WiSoSuper: Benchmarking Super-Resolution Methods on Wind and Solar Data

Rupa Kurinchi-Vendhan
California Institute of Technology
rkurinch@caltech.edu

Björn Lütjens[*]
Massachusetts Institute of Technology
lutjens@mit.edu

Ritwik Gupta[*]
University of California, Berkeley
Defense Innovation Unit
ritwikgupta@berkeley.edu

Lucien Werner[*]
California Institute of Technology
lwerner@caltech.edu

Dava Newman
Massachusetts Institute of Technology

Abstract

The transition to green energy grids depends on detailed wind and solar forecasts to optimize the siting and scheduling of renewable energy generation. Operational forecasts from numerical weather prediction models, however, only have a spatial resolution of 10 to 20-km, which leads to sub-optimal usage and development of renewable energy farms. Weather scientists have been developing super-resolution (SR) methods to increase the resolution, but often rely on simple interpolation techniques or computationally expensive differential equation-based models. Recently, machine learning-based models, specifically the physics-informed resolution-enhancing generative adversarial network (PhIREGAN), have outperformed traditional downscaling methods. We provide a thorough and extensible benchmark of leading deep learning-based super-resolution techniques, including the enhanced super-resolution generative adversarial network (ESRGAN) and an enhanced deep super-resolution (EDSR) network, on wind and solar data. We accompany the benchmark with a novel public, processed, and machine learning-ready dataset for benchmarking super-resolution methods on wind and solar data.

1 Introduction

In the United States, the national Energy Information Administration (EIA) predicts that renewable energy, predominantly wind and solar power, will contribute 42% of the country's electricity generation by 2050 [1]. To achieve this goal, operational decision-makers must integrate forecasting models into local power systems to address the spatial variability of these clean energy forms. However, current climate simulations used to obtain high-resolution data are unable to resolve the spatial characteristics necessary for accurate future local energy assessments, as increasing their spatial resolution is computationally expensive and provides insufficient accuracy [2]. Numerical weather predictions (NWPs) provide short-term climatological forecasting data at a horizontal resolution of 10 to 20-km [3, 4, 5], while energy planning requires wind and solar data at a smaller, more local scale, on the order of 2-km [6].

In the field of computer vision, researchers enhance the resolution of a data field through single-image super-resolution [7]. Although the problem of super-resolution is inherently ill-posed—coarsened low-resolution (LR) input data can map to infinitely many high-resolution (HR) outputs—machine learning-based approaches offer a computationally efficient method of generating the high-resolution data needed to predict the effect climate has on power generation [8, 9, 10].

Super-resolution techniques have increasingly been applied to climate data in recent years [11, 9, 12, 13]. As these methods have the potential to outperform traditional spatial SR for wind and solar predictions[14, 15], it is important to rigorously validate them to ensure that their SR outputs are accurate and realistic.

In this work we contribute:

---

[*]These authors contributed equally

- an extensible benchmark for determining accurate and physically consistent super-resolution models for wind and solar data;

- a novel application of state-of-the-art convolutional neural network (CNN)- and generative adversarial network (GAN)-based super-resolution techniques to a task from the physical sciences;

- and a novel publicly available machine learning-ready dataset for the super-resolution of wind speeds and solar irradiance fields.

2   Data

The training data was obtained from the National Renewable Energy Laboratory's (NREL's) Wind Integration National Database (WIND) Toolkit and the National Solar Radiation Database (NSRDB), with a focus on the continental United States. Wind velocity data is comprised of westward (ua) and southward (va) wind components, calculated from wind speeds and directions 100-m from Earth's surface. The WIND Toolkit has a spatial resolution of 2km $\times$ 1hr spatiotemporal resolution [16, 17, 18, 19]. Our wind dataset contains data sampled at a 4-hourly temporal resolution (every fourth data point) for the years 2007 to 2013. Wind test data are sampled at a 4-hourly temporal resolution for the year 2014.

Additionally, we consider solar irradiance data from the NSRDB in terms of direct normal irradiance (DNI) and diffused horizontal irradiance (DHI) at an approximately $4\text{km} \times 1/2\text{hr}$ spatiotemporal resolution [20]. The solar dataset produced for this work samples data at an hourly temporal resolution from 6 am to 6 pm for the years 2007 to 2013. Solar test data are sampled at an hourly temporal resolution (every other data point) from 6 am to 6 pm for the years 2014 to 2018. More information about the wind and solar datasets is available in the Appendix.

3   Models

We examined five super-resolution techniques—PhIREGAN [21], ESRGAN [22], EDSR [23], SR CNN, and bicubic interpolation—on their ability to perform $5\times$ spatial super-resolution, e.g., 10-km to 2-km spatial resolution, on wind and solar data fields. The PhIREGAN is structured in a two-phase process: an LR to medium-resolution (MR) step and an MR to HR step. In this work, we examine the MR $\to$ HR step of this model for both wind and solar data, which corresponds to a deep fully convolutional neural network based on the super-resolution generative adversarial network (SRGAN) [24]. Like the PhIREGAN, ESRGAN improves upon SRGAN's network. The ESRGAN has been shown to reproduce the simulated power spectral density of near-surface winds [25, 26], which encourages its application to climate data. The EDSR model has an architecture similar to that of the SRResNet [24], with simplifications to preserve high-frequency features. We modify the upsampling block of the ESRGAN and the EDSR to be compatible with $5\times$ super-resolution. We included the pre-trained network of the PhIREGAN—the SR CNN (distinct from the SRCNN [27]), which minimizes content loss—and bicubic interpolation as baselines for this benchmark.

4   Benchmarking Results

Figure 1 compare sample outputs from each model. Qualitatively, the results in Figure 1 show that deep learning models—most noticeably the PhIREGAN—produce outputs with sharper structures and small-scale details. This is a result of the fact that models such as EDSR that use an L1 loss function do not account for perceptual or adversarial loss like the ESRGAN and PhIREGAN do. Enlarged versions of the images in Figure 1 are provided in Figures 6-9 in the Appendix.

