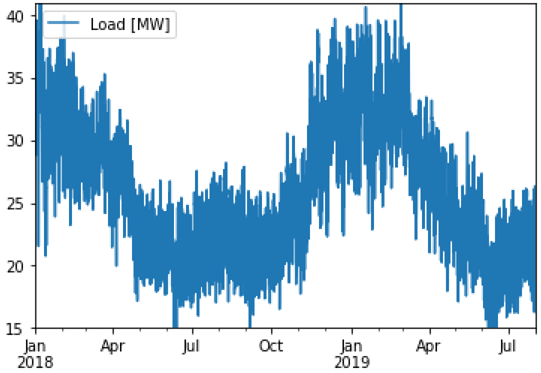
Probabilistic Net-Load Forecasting Using Machine Learning

Daniel Kirk-Davidoff, 5/24/2021

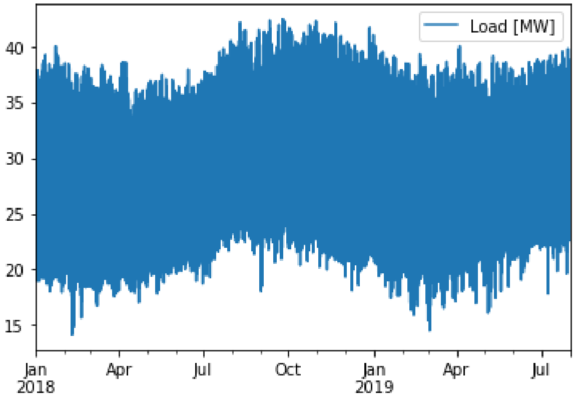
Acknowledgements: Some of the research was supported by the Department of Energy Solar Forecast Improvement Project II, in a project led by EPRI, PI Aidan Tuohy.

Load and Net-Load

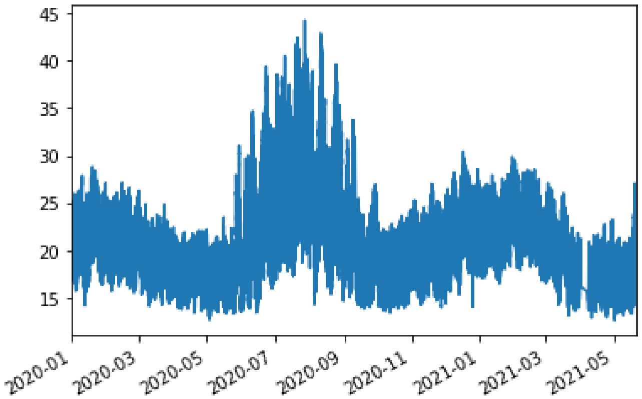
Grid A



Grid B



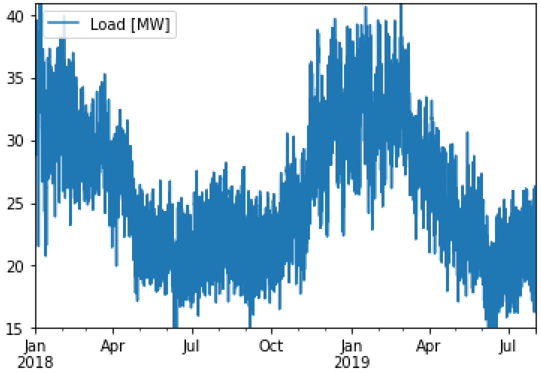
Grid C



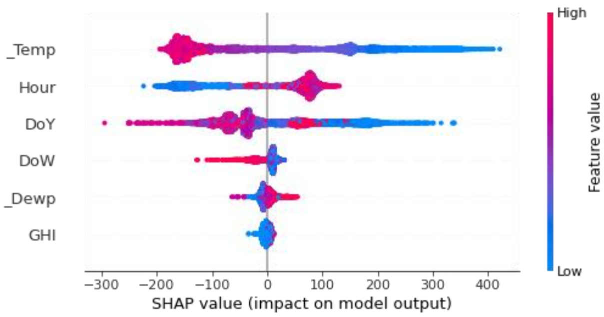
Here are three timeseries of electrical demand from three different electrical markets with different climates and infrastructures. We will examine them with a commonly used machine learning algorithm, XGBoost, and Shapely diagnostics to see what we can learn about their characteristics

