AIRU-WRF: AI-powered Physics-based Tool for Offshore Wind Forecasting & Grid Integration

NOWRDC Project #133/192900

Final Report

Prepared for:

Melanie Schultz, Program Manager

National Offshore Wind Research and Development Consortium

Prepared by:

Ahmed Aziz Ezzat[1], Yazhou "Leo" Jiang[2], Travis Miles[1]

[1]Rutgers, The State University of New Jersey, Piscataway, NJ

[2]Clarkson University, Potsdam, NY

Project Manager: Ahmed Aziz Ezzat, Ph.D.
Assistant Professor of Industrial & Systems Engineering,
Rutgers, The State University of New Jersey, Piscataway, NJ

Table of Contents

List of Figures

List of Tables

List of Acronyms

Acronym	Description
NOWRDC	National Offshore Wind Research & Development Consortium
AIRU-WRF	Artificial Intelligence-Powered Rutgers University Weather Research & Forecasting
OSW	Offshore Wind
AI	Artificial Intelligence
ML	Machine Learning
NWP	Numerical Weather Prediction
NYSERDA	New York State Energy Research & Development Agency
ASOW	Atlantic Shores Offshore Wind
E05N	NYSERDA's floating buoy at [39.97, -72.72]
E05S	NYSERDA's floating buoy at [39.48, -73.59]
E06	NYSERDA's floating buoy at [39.55, -73.43]
ASOW	Atlantic Shores Offshore Wind
ASOW-1	ASOW's floating buoy at [39.31, -74.11]
ASOW-4	ASOW's floating buoy at [39.20, -74.08]
ASOW-6	ASOW's floating buoy at [39.27, -73.89]
RUCOOL	The Rutgers University Center for Ocean Observing Leadership
WRF	Weather Research & Forecasting model
RU-WRF	Rutgers University Weather Research & Forecasting model
NY	New York
NJ	New Jersey
NJBPU	New Jersey Board of Public Utilities
GP	Gaussian Process
POE	Probability of Exceedance
CDF	Cumulative Density Function
LSTM	Long Short-Term Memory Model
ARIMA	Autoregressive Integrated Moving Average
MAE	Mean Absolute Error
CRPS	Continuous Ranked Probability Score
PI	Predictive Interval
MW	Megawatt
GW	Gigawatt
NYISO	New York Independent System Operator
NYCA	New York Control Area
IPP	Independent Power Producer
SCR	Solvency Capital Requirement
EUE	Expected Unserved Energy
LOLP	Loss of Load Probability
LOLH	Loss of Load Hours

Summary

This is the final report of Project #133/192900, titled: "*AIRU-WRF: AI-powered Physics-based Tool for Offshore Wind Forecasting & Grid Integration*." The project started on 09/2023 and officially ended in 07/2025. The project was led by Rutgers, The State University of New Jersey (the lead contractor) and included Clarkson University and the Electric Power Research Institute (EPRI) as sub-contractors.

The report comprises seven main sections. Section 1 provides a brief background about the objectives and scope of the project. This is followed by Section 2 where the extraction, processing, and analysis of offshore wind datasets are described. Sections 3 and 4 discuss the developed models, as well as the key findings and results of the project. Section 5 concludes the report and highlights plans for future research and development. Section 6 presents the efforts made by the project team to disseminate the results of this project. Finally, Section 7 includes the relevant references.

1 Background and Problem Statement

This NOWRDC project aimed to develop and demonstrate a wind forecasting technology, called AIRU-WRF, to predict, for short-term horizons (sub-hourly to a day-ahead), the offshore wind (OSW) resource and power production at key offshore wind energy areas along the U.S. East Coast. AIRU-WRF is a hybrid forecasting technology which combines state-of-the-art physics-based modeling with probabilistic machine learning (ML) capabilities. The motivation of this project is the lack of customized forecasting products that are specifically tailored to the offshore wind energy areas along the U.S. East Coast. Accurate short-term forecasts will be necessary for the reliable and cost-effective operation of future OSW farms in this geographical region, as well for the integration of OSW energy into the power grid.

Towards that, AIRU-WRF (short for the AI-powered Rutgers University Weather Research & Forecasting model) was developed to combine region-specific numerical weather predictions (NWPs) with local measurements to make wind speed and power forecasts that are short-term (sub-hourly to a day-ahead), and of high resolution, both spatially (site- or farm-level) and temporally (sub-hourly). Such high-resolution forecasts—practically unattainable by stand-alone physics-based models—are key to support the reliable and cost-effective operational decision-making in OSW farms and power grids, including resource and production estimation, economic dispatch and reserve procurement, and day-to-day OSW farm operations. The project is especially motivated by the lack of demonstrated OSW forecasting technologies that are tailored to the OSW energy areas along the U.S. East Coast.

Obtaining high-resolution, short-term wind forecasts is challenging, mainly due to the limitations of meso-scale NWPs at fine spatial and temporal scales. Hence, this project aims to unlock the potential of ML technologies to complement NWPs in OSW forecasting, with the following two objectives in mind:

A. To develop an ML-enhanced OSW forecasting technology which combines region-specific NWPs with ML tools for superior, high-resolution, short-term OSW resource and power forecasting.

B. To extensively evaluate the technology on two main fronts: (*i*) forecast quality (i.e., how accurate are the forecasts compared to benchmark models in the wind forecasting landscape); and (*ii*) forecast value (i.e., what is the added economic and reliability value of using these forecasts in grid operation and offshore wind integration).

2 Data Extraction and Analysis

High-quality forecasts require access to high-fidelity measurements at multiple spatial sites, for sufficiently long periods of time, and at altitudes that are relevant to the hub height of modern-day wind turbines. The proposed technology makes use of two sources of data, described in Sections 2.1 and 2.2., respectively.

2.1 Local Measurements from the U.S. East Coast

Hub-height wind observations that are publicly available have been extracted from six buoys in the U.S. Mid/North Atlantic region. The details and data coverage of these datasets are shown in Table 1.

Table 1: Publicly available data identified by the project team for AIRU-WRF model development. NYSERDA = New York State Energy Research & Development Authority. ASOW = Atlantic Shores Offshore Wind.

Buoy	Start Date	End Date	Latitude	Longitude	Resolution	Owner
E05 North (E05N)	08/2019	09/2021	39.97	-72.72	10-min	NYSERDA
E05 South (E05)	01/2022	01/2023	39.48	-73.59	10-min	NYSERDA
E06	09/2019	03/2022	39.55	-73.43	10-min	NYSERDA
ASOW-1	10/2021	09/2023	39.31	-74.11	10-min	ASOW
ASOW-4	05/2021	06/2022	39.20	-74.08	10-min	ASOW
ASOW-6	02/2020	05/2021	39.27	-73.88	10-min	ASOW

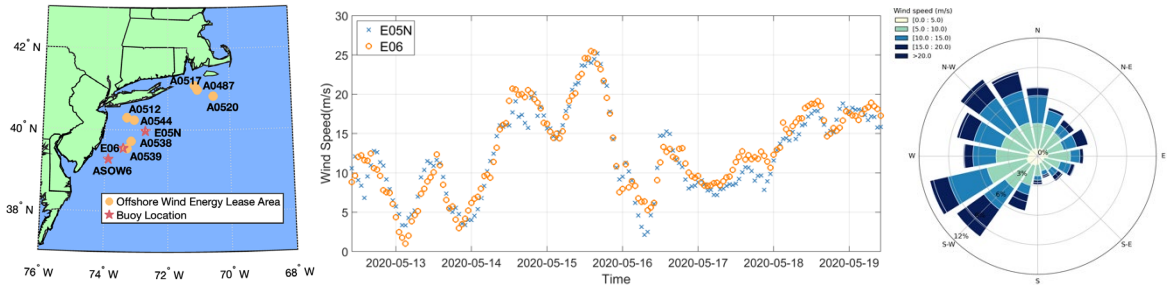

Figure 1: Left: Geographical map of the buoys surveyed, and OSW energy areas considered. Middle: A week-worth of hub-height wind speeds at locations E05N and E06. Right: Spatially averaged wind rose plot of hub-height winds.

Figure 1 provides an overview and demonstration of the data used in this work. The left panel shows a geographical map of the buoys surveyed (some buoys omitted in the figure due to spatial proximity), on top of the OSW energy areas considered. The middle panel shows a week-worth of hub-height wind speed measurements at two buoy locations (E05N and E06), suggesting strong spatial dependencies between the hub-height wind speeds at different sites. This strong spatial correlation will be explicitly modeled and leveraged in AIRU-WRF for improved forecasting performance. The wind rose plot in the right panel suggests that the prevailing winds are mostly westerly in this geographical region. Season-specific rose plots (not shown here) further reveal distinct seasonal variations in the magnitude and direction of wind conditions. Both NYSERDA and ASOW datasets are available publicly and careful data processing has been carried out to fill out missing values and/or correct anomalous measurements [1, 2].

2.2 Numerical Weather Predictions from RU-WRF:

The Rutgers University Center for Ocean Observing Leadership (RUCOOL) has run a daily real-time version of WRF 4.1., called RU-WRF, which is tailored to the meteorological phenomena along the U.S. East Coast and across the NY/NJ Bight. RU-WRF runs a parent nest at 9 km resolution out to 120 hours, and a child nest centered on the NJ shelf at 3 km resolution out to 48 hours, generating hourly forecasts of multiple meteorological variables. The model data archived from December 2019 to March 2023 includes hourly output compiled from the 3-km domain model. A 1-km resolution is run occasionally for scientific experiment purposes with parametrized turbine inputs. This data was generated through support of the NJ Board of Public Utilities (NJBPU) and has been made publicly available [3].

Once the data requirements have been met, the project team then transitioned to the development of the forecasting technology and evaluation experiments. This is presented in Section 3.

3 AIRU-WRF: Modeling and Evaluations

This Section begins by describing the methodological components of the AIRU-WRF technology, and their key technical details (Section 3.1.). This is then followed by Section 3.2. which presents an overview of the key results from the forecasting evaluation of AIRU-WRF.

3.1 AIRU-WRF – Methodology Development

AIRU-WRF is a hybrid forecasting model – it integrates physics-based outputs from an NWP model and ML components, with the goal of producing forecasts that are superior to those obtained using NWP or purely ML models, independently. Table 2 summarizes the three core modules comprising AIRU-WRF, namely: The NWP down-scaler; the ML-based wind forecaster; and the ML-based power predictor.

Below, we conceptually summarize the key technical details of these modules, but a complete description of their methodological and implementation details can be found in our recent publications [4-7]. Figure 3 shows how the three methodological modules of AIRU-WRF are connected, their inputs/outputs, as well as key methodological details.

Table 2: The three modules comprising AIRU-WRF and their description.

Module	**Module Name**	**Description**
1	NWP Down-scaler	A data-driven calibration of NWP, aimed at correcting the multi-type biases of NWP outputs when downscaled at higher spatial and temporal resolutions.
2	ML-based Wind Forecaster	A probabilistic ML model that learns the spatial and temporal dependencies for enhanced extrapolation, along with a time-dependent ensemble for improving forecast skill across multiple time scales.
3	Power Predictor	A probabilistic ML model (considering input uncertainty) to convert, probabilistically, the hub-height wind speed forecast distributions into wind power production forecasts, enabling scenario generation and uncertainty quantification.

