SAND2017-5645C

Targeted Evaluation of Utility-Scale and Distributed Solar Forecasting

Matthew Lave, Robert J. Broderick, Laurie Burnham
Sandia National Laboratories, Livermore, CA and Albuquerque, NM, 94550 and 87185, USA

***Abstract* — We evaluate the performance of commercial renewable generation and demand forecasts. We define measures of forecast performance that directly relate to the grid operations of the forecast area, and evaluate the forecasts based on these metrics. We also identify opportunities for forecast improvement by examining systematic deviations between forecast and actual quantities by looking for spatial and temporal patterns in forecast errors**.

I. Introduction

Solar and wind generation is variable due to changing weather. As penetrations increase, scheduling and operation of the electric grid becomes more challenging, motivating forecasting. While utility-scale forecasts of wind and solar generation are commercially available, their value has typically been demonstrated on a plant-by-plant basis using traditional metrics such as the root mean squared error (RMSE) which are best suited for stationary statistical processes. Wind and solar timeseries, however, have time-varying statistics. And, forecasting is most valuable when including distributed energy resources (DERs) in aggregations of wind and solar, and when comparing renewable production to load to understand the net impact to the system.

Utility-scale wind and solar plants have the largest capacities on a system-by-system basis. Intra-hour forecasts, critical for making real-time grid decisions, are generally based on statistical techniques. For hour-ahead scheduling, satellite-based methods are often used. Finally, for day-ahead or longer forecasts, numerical weather prediction (NWP) is typically relied upon. Most recent solar forecast research has focused on improving irradiance forecasting. For instance, using the infrared satellite channel in addition to the visible channel to enhance cloud detection [1], adding bias corrections to the NWP forecasts (e.g., [2]), and improving the physics of NWP models have all helped increase irradiance forecast accuracy. Additionally, power conversion for utility-scale plants is typically accurate. Existing models such as the Sandia Array Performance Model [3] can be used in conjunction with known plant specifications (capacity, tilt, module efficiency, cleaning schedule, etc.) to accurately convert irradiance to power. Furthermore, most utility-scale plants have high-quality instrumentation and detailed power output measurements. This allows for robust forecast training using machine learning techniques to further improve power output forecasting [4].

In aggregate, DERs are also a major contributor to electric grid operations: for example, distributed PV made up 32% of solar power capacity in California in 2014 [5]. Some efforts have focused on improving distributed PV forecasts. Hoff, et al. [6] used satellite-derived irradiance and simple irradiance to power models to simulate the power output of 130,000 systems in California from satellite irradiance forecasts. Lonij, et al. [7] created distributed renewable generation power forecasts for the state of Vermont using machine learning by ingesting massive amounts of weather, SCADA, and smart meter data.

In this work, we evaluated both distributed and utility-scale PV forecasts, as well as distributed load forecasts. The forecasts are commercial forecasts that are provided to a system operator. We look at forecast performance dependence on time of day, sky conditions (cloudy/clear) and season. Analyses conditional on these partitions identifies opportunities for forecast improvement.

II. Data

The forecasts evaluated were commercial forecasts that are delivered in near real time to a system operator. The forecast is based on machine learning of both weather trends but also trends in the PV performance and load consumption patterns. The machine learning algorithms are retrained on a regular basis as the historical period of recorded data becomes longer.

Data used for this analysis was:

1. 1-year of measured and forecasted distributed PV, distributed load, and distributed net load for (a) one substation with high PV penetration and (b) an aggregate of many substations covering a large geographic area. All forecasts are 1-24 hours ahead.
2. 1-year of measured and forecasted PV power output for a 2.1MW PV farm. Forecasts are included at 1-24 hour, 25-48 hour, and 49-72 hour ahead intervals.
3. 4-months of measured and forecasted wind power output for 4 wind farms.

All forecasted and measured data is at 1-hour resolution. Additional data is expected to be obtained in the near future and will be analyzed in the final paper.

III. Forecast Performance

A. One High Penetration Substation

The high penetration substation had maximum PV production of 4.7MW and maximum load of 5.7MW (i.e., an 82% PV penetration). Reverse power flow occurred over 10% of the time, with a maximum reverse flow of 3.5MW.