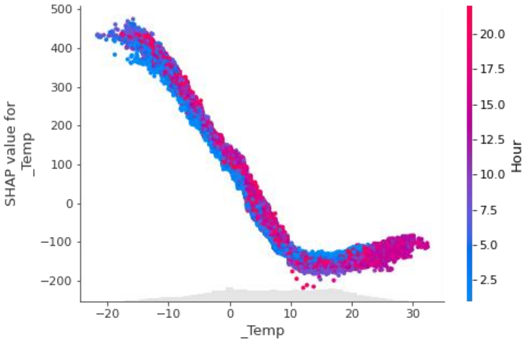
Load and Net-Load

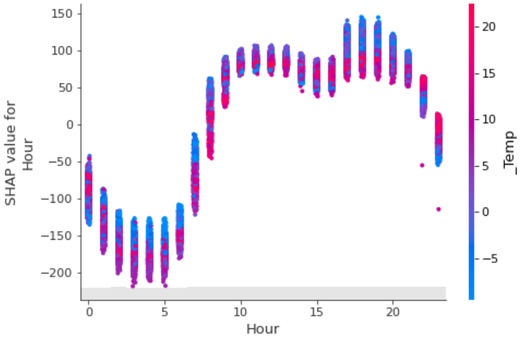
Grid A

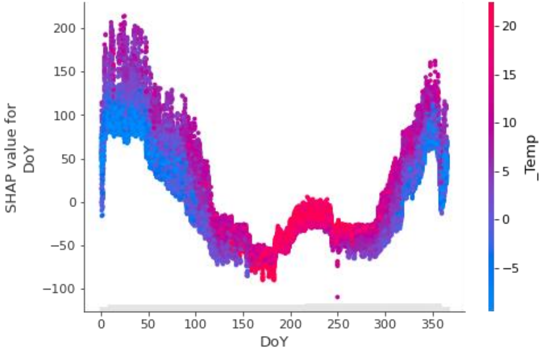


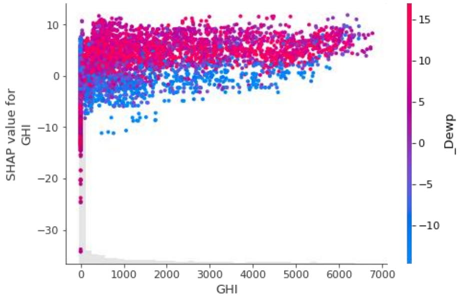
In this grid, the machine learning code determines that Temperature has the largest role in explaining the variations in electrical demand, followed by Hour of the Day, Day of the Year, Day of the Week and Wind Speed.

