

Intra-hour Forecasts of Solar Power and Ramp Events[☆]

Mohamed Abuella and Badrul Chowdhury

Department of Electrical and Computer Engineering, Energy Production and Infrastructure Center, University of North Carolina at Charlotte, NC 28223 USA

Abstract

In this study an adjusting post-processing approach is implemented for improving intra-hourly forecasts of solar power and ramp events of PV solar power systems at different locations in the United States. This study also serves as an out-of-sample test to evaluate the performance of the adjusting approach with different locations and timescales. Thus, various individual intra-hourly forecasts of solar power are combined and adjusted by applying the adjusting approach. Both point and probabilistic forecasts of solar power are included. After that, solar power ramp event forecasting by the adjusting approach is carried out.

Keywords: Adjusting, classification, ensemble, post-processing, ramp events, solar power forecast.

1. Data Description

The data description and the specifications of PV solar systems are presented in Table 1, the data were acquired from the National Renewable Energy Laboratory, NREL [1]. Another dataset of lower temporal resolution with hourly observations of solar power is also included in the rightmost column of Table 1. The latter dataset is from Australia [2], which was previously used for the proposed adjusting approach to improve the hourly forecasts of solar power and

[☆]This preprint is a part of a PhD dissertation, which may be cited as:
M. Abuella, A Post-processing Approach for Solar Power Combined Forecasts of Ramp Events. PhD Dissertation, The University of North Carolina at Charlotte, 2018, pp. 94 - 106.

1. DATA DESCRIPTION

ramp events, the results were reported in [3, 4].

The datasets of PV systems at the U.S. sites have higher temporal resolutions, wherein the original resolutions are 15-min for Golden, CO, 5-min for Cocoa, FL, and Eugene, OR. The U.S. data is adopted for evaluating the adjusting approach with intra-hour forecasts, besides of being as an out-of-sample test of the forecasting accuracy [5].

Table 1: Data description and specifications of PV solar systems

Dataset	Golden, CO	Cocoa, FL	Eugene, OR	Canberra
Country	USA	USA	USA	Australia
Climate type	Semi-arid	Subtropical	Marine coast	Oceanic
Latitude ($^{\circ}$, -S)	39.74	28.39	44.05	-35.16
Longitude ($^{\circ}$, -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 ($^{\circ}$) from horizontal	40	28.5	44	36
Panel orientation ($^{\circ}$) 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

Unfortunately, there are missing values of the intra-hourly data, conversely

1. DATA DESCRIPTION

to the previously used high quality data of the PV site in Australia, which is only with hourly observations of solar power, and hence, not suitable for intra-hour forecasts. If the missing values in the data of U.S. sites are neglected, this can impact the ramp events modeling and forecasting. Therefore, those missing values are interpolated to fill the temporal gaps in the solar power time-series.

To obtain a consistent duration of the ramp rates, 3 durations are chosen (15-min, 30-min, and 60-min), and these durations are also the rolling windows for intra-hour forecasts of solar power and ramp events. It can be noticed in Table 1 that the variability (i.e., standard deviation) of the U.S. data decreases as the data resolution becomes lower. However, the Australian data with only 1-hour resolution has the highest variability (st.div.=0.259).

Using the data from the U.S. sites with 3 forecast horizons (15, 30, and 60-min) for each site of the available sites, the number of case studies is 9.

The U.S. data are associated with measurements of several weather variables listed in Table 2.

Table 2: Measured weather variables that are associated with data of PV systems at the U.S. sites

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 ($^{\circ}C$)	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

2. METHODOLOGY

2. Methodology

The adjusting approach, as described in the former published paper [3], is now modified to include some adjustments for determining intra-hourly forecasts of solar power and ramp events. The procedure is depicted in Figure 1.

