A Methodology to Improve the Predictability of Solar Energy Generation with Confirmatory Evidence from Germany

Kevin F. Forbes, Ph.D.
Senior Energy Systems Researcher
UCD Energy Institute
School of Electrical and Electronic Engineering
University College Dublin
Kevin.Forbes@UCD.ie

ENERDAY 15th Conference on Energy Economics and Technology

"Creative Destruction?

Disruptive innovation in times of Covid and an accelerated energy system transformation

Technische Universität Dresden

9 April 2021

The Organization of this Talk

1) Why is forecasting so important? Evidence from New York and Great Britain
2) The Literature on Solar Energy Forecasting
3) Solar Energy in Germany
4) The Prospects for Improving the Solar Energy Forecasts: The Case of the 50Hertz TSO in Germany
5) A Time-Series Econometric Model of Solar PV Generation
6) Out-of sample evaluation of the model
7) Summary and conclusion

1)Why is Forecasting Important?

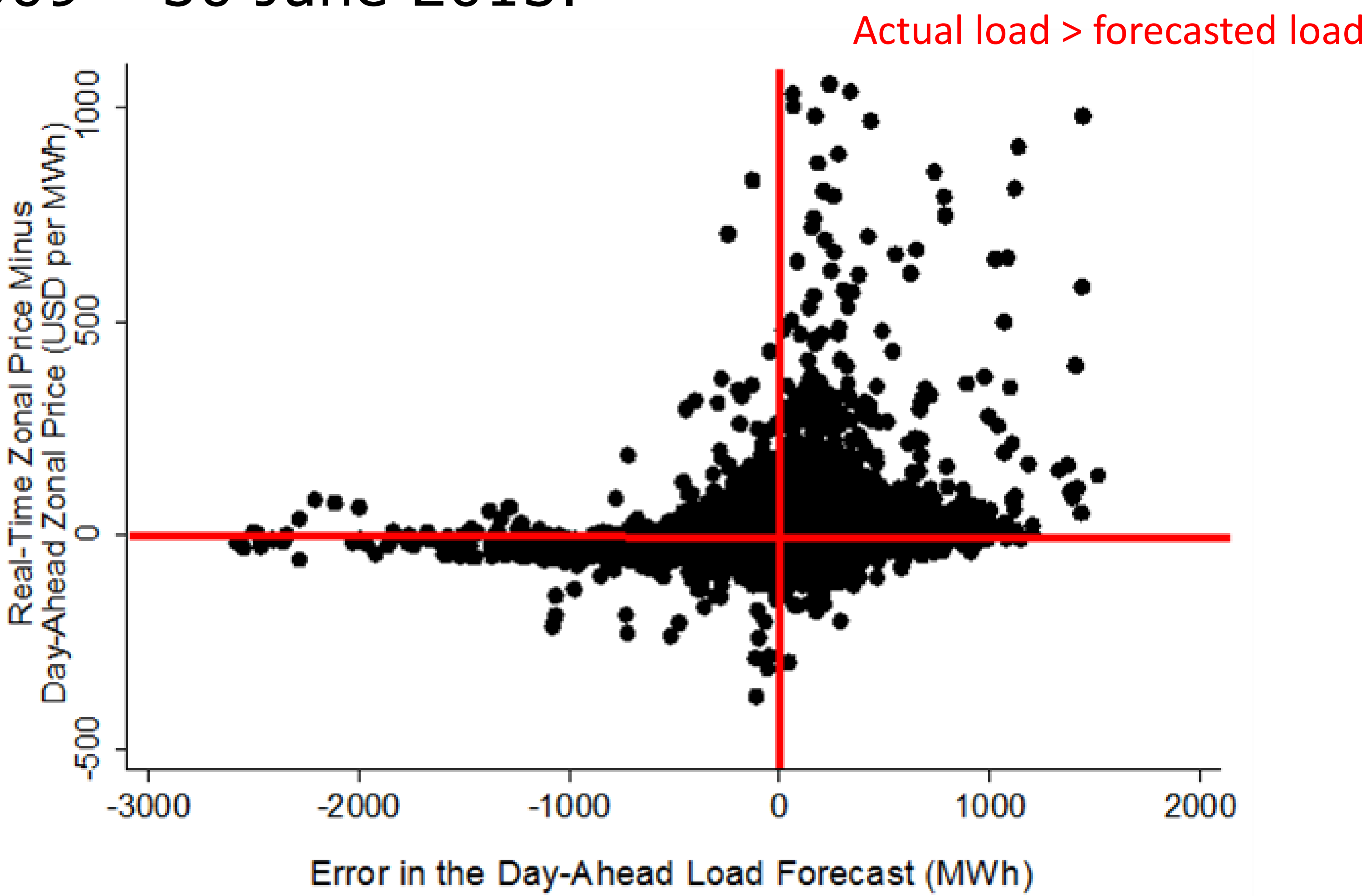
A) Forecast errors have asymmetric consequences. Errors that result in excess supply are easily resolved. Errors that result in excess demand can have significant consequences.

B)The dispatch of balancing power may be unable to resolve the error which may leave the system vulnerable to exogenous shocks

1.A Forecast Errors have Asymmetric Consequences

Errors in the Day-Ahead Load Forecast for New York City and the Differential between the Real-Time and Day-Ahead Prices in New York City, 6 August 2009 – 30 June 2013.

