21, rue d'Artois, F-75008 PARIS
http : //www.cigre.org

CIGRE US National Committee
2021 Grid of the Future Symposium

Stochastic Analysis of Wind, Solar, and Load across New England

S.L. JUDD
ISO New England, Inc.
USA

C. HAYES
DNV
USA

SUMMARY

To accurately model future system conditions of the power grid, system planners need a robust, detailed historical data set of load and resource performance to develop a range of conditions that can be analyzed for reliability. For resources, conventional generation (nuclear, hydro, fossil, etc.) have been around for decades and sophisticated models have been developed to predict future performance. More recent inverter-based resources (wind and solar) have shorter histories and the technology is rapidly evolving as it matures. Of particular interest for ISO New England (ISO-NE) were the proposed offshore wind facilities in the Federal waters off the coasts of Massachusetts and Rhode Island. In 2019, the only available historical data for offshore wind (hub-height, turbine technology, farm size, and location) was not representative of future proposed projects. To help ISO-NE better understand these new resources and properly model expected future conditions of the offshore wind farms, the ISO engaged DNV to combine their mesoscale modeling capabilities with an innovative stochastic model to quantify the frequency and magnitude of impactful events to system planners, including large wind ramps, high wind turbine shutdown, or unique weather events that may cause resource scarcity, such as extended periods of cold temperatures and low wind generation with coincident periods of high system load. Using the stochastic model, we are able to synthesize thousands of years of realistically plausible time series of energy production while maintaining the spatial coherency and inter-site correlations of key variables such as wind and solar production, ambient temperature and system load. This capability allows for a better projection of future variability as it allowed the team to evaluate the full spectrum of weather conditions that drive power production instead of a simple evaluation of purely historical behavior. The study modeled the wind generation at all existing onshore wind plants and nearly 12 GW of potential offshore wind in the BOEM lease area south of Nantucket. This portion of the study investigated the correlation of onshore and offshore wind generation, solar generation and regional electric load. The primary drivers for the modeled variability and correlation between generation sources and locations are discussed. Analysis of the results indicated that certain meteorological drivers can anti-correlate wind and solar generation at some time scales. We also found that offshore wind generation is only moderately correlated with onshore wind generation.

KEYWORDS

System Reliability, Risk-based Planning, Stochastic, Offshore Wind, Solar, Correlation.

sjudd@iso-ne.com

1 INTRODUCTION

In 2019, ISO New England (ISO-NE), the regional grid operator for the six-state region in the northeastern United States, had three major study initiatives that were analyzing the reliability of the future installation of large amounts of offshore wind turbines off the coasts of Massachusetts and Rhode Island. Three different upcoming studies needed locations of resources in the Bureau of Ocean Energy Management (BOEM) offshore lease areas, time spans from 2012-2018, and power output for the proposed turbines to accurately analyze expected performance. Previous studies sourced offshore wind data from the National Renewable Energy Laboratory (NREL) Wind Toolkit [1]. That data set contained wind speeds for specific locations from 2007-2013. ISO New England did not have the resources or the expertise to create these historical offshore wind profiles, so they teamed up with DNV to first generate historical profiles for 2000-2019 based on weather models and then augment that data set using the DNV stochastic model to create a 20,000 year data set and perform probabilistic analysis on the stochastic data set. This paper will detail the work involved in creating the historical and stochastic data sets and then some of the results of the probabilistic analysis of the data set.

2 CREATION OF WIND, SOLAR AND LOAD DATASET

To quantify the relationships between wind, solar and load across New England required a consistent long-term dataset of coincident onshore and offshore wind generation, regional solar generation and regional load. Previously available datasets of wind, solar and load either had little overlap, were not representative of current proposed offshore wind projects or were too short to help frame future risk. To address this shortfall, the DNV/ISO-NE team used its wind, solar and load modeling expertise to create an initial 20-year dataset of historical hourly onshore and offshore wind generation, regional behind-the-meter solar generation and system load. Following the creation of this initial 20-year dataset, we employed the use of DNV's Stochastic Engine (SE) model to generate the equivalent of 20,000 years of hourly wind, solar, load and weather data. The following sections detail the methodologies used to create these correlated datasets.

