Solar PV, Wind Generation, and Load Forecasting Dataset for ERCOT 2018: Performance-Based Energy Resource Feedback, Optimization, and Risk Management (P.E.R.F.O.R.M.)

Richard Bryce, Grant Buster, Kate Doubleday, Cong Feng, Ross Ring-Jarvi, Michael Rossol, Flora Zhang, and Bri-Mathias Hodge

National Renewable Energy Laboratory

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-5D00-79498
May 2023

Solar PV, Wind Generation, and Load Forecasting Dataset for ERCOT 2018: Performance-Based Energy Resource Feedback, Optimization, and Risk Management (P.E.R.F.O.R.M.)

Richard Bryce, Grant Buster, Kate Doubleday, Cong Feng, Ross Ring-Jarvi, Michael Rossol, Flora Zhang, and Bri-Mathias Hodge

National Renewable Energy Laboratory

Suggested Citation
Bryce, Richard, Grant Buster, Kate Doubleday, Cong Feng, Ross Ring-Jarvi, Michael Rossol, Flora Zhang, and Bri-Mathias Hodge. 2023. *ERCOT System-wide Renewable Generation and Load Forecasting: Performance-Based Energy Resource Feedback, Optimization, and Risk Management (P.E.R.F.O.R.M.)*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-5D00-79498. https://www.nrel.gov/docs/fy23osti/79498.pdf.

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-5D00-79498
May 2023

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the U.S. Department of Energy Advanced Research Projects Agency–Energy. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government.

Executive Summary

This report describes the Advanced Research Projects Agency-Energy Performance-Based Energy Resource Feedback, Optimization, and Risk Management (PERFORM) Electric Reliability Council of Texas (ERCOT) dataset consisting of load, solar, and wind deterministic and probabilistic forecasts at three timescales. This dataset consists of 1 year of time-coincident load, wind, and solar actuals and probabilistic forecasts for a region similar to ERCOT.

All the data are stored in Hierarchical Data Format 5 (HDF5) files and have been uploaded to an Amazon Web Services repository[1]. The ERCOT data set has 2 years (2017, 2018) of actuals and 1 year (2018) of probabilistic forecasts. These data are provided at various spatial (i.e., site-level, zone-level, and system-level) and temporal scales (i.e., day-ahead, intraday, and intra-hour). Specifically, data are provided for 125 existing wind sites, 22 existing solar sites, 139 proposed wind sites, and 204 proposed solar sites. The following variables are provided for ERCOT:

- Metadata (coordinates, capacity, and other configuration data):
    - 125 actual wind sites
    - 139 proposed wind sites
    - 22 actual solar sites
    - 204 proposed solar sites.
- Actual data (power [MW]):
    - Wind power (site level, zone level, and system level)
    - Solar power (site level, zone level, and system level)
    - Load (zone level and system level).
- Probabilistic point forecasts(power [MW]):
    - Wind power (site level, zone level, and system level)
    - Solar power (site level, zone level, and system level)
    - Load (zone level and system level).

Code examples are also provided to extract data from the HDF5 files, as shown in Figure ES-1.

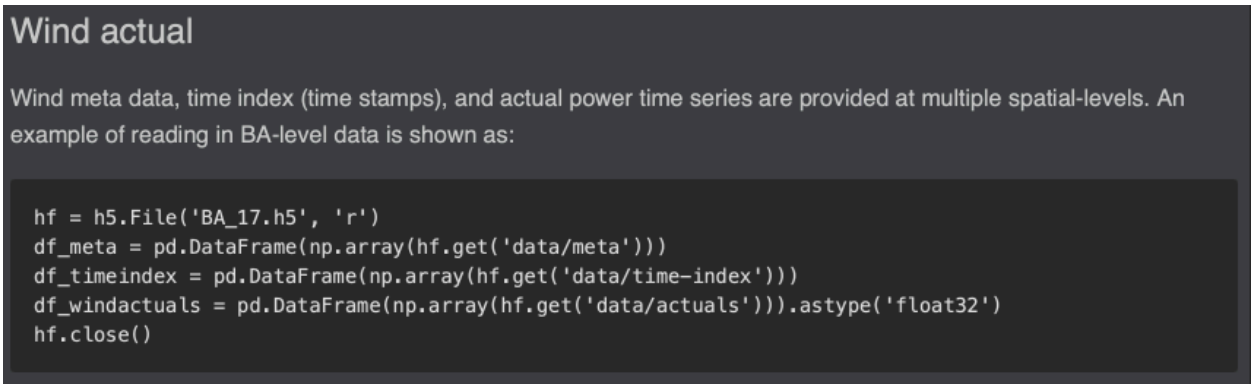

Figure ES-1. A code segment to extract actual wind power

[1] https://registry.opendata.aws/arpa-e-perform/

Table of Contents

List of Figures

List of Tables

1 Chapter 1: Renewable Generation and Load Basis Datasets

Wind and solar PV generation datasets are developed to serve as an uncertainty qunaitifcation basis for the Performance-Based Energy Resource Feedback, Optimization, and Risk Management (PERFORM) program. The basis solar generation datasets are produced from the meteorological data available in the National Solar Radiation Database (NSRDB) for 204 utility-scale proposed and 22 existing solar power plant locations in the Electric Reliability Council of Texas (ERCOT) service territory. These 226 existing and proposed plants comprise a high solar penetration scenario of 42 GW of solar photovoltaic (PV) capacity. Figure 1 shows a map of the 226 plant locations, taken from geolocation data from the ERCOT interconnection queue as of May 2019. The basis wind datasets are produced though a similar method, where the Wind Integration National Dataset (WIND Toolkit) [1] meteorological data was utilized. The load actual dataset is was produced in collaboration with the team that created the ACTIVSg2000 2,000-bus synthetic grid [2]. Additionally, the National Renewable Energy Laboratory's (NREL's) ResStock® and ComStock® platforms [3] are leveraged to produce the basis time-series load datasets within ERCOT. The following sections provide details into the development of the Wind and solar PV generation and load basis datasets for ERCOT.

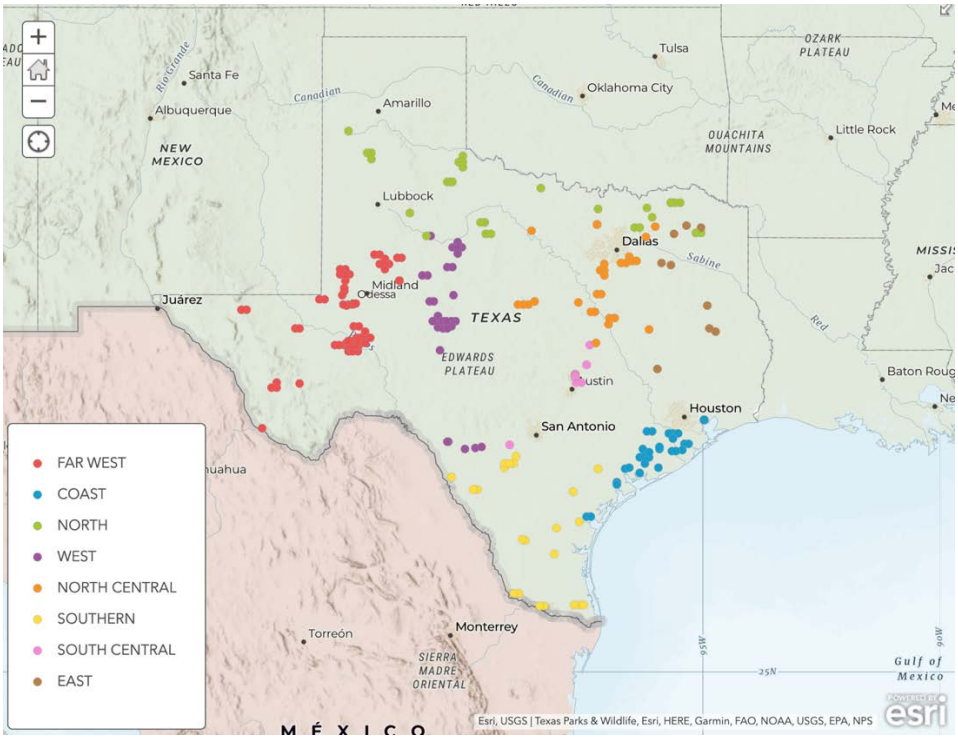

Figure 1. A map of the 266 solar power plant locations considered in the high solar penetration scenario

1.1 Basis Solar Datasets

The solar actuals were up-sampled from 30-minute to 5-minute temporal resolutions and from 4-km x 4-km to 2-km x 2-km geographic resolutions for the year 2017 using a method developed internally by NREL's NSRDB team, whereas the 2018 data were natively at the desired

resolution [4]. Consistent geographic resolution, temporal resolution, and meteorological years will be used for the wind, solar, and load data. The data are provided for the years 2017 and 2018 and are organized based on three geographic scales: site level, zone level, and total ERCOT system balancing area.

The 226 existing and proposed solar plant in the ERCOT service territory represent a total of 42 GW of AC capacity and feature diverse configurations; one of 30 potential solar energy system configurations were assigned to each plant at random. All profiles assumed a single-axis tracking system with zero-degree tilt in the north-south direction and differed with respect to thin-film versus silicon panel material, azimuthal angle (systems were assigned either 165-, 180-, or 195-degree azimuth), and DC-to-AC ratios (ranging from 1.2 to 1.4). This process resulted in 226 unique hypothetical solar energy plants, with the distribution of these attribute values shown in Table 1.

Table 1. Module Type, DC-to-AC Ratio, and Azimuth Were Randomly Assigned to the 226 Solar Power Plant Sites in the Numbers and Percentages

Module Type	Count	%	DC-to-AC Ratio	Count	%	Azimuth	Count	%
Silicon	177	78%	1.2	26	12%	165	10	4%
Thin Film	49	22%	1.25	50	22%	180	205	91%
			1.3	59	26%	195	11	5%
			1.35	51	23%			
			1.4	40	18%			

These hypothetical plants served as inputs to the System Advisor Model's (SAM's) PVWatts® platform [5], generating plant-level energy production data for the years 2017 and 2018 with the up-sampled time resolution of 5 minutes and a geographic resolution of 2 km x 2 km. The PVWatts preset system losses totaled 14%, so, initially, the output results featured very little clipping (Figure 2). Because of the high solar resource in Texas, more clipping is expected than is exhibited using these presets, so we reduced the expected loss inputs; After eliminating all losses except for the inverter efficiency losses, the power generation profile, shown in Figure 3 and Figure $, was produced which illustrates an unreasonably high level of clipping.

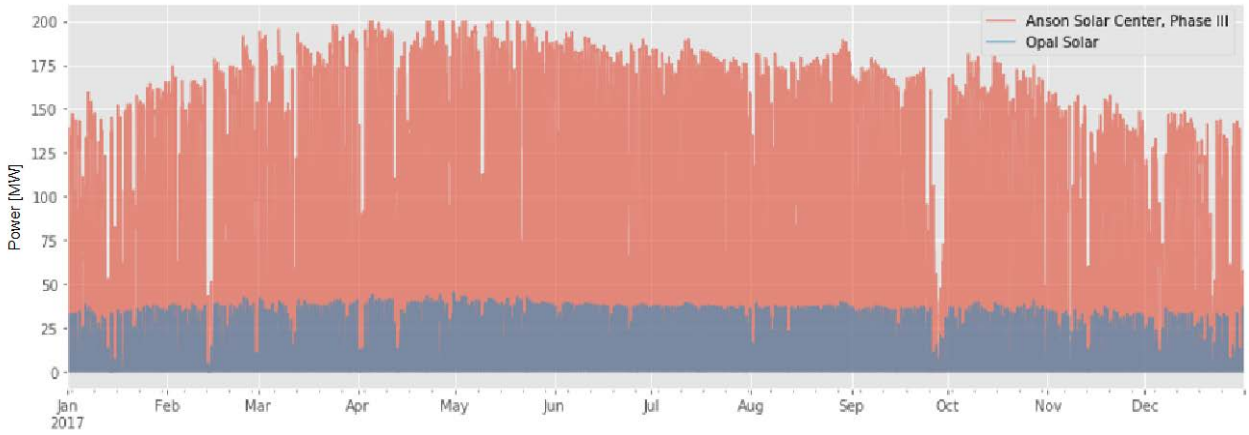

Figure 2. The output of two hypothetical plants with 14% system losses

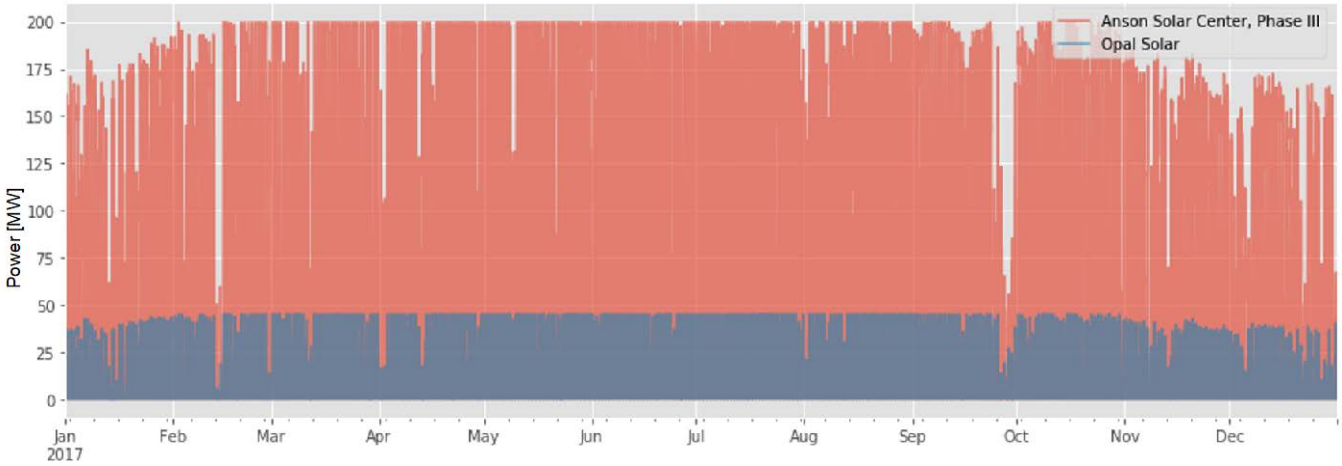

Figure 3. Very high clipping due to low system losses. All losses except for inverter efficiency losses were set to zero.