Figure 1 shows measured and forecasted PV power, load, and net load for a sample week at the substation with high PV penetration. On clear days, the PV power forecast is lower than the measured PV power, which may indicate that the forecast did not account for all of the installed PV. The load forecast typically matches the magnitude of the measured load, though does not capture midday spikes in demand. In this case, negative PV forecast errors (under predicting PV production) partially mitigate negative load forecast errors (under predicting load), such that errors in net load are smaller.

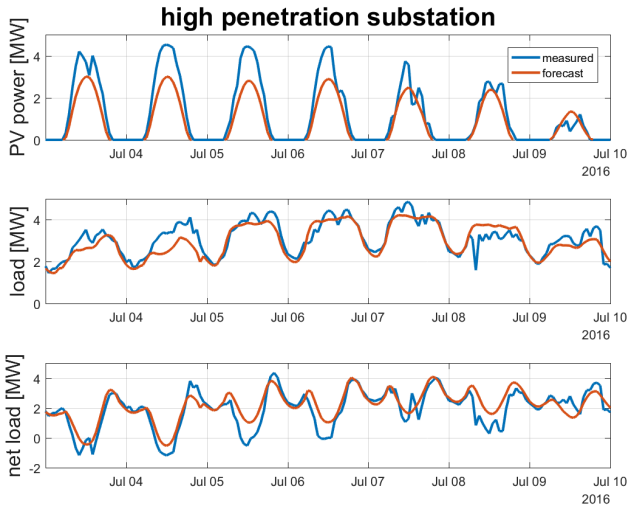

Figure 1: Measured (blue) and forecasted (red): [Top] PV power, [Middle] load, and [Bottom] net load for a substation with high PV penetration

Figure 2 shows the forecast versus measured daily PV production for all days in the August 2015 to July 2016 period of record. Consistent with Figure 1, the forecast under predicts, especially at high measured PV production. The plot is color coded into data before and after May 16, 2016. Both colors of data follow roughly linear trends, but the slope of the line are far from the 1:1 line we would expect in a perfect forecast. These linear trends indicate that when the forecast was high, the measured was also high, and vice-versa. However, it appears that the amount of installed PV capacity was incorrect in the forecast model. The measured PV production exceeds 40MWh, yet the forecasts have maximum values around only 25MWh (after May 16) or only 12MWh (before May 16).

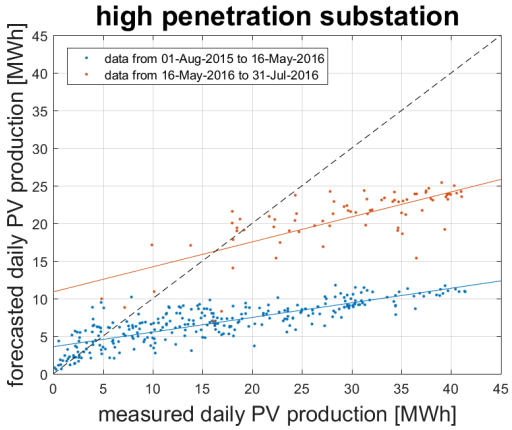

Figure 2: Forecasted versus measured daily PV production on the high penetration substation before May 16 (blue dots) and after May 16 (red dots). The dashed black line is the 1:1 line that a perfect forecast would follow.

The load forecast performed better than the PV forecast, as seen in Figure 3, where the load forecast closely matches the 1:1 with measured load. There is a small amount of under prediction of the highest loads, consistent with the forecast missing the midday spikes seen in Figure 1.

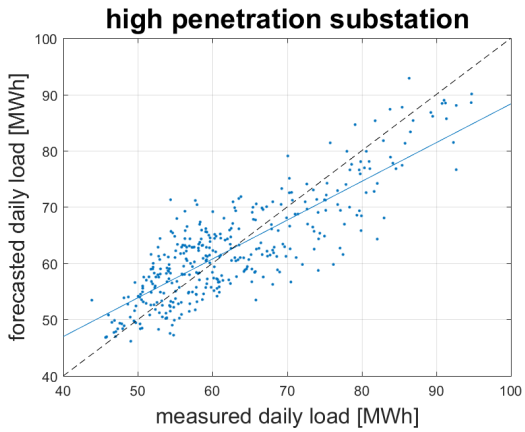

Figure 3: Forecasted versus measured daily load on the high penetration substation. The dashed black 1:1 line is what a perfect forecast would follow.