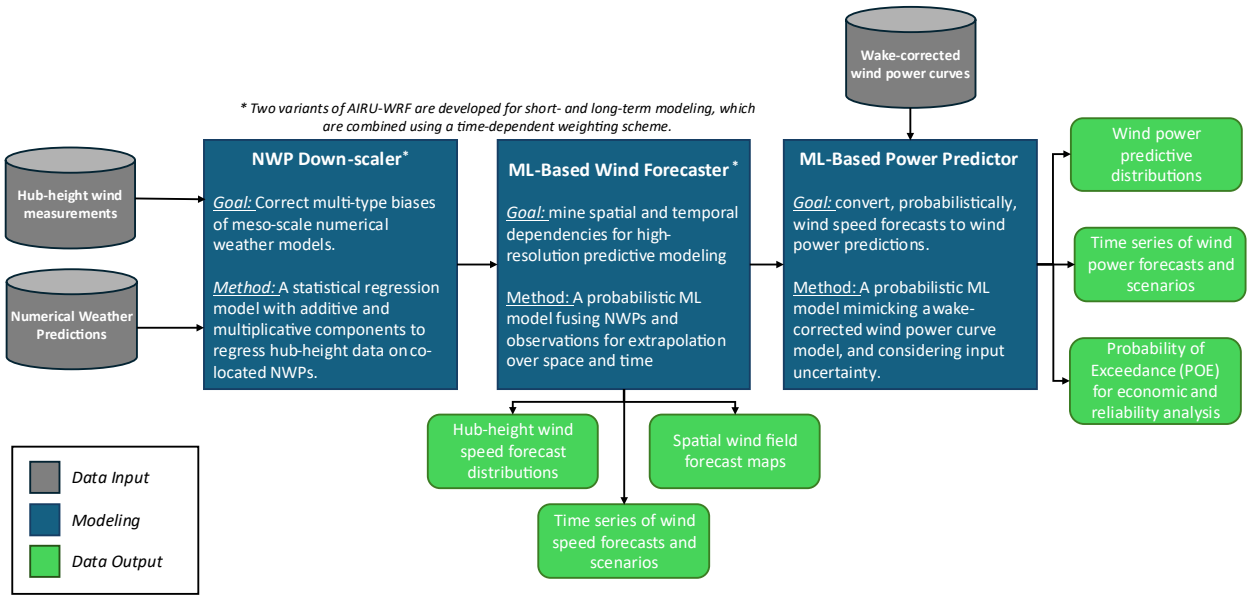

Figure 2: Modules of AIRU-WRF, including inputs, modeling details, and outputs

a. **NWP Down-scaler:** When downscaled at higher spatial and temporal resolutions, NWPs can suffer from noticeable biases and imprecisions. Figure 3 shows two separate days of hourly wind speed forecasts from the base numerical weather model RU-WRF, along with 10-min co-located hub-height measurements from the nearest buoy. When downscaled to the 10-minute, site-specific resolution at which the forecasts are needed, it is evident that NWPs, despite being generally able to predict the large-scale patterns of wind propagation, exhibit notable errors and multiple types of biases, including temporal bias (early/late predictions) and shift biases (under- or over-predictions). These are not specific to RU-WRF and are commonly observed when assessing even the best in-class NWP models [8]. This calls for a rigorous statistical treatment to explicitly correct these multi-type biases and produce precise predictions of satisfactory accuracy and operational value to support high-fidelity decision-making.

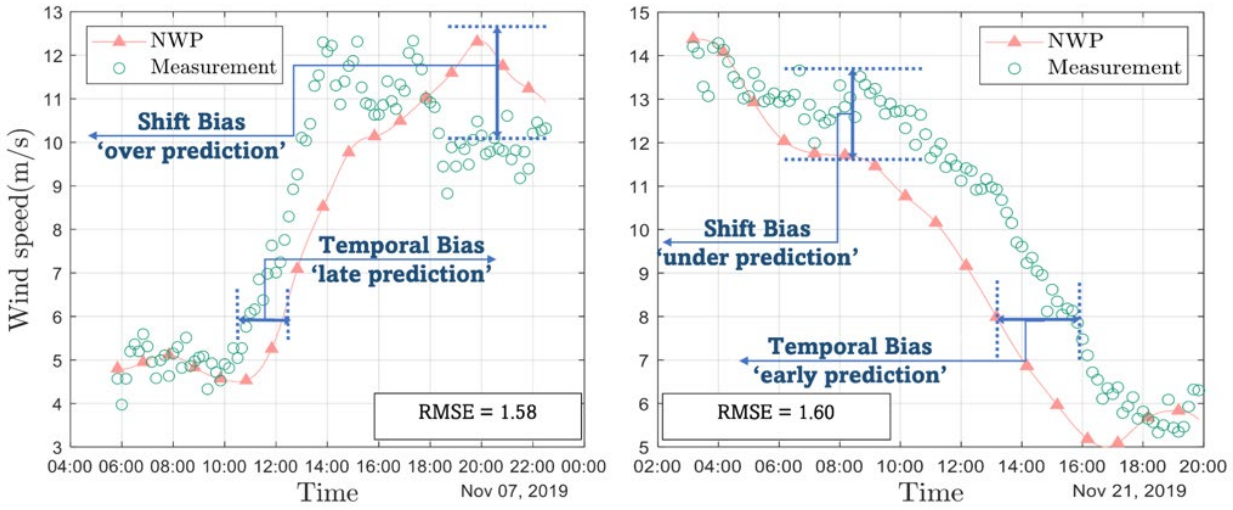

Figure 3: Two days of hub-height wind speeds and co-located NWPs from RU-WRF, suggesting overall reasonable accuracy but also significant downscaling biases.

To address this, the NWP down-scaler consists of a statistical regression function that estimates a relationship between the site-specific hub-height measurements and the set of co-located NWPs. The following overarching parametric form is adopted: $a^T\tilde{Y}^l(s,t) + b^T G(s,t) + c^T G(s,t)\tilde{Y}(s,t)$, where a, b, and c are parameter vectors representing the regression coefficients, whereas $\tilde{Y}^l(s,t) = [\tilde{Y}(s,t), \dots, \tilde{Y}(s,t-l)]^T$ are lagged NWP forecasts of the hub-height wind speeds at location s and up to lag $t-l$. Here, $G(s,t)$ is a regressor matrix, which includes standard outputs of NWP forecasts for various meteorological covariates (e.g., temperature, pressure, humidity), as well as other "engineered" features that are not readily forecast by NWP systems but can have explanatory power in predicting hub-height winds, such as geostrophic winds, pressure gradients, and sea-land temperature differences. Note that the NWP down-scaler includes regression and interaction terms that are designed to correct for both additive and multiplicative biases. Furthermore, a dynamic feature selection strategy, implemented in a sliding window fashion, continuously identifies and updates a concise subset of informative exogenous variables for inclusion in the regressor matrix $G(s,t)$. This feature selection procedure equips the forecasting model with the capability to dynamically determine the subset of meteorological covariates that are most predictive of NWP biases, which naturally are expected to change over space and time. Figure 4 shows the output of this dynamic feature selection procedure, where panel k (Bottom panel) shows the number of covariates included in the regressor matrix $G(s,t)$ over time. More details on this procedure and on the NWP downscaler is included in our published manuscript [4].

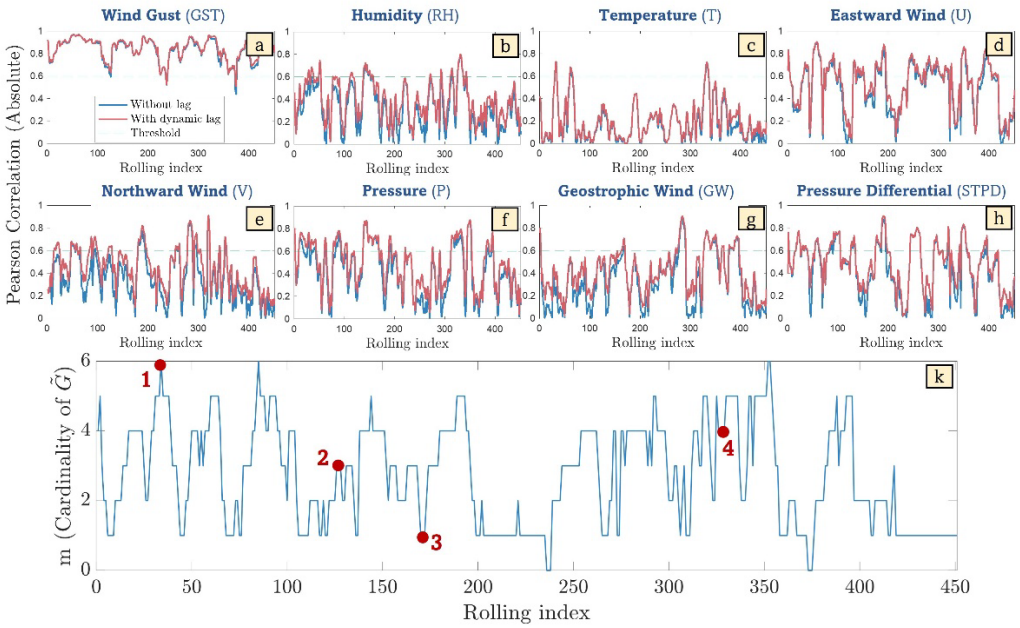

Figure 4: (a)-(h): Pearson Correlation of various meteorological NWP covariates with the hub-height measurements, showing how the explanatory power of various covariates can dynamically change over time. (k): The number of features selected in the AIRU-WRF downscaler over time. Rolling index refers to a sliding window (for this figure, 1 roll index = 6 hours).

In addition, we also developed a variant of the NWP down-scaler that acknowledges the heterogeneity of hub-height winds and power generation due to dynamic meteorological phenomenon along the U.S. East Coast, such as coastal upwelling – See Figure 5. Coastal upwelling is a common oceanographic process that occurs along stratified continental shelves, including in the U.S. Mid-Atlantic. Along-shore frictional wind stress, with the coastline to the left of the wind in the northern hemisphere (and right in the southern hemisphere) advects warm surface waters offshore. These warm surface waters are replaced through the upwelling of colder deeper waters, which directly interact with the overlying atmosphere. This phenomenon is well-known has been recently shown to directly impact the air-sea heat flux, the shape of the marine boundary layer, atmospheric stability, as well as the timing and location of important atmospheric phenomena such as sea breezes—all factors known to affect offshore wind power production. In fact, recent evidence from multiple geographical locations suggests that upwelling events can be associated with considerable variations in hub-height wind speeds and power production of offshore wind turbines. The central idea of our upwelling-informed AIRU-WRF model is to define a set of regimes (or states) based on the information of whether coastal upwelling is active or not, as derived from satellite imagery, and then invoke regime-specific NWP down-scaler functions that are trained on observations specific to the currently active regime. More details on this regime-switching variant of AIRU-WRF are available in our published manuscript [5].