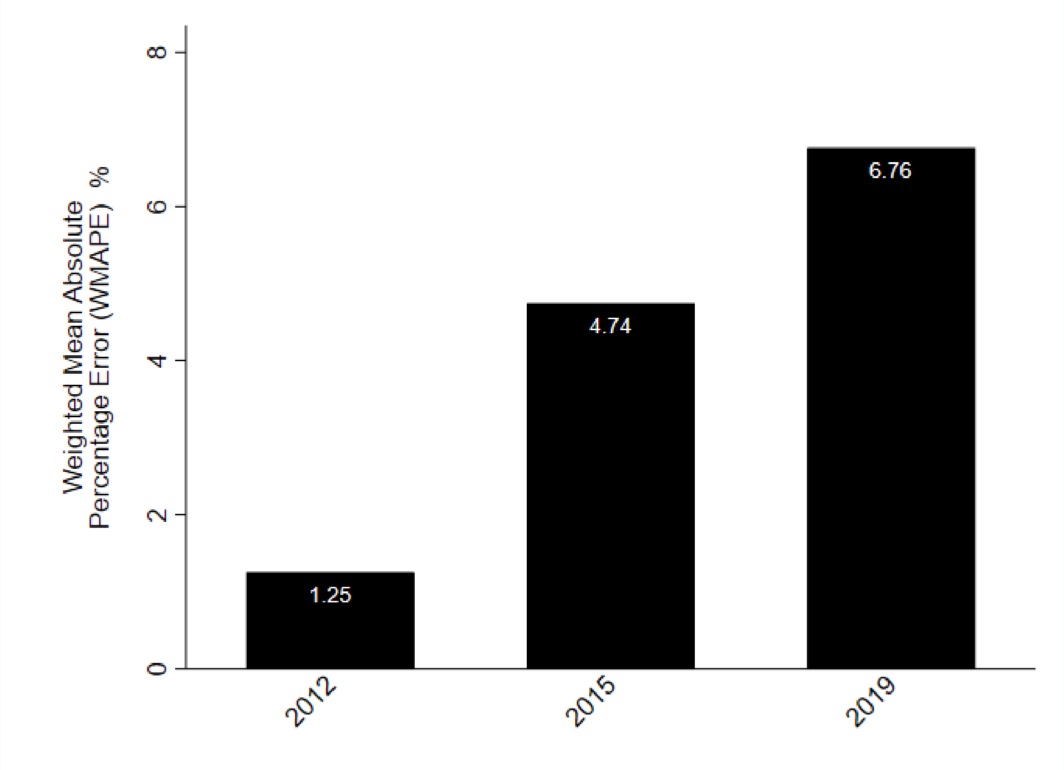
The day-ahead load forecasts in NYC have error of about 2.4 % which seems to suggest that the forecasts are accurate. Yet, the figure on the right indicates that that the forecast errors have economic consequences



Note: Excludes the period of time when operations were affected by Superstorm Sandy in late October 2012

1.B Forecast errors, if not fully resolved by balancing actions, can make the system vulnerable to exogenous shocks

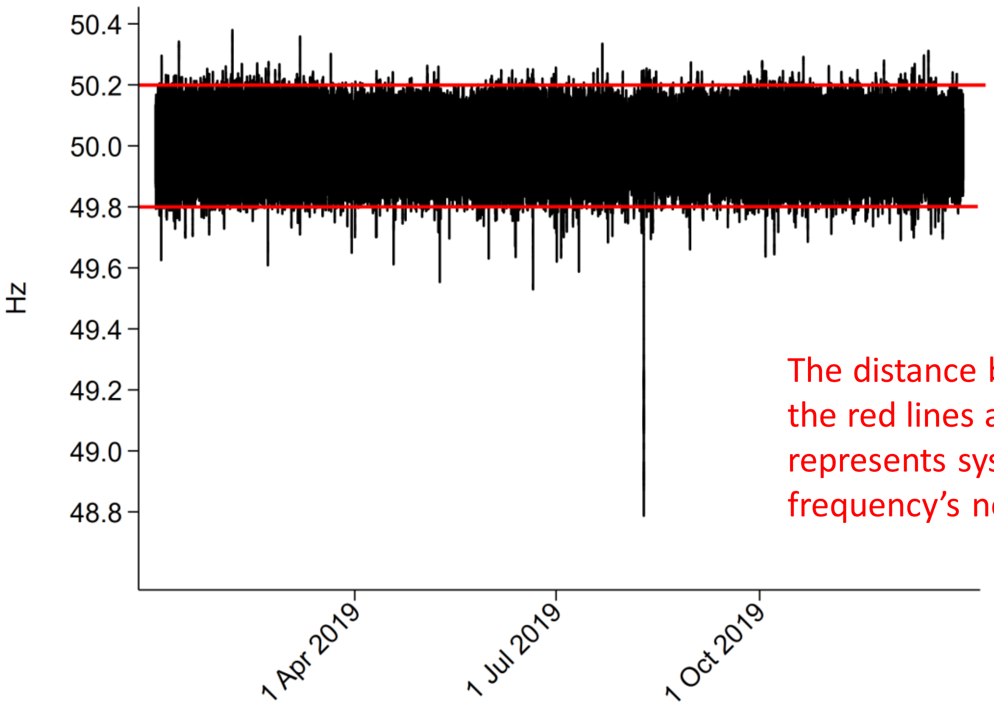
The error in the load forecasts in Great Britain has increased in recent years. It should also be noted that the errors in the solar forecasts are very large.



Source: https://www.entsoe.eu/

Not surprisingly, the decline in forecast accuracy along with other issues has had consequences for system frequency. So much so that a lighting strike in August 2019 led to a large blackout.