To assess the ability of each model to accurately recreate the ground truth data, we applied image quality metrics such as the peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), relative mean-squared error (MSE), and mean absolute error (MAE). Table 1 shows the average values of each metric across all test data, for both wind and solar data fields. All deep learning methods assessed outperform bicubic interpolation. Additionally, models which are trained to minimize content loss, such as the PhIREGAN's pre-trained CNN and EDSR have stronger metric values than the PhIREGAN, which is trained to minimize adversarial loss. EDSR outperforms all other models on image quality metrics for both wind and solar data, followed by ESRGAN.

Image quality metrics are limited in scope and do not offer a comprehensive method of evaluating wind and solar data fields as they cannot capture the ability to replicate high-frequency features. We validated the super-resolved wind speed outputs by generating kinetic energy spectra for each model which measure the distribution of energy across the various wavenumbers, k [28, 29].

Consistent with the perception-distortion trade-off [30], although the downscaled wind maps generated by the PhIREGAN performed slightly worse on accuracy metrics, they outperformed other HR outputs in terms of visual and physical fidelity. In Figure 2, energy is conserved in the inertial range of each energy spectrum and cascades at higher
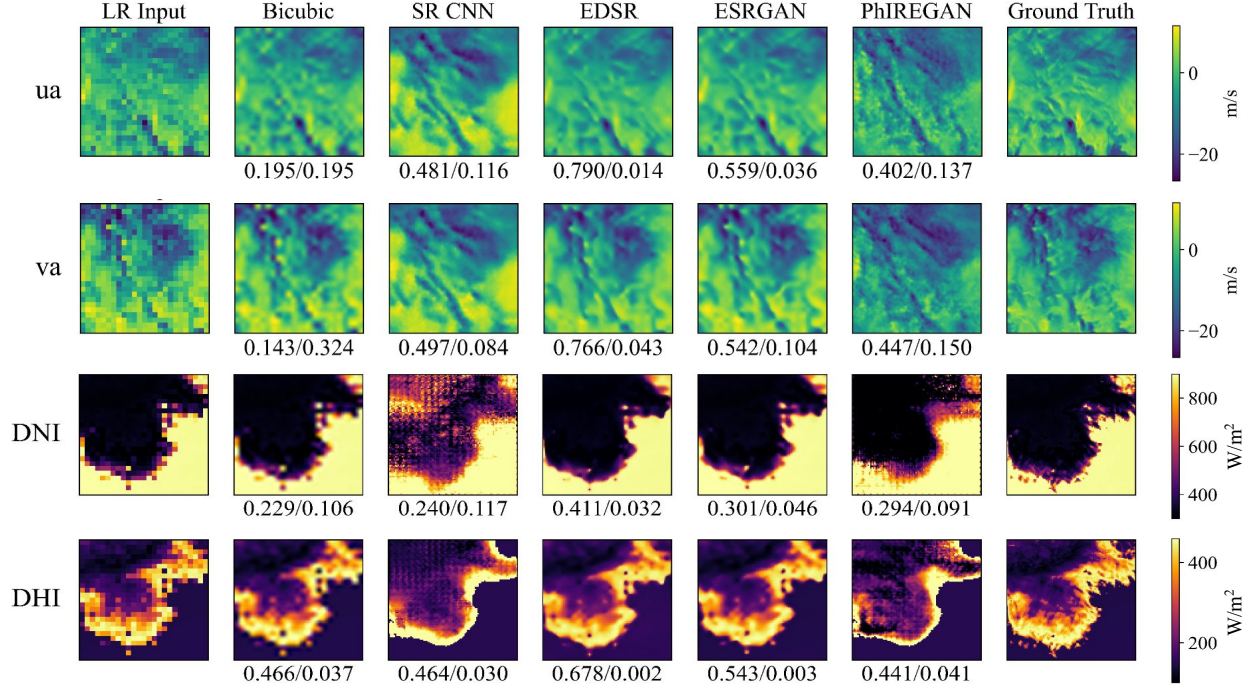
Figure 1: A comparison of wind (top) and solar (bottom) outputs from each model, with reported SSIM/MSE values. While the outputs of current models are pixelated (bicubic), blurry (SR CNN), or detailed (PhIREGAN), our contributed models (EDSR, ESRGAN) are the most accurate.

Model	Wind				Solar			
	PSNR↑	SSIM↑	MSE↓	MAE↓	PSNR↑	SSIM↑	MSE↓	MAE↓
PhIREGAN	29.11	0.46	0.12	0.27	28.85	0.44	0.34	0.56
EDSR	**32.25**	**0.83**	**0.02**	**0.09**	**32.88**	**0.63**	**0.29**	**0.19**
ESRGAN	*30.96*	*0.69*	*0.03*	*0.12*	*31.82*	*0.56*	*0.31*	*0.25*
SR CNN	28.94	0.52	0.15	0.31	29.05	0.52	0.33	0.42
Bicubic	28.83	0.36	0.21	0.36	28.78	0.40	0.46	0.61

Table 1: Summary of Average Metric Values. The models we introduce to this field (EDSR & ESRGAN) are most accurate.

wavenumbers. Bicubic interpolation and the SR CNN, which have stronger metric values, perform visibly worse than the PhIREGAN in capturing high-frequency data consistent with turbulence theory. The EDSR and ESRGAN outputs remain fairly consistent with the ground truth data especially at larger wavenumbers, despite performing optimally on the image quality metrics. As seen in Figure 2, both PhIREGAN and ESRGAN most closely match the energy spectrum of the ground truth data, which suggests that GAN-based wind downscaling approaches most successfully learn physical relationships across various frequencies that classical techniques are less capable of recreating.

For solar data, we compare super-resolved outputs by generating normalized semivariograms that show the directionally-averaged spatial autocorrelation of gradients in an image as a function of a radius r [31]. Figure 3 focus on the 4-km $< r <$ 20-km regime where each model generates high-frequency data. GAN-based models such as the PhIREGAN and ESRGAN outperform other methods on generating spectrally-consistent solar data. In both the DNI and DHI semivariograms, GAN-based models such the ESRGAN and the PhIREGAN (but primarily the ESRGAN) has the lowest deviation from the ground truth data, visibly outperforming other models.

