cnMAE

The Right Metric for Evaluating Solar Forecasts

The Right Metric for Evaluating Solar Forecasts

With increasing number of solar photovoltaic power plants coming online, solar forecasting is becoming more important for many players within the power and utility industry. Solar forecasting uses high-resolution weather data and AI-powered statistical modeling to predict the amount of solar energy that will be available at a specific time and location. This way:

- Independent power producers (IPPs) and energy traders can use short-term solar forecasts to determine how much power to sell into day-ahead markets, and whether to reserve any for real-time markets. This helps them determine whether the market is currently over- or underpriced when making decisions about mid- and long-term hedging.

- Similarly, a growing number of utilities use solar forecasts to predict what their net demand will be. These utilities have solar generation capacity, as well as consumer load, and they need to know in advance how well their solar assets will perform to avoid buying additional power in expensive, real-time markets.

With these growing use cases, it's just as important to understand how accurate your forecasts are. The best solar forecasts are highly accurate. However, any forecast is going to have a margin of predictive error. "If you're using solar forecasts, it's important to know just how accurate they are, because errors can be magnified in your trading strategies," explained Amperon's Executive Vice President Elliot Chorn. "For instance, if you decide to sell 100 percent of your solar generation, but your forecast is significantly off, you may only sell 90 percent of it. Or worse, you may sell 110 percent."

The utility industry has several conventional metrics for assessing predictive error in load forecasting. But none of them are the right fit for the unique characteristics of solar and solar forecasting. Amperon proposes the industry standardize around a new metric, cnMAE, to overcome limitations with the following metrics.

MAPE Is Mathematically Unstable

Mean absolute percentage error (MAPE) is perhaps the most common accuracy metric for load forecasting. MAPE is calculated by first subtracting actual load or generation from the predicted amount to determine the size of the difference — or "error." Then the (absolute) error value is divided by the actual observed amount (or percentage) error. This is calculated for each hour and then averaged over a period of weeks or months to determine MAPE.

$$\text{MAPE} = 100\frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

So, for example, a demand forecast may predict electricity consumption of 20 units for tomorrow between 12:00-1:00 pm. If the actual consumption turns out to be 25 units, the first step in calculating MAPE is to determine the absolute difference between the forecast and the actual consumption for that hour. In this instance, it is 5 units. Next, the difference (or error) is stated as a percentage of the predicted amount. Five divided by 25 gives the percentage: 20%. You would then do the same calculation for each hour and average them out to get the MAPE for the period you are evaluating.

MAPE provides a clear measure of load forecasting accuracy that is easy to understand and interpret. Lower MAPE values indicate more accurate forecasts — and by extension, better forecasting models and/or data. However, the calculation can fail when applied to solar generation forecasting.

With solar generation, it's not unusual for actual figures to be zero. In those instances, the standard MAPE calculation would then require dividing the error amount by zero, which is mathematically impossible. Very small figures approaching zero also impair the usefulness of the metric by hyper-inflating the percentage calculation.

nMAE Exaggerates Seasonal Variation

Normalized mean absolute error (nMAE) solves the instability problem that can occur when using MAPE for solar forecasting, but it introduces a separate problem.

$$\text{NMAE} = \frac{100\%}{\text{norm}} \cdot \frac{1}{n} \sum_{i=1}^{n} |F_i - O_i|$$

Normalization does away with zero- and near-zero figures before determining the percentage error. Typically, normalization is done by averaging the actual observed figures from the week or month, before using it to calculate daily-error percentages. Normalizing actual solar generation figures keeps them from approaching zero and throwing off the metric. However, nMAE doesn't account for seasonal variations in solar output capacity. So, in the winter, when average solar generation is minimal, it often exaggerates forecast errors, and when generation is at its highest in the summer, nMAE can depreciate errors.

Consider, for example, a winter forecast that is for two units of solar production, but only one actual unit (on average) is generated. The error would be 1 unit, and the nMAE would be 100%. Now consider a summer forecast that is for 11 units of solar production, but only 10 units are produced (on average). The error is the same: 1 unit. But now the nMAE would be only 10%.

nMAE is good for comparing forecasting models that deal with different types of data. But because of seasonal variations in solar radiance, it requires further normalization to make it a useful benchmark for solar forecasting.