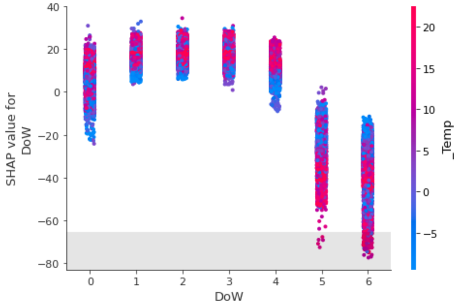






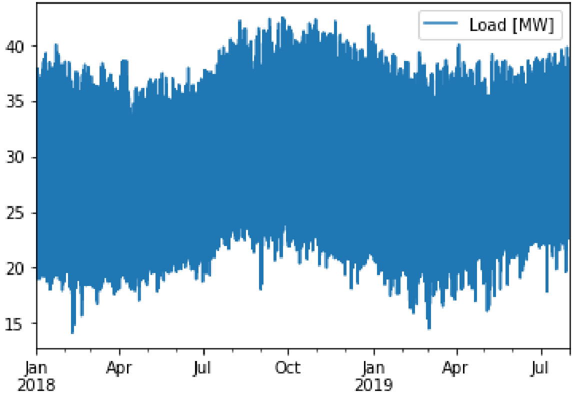


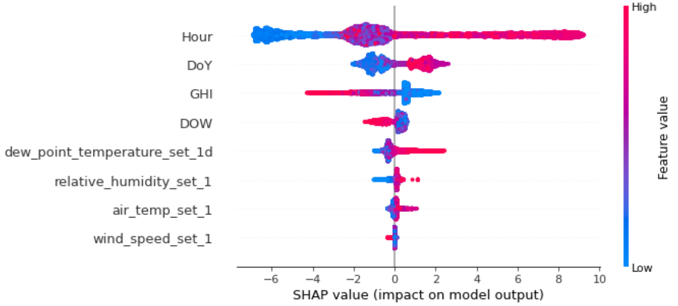


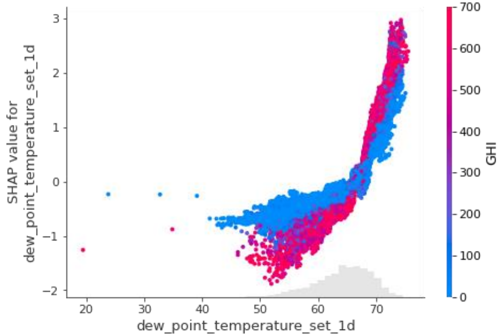


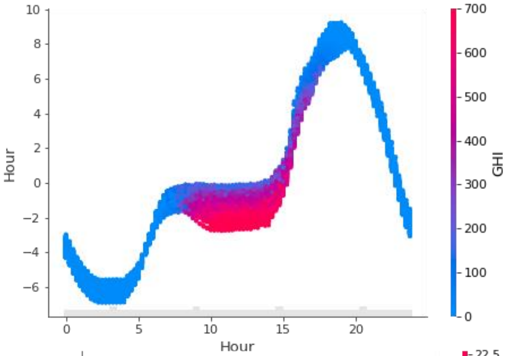
Load and Net-Load

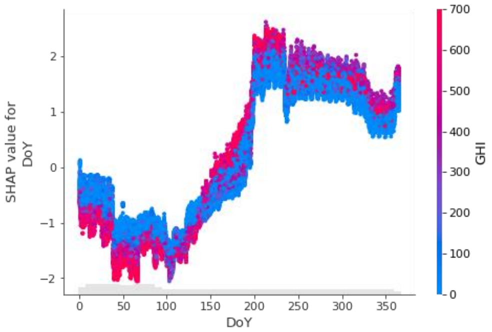
Grid B

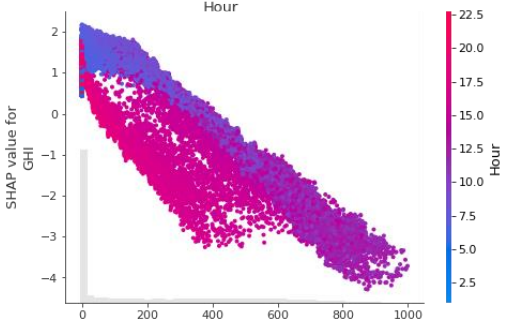


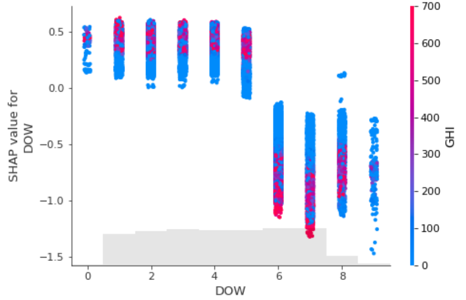








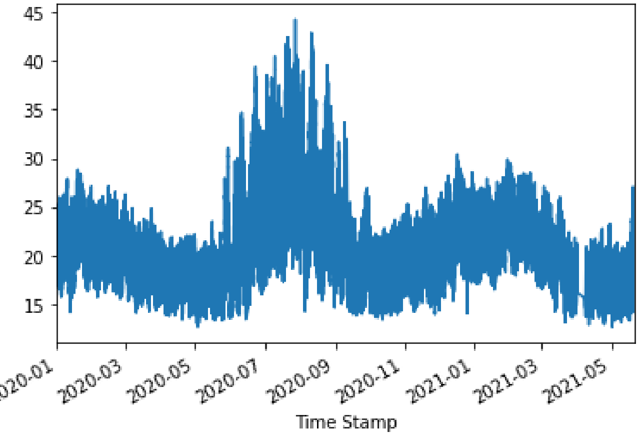