Remark The advancements of High-Resolution Rapid Refresh (HRRR) model, which is run by the National Oceanic and Atmospheric Administration (NOAA), made it possible to produce hourly forecasts of weather variables. However, powerful computing equipment and big data tools are required for modeling solar ramp events efficiently with those HRRR forecasts in terms of storage size and computation speed. Moreover, the weather forecasts are not yet available in intra-hourly timescale, and despite the high accuracy of HRRR weather forecasts, some of the extreme ramp events are still unpredictable [6, 7].

Since this study is focused on very short-term forecasts with U.S. data for horizons up to 1-hour, the available meteorological measurements in the U.S. may be used as an alternative to the weather predictions in the Australian data, which were used for hourly forecasts.

Statistical time-series models are employed to generate the individual forecasts of solar power, including ARIMA, NAR, ANN, and Extreme Learning Machine (ELM).

At the combining stage of the adjusting approach, as shown in Figure 1, the double target-horizon forecasts are combined with target-horizon forecasts, for which the adjusting is performed. Some available meteorological measurements are assimilated in the adjusting approach. The combined meteorological data are the temperature of PV panel, the relative humidity, and the direct normal irradiance (DNI). The intra-hourly information of the cloud cover can be delivered by the DNI [8].

3. Results

The performance of the individual forecasts are evaluated by using the RMSE, MAE, and MBE, for the 3 forecast horizons of the 3 sites of the U.S.

3. RESULTS

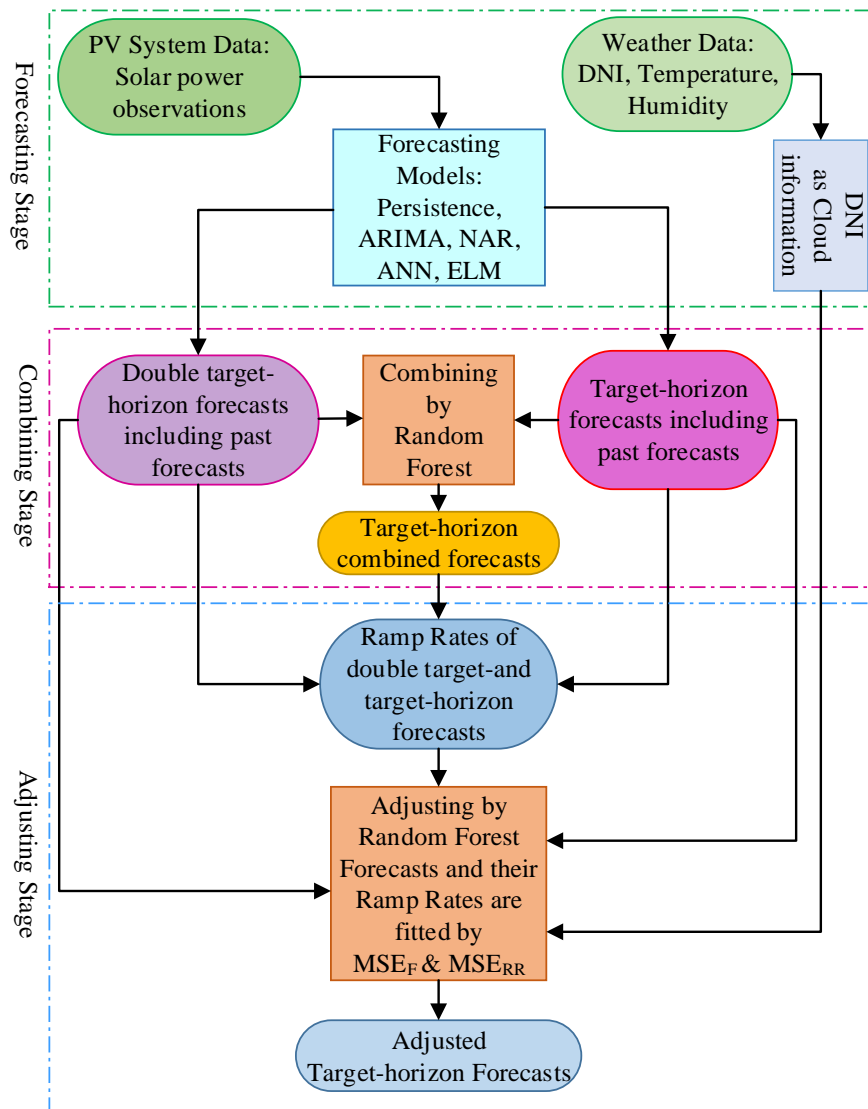


Figure 1: Block diagram of the adjusting approach for intra-hour forecasts of solar power and ramp events

data, as shown in Table 3.

