Advanced Energy Partnership for Asia

WIND RESOURCE DATA FOR SOUTHEAST ASIA USING A HYBRID NUMERICAL WEATHER PREDICTION WITH MACHINE LEARNING SUPER RESOLUTION APPROACH

David Rosencrans, Brandon Benton, Grant Buster, Andrew Glaws, Ryan King, Julie K. Lundquist, Jianyu Gu, Galen Maclaurin

National Renewable Energy Laboratory

June 2023

A product of the USAID-NREL Partnership

Contract No. IAG-19-2115

NOTICE

This work was authored, in part, by the National Renewable Energy Laboratory (NREL), operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the United States Agency for International Development (USAID) under Contract No. IAG-19-2115. The views expressed in this report do not necessarily represent the views of the DOE or the U.S. Government, or any agency thereof, including USAID.

Acknowledgments

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the U.S. Agency for International Development, the DOE Office of Grid Deployment, and the DOE Solar Energy Technologies Office. The research was performed using computational resources sponsored by the Department of Energy's Office of Energy Efficiency and Renewable Energy and located at the National Renewable Energy Laboratory (NREL). The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes. The authors thank Scott Bartos and Amanda Van den Dool (U.S. Agency for International Development) for their support of this research. The authors also thank Billy Roberts (NREL) for producing cartographic figures for this paper, Evan Rosenlieb, Avi Purkayastha, Paul Edwards, Nick Gilroy, and Reid Olson (NREL) for their support of the project, and Caroline Draxl and Nicola Bodini (NREL) for their reviews and feedback to improve this paper.

List of Acronyms

cRMSE	centered root mean square error
EEZ	Exclusive Economic Zone
EMD	earth mover's distance
GANs	generative adversarial networks
IBC	initial and boundary condition
MAE	mean absolute error
NREL	National Renewable Energy Laboratory
PBL	planetary boundary layer
RMSE	root mean square error
SODAR	sound detection and ranging
sup3r	Super Resolution for Renewable Resource Data
WRF	Weather Research and Forecasting

Executive Summary

This project developed a new, computationally efficient paradigm for wind resource assessment using the Weather Research and Forecasting (WRF) model coupled with a novel machine learning framework to produce large spatiotemporal meteorological data sets. This machine learning framework leverages generative adversarial networks (GANs) to downscale, or super-resolve, a high-accuracy yet coarse-resolution wind resource assessment using WRF to fine spatial and temporal resolutions. This hybrid Numerical Weather Prediction + GANs paradigm was used to generate 15-year high-resolution wind, temperature, and pressure data from January 2007 through December 2021 at multiple hub heights over Southeast Asia (Figure ES- 1). The resolutions of this novel data set are 3-km horizontally and 15-minute temporally.

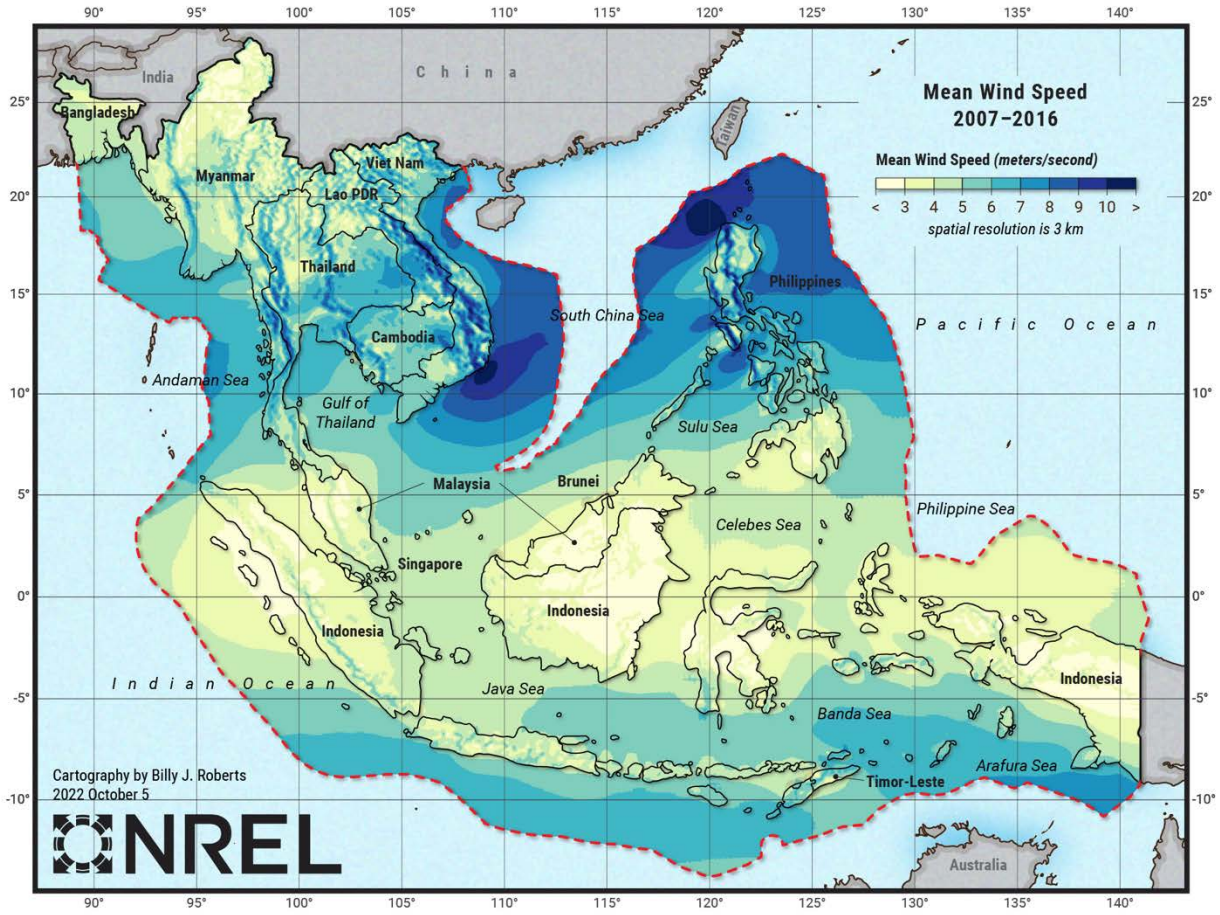

Figure ES- 1. Long-term mean wind speed of the final data set (100 m above ground level)

We made an optimal selection of boundary conditions and physics options for WRF simulations based on comparisons to tower observations in Bangladesh for a 21-day period during February 2016. Using these physics options, we performed low-resolution (9-km grid) WRF simulations over Southeast Asia. The domains for these simulations were stitched together to provide a single final domain. The GANs were trained on coarsened WIND Toolkit data, and the stitched domain was enhanced by 3x along each spatial dimension and 4x along the temporal dimension. These data can serve as essential inputs to many wind energy analysis tools and models and benefit a broad range of stakeholders working across Southeast Asia

to deploy wind energy, including developers, policymakers, governments, utilities, system operators, technical institutes, and energy consultants.

This hybrid model, and the methods described in this report, substantially reduce computational costs while producing high accuracy wind data appropriate for modeling and site evaluation. The hybrid approach using GANs downscaling to 3-km 15-minute resolution provides a 16x speedup when compared to carrying out WRF simulations at 3-km 15-minute resolution directly. Python code developed through this project for feature engineering, data handling, model training, and inference is publicly available on GitHub as the Super Resolution for Renewable Resource Data (sup3r) package (Buster, Benton, Glaws, and King 2022).

Table of Contents

List of Figures

List of Tables

1 Introduction

The National Renewable Energy Laboratory (NREL) has developed a new, computationally efficient paradigm for wind resource assessment using the Weather Research and Forecasting (WRF) model coupled with a novel machine learning framework (generative adversarial networks, or GANs) to produce large spatiotemporal meteorological data sets. This project developed a trainable algorithm to downscale, or super-resolve, a high-accuracy yet coarse-resolution wind resource assessment, using WRF to fine spatial and temporal resolutions with a GANs model. With the growing need to support wind energy planning and development from small to utility-scale across Southeast Asia, this data set provides publicly available, 15-year wind resource data from January 2007 through December 2021 at multiple hub heights. The grid resolutions of this novel data set are 3 km spatially and 15 minute temporally.

These data can serve as essential inputs to many wind energy analysis tools and models and benefit a broad range of stakeholders working across Southeast Asia to deploy wind energy, including developers, policymakers, governments, utilities, system operators, technical institutes, and energy consultants. Novel renewable energy resource assessments enable the next generation of high-fidelity energy system modeling at regional, continental, and global scales. This wind data set will complement the timeseries solar irradiance data recently produced by NREL.[1] This report describes the WRF setup, the motivation for the choices made for the WRF production runs, and a description of the GANs setup and generation of the final downscaled data set for Southeast Asia.

[1] For more information, see https://www.nrel.gov/docs/fy22osti/81799.pdf.

2 WRF Model Setup

Numerical weather prediction simulations were carried out using the WRF model (Skamarock 2019), version 4.2.1. The setup included a 50-hPa top, 61 vertical levels, an hourly output frequency, and monthly reinitializations discarding 1 day of spin-up from the last day of the preceding month.

The length of time a simulation is executed before reinitialization can affect the accuracy of the output. Shorter simulations with more frequent reinitializations moderate error growth but may introduce sudden discontinuous changes to physical properties. Longer continuous runs are more susceptible to error growth but may more accurately resolve the evolution of observed quantities (Xia et al. 2017). While large-scale wind resource assessments have provided accurate results, incorporating monthly reinitializations with 2 days of spin-up (Draxl et al. 2015), it may only take 24 hours to exploit most of the added benefit (Gómez and Miguez-Macho 2017). To provide highly accurate results while conserving computational cost, we submitted monthly reinitializations with 1 day of spin-up.

A 27-km parent domain covers the entire Southeast Asia region, with three nested domains consisting of 9-km grid resolution and one with 3-km grid resolution (Figure 1). The effective resolution of the model, however, is roughly 7 times coarser than the grid resolution (Skamarock 2004). Dissipation is required at the highest wavenumbers to prevent energy pooling from the energy cascade. This threshold is delineated by the wavenumber at which spectral energy density begins to decay in comparison to the true energy density. The extents of the 9-km domains cover the land and offshore areas of member countries[2] from the Association of Southeast Asian Nations with overlapping areas to mitigate artificial spatial seams. Offshore areas are determined by the Exclusive Economic Zone (EEZ)[3] of each country covered by the data set. The use of three smaller 9-km domains reduces the computational expense inherent in resolving meteorology for offshore areas beyond the EEZ. The 27- and 9-km hourly WRF outputs serve as inputs to the GANs model (Stengel et al. 2020). The GANs model downscales fields for wind speed, wind direction, temperature, and atmospheric pressure to the final 3-km, 15-minute resolutions. A separate 3-km, 15-minute WRF domain over Bangladesh serves as a training data set for comparison with the GANs super-resolved output. Observations for validation were only available to this project over Bangladesh.

[2] For a list of member countries, see https://asean.org/member-states/.
[3] For more information, see https://www.un.org/depts/los/convention_agreements/texts/unclos/part5.htm.

