12 important features out of 66; (b) selected features and parameters for each model

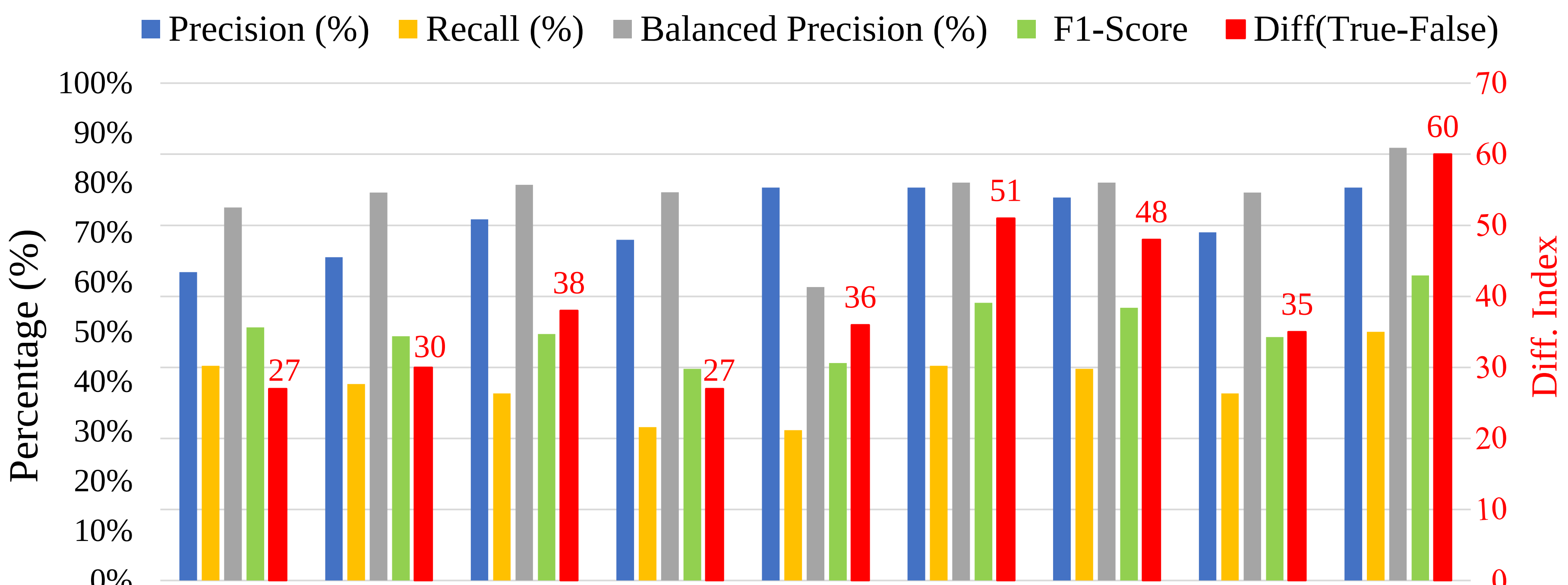
(a)

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 of 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

Modeling and Results

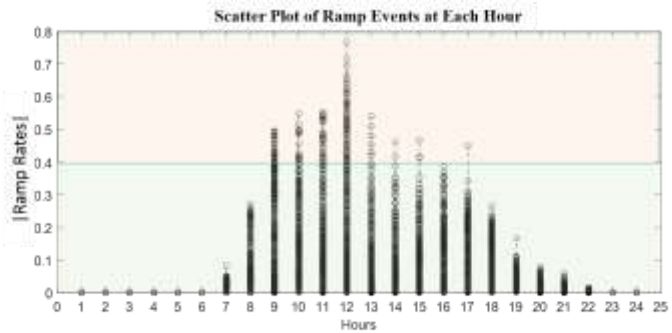


Method	Naïve Bayes	LDA	Decision Trees	kNN	Logistic Regression	Random Forest	SVM	ANN	Combined Classifiers
Precision (%)	62%	65%	73%	68%	79%	79%	77%	70%	79%
Recall (%)	43%	40%	38%	31%	30%	43%	43%	38%	50%
Balanced Precision (%)	75%	78%	80%	78%	59%	80%	80%	78%	87%
F1-Score (%)	51%	49%	50%	43%	44%	56%	55%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

Solar power ramp event forecasts of the high-rate ramp events, ($|Rate| \geq 0.4\ pu/hr$ = 162 events) by the classification techniques.