The last row in Table 3 provides the aggregated evaluation of each individual forecast, by taking the average of the evaluating values of all 3 horizons and 3

3. RESULTS

Table 3: The individual intra-hourly forecasts of solar power

Location	Forecast Horizon	Persistence			NAR			ARIMA			ANN			ELM		
		RMSE	MAE	MBE	RMSE	MAE	MBE	RMSE	MAE	MBE	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	0.034	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	-6.07	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	0.047	0.0334	0.2949	0.0451	0.0307	2.389	0.0492	0.0325	-0.482
	60min	0.087	0.0678	0.4302	0.0587	0.042	-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	0.036	0.0238	0.939	0.0355	0.0219	0.2201	0.035	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	0.027	3.761	0.0421	0.027	-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
Average		0.055	0.041	0.268	0.046	0.032	-3.914	0.048	0.034	-0.203	0.046	0.032	-3.567	0.047	0.033	-1.087

sites (i.e., averaging each column), which are rearranged and represented by $RMSE_{agg}$, MAE_{agg} , and MBE_{agg} in Table 4.

Table 4: The aggregated evaluation of the individual intra-hourly forecasts of solar power

Forecast	Persistence	NAR	ARIMA	ANN	ELM
$RMSE_{agg}$	0.0550	0.0464	0.0477	0.0455	0.0474
MAE_{agg}	0.0411	0.0323	0.0336	0.0319	0.0332
MBE_{agg}	0.2676	-3.9143	-0.2026	-3.5665	-1.0871

The RMSE and MAE have the same trends, and they indicate that in some cases of the individual forecasts, especially at the shorter horizon (15-min), they do not always outperform the persistence forecasts. The ANN produces the most accurate forecasts (RMSE=0.0455) with an average RMSE improvement equalling to 17% over the persistence forecasts. Meanwhile, the ARIMA forecasts (RMSE=0.0477) obtain 13% of an average RMSE improvement over the persistence forecasts (RMSE=0.0550). The MBE values of the individual forecasts are different in magnitude and sign, and so, the individual forecasts are diverse, and this is crucial in order to efficiently combine those forecasts by the ensemble learning in the adjusting approach.

The simple average method is also employed to combine the individual forecasts for a comparison with the intra-hourly combined forecasts of solar power

3. RESULTS

by applying the adjusting approach. The diagram of the combining method by the simple average is shown in Figure 2.