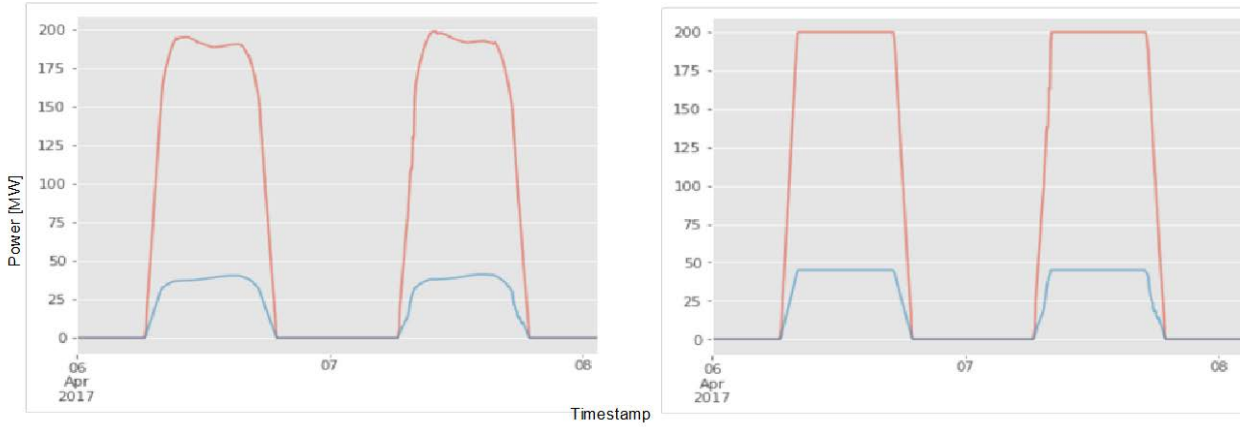

Figure 4. Production curves for April 7–8, 2017, under the high-loss and low-loss settings on the SAM platform

To produce a more reasonable solar generation dataset, losses were modulated from the high-loss setting by incorporating all modeled losses featured in PVWatts except for soiling, shading, and

availability losses as these loss vectors are crudely applied uniformly over time, which would be unrealistic. Utility-scale systems do not typically experience shading except for interrow shading, so we are assuming the plant will be installed to avoid this. With the panels typically non-horizontal because of the tracking, debris and dust should not accumulate too much, and soiling should not be a significant issue. Finally, we decided not to incorporate availability because the way this is modeled in the SAM system is to decrease daily energy production by a percentage—instead of modeling plant failure realistically as a few hours or days at a time—and would produce inaccurate results. Eliminating all three of these losses resulted in a total loss of 6.81%.

1.2 Basis Wind Data Sets

A total of 125 existing wind plants and 139 proposed wind plants wind power plant locations in ERCOT were identified, as shown in Figure 5. The wind speed time series was converted to a basis wind power time series using NREL's SAM. All wind turbines were assumed to have a height of 100 m. Three parameters—including wind speed at 100 m, surface temperature, and surface pressure—were used to generate wind power. Then, wind speeds were converted to power by applying normalized power curves for turbine classes 1–3. When wind speeds above cut-out wind speed occurred, wind turbines shut down and would not restart until the wind speed was 5 meters per second (m/s) slower than the cut-out speed [1]. The power curves used are composites of three to four commercially available wind turbines commonly used for each wind class. Examples of wind speed-to-wind power conversion are shown in Figure 6.

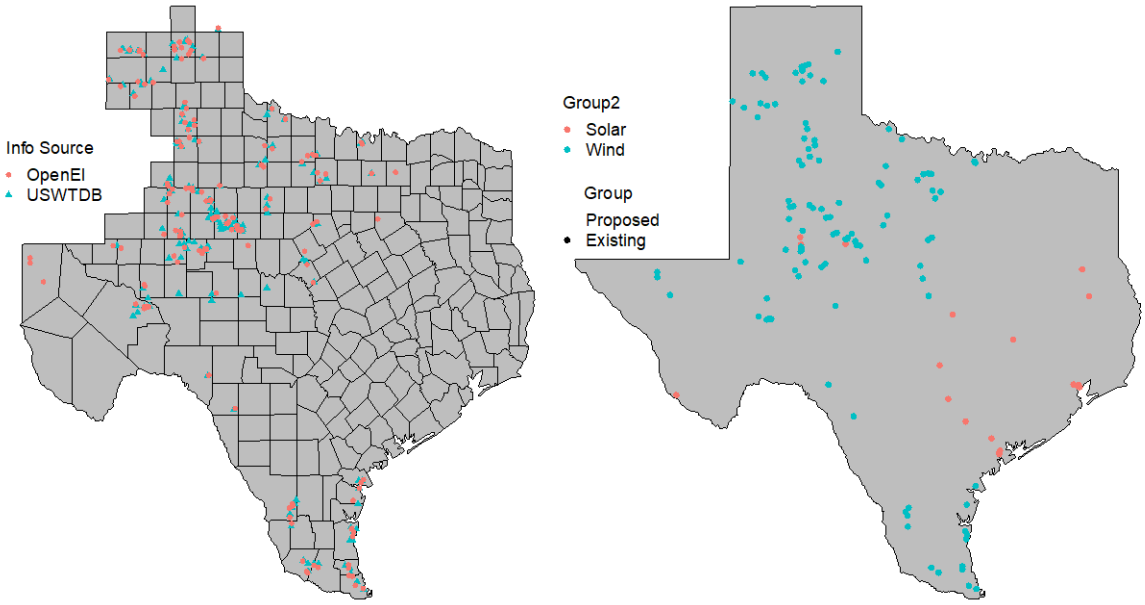

Figure 5. Wind and solar PV plant geospatial locations: (left) existing wind plants in ERCOT based on the U.S. Wind Turbine Database and the OpenEI data set and (right) the existing and proposed solar and wind plants

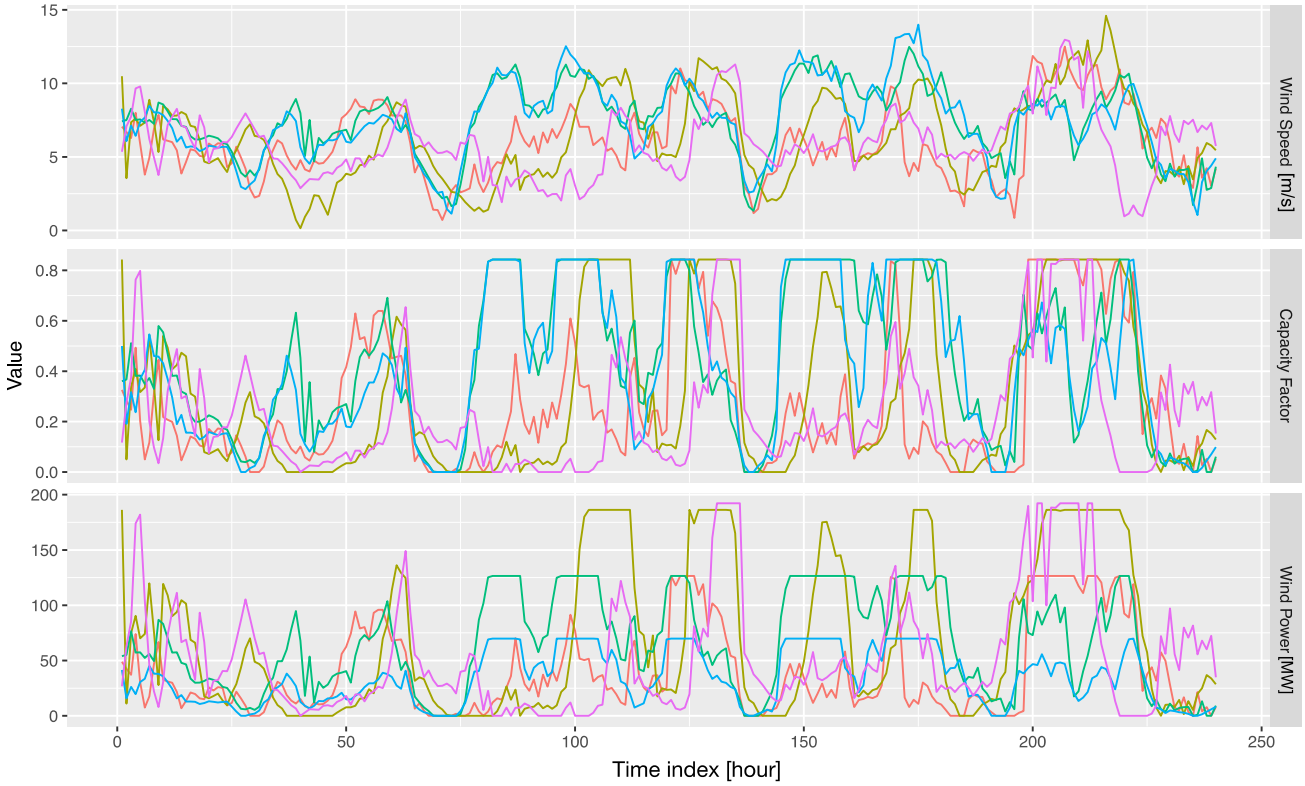

Figure 6. Wind speed-to-wind power conversion examples of five wind plants

2 Chapter 2: Forecasting Wind, Solar, and Load Datasets

2.1 Forecast Timescales

The probabilistic solar forecasts for the PERFORM program were produced on three timescales—day ahead, intraday, and hour ahead—as detailed in Table 2. The lead times, defined as the length of time between the forecast issuance and the beginning of the forecasted time period, are also defined in Table 2. The horizon is defined as how far into the future the forecast extends from the initial forecast time. The temporal resolution is the resolution of the forecast itself. Finally, the update rate is the time between subsequent issue times. A timeline illustrating these forecast attributes is shown in Figure 7.

Table 2. Forecasts Are Provided on Three Timescales at the Lead Time, Horizon, Resolution, and Update Rates.

Forecast Run	Lead Time	Horizon	Resolution	Update Rate	Method
Day-ahead (Medium-Term)	12 hours	48 hours	Hourly	Daily	BMA (using ECMWF ensemble data)
Intra-day (Short-Term)	6 hours	6 hours	Hourly	6 hours	M3 and/or BMA (using ECMWF ensemble data)
1 hour-ahead (Very Short-Term)	1 hour	2 hours	15 minutes	Hourly	M3 Machine learning/time series

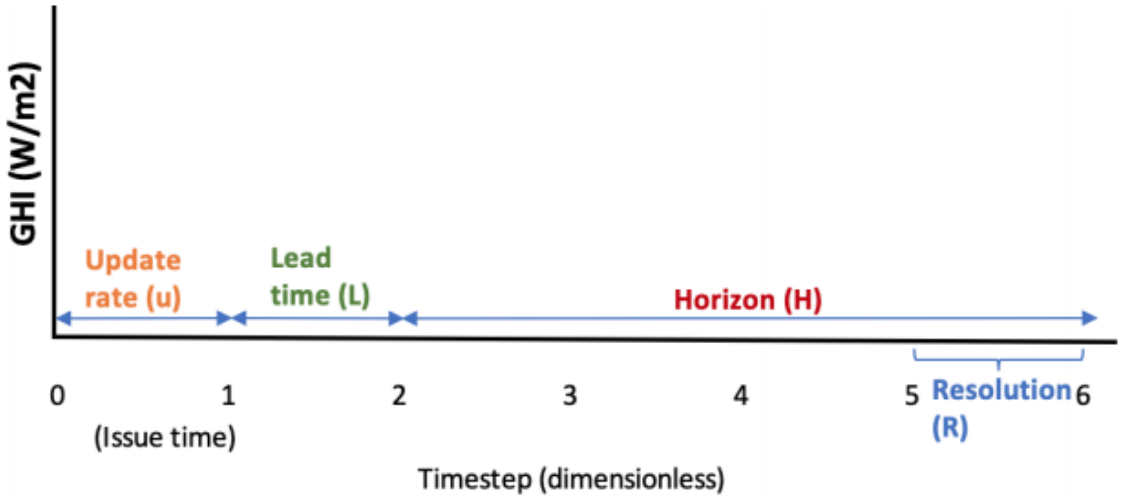

Figure 7. An illustration of the relationships among forecast attributes

2.1.1 Operational Time Constraints and Timesteps

For each forecast, the start and end times are constrained by operational timelines. The ERCOT system operator must receive the day-ahead forecast by noon CST (1:00 p.m. CDT), with the time horizon spanning from midnight the following day to the following midnight. In some cases, the system operator might wish to have a day-ahead forecast with a 1-day look-ahead (a 2-day-ahead forecast), hence the longer time horizon for this forecast. The issue times and start/end

times for these forecasts are detailed in Table 3. Note that the European Centre for Medium-Range Weather Forecasts (ECMWF) data are given in UTC, so the time steps chosen for these forecasts vary depending on whether the local time is standard time (early November through mid-March) or daylight savings time (mid-March through early-November) as detailed in Table 3.