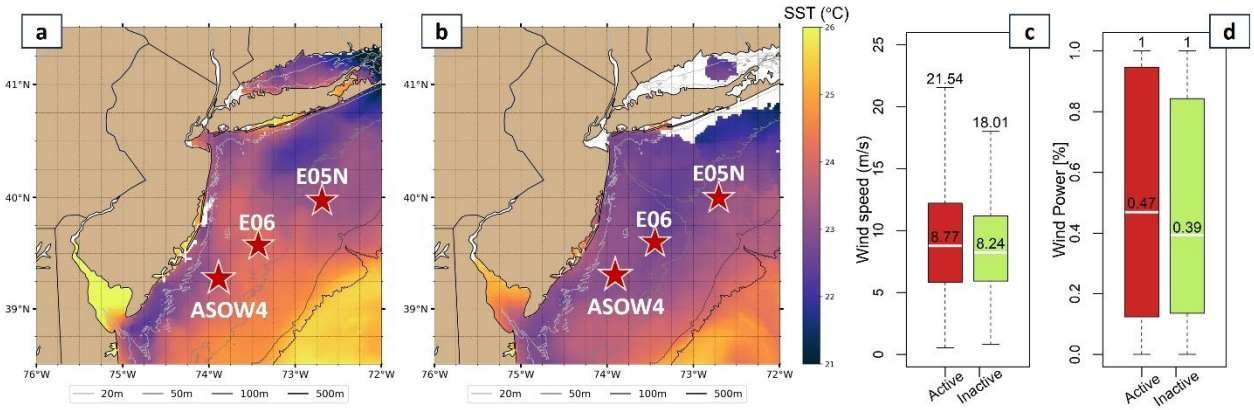

Figure 5: (a) SST on 07/28/2021 (upwelling active) with visible temperature gradient between nearshore and offshore areas. (b) SST on 08/10/2021 (upwelling inactive). (c)-(d): Boxplots of hub-height wind speeds and scaled power, suggesting distinct distributions when upwelling is active (red) or inactive (green).

b. **ML-Based Forecaster:** The ML-based forecaster directly acts on the residuals of the NWP down-scaler and is intended to capture site-specific, high-frequency variations that are difficult to capture using meso-scale NWP information. Specifically, the ML-based forecaster consists of a spatio-temporal Gaussian Process (GP) model. GPs are probabilistic, kernel-based ML models that model spatial and temporal dependencies using a kernel function. The kernel adopted in AIRU-WRF is a weighted combination of a radial basis function (RBF) and a physically motivated Lagrangian kernel function. This kernel function takes both NWP and observational data as inputs, and generates weights, based on spatial and temporal proximity, which enable extrapolations over space and time. One major advantage of GP models is their ability to output predictive distributions, rather than point/deterministic forecasts, that naturally embed spatial and temporal dependencies.

To produce day-ahead forecasts for every hour of the day, we developed two separate variants of AIRU-WRF for short- and long-term forecasting, respectively. This is because forecasting models perform well when tailored to specific forecast horizons, that is, methods for short-term forecasting (e.g., a few hours ahead) quickly degrade in performance as the forecast horizon extends (e.g., day-ahead) and vice versa. While sharing the same model formulation, the short- and long-term models differ primarily in their input features, their training data requirements, as well as their parameter settings, with the long-term model favoring the inclusion of meso-scale meteorological features that are able predictors of day-ahead wind patterns. To name one example of such meso-scale features, we define pressure gradients as the difference in atmospheric pressure between two locations at time t, which are known to be a meteorologically relevant indicator of weather fronts. We implemented a bivariate clustering analysis wherein we partition the wider

spatial region (roughly $300 \times 300\ km^2$ area) into a number of clusters, based on the NWP forecasts of air pressure and wind direction. Specifically, let $C_t = \{C_{t,1}, C_{t,2}, \cdots, C_{t,k}\}$ denote the set of clusters identified on a given day t, for each pair of clusters $(C_{t,i}, C_{t,j})$, we can compute the pairwise pressure gradient as:

$$\Delta P_{t,ij} = \left|P_{t,i} - P_{t,j}\right|, \forall i,j \in \{1,2,\cdots,k\}, i \neq j$$

where $P_{t,i}$ is the spatially averaged NWP forecast of air pressure at time t and cluster i. The selected pressure difference is then integrated into the regressor matrix of the NWP down-scaler as an additional predictive feature to enhance the forecast accuracy for the long-term. Figure 6 illustrates the spatial clustering results combined with wind vectors and corresponding air pressure contours for two consecutive days. Generally, the long-term model adopts a more "liberal" feature selection procedure, allowing the inclusion of more physics-based covariates based on the knowledge that large-scale meteorological patterns should intuitively play a more decisive role in day-to-day variations of wind propagation patterns. Meanwhile, as the forecast horizon becomes shorter (< 6 hours), more reliance on the near history of the process (and less on physics-based covariates) is expected to yield better forecasting performance.

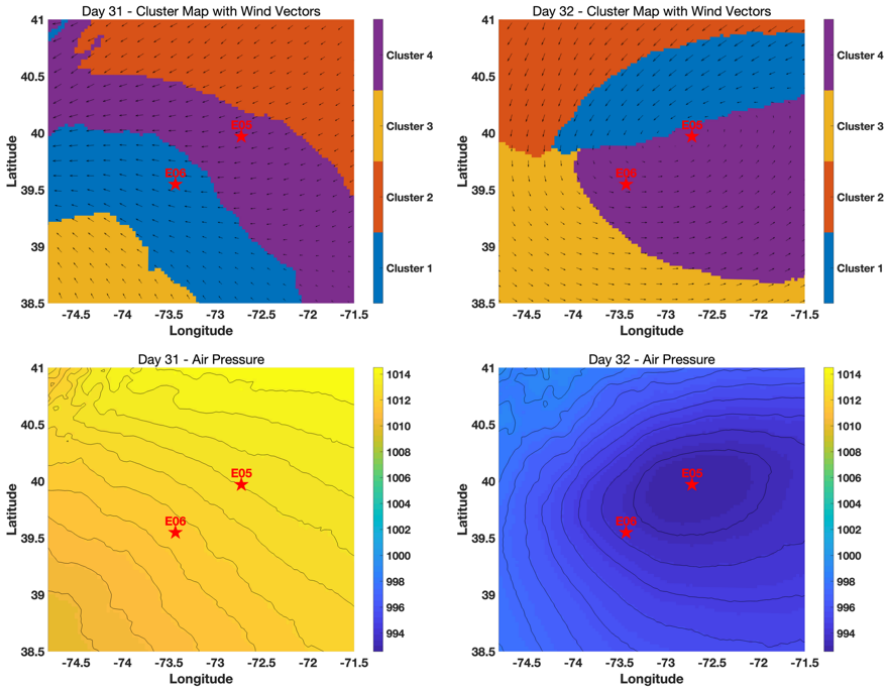

Figure 6: Cluster map with wind vector and air pressure information. The top panel is the cluster map overlaid with wind vectors; different colors represent different clusters. The bottom panel is the air pressure contour plot.

To reconcile the short- and long-term variants of AIRU-WRF, an ensemble approach combines the outputs from the two forecast models using a time-dependent weighting function which dynamically adapts to each model's degrading performance over time, ensuring temporally continuous forecast outputs. This enables the derivation, in closed form, of a probabilistic characterization of the forecast distribution of hub-height wind speeds, $\mathbb{P}^{\text{AIRUWRF}}(\boldsymbol{Y})$, where $\boldsymbol{Y}$ is the multivariate random variable denoting the hub-height wind speed forecasts for the full forecast horizon, i.e., from $t+1, \ldots, t+24$ hours ahead. This predictive distribution, $\mathbb{P}^{\text{AIRUWRF}}(\boldsymbol{Y})$, can be used to generate point/deterministic forecasts, by using the mean of the distribution, referred to as $\widehat{\boldsymbol{y}}_{\text{airuwrf}}$. The specific methodological details of this ensemble approach, along with the data and codes, are currently pending peer review and will be made publicly available soon [7].

c. **ML-Based Power Predictor:** AIRU-WRF outputs a forecast distribution of hub-height offshore wind speeds, $\mathbb{P}^{\text{AIRUWRF}}(\boldsymbol{Y})$. A probabilistic approach is then developed to convert the wind speed forecast distributions, $\mathbb{P}^{\text{AIRUWRF}}(\boldsymbol{Y})$, into closed-form predictive distributions of wind power, $\mathbb{P}^{\text{AIRUWRF}}(\boldsymbol{Z}|\boldsymbol{Y})$, where $\boldsymbol{Z}$ is the multivariate random variable denoting the offshore wind power production forecasts for the full forecast horizon, from $t+1, \ldots, t+24$ hours ahead. These predictive distributions can then utilized to: (1) efficiently generate trajectories (or scenarios) of wind power production at key offshore wind farms that naturally embed temporal dependencies; and (2) analytically derive important statistical measures, such as probability of exceedance (POE), which are key inputs to the subsequent economic and reliability analysis.

A probabilistic wind power curve model, based on a probabilistic GP model, is constructed using the 15 MW, wake-corrected wind power curve data adopted from NY-ISO [Link: https://www.nyiso.com/documents/20142/36079056/4%2023_02_07_ICAPWG_OffshoreWindProfileDevelopment.pdf]. Using the GP model, we can derive a closed-form expression for the predictive distribution of $\boldsymbol{Z}$ integrating over the wind speed forecast distribution $\boldsymbol{Y}$. A Monte-Carlo approximation is invoked to approximate this predictive distribution, $\mathbb{P}^{\text{AIRUWRF}}(\boldsymbol{Z}|\boldsymbol{Y})$. One can then leverage $\mathbb{P}^{\text{AIRUWRF}}(\boldsymbol{Z}|\boldsymbol{Y})$ to derive a point/deterministic forecast of offshore wind power production as the mean vector of this distribution, which would be denoted as $\widehat{\boldsymbol{\mu}}_{\boldsymbol{Z}}$.

Furthermore, the predictive distribution of offshore wind power production, $\mathbb{P}^{\text{AIRUWRF}}(\boldsymbol{Z}|\boldsymbol{Y})$, can be used to efficiently sample scenarios (or trajectories) of offshore wind power production and compute important statistical metrics such as the probability of exceedance (POE). POE denotes the probability that the wind power will exceed a certain critical threshold. Given the predictive distribution of Z, an analytical expression for POE can be derived as follows:

$$POE_Z(z) = P(Z \geq z) = 1 - F(z),$$

where $F(z)$ is the cumulative density function (CDF). Any quantile of offshore wind power can also be computed analytically. This POE metric will be a critical input to the subsequent economic and reliability analysis. More details on the derivation of the offshore wind power predictive distribution and the simulation of offshore wind power scenarios are currently under peer review and will be made publicly available soon [15].

3.2 Forecasting Experiments and Results

We focus on day-ahead forecasts in hourly resolution. We undertake a rolling/sliding window scheme for the model training and testing, where for each roll/sliding window, we train our model using d days of historical data and obtain the day-ahead forecast (one hour up to 24 hours ahead), and then slide by 24 hours and repeat this process, until all test data is exhausted. For the short-term model, we set $d = 10$ days, while for the long-term model, we set $d = 60$ days. This rolling/sliding window procedure is repeated for a total of 183 days. As a result, we obtain 183 * 24 = 4392 hourly testing instances for model evaluation at each location. We evaluate the performance of AIRU-WRF against the following benchmarks:

- **NWP:** These are the hourly day-ahead forecasts of hub-height wind speed from RU-WRF (the base NWP model on which the AIRU-WRF model is based).

- **PER:** refers to the persistence model, and is a standard benchmark that assumes the most recently observed wind speed value will persist into the future. It is highly competitive in very short-term horizons, but quickly degrade in performance as the forecast horizon extends.