2.1 Mesoscale wind flow modeling

The DNV Wind Mapping System (WMS), a dynamical downscaling system based on the Weather Research and Forecasting (WRF) model [2] and a well-validated analog ensemble method [3] [4], was used to generate historical wind speed, wind direction, temperature, pressure and relative humidity time series data for the period 01 January 2000 through 31 December 2019. The model domain extended across all of New England and 200 km offshore and utilized a horizontal resolution of 5 km. NASA's state-of-the-art third generation global reanalysis Modern Era Retrospective-analysis for Research and Applications Version 2 (MERRA-2) [5] [6] was chosen as the input reanalysis dataset for the WRF model.

Hourly time series of wind speed, wind direction, temperature and pressure data at the locations of 38 existing wind plants and 12 hypothetical/potential offshore plants were extracted from the WMS model output. The locations of the wind plants are shown in Figure 2-1.

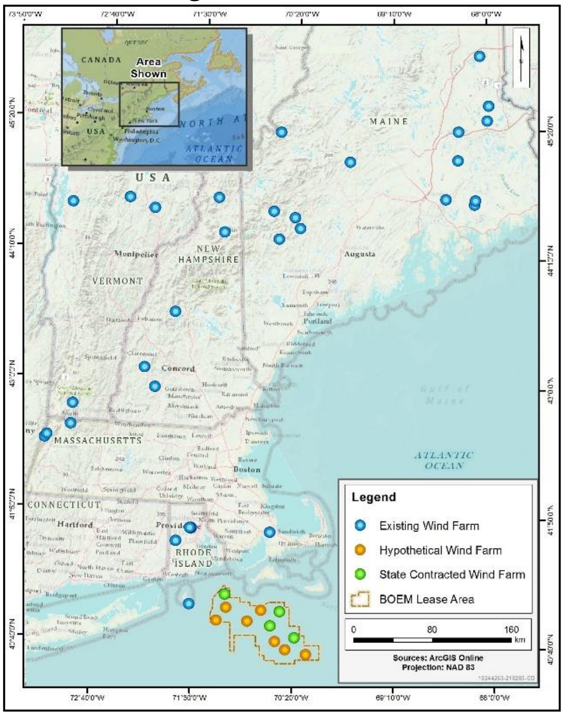


Figure 2-1 Locations of existing, state contracted and hypothetical wind plants

In order to better match the 5-km resolution modeled wind speed data to the observed wind speeds at the wind plants, the modeled hourly wind speed data were calibrated using measurements collected at or near the wind plant sites. Available wind speed measurements were cleaned to remove instances of erroneous or suspicious values. The calibration procedure required there be at least one year of valid measurements. Additional information on data cleaning and calibration can be found in [7].

2.2 Wind Power time series modeling

Following the mesoscale wind speed calibration procedure, we performed power time series modeling by applying wind plant specific power models, obtained from the DNV real-time forecast service and trained by power measurements, to the calibrated wind speed time series for the existing wind plants.

Wind plant specific layouts and turbine models for the potential offshore wind plants were not publicly available. Therefore, we created hypothetical turbine layouts with a minimum turbine spacing of one nautical mile [8] and hub heights ranging from 119 m to 150 m to meet the desired wind plant capacities, presented in [7] and [9], for each of the 12 hypothetical wind plants. Power modeling of the hypothetical offshore wind plants was performed using the DNV WindFarmer software to appropriately capture the internal and external wake interactions and their impact on power production for each proposed layout. Information on the specific offshore turbine models and projects layouts, including the four wind plants selected for long-term contracts, can be found in [7].

An estimated electrical efficiency value of 97.5% was chosen for all wind plants as this is a standard value assumed during most preconstruction energy production estimates. A wind plant availability value of 94.5% was used for the existing wind plants. This value accounts for the expected turbine availability, balance of plant and grid availability and was chosen based on our review of pre-construction estimates throughout New England. For the hypothetical offshore wind plants an availability value of 93.0% was used. This offshore preconstruction availability assumption is based on DNV's review of offshore project availability and their global offshore experience.