System Frequency in Great Britain, 1 Jan – 31 Dec 2019



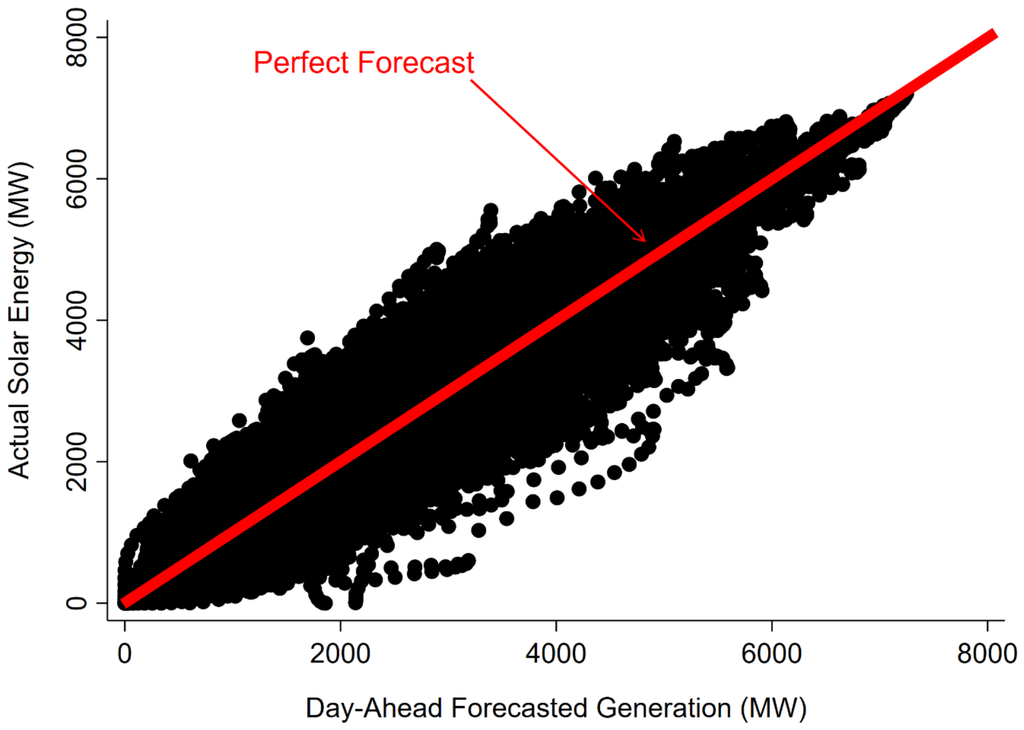
The distance between the red lines above represents system frequency's normal range

Source: https://www.nationalgrideso.com/balancing-services/frequency-response-services/historic-frequency-data

2)The Literature on Solar Energy Forecast Accuracy

At one time, it was believed that solar energy forecasts were moderately accurate. For example, it was once reported that solar energy forecasts had capacity weighted forecast accuracy of less than 10% and thus could be considered moderately accurate.

Unfortunately, this approach to measuring forecast accuracy has proven to be problematic. Indicative of this, the capacity weighted error associated with the scatter diagram on the right is about 2.9%, suggesting that the forecast is accurate even though visual inspection of the figure indicates otherwise.



The forecast accuracy metric employed in this paper

This research makes use of the weighted-mean-absolute-percentage-error (WMAPE), which is defined as follows:

$$WMAPE = \frac{\frac{1}{T}\sum_{t=1}^{T}|Actual_t - Forecast_t|}{\frac{1}{T}\sum_{t=1}^{T} Actual_t} \quad *100\%$$

WMAPE represents the mean absolute error divided by the mean of the actual outcome (Kolassa and Schütz, 2007). In contrast to capacity weighted forecast accuracy metrics, it is easily understood. It has the characteristic of being scale-free, and thus, it facilitates comparison of forecast accuracy among forecast variables of interest such as solar energy, wind energy, and load.

Using this metric, the error in the forecast on the previous slide is 13.2% over the daylight periods.

It is now recognized that solar energy forecasting is quite challenging. Indicative of this, the United States Department of Energy in 2020 issued a request for information on how to improve matters - see the webpage on the right.

U.S. DEPARTMENT OF ENERGY | *Office of* ENERGY EFFICIENCY & RENEWABLE ENERGY

Feedback Request: Solar Variability Prediction

DOE Office of Energy Efficiency and Renewable Energy sent this bulletin at 01/29/2020 01:00 PM EST

Having trouble viewing this email? View it as a Web page.

ENERGY.GOV

Office of
ENERGY EFFICIENCY & RENEWABLE ENERGY

Solar Energy Technologies Office

As solar becomes more integrated into electric transmission and distribution systems, the Office of Energy Efficiency and Renewable Energy's Solar Energy Technologies Office has issued a request for information (RFI) to inform its strategic planning efforts. This RFI seeks insights about predicting how much solar irradiance, or energy from the sun, reaches the earth's surface and the amount of power output from solar generation plants using photovoltaic (PV) or concentrating solar-thermal power technologies.

These prediction capabilities will help grid operators as they consider the impacts of solar power variability on grid planning and operations, as well as owners and operators of utility-scale plants and aggregators of distributed PV systems.

Questions in the RFI fall under the following broad categories:

- Solar power prediction
- Solar irradiance prediction
- Model evaluation and validation

Please email your responses as Microsoft Word attachments to SETO.RFI.SI@ee.doe.gov by **February 29, 2020, at 12:00 p.m. ET.** Download the RFI for more details.