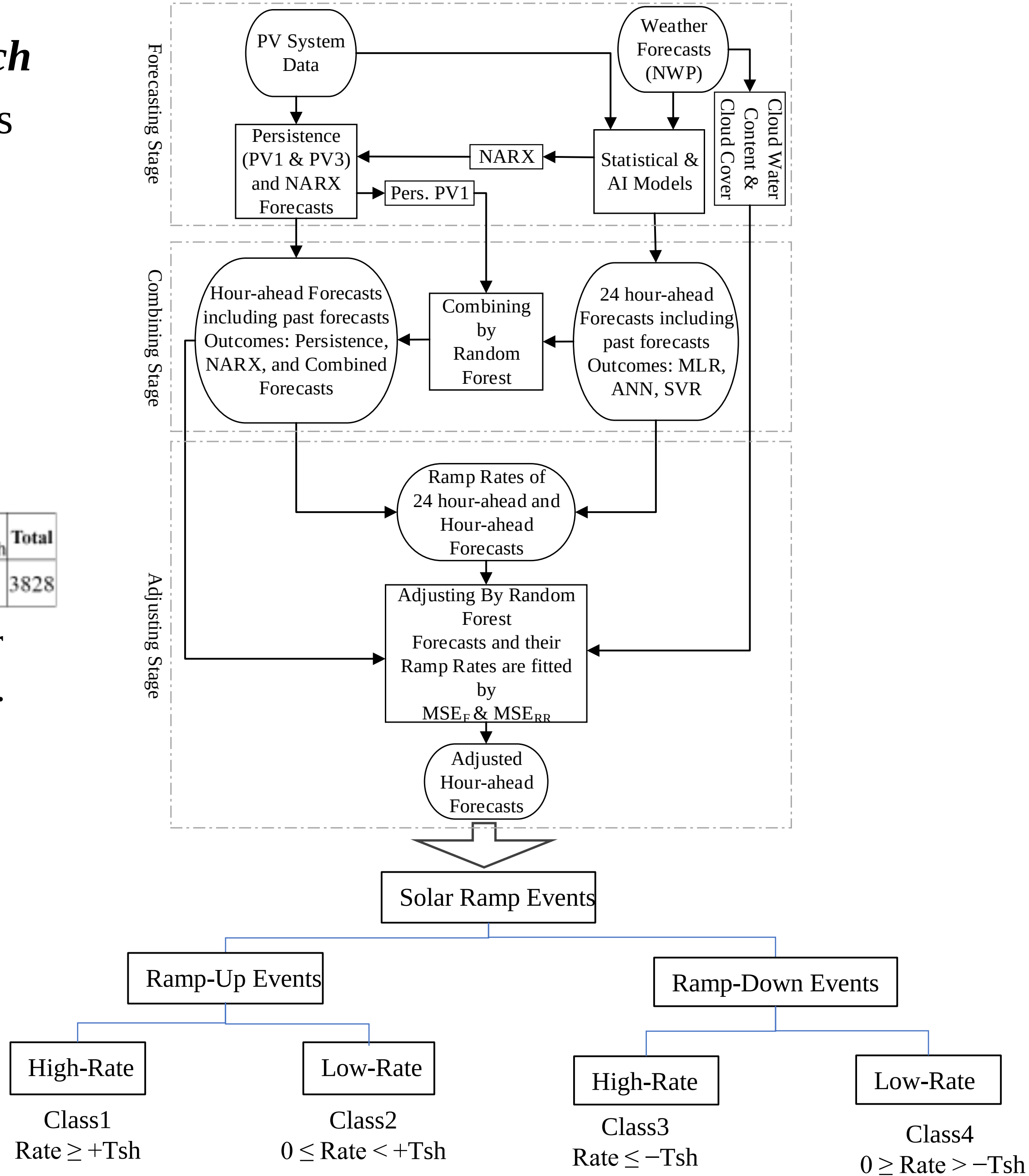
Modeling and Results

Implementing the ***adjusting approach*** to forecasts solar power ramp events



Class	Class1 Rate ≥ +Tsh	Class2 0 ≤ Rate < +Tsh	Class3 Rate ≤ −Tsh	Class 4 0 ≥ Rate > −Tsh	Total
Ramp Events at Tsh = 0.4pu/hr	131	1290	31	2376	3828

Distribution of the classes of solar power ramp events at threshold (Tsh) =0.4pu/hr.