5 Discussion and Future Works

These results indicate that the perception-distortion trade-off[30] holds for the super-resolution of wind and solar data. Our benchmark shows that CNN-based image-processing techniques (e.g., EDSR) are likely to achieve higher image similarity (in MAE, RMSE, PSNR, SSIM) while GAN-based methods achieve higher spectral similarity (see Figures 2 and 3). We show that GAN-based models have significant applications in climate scenarios, as they most reliably
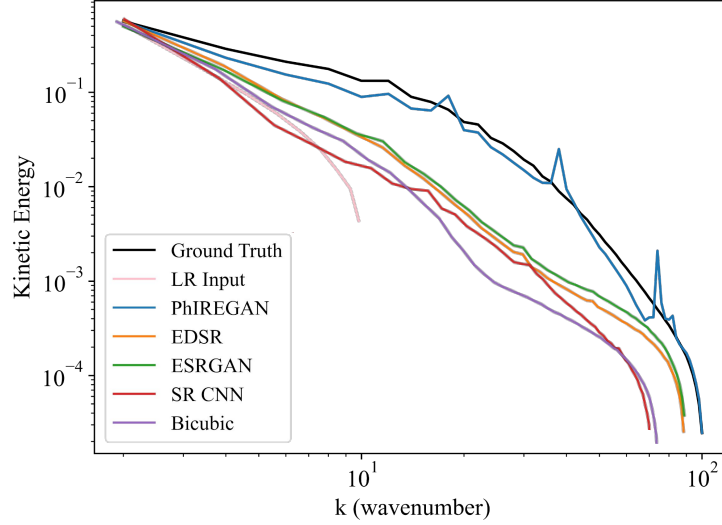
Figure 2: Kinetic energy spectra for the data fields corresponding to each HR output, averaged over all wind test data. The PhIREGAN most closely matches the turbulent physics of the ground truth data.
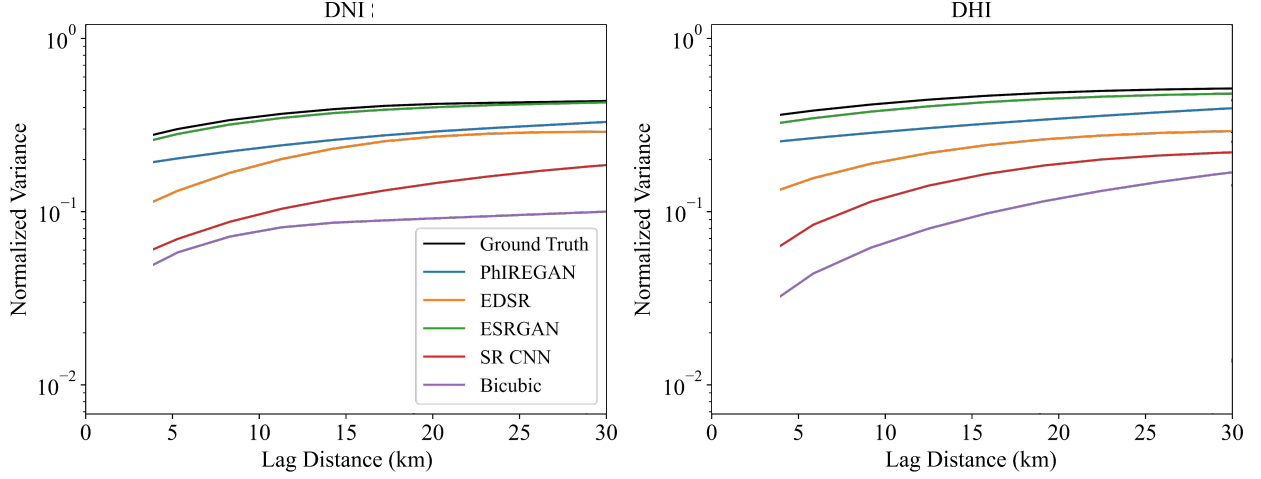

Figure 3: Normalized semivariograms for each solar output. The ESRGAN most closely matches the semivariance of the ground truth data for both DNI and DHI. This means that ESRGAN best captures the spatial correlations between pixels.

generate results that match the spectral dynamics of the ground truth. Of the two GAN-based models benchmarked, ESRGAN performs best in super-resolving solar data whereas PhIREGAN performs best for wind data.

In future works, we will extend the benchmark to stochastic super-resolution techniques such as variational auto-encoders (VAEs) [32], normalizing flows [33], and diffusion-based [34] models and metrics such as CRPS [10]. Additionally, assessing these models on their ability to preserve temporal consistency across consecutive frames by focusing on spatiotemporal super-resolution will more comprehensively represent the physical fidelity of outputs. Validating these models on data outside of NREL's WIND Toolkit and NSRDB will also widen the scope of this study and evaluate their generalization to various locations and timespans.

6 Conclusion

In this paper, several state-of-the-art super-resolution methods are thoroughly evaluated on national wind and solar data. Comprehensive experimental results demonstrate how these methods perform with respect to accuracy and plausibility. The benchmarking assessments show the qualitative and quantitative performance and limitations of each model. Our GitHub repository[2] provides machine learning-ready datasets and detailed instructions for generating a similar one, compiled implementations of each model examined in this work, and an accessible platform for assessing the turbulent flow and spatial autocorrelation of any super-resolved wind and solar output.

---

[2]https://github.com/RupaKurinchiVendhan/WiSoSuper

Acknowledgements

References

[1] *Annual Energy Outlook 2021.* U.S. Energy Information Administration, 2021.

[2] W. J. Gutowski, P. A. Ullrich, A. Hall, L. R. Leung, T. A. O'Brien, C. M. Patricola, R. Arritt, M. Bukovsky, K. V. Calvin, Z. Feng *et al.*, "The ongoing need for high-resolution regional climate models: Process understanding and stakeholder information," *Bulletin of the American Meteorological Society*, vol. 101, no. 5, pp. E664–E683, 2020.