The time series availability model uses a Poisson distribution where groups of turbines become unavailable for several consecutive timesteps until they come back online. This allows for the creation of a more realistic time series with downtime events that last for several hours or days. These downtime events are applied randomly throughout the time series such that the sum of the production time series with availability applied will be 94.5% of the sum without availability for the onshore plants and 93.0% for offshore wind plants.

There are several loss factors often considered during preconstruction energy assessments that have not been applied to the power time series. These include Turbine Icing Losses, Environmental Losses, Turbine Performance Losses, and Curtailments.

2.3 Behind-The-Meter Solar Capacity Time Series

Historical behind-the-meter (BTM) solar generation for each of the 8 ISO New England Load Zones needed to be coincident to the modeled wind generation and scalable to any future BTM solar capacity scenario.

The geographic extent, varied weather regimes (coastal vs inland), terrain variation and population center locations across each Load Zone made the use of a single time series of solar irradiance inappropriate to characterize the average solar generation profile for the entire Load Zone. To properly account for the variability in solar resource and BTM solar capacity we first selected 20 representative locations across each Load Zone. Hourly solar irradiance and weather data for the 2000 through 2019 period were created for all locations using DNV's SunSpot satellite irradiance product.

DNV SunSpot was created using a semi-empirical satellite-to-irradiance conversion model based on an extended version of the Perez model utilizing infrared and visible imagery from NASA's Geostationary Operational Environmental Satellites (GOES). SunSpot irradiance data are produced on a high-resolution 1-km grid and can resolve cloud and local weather impacts on the surface radiation budget.

To convert the solar irradiance time series to energy we developed a solar power conversion model using open source algorithms published by Sandia National Labs [10]. These algorithms utilize labmeasured efficiency curves for actual panels and inverters selected as representative for the available choices. Calculations for losses due to temperature, wire resistance, inverter clipping, and other factors are also included.

The irradiance and weather time series at each location were passed through the solar PV model to provide an estimate of how much solar energy would be available after accounting for the estimated solar capacity, equipment mounting orientation, equipment selection, and soiling due to dust and snow coverage on panels. The result is an hourly PV production time series for each location, which includes consideration of project location, panel tilt angle, azimuth angle, module type, inverter type and mounting system.

The 20 solar-PV time series for each Load Zone were normalized and then weighted together based on their modeled capacity. The resulting time series represented the normalized BTM solar generation for the Load Zone.

Hourly behind-the-meter regional solar generation capacity measurements were used to calibrate the 20-years of modeled regional solar capacity time series. Comparisons of calibrated and measured data indicated good agreement with an average R-squared value of 0.97.

2.4 Load Time Series Extension

Load observations were available for the period March 2003 through December 2019 [11]. The real-time load observations include the impacts of energy efficiency and behind-the-meter (BTM) solar. For the purposes of this study measured hourly BTM solar generation [12] was removed from the real-time load observations to create a "gross load" time series where gross load is defined as total consumption of load minus energy efficiency with BTM solar reconstituted. Load modeling techniques were used to extend the gross load data back to January 2000 for each Load Zone.

2.5 Stochastic Time Series Modeling

A next-generation probabilistic model was developed to provide a comprehensive assessment of both typical and unusual wind, solar and load scenarios across the portfolio [13]. The DNV Stochastic Engine (SE) model employs a non-parametric bootstrap resampling method to generate synthetic sequences of time series data based on the historical record. It uses historical observation trajectories as multivariate "dependence templates" onto which the stochastic realizations are assembled to duplicate the pairwise rank correlation structure in the historical record. Like other empirical copula methods, the dependence templates used in the SE contain all the information about the inter-site and inter-variable dependencies within the data (both linear and non-linear dependencies). The daily, seasonal, and annual cycles of the original dataset are fully preserved, along with the spatial coherency of weather, wind and solar generation, and load across the entire portfolio of sites (inter-site correlations). The preservation of these inter-site correlations is very important for understanding the relationship of projects across a region. The true value of the SE is its ability to represent the full spectrum of possible weather conditions that drive wind and solar power production, thereby allowing a comprehensive assessment of all possible scenarios across the entire portfolio. Each individual synthetic time series closely mimics the characteristics of the weather that could occur at each project location, based on the historical record. The output is a synthetic series of hourly data with the same statistical properties as the observations. Relevant examples include but are not limited to understanding risks associated with low generation and high demand, revenue risk where time of day or seasonal characteristic are important and robust probabilistic estimates of events, such as low wind years, high-wind shutdown events, large wind ramp events or periods of resource constraints.