RMSE Isn't Relative to Size

Root Mean Squared Error (RMSE) is another common forecasting metric with only limited usefulness in solar forecasting.

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

Like the metrics above, RMSE is calculated by first determining the difference between predicted amounts and the actual observed amounts for a given period. For RMSE, each difference is then squared to remove negative values and to give greater weight to larger errors. Next, the squared differences are averaged, and the final step is to take the square root, which returns the value to the original unit-scale of the data.

RMSE, like nMAE, overcomes the mathematical instability of dividing by zero or near-zero values. RMSE is more consistent than nMAE across seasonal variations in solar capacity, however, it is sensitive to scale. Because RMSE is stated in the original units of measurement, it doesn't account for the size of the solar installation or the average yield from that asset over the period we're examining.

RMSE of 250 kilowatts (kw) for a 20 megawatts (MW) solar farm would be an order of magnitude more accurate than the same RMSE for a 2MW solar farm. But because the relative size of the installation is not represented in the metric, RMSE is inappropriate for comparing forecasts across a range of installation sizes.

cnMAE: The Goldilocks Metric

After considering numerous metrics, including many proposed at Solar Forecast Arbiter, Amperon concluded that the most useful measurement of predictive error for solar forecasting is nMAE with an additional normalization step for capacity: cnMAE, or capacity normalized mean absolute error.

"cnMAE doesn't have the shortcomings that MAPE and nMAE have for market participants that need to rely on solar forecasting," Chorn said. "Plus, it's a more valuable tool than RMSE for power producers that want to make decisions based on comparisons of assets within a varied portfolio of solar installations."

The benefits of cnMAE make it the just-right metric for solar and wind forecasting:
- Prevents Unintended Bias: Unlike other measurements, cnMAE maintains an unbiased context by providing capacity-normalized insights and reducing sensitivity to seasonality.
- Supports Comparability: cnMAE enables performance comparison between systems with different capacities. This makes it an ideal metric for diversified portfolios of single-site installations.
- Enhances Planning and Budgeting: Power producers can better prioritize maintenance and operations with the capacity-adjust insights of cnMAE.

cnMAE normalizes for capacity by scaling the forecast error relative to a solar system's rated nameplate capacity. Simply put, cnMAE = MAE/Capacity.

$$cnMAE = \frac{100\%}{nC} \sum_{i=1}^{n} |F_i - O_i|$$

cnMAE in the Context of a Day

In addition to conventional forecasting metrics, Amperon provides its solar forecasting clients with cnMAE, which has consistently proven more useful. Take for example, a 20MW installation owned by one of our Midwestern IPP clients. The table below shows the actual hourly generation for a day in April, followed by the forecasted generation, the absolute error, the MAPE and the cnMAE.

	Actual	AMP DA	AbsErr	MAPE	nMAE	RMSE	cnMAE
⊟ 14	7.77	5.94	2.37	91.55%	30.50%	3.61	11.86%
1	-0.04	-0.04	0.00	0.00%	0.00%	0.00	0.00%
2	-0.04	-0.03	0.01	50.00%	25.00%	0.02	0.08%
3	-0.04	-0.04	0.00	0.00%	0.00%	0.00	0.01%
4	-0.04	-0.02	0.02	50.00%	50.00%	0.02	0.10%
5	-0.04	-0.02	0.02	50.00%	50.00%	0.02	0.12%
6	-0.04	0.02	0.06	150.00%	150.00%	0.06	0.29%
7	-0.04	0.44	0.48	1200.00%	1,200.00%	0.48	2.39%
8	2.40	1.95	0.45	18.75%	18.75%	0.45	2.26%
9	5.57	8.88	3.31	59.43%	59.43%	3.31	16.55%
10	18.10	10.42	7.68	42.43%	42.43%	7.68	38.39%
11	20.04	13.86	6.18	30.89%	30.84%	6.19	30.95%
12	19.99	15.51	4.48	22.41%	22.41%	4.48	22.40%
13	20.03	16.58	3.45	17.22%	17.22%	3.45	17.24%
14	20.01	15.96	4.05	20.24%	20.24%	4.05	20.23%
15	19.73	14.62	5.11	25.90%	25.90%	5.11	25.54%
16	10.92	13.31	2.39	21.98%	21.89%	2.40	11.99%
17	19.82	12.14	7.68	38.75%	38.75%	7.68	38.40%
18	18.22	10.71	7.51	41.22%	41.22%	7.51	37.55%
19	9.64	7.02	2.62	27.07%	27.18%	2.61	13.07%
20	2.53	1.24	1.29	50.99%	50.99%	1.29	6.43%
21	-0.05	0.09	0.14	280.00%	280.00%	0.14	0.71%
22	-0.04	-0.04	0.00	0.00%	0.00%	0.00	0.00%
23	-0.04	-0.04	0.00	0.00%	0.00%	0.00	0.00%
24	-0.04	-0.04	0.00	0.00%	0.00%	0.00	0.01%