Table 3. Issue Time, Horizon Start Time, and Horizon End Time for Day-Ahead and 2-Day-Ahead (Medium-Term) Forecasts

Day-Ahead Times (CST)							
Issue Time		**START**			**END**		
UTC	*CST*	*Step*	*Time UTC*	*Time CST*	*Step*	*Time UTC*	*Time CST*
18:00	12:00	12	06:00 day x+1	00:00 day x+1	35	05:00 day x+2	23:00 day x+1

Day-Ahead Times (CDT)							
Issue Time		**START**			**END**		
UTC	*CDT*	*Step*	*Time UTC*	*Time CDT*	*Step*	*Time UTC*	*Time CDT*
18:00	13:00	11	05:00 day x+1	00:00 day x+1	34	04:00 day x+2	23:00 day x+1

2-Day-Ahead Times (CST)							
Issue Time		**START**			**END**		
UTC	*CST*	*Step*	*Time UTC*	*Time CST*	*Step*	*Time UTC*	*Time CST*
18:00	12:00	36	06:00 day x+2	00:00 day x+2	59	05:00 day x+3	23:00 day x+2

2-Day-Ahead Times (CDT)							
Issue Time		**START**			**END**		
UTC	*CDT*	*Step*	*Time UTC*	*Time CDT*	*Step*	*Time UTC*	*Time CDT*
18:00	13:00	35	05:00 day x+2	00:00 day x+2	59	04:00 day x+3	23:00 day x+2

For the intraday forecast, the start and end times are constrained by the weather model data update time (how quickly new data become available) and the balancing gate closure time, defined as the "point in time when submission or update of a balancing energy bid for a standard product on a common merit order list is no longer permitted" [6]. Gate closure time has been added to the schematic in Figure 8.

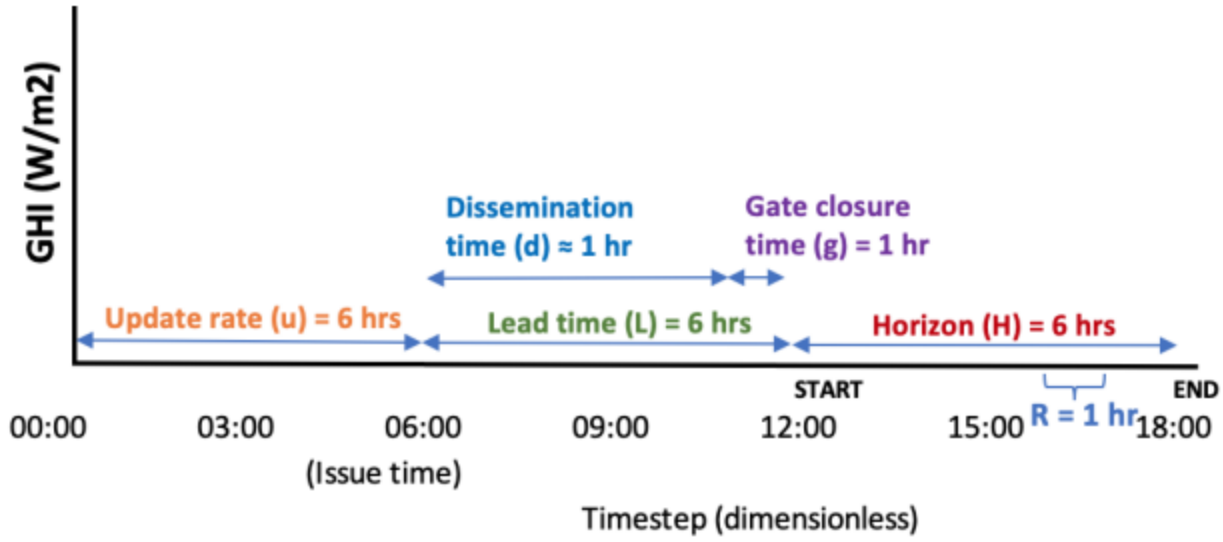

Figure 8. The forecast attributes, this time with operational times added. This figure is only reflective of the intraday forecast.

Essentially, the only interesting forecasts are those upon which the operator can base decisions. The forecast data dissemination time is 6 hours and 40 minutes for ECMWF's operational 15-day ensemble forecasts, according to their published dissemination times [7], and it is assumed to be shorter for a single day of data. With a balancing gate closure time of approximately an hour for most operators, the lead time is assumed to be approximately 6 hours total. Temporal attributes for the intraday forecasts are detailed in Table 4.

Table 4. Issue Time, Horizon Start Time, and Horizon End Time for Intraday (Short-Term) Forecasts

Intra-Day Times (CST)							
Issue Time		**START**			**END**		
UTC	*CST*	*Step*	*Time UTC*	*Time CST*	*Step*	*Time UTC*	*Time CDT*
00:00	18:00 day x-1	6	06:00	00:00	11	11:00	05:00
06:00	00:00	6	12:00	06:00	11	17:00	11:00
12:00	06:00	6	18:00	12:00	11	23:00	17:00
18:00	12:00	6	00:00 day x+1	18:00	11	05:00 day x+1	23:00

Intra-Day Times (CDT)							
Issue Time		**START**			**END**		
UTC	*CDT*	*Step*	*Time UTC*	*Time CDT*	*Step*	*Time UTC*	*Time CDT*
00:00	19:00 day x-1	5	05:00	00:00	10	10:00	05:00
06:00	01:00	5	11:00	06:00	10	16:00	11:00
12:00	07:00	5	17:00	12:00	10	22:00	17:00
18:00	13:00	5	23:00	18:00	10	04:00 day x+1	23:00

2.2 European Centre for Medium-Range Weather Forecasts Numerical Weather Predictive Datasets

The European Centre for Medium-Range Weather Forecasts (ECMWF) Meteorological Archival and Retrieval System (MARS) database provides 30 years of archival meteorological data to authorized users. MARS holds petabytes of data, mainly using the GRIB format for

meteorological fields. The data required for producing solar forecasts with the PVWatts model include surface-level global horizontal irradiance (GHI), direct normal irradiance (DNI), air temperature, and wind speed. The relevant ECMWF parameters are shown in Table 5.

Table 5. The Surface-Level Parameters Gathered from ECMWF's MARS Database

ECMWF Output	PV Watts Input
Surface Solar Radiation Downward (ssrd)	GHI
Total sky direct solar radiation at surface (fdir)	DNI
10 metre U wind component (east) (u10)	windspeed
10 metre V wind component (north) (v10)	windspeed
2 metre temperature (ambient) (t2m)	air_temperature

ECMWF data collection was divided into four queries: 2017 intraday, 2017 day ahead, 2018 intraday, and 2018 day ahead. These queries were run on NREL's Eagle high-performance computing system and a local server using NREL's ECMWF Python scripts. The queries are gathered the parameters detailed in Table 5 for 50 perturbed ensemble members and one control member. They span 60 time steps for the day-ahead forecast (with an issue time of 18:00 UTC) and 12 time steps for the intraday forecast (with issue times of 00:00, 06:00, and 12:00 UTC). The resulting ECMWF datasets are converted from the GRIB format to the Hierarchical Data Format 5 (HDF5) and converted from four-dimensional variables to single-dimensional hourly time-series data according to Table 6 to serve as inputs into the SAM PVWatts platform.

Table 6. A Comparison of the ECMWF Output Files with the PVWatts Input Files

	ECMWF Output	PV Watts Input
Files	730 Daily Output Files *(365 days per year x 2 years)*	678 Site-Specific Time-Series Files *(226 sites x 3 timescales)*
Type	GRIB	HDF5
Variables	5 [ssrd, fdir, u10, v10, t2m]	4 [GHI, DNI, windspeed, air_temperature]
Dimensions	4 [longitude, Size: 591 latitude, Size: 251 number, Size: 51 step, Size: 60]	1 [time, Size: 8760] with 51 members per file
Geographic Resolution	All of Texas (0.1 lat x 0.1 lon)	Site-level (closest lat/lon coordinates)
Temporal Resolution	Hourly	Hourly

2.3 Solar Forecast Datasets

The short-term and medium-term solar forecasts for this project were trained on hypothetical historical solar production data at the 226 locations obtained from the NSRDB and NREL's SAM PVWatts model, as detailed in Chapter 1. The inputs to the PVWatts model include both the site characteristics from the 226 sites and 2017–2018 ensemble model weather data from the ECMWF's MARS, which is widely considered to be the most accurate numerical weather prediction (NWP) model at the time resolutions that this project requires. The forecast error metrics will be verified with the NSRDB 2018 "actuals." The overall process for the solar portion of this project is detailed in Figure 9.

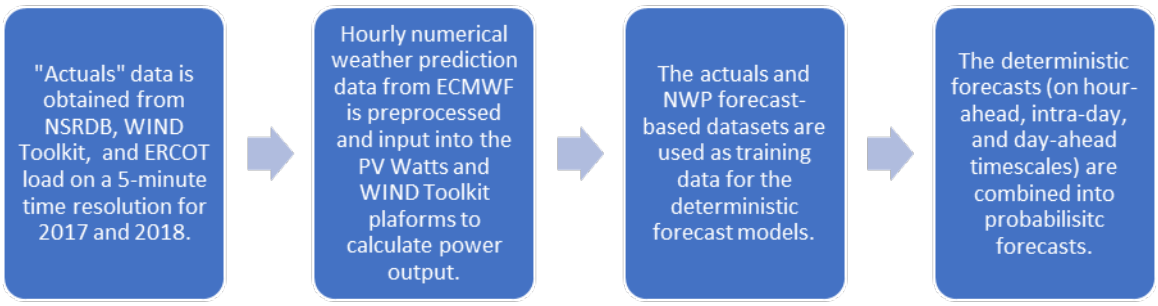

Figure 9. A schematic of the overall data collection and forecasting process

2.3.1 ECMWF-Based Solar Power Forecasts

The daily weather files queried from ECMWF for 2017 and 2018 were converted to site-level files for all 226 utility-scale solar power plants considered in the high renewable penetration scenario via the process detailed in the proceeding section. These site-level files, containing 51 ensemble members, were then converted to power using the SAM PVWatts platform. The latest version of the SAM PVWatts can calculate diffuse horizontal irradiance based on geographic location and solar zenith angle. PVWatts takes as input GHI, DNI, and temperature and calculates capacity factor, which is then converted to megawatt output by multiplying by the AC capacity of the plant.

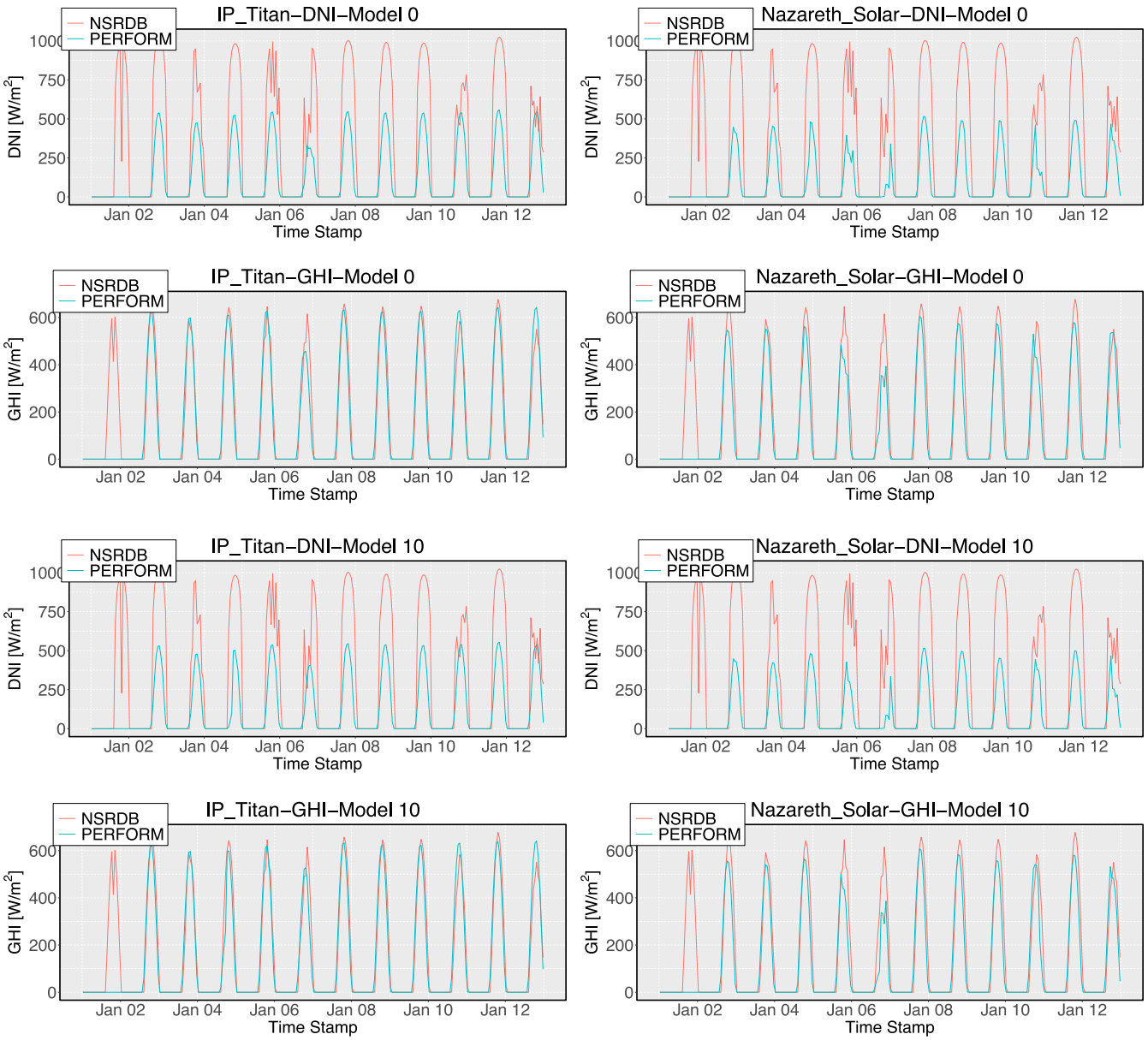

Figure 10. ECMWF direct flux vs. NSRDB DNI. Three solar sites and two ensemble models are randomly selected for comparison.