The SE was used to model 1,000 realistically plausible historical 20-year time series of hourly wind generation for each wind plant, and BTM solar photovoltaic (PV) generation and load for each Load Zone, to capture the full range of meteorological conditions that can occur across the ISO-NE service area [9]. This allows quantification of the variability in the wind and solar resources, and any associated risks of underproduction or rare weather events. The input dataset of wind, solar and load covering January 2000 through December 2019, as described above, was used to initialize the SE model. This amounts to the equivalent of 20,000 years (20 years × 1,000 synthetic sequences) of hourly wind and solar production data, load, temperature, relative humidity, wind speed, and solar irradiance.

The SE preserves all trends present in the original dataset. Figure 2-2 presents the distribution of monthly mean gross load values calculated from the 20,000-year stochastic dataset. The original 20-years of input gross load data exhibited a downward trend due to the implementation of energy efficiency programs in recent years. The stochastic dataset preserves this trend.

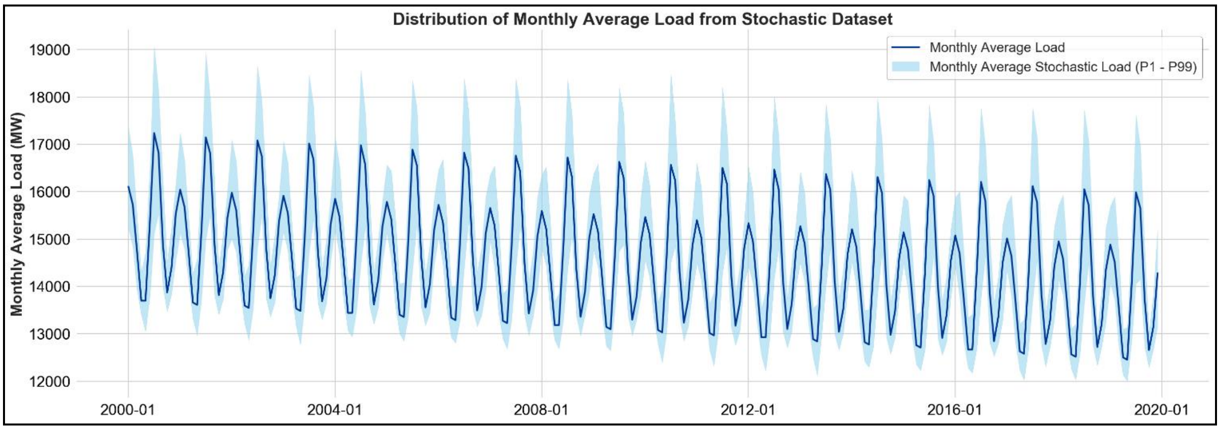


Figure 2-2 Monthly average gross load trend in stochastic dataset

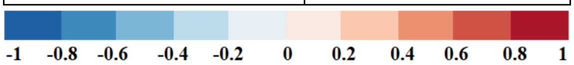
A robust correlation analysis has been performed using the stochastic dataset. The results are described below.

3 CORRELATION OF WIND, SOLAR AND LOAD DATA

The wind, solar, and load time series for each stochastic realization were aggregated for each region and correlation analysis performed to quantify the strength of linear association between the datasets. Pearson correlation coefficients were computed based on hourly, monthly, and annual values. A Pearson correlation coefficient measures the statistical relationship between two continuous variables, and is based on the method of covariance. It gives information about both the magnitude and direction of correlation between two variables. A negative correlation coefficient means the two variables are negatively correlated (one goes up while the other goes down), while a positive coefficient indicates the variables are positively correlated (both go up or down at the same time). Table 3-1 presents a description of the degree of correlation quality as it relates to the Pearson coefficient [14].