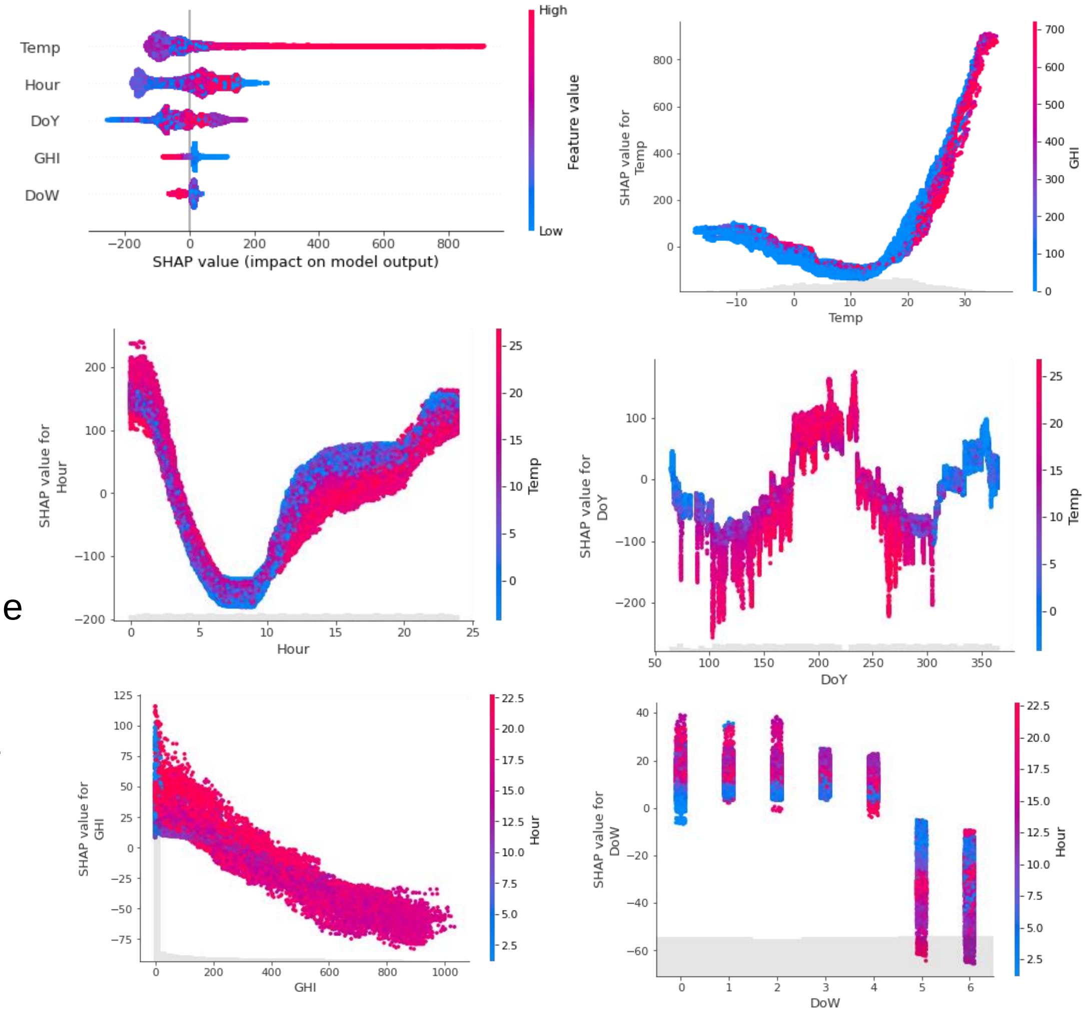
In the warm climate of this grid, the machine learning code determines that calendar factors (hour and Julian day) are most important, followed by global horizontal irradiance and dew point temperature: the algorithm has diagnosed the role of solar generation in the load noticed by the grid (the *net load*, demand, less local generation).