The ECMWF forecasts provide direct flux (defined as DNI*cos(zenith angle)), which is significantly smaller than DNI. This is shown in Figure 10, where comparisons of the ECMWF irradiance and NSRDB irradiance are visualized. It is observed that the GHI forecast error is small, whereas the DNI forecast error (direct flux forecast – actual DNI) is much larger; therefore, we first modified the DNI forecasts in the ECMWF by converting direct flux to DNI, which is shown in Figure 11. After improving the DNI and using the same PV system loss (6.81%), the updated ECMWF-based day-ahead and intraday solar power forecasts were generated. Figure 12 shows the ECMWF solar power forecasts from an ensemble member (discussed above) and their comparisons to the NSRDB-based actual solar power. The forecast errors are reduced.

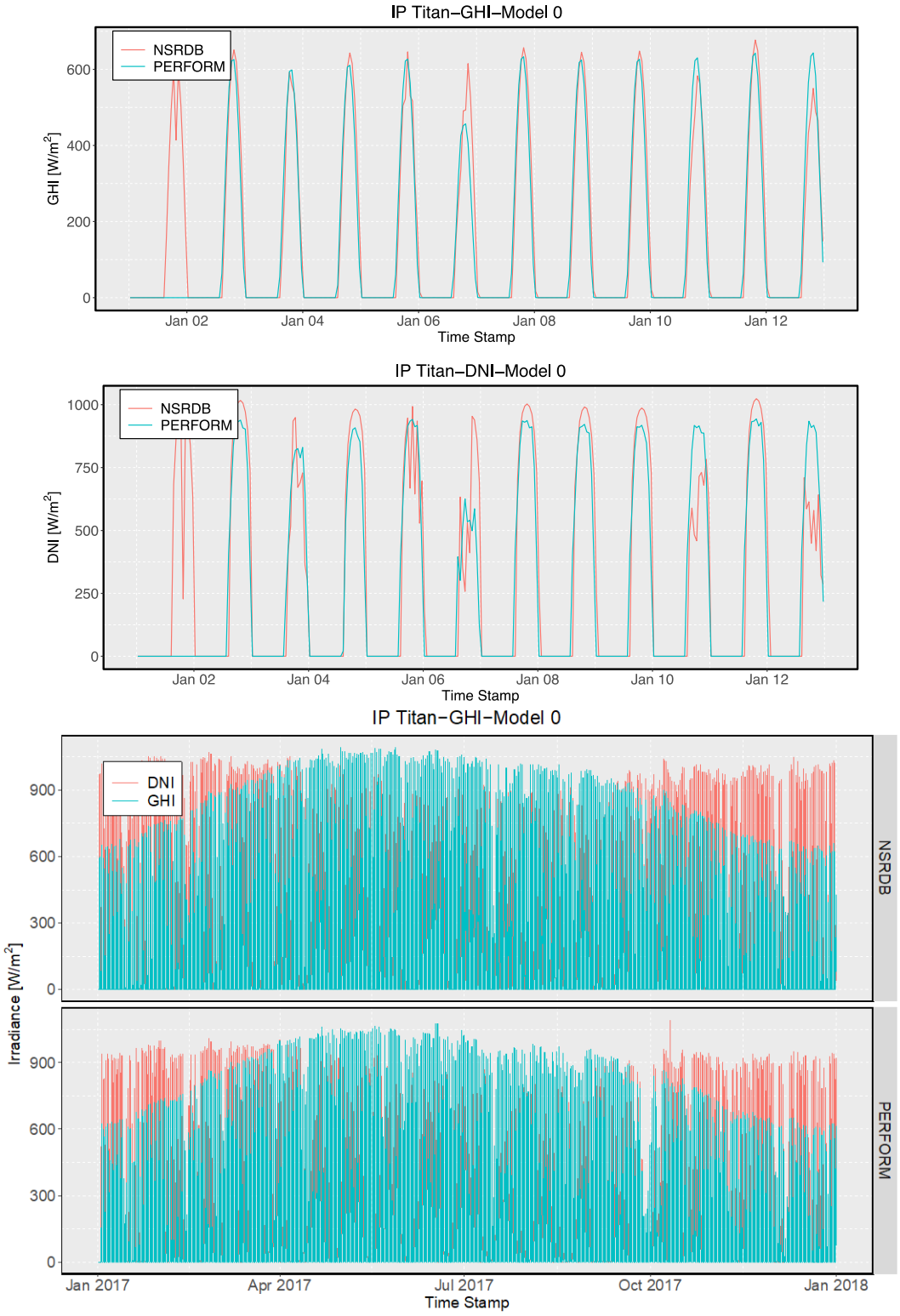

Figure 11. Comparison of the NSRDB DNI and the modified ECMWF DNI forecasts. The upper plots show the 10-day DNI and GHI curves in the ECMWF and NSRDB. The lower plots show the 1-year DNI and GHI curves in the two databases.

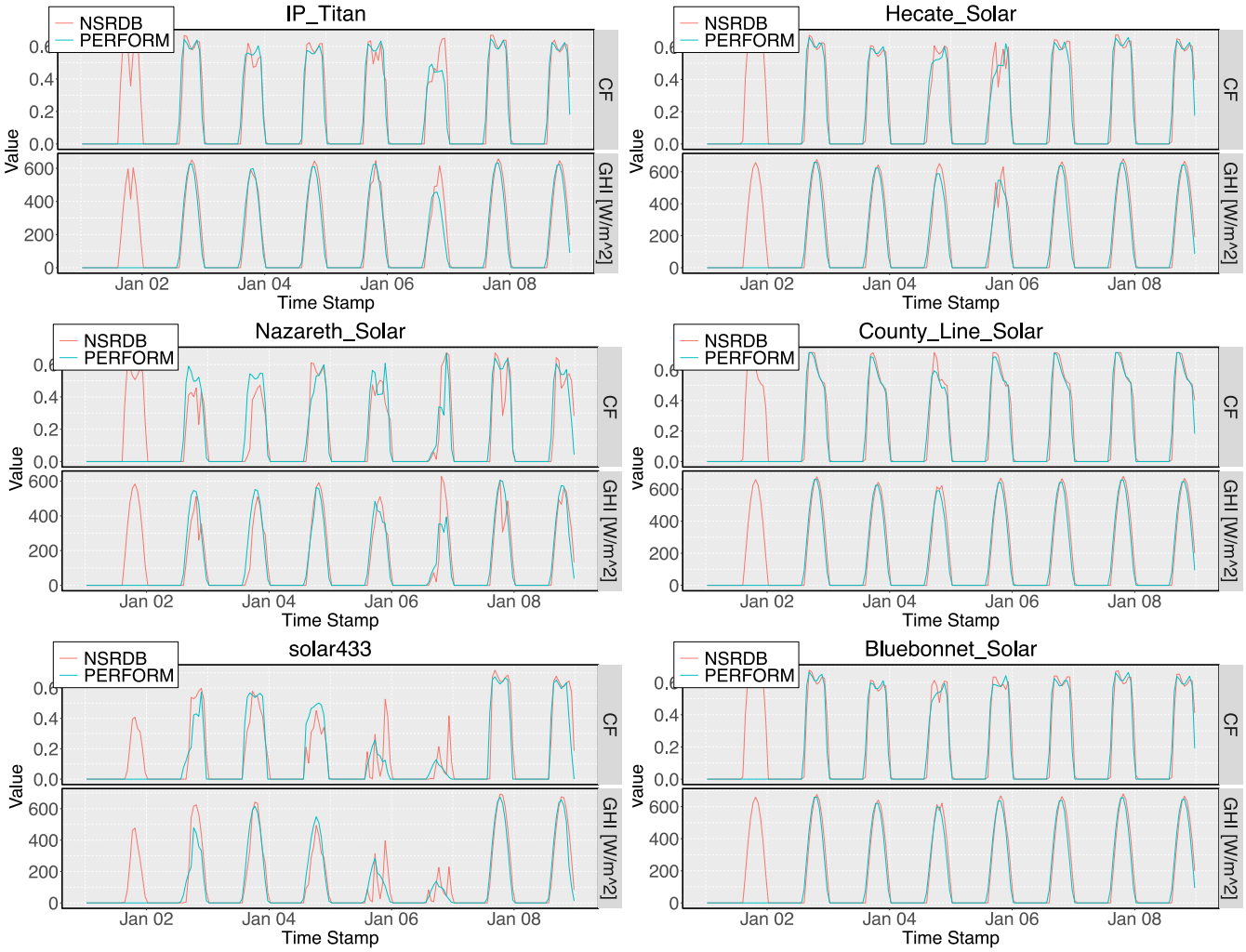

Figure 12. The updated ECMWF solar power forecasts. The first 200 hours of data of the randomly-selected 9 PV sites are used for demonstration.

2.3.2 Probabilistic Solar Power Forecasts

Two probabilistic methods were used to develop the solar power forecasts: the machine learning-based multi-model method (M3) and the Bayesian model average (BMA).

2.3.2.1 M3 Method

The M3 forecasting framework is a two-step data-driven methodology that provides both point and probabilistic forecasts for very short-term wind [8], solar [9], and load forecasting [10]. The M3 method generates point forecasts in the first step with a two-layer machine learning ensemble algorithm, which serves as the input to the pinball loss optimization-based predictive distribution model to generate the probabilistic forecasts in the second step.

Specifically, in the point forecasting step, a collection of machine learning models form the first layer, including three artificial neural networks with backpropagation, three support vector regression models with different kernels, three gradient boosting machine models with three different distribution functions, and a random forest model. These models generate independent point forecasts, which were ensembled in the second layer by another machine learning model to

generate the final point forecasts. The ensemble model is expected to provide more accurate and robust point forecasts than single-algorithm models.

Then, in the probabilistic forecasting step, the genetic algorithm was used to optimize the standard deviation of the predictive distributions, sigma, given the mean value (assumed to be the point forecast from the first step). A support vector regression surrogate model was first trained based on the point forecasts and sigma values, which was used to estimate the pseudo sigma to generate the quantiles.

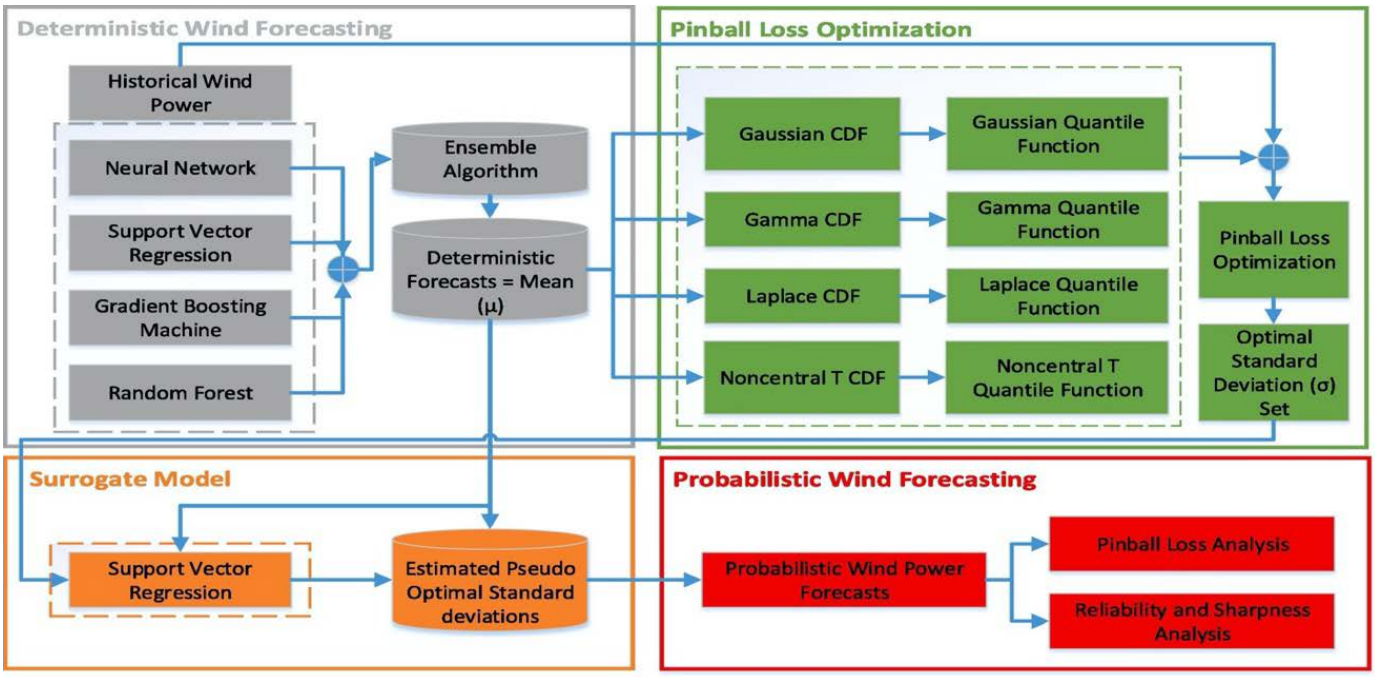

Figure 13. Overall framework of the M3 method

During the development stage, the M3 model was trained for a total of 22 ERCOT solar sites with the 2017 data and tested with 10 representative days. The M3 method outputs point forecasting time series and 1^{th}–99^{th} quantiles to form the probabilistic forecast. Table 7 shows the point and probabilistic forecast accuracy of the M3 method. Further development is planned to improve the M3 performance by applying a zenith angle filter and emphasizing reliability optimization. Figure 14 shows a probabilistic forecasting time series in the form of quantiles (i.e., 1^{st}, 5^{th}, 95^{th}, and 99^{th} quantiles). Results indicate satisfying reliability and sharpness.