- **LSTM:** Long short-term memory is a deep learning model based on recurrent neural networks, particularly well-suited for modeling time series data. A separate model is trained for each location using stochastic gradient descent (SGD) to optimize the network parameters through the Deep Learning Toolbox in Matlab. A grid search was used to tune the hyperparameters of the model. The model consists of 100 hidden units and is trained for 100 epochs. The initial learning rate is set to 0.005 and is reduced by a factor of 0.1 every 30 epochs. The LSTM model here does not utilize NWP information as input, and hence is designed so as a representative of a purely data-driven, deep learning model.

- **ARIMA:** This is the autoregressive integrated moving average with exogenous inputs (ARIMA), and is trained separately for each location. The model parameters and the autoregressive coefficients are optimized using the pmdarima package in Python.

Table 3 shows the Mean Absolute Error (MAE) and Continuous Ranked Probability Score (CRPS) value of all models for hub-height offshore wind speed forecasting. The results in this Table are aggregated across all locations and forecast horizons. Overall, AIRU-WRF achieves considerable improvements relative to all forecasting benchmarks. Specifically, it outperforms data-driven methods (LSTM, ARIMA-X, and PER) by $49.4 - 64.8\%$ and physics-based models (RU-WRF) by $7.4 - 7.6\%$.

Table 3: Evaluation of offshore 24-hour ahead wind speed forecast (hub-height), averaged across all spatial locations and forecast horizons (1-hr to 24-hrs ahead). Bold-faced values denote best performance.

	AIRU-WRF	LSTM	ARIMA	NWP (RU-WRF)	PER
Average MAE	**1.4710**	3.4923	2.9092	1.5881	4.1328
% Improvement	-	57.9%	49.4%	7.4%	64.4%
Average CRPS	**1.1201**	-	2.2736	1.2120	3.1865
% Improvement	-	-	50.7%	7.6%	64.8%

Figure 7 displays the MAE values, across the forecast horizon [1, …, 24] hours ahead, for AIRU-WRF (orange triangles) versus RU-WRF (blue squares). The forecasting results in this Figure show a clear superiority of AIRU-WRF over RU-WRF across all forecast horizons. The ability of AIRU-WRF to combine data-driven insights with physics-based NWP information, as well as to dynamically adapts to each source of information's degrading performance over time (through its time-dependent ensemble) proves to be key in producing superior forecasts across the look-ahead horizon.

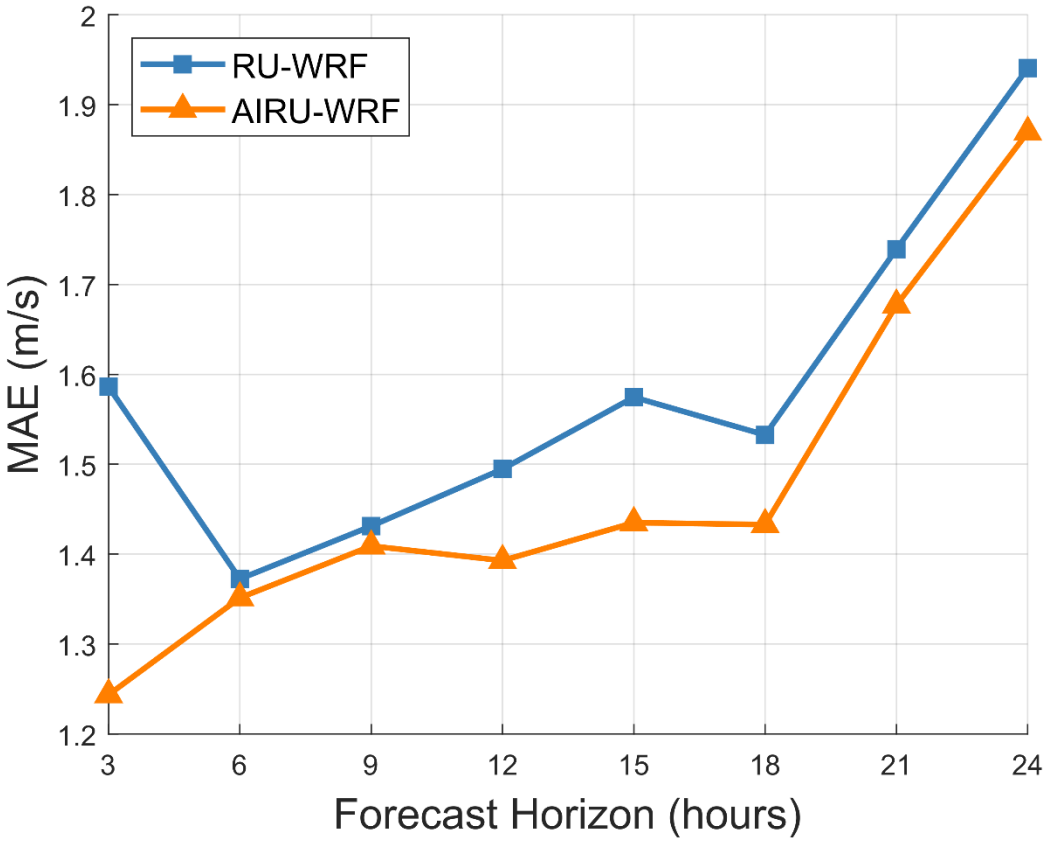

Figure 7: Forecast error (in MAE) of AIRU-WRF vs. that of RU-WRF plotted versus the forecast horizon in hours. The errors have been averaged across all spatial locations. Clearly, AIRU-WRF can improve upon RU-WRF across all forecast horizons.

Figure 8 shows the time series of AIRU-WRF's offshore wind speed forecasts for a select period in May, 2020, suggesting a faithful alignment with actual observations. The 90% prediction intervals (PI) produced by AIRU-WRF are also displayed. Most of the observations fall within the 90% PI, indicating that AIRU-WRF's uncertainty quantification has adequate coverage. The ability of AIRU-WRF to adjust forecast biases is demonstrated in Figure 9(a)-(b), where the true versus forecast values for AIRU-WRF show a more symmetric clustering around the 45^{o} line, relative to those from RU-WRF.

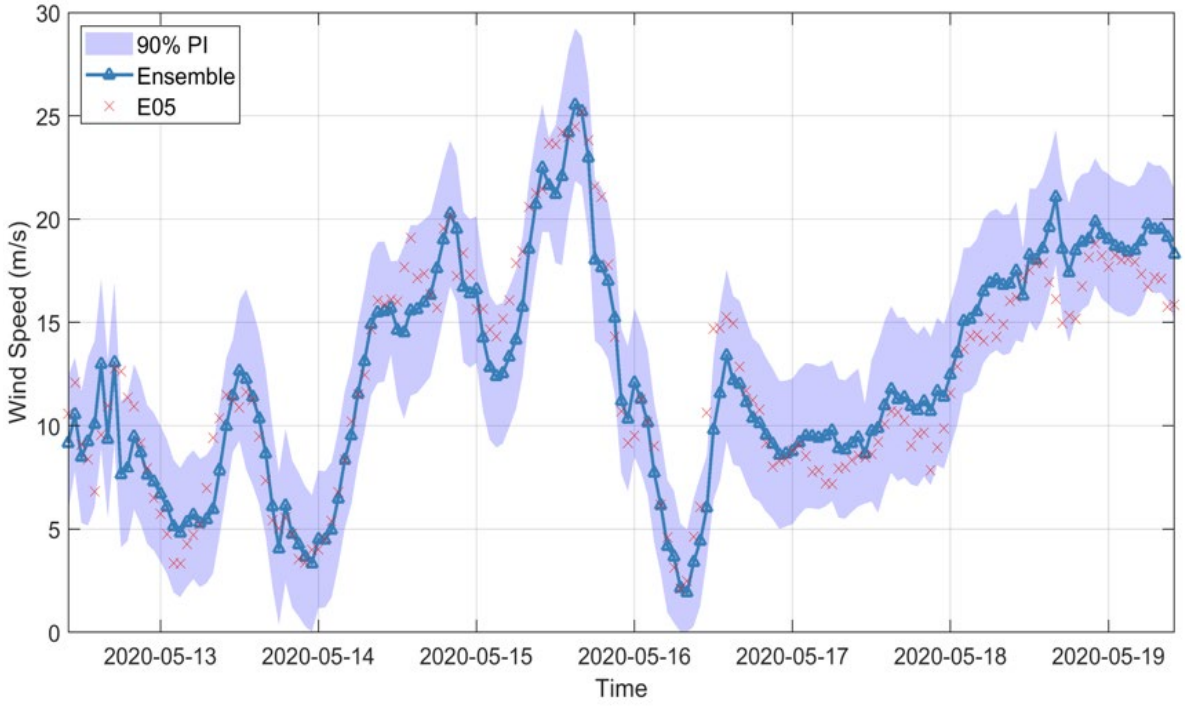

Figure 8: AIRU-WRF forecasts for a sample week in May 2020.

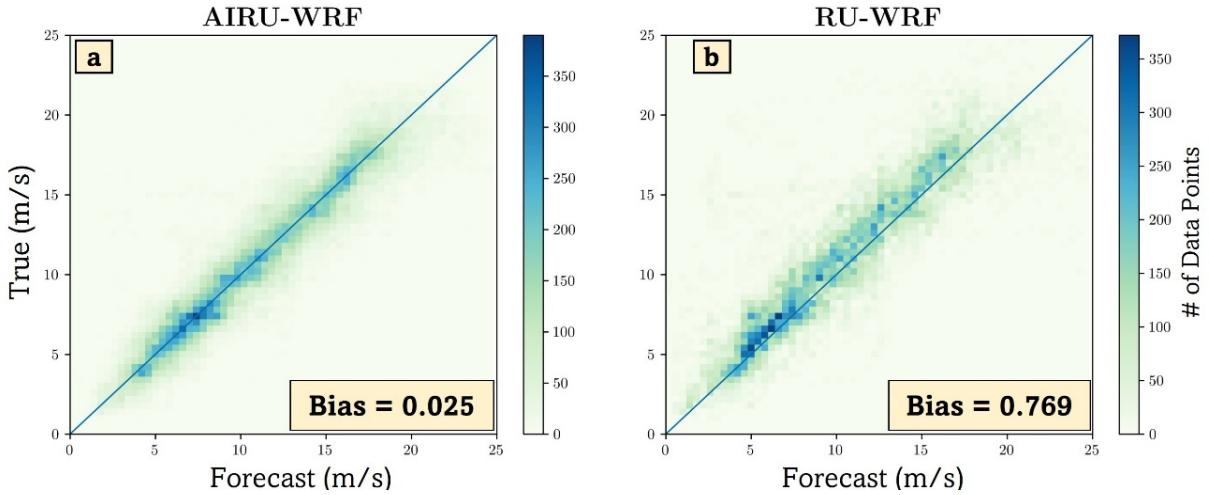

Figure 9: (a)-(b) Comparing AIRU-WRF and RU-WRF: On average, AIRU-WRF's forecasts are noticeably closer to the true values, yielding a significant reduction in forecast bias.