It is also important to understand how errors in PV and load forecasts interact with one another to affect errors in net load forecasts. Figure 3 plots the hourly comparison of measured to forecasted net load (rather than the daily comparisons seen in Figure 1 and Figure 2). This hourly comparison allows us to examine the forecast's ability to predict reverse flows on the feeder, and Figure 3 is color coded to show when there was measured reverse power flow and if the forecast correctly predicted. The large number of red dots indicate periods when the measured net load was negative but the forecast was positive. There are many more of these red dots than there are magenta dots which signify a correct reverse power flow prediction, showing that the forecast does not predict reverse power flow often enough.

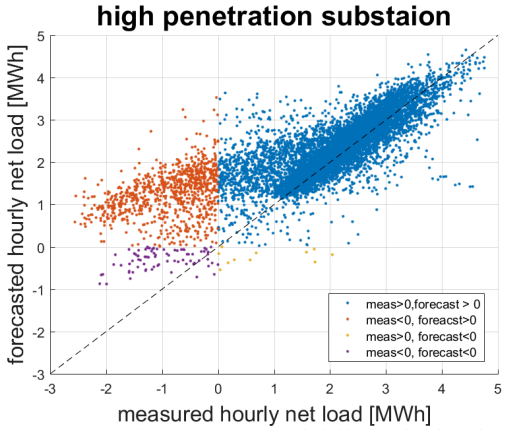

Figure 4: Forecasted versus measured daily load (blue dots) on the high penetration substation (blue dots). The dashed black 1:1 line is what a perfect forecast would follow.

B. *Aggregate of Many Substations*

When many substations were aggregated, the PV penetration was smaller (12.6%) than the single high penetration substation, and reverse flow never occurred. Thus, the (relatively accurate) load forecasts dominated net load forecast accuracy. As PV penetrations increase, though, accuracy of the PV forecast will become more important.

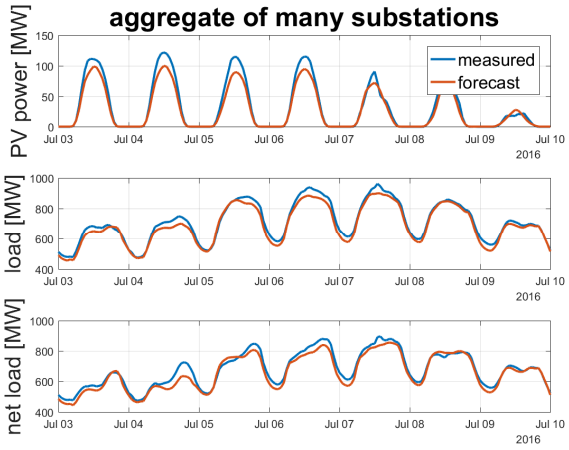

Figure 5: Measured (blue) and forecasted (red): [Top] PV power, [Middle] load, and [Bottom] net load for the aggregate of many substations.

The PV forecast for this aggregate of many substations was better than for the single substation, as seen in Figure 6. There are likely two factors for this accuracy: (a) this aggregate of many substations contains several of the PV farms for which forecasting accuracy is generally very good (section III. C.), and (b), forecasting over large spatial areas is easier since forecast errors tend to average out one another (e.g., the forecast may be low in one region but high in another region). The aggregate PV forecast does still under predict PV production when measured production is high. As for the single substation, this is likely due to not adjusting to the correct PV capacity.

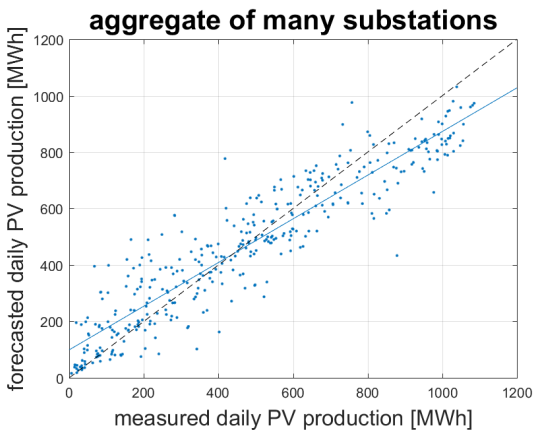

Figure 6: Forecasted versus measured daily PV production for the aggregate of many substations. The dashed black line is the 1:1 line that a perfect forecast would follow.