Elia, one of leading system operators in Europe, is clearly cognitive of the forecasting challenge:

A Statement by Elia on forecasting:

Generation forecast is becoming more and more challenging

Conventional assets like nuclear, gas or coal power plants are being replaced by intermittent renewable assets like onshore and offshore windmills or solar panels. Forecasts of conventional generation are based on financial calculations as well as a number of parameters. In contrast, renewable assets generate energy when the wind blows or the sun shines. The owners of such assets do not know for certain how much the sun will shine the next day or, even worse, 300 days ahead. **Weather forecasting is not an easy task and has a huge impact on generation** - just think about ice on a windmill's blades, snow on PV panels, or storms at offshore wind farms.

https://www.agorize.com/en/challenges/elia-start-up-challenge-edition-2/pages/learn-more-about-forecasting?lang=en

The Academic Literature

- There has been an exponential increase in the number of peer-reviewed publications on the topic of solar energy forecasting in recent years.

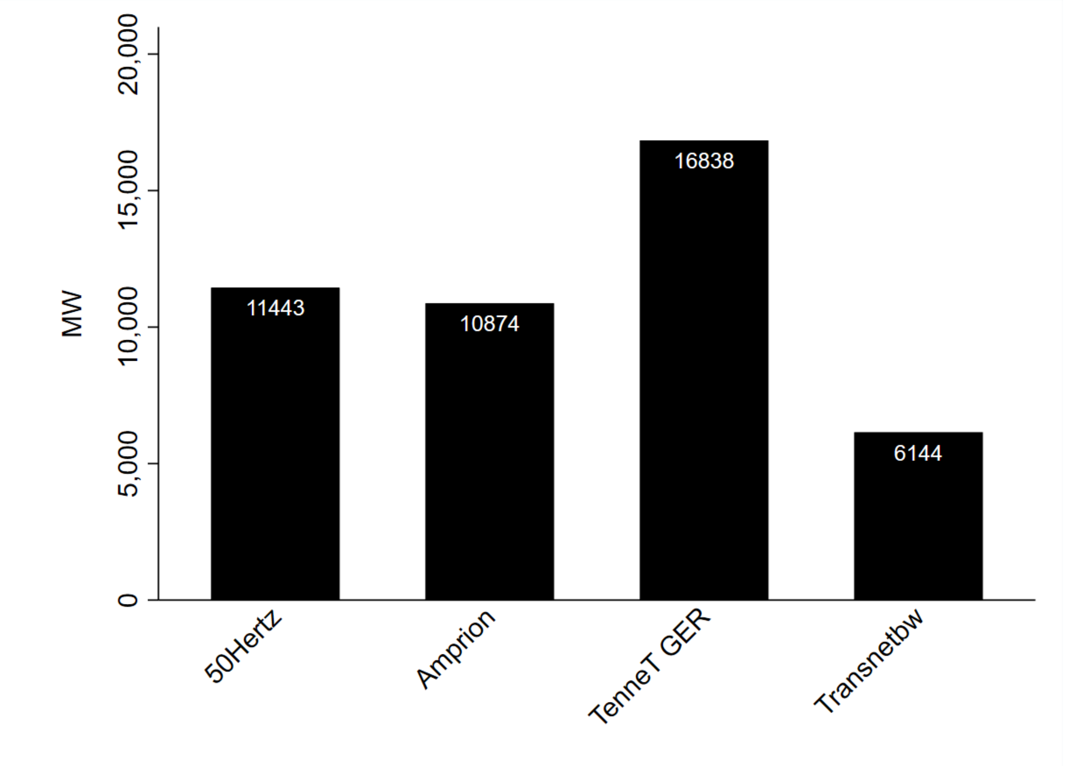
- Unfortunately, there is no consensus on the best approach.

- The most informative article was one that found a correlation between air density and solar radiation (https://link.springer.com/article/10.3103/S1068373918010077)

3) Solar Energy in Germany, the Highest Ranked Country in Europe in terms of Solar Energy Capacity.

The Electricity Control Areas in Germany

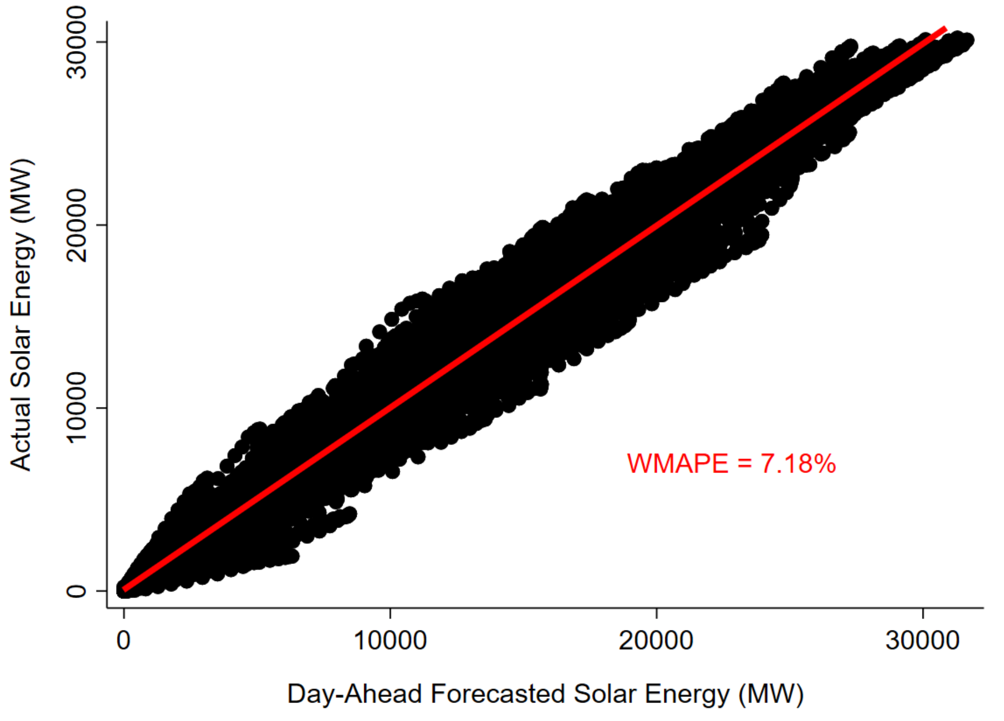
2019 Solar Energy Capacity in each of the Control Areas