Unique to AIRU-WRF is its ability to make forecasts at locations where no observations are available. Those are used to produce spatial wind field forecast "maps" in the form of evolving two-dimensional images for a region of interest. Figure 10 shows examples of those wind field forecast maps on a select day, on top of the OSW energy lease areas in the NY/NJ Bight. A video showing the evolution

of these wind field forecast maps can found at the following webpage: https://sites.rutgers.edu/azizezzat/airu-wrf-data-science-based-offshore-wind-forecasting-model-for-the-u-s-east-coast/. These maps can be effective in communicating AIRU-WRF's outputs to stakeholders and forecast users in the OSW energy industry.

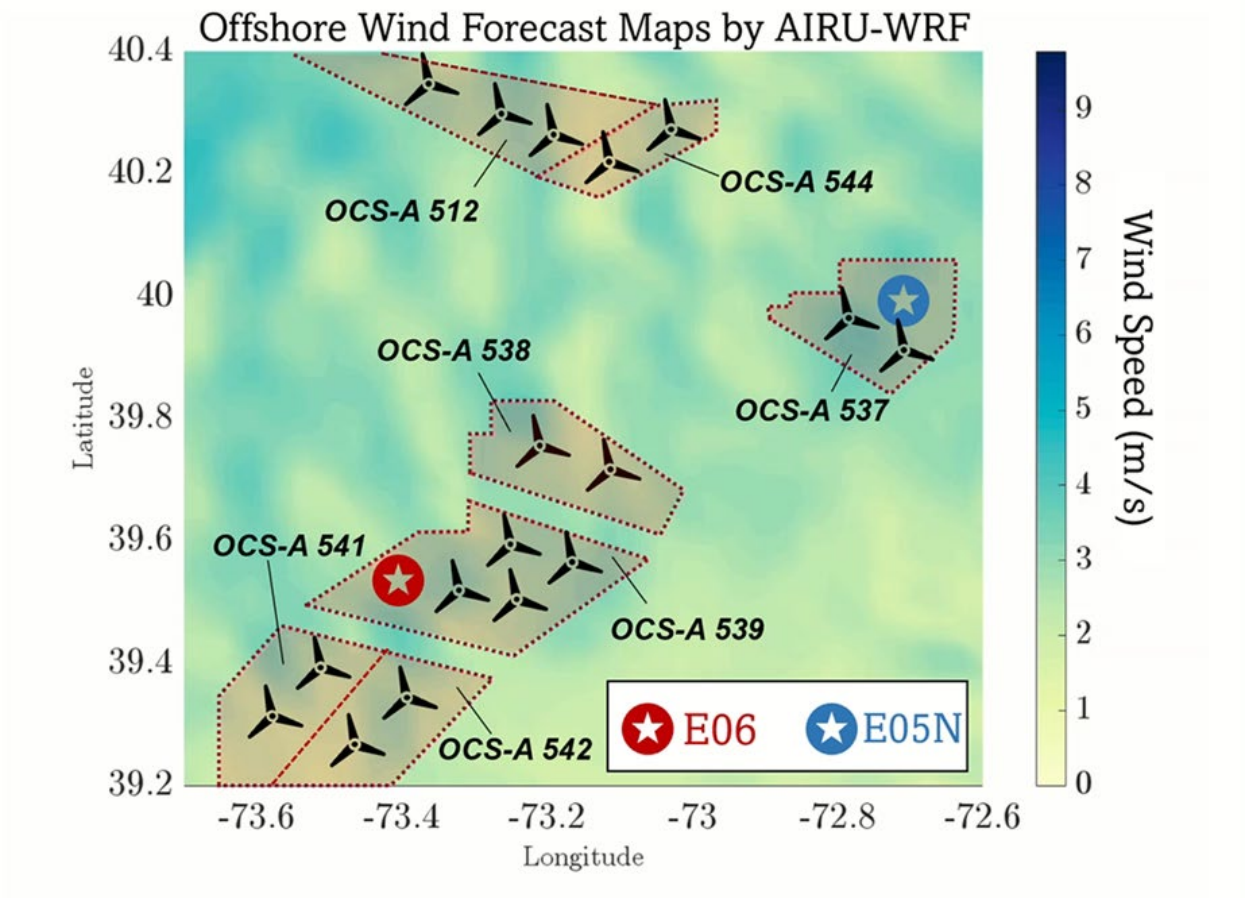

Figure 10: AIRU-WRF's spatial wind field forecast maps as evolving 2-dimensional images. A dynamic video illustration can be found at this webpage: https://sites.rutgers.edu/azizezzat/airu-wrf-data-science-based-offshore-wind-forecasting-model-for-the-u-s-east-coast/

Finally, the probabilistic nature of AIRU-WRF enables it to generate, in a matter of seconds, temporal trajectories (scenarios) of hub-height wind speeds and offshore wind production to support decision-making under uncertainty. This is demonstrated in Figure 11 which shows ten sampled OSW wind speed (left) and power production (right) scenarios generated by the AIRU-WRF day-ahead model (left panel) for a select period in February 2020.

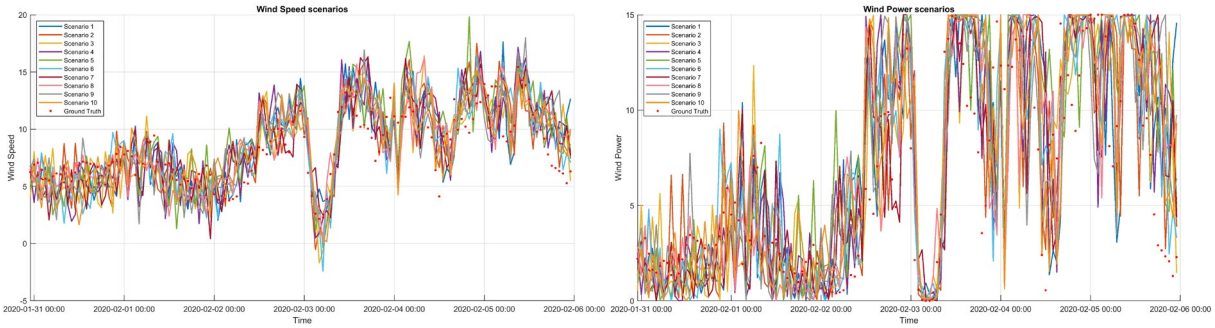

Figure 11: Scenario generation using AIRU-WRF for a week in February 2020 (10 scenarios shown).

4 Economic and Reliability Evaluation

Traditional forecast evaluations rely on statistical accuracy metrics, which can be insufficient to inform grid operations. As offshore wind deployment grows, it is increasingly important to assess forecasts based on their operational value. This project seeks to evaluate improved offshore wind forecasts through two major categories: (1) economic evaluation, which examines reductions in operational production costs enabled by more accurate forecasts, and (2) reliability evaluation, which measures the impact on system operational adequacy and loss-of-load risk under uncertainty [15].

4.1 Economic Evaluation

The economic benefit of improved offshore wind forecasting on grid operation is evaluated and quantified in terms of reduced production cost through the reduction of dynamic reserve MW. Dynamic reserve technology involves the real-time procurement of operating reserve resources to accommodate system contingencies and renewable forecast uncertainties. More accurate offshore wind forecasts enable grid operators to better manage uncertainty, optimize reserve procurement, and reduce overall production costs by minimizing operating reserve requirements.

4.1.1 Dynamic Reserve Calculation

In our study, we adopt NYISO's dynamic reserve procurement approach, currently transitioning from market design to implementation. It accounts for generation and transmission contingencies and simultaneous loss from geographically clustered intermittent resources.

- **Loss of Generation Resources**: Dynamic reserves are designed to cover the largest generation loss, minus available headroom, which represents the system's ability to import reserves. To manage the loss of generation in dynamic reserves for a 30-minute product, NYISO suggests the following formulation:

$$Res_{RA_{a_i}}^{30Total} \geq Mult_{RA_a}^{30Total} \cdot \left\{ \max_{k \epsilon Gen_{RA_a}} \left(gen_{k_i} + res_{k_i}^{30Total} \right) \right\} - RA_{aHeadroom} \quad (4.1.1)$$

where, in the reserve area a at time step t, $Res_{RA_{a_i}}^{30Total}$ is the 30-minute total MW reserve requirement; $Mult_{RA_a}^{30Total}$ is a multiplier; gen_{k_i} is the generation capacity of generator k; $res_{k_i}^{30Total}$ is the 30-minute reserve assigned to generator k; $RA_{aHeadroom}$ is the available headroom.

- **Loss of Transmission Resources:** Dynamic reserves account for lost energy imports due to transmission outages. The requirement is based on the difference between post-contingency interface limits and actual flows after the largest line loss. External proxies are treated as generators in NYCA reserve calculations.

$$30\min u\,te_{PostCon_{Import_{RA_{a_i}}}} = Limit_{N-2EmerRA_{a_i}} - RA_{Flow_{a_i}} \quad (4.1.2)$$

where, in the reserve area a at time step t, $30\min u\,te_{PostCon_{Import_{RA_{a_i}}}}$ is the post-contingency 30-minute import capability; $Limit_{N-2EmerRA_{a_i}}$ is the emergency limit for the $N-2$ contingency and $RA_{Flow_{a_i}}$ is the current flow on the interface.

- **Uncertainties from Renewable Resources:** To address variability from weather-dependent renewables and under-forecasting, NYISO uses a POE forecast, typically POE90 or POE95, to quantify potential energy at risk. Dynamic reserve requirements are based on the difference between scheduled IPP output and POE forecasts, adjusted for external reserve capacity, as defined by:

$$Res_{RA_{a_i}}^{30Total} \geq \left(\sum_{RA_{a_i}} IPP_{Schedule_i} - \sum_{RA_{a_i}} XX\,\%\,POE_{Forecast_i} \right) - RA_{aRES_{Capability_i}} \quad (4.1.3)$$

where, in the reserve area a at time step t, $IPP_{Schedule_i}$ is the scheduled energy output from IPP, which is usually POE50 value of forecast renewable generation. $POE_{Forecast_i}$ is the selected POE of renewable production uncertainties up to which the operating reserve can accommodate; and $RA_{aRES_{Capability_i}}$ is the headroom of the intertie transmission lines connecting with neighboring areas that provide the operating reserve.

4.1.2 Correlation Effect on Dynamic Reserves

Wind power uncertainty arises from the nonlinear wind speed–power curve, leading to skewed and variable forecast errors. Small changes in wind speed can cause large output deviations, making accurate reserve sizing essential. When aggregating reserves across multiple wind farms, spatial correlation becomes a key factor. Wind farms spreading over a wide area experience only partially coincident fluctuations: output correlation generally decreases as the distance between sites increases [9]. Distant wind farms are less likely to experience simultaneous power drops due to varying weather systems, while closely located

farms under the same weather front can exhibit high output correlation, increasing aggregate uncertainty. Uncorrelated or weakly correlated sites offer diversity, reducing total output variability. Ignoring spatial correlation can lead to significant reserve misestimation—assuming independence underestimates reserve needs, risking underproduction, while assuming perfect correlation overestimates requirements, increasing costs. Accurate correlation modeling is essential for risk-aware, cost-effective dynamic reserve sizing based on probabilistic forecasting.