The load forecast closely matches the measured load, as see in Figure 7; the load forecast for the aggregate appears to be better than for the single substation. Just as for the PV, this is expected since the aggregation of more loads make it easier to predict.

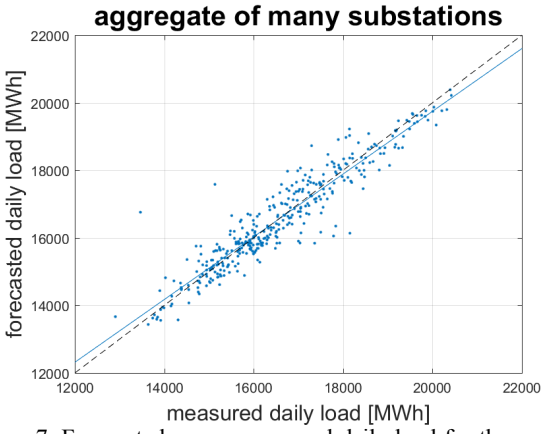

Figure 7: Forecasted versus measured daily load for the aggregate of many substations. The dashed black 1:1 line is what a perfect forecast would follow.

The load is always much greater than the PV generation, such that negative net loads never occur and the net load forecast errors are very similar to the load forecast errors.

C. *PV Farm*

In addition to distributed forecast analysis, we also examined utility-scale PV farm forecasts. Figure 8 shows the measured and forecasted power output for a week at a PV farm. Included in Figure 8 are forecasts at three different time horizons: 1-24 hours ahead, 25-48 hours ahead, and 49-72 hours ahead. These were generated by forecasts published as midnight UTC one, two, or three days ahead.

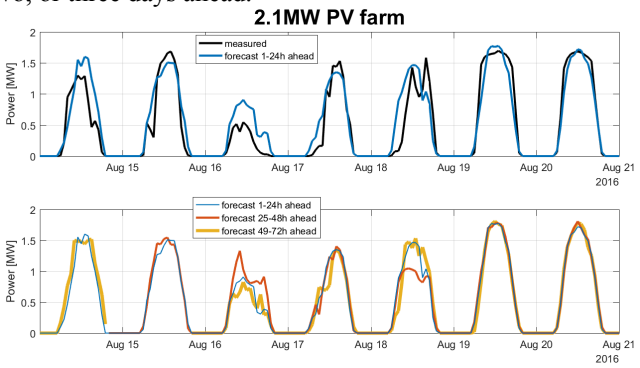

Figure 8: [Top] Measured and most recently forecasted PV farm power output. [Bottom] Comparison of forecasts at different time horizons

The PV farm forecast does good job at forecasting clear versus cloudy days. Clear days August 19th and 20th are correctly forecast as clear. The cloudy day August 16th is correctly forecasted as cloudy, though the forecast over predicts

the measured power. No major difference in accuracy was seen between one, two, or three day ahead forecasts: all three time horizons captured the general trends (clear/cloudy/partly cloudy) of the PV production.

Figure 9 shows the scatter plot of forecasted versus measured daily PV production. This PV farm forecast matches the measured data better than either the single or the aggregate of multiple substations. A main reason for this is likely that training of the forecast is much easier for the PV farm: trends such as temperature dependence, soiling, shading, etc. can be learned from the historical data and are likely to continue into the future. The substations include several tens or hundreds of PV systems aggregated together such that trends of a single system are difficult to detect from the aggregate data. Additionally, the substations have ever-increasing amounts of PV as new residential systems are constantly added while the PV farm capacity is fixed.