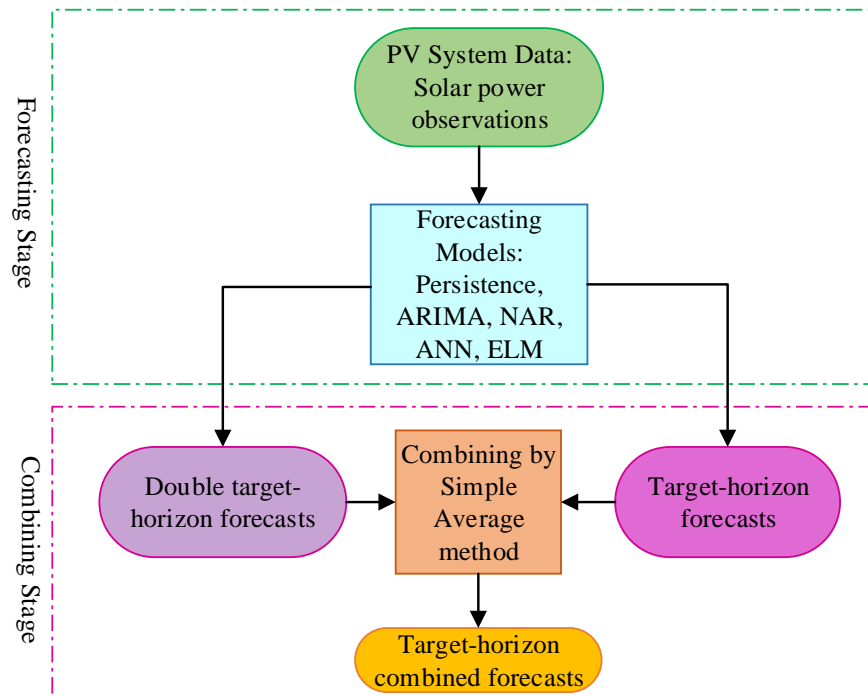


Figure 2: Block diagram of the simple average method for combining intra-hour forecasts of solar power and ramp events

Table 5 presents the combined forecasts of solar power by the simple average (Simple Average) and the combined forecasts (Adjusting Approach) by the adjusting approach.

As expected, the combined forecasts even by the simple average outperform the individual forecasts, and an additional improvement is achieved by applying the adjusting approach. From the last row in Table 5, the average RMSE improvement of the combined forecasts (RMSE=0.0310) from the adjusting approach is about 16% over the combined forecasts (RMSE=0.0368) by the simple average and 44% over the persistence forecasts (RMSE=0.0550). A graph of average improvements of the combined forecasts by the adjusting approach over

3. RESULTS

Table 5: Individual and combined forecasts of solar power

Location	Forecast Horizon	RMSE						
		Persistence	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

other forecasts is shown in Figure 3.

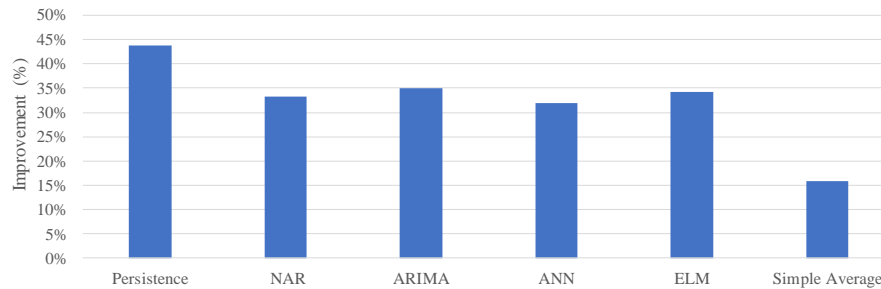


Figure 3: Average improvements of the combined forecasts by the adjusting approach with respect to other forecasts

As in the previous study, the probabilistic forecasts are also utilized to quantify the uncertainty of the intra-hourly point forecasts of solar power. The probabilistic forecasts including ensemble-based probabilistic forecasts (Ensemble), Analog Ensemble (AnEn), and the persistence probabilistic forecasts (Persistence).

The intra-hourly combined forecasts from the adjusting approach for different locations and timescales are also evaluated by the DM test [9]. Table 6

3. RESULTS

indicates the adjusted combined forecasts outperform all other time-series forecasts, as demonstrated by the DM test, which evaluates the significant accuracy differences of the adjusted combined forecasts with respect to other forecasts.