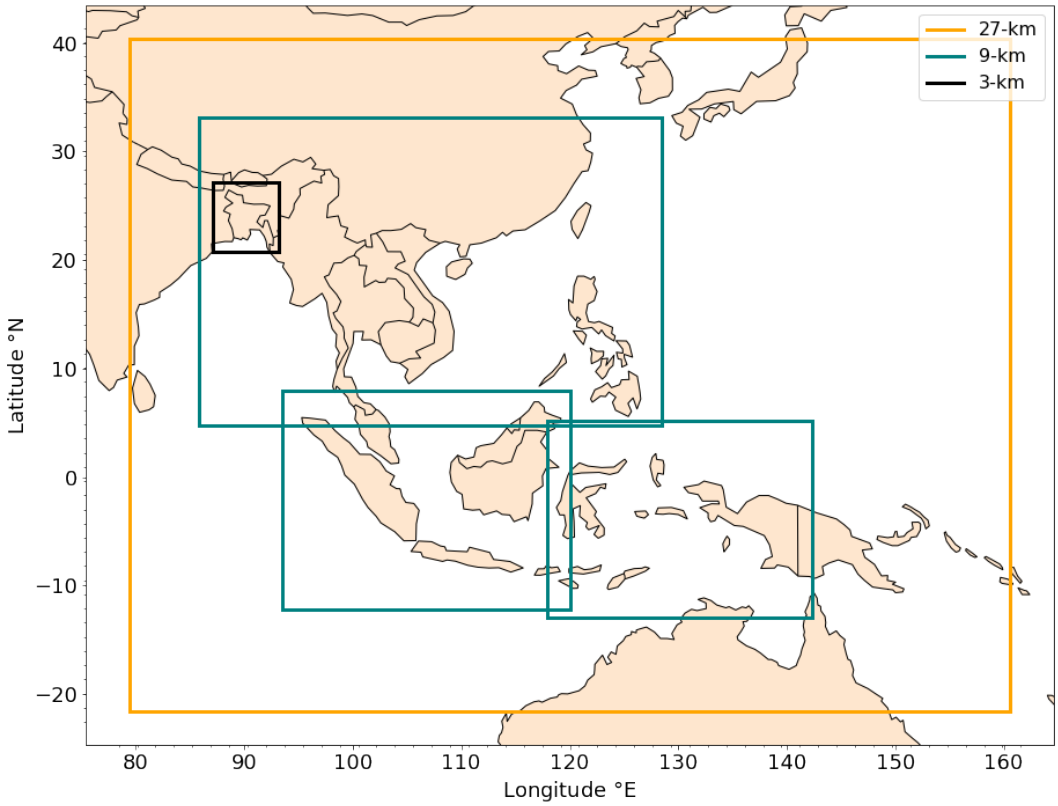

Figure 1. WRF simulation domains

Note: The parent domain is outlined in orange and utilizes 27-km grid resolution. Three nested domains are outlined in green and utilize 9-km grid resolution. The inner domain outlined in black is a training domain at 3-km grid resolution.

Spectral nudging was applied to relax model outputs toward the boundary conditions. Fields were transformed to spectral space using a Fast Fourier Transform. Wavelengths smaller than a designated threshold, the cutoff wavelength, were set to zero and allowed to develop freely, while those longer than the cutoff wavelength were guided toward the boundary conditions. We used a cutoff wavelength that corresponds with the Rossby Radius of Deformation. The Rossby Radius of Deformation is roughly 100 times the vertical scale (Gill 1982) and delineates the transition between convective and synoptic scales (Gómez and Miguez-Macho 2017). Prior work has found the best results incorporate spectral nudging in the mid-latitudes using a cutoff wavelength of 1,000 km, which corresponds with a typical 10-km tropopause height (Gómez and Miguez-Macho 2017; Yang et al. 2019). With tropopause heights ranging from 10 km to 16 km between the midlatitudes and the equator, we assumed a Rossby Radius of Deformation of 1,300 km and used nudging coefficients of 3×10^{-4} (Hahmann et al. 2015; Vincent and Hahmann 2015; Gómez and Miguez-Macho 2017; Silva and Camargo 2018; Yang et al. 2019). No nudging was performed beneath the planetary boundary layer (PBL) height so that the model could resolve boundary-layer processes explicitly.

Table 1. Summary of Model Initial and Boundary Condition (IBC) Data Sets Assessed

Note: From left to right, columns indicate the name of the data set, the center providing that data set, the horizontal resolution, the number of pressure levels, and the temporal output frequency.

Product	Center	Horizontal Resolution	Number of Levels	Temporal Frequency
FNL	NCEP	1° ×1°	26 levels	6-hourly
ERA5	ECMWF	0.25° ×0.25°	37 levels	1-hourly
MERRA2	NASA	0.5° × 0.63°	72 levels	3-hourly

Before the final production run calculations can be made, choices of IBCs must be selected. We used a set of 21-day sensitivity simulations, which we compared to observations in Bangladesh to determine which set of IBCs should be used. IBCs incorporate a blend of instrument measurements and model outputs to provide an accurate, gridded representation of the state of the surface and the atmosphere. These datasets are ingested by models to guide simulated output toward "truth" and reduce error growth. Here, we considered FNL (NCEP 2000), ERA5 (Hersbach et al. 2020), and MERRA2 (Gelaro et al. 2017). Prior work documented accurate performance using CFSR input conditions in Bangladesh (Jacobson et al. 2018). However, we removed the CFSR data set from consideration, as it underwent a version change during the production run period. GFS also performs well in the region (Draxl et al. 2014; Singh et al. 2015; Du Duc et al. 2019; Nuryanto et al. 2019). Because FNL uses the same data assimilation procedures as GFS, but with 10% more observations, we chose FNL instead. MERRA2 performed well regarding wind speed sensitivity (Jacobson et al. 2018), and we added the ERA5 dataset for breadth considering its strong performance in Europe (Hahmann et al. 2020).

Table 2. Summary of Sensitivity Simulations

Note: Columns delineate separate ensemble members. The first row specifies the IBC dataset, the second row the PBL scheme used, and the third row the horizontal resolution of each domain.

	Sim. 1	**Sim. 2**	**Sim. 3**	**Sim. 4**	**Sim. 5**	**Sim. 6**	**Sim. 7**	**Sim. 8**	**Sim. 9**
IBC	ERA5	ERA5	ERA5	FNL	FNL	FNL	MERRA2	MERRA2	MERRA2
PBL	MYJ	MYNN	YSU	MYJ	MYNN	YSU	MYJ	MYNN	YSU
DX [km]	27, 9	27, 9	27, 9, 3	27, 9	27, 9	27, 9	27, 9	27, 9	27, 9

We ran an ensemble of simulations to test the sensitivity to chosen model parameters. Each ensemble member integrates a different combination of input conditions and PBL scheme (Table 2). We tested the accuracy of the MYJ (Mellor and Yamada 1974; Janjić 1994), YSU (Hong et al. 2006), and MYNN (Nakanishi and Niino 2006) PBL schemes for their reliability of resolving winds in the region (Draxl et al. 2014; Jacobson et al. 2018). While all simulations supported sensitivity evaluation with 27-km and 9-km grid resolution domains, ERA5 with YSU additionally provided a training dataset for the GANs model by including a 3-km domain over Bangladesh (Figure 1). In general, PBL schemes can be categorized as either local or non-local turbulence closure schemes. MYJ and MYNN incorporate local turbulence closure, while YSU uses non-local turbulence closure. Local schemes compute grid-cell quantities from the turbulent flux in vertically neighboring grid cells only. This flux is estimated from the physical gradient. Non-local schemes represent turbulence as a superposition of eddies within the entire column using the K-profile method and can explicitly account for entrainment at the capping inversion (Stull B. 1988).

Due to limited computational resources, we held the remaining model physics constant. Because the validation component of this work only had observations from Bangladesh available, we relied on past sensitivity evaluations made throughout the entire region of study to make informed decisions. These included sensitivity studies in Bangladesh, Thailand, India, Singapore, the Philippines, Indonesia, Laos, and Vietnam. We chose the BMJ cumulus parameterization in the 27-km and 9-km domains. Although the Kain-Fritsch scheme returns better performance for isolated heavy rainfall events, BMJ may better resolve rainfall throughout longer periods, averaged over larger spatial extents, and in conjunction with the YSU scheme in both the wet and dry seasons in the region (Draxl et al. 2014; Du Duc et al. 2019; Athukorala et al. 2021). We chose the WSM6 microphysics package for its reliability throughout Southeast Asia (Tolentino et al. 2016; Nuryanto et al. 2019; Rai and Pattnaik 2019) and performance in sensitivity analysis (Du Duc et al. 2019). We chose the Dudhia shortwave, the RRTMG longwave, the Noah Land Surface Model, and the MM5 Surface Layer parameterizations for their regular use by modeling experts in the region and performance in sensitivity evaluation (Tolentino et al. 2016; Jacobson et al. 2018; Du Duc et al. 2019; Nuryanto et al. 2019; Rai and Pattnaik 2019). The MM5 surface layer scheme is not compatible with the MYJ PBL scheme, so we selected the Eta similarity scheme instead.

Each of the sensitivity simulations ran for a 3-week period from February 1 to February 21, 2016. We chose this period for overlap of the measurement data between the nine independent measurement sites, although the average wind speeds are generally slow. We validated the 9-km, hourly output against observations because this resolution and frequency was supplied to the GANs downscaling algorithm during production. The 3-km output was used to train the GANs discriminator (described in Section 4).

3 WRF Validation

3.1 Validation Methods

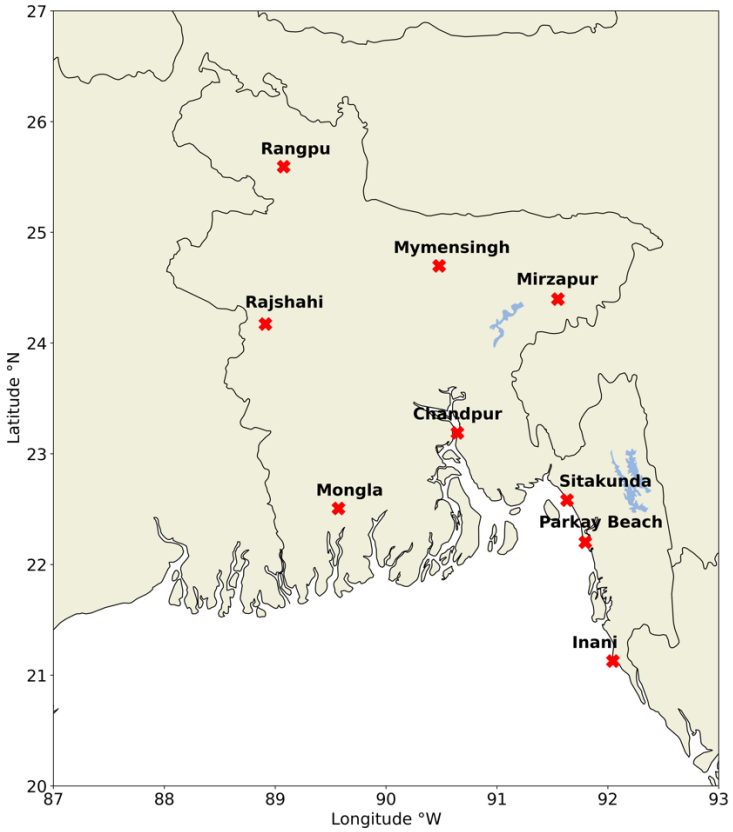

Figure 2. Bangladesh measurement sites

Note: Red Xs indicate the locations of each site.

Atmospheric observations were obtained from nine different measurement sites across Bangladesh (Figure 2). Two sites, Rangpur and Inani, provided sound detection and ranging (SODAR) observations at 10-m intervals from 40 m to 60 m and then 20-m intervals up to 200 m AGL. The remaining seven sites provided meteorological tower measurements with anemometers at 20-m intervals up to 80 m. Wind speeds were collected at the meteorological towers using a mix of NRG class 1 and NRG 40C anemometers. We validated model-output wind speeds from each ensemble member at the measurement heights, respective to the site. Heights for which greater than 50% of the data were missing throughout the 3-week period from February 1 to February 21, 2016, were discarded from the analysis. For this reason, no data were considered from the Inani SODAR or the Parkay Beach meteorological towers.

We compared WRF wind speeds to observed wind speeds at each measurement height using various metrics, including bias, centered root mean square error (cRMSE), RMSE, earth mover's distance (EMD), and the correlation coefficient (R):

$$\text{Bias} = \frac{\sum_i^N \left(V_{WRF_i} - V_{obs_i}\right)}{N} \quad (1)$$

where V is the wind speed and N is the total number of values. Negative bias indicates an underestimation by WRF of the wind speed.

$$\mathrm{cRMSE} = \sqrt{\frac{\sum_{\mathrm{i}}^{\mathrm{N}} \left(\left(\mathrm{V}_{\mathrm{WRF_i}} - \overline{\mathrm{V}_{\mathrm{WRF}}} \right) - \left(\mathrm{V}_{\mathrm{obs_i}} - \overline{V_{obs}} \right) \right)^2}{\mathrm{N}}} \tag{2}$$

$$RMSE = \sqrt{\frac{\sum_i^N \left(V_{WRF_i} - V_{obs_i} \right)^2}{N}} \tag{3}$$

A RMSE score of 0 indicates that all values lie on a 1:1 regression line. Centered RMSE removes the bias from the RMSE. We further considered the correlation coefficient:

$$\mathrm{R} = \frac{\sum_{\mathrm{i}}^{\mathrm{N}} (\mathrm{V}_{\mathrm{WRF_i}} - \overline{\mathrm{V}_{\mathrm{WRF}}})(\mathrm{V}_{\mathrm{obs_i}} - \overline{\mathrm{V}_{\mathrm{obs}}})}{\mathrm{N}\sigma_{\mathrm{WRF}}\sigma_{\mathrm{obs}}} \tag{4}$$

where σ is the standard deviation. A correlation value of 1 indicates a perfect correlation between model-derived and measured values.