Table 3-1 Quality of correlation for Pearson coefficient ranges

Quality of correlation	Pearson coefficient
Perfect	±1.00
High	±0.50 to ±0.99
Moderate	±0.30 to ±0.49
Low	±0.01 to ±0.29
None	0.00



Load, wind, and solar generation within each Load Zone were correlated to determine the strength of linear association. These correlations are presented in Table 3-2 to Table 3-4 for hourly, monthly, and annual data. Note that CT and NEMA do not have any modeled wind generation. Also, RI and SEMA wind generation includes the combined onshore and offshore generation associated with each Load Zone. Based on these results, there is a lack of correlation between load and wind generation, whereas a weak positive correlation exists between load and solar generation. Hourly wind and solar generation exhibit a very weak negative correlation, and this is likely due to lower average wind generation during the afternoon when solar generation is at its peak. The opposite relationship exists during the early morning and evening hours, when solar generation is at its minimum.

The positive net load to solar correlations for NH and NEMA indicate that more solar could be added to these Load Zones that would correlate with existing load while the moderate to strong negative net load to solar correlations for VT, SEMA, and WCMA indicates an abundance of BTM solar as compared to those individual zones' load.

Table 3-2 Pearson correlation coefficients of hourly data between load, wind, and solar generation within each Load Zone

	ME	NH	VT	CT	RI	SEMA	WCMA	NEMA	ISO-NE
Load - Wind	-0.01	0.00	0.01		-0.03	-0.03	-0.04		-0.02
Net Load - Wind	0.05	0.02	0.22		0.09	0.11	0.14		0.10
Solar - Wind	-0.12	-0.08	-0.24		-0.18	-0.16	-0.19		-0.18
Load - Solar	0.36	0.37	0.30	0.36	0.38	0.36	0.36	0.39	0.39
Net Load - Solar	-0.20	0.14	-0.66	-0.03	-0.32	-0.54	-0.65	0.08	-0.23
Load - Solar+Wind	0.18	0.31	0.31		0.18	0.13	0.36		0.34
Net Load - Solar+Wind	-0.06	0.13	-0.62		-0.09	-0.14	-0.65		-0.15
Load - Offshore Wind									-0.00
Net Load - Offshore Wind									0.10
Solar - Offshore Wind									-0.17
Onshore Wind - Offshore Wind									0.51

Monthly wind and solar generation exhibit a moderate to strong negative correlation within each Load Zone, which again is due to the seasonal variations in wind and solar generation. Solar generation is highest during the summer months when wind speeds tend to be lower, whereas during winter wind generation tends to be higher and solar generation lower (solar insolation is greatly decreased during winter).

Table 3-3 Pearson correlation coefficients of monthly average data between load, wind, and solar within each Load Zone

	ME	NH	VT	CT	RI	SEMA	WCMA	NEMA	ISO-NE
Load - Wind	0.07	-0.08	0.32		-0.42	-0.43	-0.08		-0.24
Net Load - Wind	0.28	0.01	0.59		-0.24	-0.15	0.33		-0.03
Solar - Wind	-0.73	-0.66	-0.78		-0.66	-0.65	-0.80		-0.70
Load - Solar	-0.24	-0.08	-0.51	0.08	0.25	0.23	-0.03	0.13	0.06
Net Load - Solar	-0.51	-0.22	-0.81	-0.09	-0.03	-0.20	-0.53	-0.03	-0.24
Load - Solar+Wind	-0.02	-0.19	-0.52		-0.40	-0.43	-0.04		-0.26
Net Load - Solar+Wind	0.13	-0.21	-0.80		-0.31	-0.26	-0.53		-0.33
Load - Offshore Wind									-0.25
Net Load - Offshore Wind									-0.05
Solar - Offshore Wind									-0.66
Onshore Wind - Offshore Wind									0.88

Annual wind and solar generation within VT and SEMA have a moderate negative correlation, such that years with lower average wind generation tend to have higher than average solar generation. One potential explanation is that "sunnier" than normal years over New England correspond to frequent and persistent high-pressure systems. High-pressure systems are typified by clear and calm conditions, with suppressed wind speeds. There does not appear to be a meaningful relationship between wind and solar generation for NH, ME, and WCMA as indicated by the near zero correlation coefficients.