Table 6: The DM test of the intra-hourly combined forecasts by the adjusting approach over other forecasts

Location	Forecast Horizon	DM Test (%)					
		Persist.	NAR	ARIMA	ANN	ELM	Simple Average
Golden	15min	25.22	17.4	25.20	25.74	24.90	33.34
	30min	28.14	22.35	26.12	22.64	25.35	15.15
	60min	19.48	10.43	15.56	11.49	10.70	5.87
Cocoa	15min	34.81	25.06	30.77	31.46	34.63	20.44
	30min	28.40	13.70	22.01	16.29	11.71	12.61
	60min	28.12	13.34	15.46	12.92	13.51	15.47
Eugene	15min	25.47	26.00	22.72	22.49	23.32	16.53
	30min	24.96	15.90	16.94	16.94	16.95	9.85
	60min	24.90	12.30	11.70	13.06	12.90	7.20
Average		26.61	17.39	20.72	19.23	19.33	15.16

As shown in Table 7, the AnEn probabilistic forecasts are used to quantify the uncertainty of the combined forecasts from two combining methods - the simple average and the adjusting approach. The uncertainty of the combined forecasts by the adjusting approach is also quantified by the ensemble-based probabilistic forecasts, which are provided in the rightmost column of the table.

Table 8, presents the evaluation of intra-hourly probabilistic forecasts of solar power by using CRPS instead of pinball as an evaluation metric for the probabilistic forecasts. Although the evaluating values of pinball and CRPS are different, the improvements by pinball and CRPS are almost the same. However, in terms of pinball (PB), the average improvements of the ensemble-based

3. RESULTS

Table 7: Pinball of the intra-hourly probabilistic forecasts of solar power

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

probabilistic forecasts (PB=0.0080) by the adjusting approach over the EnAn of adjusting approach (PB=0.0091) and the EnAn of simple average (PB=0.0113) are 12% and 29% respectively. It should be noted that the average improvement of (PB=0.0080) is about 74% over the persistence probabilistic forecasts (PB=0.0311).

Forecasting of solar power ramp events is also carried out with these intra-hourly data at sites in the U.S. Table 9 shows the statistics of solar power ramp rates with different thresholds to define the ramp rate as high or low, we observed that the maximum ramp rate is 0.487 pu/dt occurs at the Cocoa, FL, site with a temporal resolution equalling to 30-min. In provirus study with the Australian data, the maximum ramp rate was about 0.8 pu/hr. The number of high-rate ramp events is reduced significantly by increasing the threshold. For instance, at threshold=0.4 pu/dt, the total number of high-rate ramp events is 6 events only. Whereas, in the Australian data, when using the same threshold ($Tsh =$

3. RESULTS

Table 8: CRPS of the intra-hourly probabilistic forecasts of solar power

Location	Forecast Horizon	Continuous Ranked Probability Score (CRPS)			
		Persistence	AnEn Simple Average	AnEn Adjusting Approach	Ensemble Adjusting Approach
Golden	15min	0.0466	0.0196	0.0141	0.0127
	30min	0.0535	0.0202	0.0167	0.0153
	60min	0.0572	0.0322	0.0281	0.0245
Cocoa	15min	0.0546	0.0164	0.0136	0.0125
	30min	0.0600	0.0201	0.0161	0.0145
	60min	0.0630	0.0330	0.0245	0.0216
Eugene	15min	0.0623	0.0134	0.0106	0.0091
	30min	0.0731	0.0173	0.0143	0.0123
	60min	0.0819	0.0294	0.0251	0.0210
Average		0.0614	0.0224	0.0181	0.0159

0.4 pu/hr), the number of high-rate events is 162. The lower temporal resolution of the U.S. data are obtained by aggregating the higher original resolution data of each location, with 15 min, 30 min, and 60 min averages over the entire year. In addition, as shown in Table 1, the Australian data have the highest variability (i.e., Std. Dev.=0.251 at 1 hr temporal resolution). The medium value of the solar power ramp rates of those data is about 0.2 pu/dt , but it is about 0.4 pu/hr in the Australian data.

The intra-hourly forecasting of solar power ramp events is conducted with two thresholds to define the high and low ramp events, $|Rate| \geq 0.1 pu/dt$ and $|Rate| \geq 0.2 pu/dt$, as shown in Figure 4 and 5, respectively.