The EMD, also known as the Wasserstein Metric, provided an additional quantification of error, for which a built-in SciPy function was utilized (Virtanen et al. 2020). A value of 0 for EMD indicated that the probability density functions for modeled and measured distributions are equivalent.

Both modeled and measured wind speeds were averaged at each hour of the day throughout the period to capture the diurnal evolution. This process was repeated at each measurement height. Separately, wind speeds and validation metrics were averaged at each height in a vertical profile.

3.2 Validation Results

We evaluated the performance of each ensemble member individually throughout the 3-week period. Individual performance metrics were averaged in time to calculate one value for each combination. An ensemble member was awarded a point for having the “best” metric for a given height and location.

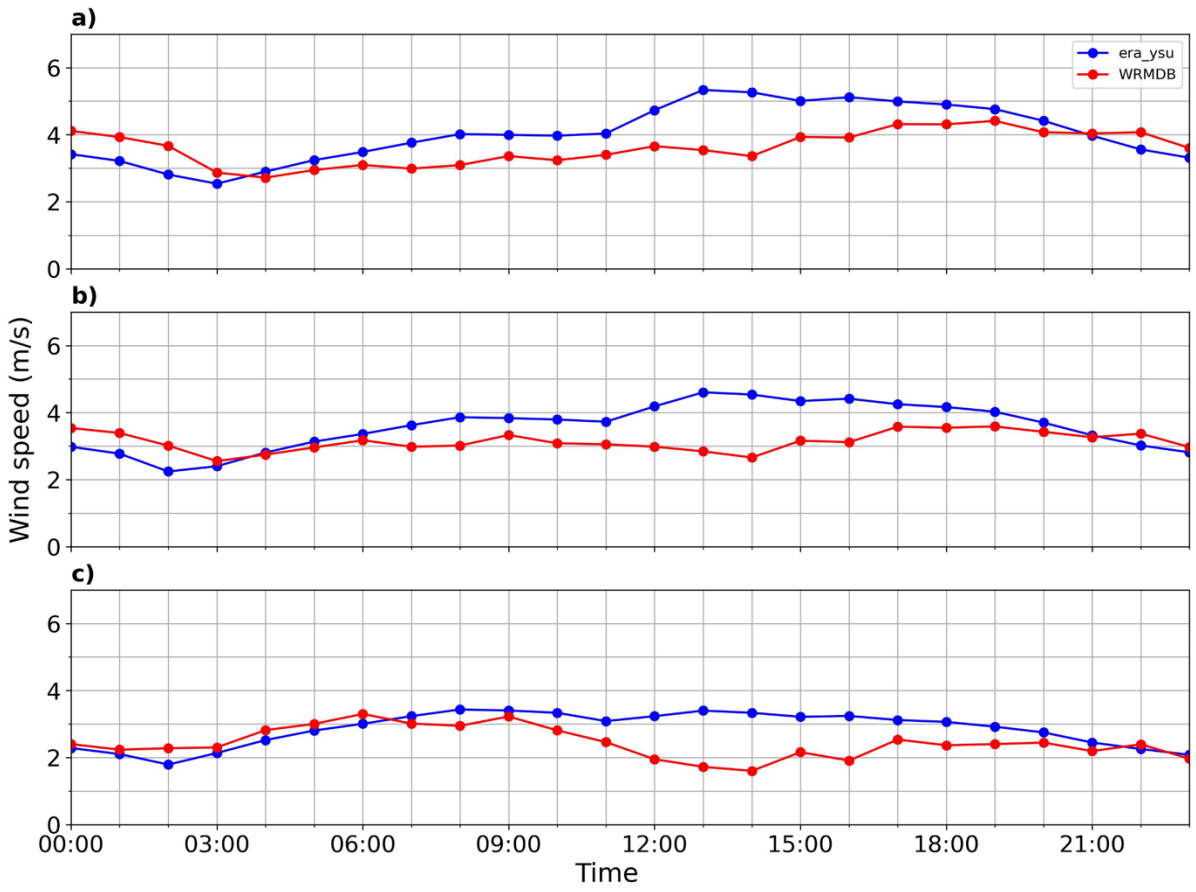

Figure 3. Diurnally averaged wind speeds at Chandpur

Note: The x-axis shows time in UTC, based on hourly values from February 1–21, 2016. The top row is the 60-m measurement height, the middle row the 40-m measurement height, and the bottom row the 20-m measurement height. Wind speeds from WRF using ERA5 with the YSU PBL scheme are shown in blue and from measurements in red.

Wind speeds remained slow throughout the 3-week period and lacked a prominent diurnal cycle (Figure 3). In general, WRF captures the initial wind speed variability between 0000 and 0700 UTC well. However, WRF fails to capture the successive decrease in wind speeds between 1100 and 1500 UTC, or 7 p.m. to 11 p.m. local time, resulting in overestimation by over 1 m s^{-1}. As a nonlocal scheme, YSU may not capture the evening decay of small-scale eddies with the precision of a local scheme. Thus, the overestimation of evening wind speeds may result from a diminished representation of leftover buoyant turbulence from daytime heating.

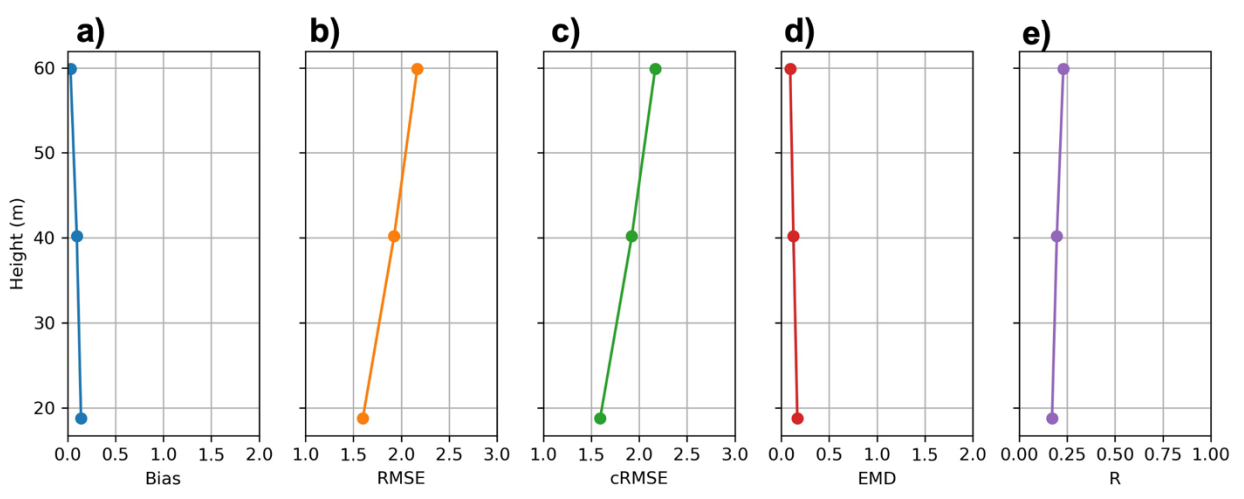

Figure 4. Average metrics by height at Chandpur using ERA5 IBCs with the YSU PBL scheme

Note: From left to right, profiles show bias, RMSE, cRMSE, EMD, and R.

Despite the difficulty in resolving the diurnal evolution of wind speeds, WRF performed well for the purposes of this work. In creating a low-error climatology as opposed to case study analysis, we targeted minimizing bias. For instance, the average bias remained less than 0.25 m s^{-1} and reduced to nearly 0 m s^{-1} at 60 m at the Chandpur measurement site (Figure 4a). Average centered and RMSE increased with height between 1.5 and 2.5 m s^{-1} and the profiles remained nearly indistinct because cRMSE removed bias, which was small to begin with (Figure 4b,c). The mean profile of EMD was small and approached 0 m s^{-1} at 60 m, which indicated that the distribution of wind speeds was properly captured (Figure 4d). However, due to WRF struggling to resolve the meandering of slow wind speeds well (Figure 3), correlation was poor and did not exceed 0.25 (Figure 4e).

To determine the best performing ensemble member, we did not consider the correlation coefficient metric, as the timeseries lacked a prominent diurnal cycle (Figure 3) and because correlation was poor on average (Figure 4e). The tally of points for best-performing metrics revealed that the combination of ERA5 with the YSU PBL scheme has the greatest model skill. ERA5 with the YSU PBL scheme attained a total count of 47 of a possible 100. The next best-performing simulation was ERA5 with the MYJ PBL scheme, with a total count of 23.

We separately considered the 80-m and 140-m heights. 80 m is the most common high-altitude measurement with greater than 50% data availability, existing at six of the nine sites. 140 m is the absolute highest, existing only at the Rangpur SODAR location and is considered for its proximity to modern turbine hub heights. Of note considering the importance of hub-height winds, ERA5 with YSU performed best at both 80-m and 140-m heights.

As a second method for determining the optimal configuration, we determined which simulation performs best throughout Bangladesh by finding the mean of each metric by simulation type across all locations and heights. ERA5 with the YSU PBL scheme remains the top-performing simulation, with an average bias of 0.587 m s^{-1}, cRMSE of 1.951 m s^{-1}, RMSE of 2.070 m s^{-1}, EMD of 0.603 m s^{-1}, and a correlation of 0.230 (Figure 5). The average correlation remained poor for all simulations, with a range from 0.180 to 0.296.

	Bias [m/s]	cRMSE [m/s]	RMSE [m/s]	EMD [m/s]	R
ERA_MYJ	0.802	2.015	2.213	0.832	0.264
ERA_MYNN	0.703	2.152	2.309	0.771	0.246
ERA_YSU	0.587	1.951	2.070	0.603	0.23
FNL_MYJ	0.918	2.108	2.324	0.923	0.193
FNL_MYNN	0.983	2.269	2.501	0.993	0.198
FNL_YSU	0.683	2.065	2.190	0.685	0.18
MERRA_MYJ	0.818	2.099	2.276	0.828	0.244
MERRA_MYNN	0.949	2.179	2.398	0.966	0.28
MERRA_YSU	0.747	2.025	2.169	0.750	0.296

Figure 5. Mean metrics for each ensemble member

Note: Rows delineate each ensemble member with titles showing the IBC (ERA, FNL, or MERRA) and PBL scheme (MYJ, MYNN, or YSU). Columns delineate the metric applied. Darker greens indicate lower error, and darker blues indicate higher correlation.

Model skill varies regionally. To determine which measurement location may represent the rest of the domain, each site is categorized by land use type. Land use affects the physical properties of the surface, including the roughness length, and wind speed differences between PBL schemes can be amplified in regions with a higher roughness length (Hahmann et al. 2020). Each model grid cell consists of a variety of land use types that total to 100% coverage. The most frequently occurring land use index throughout the domain, apart from ocean, is Index 1, which corresponds with Evergreen Broadleaf Forest. Accordingly, the model grid cell closest to the Chandpur meteorological tower may act as a proxy for the rest of the domain by having both the highest percentage of Evergreen Broadleaf Forest and less than 50% missing data. The highest measurement height at Chandpur considered here was 60 m, and results indicated that ERA5 with both the MYJ and YSU PBL schemes performed well (Table 3). The simulation using MYJ had lower bias by 0.026 m s^{-1} and lower cRMSE and RMSE by 0.048 m s^{-1}. Alternatively, the simulation using the YSU PBL scheme had lower EMD by 0.096 m s^{-1}. These differences between ERA_MYJ and ERA_YSU were very small.

Table 3. Mean Metrics at the 60-m Measurement Height at the Chandpur Meteorological Tower

Note: Rows delineate the ensemble member by IBC and PBL scheme. Columns indicate the metric calculated.

Ens. Mem.	Bias [m s^{-1}]	RMSE [m s^{-1}]	cRMSE [m s^{-1}]	EMD [m s^{-1}]	R
ERA_MYJ	0.004	2.119	2.119	0.189	0.300
ERA_YSU	0.030	2.167	2.167	0.093	0.230

Because two of these three approaches clearly preferred ERA_YSU, and the third had an effective tie between ERA_YSU and ERA_MYJ, we selected ERA5 with the YSU PBL scheme for production runs. Thus, this combination was simulated for the 3-week period from February 1–21, 2016, at 3-km grid resolution with 15-min output temporal resolution, for GANs discriminator training. Additionally, we ran production-level simulations for January 1, 2007, to December 31, 2021, using the 27-km and 9-km domains with an hourly output frequency.