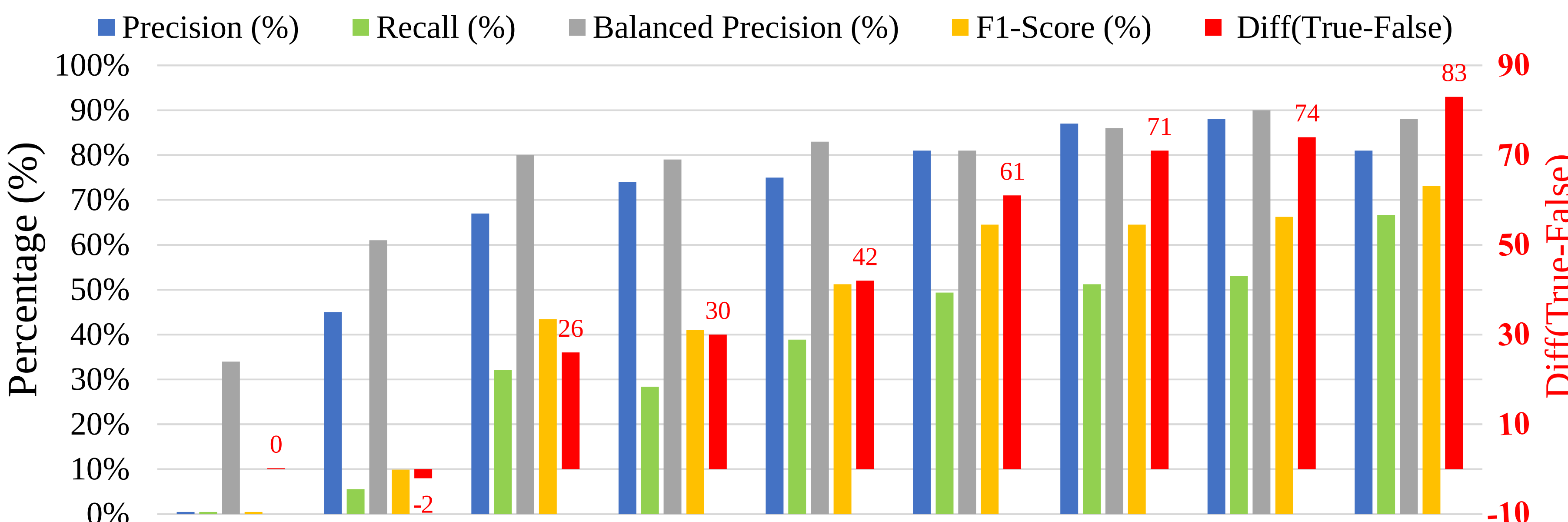


Block diagram of the **adjusting approach** of solar power ramp event forecasting

Modeling and Results

Implementing the ***adjusting approach*** to forecasts solar power ramp events

Evaluation of Solar Power Ramp Event Forecasts by Using Different Evaluation Metrics



	Persistence, PV1	NARX	Simple Average (NARX, MLR, ANN, SVR)	Combined (Pers.1, MLR, ANN, SVR)	Adjust Approach (Pers.1, MLR, ANN, SVR)	Combined (Pers.1&3 +NARX, Clouds, MLR, ANN, SVR)	Adjust Approach (Pers.1&3 +NARX, Clouds, MLR, ANN, SVR)	Probabilistic Forecasts (Q1 to Q99)	Probabilistic Forecasts (Q25 to Q75)
Precision (%)	0%	45%	67%	74%	75%	81%	87%	88%	81%
Recall (%)	0%	6%	32%	28%	39%	49%	51%	53%	67%
Balanced Precision (%)	34%	61%	80%	79%	83%	81%	86%	90%	88%
F1-Score (%)	0%	10%	43%	41%	51%	64%	64%	66%	73%
Diff(True-False)	0	-2	26	30	42	61	71	74	83

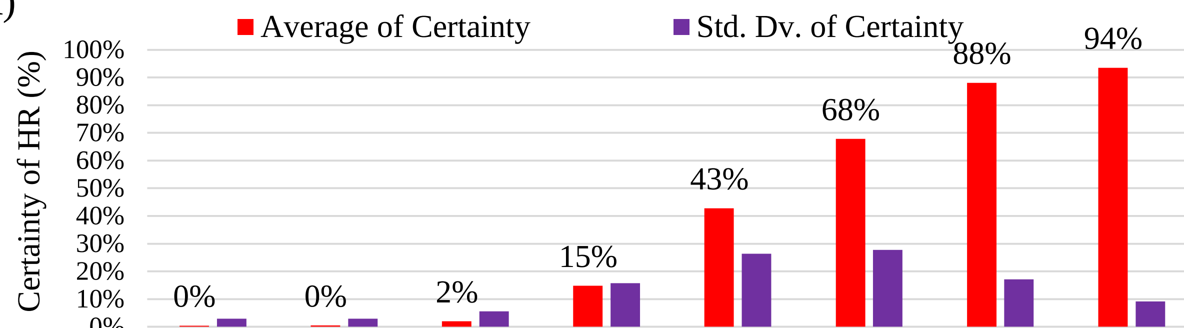
Solar power ramp event forecasts of the high-rate ramp events, ($|Rate| \geq 0.4\ pu/hr$ = 162 events) by the ***adjusting approach***