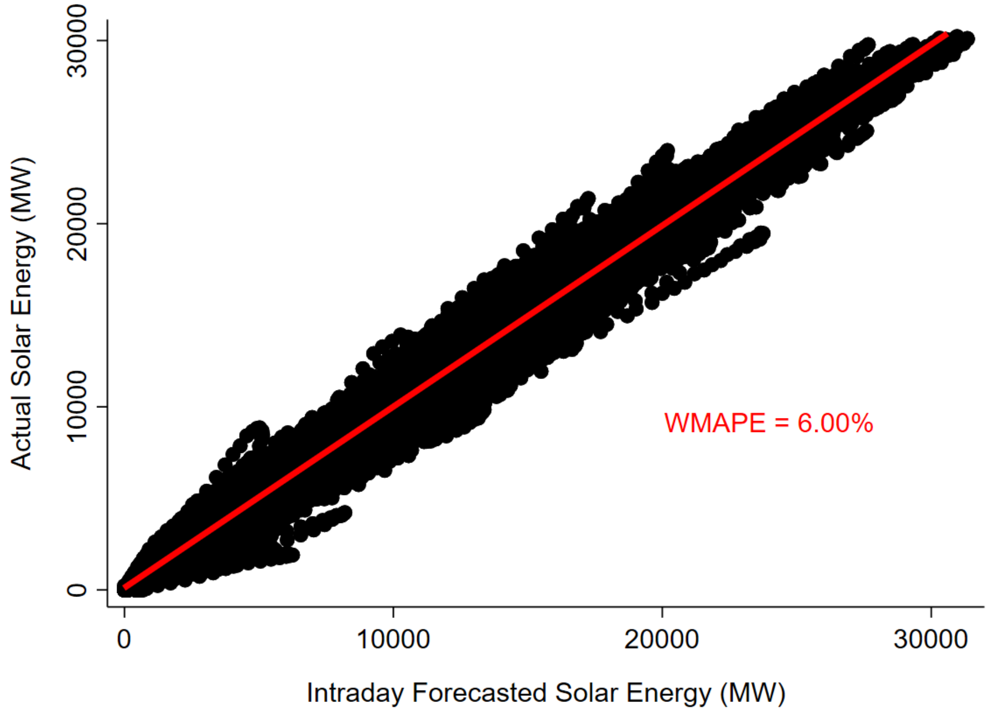
Source: https://www.entsoe.eu/

For All Four TSOs Combined….

Day – Ahead Forecasted and Actual Solar Energy, 1 – Jan 2018 – 31 December 2019



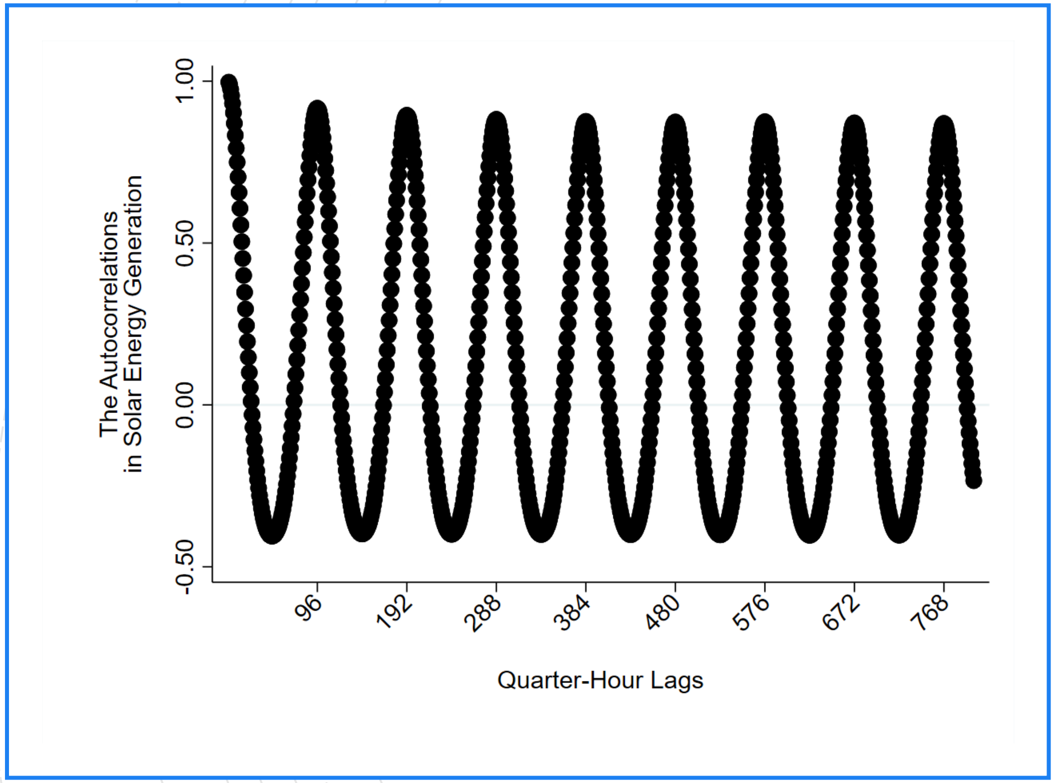
Intraday Forecasted and Actual Solar Energy, 1 – Jan 2018 – 31 December 2019



Sources: https://www.entsoe.eu/
https://www.50hertz.com/en/Transparency/GridData
https://www.amprion.net/Grid-Data/Photovoltaic-Infeed/
https://www.tennet.eu/electricity-market/transparency-pages/transparency-germany/network-figures/actual-and-forecast-solar-energy-feed-in/
https://www.transnetbw.com/en/transparency/market-data/key-figures

Observation: These solar forecasts are likely to be among the most accurate in Europe.