4 GANs Setup

Our GANs model generally follows the description in Stengel et al. (2020) with expanded convolutional kernels for handling higher dimensional data with temporal super resolution. The GAN consists of a generator, which generates enhanced output from lower resolution input, and a discriminator, which discriminates between simulated output and reference high-resolution data. During training, reference high-resolution data is coarsened and fed into the generator. The generator then outputs simulated high-resolution data. Meanwhile, the discriminator is provided with balanced sets of reference and simulated high-resolution data and learns to discern between the two (Figure 6).

The training performance of the generator is measured using a combination of content loss and adversarial loss. Content loss is computed from the mean absolute error (MAE) between the simulated high-resolution data and reference high-resolution data. The adversarial loss for the generator is computed as the sigmoid cross entropy (Wang et al. 2022) of the misclassification rate of the discriminator when provided with a set of simulated data from the generator (the generator performs best when the discriminator has a high misclassification rate). The discriminator loss is the sigmoid cross entropy on a balanced set of reference and simulated observations.

We provided mid-network injection of high-resolution topography in the GANs so that high-resolution rather than low-resolution topography can be used for training. This mid-network injection concatenates the high-resolution topography to the training data immediately following the spatial enhancement portion of the network.

We maintain an extensive codebase to implement easily customizable GAN architectures, handle data extraction, batch handling, and model training, and to distribute the forward passes of input data through the GAN across multiple nodes. This codebase is released publicly as the Super Resolution for Renewable Resource Data (sup3r) package and is installable through the Python package index (Buster et al. 2022).

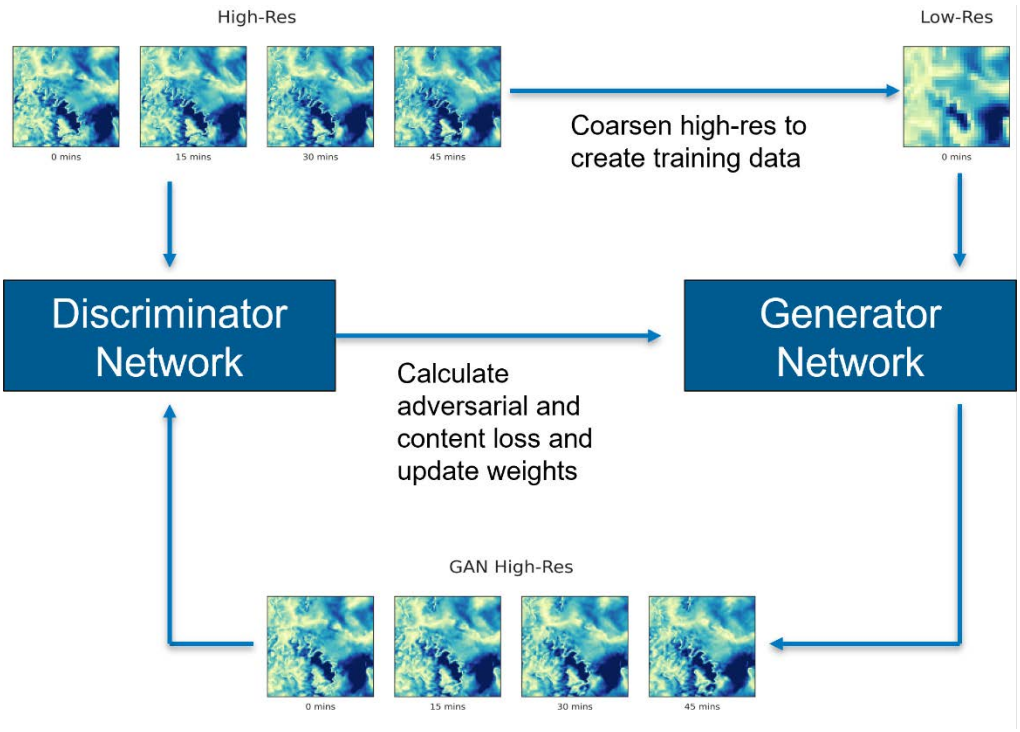

Figure 6. GAN training flow

Note: Reference high-resolution data is coarsened, this low-resolution data is fed to the GAN, the GAN generates simulated high-resolution data. Content and adversarial loss is computed and used to update the model weights.

We used three separate models for: (1) wind speed and wind direction (2) pressure, and (3) temperature, due to memory limitations. The first model was trained on zonal/meridional (U/V) wind components and surface topography. The second model was trained on pressure and topography. The third model was trained on temperature and topography. Each of these main models consists of a spatial enhancement GAN and a temporal enhancement GAN, except for the pressure model, which employs linear interpolation for temporal enhancement. The spatial and temporal GANs were trained separately but were linked when performing the forward passes. The linear interpolation model for pressure was selected for temporal enhancement because the primary use of pressure data in wind resource analyses is to capture low-frequency oscillations. A GAN model would mainly help capture small-magnitude, high-frequency fluctuations that have a negligible effect on wind power. A summary of the wind, pressure, and temperature models is shown in Table 4.

All models were trained with WIND Toolkit data (Draxl et al. 2015) for 2007, using observations sampled from a 500x500 grid (nominally 1,000x1,000 km) (Figure 7). Although the training observations were selected from North America, they included sufficiently diverse conditions to enable model prediction outside of the training domain. This approach is supported by the out-of-sample statistical comparisons in Section 6. Training was performed on the Eagle high-performance computing system at NREL. Each training epoch consisted of 100 batches, with 16 observations per batch. Spatial enhancement GANs use observations with a single timestep and 30x30 high-resolution spatial extent. Temporal enhancement GANs use observations with 72 timesteps and 30x30 spatial extent. The spatial regions for these observations are randomly selected from the full spatial domain with uniform probability. The temporal region for each observation is randomly selected with monthly weights based on the validation loss. For instance, if during a given training epoch the model is performing poorly on summer observations, more observations will be selected from the summer for the next epoch. This data-

centric training approach helps ensure that the model will perform over a wide range of season-specific weather conditions.

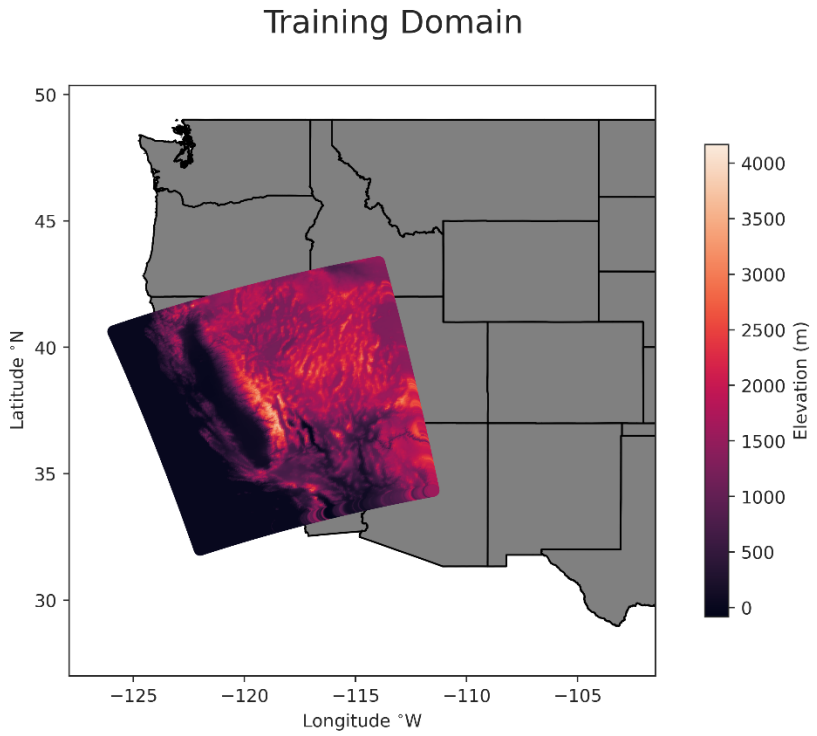

Figure 7. Training domain

Note: All training samples were taken from the region shown.

Prior to training, observations were coarsened according to the planned spatiotemporal enhancement of the WRF data over Southeast Asia. The native resolution of the WIND Toolkit data is 2-km 5-minute, so every third timestep was selected to obtain 2-km 15-minute data. The spatial domain was coarsened by a factor of three, and the temporal domain was coarsened by a factor of four, resulting in a resolution of 6 km hourly. Thus, each spatial enhancement batch has a 16x10x10 low-resolution data volume per input feature and each temporal enhancement batch has a 16x10x10x18 low-resolution data volume per feature, where the first dimension was the number of observations per batch, the second was the north-south spatial axis, the third was the east-west spatial axis, and the fourth dimension (only for the temporal enhancement model) was the temporal axis.

The coarsened data was also spatially smoothed using a Gaussian filter to more accurately represent the kinetic energy spectrum of a reference low-resolution simulation (Skamarock 2004). The smoothing parameter was determined by matching the kinetic energy spectrum of reference low-resolution WRF simulations with coarsened and smoothed high-resolution WRF simulations over the same time period and spatial domain. We executed this approach due to the mismatch between the effective resolution and the grid resolution of WRF simulations (Skamarock 2004). All input features were smoothed except for topography.

Table 4. Model Summary

Note: All models were trained on the Eagle high-performance computing system at NREL

Model Type	Training Features	Spatial GAN Training Time	Temporal GAN Training Time
Wind	U/V at 10, 40, 80, 100, 120, 160, 200 meters + topography	24 hours, 10,099 epochs	1 week, 7,000 epochs
Pressure	Pressure at 0, 100, 200 meters + topography	5 hours, 10,099 epochs	N/A
Temperature	Temperature at 10, 40, 80, 100, 120, 160, 200 meters + topography	10 hours, 10,099 epochs	1 week, 7,000 epochs

5 GAN Data Generation

To generate the final super-resolved data, we combined the WRF data for all domains (Figure 1). This synthesis provided a single super-resolved domain as output. The 27-km parent domain was first bilinearly interpolated to 9 km. The data from each smaller domain was inserted into the larger domain with spatial bilinear interpolation around the domain boundaries to blend any seams. This border interpolation was performed over a padding width of 100 grid points from the edge of all the smaller domains. The final 9 km combined domain consisted of 807x987 grid points.

Each year of WRF data, consisting of 807x987 grid points and 8,760 timesteps, was passed through the GAN. This data passage required splitting the full data volume into overlapping chunks with 150x150 grid points and 30 timesteps. These chunks were made to overlap in both space (15 grid points) and time (5 timesteps) so they can be passed through the GAN separately and then stitched together to construct the full output domain. Approximately 28,000 data chunks were distributed across 2,000 nodes and fed through the GAN. This processing required approximately 4,000 node hours and 8 hours of total wall-clock time per year of input data. The overlapping regions of the output chunks were cropped, and the cropped chunks were combined into the full output domain. After combining each output chunk, the enhanced data consisted of 2,421 x 2,961 grid points and 35,040 timesteps per year.

This output data was trimmed to exclude areas where the 27-km data was interpolated to 9 km, as well as areas outside the offshore EEZ. This final domain is shown in Figure 8.

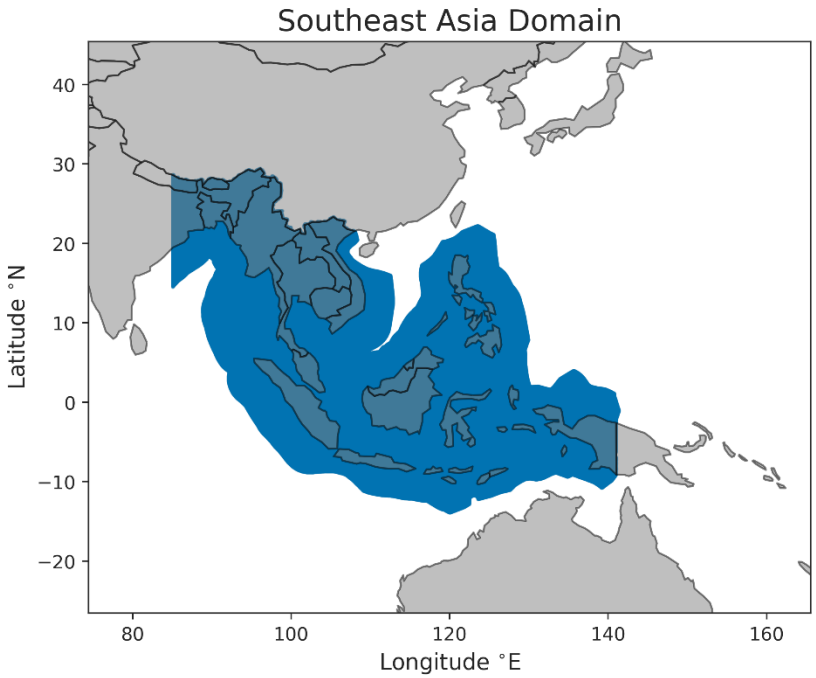

Figure 8. The final trimmed domain, including the offshore EEZ (blue shading)

6 GAN vs. WRF Validation

As described in Section 2, overlapping low-resolution (9 km, hourly) and high-resolution (3 km, 15-minute) datasets using two separate WRF runs at different grid resolutions were generated over Bangladesh to use for model validation. The low-resolution WRF data for this domain was fed through the GAN to generate simulated high-resolution data to use for comparison with the reference high-resolution WRF simulation. Because the GAN downscaling is an inherently under-constrained problem, with many high-resolution solutions that can be coarsened to the same low-resolution input, we used statistical metrics for validation. The complete set of metrics included the turbulent kinetic energy spectrum (Figure 9), wind direction time derivative distribution, wind speed time derivative distribution, longitudinal velocity gradient distribution, and vorticity distribution (Figure 10). We also evaluated the pressure frequency spectrum and longitudinal gradient distribution (Figure 11) and the temperature time derivative distribution and longitudinal gradient distribution (Figure 12).