Load and Net-Load

Grid C



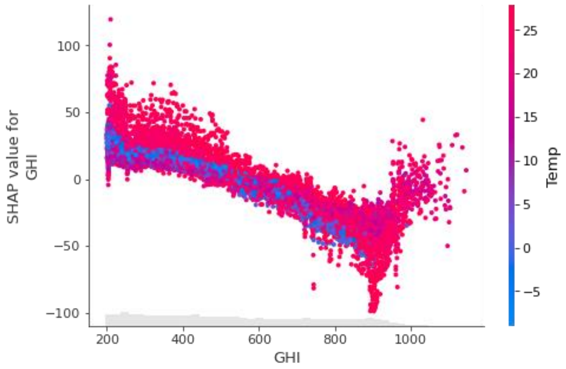
In our third grid, the machine learning code determines that Temperature has a dominant role in explaining the variations in electrical demand (despite only modest levels of heating load), followed by calendar factors, but with GHI playing a smaller but still readily detectable load.

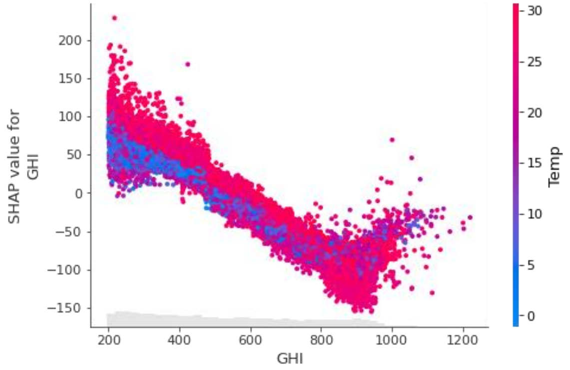


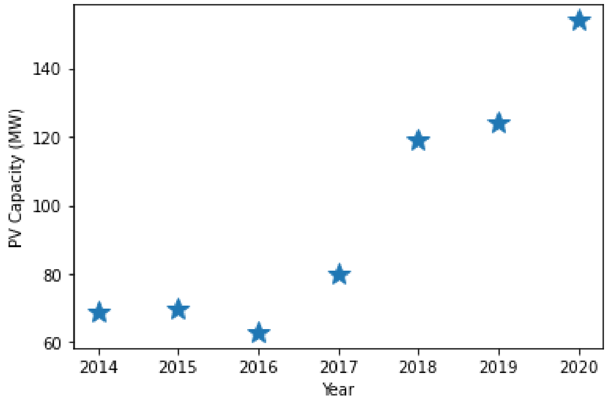