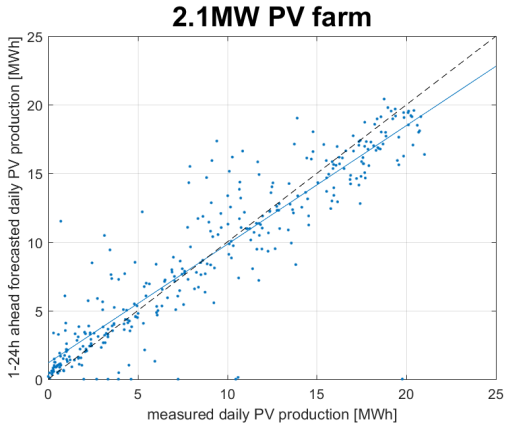

Figure 9: Forecasted versus measured daily PV production for the 2.1MW PV farm. The dashed black line is the 1:1 line that a perfect forecast would follow.

Seasonal trends are shown in Figure 10 for each of the three time horizons. Overall, the forecasts tend to under predict in the winter months and over predict in June and July. These seasonal changes are sharp, ranging from a -6% bias in winter to a +7-11% bias in July. The 1-24 hour ahead forecast does not always have the smallest monthly mean bias error, showing that sometimes forecast updates lead to larger errors.

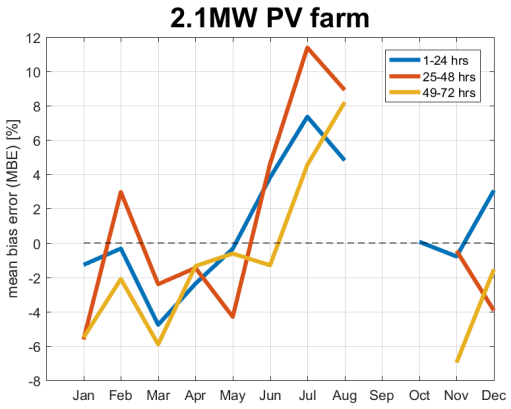

Figure 10: Mean bias error for each forecast horizon (colors), sorted by month of the year. No data was available for September, and only 1-24 hrs ahead forecasts were available in October.

IV. Suggested Improvements

The PV forecasts at times did not match the measured power output during clear-sky conditions. Improvements in forecast performance could be achieved by using clear-sky models which (a) accurately model the PV capacity, and (b) account for the tilt and azimuth of the PV modules.

A. Faster Updates on PV capacity

The PV capacity at substations is a moving target. As seen in Figure 1, the distributed forecast sometimes under-predicted the PV production on clear days, which may be caused by additional PV installations being added to the system. The forecast does update for increased PV penetrations, but the update cycle is slow and depends on accurate numbers on PV installations and interconnection dates.

One simple solution to more quickly account for this changing capacity could be to scale the forecast by the ratio of maximum measured to maximum forecasted power from the previous week:

The impact of using this adjusted forecast on the clear days for the high penetration substation is shown in **Error! Reference source not found.**. Specifically, the mean bias error of the forecast is reduced from -39% to -8% for these clear days. However, care must be taken to only apply this correction to clear days, as the forecast does not have the same (low) bias on cloudy days.

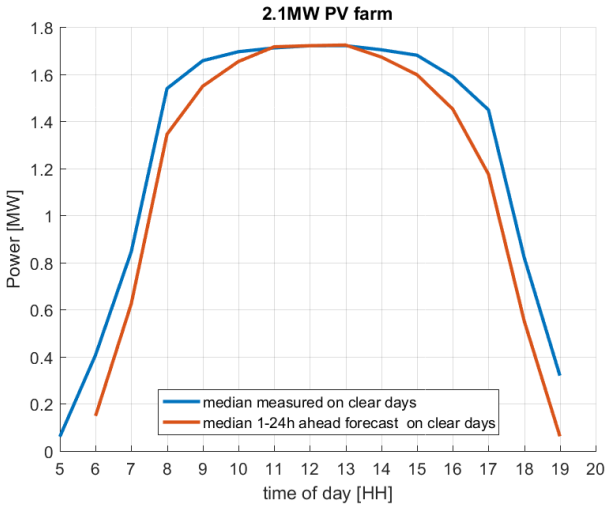

Figure 12: [Left] average measured (blue) and forecasted (red) power output on clear days at 2.1MW PV farm. [Right] Picture from the 2.1MW PV farm, showing that the modules are dual-axis tracking.