Modeling and Results

The uncertainty analysis

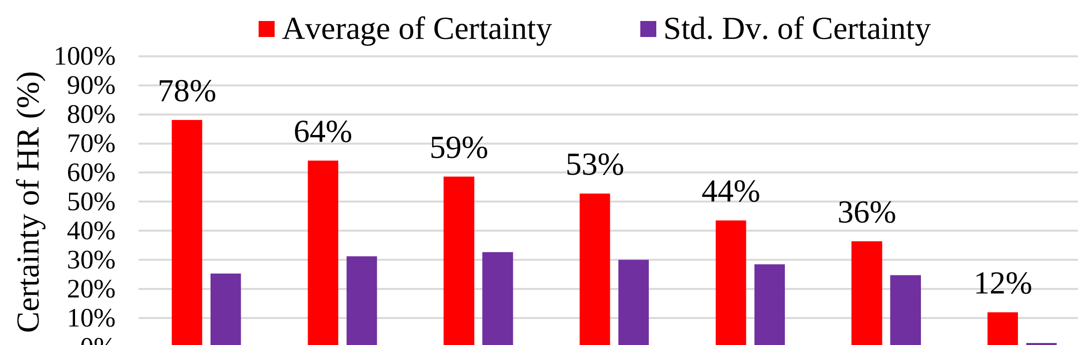
$$CFI(RC_n) = \frac{1}{99}\sum_{i=1}^{i=99} Q_n^i \ , \quad Q_n^i = \begin{cases} 1, & RC_n^{F(Q_n^i)} = RC_n^{Obs} \\ 0, & Otherwise \end{cases} , \qquad CFI_{avg} = \frac{1}{N}\sum_{i=1}^{N} CFI(RC_n)$$

The certainty forecasts for high-rate (HR) events by the adjusting approach for different ranges of ramp rates when threshold=0.4 pu/hr



Range of Rates, |p.u/hr|	[0-0.1)	[0.1-0.2)	[0.2-0.3)	[0.3-0.4)	[0.4-0.5)	[0.5-0.6)	[0.6-0.7)	[0.7-0.8)
Ramp Events	1787	1183	478	218	108	41	11	2

The certainty of forecasts for high-rate (HR) events by the adjusting approach with different thresholds, Tsh=0.1 to 0.7 pu/hr



Threshold |pu/hr| ≥	0.1	0.2	0.3	0.4	0.5	0.6	0.7
Ramp Events	2041	858	380	162	54	13	2

Modeling and Results

Data Description:

In this study the adjusting approach is implemented for intra-hour forecasts.
This study also serves as an out-of-sample test* to *verify* the adjusting approach performance.

- Golden – August 14, 2012, through September 24, 2013, semi-arid climate.
- Cocoa – January 21, 2011, through March 4, 2012, subtropical climate.
- Eugene – December 20, 2012, through January 20, 2014, marine west coast climate.

Data partition into training and testing sets:
The cross-validation strategy is adopted

Available data

Weather variables (measurements)	
Plane-of-Array (POA) Irradiance (W/m^2)	Amount of solar irradiance received on the PV panel surface
Back-Surface Temperature of PV Panel (°C)	PV panel back-surface temperature, measured behind the center of PV panel
Relative Humidity (%)	Relative humidity at the site
Precipitation (mm)	Accumulated daily total precipitation in millimeter
Direct Normal Irradiance (DNI) (W/m^2)	Amount of solar irradiance received within a 5.7° field-of-view centered on the sun
Global Horizontal Irradiance (GHI) (W/m^2)	Total amount of direct and diffuse solar irradiance received on a horizontal surface
Diffuse Horizontal Irradiance (DHI) (W/m^2)	Amount of solar irradiance received from the sky (excluding the solar disk) on a horizontal surface

3 Forecast horizons (15, 30, 60min) for 3 sites, (Golden, CO, Cocoa, FL, Eugene, OR) ➔9 cases.

* L. J. Tashman, Out-of-sample tests of forecasting accuracy: an analysis and review," International journal of forecasting, vol. 16, no. 4, pp. 437-450, 2000.

Modeling and Results

The individual intra-hourly forecasts of solar power

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - F_i)^2} \qquad MAE = \frac{1}{n}\sum_{i=1}^{n}|P_i - F_i| \qquad BIAS = \sum_{i=1}^{n}(P_i - F_i)$$

