

MODELING TOOLS FOR WIND AND SOLAR INTEGRATION AND AIR QUALITY CO-BENEFITS IN A ZERO-CARBON FUTURE

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Jeffrey Aaron Sward

May 2022

© 2022 Jeffrey Aaron Sward

ALL RIGHTS RESERVED

MODELING TOOLS FOR WIND AND SOLAR INTEGRATION AND AIR QUALITY CO-BENEFITS IN A ZERO-CARBON FUTURE

Jeffrey Aaron Sward, Ph.D.

Cornell University 2022

Wind and solar can provide an endless supply of clean electricity, affording us all the benefits that modern society has to offer without the debilitating and inequitable effects of pollution. Unfortunately, these distributed energy resources stand in stark contrast to the central-station synchronous power plants of the past, and we have yet to remember how to work with naturally occurring flows of energy. In this future, both electricity demand and supply are inextricably linked to the weather. In response, I present a collection of open-source tools that center around meteorology – the underlying driver of future electrical grid and air quality uncertainty. I begin with a spatial study focusing on solar development and show how sunny winter days might cause as many problems as cloudy summer ones. I then showcase novel tools that will lower the barrier to entry for meteorological modeling and are aimed at giving each government and non-profit agency access to in-house wind and solar forecasts. Building upon these, I propose an integrated framework for quantifying air-quality co-benefits associated with renewable energy development, which improves the case for further investment.

BIOGRAPHICAL SKETCH

Jeffrey Sward grew up near Santa Fe, New Mexico. The still wild areas led him to develop a connection with the outdoors and a distaste for destructive extractive industries, which the state's budget still relies upon. Still, his academic success led him away from a path of "hands-on" environmental activism advocated by Edward Abbey, and he enrolled as an undergraduate in Physics at the University of New Mexico. After less than a year, he decided that physics had departed from questions relevant to those living on Earth and switched to Mechanical Engineering. He wrote his undergraduate thesis on wind energy harvesting from piezoelectrics with Dr. Svetlana Poroseva, improved his climbing skills on the crags in Central New Mexico, and managed to get a decent amount of snowboarding done in the Sangre de Cristos.

Deliberately looking for analytical tools to stem the flows of emissions across the world, he joined Dr. Max Zhang's Energy and Environment Research Laboratory at Cornell. Jeff's research spanned solar siting, wind and solar forecasting, grid integration, and air quality modeling. During his second year, he received the NSF Graduate Research Fellowship, which allowed him the flexibility to pursue new multidisciplinary research questions. But he also wanted to do something to connect the clean energy community at Cornell, so he founded the Cornell Energy Systems Club, which now holds an annual conference near Earth Day. Throughout his six years in Ithaca, he came to appreciate the difficult weather and the changing seasons of the Northeast. He took up skiing, enjoyed some bitterly cold nights at Greek Peak learning how to slide across objectively sketchy pieces of metal, and now has a new definition for "poor conditions."

Jeff will join RMI's Carbon-Free Electricity team based in Boulder, Colorado to continue his work on a cleaner grid while spending plenty of time outdoors.

To my niece Zoe Sward, my nephews Harrison and Max Sweezey, and all the others occupying future generations who will know the impacts of climate change.

ACKNOWLEDGEMENTS

As I think back over my education, I feel that a staggering number of people and events all conspired in my favor. Perhaps I cannot fully explain this feeling, but I can thank a subset of the people who actively helped me stand where I do today.

As far back as I can remember, I felt enthralled with understanding the natural world, but certain teachers helped contextualize this understanding for me, especially Anita Nugent. Her physics class convinced me that I should pursue STEM in college. Many thanks to several professors at UNM. Jose Cerrato taught one of the first engineering classes that I took immediately convincing me that I had made the correct choice to switch. Peter Vorobieff taught me the nuts and bolts of engineering experimentation but often slipped in a panda video or an unexpected joke making the subject infinitely more entertaining. Randy Truman peaked my interest in fluid mechanics with a difficult class that I left only wanting to know more. Andrea Mammoli's class on energy systems catalyzed what I already suspected – that I wanted a career working on clean energy. However, I'm especially indebted to Svetlana Poroseva who not only brought me into her lab but trusted me with my own project. This leap of faith on her part made graduate school an option for me. Still, her advice to attend graduate school in a place that I felt at home feels equally important looking back. Finally, I must mention an old friend, Jack Morrison, who one day casually put the idea into my head that I should pursue a PhD. Without this nudge, I'm not sure that I would have seriously considered the path.

I would like to extend my deepest gratitude to my advisor Max Zhang who has effectively challenged and molded my research craft during these past six years. His thoughtful questioning and willingness to entertain new ideas – even

treading upon whole new research topics – greatly contributed to my professional and personal development. Much of the work in the latter part of this thesis would have been impossible without the input from Toby Ault who’s perspective and encouraging words helped me complete projects that I would have never imagined possible just a few years ago. Prof. Lindsay Anderson informed my thinking about optimization as it pertains to the grid and policies that can shape its future. Her outgoing nature and reliable connections were instrumental in helping me acquire an internship at NREL – something that had been on my list for the better part of a decade. Finally, I want to reserve a special thank you for Prof. Jefferson Tester. His expansive knowledge and experience in the field of energy systems inspired me to think about the larger system- and societal-level questions surrounding energy transitions.