For the hour ending (HE) at 1a.m. in the middle of the night, the prediction was spot on — no solar output. For the next hour — still in the middle of the night, the absolute error was miniscule: 0.01. But due to its mathematical instability, the MAPE is 50%. By comparison, the normalized cnMAE is a mere 0.08% and better reflects the insignificance of the error. The instability is much more noticeable at 7 a.m., when the sun is just beginning to come up. The forecast predicted a small amount of generation (0.44 MW), but there was none. In fact, demand can be <u>negative at night</u> when the photovoltaic inverters consume electricity. cnMAE shows the error as 2.39%. But because MAPE divides the predicted amount by the near-zero actual (-0.04), it calculates an enormous MAPE of 1,200%.

Over the course of the day, the absolute errors range from 0.0 to a high of 7.68. cnMAE tracks closely to the ups and downs in absolute error, while ranging from 0.0% to a high of 38.4%. By comparison, MAPE ranges from 0.0% to 1,2000% and does not correlate to the ups and downs of the absolute errors over the course of the day. The correlation and moderate range of cnMAE is radically more useful as an hour-to-hour measurement, and as a result, the average cnMAE for the day (11.86%) is inherently more accurate than the average MAPE (91.55%).

cnMAE in Seasonal Context

The advantages of cnMAE over nMAE is particularly obvious when looking at examples from two different times of the year. Table 2 shows averages for the same solar installation for the month of January and the month of April.

❶

	Actual	AMP DA	AbsErr	MAPE	nMAE	RMSE	cnMAE	MAE
2024	3.26	3.17	1.09	93.38%	33.44%	2.25	5.46%	1.09
Jan	1.07	1.01	0.58	55.24%	54.21%	1.41	2.88%	0.58
Apr	5.53	5.40	1.63	133.01%	29.48%	2.87	8.14%	1.63

❷

Date	Capacity	Monthly avg. forecast generation	Monthly avg. actual generation	Monthly avg. abs. error1	nMAE	cnMAE
Jan 2024	20 MW	1.01 MW/h	1.07 MW/h	0.58	54.21%	2.88%
Apr 2024	20 MW	5.40 MW/h	5.53 MW/h	1.63	29.48%	8.14%

For January, when solar radiation is typically low in the northern hemisphere, the average hourly forecast was for 1.01 MW of generation. This was very close to the actual average hourly generation of 1.07 MW. The monthly average of hourly absolute errors was 0.58, which is a small figure, but nonetheless produces a significant nMAE (54.21%) when divided by the low actual generation figure.

In April, the monthly average hourly absolute error was roughly three times greater (1.63) than in January, but because the actual generation figure was so much larger in April (5.53 MW/h) the nMAE score was minimized (29.48%).

By comparison, the cnMAE score normalizes these error figures against the capacity of the installation (20 MW) instead of the actual generation figure. Relative to the large size of the solar farm, the forecasting error in January is quite small, and this is reflected in the cnMAE score (2.88%). The April forecasting error was larger than the January error, and again, the cnMAE score (8.14%) reflects this.

cnMAE in Site Comparison

cnMAE also proves more useful than RMSE when comparing the accuracy of forecasts between solar installations of different sizes. Table 3 figures for the 20 MW solar farm from the examples above. Table 4 shows figures for a smaller installation, owned by the same IPP, that has a capacity of only 1.75 MW.