Changing PV Capacity in Grid C

2014

2020



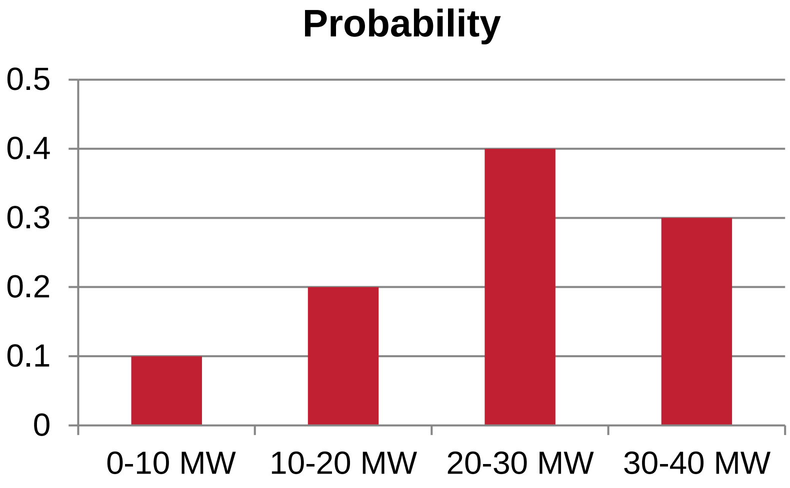


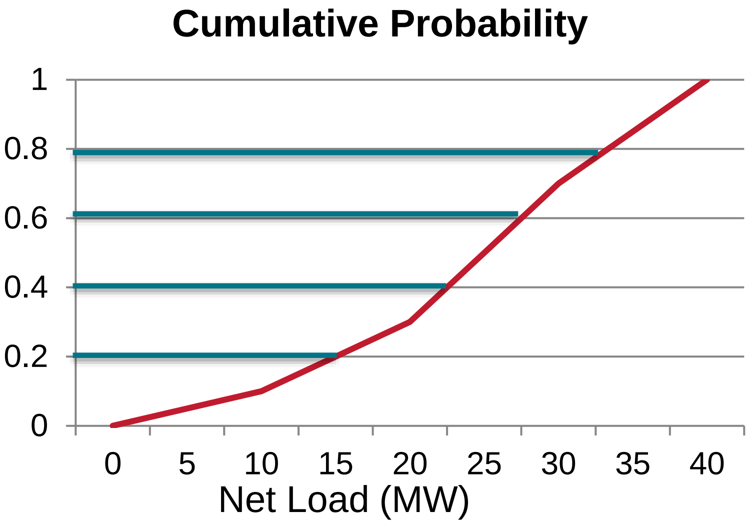


By perturbing the GHI in the trained XGBoost model, we can estimate the changing PV Capacity in the grid. These numbers are in reasonable agreement with official BTM capacity numbers.

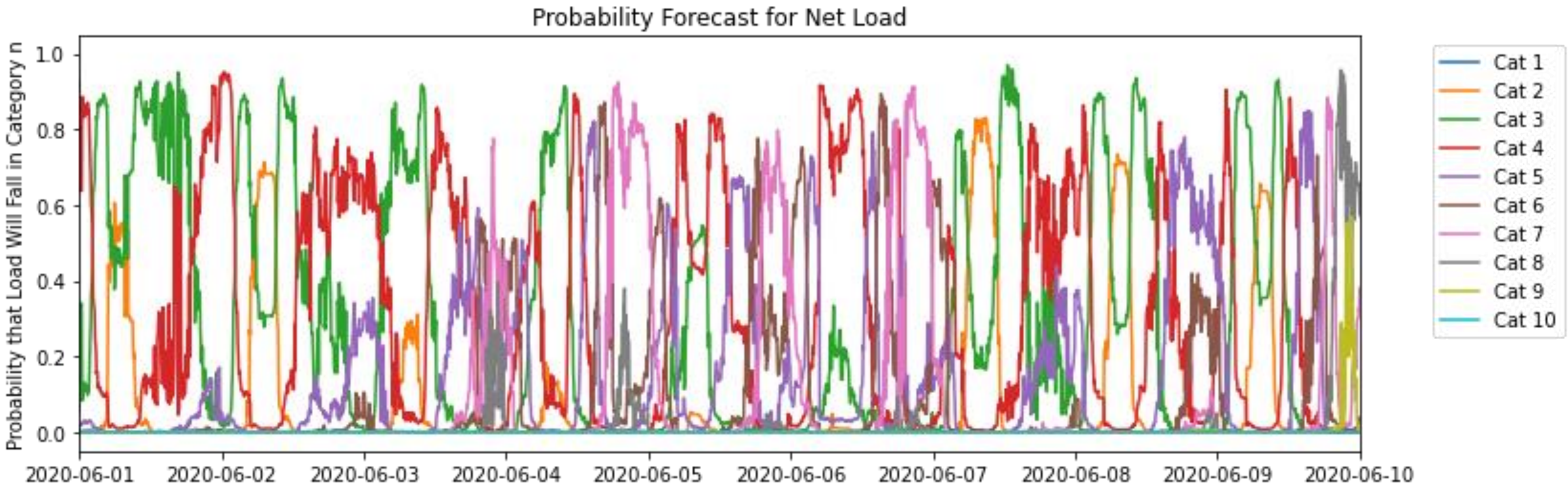
Probabilistic Forecasting Using Machine Learning