For each statistical metric, we compared the low-resolution input, GAN output, reference high-resolution WRF output, coarsened high-resolution WRF output, and spatiotemporal bilinear interpolation of the low-resolution WRF output. In all cases, an ideal result would be for the GAN high-resolution output to match the true high-resolution field. The frequency spectrum over time derivative distribution for pressure is important because it represents the large-magnitude low-frequency changes rather than small-magnitude high-frequency changes. The primary use of pressure data to estimate air density for wind power estimation makes large-magnitude lower-frequency content more important.

Across all metrics, the GAN downscaling outperformed the baseline spatiotemporal bilinear interpolation. The interpolation failed to capture the kinetic energy across wavenumbers absent in the coarse data and the broadness of the time derivative, velocity, and vorticity distributions. In contrast, the GAN output better captured this broadening and the kinetic energy at high wavenumbers. Spatial and temporal variations were well captured across wind, pressure, and temperature fields by the GAN. A summary of the errors between the GAN output and reference high-resolution WRF simulation is shown in Table 5. This table includes the MAE between the GAN output and the reference high-resolution data and the MAE between the interpolated baseline and the reference high-resolution data. The improvement of GANs over the interpolated baseline ranged between 20% and 80%.

The metrics involving spatial derivatives consistently showed smaller errors than the metrics involving temporal derivatives. This higher difficulty in capturing temporal dynamics is reflected in the slower training and increased network complexity for temporal GANs.

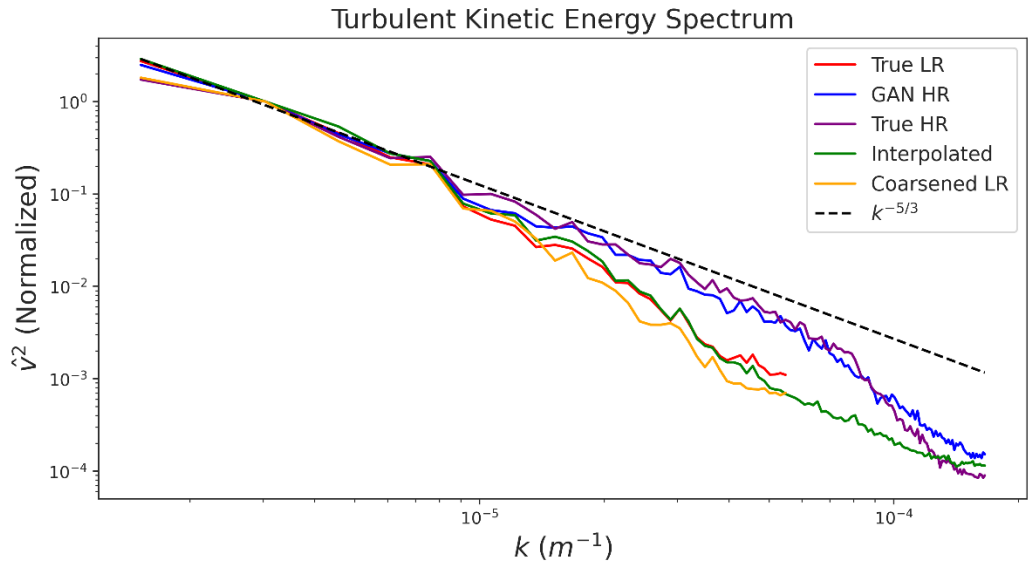

Figure 9. Turbulent Kinetic Energy spectrum for windspeed at 100 m ($\hat{v}^2$ is the Fourier transform of v^2)

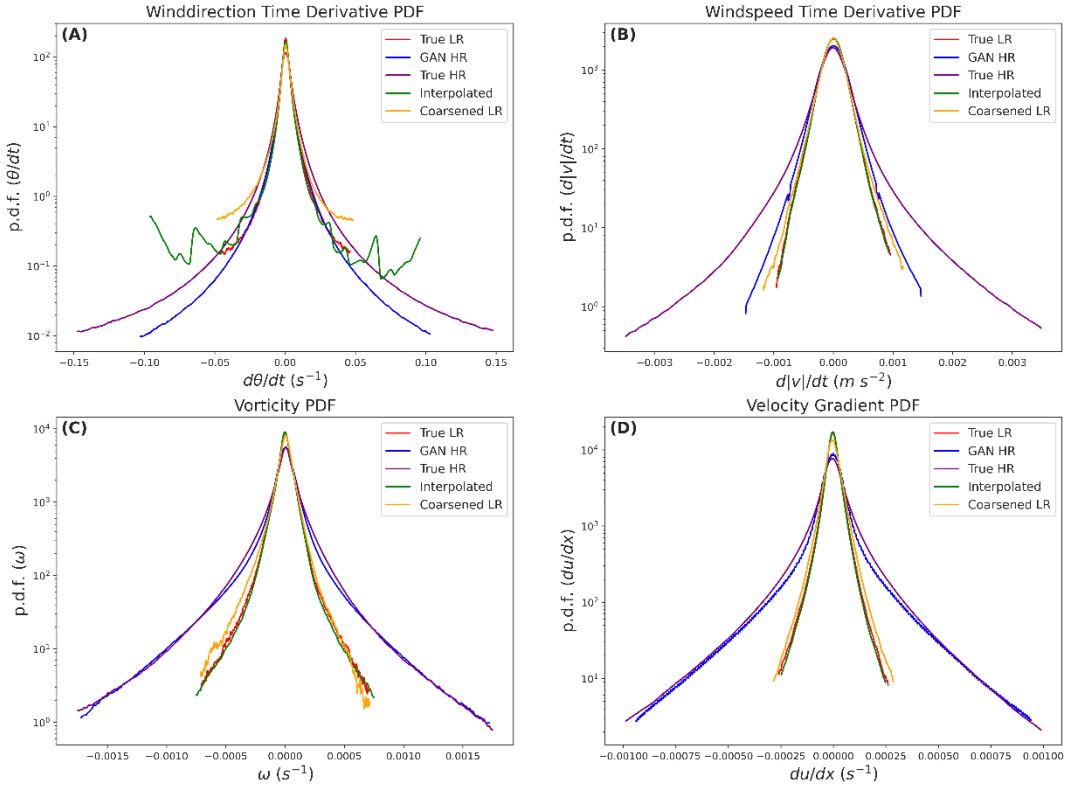

Figure 10. (A) Wind direction time derivative distribution, (B) windspeed time derivative distribution, (C) vorticity distribution, (D) velocity gradient distribution

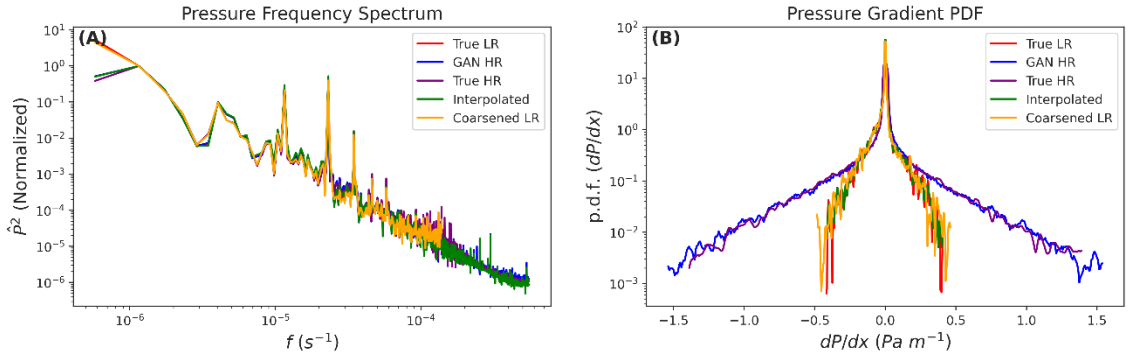

Figure 11. (A) The pressure frequency spectrum, and (B) the pressure gradient distribution

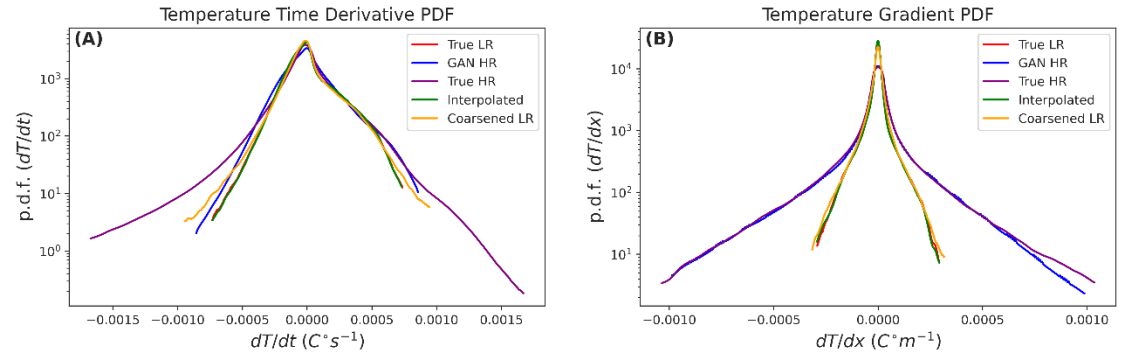

Figure 12. (A) Temperature time derivative distribution, and (B) temperature gradient distribution

Table 5. MAE for GAN vs. Reference High-Resolution (middle) and MAE for Interpolated Baseline vs. Reference High-Resolution (right)

Metric	MAE (GAN)	MAE (Interpolation)
TKE Spectrum (Normalized)	4.7e-3	6.6e-3
Velocity Gradient (s^{-1})	3.68e-8	8.76e-8
Vorticity (s^{-1})	4.36e-8	1.51e-7
Wind direction time derivative (deg/s)	3.99e-6	5.5e-6
Windspeed time derivative (m/s^2)	1.34e-7	1.78e-7
Pressure Frequency Spectrum (Pa^2)	3.6e-4	8.2e-4
Pressure Gradient (Pa/m)	1.98e-5	4.34e-5
Temperature time derivative (K/s)	6.83e-8	7.02e-8
Temperature Gradient (K/m)	9.6e-9	5.5e-8

7 GAN vs. Site Measurements

We used the same atmospheric observations obtained from Bangladesh to validate the wind speed results from GANs (Jacobson et al. 2018). The validation utilized temporally overlapped data between January and December 2017. We performed the same analysis described in Section 3.1. We did not consider the measurements at Inani Beach and Sitakunda, due to the lack of temporal overlap. We did not include the measurements taken at Rangpur with Triton Wind Profilers,[4] due to greater than 50% of missing data at all heights for 2017. Therefore, the observations from six sites (Chandpur, Mirzapur, Mongla, Mymensingh, Parkay Beach and Rejshahi) contributed to validating the GANs results, and each tower provides measurements at multiple heights above ground level.

Diurnally averaged wind speeds were compared by height at each six sites (Figure 13 and Figure 14). We only show the results at Rajshahi and Parkay Beach, as they represent coastal and land-based patterns, which are similar at others. Wind speeds remained moderate throughout 2017. Most sites did not have a strong diurnal cycle except for Parkay Beach, which had a slight acceleration between 0 and 12 hours UTC (6 a.m. to 6 p.m. local time). In general, GANs captures wind speed variations well compared to in-situ observations.

The average metrics by height indicated that GANs had a small deviation from the actual measurements and matched the distribution of wind speed (Figure 15–Figure 20). The bias at most sites remained between -0.5 and 0.5 m s^{-1} except for the wind speed at 40 m on Parkay Beach. RMSE and cRMSE grew with height, and the values fluctuated between 1.5 and 2.5 m s^{-1}. Most EMD values were smaller than 0.5 at 60 m and around 0.25 at 80 m. Most correlation values were higher than 0.5 except for at Mirzapur and Rajshahi.