To quantify the aggregated dynamic reserve requirement under correlated uncertainty, we adopt a correlation-based risk aggregation approach—commonly applied in financial risk management for the calculation of the Solvency Capital Requirement (SCR) [10]. The total dynamic reserve requirement is computed as:

$$DR_{total,Corr} = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} R_i \times R_j \times \rho_{ij}} \quad ,\forall i,j \in N \tag{4.1.4}$$

where, R_i and R_j are the individual plant reserve requirements; $\rho_{i,j}$ is the correlation in between plant i and plant j; and N is the number of plants.

This formulation captures the spatial dependency in generation uncertainty, enabling operators to leverage geographic diversity for optimal reserve procurement. Wider uncertainty from one site can be offset by narrower ranges from another, improving reserve efficiency—especially when aggregating across regions. As site correlation decreases, the aggregate reserve requirement declines, reflecting reduced system risk through spatial diversification.

To demonstrate the impact of correlation, two offshore wind farms (A0487 and A0512) with probabilistic forecasts from the AIRU-WRF model are analyzed. Table 4 presents their POE90, POE50 values, and individual reserve requirements.

Table 4: Dynamic reserve calculated from probabilistic wind power forecasting.

A0487 (Capacity 930 MW)			**A0512 (Capacity 2070 MW)**		
POE90 (MW)	POE50 (MW)	Individual Reserve (MW)	POE90 (MW)	POE50 (MW)	Individual Reserve (MW)
344.08	454.12	110.04	686.93	951.97	265.03

Equation (4.1.4) is used to compute total reserve requirements for the combined system under varying correlation levels. Table 4 summarizes the aggregated reserve for different correlation coefficients between the two sites.

Table 5: Aggregated dynamic reserve with varying correlation coefficient.

Correlation Coefficient	1.00	0.80	0.60	0.40	0.20	0.00
Reserve (MW)	375.10	359.20	342.60	325.10	306.60	286.97
Reserve (% of Nameplate Cap.)	12.50%	11.97%	11.42%	10.84%	10.22%	9.57%

Result shows that, when the correlation is high (1.0), the reserve nearly equals the sum of individual reserves (375.07 MW or 12.5% of capacity). As correlation decreases, the required reserve drops significantly—down to 286.97 MW (9.57%) at zero correlation—demonstrating the benefit of spatial diversity. Lower correlation reduces reserve needs by offsetting fluctuations across sites. Incorporating this effect improves reserve accuracy and reduces system-wide costs.

4.1.3 Production Cost Calculation

In this study, the historical operating reserve price from NYISO is analyzed and the production cost for reserve procurement is calculated by multiplying the reserve MW procurement requirement and the price as

$$Cost = Res_{RA_{a_i}}^{30Total} \cdot \mathrm{Pr}\, i\, ce_{a_i} \tag{4.1.5}$$

A more robust approach is to use price quantiles, which offer probabilistic insights and more reliable results. This leads to a modified formula that incorporates the price at a given percentile.

$$Cost = Res_{RA_{a_i}}^{30Total} \cdot \mathrm{Pr}\, i\, ce_{a_i}(P), \quad \forall P \in \{1,2,\dots,99\} \tag{4.1.6}$$

4.1.4 Case Study on Economic Evaluation

In this study, we assess the economic impact of improved offshore wind forecasting on dynamic reserve requirements using the New York power system as a case study. In this study, we evaluate our improved offshore wind forecast model, AIRU-WRF, which aims to reduce the forecast band gap for dynamic reserve calculations. The performance of the model is compared to two benchmark models: the

NWP and Persistence model. For both benchmarks, forecasted quantiles are generated by superimposing historical residuals onto the predictions.

The dynamic reserve requirements vary across different plant locations due to site-specific variations in wind power generation. Within a given site, reserve needs also exhibit seasonal and diurnal fluctuations, reflecting the temporal variability in wind generation uncertainty. We present the seasonal average reserve requirements as a percentage of nameplate capacity for various plants locations. Table 6 summarizes these values based on the AIRU-WRF model. Overall, the summer and winter seasons generally require higher reserve margins compared to Spring and Fall.

Table 6: Seasonal dynamic reserve requirements as percentage of nameplate capacity (Model: AIRU-WRF; temporal correlation preserved during sampling).

Plants	Winter	Spring	Summer	Fall	Capacity (MW)
A0487	15.65%	12.91%	16.91%	13.96%	924
A0512	15.67%	13.61%	17.06%	14.80%	130
A0517	15.84%	13.14%	17.40%	14.59%	1230
A0520	16.07%	12.72%	17.05%	14.09%	2076
A0538	15.39%	12.93%	16.63%	14.39%	1404
A0539	15.28%	12.87%	16.20%	14.47%	1314
A0544	15.65%	13.16%	17.20%	14.68%	1314
E05N	15.69%	12.91%	17.20%	14.33%	5674
E06/ASOW6	15.17%	12.93%	16.16%	14.61%	2718

To evaluate the performance of the improved AIRU-WRF forecasting model, we focus on sites where direct wind speed measurements are available—like, locations E05N and E06/ASOW6. Since observational data is limited to these positions, we validate model performances by aggregating wind farm capacities within the proximity of E05N or E06/ASOW6. We computed dynamic reserves for these two locations to evaluate the improved model (AIRU-WRF) performance over two benchmark models and the results are summarized in Table 7.

Table 7: Performance analysis of models in terms of dynamic reserve requirements (percentage of nameplate capacity).

Seasons	E05			E06/ASW6		
	AIRU-WRF	NWP	PER	AIRU-WRF	NWP	PER
Winter	15.69%	17.56%	33.20%	15.17%	18.33%	31.32%
Spring	12.91%	15.96%	32.93%	12.93%	16.39%	32.90%
Summer	17.20%	19.87%	32.68%	16.16%	21.41%	32.38%
Fall	14.33%	17.70%	32.64%	14.61%	17.38%	30.53%

The AIRU-WRF model consistently provides lower seasonal dynamic reserve requirements (expressed as a percentage of nameplate capacity), reflecting reduced forecasting uncertainty. Additionally, seasonal analysis shows that reserve requirements are lowest in Spring and highest in Summer, aligning with observed wind power variability patterns. Therefore, dynamic reserve design must explicitly account for seasonal variability.

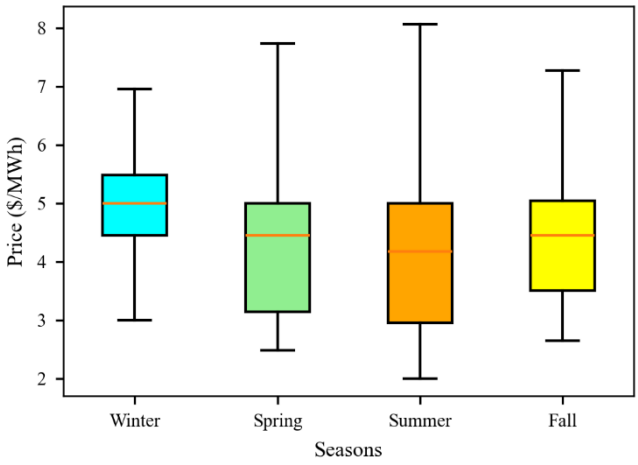

Figure 12: 10 min Spinning Reserve Average Price (2019 – 2023).

To calculate reserve procurement cost, we analyzed five years of 10 minutes and 30 minutes operating reserve price data (2019-2023) across various load zones from NYISO. The specific zone data can be used for offshore wind interconnection points. As illustrated in Figure 12, the reserve prices also exhibit seasonality. Summer and Spring exhibit a wider price distribution compared to Winter and Fall. Notably, the median price (P50) is typically higher during the Winter season. In this study, the P50 price for each season is used to ensure a representative measure of typical reserve costs across different periods.

Figure 13: Annual reserve procurement cost comparison a) without temporal correlation b) with temporal correlation consideration.

Seasonal dynamic reserve procurement cost with and without considering temporal correlation is demonstrated in Figure 13. Although reserve requirements are highest in Summer, procurement costs peak in Winter due to higher reserve prices. Thus, identical reserve percentages can incur different costs across seasons due to price variability. By calculating annual reserve procurement cost for all models, it is evident that the AIRU-WRF model consistently outperforms both benchmark models (NWP and PER), regardless of temporal correlation inclusion as shown in Figure 13. Without temporal correlation, AIRU-WRF achieves reserve procurement cost savings of approximately $1.65 million over NWP model annually. Incorporating temporal correlation further enhances these savings, reaching up to $3.20 million annually, emphasizing the significant economic benefits of improved forecasting methodologies in dynamic reserve design.

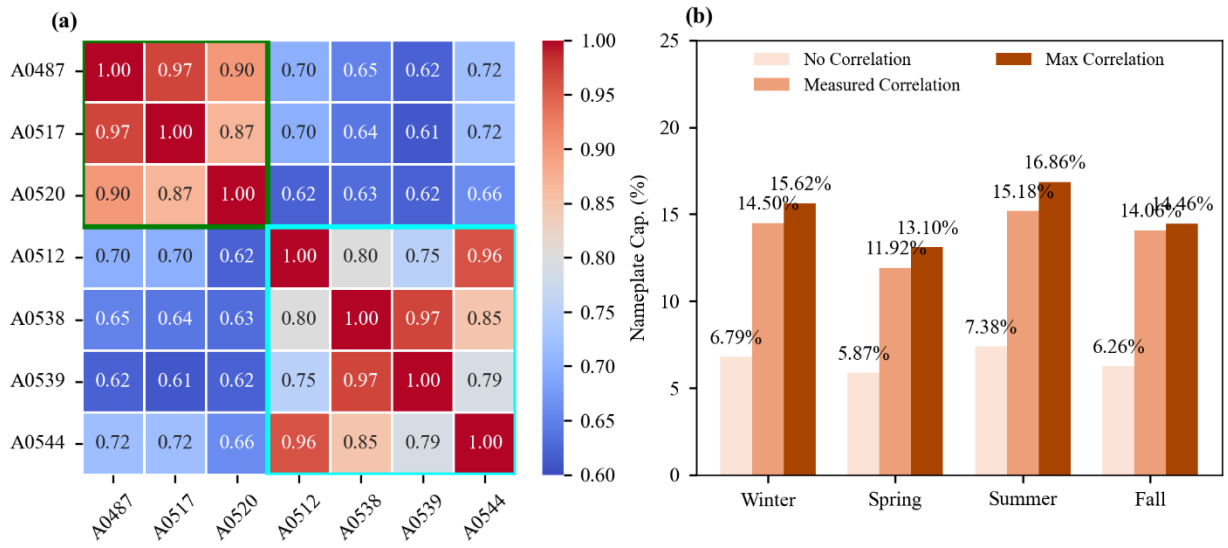

Figure 14: Correlation coefficient among sites for summer season b) Aggregated dynamic reserve considering correlation.

In Section 4.1.2, we incorporated correlation considerations into the aggregated dynamic reserve calculation. Pearson partial correlation coefficients for generated wind power profiles were considered. For the summer season, the correlation of wind power generation across different sites is shown in Figure 14(a). As evident, closely situated sites (e.g., A0487 and A0517) exhibit high correlation, indicating simultaneous wind generation fluctuations and thus higher operational risks. Conversely, geographically distant sites show lower correlations, reducing aggregated variability. For instance, A0487 and A0539 have a correlation coefficient of 0.62, significantly lower than 0.97 between A0487 and A0517. This spatial correlation is essential for accurately capturing the risk-based aggregation of dynamic reserves. Using Equation (4.1.4), dynamic reserve calculations considering correlations based on the AIRU-WRF forecast model are summarized in Figure 14(b). When correlation is assumed to be maximal (1.00), dynamic reserves equal the simple summation across individual sites, yielding no diversity benefits. Conversely, when assuming zero correlation, maximum diversity benefits reduce the aggregated reserve requirements.