[3] C. Sweeney and P. Lynch, "Adaptive post-processing of short-term wind forecasts for energy applications," *Wind Energy*, vol. 14, no. 3, pp. 317–325, 2011.

[4] A. Clifton, B.-M. Hodge, C. Draxl, J. Badger, and A. Habte, "Wind and solar resource data sets," *Wiley Interdisciplinary Reviews: Energy and Environment*, vol. 7, no. 2, p. e276, 2018.

[5] M. Diagne, M. David, P. Lauret, J. Boland, and N. Schmutz, "Review of solar irradiance forecasting methods and a proposition for small-scale insular grids," *Renewable and Sustainable Energy Reviews*, vol. 27, pp. 65–76, 2013.

[6] S. L. Cox, A. J. Lopez, A. C. Watson, N. W. Grue, and J. E. Leisch, "Renewable energy data, analysis, and decisions: A guide for practitioners," 3 2018. [Online]. Available: https://www.osti.gov/biblio/1427970

[7] C.-Y. Yang, C. Ma, and M.-H. Yang, "Single-image super-resolution: A benchmark," in *European conference on computer vision.* Springer, 2014, pp. 372–386.

[8] J. Baño-Medina, R. Manzanas, and J. M. Gutiérrez, "Configuration and intercomparison of deep learning neural models for statistical downscaling," *Geoscientific Model Development*, vol. 13, no. 4, pp. 2109–2124, 2020.

[9] C. D. Watson, C. Wang, T. Lynar, and K. Weldemariam, "Investigating two super-resolution methods for downscaling precipitation: Esrgan and car," 2020.

[10] J. Leinonen, D. Nerini, and A. Berne, "Stochastic super-resolution for downscaling time-evolving atmospheric fields with a generative adversarial network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 9, 2021.

[11] J. Cheng, Q. Kuang, C. Shen, J. Liu, X. Tan, and W. Liu, "Reslap: Generating high-resolution climate prediction through image super-resolution," *IEEE Access*, vol. 8, pp. 39 623–39 634, 2020.

[12] T. Vandal, E. Kodra, S. Ganguly, A. Michaelis, R. Nemani, and A. R. Ganguly, "Deepsd: Generating high resolution climate change projections through single image super-resolution," in *Proceedings of the 23rd acm sigkdd international conference on knowledge discovery and data mining*, 2017, pp. 1663–1672.

[13] R. Kaltenboeck, G. Croonen, H. Ganster, M. Gruber, K. Hennermann, M. Kerschbaum, C. Nowak, H. Mayer, S. Mayer, and M. Uray, "Image processing for weather radar data correction for aeronautical meteorology," 01 2012.

[14] T. Hoar and D. Nychka, "Statistical downscaling of the community climate system model (ccsm) monthly temperature and precipitation projections," April 2008. [Online]. Available: https://gisclimatechange.ucar.edu/sites/default/files/users/Downscaling.pdf

[15] W. Zhang, W. Kleiber, A. R. Florita, B.-M. Hodge, and B. Mather, "Modeling and simulation of high-frequency solar irradiance," *IEEE Journal of Photovoltaics*, vol. 9, no. 1, pp. 124–131, 2019.

[16] C. Draxl, B. Hodge, A. Clifton, and J. McCaa, "Overview and meteorological validation of the wind integration national dataset toolkit," National Renewable Energy Laboratory, Tech. Rep., 2015.

[17] C. Draxl, A. Clifton, B.-M. Hodge, and J. McCaa, "The wind integration national dataset (wind) toolkit," *Applied Energy*, vol. 151, p. 355–366, 2015.

[18] W. Lieberman-Cribbin, C. Draxl, and A. Clifton, "Guide to using the wind toolkit validation code," National Renewable Energy Laboratory, Tech. Rep., 2015.

[19] J. King, A. Clifton, and B. Hodge, "Validation of power output for the wind toolkit," National Renewable Energy Laboratory, Tech. Rep., 2014.

[20] M. Sengupta, Y. Xie, A. Lopez, A. Habte, G. Maclaurin, and J. Shelby, "The national solar radiation data base (nsrdb)," *Renewable and Sustainable Energy Reviews*, vol. 89, pp. 51–60, 2018.

[21] K. Stengel, A. Glaws, D. Hettinger, and R. N. King, "Adversarial super-resolution of climatological wind and solar data," *Proceedings of the National Academy of Sciences*, vol. 117, no. 29, pp. 16 805–16 815, 2020.

[22] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. Change Loy, "Esrgan: Enhanced super-resolution generative adversarial networks," in *Proceedings of the European conference on computer vision (ECCV) workshops*, 2018, pp. 0–0.

[23] B. Lim, S. Son, H. Kim, S. Nah, and K. Mu Lee, "Enhanced deep residual networks for single image super-resolution," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2017, pp. 136–144.

[24] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 105–114.

[25] A. Singh, B. White, and A. Albert, "Numerical weather model super-resolution."

[26] A. Manepalli, A. Singh, M. Mudigonda, B. White, and A. Albert, "Generalization properties of machine learning based weather model downscaling," in *Published as a conference paper at ICLR 2020*, 2020.

[27] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," 2015.

[28] A. N. Kolmogorov, "The local structure of turbulence in incompressible viscous fluid for very large reynolds numbers," *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, vol. 434, no. 1890, pp. 9–13, 1991.

[29] A. N. Kolmogorov, V. Levin, J. C. R. Hunt, O. M. Phillips, and D. Williams, "Dissipation of energy in the locally isotropic turbulence," *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, vol. 434, no. 1890, pp. 15–17, 1991.

[30] Y. Blau and T. Michaeli, "The perception-distortion tradeoff," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6228–6237.

[31] G. Matheron, "Principles of geostatistics," *Economic geology*, vol. 58, no. 8, pp. 1246–1266, 1963.

[32] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," in *2nd International Conference on Learning Representations, ICLR*, 2014.

[33] A. Lugmayr, M. Danelljan, L. Van Gool, and R. Timofte, "Srflow: Learning the super-resolution space with normalizing flow," in *ECCV*, 2020.