The annual wind speed profile and its standard deviation show how GANs matched the reference wind speeds at a specific height (Figure 21). The profile of GANs data overlaps well with the measurements, demonstrating that the GANs model captured the mean and variation of wind speeds at multiple heights.

[4] Vaisala Triton Sonic Wind Profiler.

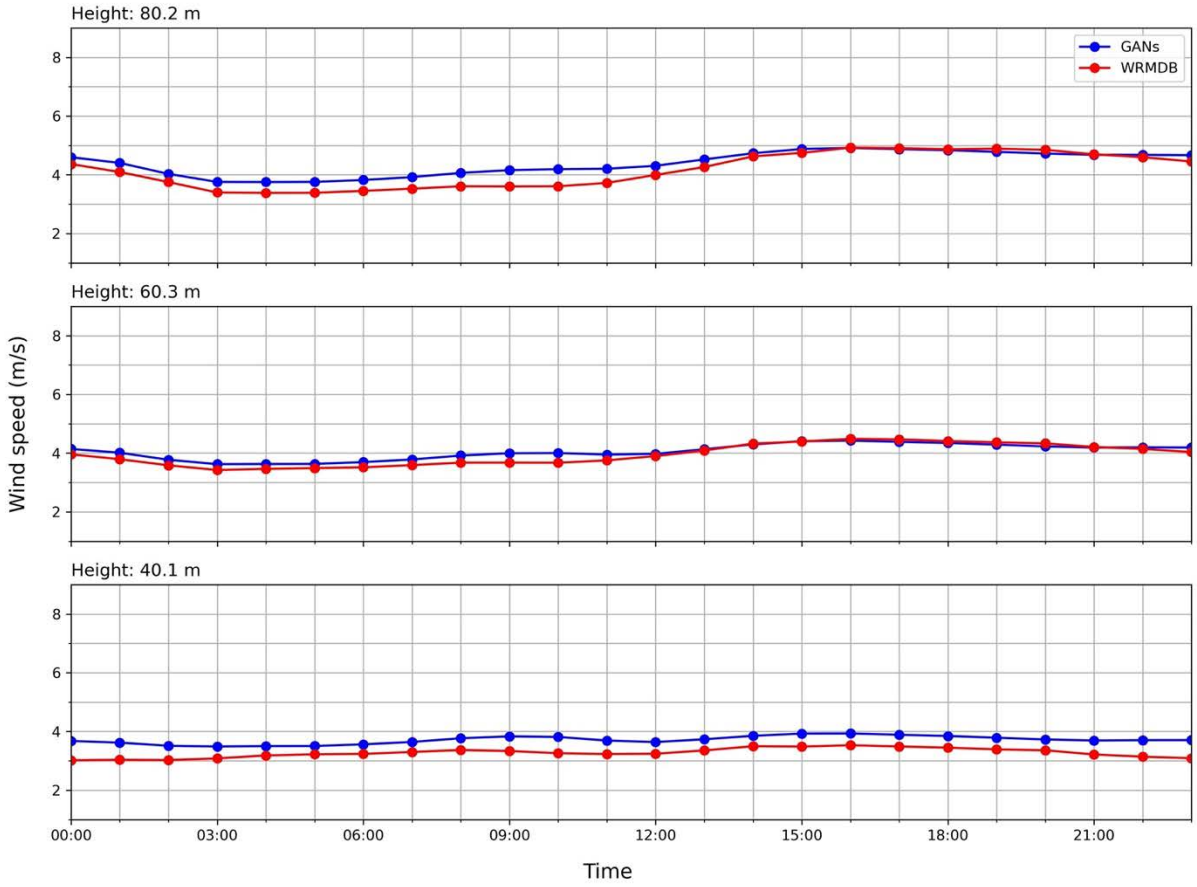

Figure 13. Diurnally averaged wind speeds at Rajshahi

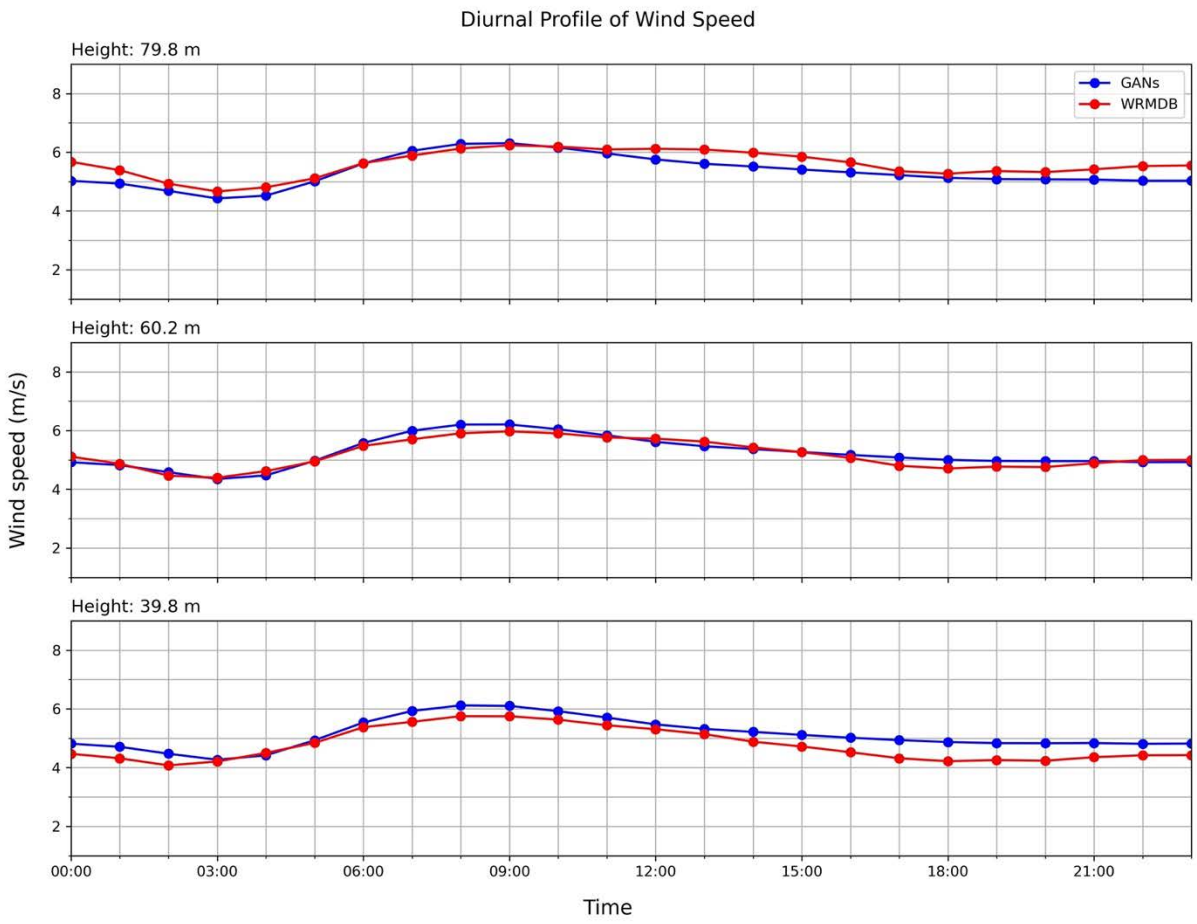

Figure 14. Diurnally averaged wind speeds at Parkay Beach

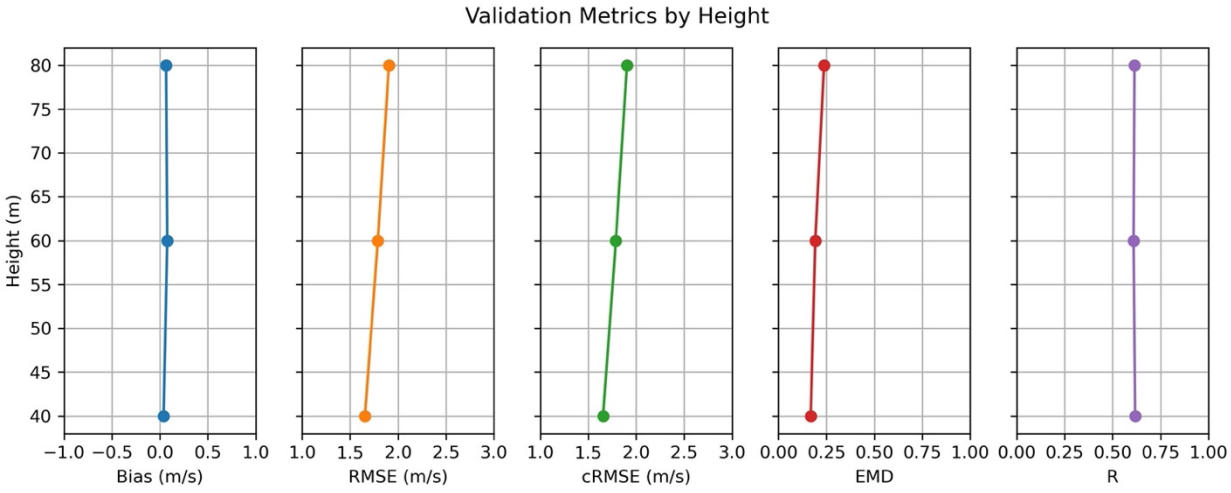

Figure 15. Average metrics by height at Mongla

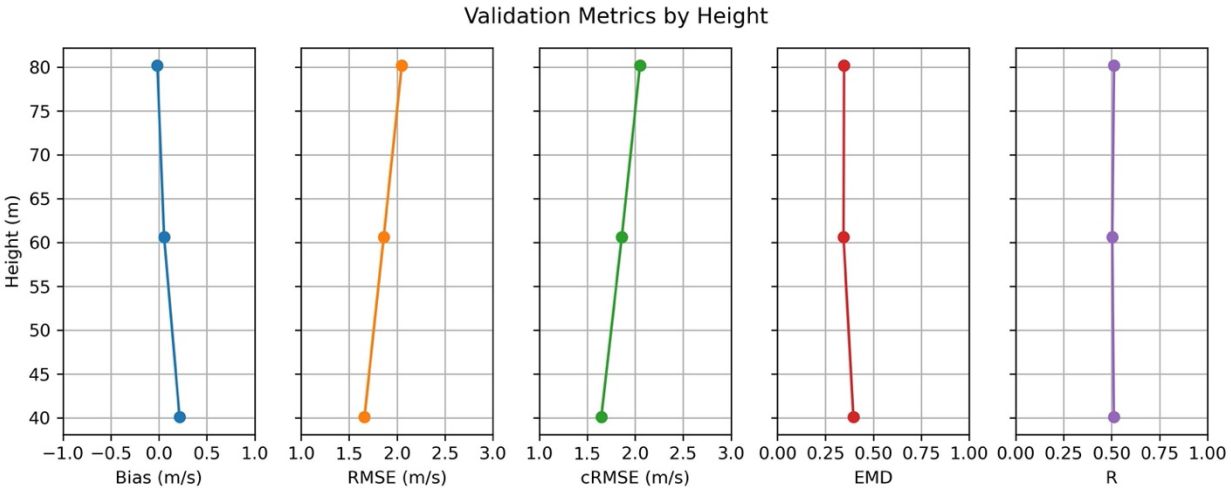

Figure 16. Average metrics by height at Mymensingh

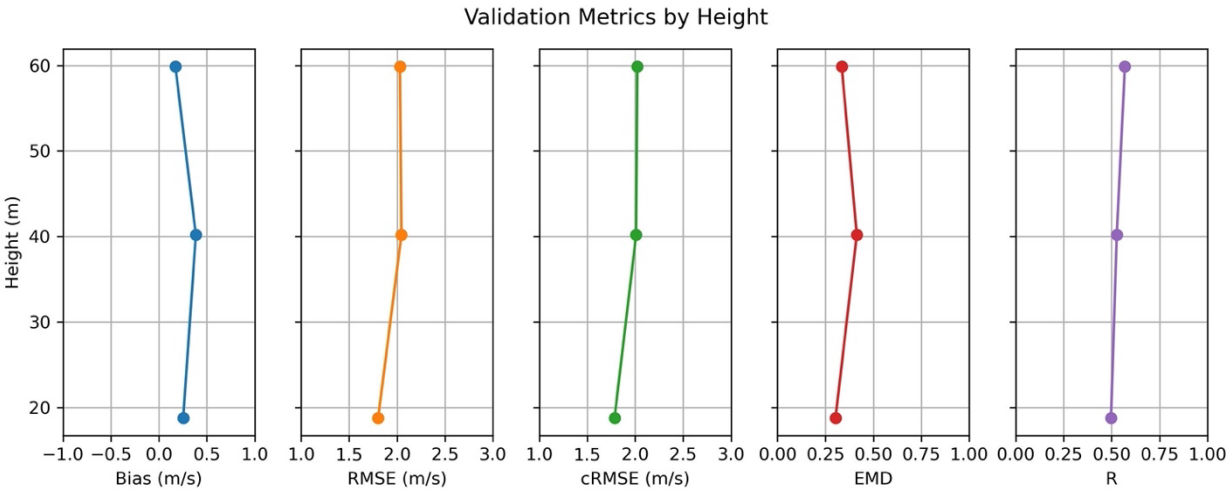

Figure 17. Averaged metrics by height at Chandpur

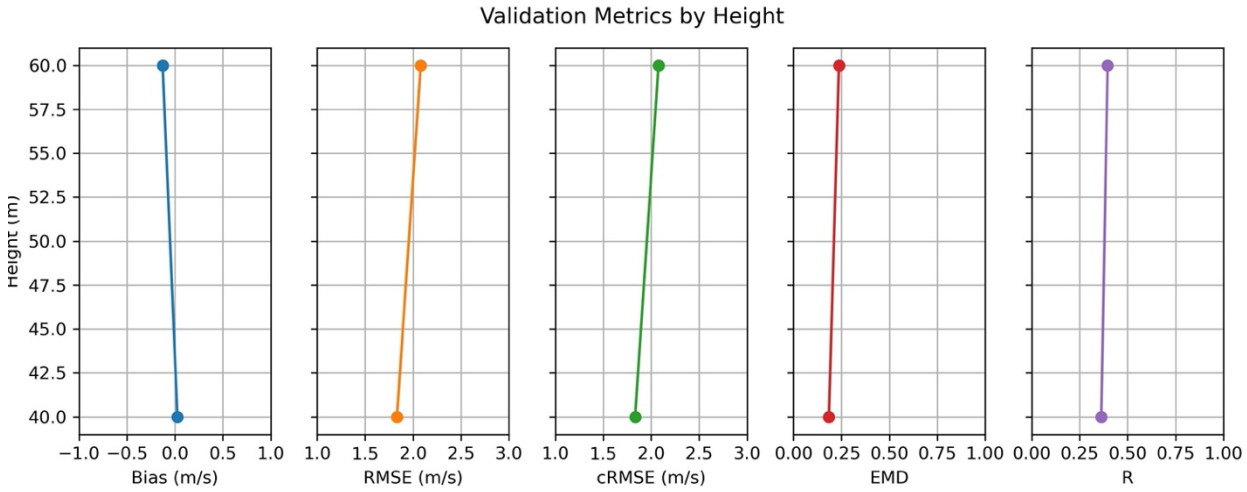

Figure 18. Averaged metrics by height at Mirzapur

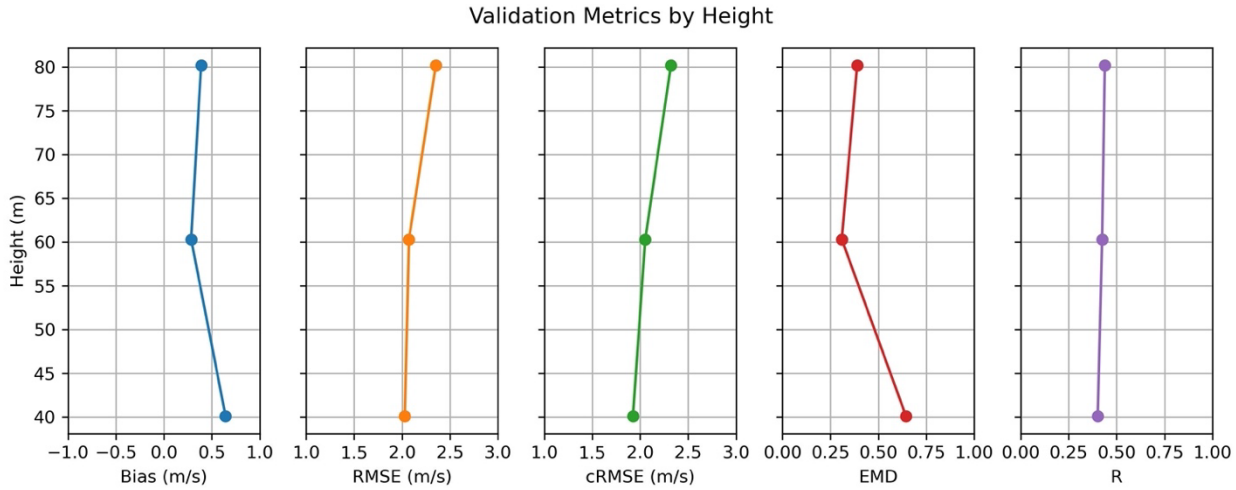

Figure 19. Averaged metrics by height at Rajshahi

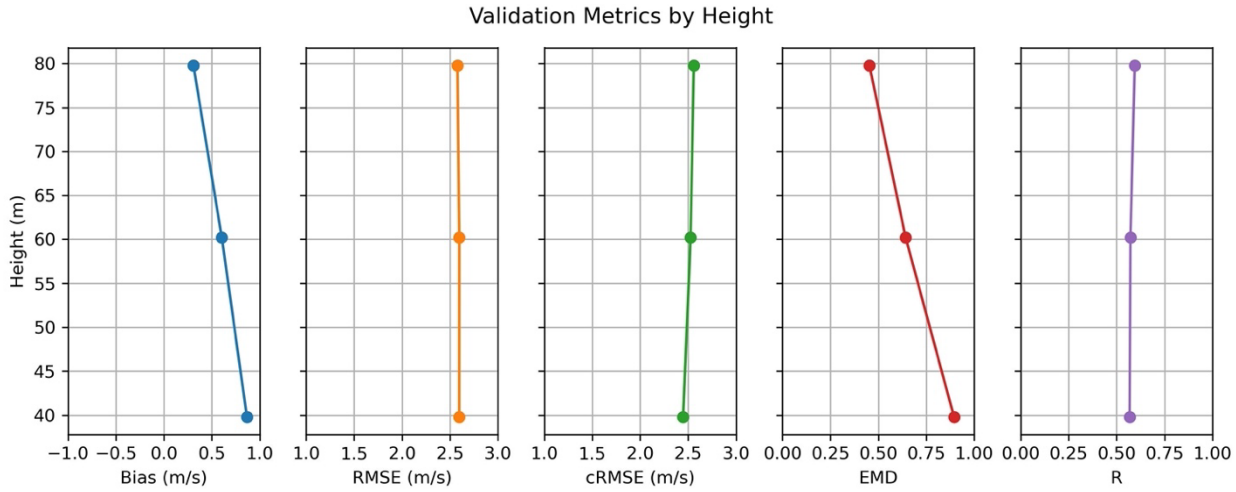

Figure 20. Averaged metrics by height at Parkay Beach

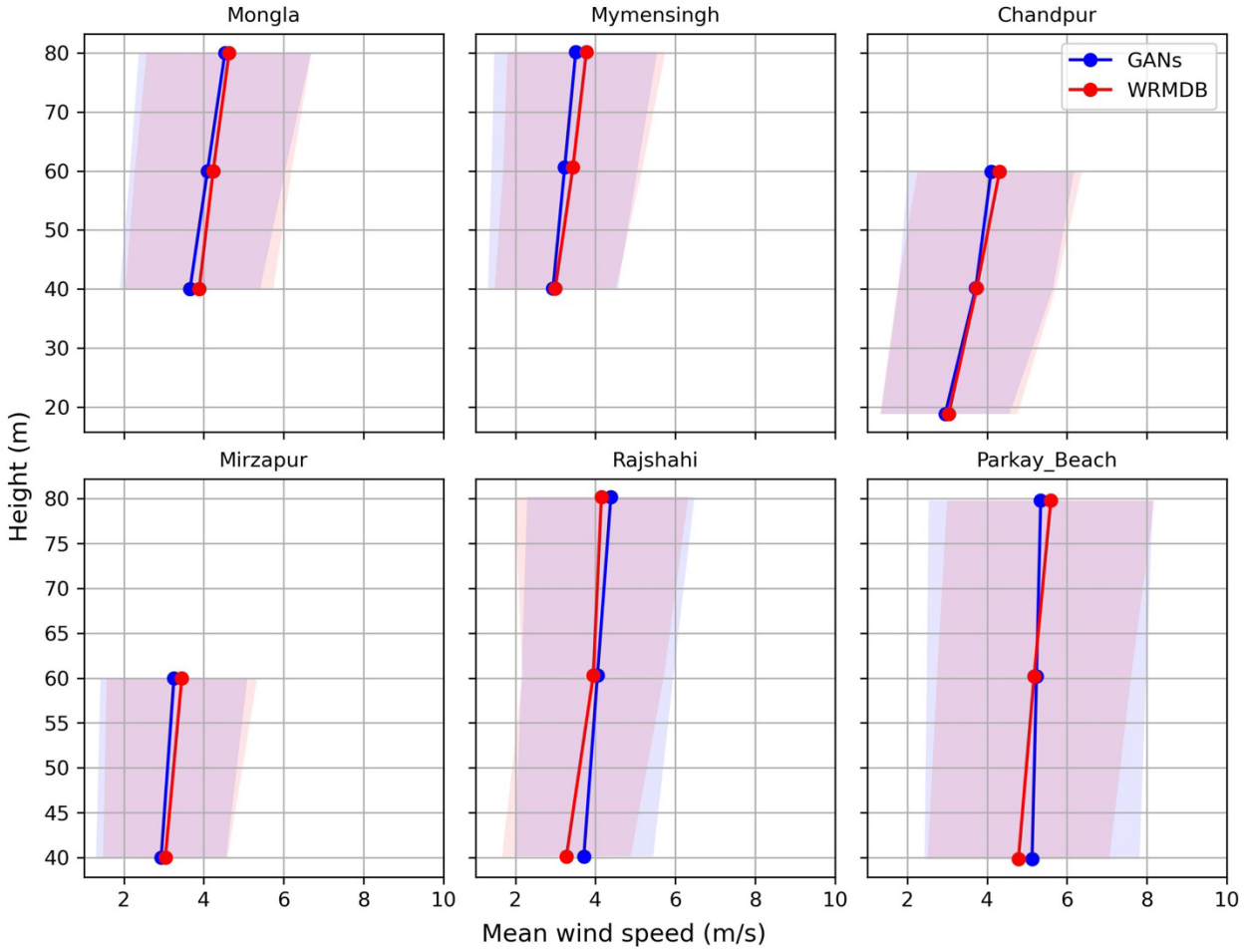

Figure 21. Annual wind speed profile by height where shaded area shows the standard deviation

8 Conclusions

The results herein support the use of the wind data created by the hybrid Numerical Weather Prediction + generative machine learning approach for the Southeast Asia domain. An optimal selection of boundary conditions and physics options was determined for WRF simulations based on comparisons to tower and SODAR observations in Bangladesh for a 21-day period during February 2016. Using these physics options, low-resolution (9-km grid) WRF simulations were