B. Directly Account for Tilt and Azimuth Angles

The current PV farm forecast only accounts for a possible tilt angle of PV modules. This leads to inaccuracies in the clear sky model. For the 2.1MW solar farm, the clear-sky profile did not follow the same shape as the measured, as seen in Figure 12. The modules at this solar farm are dual-axis tracking, and the measured power output on clear days has a nearly flat, high power output for most of the day, consistent with dual-axis tracking. The forecast, however, has a "U" shape which more resembles a fixed-tilt system. The underlying clear-sky model used in the forecast could be altered to better resemble the dual-axis tracking.

C. Separate Forecast Training for Clear vs. Other Days

To test performance during clear vs. not clear days, we used a simple clear-sky detection. We compared measured PV output to a clear-sky model [8], and a day had to pass two tests to be considered a clear day:

1. Ratio of measured daily energy to clear-sky daily energy greater than 0.9
   - Clear days must have similar energy output to the clear-sky model
2. Midday correlation (10A-2P) greater than 0.95
   - Clear days must have the same shape as the clear-sky model

An example of this filter applied to a week of PV data from the aggregate of many substations is shown in Figure 13.

Figure 14 shows the results of segregating clear days for the PV forecast of the aggregate of many substations. We see that the forecast under predicts the PV production on clear days, and slightly over predicts the PV production on other days. This result is consistent with Figures 2, 6, and 9, where the forecast often under predicted at high power outputs (e.g., clear days) and over predicted at low power outputs.

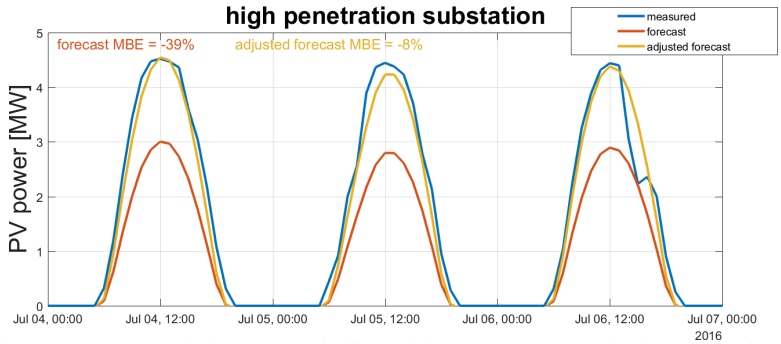

Figure 11: Measured, forecast, and adjusted forecast PV power for three clear days.

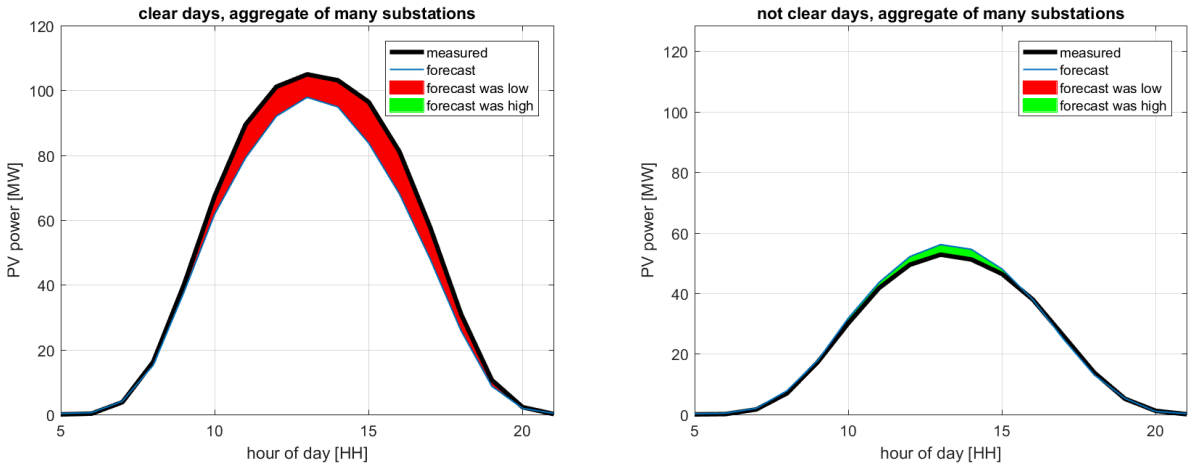

Figure 14: Average forecast performance [Left] on clear days and [Right] on no clear days.