Figure 6 illustrates the forecasts of solar power ramp events by implementing the classification techniques. The SVM and RF techniques achieve the most accurate forecasts. Since the combined forecasts of solar power ramp events by the adjusting approach (Diff. Index=79) are included as input variable in

3. RESULTS

Table 9: Statistics of intra-hourly data of the solar power observations for solar power ramp rates. (a) at different thresholds, $0.1 pu/dt$ to $0.4 pu/dt$, (b) some statistical measures of the ramp rates

(a)

Location	Temporal Resolution	Total Ramps	Ramp Rates at Different Thresholds							
			$ Rate = 0.1 pu/dt$		$ Rate = 0.2 pu/dt$		$ Rate = 0.3 pu/dt$		$ Rate = 0.4 pu/dt$	
			High-Rate Ramps	Low-Rate Ramps	High-Rate Ramps	Low-Rate Ramps	High-Rate Ramps	Low-Rate Ramps	High-Rate Ramps	Low-Rate Ramps
Golden	15min	13142	170	12972	9	13133	7	3282	0	3289
	30min	6575	246	6329	21	6554	2	6573	0	6575
	60min	3289	472	2817	40	3249	2	13140	0	13142
Cocoa	15min	13367	351	13016	23	13344	13	3349	2	3360
	30min	6694	382	6312	39	6655	11	6683	2	6692
	60min	3362	734	2628	89	3273	6	13361	0	13367
Eugene	15min	14615	305	14310	44	14571	5	3672	0	3677
	30min	7332	303	7029	41	7291	7	7325	2	7330
	60min	3677	600	3077	60	3617	13	14602	0	14615

(b)

Location	Temporal Resolution	Total Ramps	Statistical Measures of Ramp Rates, $ Rate $					
			Min. Value	Max. Value	Medium Value	Median	Mean	Std. Div.
Golden	15min	13142	0	0.373	0.187	0.023	0.030	0.026
	30min	6575	0	0.353	0.176	0.031	0.039	0.033
	60min	3289	0	0.398	0.199	0.042	0.054	0.047
Cocoa	15min	13367	0	0.364	0.182	0.025	0.033	0.030
	30min	6694	0	0.487	0.243	0.027	0.035	0.033
	60min	3362	0	0.411	0.205	0.025	0.035	0.034
Eugene	15min	14615	0	0.387	0.193	0.024	0.032	0.031
	30min	7332	0	0.422	0.211	0.026	0.035	0.032
	60min	3677	0	0.384	0.192	0.026	0.035	0.032

the classification techniques, this leads to slightly more accurate forecasts of solar power ramp events by the SVM model (Diff. Index=82). Excluding the adjusted combined forecasts of solar power ramp events from the classification methods yields lower Diff. Index compared to the adjusting approach.

It may be useful to compare all datasets from Australia and U.S. sites within the same timescale. Table 10 shows the comparison of hourly probabilistic forecasts by the adjusting approach when using different datasets. The forecasts