The Autoregressive Nature of Solar Energy Generation in Germany, 1 Jan 2018- 31 Dec 2019



- Considering that each day in the German TSOs has 96 operating periods, the figure on the left makes it fairly clear that solar energy generation in period t is highly correlated with the generation levels in t-96, t-192, t-288, t-384, t-480 etc.
- This characteristic (not fully recognized by previous literature) is quite useful in predicting solar energy given that the past values of solar energy generation are known prior to period t.

4) The Prospects for Improving the Solar Energy Forecasts: The Case of 50Hertz

- Significant improvements in the forecasts are possible because there is evidence that the solar energy forecasts do not fully reflect the information contained in the weather forecasts.

- Significant improvements in very short run forecasts are possible by exploiting the autoregressive nature of solar energy generation.

- In short, accuracy can be improved.

5) An ARMAX (autoregressive moving average with exogenous inputs) Model of Solar PV Generation

- Following Box and Cox (1964), the dependent variable, the level of solar energy generation in 50Hertz, is transformed.

- The explanatory variables include the level of forecasted generation reported by 50Hertz and various measures of forecasted/simulated meteorological conditions. Linearity is not presumed. Some interactions among the regressors are modeled (e.g. air density and forecasted PV generation)

- Seasonality is modeled using a series of 72 binary variables.

- The model makes use of 44 time-series variables that control for the autoregressive nature of the data.

Estimation and Results

The model was estimated over the period 1 Jan 2014 through 31 December 2017. There are 70,189 daylight periods in the sample.

The statistically significant variables include:

- Day-Ahead Forecasted PV Generation
- Forecasted Air Density (both directly as well as in conjunction with other variables)
- Forecasted Temperature
- Forecasted Direct Radiation
- Forecasted cloud cover at moderate height
- Forecasted Dewpoint
- Forecasted Air Pressure
- 36 of the 72 variables representing the "season" of the year.

Estimation and Results (Continued)

- 13 of the 22 AR terms are statistically significant.
- 8 of the 22 MA terms are statistically significant.
- 69 of the 78 variables (e.g., $\sqrt{AirDensity}$) that model the conditional variance are statistically significant.
- The random error term is not Gaussian. Instead, the data indicates it follows a Student t-distribution with about five degrees of freedom. This indicates that the level of kurtosis in the error term is substantially higher than the Gaussian level of three.

Estimation and Results (Continued)

- The full model has an explanatory power that is equivalent to an R-Square of 0.9984.

- The explanatory power is equivalent of an R-Square of 0.9451 if the ARMA terms are disregarded.

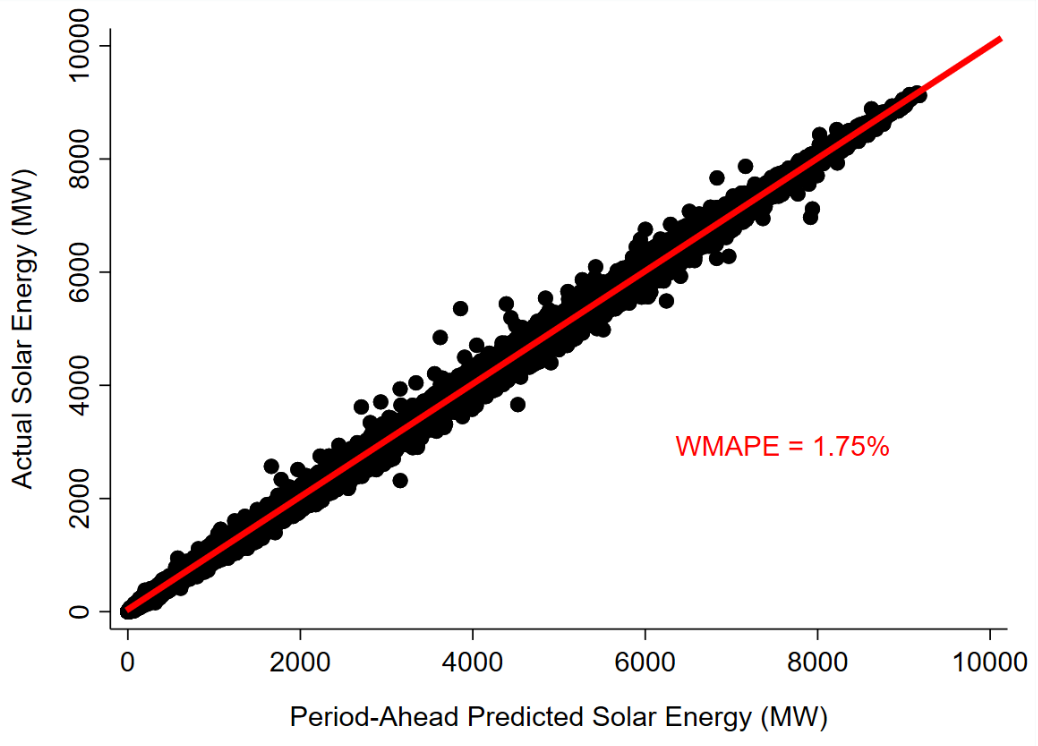
- These R-Square measures are encouraging but it is noted that the true adequacy of a model can only be determined by considering how well it performs on data that were not used in its estimation.