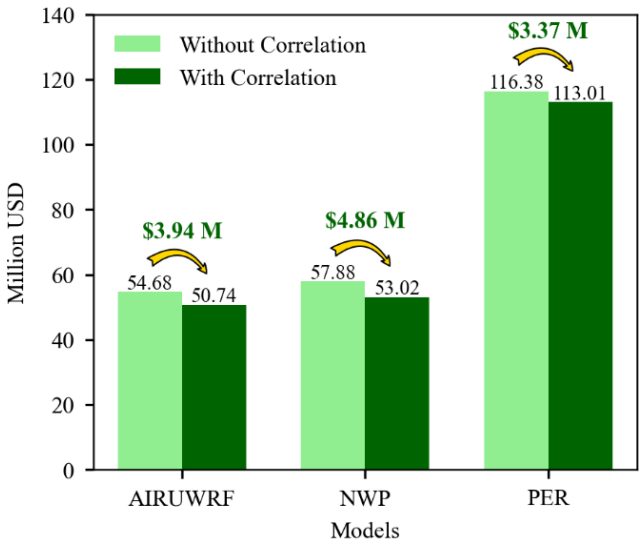

Figure 15: Yearly dynamic reserve procurement cost considering spatial correlation.

We evaluated total seasonal reserve procurement costs across all plant locations using seasonal price information. The AIRU-WRF model consistently outperformed both the NWP and PER benchmark models for all seasons. While reserve requirements and associated costs demonstrate notable seasonal variability, AIRU-WRF consistently minimizes procurement expenses. Furthermore, considering temporal correlation alone, annual reserve procurement costs using AIRU-WRF total approximately $54.68 million, resulting in annual savings of about $3.2 million compared to NWP, and over $61 million compared to the PER approach. Incorporating spatial correlation further enhances these savings. As illustrated in Figure 15, accounting for spatial correlation reduces AIRU-WRF based annual procurement costs to approximately $50.74 million. This spatial correlation consideration provides additional annual savings of around $3.94 million compared to the nonspatial scenario. Thus, explicitly accounting for both temporal and spatial correlations significantly improve reserve allocation efficiency and procurement cost savings.

4.2 Reliability Assessment through Operational Resource Adequacy Model

Power grids face increasing challenges from the uncertainties introduced by rapid integration of variable renewable energy sources like land-based wind (LBW) and offshore wind (OSW). Current grid operations, focused mainly on single contingencies, lack tools to assess combined uncertainties from load and renewable forecasting, leading to gaps in managing tail risks and costly uncertainty reserves. The absence of regulatory standards further limits operators' ability to quantify and respond to operational risk. To address these issues, we propose Operational Resource Adequacy, a control-room tool that quantifies

real-time uncertainties and assesses loss of load risks. The tool consists of two main components: a unit commitment model to identify committed generation resources, and a scenario-based DC Optimal Power Flow (DC-OPF) model to detect potential resource shortfalls. It incorporates uncertainties from load, LBW, and OSW forecasts using Markov Chain Monte Carlo simulations. System responses are evaluated under transmission and generation constraints, and key reliability metrics—Loss of Load Probability (LOLP), Loss of Load Hours (LOLH), and Expected Unserved Energy (EUE)—are computed to quantify resource adequacy.

To assess grid reliability, forecasting data with uncertainties are required for load ($N(L_t^f, \sigma_{ld,t}^2)$), land-based wind ($N(P_{l,t}^{LBW}, \sigma_{LBW,t}^2)$), and offshore wind ($N(P_{o,t}^{OSW}, \sigma_{OSW,t}^2)$). For simplicity, these uncertainties are modeled as normal distributions $N(\mu, \sigma^2)$, at each time step t. The standard deviation $\sigma_{i,t}$,where $i\epsilon\{load, LBW, OSW\}$, represent the uncertainty range for each resource. In alignment with standard RTO/ISO practices, generation resources are deterministically committed based on forecast values for load (L_t^f), land-based wind ($P_{l,t}^{LBW}$), and offshore wind ($P_{o,t}^{OSW}$). Generator commitments are established before executing a DC-OPF model, which assesses shortfalls and grid reliability. The unit commitment model aims to minimize daily production costs, formulated as follows:

$$min \sum_{t \in T} \left(\sum_{i \in \Omega_{Gen}} C_i^{GEN} * g_{i,t} + SU_i * C_i^{SU} + SD_i * C_i^{SD} + \sum_{s \in \Omega_{ESS}} \left(C_{s,t}^{ESR} * \left(P_{s,t}^{C} - P_{s,t}^{D} \right) \right) \right) \quad (4.2.1)$$

subject to

$$\sum_{i \in \Omega_{Gen}} g_{i,t} + \sum_{s \in \Omega_{ESS}} \left(P_{s,t}^{D} - P_{s,t}^{C} \right) + \sum_{l \in \Omega_{LBW}} P_{l,t}^{LBW} + \sum_{o \in \Omega_{OSW}} P_{o,t}^{OSW} = L_t^f, \forall t \in T \quad (4.2.2)$$

where the optimization model spans over time periods T and includes traditional generators Ω_{Gen}, OSW farms Ω_{OSW}, LBW farms Ω_{LBW}, and energy storage systems Ω_{ESS}. Each generator $i \in \Omega_{Gen}$has a marginal cost C_i, startup cost C_i^{SU}, $shutdown\ cost\ C_i^{SD}$, with power output $g_{i,t}$. For the OSW and LBW generated power are $P_{o,t}^{OSW}$ and $P_{l,t}^{LBW}$. Each storage unit $s \in \Omega_{ESS}$, has marginal cost of $C_{s,t}^{ESR}$, charging and discharging power are $P_{s,t}^{C}$ and $P_{s,t}^{D}$ respectively. Constraint (4.2.2.) ensures that total generation from conventional units, wind farms, and net energy storage contributions meet the forecasted load L_t^f at all times.

Once the unit commitment signal is derived from the modeling above, the DC-OPF model can be initialized for assessing the shortfall. In contrast to the traditional DC-OPF with an objective of minimizing

the operational generation cost, the DC-OPF for operational resource adequacy assessment is to maximize the total load served subjected to the generation operational constraints and line thermal limits. Note that the forced outage rate (FOR) of generation resources and transmission lines is considered in the loss of load risk quantification, and the committed generation and/or transmission lines may be out of service in real time. After running the DC-OPF for a power grid, the maximum load-serving capacity for each bus is calculated by subtracting the load demand at each bus by the served load. Based on the shortfall amount and instance, operational reliability metrics, i.e., LOLH, LOLP, and EUE are calculated.

4.2.1 Case Study on Reliability Assessment with Improved Offshore Wind Forecasts

In this study, we apply the model to assess the performance of the improved offshore wind forecasting model, AIRU-WRF, alongside two benchmark models—NWP and PER. The analysis is based on the 2022 New York State generation fleet, using publicly available data from the NYISO Gold Book [11]. In 2022, the fleet included approximately 587 in-market generation units, with a total installed capacity of 36.30 GW and a peak demand of 30.50 GW. At that time, no offshore wind capacity was operational, and land-based wind contributed 1,739 MW. For future projections, we consider NYISO's System and Resource Outlook (Policy Case S1 and S2) [12].

We align with New York State's Climate Leadership and Community Protection Act (CLCPA), which targets 9,000 MW of offshore wind capacity by 2035. Our analysis uses the AIRU-WRF probabilistic forecasting model, benchmarked against NWP and PER models. While AIRU-WRF supports hour-ahead probabilistic forecasts, validation is limited to 2020 buoy data (E05N and E06). For 2030 and 2035 projections, forecast error characteristics from 2020 are assumed to remain constant, with adjustments only to nameplate capacity.

Table 8: Generation capacity in different years.

Gen Info/Year	**2022**	**2030**	**2035**
Bio	327	327	327
GasCC	9686	9686	9686
GasCT	3282	3282	3282
GasST	7923	7923	7923
Hydro	4437	4437	4437
Nuclear	3343	3343	3343
OffshoreWind		5764	8392
OilCT	1549	1549	1549
OilST	2825	2825	2825
OnshoreWind	1739	9086	12612
PSH	1167	1167	1167
Solar	32	32	32
Total Capacity	**36310**	**49421**	**55575**
No of Gen Fleet	***587***	***592***	***594***

The generation mix for each scenario is detailed in Table 8. Due to data confidentiality, generator specific FOR, MTTF, and MTTR are not publicly available; we use aggregated reliability data from the 2024 NREL Annual Technology Baseline (ATB) [13] and IEEE RTS [14]. To reduce computational complexity and maintain illustrative clarity, one representative week per season is sampled (Winter: Dec–Feb; Spring: Mar–May; Summer: Jun–Aug; Fall: Sep–Nov). Results for 2030 are shown in Table 9, where seasonal peak loads reflect the highest hourly demand within each sampled week. Variations across rows highlight the performance differences among AIRU-WRF, NWP, and PER models, the primary focus of this analysis.

Table 9: System reliability result comparison considering different offshore wind forecast model.

Seasons	Model	LOLP (%)	LOLH (hour/day)	EUE (MWh)
Winter	AIRUWRF	0.609	0.146	71.35
	NWP	1.151	0.276	147.78
	PER	2.695	0.647	610.23
Spring	AIRUWRF	1.399	0.336	226.86
	NWP	1.876	0.450	349.77
	PER	7.487	1.797	2090.30
Summer	AIRUWRF	0.798	0.192	97.59
	NWP	1.022	0.245	116.20
	PER	0.174	0.042	21.35
Fall	AIRUWRF	21.880	5.251	9979.55
	NWP	23.913	5.739	10676.62
	PER	23.888	5.733	7425.80

For the year 2030, we assume an offshore wind capacity of 5,764 MW and onshore wind capacity of 9,086 MW will be operating. Offshore wind exhibits a higher capacity factor (50–60%) compared to onshore wind (25–30%), contributing significantly to overall system adequacy. Table 9 presents seasonal reliability metrics under three offshore wind forecast models. The improved AIRU-WRF model consistently outperforms the benchmark models (NWP and PER) in reducing system risk. For instance, in Winter, AIRU-WRF achieves a LOLP of 0.609%, significantly lower than NWP (1.151%) and PER (2.695%). Similarly, AIRU-WRF achieves lower LOLH and EUE values, indicating fewer hours and lower energy shortfalls. These metrics can be extended to seasonal totals; for instance, a Winter LOLH of 0.146 h/day under AIRU-WRF results in approximately 13.14 total hours of load loss over a 90-day season. Notably, the Fall season exhibits the poorest reliability across all models, with AIRU-WRF showing a LOLP of 21.88%, suggesting increased system vulnerability during this period despite higher offshore wind production. This highlights the seasonal variation in reliability outcomes, especially under high offshore

wind penetration scenarios. It is important to note that no additional traditional generation or storage resources are included in this analysis.

Table 10: Reliability results comparison across different years for AIRU-WRF offshore wind.