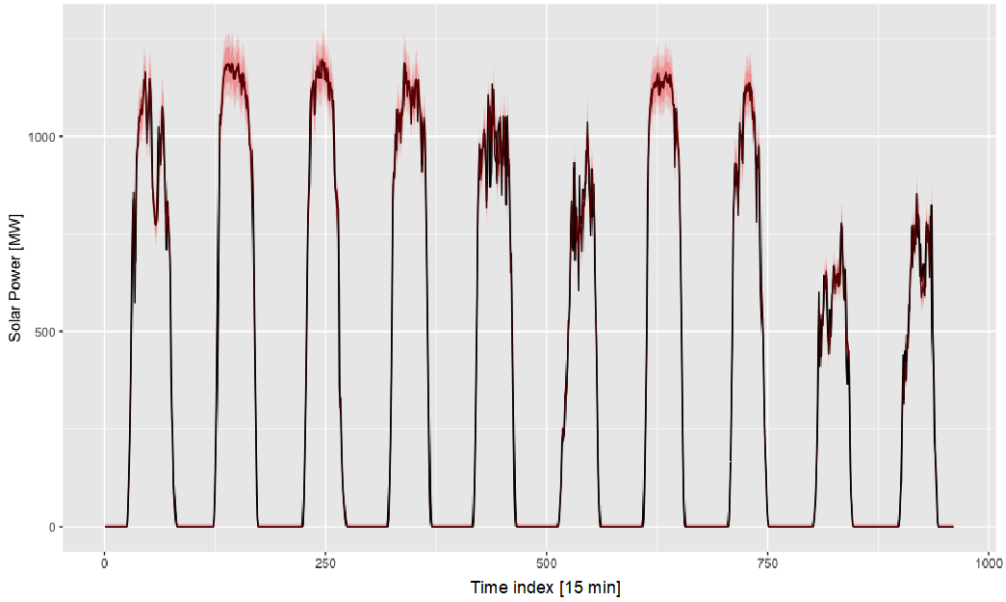

Figure 14. Probabilistic forecasting time series of 22 solar power plants

Table 7. M3 Point and Probabilistic Forecast Accuracy

nRMSE [%]	nMAE [%]	MBE [MW]	CRPS [MW]	Mean Pinball Loss [MW]
2.45	1.27	-1.9E-19	12.57	5.72

In addition to balancing area-level forecasts, intra-hour forecasts at the solar site level were also generated. The normalized mean absolute error (nMAE) of the M3 method and the smart persistence benchmark are listed in Table 8 and Table 9, respectively. Different columns indicate different lead times (e.g., 4-SA means 4-step-ahead, which is 1 hour ahead for 15-min resolution data), whereas different rows indicate the 22 solar plants. It was found that the proposed M3 method provides more accurate point forecasts. Also, probabilistic forecast normalized continuous ranked probability score (nCRPS) was used to assess the probabilistic forecast accuracy, which is detailed in Table 10. By comparing the nCRPSs to state-of-the-art probabilistic forecasting methods that applied to similar solar sites, the M3 probabilistic forecasts show competitive to better accuracy.

Table 8. Point Forecast nMAE of the M3 Method (Decomposed to Different Steps Ahead)

4-SA	5-SA	6-SA	7-SA	8-SA	9-SA	10-SA	11-SA	4-7SA	8-11SA	Overall
5.72%	6.18%	6.88%	7.34%	7.80%	8.74%	8.30%	8.57%	6.53%	8.35%	7.60%
5.94%	6.62%	6.92%	8.07%	8.88%	9.69%	9.43%	9.70%	6.89%	9.42%	8.37%
6.72%	7.30%	7.63%	8.31%	8.65%	9.19%	9.61%	9.92%	7.49%	9.34%	8.61%
6.16%	6.77%	7.17%	7.62%	8.12%	8.42%	8.88%	9.05%	6.93%	8.62%	7.95%
6.29%	6.91%	7.48%	7.96%	8.24%	8.55%	8.77%	9.29%	7.16%	8.71%	8.12%
7.01%	7.50%	8.07%	8.36%	8.91%	9.15%	9.55%	9.86%	7.74%	9.37%	8.76%
6.84%	7.33%	7.88%	8.35%	8.66%	8.81%	9.13%	9.51%	7.60%	9.03%	8.52%
6.65%	7.03%	7.57%	8.07%	8.33%	8.63%	8.84%	9.37%	7.33%	8.79%	8.28%
5.83%	6.48%	6.88%	7.71%	7.86%	8.08%	8.27%	8.56%	6.72%	8.19%	7.62%
6.32%	7.03%	7.37%	7.78%	8.15%	8.89%	9.13%	9.12%	7.12%	8.82%	8.13%
6.62%	7.41%	8.14%	8.41%	9.30%	9.37%	9.40%	9.90%	7.64%	9.49%	8.71%
6.30%	6.96%	7.69%	8.09%	8.48%	8.65%	9.17%	10.11%	7.26%	9.11%	8.35%
6.57%	7.24%	7.92%	8.28%	8.84%	9.33%	9.95%	10.38%	7.50%	9.63%	8.81%
6.47%	6.97%	7.33%	7.85%	8.42%	8.90%	9.32%	9.70%	7.16%	9.08%	8.34%
6.85%	7.24%	7.86%	8.23%	8.76%	9.27%	9.52%	10.01%	7.55%	9.39%	8.69%
6.37%	6.82%	7.68%	7.87%	8.28%	8.65%	8.95%	9.31%	7.18%	8.80%	8.21%
6.75%	6.84%	7.42%	7.97%	8.48%	8.97%	9.21%	9.74%	7.24%	9.10%	8.39%
6.60%	7.01%	7.58%	8.10%	8.49%	9.14%	9.38%	9.85%	7.32%	9.22%	8.48%
6.27%	6.68%	7.21%	7.73%	8.11%	8.57%	8.92%	9.29%	6.97%	8.72%	8.05%
6.85%	7.49%	7.96%	8.39%	8.86%	9.25%	9.53%	9.82%	7.67%	9.36%	8.70%
6.42%	7.05%	7.66%	8.07%	8.45%	8.87%	9.10%	9.51%	7.30%	8.98%	8.35%

Table 9. Point forecast nMAE of the Smart Persistence Benchmark (Decomposed to Different Steps Ahead)

4-SA	5-SA	6-SA	7-SA	8-SA	9-SA	10-SA	11-SA	4-7SA	8-11SA	Overall
9.24%	11.33%	13.38%	15.42%	17.37%	19.14%	20.79%	22.29%	12.34%	19.90%	16.12%
7.27%	8.47%	9.62%	10.87%	12.12%	13.29%	14.40%	15.47%	9.06%	13.82%	11.44%
8.46%	9.89%	11.36%	12.90%	14.38%	15.70%	16.93%	18.07%	10.65%	16.27%	13.46%
8.02%	9.60%	11.08%	12.42%	13.70%	14.91%	15.96%	16.94%	10.28%	15.38%	12.83%
9.89%	11.90%	13.91%	15.77%	17.56%	19.21%	20.77%	22.22%	12.87%	19.94%	16.40%
10.61%	12.74%	14.87%	16.85%	18.75%	20.50%	22.15%	23.69%	13.77%	21.27%	17.52%
14.23%	17.80%	21.24%	24.50%	27.61%	30.56%	33.32%	35.87%	19.44%	31.84%	25.64%
12.84%	15.99%	19.04%	21.95%	24.74%	27.38%	29.86%	32.16%	17.46%	28.53%	22.99%
12.66%	15.84%	18.93%	21.89%	24.75%	27.46%	30.01%	32.39%	17.33%	28.65%	22.99%
12.84%	15.99%	19.04%	21.95%	24.74%	27.38%	29.86%	32.16%	17.46%	28.53%	22.99%
6.87%	8.11%	9.22%	10.25%	11.21%	12.12%	12.95%	13.72%	8.61%	12.50%	10.56%
7.41%	8.76%	9.95%	11.07%	12.11%	13.09%	13.99%	14.82%	9.30%	13.50%	11.40%
6.56%	7.66%	8.70%	9.65%	10.55%	11.39%	12.25%	12.94%	8.15%	11.78%	9.96%
6.74%	7.87%	8.93%	9.89%	10.80%	11.67%	12.54%	13.25%	8.36%	12.07%	10.21%
5.95%	7.03%	8.07%	9.05%	9.95%	10.76%	11.52%	12.16%	7.53%	11.10%	9.31%
6.42%	7.58%	8.70%	9.76%	10.73%	11.60%	12.42%	13.12%	8.11%	11.97%	10.04%
5.80%	6.83%	7.78%	8.67%	9.48%	10.20%	10.85%	11.41%	7.27%	10.49%	8.88%
6.11%	7.24%	8.31%	9.38%	10.38%	11.29%	12.16%	12.91%	7.76%	11.68%	9.72%
6.19%	7.31%	8.39%	9.41%	10.35%	11.19%	11.98%	12.65%	7.83%	11.54%	9.68%
6.06%	7.16%	8.22%	9.23%	10.15%	10.97%	11.76%	12.42%	7.67%	11.33%	9.50%
7.90%	9.26%	10.62%	11.95%	13.16%	14.29%	15.43%	16.39%	9.93%	14.82%	12.38%

Table 10. Probabilistic Forecast nCRPS of the M3 Method (Decomposed to Different Steps Ahead)

4-SA	5-SA	6-SA	7-SA	8-SA	9-SA	10-SA	11-SA	4-7SA	8-11SA	Overall
4.73%	5.03%	5.54%	5.85%	6.16%	7.04%	6.48%	6.81%	5.29%	6.62%	6.06%
13.12%	13.23%	13.26%	13.52%	13.73%	13.94%	13.78%	13.83%	13.28%	13.82%	13.59%
6.33%	6.91%	7.23%	7.91%	8.25%	8.79%	9.21%	9.52%	7.10%	8.94%	8.21%
5.48%	6.02%	6.41%	6.85%	7.32%	7.63%	8.13%	8.31%	6.19%	7.85%	7.20%
18.29%	18.37%	18.44%	18.53%	18.55%	18.57%	18.61%	18.66%	18.41%	18.59%	18.52%
6.04%	6.50%	7.02%	7.30%	7.81%	8.05%	8.44%	8.77%	6.71%	8.27%	7.70%
6.34%	6.55%	6.85%	7.03%	7.19%	7.27%	7.40%	7.54%	6.69%	7.35%	7.11%
5.29%	5.51%	5.88%	6.35%	6.42%	6.54%	6.72%	7.07%	5.76%	6.69%	6.38%
4.62%	5.00%	5.47%	5.99%	6.06%	6.20%	6.44%	6.71%	5.27%	6.36%	5.95%
5.56%	6.20%	6.53%	6.93%	7.31%	8.06%	8.21%	8.17%	6.30%	7.94%	7.27%
5.40%	5.78%	6.31%	6.47%	6.87%	7.07%	7.17%	7.35%	5.99%	7.12%	6.64%
5.24%	5.56%	6.03%	6.42%	6.40%	6.58%	6.81%	7.72%	5.81%	6.87%	6.45%
6.54%	6.80%	7.05%	7.18%	7.43%	7.74%	8.03%	8.28%	6.89%	7.87%	7.50%
5.29%	5.63%	5.81%	6.17%	6.54%	6.79%	7.10%	7.28%	5.73%	6.93%	6.46%
5.52%	5.84%	6.37%	6.70%	7.18%	7.63%	7.87%	8.30%	6.11%	7.74%	7.12%
5.08%	5.45%	6.14%	6.31%	6.66%	6.97%	7.24%	7.57%	5.74%	7.11%	6.61%
5.42%	5.50%	5.98%	6.45%	6.91%	7.35%	7.60%	8.04%	5.84%	7.48%	6.85%
5.26%	5.63%	6.09%	6.56%	6.94%	7.47%	7.71%	8.10%	5.88%	7.56%	6.90%
5.39%	5.76%	6.26%	6.76%	7.12%	7.56%	7.93%	8.27%	6.04%	7.72%	7.08%
5.96%	6.56%	6.98%	7.40%	7.83%	8.23%	8.50%	8.75%	6.72%	8.33%	7.70%
5.98%	6.61%	7.20%	7.61%	7.99%	8.40%	8.63%	9.07%	6.85%	8.52%	7.89%

2.3.2.2 Bayesian Model Averaging Method

For the intraday and day-ahead probabilistic solar power forecasting, a Bayesian Model Averaging (BMA) method was implemented. BMA is a kernel dressing technique that applies a probability density to each member of an NWP ensemble, with each member dressed in a mixture model: each model includes a discrete component forecasting power clipped at the inverter rating plus a continuous kernel for outputs that are less than the rated maximum [11]. As is appropriate for this method, we gathered perturbed forecast data and control forecast NWP ensemble data from the MARS' ENFO (Ensemble Prediction System) model. These NWP ensembles are then post-processed with BMA to address weaknesses and to smooth the ensemble from a discrete set of points to a full cumulative distribution function, mitigating the sunny bias and underdispersion typically found in these ensembles. We converted the site-level files to PVWatts input files.