A lot of the heritage of machine learning techniques involves categorical prediction (is the image more likely of a cat or a dog?)   This means that many of the popular techniques are well-suited to probabilistic forecasts. In the case of XGBoost, we have only to switch from a regression

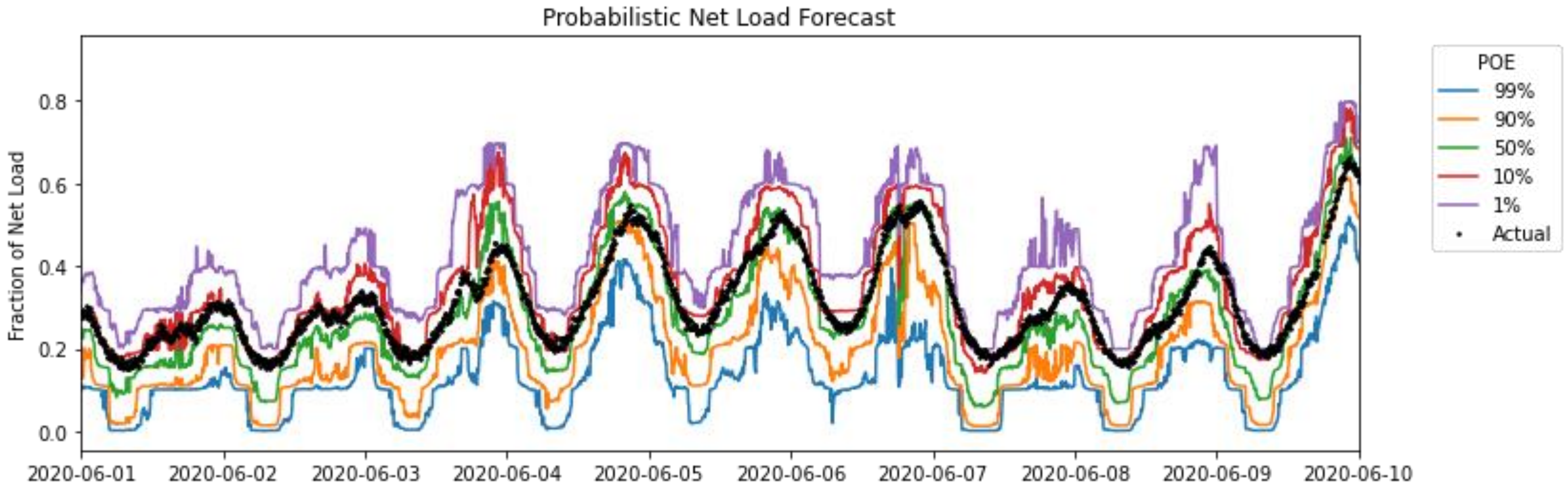




Probabilistic Forecasting Using Machine Learning

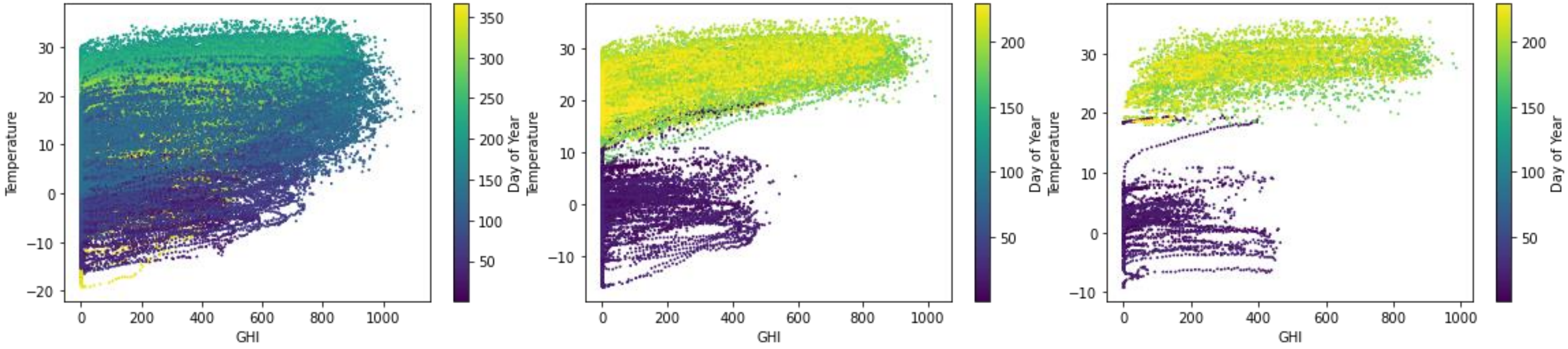


Probabilistic Forecasting Using Machine Learning



Error Covariance and Net-Load Forecasting

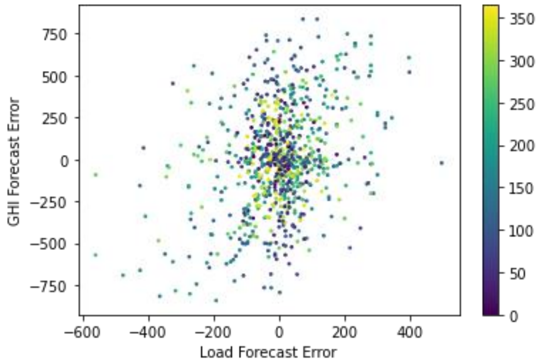
Mid-day data only



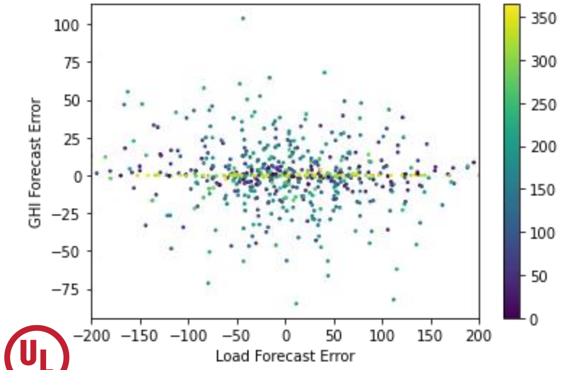
On a day-to-day basis, the correlation of GHI and temperature mostly vanishes – warm days can be sunny or cloudy, cold days too.

Error Covariance and Net-Load Forecasting

Grid C



Grid B



Forecast errors in GHI and the Temperature-dependent load component are slightly positively correlated in Grid C which will lead to increasing error cancelation as solar generation grows, but completely uncorrelated in Grid B.

Conclusions and Ongoing Work

- Our imperfect ability to forecast weather causes uncertainty in both load and renewable generation.
- As renewable generation increases as a fraction of demand, the covariance of solar radiation, wind and weather-dependent electrical demand becomes ever more important in determining the load that grid operators must balance with controllable sources of power such as hydropower and batteries.
- Error covariance will vary by location and may vary by season.
- Capturing this covariance correctly, to have as accurate as possible a forecast of forecast error requires the use of a consistent forecast system for all components of net load.