Location	Horizon	Persistence			NAR			ARIMA			ANN			ELM		
		RMSE	MAE	MBE	RMSE	MAE	MBE	RMSE	MAE	BIAS	RMSE	MAE	MBE	RMSE	MAE	MBE
Golden	15min	0.0344	0.0255	0.2394	0.0327	0.0235	-4.647	0.0346	0.0254	-0.0007	0.0344	0.0257	-5.276	0.0340	0.0254	-0.969
	30min	0.0481	0.0365	0.2113	0.0434	0.0322	-2.108	0.0459	0.0346	-0.3162	0.0432	0.0322	-7.525	0.0464	0.0345	-2.022
	60min	0.0715	0.0541	0.2113	0.0586	0.0438	-6.070	0.0608	0.0462	0.8859	0.0571	0.0431	-3.672	0.0646	0.0465	3.773
Cocoa	15min	0.0411	0.0303	0.3746	0.0390	0.0269	-6.028	0.0405	0.0287	0.3298	0.0384	0.0274	-9.703	0.0408	0.0302	-0.904
	30min	0.0553	0.0417	0.3005	0.0479	0.0315	-2.244	0.0470	0.0334	0.2949	0.0451	0.0307	2.389	0.0492	0.0325	-0.482
	60min	0.0870	0.0678	0.4302	0.0587	0.0420	-7.261	0.0595	0.0427	-0.1683	0.0562	0.0394	-2.768	0.0579	0.0413	0.107
Eugene	15min	0.0358	0.0235	0.2126	0.0360	0.0238	0.939	0.0355	0.0219	0.2201	0.0350	0.0215	-5.825	0.0344	0.0215	-2.752
	30min	0.0483	0.0344	0.2144	0.0425	0.0271	-7.997	0.0442	0.0281	-2.9847	0.0425	0.0270	3.761	0.0421	0.0270	-3.048
	60min	0.0738	0.0561	0.2144	0.0586	0.0397	0.187	0.0616	0.0415	-0.0845	0.0575	0.0397	-3.480	0.0568	0.0394	-3.485
Total Average		**0.0550**	**0.0411**	**0.2676**	**0.0464**	**0.0323**	**-3.9143**	**0.0477**	**0.0336**	**-0.2026**	**0.0455**	**0.0319**	**-3.5665**	**0.0474**	**0.0332**	**-1.0871**

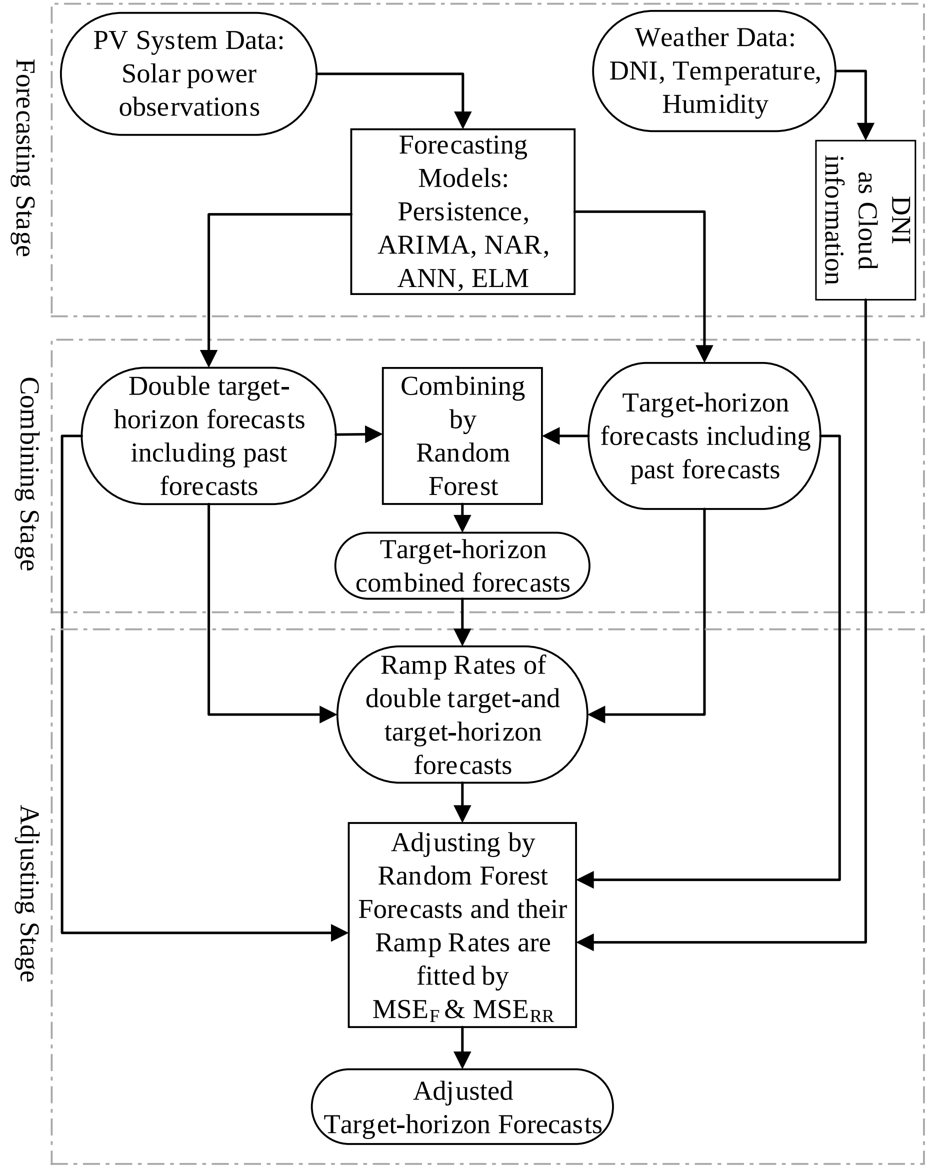
The aggregated* evaluation of the individual intra-hourly forecasts of solar power

Forecast	**Persist**	**NAR**	**ARIMA**	**ANN**	**ELM**
$RMSE_{avg}$	0.0550	0.0464	0.0477	**0.0455**	0.0474
MAE_{avg}	0.0411	0.0323	0.0336	**0.0319**	0.0332
$BIAS_{avg}$	0.2676	-3.9143	**-0.2026**	-3.5665	-1.0871

*Aggregation is conducted by taking the average of the evaluating values of all 3 horizons and 3 sites.