With research as interdisciplinary as mine, I have a long list of collaborators who have helped me along the way. Particularly, I want to thank Ellen Burkhard, our NYSERDA project manager. Her support of our lab and annual meetings that brought together many different researchers provided me with valuable feedback. Mike Swider at NYISO has been a constant invaluable resource grounding my research in the realities of true grid operations. Eric Zalewsky and Winston Hao at NYDEC provided much needed help on air quality modeling. I had the great pleasure of working with Robi Nilson, Rich Stedman, David Kay, and Jenny Ifft on a cross-disciplinary review of solar siting. I hope that this work helps to keep solar development honest. A special thanks go to Venky, Jackson, and Bo Yuan who worked directly with me on numerous projects and contributed to my ability to mentor younger researchers and peers but also as friends. On the opposite end, thanks to Bo Yang and Jiajun for helping me as I joined the lab and to Khaled as we figured things out together.

MAE is defined by a strong community, which is ultimately why I chose this program. This strength would not have been possible without the consistent and deliberate work of Marcia Sawyer, and for that, I owe her more gratitude than I can express. To Nick, Erik, Solberg, McCraney, Ronald, Gabe, Kyle, and Carlos – I wouldn't have been able to do this without you. The times we spent in the outdoors skiing, canoeing, running, and climbing have left me with a nostalgia for the Northeast that I will hold forever. I need to give an additional shout-out to Nick and his family for essentially adopting me during the pandemic when air travel home was out of the question.

Big thanks to Zoe whose inner strength and undying determination have helped me through the highs and lows of the past several years. I thank my siblings and their partners Jon, Anna, Marie, and Stephen. All of you have plotted marvelous trajectories along meaningful lives. I cannot say enough to thank my parents, Mark and Katie, whose artistic lens, humble morality, and unassuming support have instilled in me the courage to think outside society's entrenched systems.

TABLE OF CONTENTS

Biographical Sketch	iii
Dedication	iv
Acknowledgements	v
Table of Contents	viii
List of Tables	xi
List of Figures	xiii
1 Introduction	1
2 Strategic planning for utility-scale solar photovoltaic development – historical events revisited	7
2.1 Introduction	8
2.2 Method	13
2.2.1 Utility-scale solar photovoltaic farm development	14
2.2.2 Hourly solar PV output	18
2.2.3 Net load profiles, peak demand reduction, and ramping rates	20
2.3 Results and Discussion	22
2.3.1 Spatial distribution of projected solar farm sites	22
2.3.2 Implications for system flexibility and peak load reduction	24
2.3.3 Implications for installed capacity valuation	30
2.4 Conclusions	33
2.5 Supporting Information	41
3 Spatial biases revealed by LiDAR in a multiphysics WRF ensemble designed for offshore wind	47
3.1 Introduction	48
3.2 Method	51
3.2.1 Offshore Wind Observational Data	52
3.2.2 WRF Ensemble Setup	52
3.2.3 WRF Ensemble Validation	57
3.3 Results and Discussion	60
3.3.1 Ensemble Member Validation	60
3.3.2 Effect of Atmospheric Stability on Wind Speed Profiles	65
3.3.3 Performance of the WRF Ensemble for Offshore Wind Studies	67
3.4 Conclusion	69
3.5 Supporting Information	80
3.5.1 WRF Eta Levels	80
3.5.2 WRF Namelist Parameters	80
3.5.3 Bias Comparison at the South Buoy	81
3.5.4 RMSE Comparison at Both Buoys	82

3.5.5	Potential Errors from Orographic Waves	84
3.5.6	Stability by Wind Speed	85
3.5.7	Stability by Hour	86
4	Genetic Algorithm Selection of the Weather Research and Forecasting Model Physics to Support Wind and Solar Energy Integration	90
4.1	Introduction	91
4.2	Method	95
4.2.1	The Weather Research and Forecasting Model	95
4.2.2	Genetic Algorithm	98
4.3	Results	103
4.3.1	Annual Comparisons	105
4.3.2	Wind and Solar Trends	106
4.3.3	Effects of Physics Parameterizations	108
4.4	Discussion	110
4.4.1	Domain and Fitness Function	111
4.4.2	Best Physics Option Sets	113
4.5	Conclusion	116
4.6	Supporting Information	131
4.6.1	WRF Eta Levels	131
4.6.2	WRF Namelist Parameters	131
4.6.3	WRF Physics Parameterization Options	131
4.6.4	Notes on parallelism and computing	132
4.6.5	Day Length Model	136
4.6.6	Error and Fitness Scatter Plots	136
4.6.7	Shapley Value Plots for Microphysics, Land Surface, and Cumulus Schemes	136
5	An internally consistent framework for determining the air quality co-benefits of wind and solar development	141
5.1	Introduction	142
5.2	Method	147
5.2.1	Future Wind and Solar Development	147
5.2.2	WRF for Wind, Solar, and CMAQ	149
5.2.3	Generator Unit Commitment and Economic Dispatch . . .	153
5.2.4	Generator Emissions Estimation	155
5.2.5	The Community Multiscale Air Quality Model	156
5.3	Results	157
5.4	Discussion	165
5.4.1	Power Plant Dispatch and Emissions Modeling	165
5.4.2	WRF Meteorology for Air Quality	167
5.5	Conclusion	167
5.6	Supplemental Information	174
5.6.1	WRF Namelist Parameters	174