6) Out-of Sample Evaluation of the Model

- The model was evaluated between sunrise and sunset over the period 1 Jan 2018 – 30 August 2020.

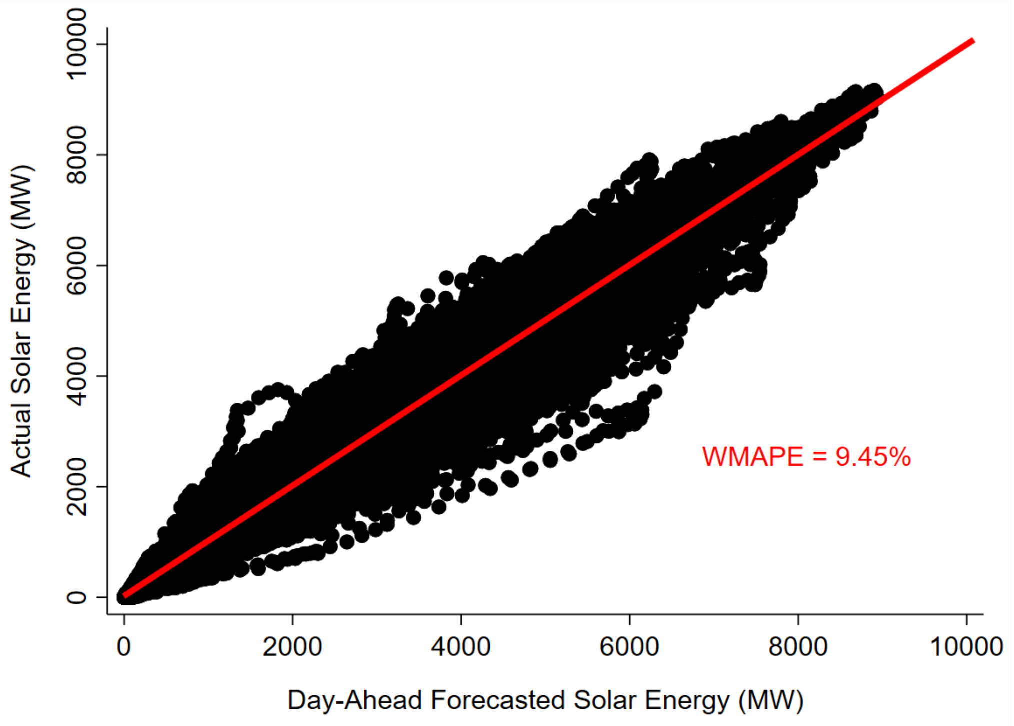
- The predictions for period t based on the information known at the end of period t-1) have a Predictive R-Square of 0.999.

- The associated WMAPE equals 1.75 %. This is an interesting result but unlikely to be of significant value to a system operator given that the amount of advance notice is small.

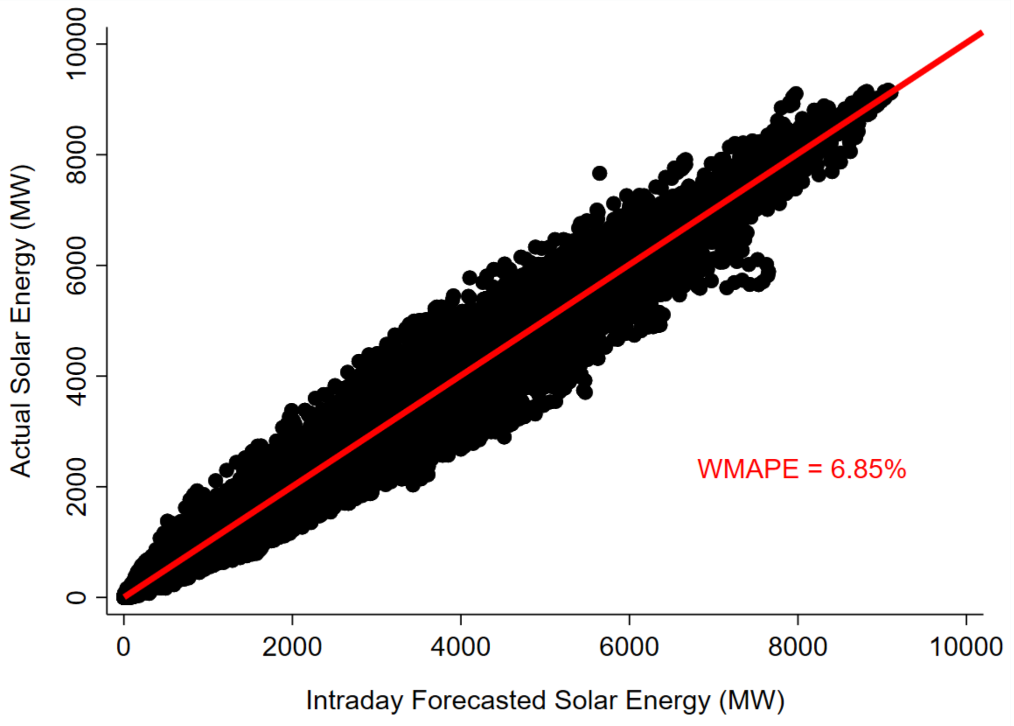


The Actual and the Forecasted Levels of Solar Energy in 50Hertz, 1 Jan 2018 – 30 August 2020