Modeling and Results

Implementing the **adjusting approach** to forecasts solar power ramp events

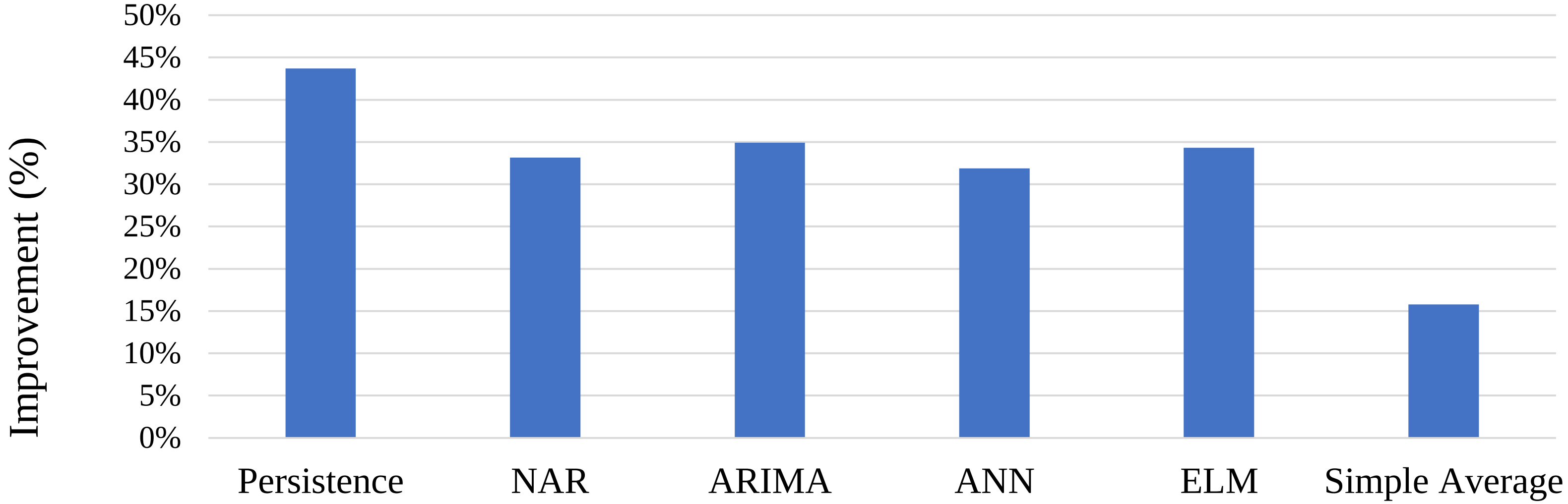


Block diagram of the **adjusting approach** for intra-hour forecasts of solar power and ramp events

Modeling and Results

Intra-Hour Forecasts

Average improvements of the combined forecasts by the adjusting approach with respect to other forecasts



Location	Forecast Horizon	RMSE						
		Persist.	NAR	ARIMA	ANN	ELM	Simple Average	Adjusting Approach
Golden	15min	0.0344	0.0327	0.0346	0.0344	0.0340	0.0322	0.0246
	30min	0.0481	0.0434	0.0459	0.0432	0.0464	0.0328	0.0280
	60min	0.0715	0.0586	0.0608	0.0571	0.0637	0.0484	0.0453
Cocoa	15min	0.0411	0.0389	0.0405	0.0384	0.0408	0.0288	0.0240
	30min	0.0553	0.0478	0.0470	0.0451	0.0484	0.0345	0.0288
	60min	0.0870	0.0587	0.0594	0.0562	0.0578	0.0511	0.0420
Eugene	15min	0.0358	0.0360	0.0355	0.0350	0.0344	0.0255	0.0193
	30min	0.0483	0.0425	0.0441	0.0425	0.0421	0.0313	0.0257
	60min	0.0738	0.0586	0.0607	0.0575	0.0568	0.0465	0.0411
Average		**0.0550**	**0.0464**	**0.0476**	**0.0455**	**0.0472**	**0.0368**	**0.0310**