3. RESULTS

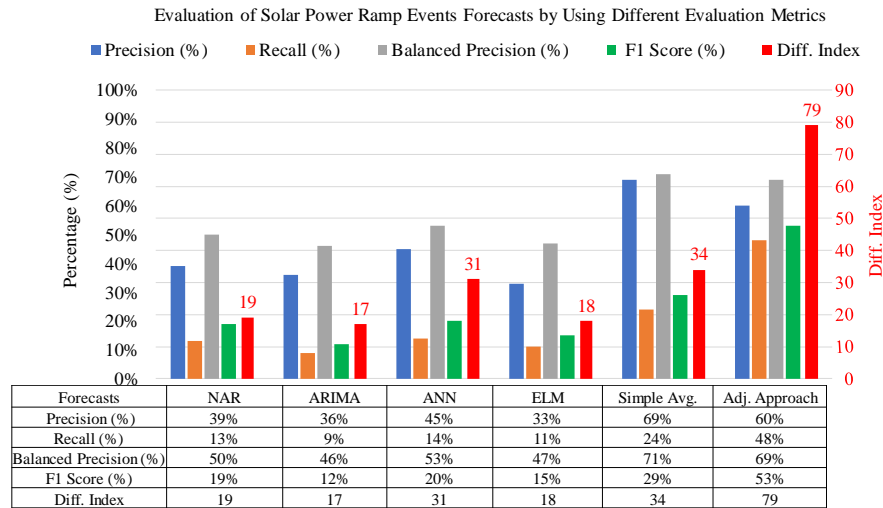


Figure 4: Forecasts of solar power ramp events with different evaluation metrics of high-rate ramp events, when $|Rate| \geq 0.1 \text{ pu/dt}$

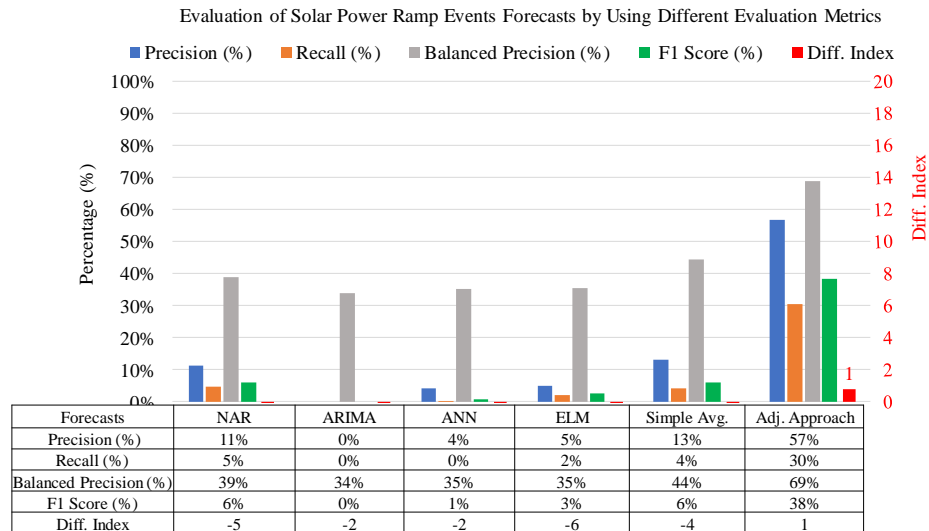


Figure 5: Forecasts of solar power ramp events with different evaluation metrics of high-rate ramp events, when $|Rate| \geq 0.2 \text{ pu/dt}$

in the Australian case are driven by NWP, which reduces the uncertainty and improves the probabilistic forecasts in the range of 20% - 33%.

4. SUMMARY

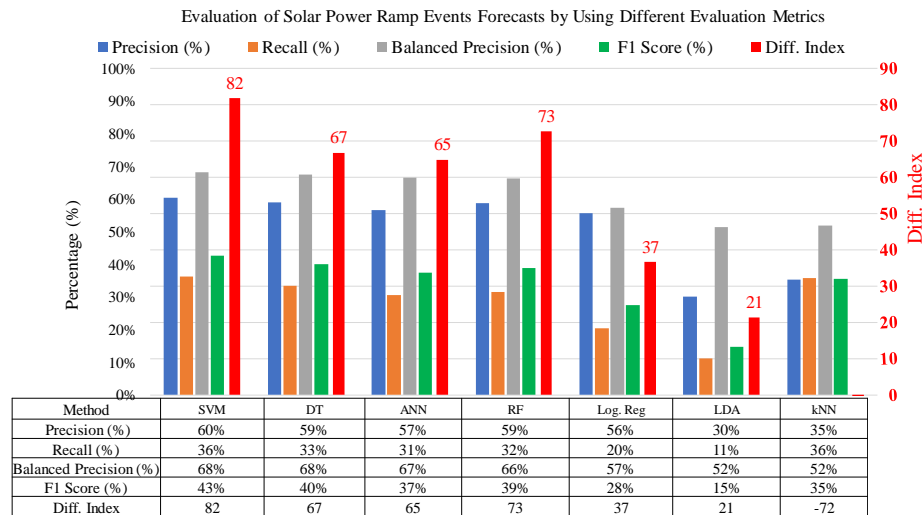


Figure 6: Classification techniques for forecasting of solar power ramp events with different evaluation metrics of high-rate ramp events, when $|Rate| \geq 0.1 pu/dt$