carried out over Southeast Asia. The domains for these simulations were stitched together to provide a single final domain. The GANs were trained on coarsened WIND Toolkit data, and the stitched domain was enhanced by 3x along each spatial dimension and 4x along the temporal dimension. High-resolution wind, temperature, and pressure data were generated over Southeast Asia to provide the necessary variables for estimating wind turbine electricity generation potential.

We demonstrated that GANs can be used to substantially reduce computational costs while producing high-accuracy wind data appropriate for modeling and site evaluation. Using WRF alone, a month of data at 15-minute, 3-km resolutions for the validation region over Bangladesh would use 7,200 node hours on NREL's high performance computing system. Our hybrid approach using WRF to downscale to hourly 9 km data followed by GANs downscaling to 15-minute, 3-km took just 450 node hours. Thus, we report a 16x speed up for GANs-based downscaling over WRF alone.

The final data set (Figure 26) is publicly available for download on the Renewable Energy Data Explorer,[5] a geospatial interface that provides options to download meteorological variables at multiple hub heights for each year in the data set. Additional functionality allows the user to select different temporal resolutions (at 15-, 30-, or 60-minute intervals) and format data to local time. NREL also provides application programming interfaces to access larger amounts of data with Python or other programming languages.[6]

[5] Renewable Data Explorer: https://www.re-explorer.org.
[6] Download at https://developer.nrel.gov/docs/wind/wind-toolkit/southeast-asia-wind-download/.