The intraday and day-ahead forecasting errors and scores using BMA with Beta kernels and three state-of-the-art benchmark models at Site Titan are detailed in Table 11 and Table 12, respectively. It is shown that the BMA outperforms the other three benchmark models and achieves the best accuracy.

In addition to the site-level probabilistic solar forecast, zone-level and balancing-level probabilistic forecasts are included in the delivery.

Table 11. Intraday Forecasting Errors and Scores at IP Titan

Method	CRPS [MW]	nMAE [%]	nRMSE [%]	MBE
BMA	10.26	4.23	7.48	-0.85
PeEn	20.45	9.58	9.67	1.27
Raw ensemble	17.25	5.42	8.64	-1.25
EMOS	12.92	5.45	7.90	-0.98

Table 12. Day-Ahead Forecasting Errors and Scores at IP Titan

Method	CRPS [MW]	nMAE [%]	nRMSE [%]	MBE
BMA	10.85	4.58	8.17	-0.94
PeEn	22.33	10.23	10.67	1.86
Raw ensemble	14.72	5.93	9.24	-1.30
EMOS	13.36	6.04	9.28	-1.13

Figure 15 shows a reliability plot of empirical coverage against nominal levels. For a perfectly calibrated simulation, the dotted line would follow the identity line, which is the red line shown in each plot. As can be seen, the median is generally close to the identity line. There is underdispersion at higher nominal level for BA levels solar as shown in Figure 16.

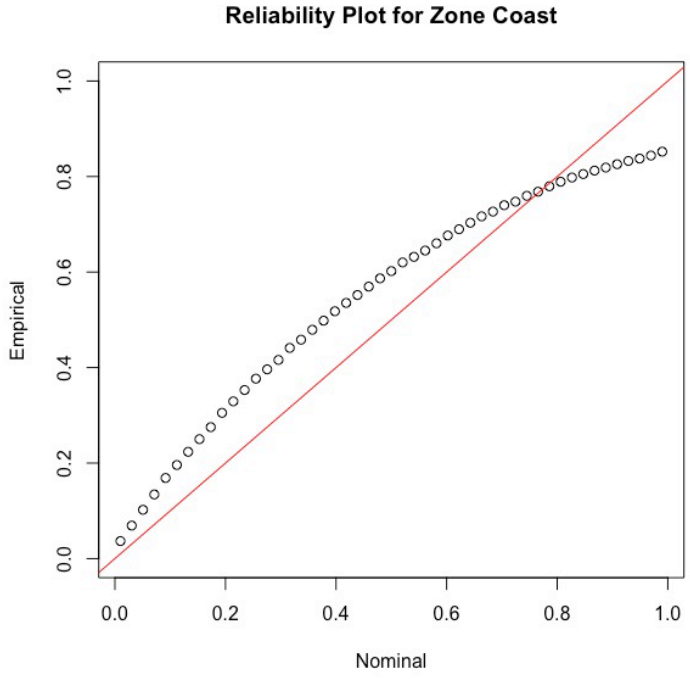

Figure 15: Reliability plot for day ahead solar forecasts for zone Coast

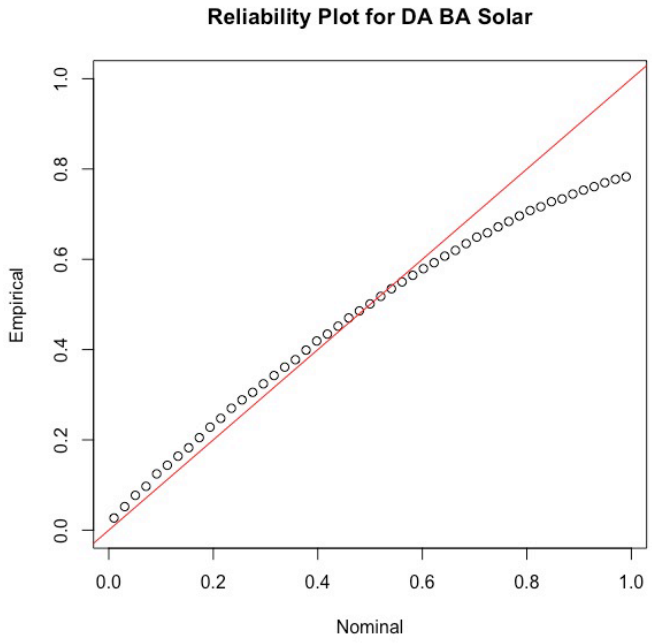

Figure 16: Reliability plot for day ahead solar forecasts for for entire balancing area.

2.4 Wind Forecast Datasets

2.4.1 ECMWF-Based Wind Power Forecasts

Wind power forecasts were obtained through wind power conversion. Similar to actual wind power simulation, forecasts of wind speed, temperature, and pressure are extracted from the ECMWF GRIB files and constructed as time series. Then, the SAM (reV) model is used to convert the weather time series to capacity factors and wind power. The same wind plant meta configuration setups are used in forecasting wind power conversion to maintain consistency.

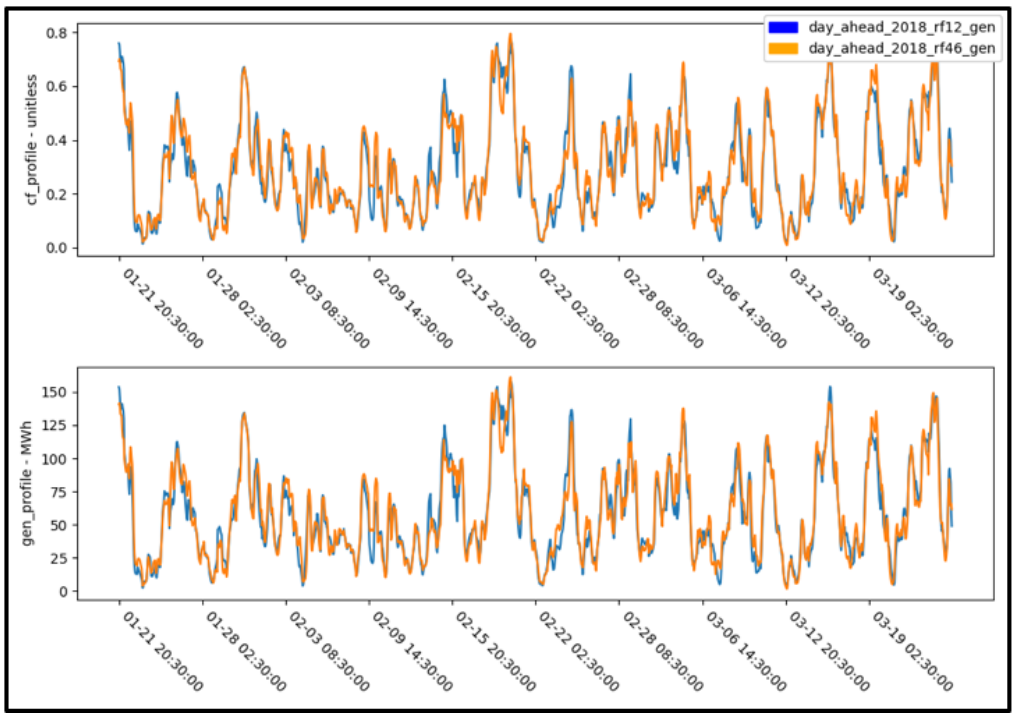

Figure 17. ECMWF wind power forecasts. The top figures show the capacity factor. The bottom figures show the wind power forecasts. Two ECMWF ensemble members were randomly selected for demonstration.

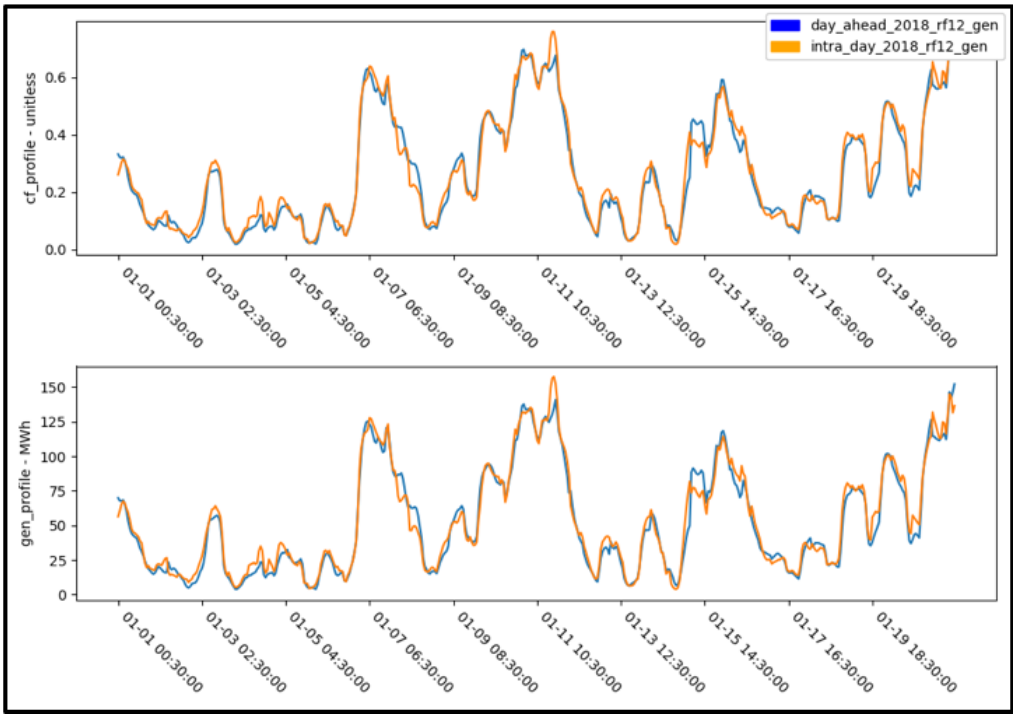

Figure 18. ECMWF wind power forecasts. The top figures show the capacity factor. The bottom figures show the wind power forecasts. Two ECMWF ensemble members were randomly selected for demonstration.

2.4.2 Probabilistic Wind Power Forecasts

Similar to the intraday and day-ahead probabilistic solar power forecasting discussed above, BMA was applied to the intraday and day-ahead probabilistic wind power forecasting. The only difference being that the Gaussian kernels perform better than the Beta kernels for wind power forecasting. The intraday and day-ahead forecasting errors and scores using BMA with Gaussian kernels and three state-of-the-art benchmark models at Site Aquilla Lake 2 Wind are detailed in Table 13 and Table 14, respectively. Both zone-level and balancing-level probabilistic wind forecasts are provided along with site-level wind forecasts. There is underdispersion at lower nominal levels for BA level wind, as shown in Figure 19.

Table 13. Intraday Forecasting Errors and Scores at Aquilla Lake 2 Wind

Method	CRPS [MW]	nMAE [%]	nRMSE [%]	MBE
BMA	7.62	5.86	18.24	4.85

Table 14. Day-Ahead Forecasting Errors and Scores at Aquilla Lake 2 Wind

Method	CRPS [MW]	nMAE [%]	nRMSE [%]	MBE
BMA	8.34	5.38	18.66	5.82

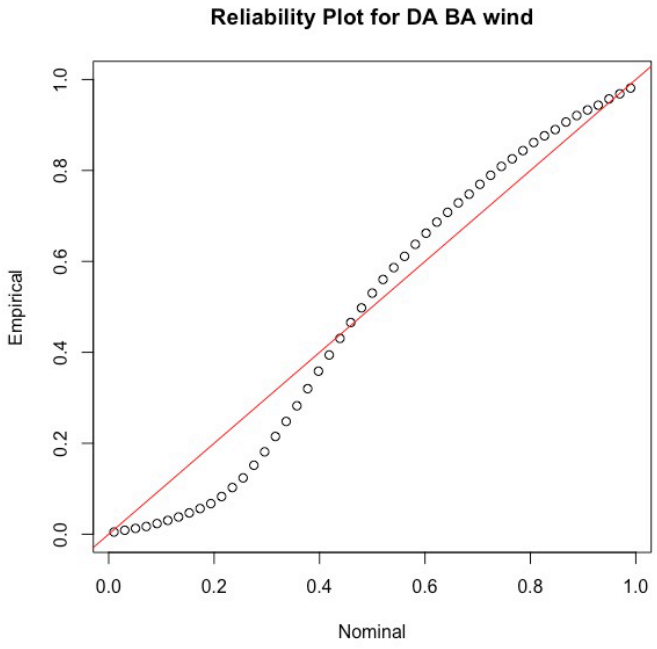

Figure 19: Reliability plot for day ahead wind forecasts for for entire balancing area.