Modeling and Results

Probabilistic Forecasts

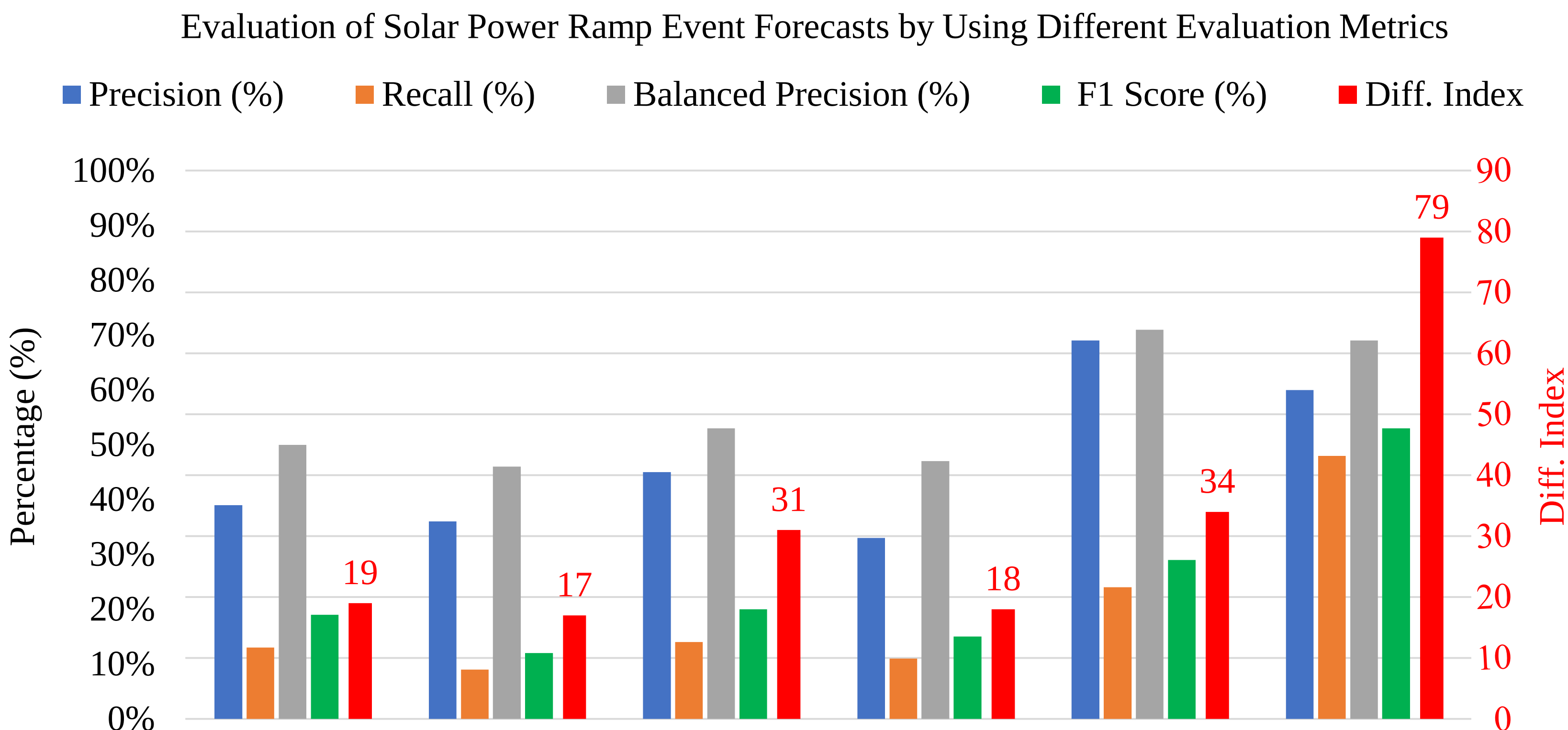
$$Pb_{q,i}(F_q, P_i) = \begin{cases} \left(1 - \frac{q}{100}\right)(F_q - P_i), & \text{if } P_i < Fq \\ \frac{q}{100}(P_i - F_q), & \text{if } P_i \geq F_q \end{cases}$$

Quantiles, $q \in [1- 99]$

$$\text{Pinball}, PB = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{99}\sum_{q=1}^{q=99} PB_{q,i}\right)$$

Location	Forecast Horizon	Pinball (PB)			
		Persistence	AnEn Simple Average	AnEn Adjusting Approach	Ensemble Adjusting Approach
Golden	15min	0.0236	0.0098	0.0071	0.0064
	30min	0.0271	0.0102	0.0084	0.0077
	60min	0.0289	0.0162	0.0141	0.0124
Cocoa	15min	0.0277	0.0082	0.0069	0.0063
	30min	0.0304	0.0101	0.0081	0.0073
	60min	0.0319	0.0166	0.0123	0.0109
Eugene	15min	0.0316	0.0067	0.0053	0.0046
	30min	0.0370	0.0087	0.0072	0.0062
	60min	0.0415	0.0148	0.0126	0.0106
Average		**0.0311**	**0.0113**	**0.0091**	**0.0080**

Modeling and Results



Forecasts	NAR	ARIMA	ANN	ELM	Simple Avg.	Adj. Approach
Precision (%)	39%	36%	45%	33%	69%	60%
Recall (%)	13%	9%	14%	11%	24%	48%
Balanced Precision (%)	50%	46%	53%	47%	71%	69%
F1 Score (%)	19%	12%	20%	15%	29%	53%
Diff. Index	19	17	31	18	34	79

Intra-hour forecasts of solar power ramp events of the high-rate ramp events, ($|Rate| \geq 0.1$ pu/dt) by the adjusting approach.