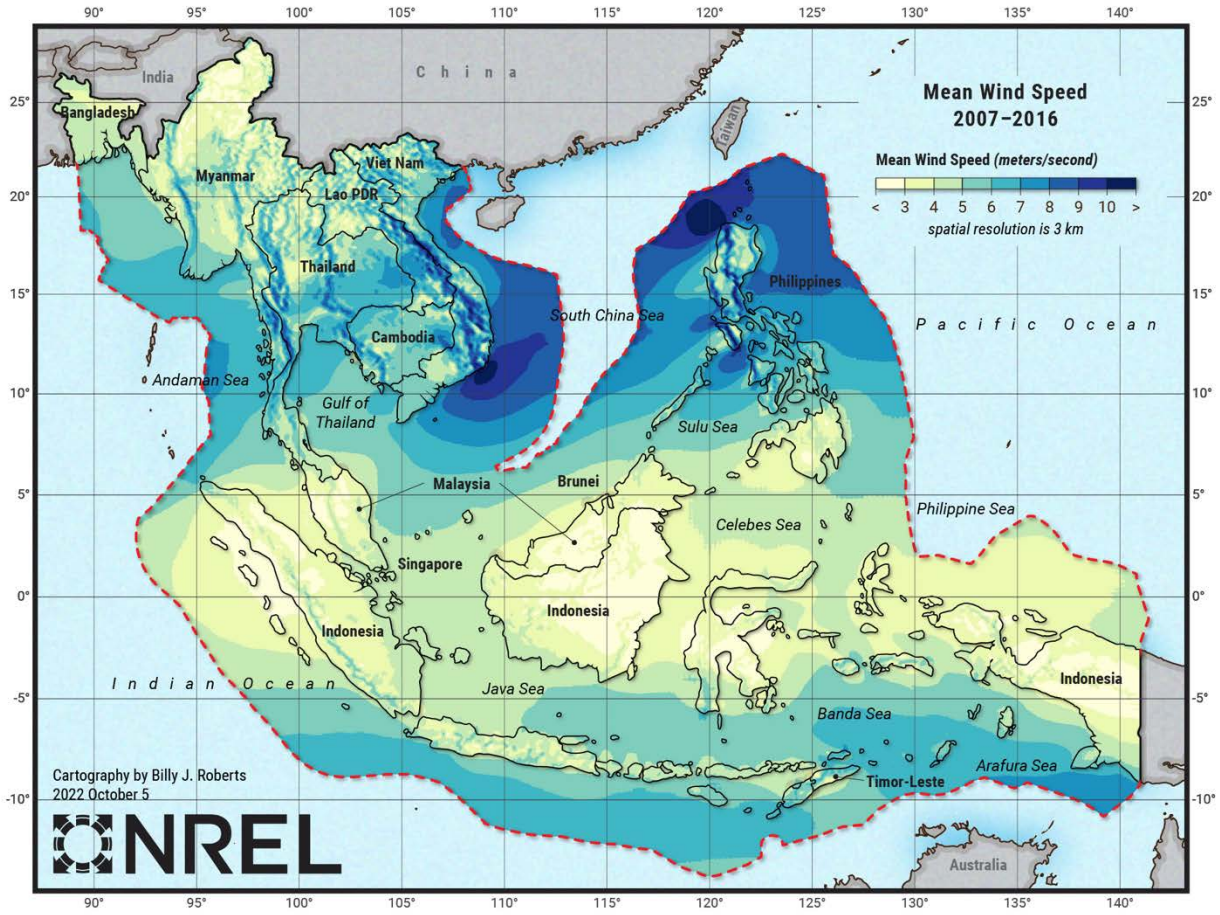

Figure 22. Long-term mean wind speed of the final data set (100 m above ground level)

Because many high-resolution fields (or realizations) can be consistent with a given low-resolution field, the appropriate interpretation of the simulated output fields from GANs is not whether they match the reference high-resolution data exactly, but rather that they are one of the possible valid high-resolution fields corresponding to the low-resolution input. The GANs output fields capture the physical and atmospheric properties and accurately match the statistical distribution of the high-resolution training data. GANs achieve this objective through the probability distribution comparisons between reference high-resolution data and simulated outputs for temporal derivatives and spatial gradients across windspeed, wind direction, temperature, and pressure. Across validation metrics, we saw 20%–80% reduction of MAE over baseline interpolation. We also saw good agreement between the GANs output and site measurements, with a bias within -0.5 and 0.5 m s^{-1} for all except 40-m wind speed at Parkay Beach.

Python code developed through this project for feature engineering, data handling, model training, and inference is made publicly available on GitHub as the sup3r package (Buster, Benton, Glaws, & King, 2022).

References

Athukorala, R., T. Thol, P. Neluwala, M. Petri, S. Sengxeu, L. Lattada, S. Keomanivong, and V. Sithivong, 2021: Evaluating the Performance of a WRF Physics Ensemble in Simulating Rainfall over Lao PDR during Wet and Dry Seasons. *Advances in Meteorology*, **2021**, e6630302, https://doi.org/10.1155/2021/6630302.

Buster, G., B. Benton, A. Glaws, and R. King, 2022: *sup3r (Super Resolution for Renewable Resource Data)*. National Renewable Energy Lab.(NREL), Golden, CO (United States),.

Draxl, C., A. Purkayastha, and Z. Parker, 2014: Wind Resource Assessment of Gujarat (India). NREL/TP-5000-61741, 1149654, https://doi.org/10.2172/1149654.

Draxl, C., A. Clifton, B.-M. Hodge, and J. McCaa, 2015: The Wind Integration National Dataset (WIND) Toolkit. *Applied Energy*, **151**, 355–366, https://doi.org/10.1016/j.apenergy.2015.03.121.

Du Duc, T., C. Hoang Duc, L. R. Hole, L. Hoang, H. Luong Thi Thanh, and H. Mai Khanh, 2019: Impacts of Different Physical Parameterization Configurations on Widespread Heavy Rain Forecast over the Northern Area of Vietnam in WRF-ARW Model. *Advances in Meteorology*, **2019**, e1010858, https://doi.org/10.1155/2019/1010858.

Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, **30**, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1.

Gill, A. E., 1982: *Atmosphere-Ocean Dynamics*. Academic Press, 684 pp.

Gómez, B., and G. Miguez-Macho, 2017: The impact of wave number selection and spin-up time in spectral nudging. *Quarterly Journal of the Royal Meteorological Society*, **143**, 1772–1786, https://doi.org/10.1002/qj.3032.

Hahmann, A. N., C. L. Vincent, A. Peña, J. Lange, and C. B. Hasager, 2015: Wind climate estimation using WRF model output: method and model sensitivities over the sea. *International Journal of Climatology*, **35**, 3422–3439, https://doi.org/10.1002/joc.4217.

——, and Coauthors, 2020: The making of the New European Wind Atlas – Part 1: Model sensitivity. *Geoscientific Model Development*, **13**, 5053–5078, https://doi.org/10.5194/gmd-13-5053-2020.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146**, 1999–2049, https://doi.org/10.1002/qj.3803.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes. *Monthly Weather Review*, **134**, 2318–2341, https://doi.org/10.1175/MWR3199.1.

Jacobson, M., C. Draxl, T. Jimenez, B. O'Neill, T. Capozzola, J. A. Lee, F. Vandenberghe, and S. E. Haupt, 2018: Assessing the Wind Energy Potential in Bangladesh: Enabling Wind Energy Development with Data Products. NREL/TP--5000-71077, 1476253, https://doi.org/10.2172/1476253.

Janjić, Z. I., 1994: The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer, and Turbulence Closure Schemes. *Monthly Weather Review*, **122**, 927–945, https://doi.org/10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2.

Mellor, G. L., and T. Yamada, 1974: A Hierarchy of Turbulence Closure Models for Planetary Boundary Layers. *Journal of the Atmospheric Sciences*, **31**, 1791–1806, https://doi.org/10.1175/1520-0469(1974)031<1791:AHOTCM>2.0.CO;2.

Nakanishi, M., and H. Niino, 2006: An Improved Mellor–Yamada Level-3 Model: Its Numerical Stability and Application to a Regional Prediction of Advection Fog. *Boundary-Layer Meteorol*, **119**, 397–407, https://doi.org/10.1007/s10546-005-9030-8.

NCEP, 2000: NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999. https://doi.org/10.5065/D6M043C6.

Nuryanto, D. E., R. Satyaningsih, T. A. Nuraini, J. Rizal, E. Heriyanto, U. A. Linarka, and A. Sopaheluwakan, 2019: Evaluation of Planetary Boundary Layer (PBL) schemes in simulating heavy rainfall events over Central Java using high resolution WRF model. *Sixth International Symposium on LAPAN-IPB Satellite*, **11372**, 225–236, https://doi.org/10.1117/12.2541817.

Rai, D., and S. Pattnaik, 2019: Evaluation of WRF planetary boundary layer parameterization schemes for simulation of monsoon depressions over India. *Meteorol Atmos Phys*, **131**, 1529–1548, https://doi.org/10.1007/s00703-019-0656-3.

Silva, N. P. da, and R. de Camargo, 2018: Impact of Wave Number Choice in Spectral Nudging Applications During a South Atlantic Convergence Zone Event. *Frontiers in Earth Science*, **6**, https://doi.org/10.3389/feart.2018.00232.

Singh, J., K. Yeo, X. Liu, R. Hosseini, and J. R. Kalagnanam, 2015: Evaluation of WRF model seasonal forecasts for tropical region of Singapore. *Advances in Science and Research*, **12**, 69–72, https://doi.org/10.5194/asr-12-69-2015.

Stengel, K., A. Glaws, D. Hettinger, and R. N. King, 2020: Adversarial super-resolution of climatological wind and solar data. *PNAS*, **117**, 16805–16815, https://doi.org/10.1073/pnas.1918964117.

Stull B., R., 1988: *An Introduction to Boundary Layer Meteorology*. Springer Science & Business Media,.

Tolentino, J. T., M. V. Rejuso, L. C. Inocencio, M. R. C. Ang, and G. Bagtasa, 2016: Effect of horizontal and vertical resolution for wind resource assessment in Metro Manila, Philippines using Weather Research and Forecasting (WRF) model. *Earth Resources and Environmental Remote Sensing/GIS Applications VII*, **10005**, 351–366, https://doi.org/10.1117/12.2241952.

Vincent, C. L., and A. N. Hahmann, 2015: The Impact of Grid and Spectral Nudging on the Variance of the Near-Surface Wind Speed. *Journal of Applied Meteorology and Climatology*, **54**, 1021–1038, https://doi.org/10.1175/JAMC-D-14-0047.1.

Virtanen, P., and Coauthors, 2020: SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*, **17**, 261–272, https://doi.org/10.1038/s41592-019-0686-2.

Wang, Q., Y. Ma, K. Zhao, and Y. Tian, 2022: A Comprehensive Survey of Loss Functions in Machine Learning. *Ann. Data. Sci.*, **9**, 187–212, https://doi.org/10.1007/s40745-020-00253-5.

Xia, G., M. C. Cervarich, S. B. Roy, L. Zhou, J. R. Minder, P. A. Jimenez, and J. M. Freedman, 2017: Simulating Impacts of Real-World Wind Farms on Land Surface Temperature Using the WRF Model: Validation with Observations. *Mon. Wea. Rev.*, **145**, 4813–4836, https://doi.org/10.1175/MWR-D-16-0401.1.

Yang, L., S. Wang, J. Tang, X. Niu, and C. Fu, 2019: Impact of Nudging Parameters on Dynamical Downscaling over CORDEX East Asia Phase II Domain: The Case of Summer 2003. *Journal of Applied Meteorology and Climatology*, **58**, 2755–2771, https://doi.org/10.1175/JAMC-D-19-0152.1.

www.nrel.gov/usaid-partnership

Scott C. Bartos
Regional Energy Advisor
USAID Regional Development Mission for Asia
U.S. Agency for International Development
Tel: +66-2-257-3000
Email: sbartos@usaid.gov

Derina Man
Technical Lead
Advanced Energy Partnership for Asia
National Renewable Energy Laboratory
Tel: +1 303-275-4744
Email: derina.man@nrel.gov

NREL/TP-5000-85481 | June 2023

The USAID-NREL Advanced Energy Partnership for Asia helps partner countries address the technical challenges of transitioning to sustainable, secure, and market-driven energy sectors across Asia. More information can be found at: **www.nrel.gov/usaid-partnership**.

USAID FROM THE AMERICAN PEOPLE

NREL Transforming ENERGY

Advanced Energy Partnership for Asia