Actual Solar Energy and the Day-Ahead Forecasted Level of Solar Energy

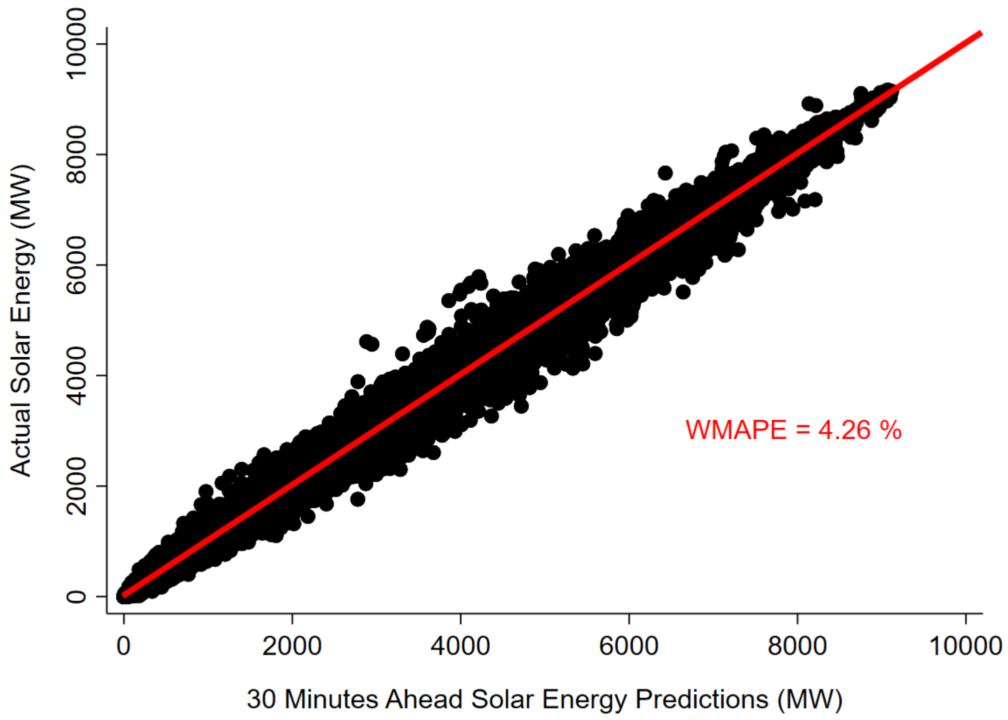


Actual Solar Energy and the Intraday Forecasted Level of Solar Energy

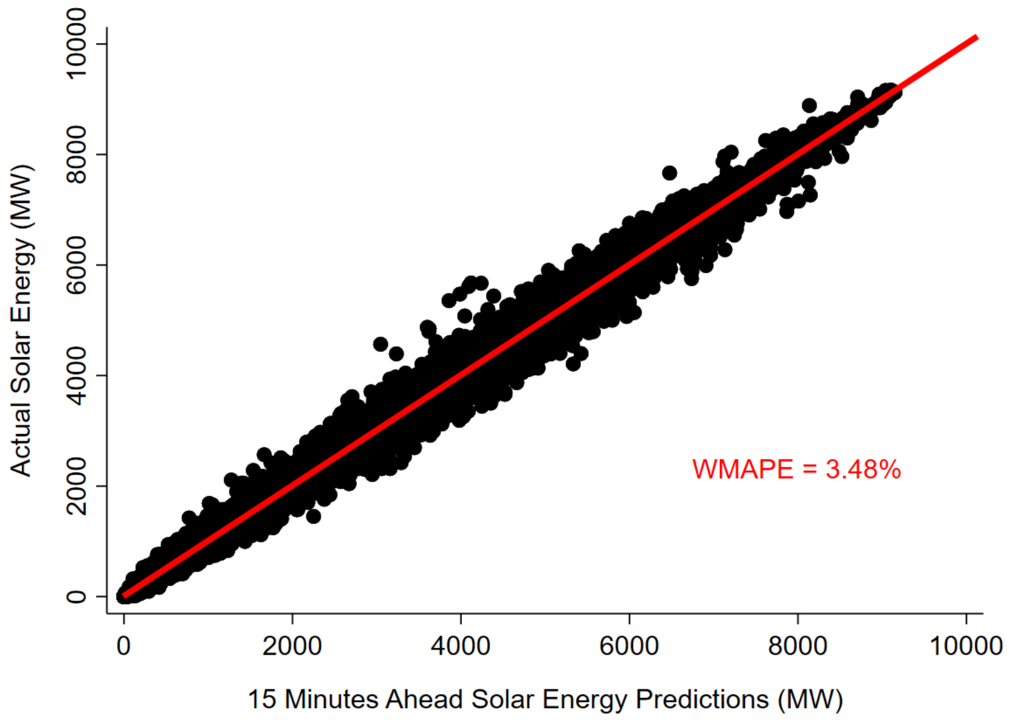


Actual Solar Energy in 50Hertz and Out-of-Sample ARMAX Solar Energy Predictions, 1 Jan 2018 – 30 August 2020

30 Minutes Ahead Predictions



15 Minutes Ahead Predictions



7) Summary and Conclusions

- **Evidence has been presented that the solar energy forecast errors have two systematic components.**
  - **The first portion of the systematic component is that the existing solar forecasts do not fully capture expected metrological and seasonal conditions.**
  - **The second portion of the systematic component is the existing solar forecasts do not capture the autoregressive nature of the solar energy**
- **Based on these findings, ARMAX modelling can yield very short-run solar energy forecasts that are significantly more accurate.**
- **The analysis presented here made use of the existing day-ahead forecasts. These forecasts are an invaluable driver in the methodology presented here. The next step in the research will make use of the intraday forecasts which are more accurate but also not fully reflect forecasted meteorological conditions. It is expected that should improve the accuracy of the ARMAX predictions.**