Modeling and Results

Comparison the hourly forecasts of solar power by the adjusting approach with different datasets

Location	Australia	Golden, CO	Cocoa, FL	Eugene, OR
RMSE	0.0523	0.0453	0.0420	0.0411
Pinball	0.0084	0.0124	0.0109	0.0106

Specifications of solar PV systems and data statistics

Dataset	**Golden, CO**	**Cocoa, FL**	**Eugene, OR**	**Canberra**
Country	USA	USA	USA	Australia
Climate type	Semi-arid	Subtropical	Marine coast	Oceanic
Latitude (°, -S)	39.74	28.39	44.05	-35.16
Longitude (°, -W)	-105.18	-80.46	-123.07	149.06
Elevation above sea (m)	1798	12	145	595
Number of panels	11	11	11	8
Panel tilt (°) from horizontal	40	28.5	44	36
Panel orientation (°) clockwise from North	180	180	180	38
Total capacity (W)	1252	1272	1290	1560
Time period of observations	Aug. 2012 to Sep. 2013	Jan. 2011 to March 2012	Dec. 2012 to Jan. 2014	April 2012 to May 2014
Data resolution	15min	5min	5min	1hr
Missing (% of observations)	18%	17%	10%	0%
Variability (data resolution) Std.Div.	(15min) 0.256 **(1hr) 0.119**	(5min) 0.251 **(1hr) 0.164**	(5min) 0.250 **(1hr) 0.161**	**(1hr) 0.259**

Conclusions

- ✓ The adjusting approach improves the combined forecasts.
- ✓ The approach is simple vs. the classification techniques.
- ✓ Ramp classes more separable with features: forecasts, ramp-rates, clouds, neighboring PVs.
- ✓ Most effective weather variables:
  - Cloud information: cloud type, height and cloud formation.
  - Clear-sky solar irradiance / Top solar irradiance at Earth's atmospheric layer.
- ✓ Diff. Index of high-rate ramps is efficient for the imbalanced classification.
- ✓ Probabilistic forecasts as a tool of situational awareness.

Future Work

Recommendations and further work for this dissertation are as follows:

- A direct comparison between the satellite-driven forecasts and the proposed approach.
- Datasets with more different levels of variability of solar power.
- Applying the adjusting approach to forecast wind power ramp events.
- Assimilating power of neighboring PV systems, sky imaging and satellite data.
- Using operational weather forecasts of high-resolution rabid refresh (HRRR) model.

https://www.researchgate.net/publication/328757565_A_Post-processing_Approach_for_Solar_Power_Combined_Forecasts_of_Ramp_Events/download

Vita

Dissertation-based Publications

1. M. Abuella and B. Chowdhury, “Solar Power Probabilistic Forecasting by Using Multiple Linear Regression Analysis,” in IEEE Southeast Con. Proceedings, 2015.
2. M. Abuella and B. Chowdhury, “Solar Power Forecasting using Artificial Neural Networks,” in North American Power Symposium (NAPS), 2015.
3. M. Abuella and B. Chowdhury, “Solar Power Forecasting Using Support Vector Regression,” in Proceedings of the American Society for Engineering Management 2016 International Annual Conference, 2016.
4. M. Abuella and B. Chowdhury, “Random forest ensemble of support vector regression models for solar power forecasting,” in 2017 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT), 2017.
5. M. Abuella and B. Chowdhury, “Hourly probabilistic forecasting of solar power,” in 2017 North American Power Symposium (NAPS), 2017.
6. 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.
7. M. Abuella and B. Chowdhury, “Qualifying Combined Solar Power Forecasts in Ramp Events’ Perspective,” in IEEE Power & Energy Society General Meeting, 2018.
8. M. Abuella, & B. Chowdhury, “Improving Combined Solar Power Forecasts Using Estimated Ramp Rates: Data-driven Post-processing Approach,” IET Renewable Power Generation Journal, 2018.
9. M. Abuella, & B. Chowdhury, “Forecasting of Solar Power Ramp Events: A post Processing Approach,” Renewable Energy, 2018.

Thanks for Your Listening

Any Question?

A Post-processing Approach for Solar Power Combined Forecasts of Ramp Events

Mohamed Ali Abuella
mabuella@uncc.edu

https://epic.uncc.edu

Energy Production and Infrastructure Center
University of North Carolina at Charlotte

UNC CHARLOTTE
Energy Production and Infrastructure Center (EPIC)