2.5 Load Forecast Data Sets

2.5.1 Deterministic Load Forecasts

Forecasted "air temperature" proved to be a strong predictor of system load and was utilized as the training features at horizon times from ECMWF data. An ensemble of machine learning-generated forecast was developed to produce the point forecasts. Initially, 12 deterministic ensemble members we developed comprising three different machine learning models: a convolutional neural network (CNN), a recurrent neural network (RNN), and an extreme gradient boosting (XGB) model. Hyperparameter optimization was run for each model to find four sets of parameters exhibiting sufficiently low error and high diversity between the ensemble members to be useful. Point forecasts were produced for load at both the zone level and balancing area level (for the entire ERCOT territory). Each machine learning model was scored based on normalized root mean square error (nRMSE) and nMAE, each normalized on the basis of the maximum load for each zone. We also included the mean bias error (MBE) for each forecast, where a negative MBE indicates that predictions are generally lower than actuals, and a positive MBE indicates that predictions are generally higher than actuals. The following sections describe the development of the deterministic load forecasts.

2.5.1.1 Convolutional Neural Network Method

A CNN is a supervised machine learning technique developed originally for images, but it has been shown to work well for time-series problems. A CNN learns a filter that is passed over input features; the power of a CNN is that the learned filter extracts patterns from the features. We then connect the extracted CNN features to a neural network that learns to predict a sequence of data, thus a Deep-CNN. We used 2015 through 2017 weather data (temperature and humidity from 24 weather stations around Texas) and ERCOT electrical load data to train the Deep-CNN on a sequence of future electrical load. We then tested the Deep-CNN on the 2018 ERCOT load data. Ranges of horizon and results are shown in Table 15.

Table 15. The CNN Model Exhibits Low Error for Very Short-Term (1-h) Forecast Lead Time, with Increasing Error for Short- and Medium-Term Forecasts, as Expected. According to the Literature, Forecast Error Expected for Load is approximately 3%.

Model	Horizon	Horizon Range nRMSE Percentage	Horizon Range nMAE Percentage	Horizon Range MBE
CNN	Very short-term	(1.08, 1.65)	(0.79, 1.26)	(-140, 52)
CNN	Short-term	(3.76, 4.01)	(2.89, 3.09)	(-589, 80)
CNN	Medium-term	(3.94, 5.37)	(3.14, 4.09)	(-1871, -901)

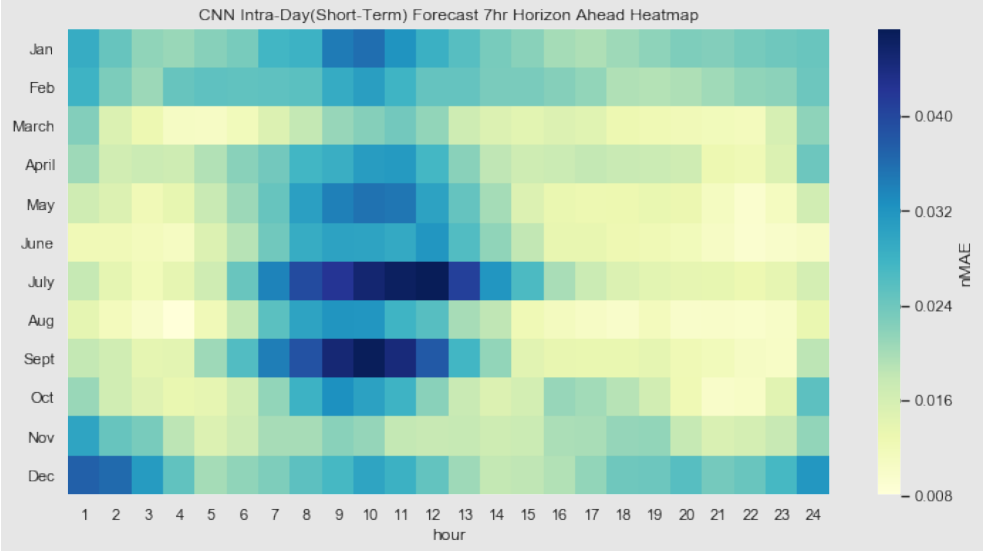

Figure 20. This heat map of errors shows where the model is performing best and where the larger errors occur, with the month on the y-axis, and the hour of the day on the x-axis. As expected, the larger errors appear at times of extreme temperatures, such as in January or during the summer months.

2.5.1.2 Recurrent Neural Network Method

An RNN is a supervised machine learning technique use for sequential data. The input to an RNN is a sequence of features, and the output is a sequence of targets. We used 2015 through 2017 weather data (temperature and humidity from 24 weather stations around Texas) and ERCOT electrical load data to train the RNN on a sequence of future electrical load. We then tested the RNN on the 2018 ERCOT load data. Ranges of horizon and results are shown in Table 16.

Table 16. The RNN Model Exhibits Low Error for Very Short-Term (1-h) Forecast Lead Time, with Increasing Error for Short- and Medium-Term Forecasts, as Expected

Model	Horizon	Horizon Range nRMSE Percentage	Horizon Range nMAE Percentage	Horizon Range MBE
RNN	Very short-term	(2.22, 3.96)	(1.81, 3.27)	(450, 573)
RNN	Short-term	(3.21, 3.33)	(2.59, 2.61)	(-530, -30)
RNN	Medium-term	(2.97, 4.26)	(2.21, 3.13)	(-1791, 1276)

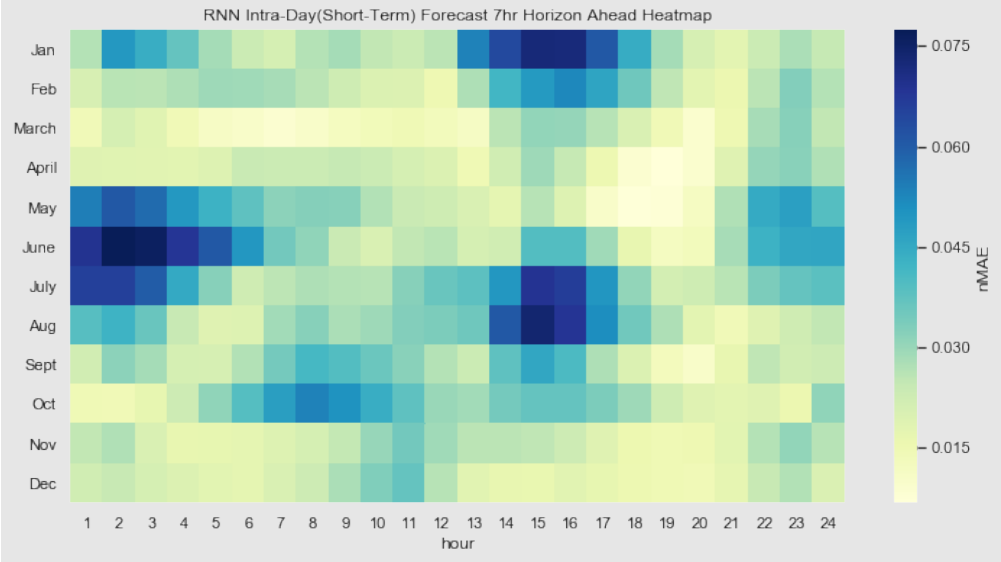

Figure 21. This heat map of errors shows where the model is performing best and where the larger errors occur, with the month on the y-axis, and the hour of the day on the x-axis. As expected, the larger errors appear at times of extreme temperatures, such as in January or during the summer months.

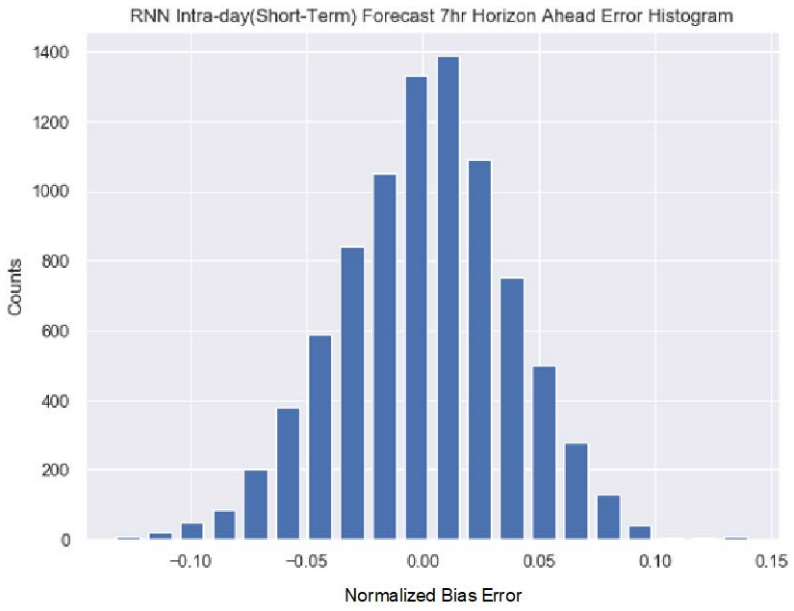

Figure 22. This chart shows that most errors are quite small with a positive bias and that the errors exhibit a normal distribution.

2.5.1.3 Extreme Gradient Boosting Method

A gradient boosting method trains a decision tree in which each observation is given equal weight, then the first tree is evaluated using gradients from the loss function to increase the weights of those observations that are difficult to classify and reduces the weights for observations that are easy to classify. Subsequent trees are grown with these new weights, and so on, until we have many trees. Predictions are then made from a weighted sum of the ensemble. The extreme in XGB comes from computational efficiencies developed to increase learning

speed and accuracy. We used 2015 through 2017 weather data (temperature and humidity from 24 weather stations around Texas) and ERCOT electrical load data to train the XGB on a sequence of future electrical load. We then tested the XGB on the 2018 ERCOT load data. Ranges of horizon and results are shown in Table 17.

Table 17. The XGB Model Exhibits Low Percentage Error for Very Short-Term (1-h) Forecast Lead Time, with Increasing Error for Short- and Medium-Term Forecasts, as Expected. Predictions Tend to Be Lower Than Observations for This Model, as Evidenced by the negative MBE.

Model	Horizon	Horizon Range nRMSE Percentage	Horizon Range nMAE Percentage	Horizon Range MBE
XGB	Very short-term	(1.2, 2.21)	(0.92, 1.65)	(-713.62, -339.65)
XGB	Short-term	(3.68, 3.77)	(2.84, 2.91)	(-1256, -1206)
XGB	Medium-term	(3.81, 5.52)	(2.79, 4.15)	(-1335, -590)

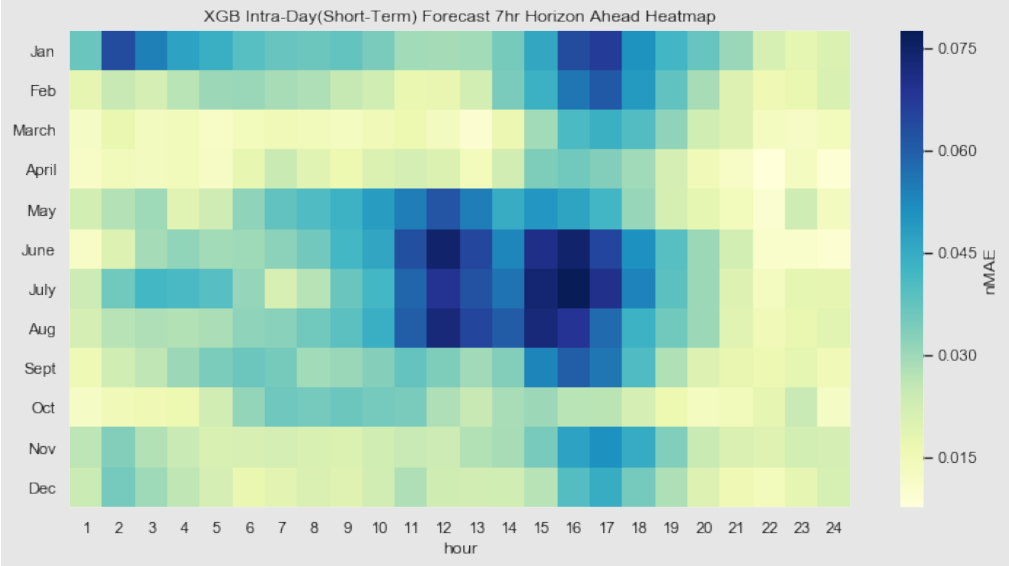

Figure 23. This heat map of errors shows where the model is performing best and where the larger errors occur, with the month on the y-axis, and the hour of the day on the x-axis. As expected, the larger errors appear at times of extreme temperatures, such as in January or during the summer months.

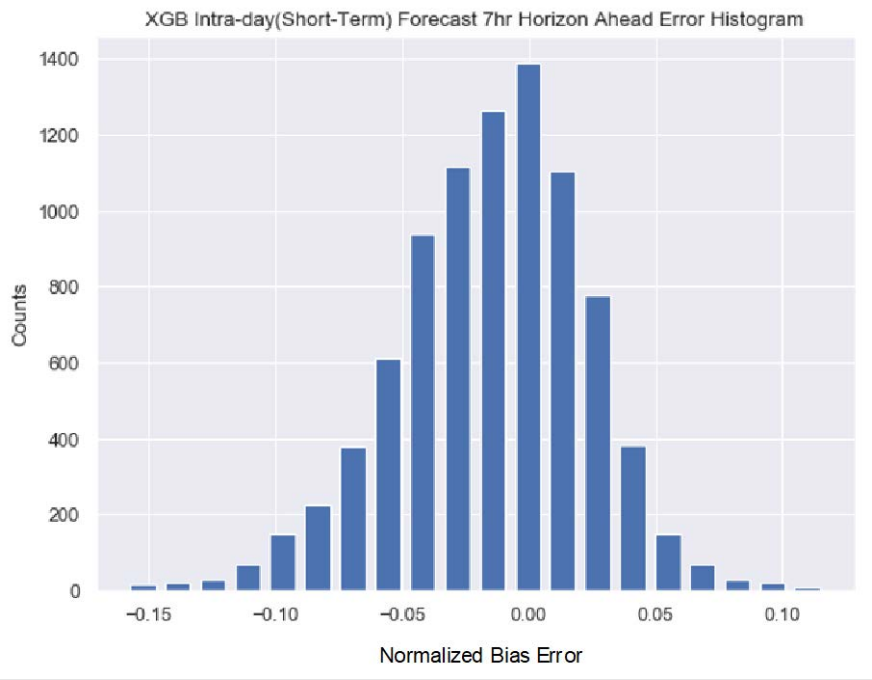

Figure 24. This chart shows that most errors are quite small with a positive bias and that the errors exhibit a normal distribution.