Table 3-4 Pearson correlation coefficients of annual average data between load, wind, and solar within each Load Zone

	ME	NH	VT	CT	RI	SEMA	WCMA	NEMA	ISO-NE
Load - Wind	0.42	0.21	-0.02		0.02	0.01	0.12		0.09
Net Load - Wind	0.42	0.22	0.01		0.03	0.04	0.12		0.10
Solar - Wind	-0.05	0.01	-0.18		-0.09	-0.21	-0.07		-0.11
Load - Solar	0.05	0.37	-0.16	0.20	0.40	0.45	0.16	0.36	0.24
Net Load - Solar	-0.02	0.32	-0.28	0.17	0.30	0.35	0.06	0.33	0.18
Load - Solar+Wind	0.42	0.32	-0.16		0.08	0.06	0.17		0.16
Net Load - Solar+Wind	0.42	0.31	-0.25		0.07	0.07	0.08		0.15
Load - Offshore Wind									0.06
Net Load - Offshore Wind									0.07
Solar - Offshore Wind									-0.18
Onshore Wind - Offshore Wind									0.63

Annual load and solar generation exhibit a moderate positive correlation for CT, NH, RI, SEMA, and NEMA, which indicates that years with high load tend to correspond to higher than average solar generation. One potential driver is that during years with higher solar generation more days with mostly clear skies occur. As a result, temperatures are slightly higher, which in turn drives load slightly higher.

The team also examined the linear relationship between onshore and offshore (existing and state-contracted) wind generation by performing a correlation analysis for hourly, monthly, and annual average generation capacity. Results in Table 3-5 indicate moderate correlation for hourly and annual records and a strong correlation for monthly generation. When compared with the monthly correlations, the weaker correlation for annual values likely represents regionally specific interannual variations in the wind resource and wind generation. Interestingly, the hourly correlation appears to be slightly weaker between the offshore and onshore aggregate wind generation than the correlation of wind generation between Load Zones.

Table 3-5 Pearson correlation coefficients of hourly, monthly, and annual onshore and offshore wind generation

	Hourly	Monthly	Annual
Offshore - Onshore Wind	0.51	0.88	0.63

Understanding the relationship of wind, solar and load across Load Zones can be important for system balancing and resource allocation. The analysis below investigates the correlation of both gross load and net load across Load Zones as well as wind and BTM solar generation. Note that gross load is defined as total consumption of gross load minus energy efficiency with BTM solar reconstituted. Net load is defined as gross load minus energy efficiency and BTM solar generation.

The correlation of hourly gross and net load across each Load Zone is shown in Figure 3-1. As expected, hourly gross load is highly correlated across Load Zones; however, due to differing amounts of BTM solar generation in each region, the net load is not as well correlated across all zones. This is likely because on an hourly basis, net load for some Load Zones such as VT can be negative due to solar generation being greater than gross load.