Table 10: Pinball and CRPS of the hourly probabilistic forecasts of solar power by the adjusting approach with different datasets

Location	Australia	Golden, CO	Cocoa, FL	Eugene, OR
Pinball	0.0084	0.0124	0.0109	0.0106
CRPS	0.0169	0.0245	0.0216	0.0210

4. Summary

In an overall evaluation, the individual solar power forecasts (NAR, ARIMA, ANN, and ELM) outperform the persistence forecasts, but they are not efficient for high-rate ramp events.

The adjusting approach with the sub-hourly resolution data at the U.S. sites, adjusts and improves the combined forecasts, especially at the ramp events that are impacted by combining individual forecasts.

The probabilistic forecasts of the shorter horizons are more accurate than those of the longer horizons. The ensemble-based probabilistic forecasts are more accurate than the analog ensemble-probabilistic forecasts (by 12% in the

4. SUMMARY

U.S. case, and 1% in the Australian case). The pinball improvements of the probabilistic forecasts are higher than the RMSE improvements that are obtained by point forecasts. For 1-hr horizon, the pinball is 0.0084 for the Australian case, 0.0124 for the data from Golden, CO, 0.0109 for Cocoa, FL, and 0.0106 for Eugene, OR. In the Australian case, the forecasts are driven by NWP, which reduces the uncertainty and gains improvements in the range of 20% - 33%.

Forecasting solar power ramp events is different and more challenging than forecasting the solar power which include both the normal and ramp events. The adjusting approach leads to the most accurate forecasts of solar power ramp events throughout various thresholds, locations and forecast horizons.

The classification techniques need a feature selection for each threshold, location, and horizon which increases the modeling complexity.

References

- [1] B. Marion, B. Smith, Photovoltaic system derived data for determining the solar resource and for modeling the performance of other photovoltaic systems, *Solar Energy* 147 (2017) 349–357.
- [2] Global Energy Forecasting Competition 2014, Probabilistic solar power forecasting, [Online]. Available: <http://www.crowdanalytix.com/contests/global-energy-forecasting-competition-2014>.
- [3] M. Abuella, B. Chowdhury, Forecasting of solar power ramp events: A post-processing approach, *Renewable Energy*.
- [4] M. A. Abuella, A Post-processing Approach for Solar Power Combined Forecast of Ramp Events, Ph.D. thesis, The University of North Carolina at Charlotte (2018).
- [5] L. J. Tashman, Out-of-sample tests of forecasting accuracy: an analysis and review, *International journal of forecasting* 16 (4) (2000) 437–450.

4. SUMMARY

- [6] K. D. Orwig, M. L. Ahlstrom, V. Banunarayanan, J. Sharp, J. M. Wilczak, J. Freedman, S. E. Haupt, J. Cline, O. Bartholomy, H. F. Hamann, et al., Recent trends in variable generation forecasting and its value to the power system, *IEEE Transactions on Sustainable Energy* 6 (3) (2015) 924–933.
- [7] R. P. Worsnop, M. Scheuerer, T. M. Hamill, J. K. Lundquist, Generating wind power scenarios for probabilistic ramp event prediction using multivariate statistical post-processing, *Wind Energy Science* 3 (1) (2018) 371–393.
- [8] S. Letendre, M. Makhyoun, M. Taylor, Predicting solar power production: Irradiance forecasting models, applications, and future prospects, Tech. rep., Solar Electric Power Association, Washington, D.C. (2014).
- [9] F. X. Diebold, R. S. Mariano, Comparing predictive accuracy, *Journal of Business & economic statistics* 20 (1) (2002) 134–144.