Seasons	Year	LOLP (%)	LOLH (hour/day)	EUE (MWh)
Winter	2022	0.001	0.000	0.05
	2030	0.609	0.146	71.35
	2035	1.940	0.466	453.27
Spring	2022	0.008	0.002	0.60
	2030	1.399	0.336	226.86
	2035	1.534	0.368	258.90
Summer	2022	0.076	0.018	8.73
	2030	0.798	0.192	97.59
	2035	5.441	1.306	1223.54
Fall	2022	0.024	0.006	1.72
	2030	21.880	5.251	9979.55
	2035	24.299	5.832	12455.82

Reliability performance over different years is evaluated and summarized in Table 10. In 2022, the system exhibited its highest Loss of Load Probability (LOLP) during the summer, primarily due to the absence of offshore wind and the limited onshore wind capacity of 1,739 MW. Reliability in this case is largely influenced by forced outages, captured through MTTF and MTTR parameters. As offshore wind capacity expands, reliability outcomes shift notably. For instance, in Winter, LOLP increases from 0.001% in 2022 to 0.609% in 2030, and further to 1.94% in 2035, reflecting growing dependence on variable wind resources. A notable shift in seasonal reliability risk is observed, with Summer being the most vulnerable season in 2022, transitioning to Fall in 2030 and 2035. This shift is driven by increased offshore wind contributions during Fall, particularly at night and in the early morning hours. Although fall offers the highest average wind output, it also exhibits greater variability, resulting in higher system reliance and risk. As offshore wind capacity expands, this combination of high contribution and uncertainty significantly increases the likelihood of shortfall events, reflected in elevated LOLP values for fall in 2030 (21.88/%) and 2035 (24.30%).

The reliability metrics are pre-risk mitigation actions taken by grid operators in real-time grid operations; such reliability risk mitigation could include increased operating reserve procurement, commitment of generation resources from in-market products, such as real-time commitment, or out of market products, such as Capacity Analysis Commitment Tool.

5 Conclusions and Potential Future Developments

This report summarized the main developments, findings, and results of the NOWRDC project #133/192900, titled: "*AIRU-WRF: AI-powered Physics-based Tool for Offshore Wind Forecasting & Grid Integration*." In this project, we have developed and demonstrated a machine-learning-enabled forecasting technology, called AIRU-WRF, that is able to produce high-resolution operational forecasts of the offshore wind resource and generation at key offshore wind energy regions in the U.S. East Coast. When tested at multiple offshore locations where hub-height data is available, our experiments show that AIRU-WRF can improve the forecast accuracy by ~7-8% relative to base numerical weather prediction models, and up to ~50-65% compared to purely data-driven models, across the full forecast horizon (from 1-hour to 24-hours ahead).

Beyond forecast accuracy, the advantages of AIRU-WRF are demonstrated, including its ability to generate probabilistic characterization of the offshore wind resource and generation, its ability to generate high-resolution spatial wind field forecast maps for effective visualization and communication of the forecast outputs, and its ability to generate temporal trajectories/scenarios of hub-height winds and power production to support risk-aware operational decision-making at both the farm- and system-levels. Furthermore, we have extensively tested the accrued benefits of utilizing AIRU-WRF forecasts in grid integration, as measured by economic value and reliability in reserve quantification and grid support. Specifically, AIRU-WRF has been shown to reduce reserve procurement costs by up to $1.65-3.2M annually compared to best-in-class numerical weather products and drive significant improvements in reliability metrics including loss of load probability, loss of load hours, and expected unserved energy.

This project raises a number of important avenues for further research, development, and application of AIRU-WRF. From a technological development perspective, a key enabler of AIRU-WRF is a dense network of synchronous and spatially distributed hub-height offshore wind data. With the scarcity and sparsity of these measurements, the team is undertaking research and development efforts to augment AIRU-WRF with generative AI capabilities to learn and generate synthetic hub-height data (using existing datasets) for extended streaks of time at "virtual locations." This synthetic data can then be integrated, in addition to the actual measurements, to train the AIRU-WRF model for improved extrapolation over space and time. Furthermore, the increased volume of data will enable the project team to explore the merit of using emerging deep learning architectures, including graph- and transformer-based neural network models.

Extensive testing across multiple locations, time periods, and environmental settings is currently being pursued. Given our geographical focus on the U.S. East Coast, the team is currently exploring avenues to conduct similar testing and grid integration analysis for other system operators that may benefit from offshore wind forecasts to support future grid integration studies and scenario analysis (e.g., New Jersey, Virginia). Finally, another important development avenue is to design an adaptive graphical user interface to further broaden the reach and usability of AIRU-WRF by a vast portfolio of stakeholders, including offshore wind developers, system operators, as well as other ocean users.

6 Project Dissemination

6.1 Publications and Patents:

1. Ye, Feng, Joseph Brodie, Travis Miles, and Ahmed Aziz Ezzat. "AIRU-WRF: A physics-guided spatio-temporal wind forecasting model and its application to the US Mid Atlantic offshore wind energy areas." *Renewable Energy* 223 (2024): 119934. https://doi.org/10.1016/j.renene.2023.119934

2. Ye, Feng, Joseph Brodie, Travis Miles, and Ahmed Aziz Ezzat. "Ultra-Short-Term Probabilistic Wind Forecasting: Can Numerical Weather Predictions Help?" *In 2023 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1-5. IEEE, 2023. DOI: 10.1109/PESGM52003.2023.10252311

3. Ye, Feng, Travis Miles, and Ahmed Aziz Ezzat. "Improved spatio-temporal offshore wind forecasting with coastal upwelling information." *Applied Energy* 380 (2025): 125010. https://doi.org/10.1016/j.apenergy.2024.125010

4. Ye, Feng, Travis Miles, and Ahmed Aziz Ezzat. "Offshore Wind Energy Prediction Using Machine Learning with Multi-Resolution Inputs." *In Multimodal and Tensor Data Analytics for Industrial Systems Improvement*, pp. 167-183. Cham: Springer International Publishing, 2024

5. Ji, Jiaxiang, Ye, Feng, Travis Miles, and Ahmed Aziz Ezzat. "AIRU-WRF: Time-Dependent Ensembles for Day-ahead Spatio-Temporal Offshore Wind Energy Forecasting," *Under Review* (2025)

6. Walid, K.B., Ye, F., Ji, J., Miles, T., Aziz Ezzat, A., and Jiang, Y., "Economic and Reliability Value of Improved Offshore Wind Forecasting in Bulk Power Grid Operation," *Under Review* (2025).

7. "*Techniques To Provide Improved Wind Input for Operating Offshore Wind Turbines*," U.S. Application 19/111. Inventors: Ahmed Aziz Ezzat, Feng Ye, Travis Miles, Joseph Brodie (supported by the Rutgers Office of Innovation Ventures)

6.2 Dataset and Software:

1. Aziz Ezzat, A., Ye, F., Ji, J., & Miles, T. (2025). AIRU-WRF: Spatial-Temporal Wind Datasets and Forecasts for Data-Science-Based Operational Offshore Wind Forecasting in the U.S. East Coast [Data set]. Zenodo. https://doi.org/10.5281/zenodo.15642047

2. AIRU-WRF Webpage [Updates, News, Software]: https://sites.rutgers.edu/azizezzat/airu-wrf-data-science-based-offshore-wind-forecasting-model-for-the-u-s-east-coast/

7 References

[1] New York State Energy Development Authority (2019), “NYSERDA Floating LiDAR Buoy Data,” https://oswbuoysny.resourcepanorama.dnvgl.com/

[2] Atlantic Shores Offshore Wind (2021), “Atlantic Shores Offshore Wind, ASOW-6, Winds Profile,” https://erddap.maracoos.org/erddap/tabledap/index.html?page=1&itemsPerPage=1000/

[3] Rutgers University Weather Research & Forecasting Model (RUWRF), https://tds.marine.rutgers.edu/thredds/catalog/cool/ruwrf/wrf_4_1_3km_processed/catalog.html

[4] Ye, Feng, Joseph Brodie, Travis Miles, and Ahmed Aziz Ezzat. "AIRU-WRF: A physics-guided spatio-temporal wind forecasting model and its application to the US Mid Atlantic offshore wind energy areas." *Renewable Energy* 223 (2024): 119934. https://doi.org/10.1016/j.renene.2023.119934

[5] Ye, Feng, Travis Miles, and Ahmed Aziz Ezzat. "Improved spatio-temporal offshore wind forecasting with coastal upwelling information." Applied Energy 380 (2025): 125010. https://doi.org/10.1016/j.apenergy.2024.125010

[6] Ye, Feng, Joseph Brodie, Travis Miles, and Ahmed Aziz Ezzat. "Ultra-short-term probabilistic wind forecasting: Can numerical weather predictions help?." In 2023 IEEE Power & Energy Society General Meeting (PESGM), pp. 1-5. IEEE, 2023. DOI: 10.1109/PESGM52003.2023.10252311

[7] Ji, Jiaxiang, Ye, Feng, Travis Miles, and Ahmed Aziz Ezzat. “Time-Dependent Ensembles for Day-ahead Spatio-Temporal Offshore Wind Energy Forecasting,” *Under Review* (2025)

[8] Sweeney, Conor, Ricardo J. Bessa, Jethro Browell, and Pierre Pinson. "The future of forecasting for renewable energy." Wiley Interdisciplinary Reviews: Energy and Environment 9, no. 2 (2020): e365. https://doi.org/10.1002/wene.365

[9] S. Pattanariyankool and L. B. Lave, “Optimizing transmission from distant wind farms,” *Energy Policy*, vol. 38, no. 6, pp. 2806–2815, Jun. 2010, doi: 10.1016/j.enpol.2010.01.012.

[10] Committee of European Insurance and Occupational Pensions Supervisors (CEIOPS), “CEIOPS’ Advice for Level 2 Implementing Measures on Solvency II: SCR Standard Formula Article 111(d) Correlations,” Committee of European Insurance and Occupational Pensions Supervisors(CEIOPS), CEIOPS-DOC-70/10, Jan. 2010. Accessed: Jun. 05, 2025. [Online]. Available: https://register.eiopa.europa.eu/CEIOPS-Archive/Documents/Advices/CEIOPS-L2-Advice-Correlation-Parameters.pdf

[11] "Library," New York Independent System Operator (NYISO). Accessed: Jul. 03, 2025. [Online]. Available: https://www.nyiso.com/library#reports

[12] NYISO, "2022 Grid in Transition Study: A Study of Expected Ramp Rates," New York Independent System Operator (NYISO), 2022. [Online]. Available: https://www.nyiso.com/documents/20142/34388803/Grid%20in%20Transition%20FINAL%201219 2022.pdf/0f7b7072-21f4-c81f-1805-d53a15eec18f

[13] Mirletz and Brian, "Open Energy Data Initiative (OEDI)." 2024. doi: 10.25984/2377191.

[14] C. Grigg *et al.*, "The IEEE Reliability Test System-1996. A report prepared by the Reliability Test System Task Force of the Application of Probability Methods Subcommittee," *IEEE Transactions on Power Systems*, vol. 14, no. 3, pp. 1010–1020, Aug. 1999, doi: 10.1109/59.780914.

[15] Khaled Bin Walid, Feng Ye, Jiaxiang Ji, Ahmed Aziz Ezzat, Travis Miles, Yazhou Leo Jiang, "Economic and Reliability Evaluation of Improved Offshore Wind Forecasting in Bulk Power Grid Operation." *Under Review* (2025).