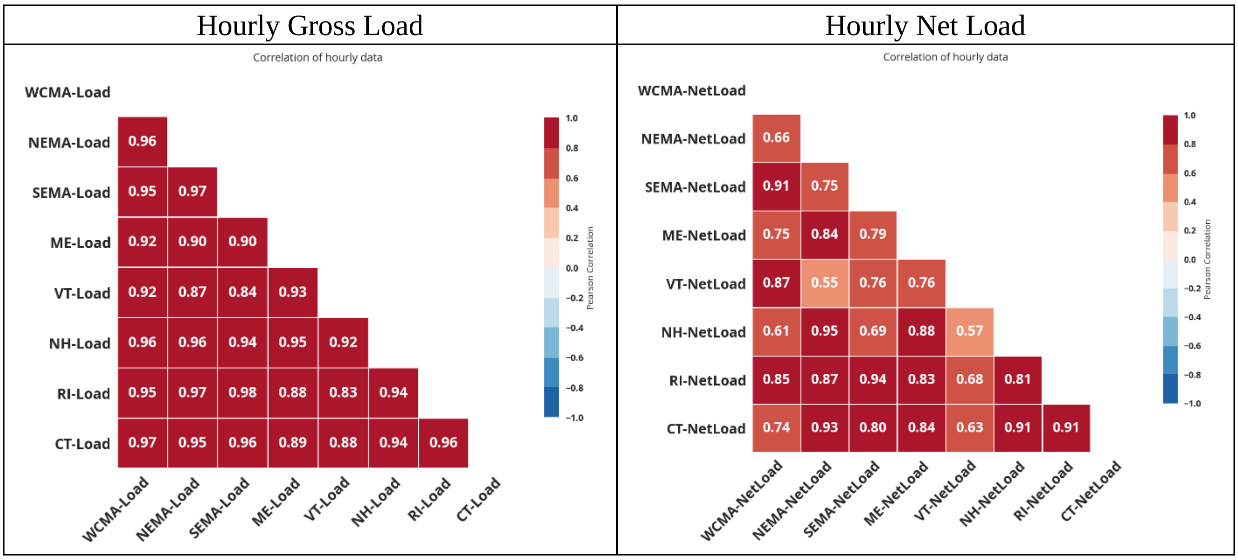


Figure 3-1 Correlation of hourly gross load and net load across Load Zones

Next, a correlation of wind and solar generation across the different Load Zones was performed. Note that there is no existing or planned wind generation in CT and NEMA. Wind generation for RI includes both existing onshore and offshore generation as well as future state-contracted offshore generation. Similarly, SEMA wind generation includes the existing onshore wind farms as well as the future state contracted offshore wind farms. As a result, both RI and SEMA wind generation are heavily weighted by the future state-contracted offshore generation. On an hourly basis there is a moderate to high correlation between wind generation for each Load Zone, as shown in Figure 3-2. Hourly solar generation exhibits a much higher correlation between load zones when compared with wind generation, which tends to be more impacted by local terrain and surface roughness impacts. Adjacent zones, such as ME and NH, and RI and SEMA, have higher Pearson coefficients than zones that are geographically separated, such as ME and WCMA.