[34] H. Li, Y. Yang, M. Chang, H. Feng, Z. Xu, Q. Li, and Y. Chen, "SRDiff: Single Image Super-Resolution with Diffusion Probabilistic Models," *arXiv e-prints*, Apr. 2021.

7 Appendix

Ethical Considerations

This work aims to increase the efficiency of integrating renewable energy forms in US power systems. Since the focus of this study is on the continental US, our results may not be generalizable to other nations and geographic regions. However, we provide a replicable codebase and dataset that will enable transition of our results to other national wind and solar datasets.

Dataset

The data that will be available for download upon publication consists of separate datasets for wind and solar. We transform 2D data arrays of wind speed and direction into corresponding ua and va wind speed components. These are chipped into 100×100 patches. Low resolution imagery is obtained by sampling high resolution data at every fifth data point as instructed by NREL's guidelines.

The NSRDB database is formatted differently from the WIND Toolkit. A 1D array of data points is provided along with latitude and longitude metadata for each point. We re-arrange this 1D array into a 2D image based on the lat/long metadata.

All data files are made available as PNGs in their respective LR or HR resolutions. The code for for retrieving and processing WIND Toolkit and NSRDB data are accessible in the code library. Additional specifications of the wind and solar data are summarized in Table 2.

Data	Wind	Solar
Institute	NREL	NREL
Model	WIND Toolkit	NSRDB
Spatial Resolution	2 km	4 km
Temporal Resolution	4-hr	1-hr
Years	2007-2013	2007-2013
Number of Files	153,600	153,600
HR Dimensions	100×100	100×100
LR Dimensions	20×20	20×20
Colormap	Viridis	Inferno

Table 2: Wind and solar data specifications.

Examples of LR and HR imagery for both wind and solar datasets are visualized in Figures 4 and 5.

Training

The following training hyperparameters were used:

- The PhIREGAN model was trained on NREL's Eagle computing system (the pre-trained model weights were used to run inference for this project ESRGAN was pre-trained for 20 epochs and GAN-trained for an additional 20 epochs.
- The EDSR network was trained for 20 epochs.
- The ESRGAN was pre-trained for 20 epochs and GAN-trained for an additional 20 epochs. The implementations offered in the GitHub repository for this project default to these training settings.

Supplementary Figures

In support of the qualitative discussion of sample SR outputs in Section 4, Figures 6-9 provide the same images at a larger scale for better viewing and comparison.

Figures 10 and 11 show the distributions of each image quality metric examined in this work across test wind and solar data, respectively.
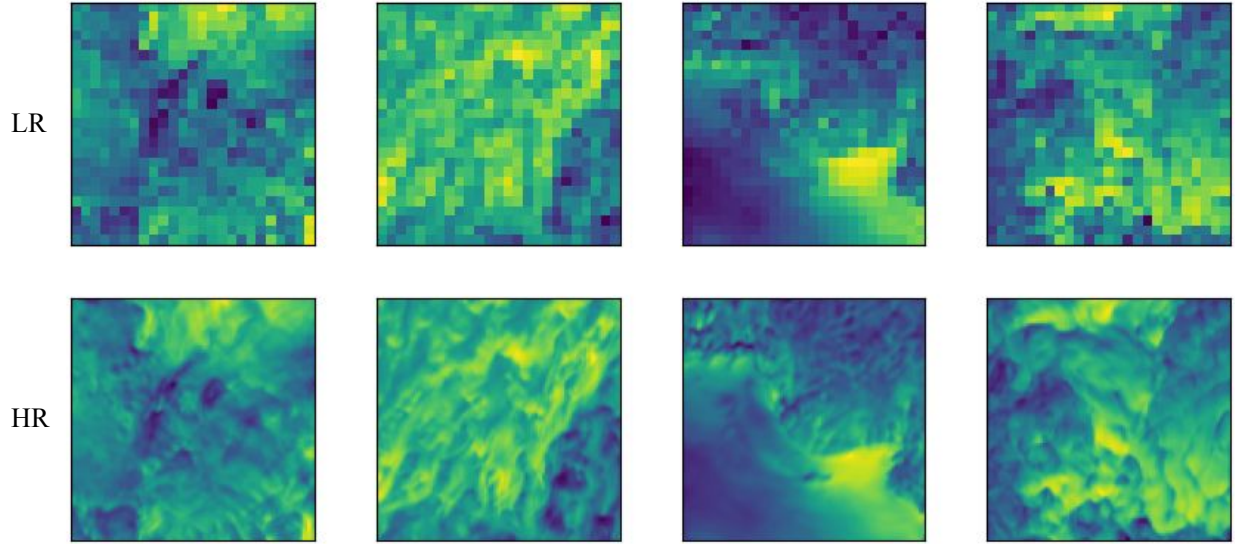
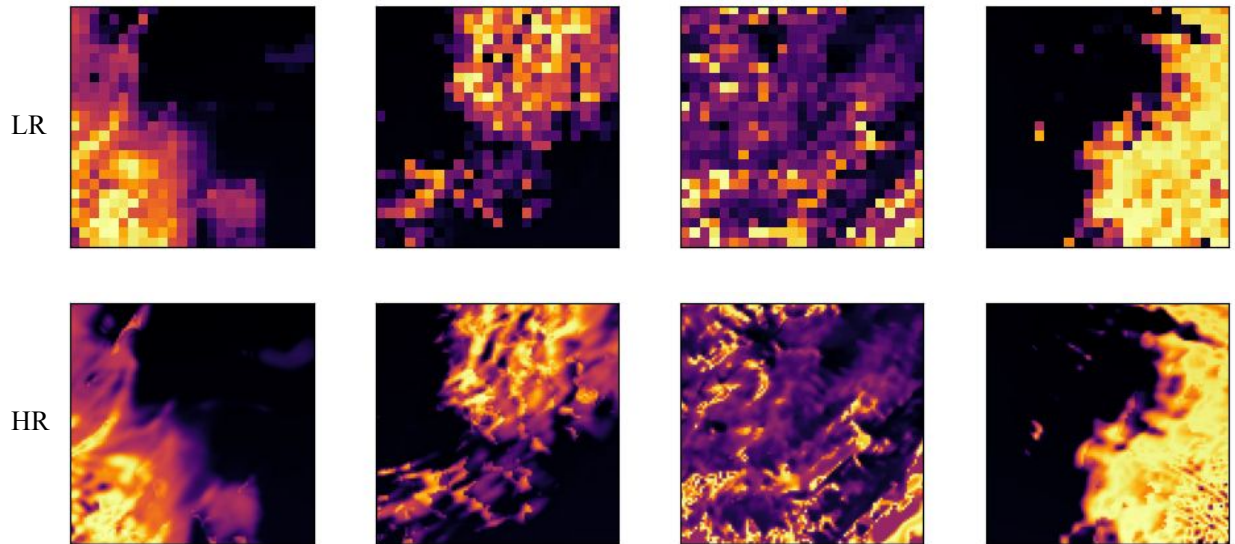Figure 4: The LR and HR pairs from the generated wind dataset, with the coarsened data being 5× downsampling.