3

Date	Capacity	Forecast generation	Actual generation	Absolute error	RMSE	cnMAE
2/9 HE9	20 MW	0.45	1.46	1.01	1.01	5.05%

4

Date	Capacity	Forecast generation	Actual generation	Absolute error	RMSE	cnMAE
2/9 HE 11	1.75 MW	0.19	1.20	1.01	1.01	57.58%

Hour (HE)	20 MW Installation					1.75 MW Installation				
	Actual	AMP DA	AbsErr	cnMAE	RMSE	Actual	AMP DA	AbsErr	cnMAE	RMSE
9	3.05	1.92	1.15	5.73%	2.37	0.36	0.13	0.24	13.30%	0.41
1	-0.04	-0.04	0.00	0.01%	0.00	-0.01	-0.01	0.00	0.02%	0.00
2	-0.04	-0.04	0.00	0.01%	0.00	-0.01	0.00	0.01	0.06%	0.00
3	-0.04	-0.04	0.00	0.01%	0.00	-0.01	-0.01	0.00	0.00%	0.00
4	-0.04	-0.04	0.00	0.00%	0.00	-0.01	-0.01	0.00	0.00%	0.00
5	-0.04	-0.04	0.00	0.00%	0.00	-0.01	-0.01	0.00	0.00%	0.00
6	-0.04	-0.04	0.00	0.01%	0.00	-0.01	-0.01	0.00	0.01%	0.00
7	-0.04	-0.02	0.02	0.09%	0.02	-0.01	0.00	0.01	0.29%	0.01
8	0.00	0.02	0.02	0.11%	0.02	0.00	0.01	0.01	0.60%	0.01
9	1.46	0.45	1.01	5.05%	1.01	0.25	0.04	0.21	11.74%	0.21
10	2.90	1.70	1.20	6.01%	1.20	0.65	0.11	0.54	30.67%	0.54
11	9.11	3.03	6.08	30.39%	6.08	1.20	0.19	1.01	57.85%	1.01
12	7.27	5.46	1.81	9.07%	1.81	1.20	0.37	0.83	47.46%	0.83
13	13.65	7.33	6.32	31.57%	6.31	1.04	0.50	0.54	30.66%	0.54
14	14.69	7.82	6.87	34.36%	6.87	1.23	0.55	0.68	38.56%	0.68
15	8.13	7.74	0.39	1.94%	0.39	1.36	0.51	0.85	48.56%	0.85
16	8.22	7.45	0.77	3.82%	0.77	1.04	0.43	0.61	34.81%	0.61
17	6.93	4.92	2.01	10.04%	2.01	0.59	0.30	0.29	16.29%	0.29
18	1.46	0.60	0.86	4.28%	0.86	0.05	0.07	0.02	0.91%	0.02

Note that the absolute error is the same but on different hours for the two different solar farms. Relative to capacity of the 20 MW installation, the error is rather small, whereas it represents a much greater proportion of the potential capacity of the 1.75 MW installation. The RMSE score, which is the same in both tables doesn't reflect this perspective at all. cnMAE however does, making it a useful metric for making comparisons and prioritizations within a portfolio.

In Conclusion

Predictive error, though inevitable, can have a profound impact on decision making and outcomes that rely on forecasting in the utility industry. Solar forecasting will continue to grow in importance for the operations of utilities and IPPs, as well as the speculative decisions of energy traders. To accelerate the maturity of solar forecasting in the market, Amperon believes the industry should coalesce around a single, well-suited measure of predictive error. In our experience, cnMAE is the best option.

About Amperon

Amperon is a technology company at the intersection of energy data and AI. Specializing in data management infrastructure, coherent AI/ML models, and leading-edge predictive analytics, Amperon helps the full spectrum of power and utility companies to improve grid reliability, optimize asset economics, and accelerate decarbonization. Amperon delivers demand and renewable forecasts for energy market in North America, Europe, and Australia.

To learn more about the benefits of Amperon's forecasting service, visit www.amperon.co or contact us here.

AMPERON

sales@amperon.co   amperon.co   @amperon