2.5.1.4 Ensemble Method

Table 18. The ERCOT Model Exhibits Low Percentage Error for Very Short-Term (1-h) Forecast Lead Time, with Increasing Error for Short- and Medium-Term Forecasts, as Expected.

Model	Horizon	Horizon Range nRMSE Percentage	Horizon Range nMAE Percentage	Horizon Range MBE
ERCOT	Short-term	(2.92, 3.05)	(2.26, 2.38)	(-67, 115)

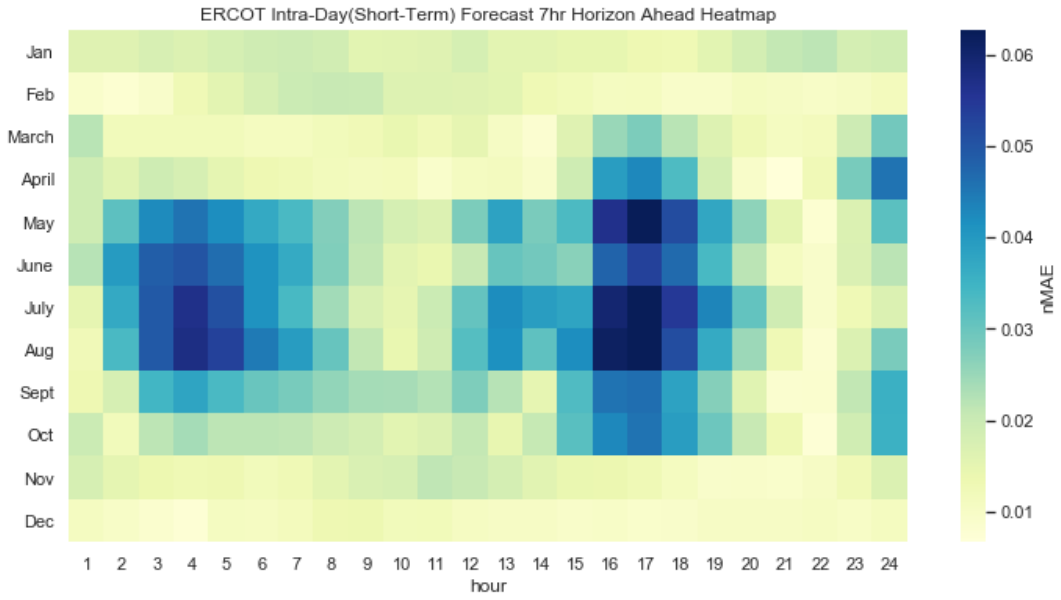

Figure 25. This heat map of errors shows where the model is performing best and where the larger errors occur, with the month on the y-axis, and the hour of the day on the x-axis. The ERCOT model does extremely well in the winter months, but then larger errors occur during the summer months as is to be expected given that air conditioning loads during the hottest days is the largest single variance factor for load.

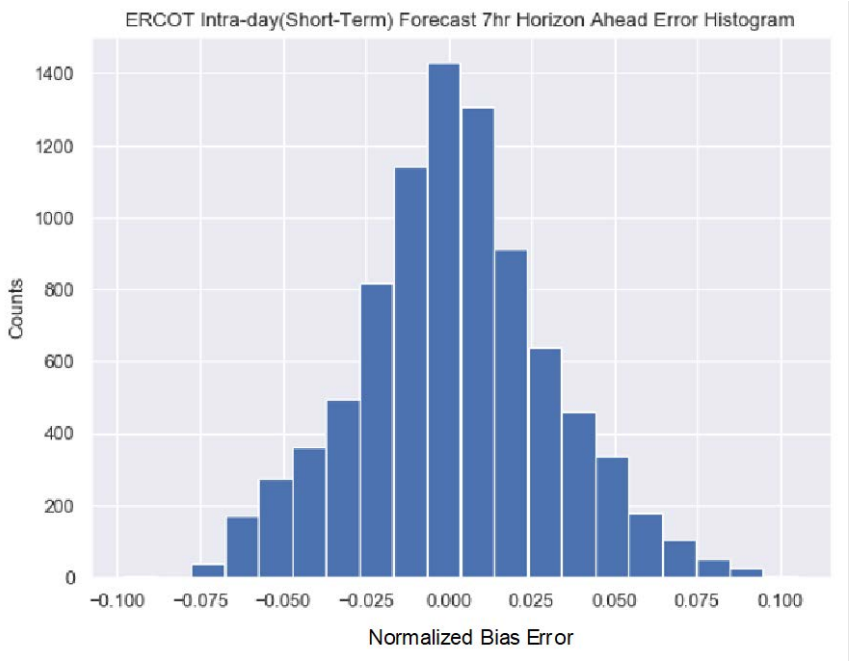

Figure 26. This chart shows that most errors are quite small with a positive bias and that the errors exhibit a distribution that is symmetric about a mean of zero.

2.5.2 2.5.2 Probabilistic Load Forecasts

We produced three sets of probabilistic forecasts—one for each forecast run—using the three methods that we compare in this section. The first two methods are based on the ensemble of machine learning-generated point forecasts, as described in the previous section, by applying a

BMA post-processing method and a Gaussian process method to produce quantiles over respective distributions.

These probabilistic forecasts were scored on the basis of sharpness (how well the point forecasts agree) and reliability (how well the ensemble agrees with the observed values). We applied the CRPS to measure both reliability and sharpness for each probabilistic forecast.

For load probabilistic forecasting, we use the three point forecasts based on CNN, RNN, and XGB as ensemble inputs for the BMA approach. Each forecast will be dressed in a Gaussian model, in which the weight for each member will be adjusted via the Ensemble Method algorithm. Using the deterministic load forecasts without weather features, the performance of the BMA approach is shown in Figure 27, and the evaluation metrics are listed in Table 19.

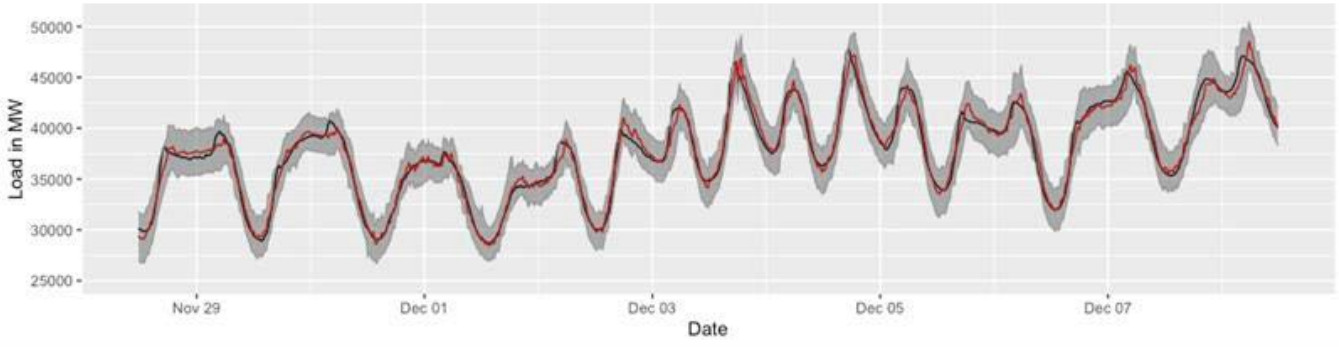

Figure 27. Probabilistic forecasting time series of system-level load using BMA

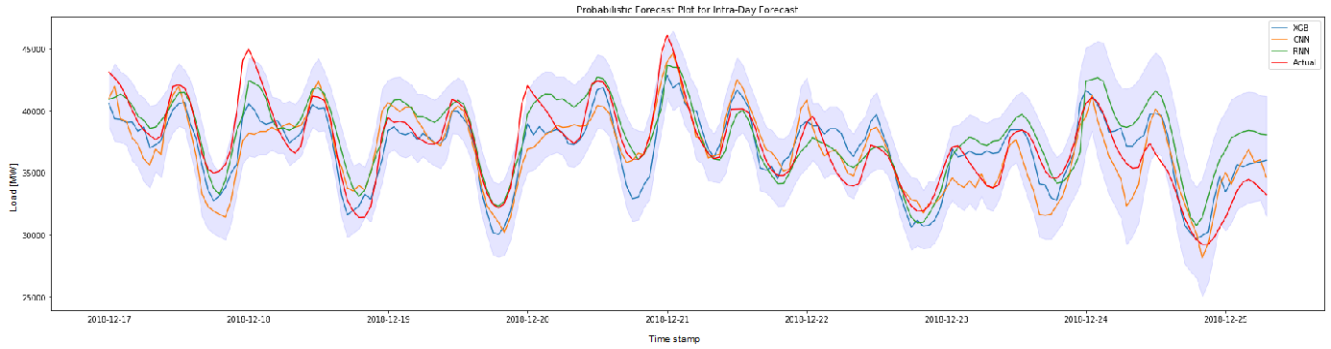

Figure 28. Probabilistic forecasting time series of system-level load using the Gaussian process

Table 19. The CRPS and Error for the Probabilistic Forecast for the Machine Learning Ensemble

Model	*CRPS*	*Mean Forecast nMAE Percentage*	*Mean Forecast nRMSE Percentage*	*MBE*
Ensemble Forecast	477.95	7.2%	9.6%	-51.04

3 Conclusions

This report described the Advanced Research Projects Agency-Energy Performance-Based Energy Resource Feedback, Optimization, and Risk Management (PERFORM) Electric Reliability Council of Texas (ERCOT) dataset consisting of load, solar, and wind deterministic and probabilistic forecasts at three timescales. This dataset consists of 1 year of time-coincident load, wind, and solar actuals and probabilistic forecasts for a region similar to ERCOT.

All the data are stored in Hierarchical Data Format 5 (HDF5) files and have been uploaded to an Amazon Web Services repository[2]. The ERCOT data set has 2 years (2017, 2018) of actuals and 1 year (2018) of probabilistic forecasts. These data are provided at various spatial (i.e., site-level, zone-level, and system-level) and temporal scales (i.e., day-ahead, intraday, and intra-hour). Specifically, data are provided for 125 existing wind sites, 22 existing solar sites, 139 proposed wind sites, and 204 proposed solar sites.

[2] https://registry.opendata.aws/arpa-e-perform/

References

[1] C. Draxl, A. Clifton, B. M. Hodge, and J. McCaa, "The Wind Integration National Dataset (WIND) Toolkit," *Appl. Energy*, vol. 151, pp. 355–366, 2015, doi: 10.1016/j.apenergy.2015.03.121.

[2] A. B. Birchfield, T. Xu, K. M. Gegner, K. S. Shetye, and T. J. Overbye, "Grid structural characteristics as validation criteria for synthetic networks," *IEEE Trans. power Syst.*, vol. 32, no. 4, pp. 3258–3265, 2016.

[3] E. Hale *et al.*, "The Demand-Side Grid (dsgrid) Model Documentation," National Renewable Energy Lab.(NREL), Golden, CO (United States), 2018.

[4] G. Buster, M. Rossol, G. Maclaurin, Y. Xie, and M. Sengupta, "A physical downscaling algorithm for the generation of high-resolution spatiotemporal solar irradiance data," *Sol. Energy*, vol. 216, pp. 508–517, 2021.

[5] P. Gilman, "SAM photovoltaic model technical reference," National Renewable Energy Lab.(NREL), Golden, CO (United States), 2015.

[6] "Europian union emission trading scheme," *Europian union emission trading scheme*. https://www.emissions-euets.com/balancing-energy-gate-closure-time.

[7] "ECMWF Dissemination Schedule." https://www.ecmwf.int/en/forecasts/documentation-and-support/data-delivery/dissemination-schedule.

[8] C. Feng, M. Cui, B. M. Hodge, and J. Zhang, "A data-driven multi-model methodology with deep feature selection for short-term wind forecasting," *Appl. Energy*, vol. 190, pp. 1245–1257, 2017, doi: 10.1016/j.apenergy.2017.01.043.

[9] C. Feng, M. Cui, B.-M. Hodge, S. Lu, H. F. Hamann, and J. Zhang, "Unsupervised clustering-based short-term solar forecasting," *IEEE Trans. Sustain. Energy*, vol. 10, no. 4, pp. 2174–2185, 2018.

[10] C. Feng and J. Zhang, "Assessment of aggregation strategies for machine-learning based short-term load forecasting," *Electr. Power Syst. Res.*, vol. 184, p. 106304, 2020.

[11] K. Doubleday, S. Jascourt, W. Kleiber, and B.-M. Hodge, "Probabilistic solar power forecasting using bayesian model averaging," *IEEE Trans. Sustain. Energy*, vol. 12, no. 1, pp. 325–337, 2020.