Figure 5: The LR and HR pairs from the generated solar dataset, with the coarsened data being 5× downsampling.
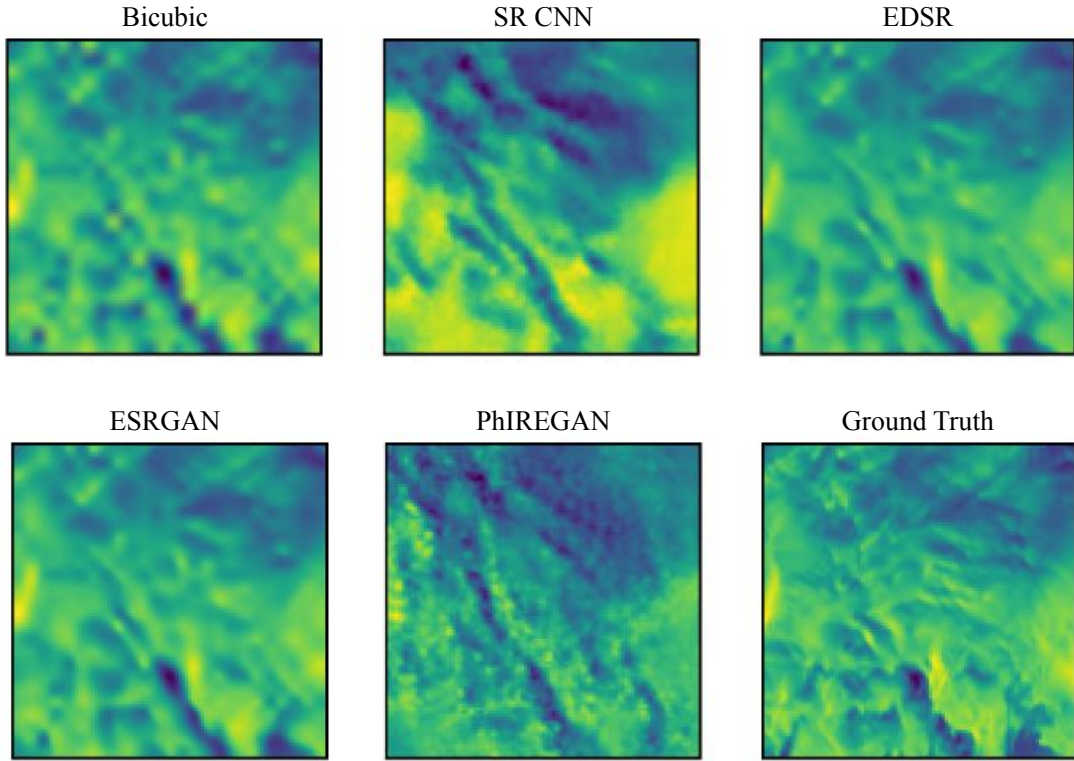
Figure 6: The ua SR outputs from each of the five models examined in this work, on a larger scale for ease of comparison.
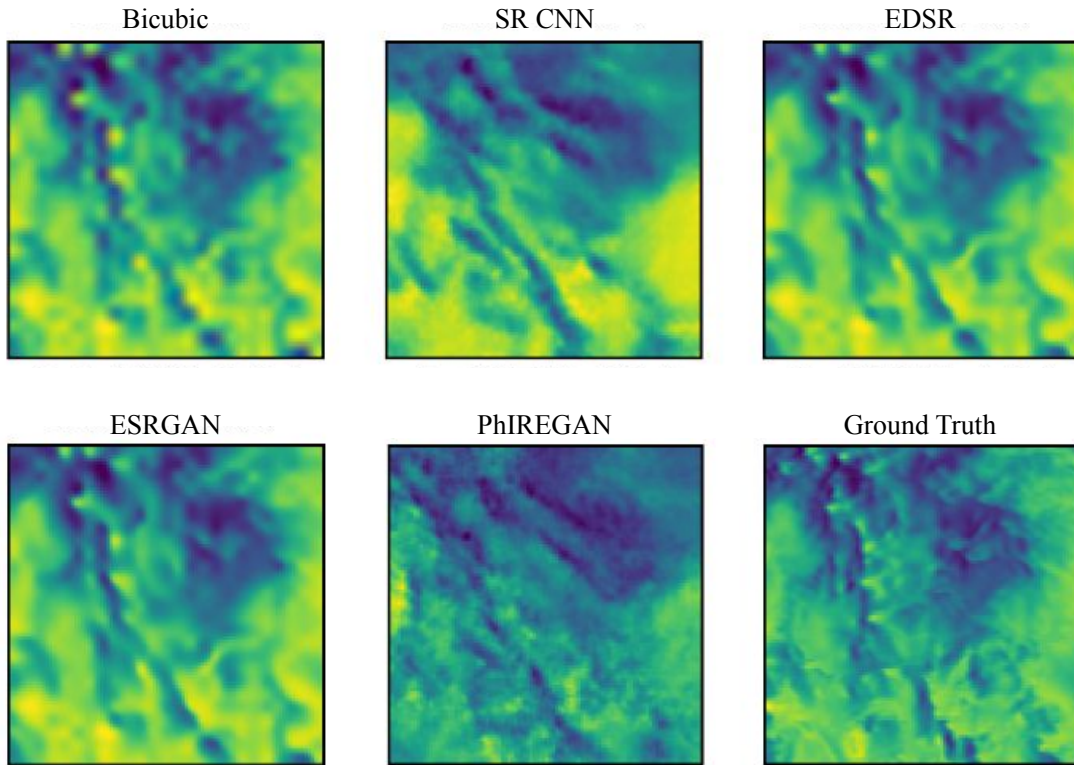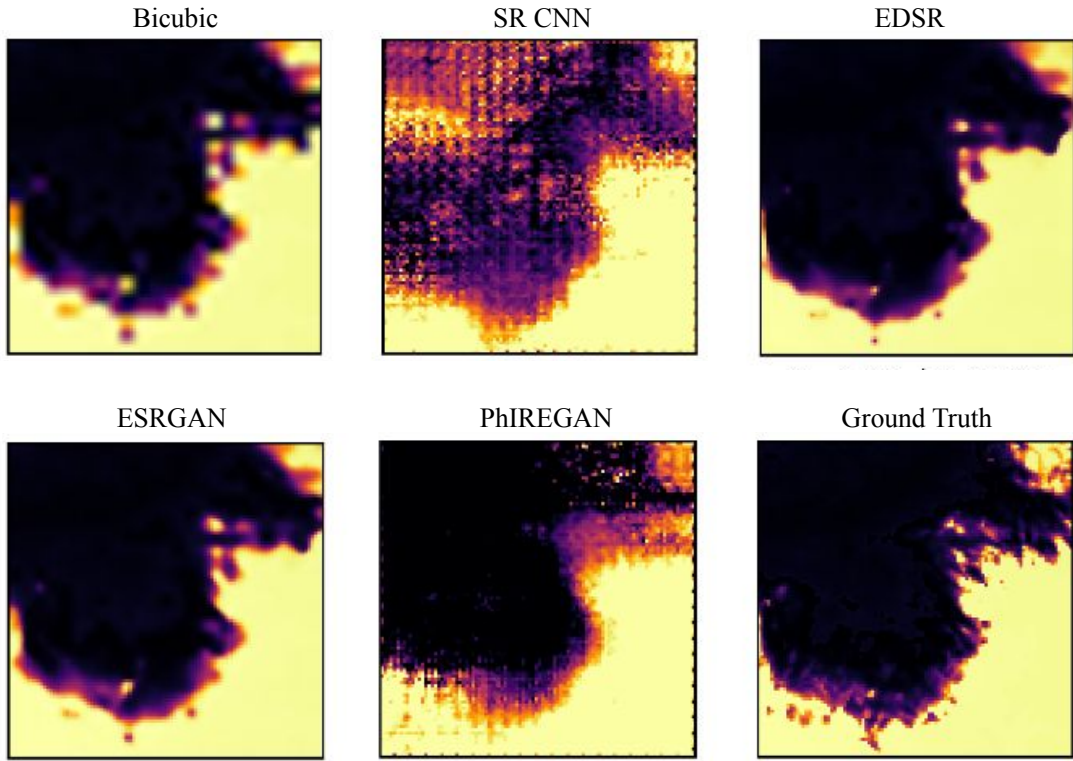