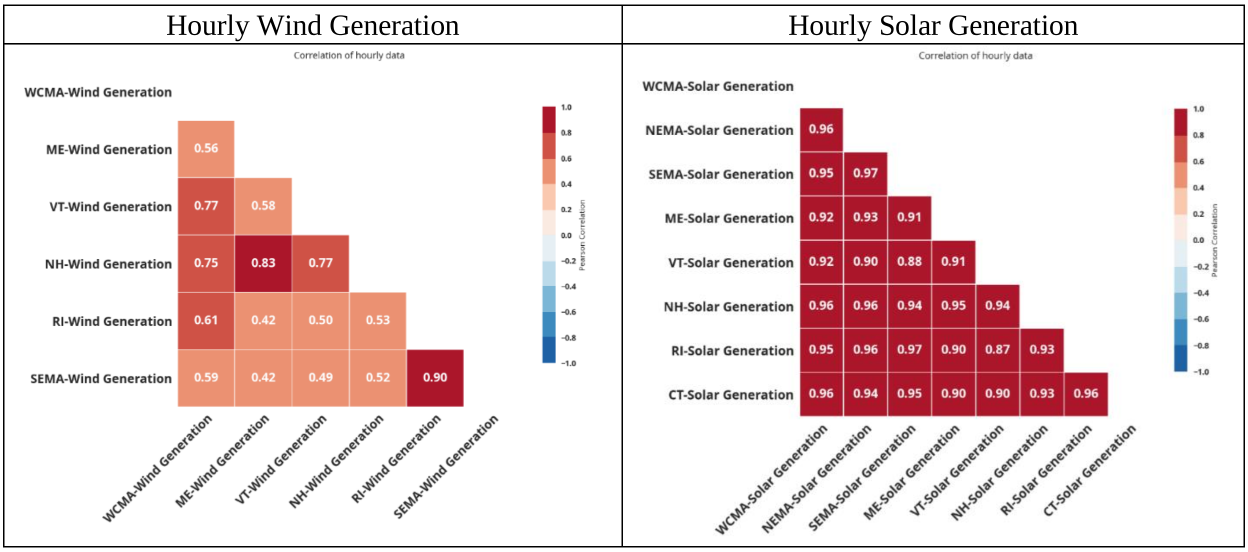


Figure 3-2 Correlation of hourly wind and solar generation across Load Zones

4 CONCLUSION

To accurately plan for the clean energy transition, system planners need technically sound, historical wind data to apply to planned wind resources, specifically offshore wind. Since historical data did not exist for resources that have not been built, ISO New England engaged DNV to develop 'historical' resource time series profiles based on available weather data and then vastly expand that data set using DNV's Stochastic Engine model to perform probabilistic analysis. Correlation of the wind, solar, and load data sets discovered some interesting results. Hourly correlations highlighted wind and solar were negatively correlated, which shows the benefit of a diversity of variable energy resources. It also revealed the NH and NEMA Load Zones in New England that may be potentially preferable for additional development of solar resources due to remaining load that still has a positive correlation. This analysis did not examine the effect of energy storage systems to increase the correlation of solar and load. Another interesting outcome was the weak-to-no-correlation of load and offshore wind in the different time frames. This was due to the high variability of offshore wind output during any given hour of the day. This new set of data will aid in planning a reliable system as we transition to cleaner, more variable resources.

5 BIBLIOGRAPHY

[1] National Renewable Energy Laboratory, "Wind Integration National Dataset Toolkit," [Online]. Available: https://www.nrel.gov/grid/wind-toolkit.html. [Accessed 2021].

[2] "Weather Research and Forecasting Model," UCAR, [Online]. Available: https://www.mmm.ucar.edu/weather-research-and-forecasting-model.

[3] L. T. L. Delle Monache, "A new and computationally efficient technique for quantifying uncertainty in a downscaled wind resource dataset.," in *AWEA Wind Resource Energy Assessment Seminar*, Pitssburgh, 2021.

[4] L. T. L. e. a. Delle Monache, "Probabilistic weather prediction with an analog ensemble," *Monthly Weather Review,* p. 3498–3516, 2013.

[5] NASA, "National Aeronautics and Space Administration, MERRA-2," 1980-present. [Online]. Available: http://disc.sci.gsfc.nasa.gov/mdisc/.

[6] e. a. Ronald Gelaro, "MERRA-2 Overview: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2)," *Journal of Climate,* p. 5419–5454, 2017.

[7] DNV GL, "10176388 HOU-T-01-C ISONE Power Time Series," DNV GL, 2020.

[8] E. V. W. e. M. Orsted, "RE: Proposal for a uniform 1x1 nm wind turbine layout for New England Offshore Wind," 2019.

[9] DNV, "10244263-HOU-T-01-F Analysis of Stochastic Dataset_ISO-NE," DNV, 2021.

[10] C. W. H. a. M. A. M. William F. Holmgren, "pvlib python: a python package for modeling solar energy systems.," *Journal of Open Source Software, https://doi.org/10.21105/joss.00884,* 2018.

[11] ISO-NE, "ISO-NE Energy, Load and Demand Reports," ISO-NE, 2021. [Online]. Available: https://www.iso-ne.com/isoexpress/web/reports/load-and-demand/-/tree/zone-info. [Accessed 2021].

[12] ISO-NE, "ISO-NE Behind-The-Meter PV data," ISO-NE, [Online]. Available: https://www.iso-ne.com/static-assets/documents/2020/07/btm_pv_data.xlsx .

[13] DNV GL, *DNV GL Renewables Energy Variability Assessment Flyer*, DNV GL, 2021.

[14] "Statistics Solutions," [Online]. Available: https://www.statisticssolutions.com/pearsons-correlation-coefficient/. [Accessed 2021].

[15] ISO-NE, "FCA14 Qualification Data for Q944," ISO-NE Sharefile, 21 November 2019.

[16] S. I.-N. Judd, "PV Data for 2020, Email Received on 02/25/2021," 2021.