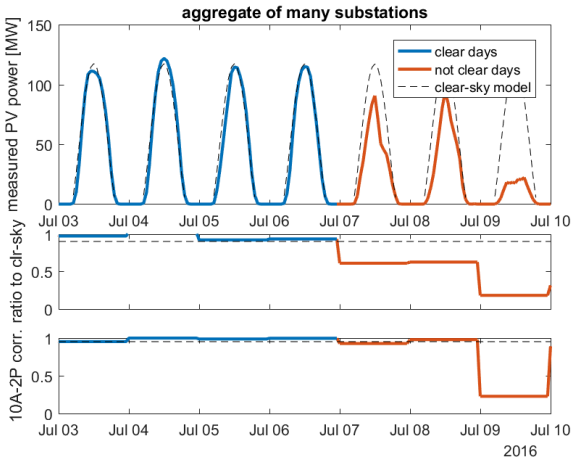

Figure 13: Example of the clear-sky filter, showing days determined to be clear-sky in blue and other days in red. The two bottom plots are the two filters applied to determine if a day was clear.

These errors appear to be caused by forecast training on all days at once. On clear days, the mean of all days, which the forecast is trained on, is less than the actual power output. On not clear days, the opposite is true. Clear days are the special case where we can make a very good prediction using simply a clear-sky model; this simple prediction would eliminate these negative forecast errors on clear days. Thus, we suggest separating training data based on whether the day was clear or cloudy. If a day is forecast to be clear (e.g., using the clear-sky detection mentioned above), its forecast should be based solely on previous clear days. Similarly, if a day is forecast to not be clear, it should be trained only on data from other not clear days.

V. Conclusions and Future Work

Evaluation of the PV forecasts has led to a better understanding of forecast performance and suggested improvements. Load forecast were generally found to be accurate, but PV forecasts were more errant. Three potential improvements to the PV forecasts were suggested: faster updating of PV capacity, directly accounting for azimuth angles, and segmenting training data by clear and not clear days. Further forecast evaluation work is looking at the performance of wind forecasts to see if additional trends in forecast performance exist in the wind forecasts.

Acknowledgment

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

REFERENCES

[1] R. Perez, S. Kivalov, A. Zelenka, J. Schlemmer, and K. Hemker Jr, "Improving the performance of satellite-to-irradiance models using the satellite's infrared sensors," in *Proceedings of the Solar 2010 Conference, Phoenix, AZ, American Solar Energy Soc*, 2010.

[2] P. Mathiesen and J. Kleissl, "Evaluation of numerical weather prediction for intra-day solar forecasting in the continental United States," *Solar Energy,* vol. 85, no. 5, pp. 967-977, 5// 2011.

[3] D. L. King, J. A. Kratochvil, and W. E. Boyson, "Photovoltaic array performance model," Sandia National LaboratoriesSAND2004-3535, 2004.

[4] IBM. *Machine Learning Helps IBM Boost Accuracy of U.S. Department of Energy Solar Forecasts by up to 30 Percent*. Available: http://www-03.ibm.com/press/us/en/pressrelease/47342.wss

[5] U.S. Energy Information Administration. *California has nearly half of the nation's solar electricity generating capacity*. Available: http://www.eia.gov/todayinenergy/detail.php?id=24852

[6] T. E. Hoff *et al. Behind-the-Meter PV Fleet Forecasting*. Available: http://cleanpower.com/wp-content/uploads/BehindMeterPVFleetForecast.pdf

[7] V. P. A. Lonij *et al.*, "A scalable demand and renewable energy forecasting system for distribution grids," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, 2016, pp. 1-5.

[8] P. Ineichen and R. Perez, "A new airmass independent formulation for the Linke turbidity coefficient," (in English), *Solar Energy,* vol. 73, no. 3, pp. 151-157, 2002.