Figure 7: The va SR outputs from each of the five models examined in this work, on a larger scale for ease of comparison.

Figure 8: The DNI SR outputs from each of the five models examined in this work, on a larger scale for ease of comparison.
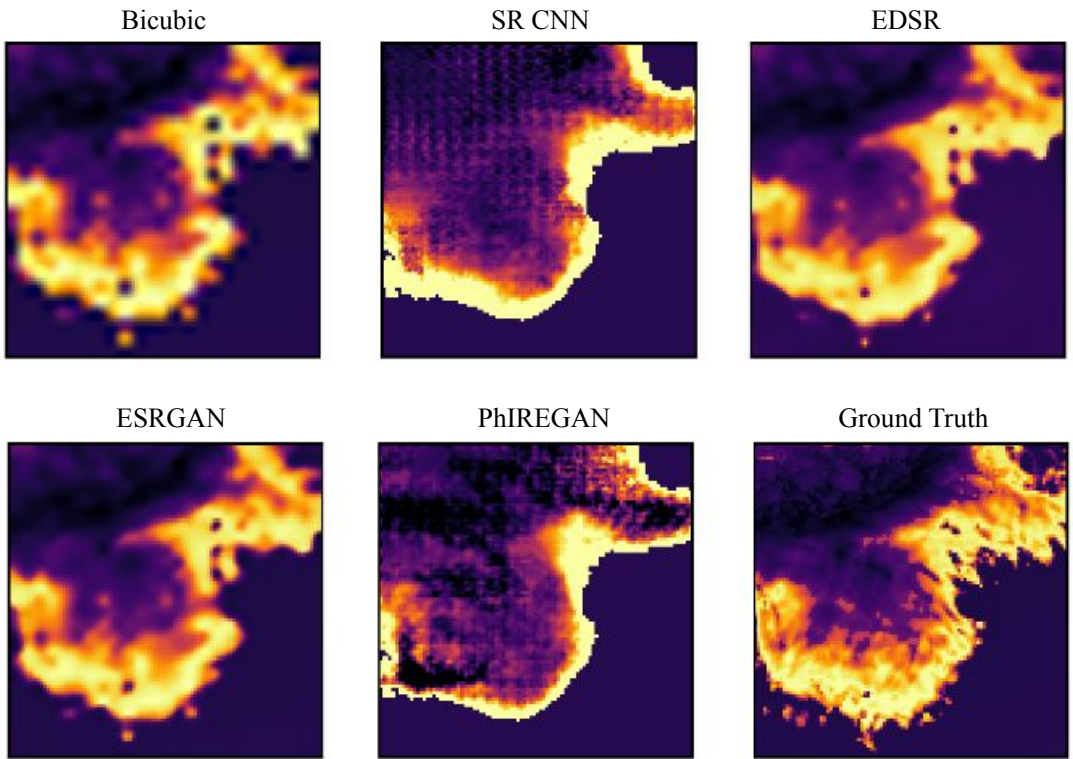

Figure 9: The DHI SR outputs from each of the five models examined in this work, on a larger scale for ease of comparison.
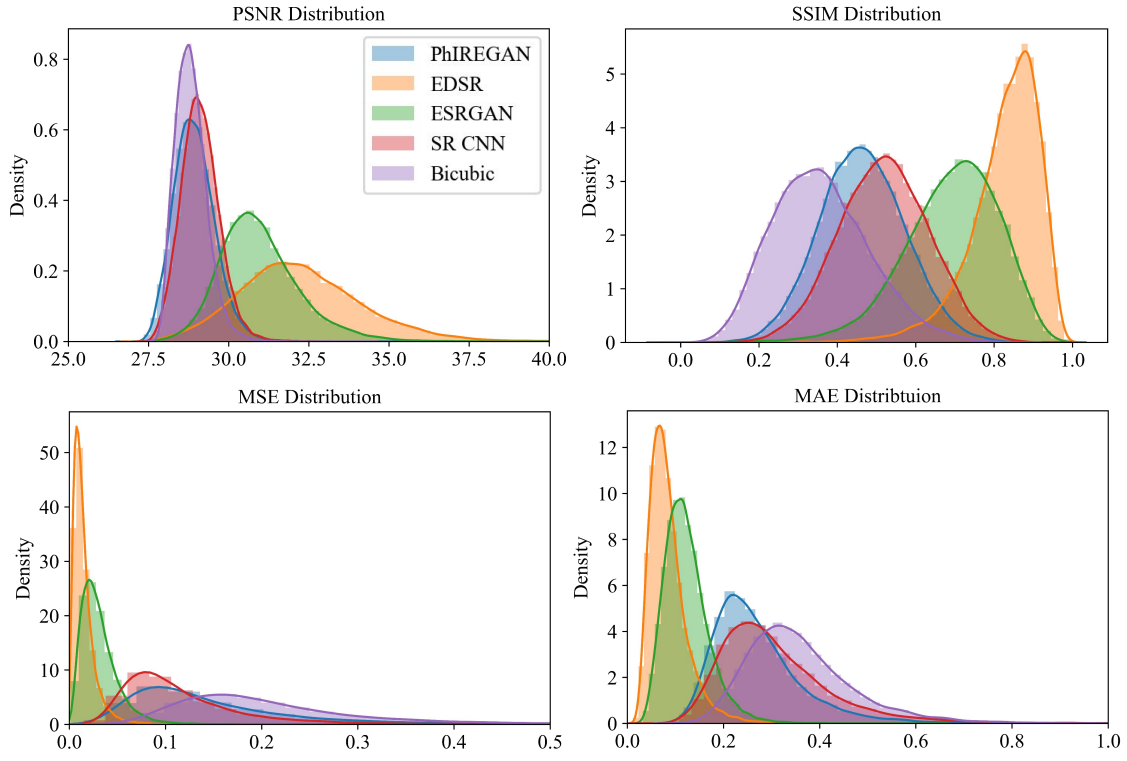
Figure 10: The distribution of each distortion metric across all wind test data, averaged over both ua and va.
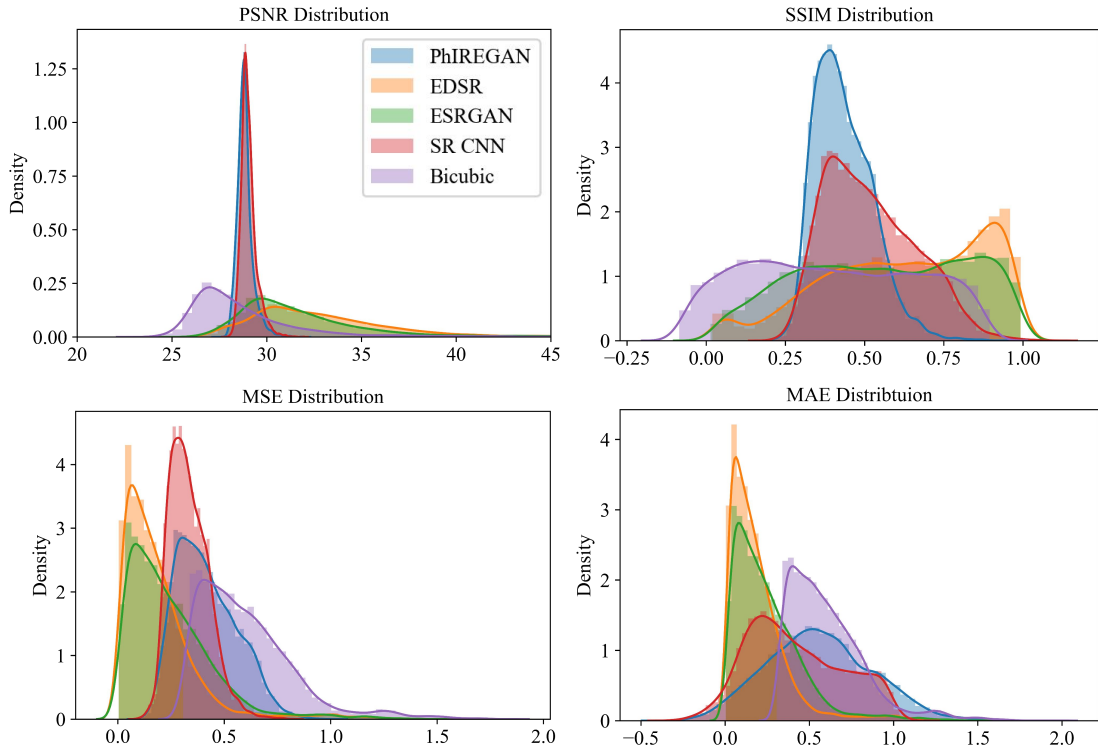

Figure 11: The distribution of each distortion metric across all solar test data, averaged over both DNI and DHI.