LIST OF TABLES

2.1	The first two columns give property class codes and descriptions for land either on which existing solar farms are located or that assumed suitable for future utility-scale solar PV farm development. The third column identifies if a property class has seen solar development, is assumed to see solar development in the future, or both. Lands classified under all codes without “Future” or “Both” listed in column three were excluded. Column four provides a justification for discrepancies between historically observed and future solar development for a property class.	43
2.2	Loss factors assigned within NREL’s System Advisor Model PV Module	44
2.3	Mean and standard deviation of peak demand reduction values for each load zone. Mean taken over the peak demand reduction on the 40 highest electricity demand days, which explains the high standard deviation values (i.e., cloud cover and therefore solar irradiance on each high electricity demand day shows a great deal of variability). All values are rounded to the nearest MW. Values in zones H,I, and K do not change because so few sites were selected that all available sites in these zones were developed under each scenario. No sites in zone J were selected using the current methodology.	44
2.4	Highest historical electricity demand days in NYS.	46
3.1	WRF Model Setups	55
3.2	Representative Weeks for Simulations (all 2020)	57
3.3	Mean Wind Profile Error Values (m s^{-1})	64
3.4	WRF Eta Levels	80
3.5	Constant Namelist Options	81
4.1	Highest Performing OptWRF and WRF User Guide-Recommended Setups. Improvements are reported with respect to the highest performing setup recommended in the WRF user guide, which is shown directly below the midrule. Plus signs indicate a performance improvement, and minus signs indicate a performance deterioration.	104
4.2	WRF Eta Levels	131
4.3	Constant Namelist Options	132
4.4	WRF Microphysics (MP) Parameterization Options	133
4.5	WRF Longwave Radiation (LW) Parameterization Options	133
4.6	WRF Shortwave (SW) Radiation Parameterization Options	134
4.7	WRF Land Surface (LSM) Parameterization Options	134
4.8	WRF PBL Parameterization Options	134

4.9	WRF Cumulus Parameterization Options	135
5.1	Constant Namelist Options	175

LIST OF FIGURES

1.1	Diagram depicting the influence of meteorology upon the power system, air quality, and public health.	2
2.1	Flow chart depicting a bottom-up approach for utilizing tax data to draw conclusions about electricity system impacts associated with future solar PV development.	13
2.2	Properties remaining after each filter has been applied are colored in dark purple while light purple properties are those eliminated by the filter. (a) shows all original properties and their boundaries, (b) shows properties excluded due to unsuitable land classes, (c) shows which properties are larger than 10 acres, (d) depicts which properties are within 1-mile of a distribution level substation, and (e) shows all properties that have an average slope below 5%.	15
2.3	NYCA load zones depicted with letters A - K. Here, Zones A - F are referred to as "upstate zones" and Zones G - K as "downstate zones". Weather station locations are shown and identified by their three character abbreviations.	19
2.4	Solar farm distribution across NYS. Blue points correspond to sites selected under only the 4500 MW scenario. Gold points correspond to sites added moving to the 6000 MW scenario, and dark red points correspond to sites added under the 9000 MW scenario.	23
2.5	(a) Shows net load profiles for upstate Zones A - F under each of the three scenarios during the heat wave spanning from July 15-19, 2013. Peak load for this episode under the no solar scenario reached 12,573 MW. (b) Depicts ramping rates for Zones A - F over the same temporal period. Solar farms mitigate a significant portion of peak load without causing significant cycling of thermal power plants. The highest ramping rate calculated ($1,171 \text{ MW h}^{-1}$) occurred under the 9000 MW scenario. (c) Gives load duration curves for Zones A - F under each of the three scenarios during the heat wave spanning from July 15-19, 2013. These confirm that peak load is reduced and shows that little to no baseload capacity is displaced by solar generation from the baseline scenario.	25

2.6	(a) Shows net load profiles for upstate Zones A - F under each of the three scenarios over a cold spell spanning from January 22-26, 2014. Peak loads as high as 10,914 MW were observed and are not reduced by solar PV generation. (b) Depicts ramping rates for Zones A - F over the same temporal period. The most drastic increases in required ramp rate occurs to meet the evening peak load, which has more than doubled on some days. The highest ramping rate – 2,356 MW h ⁻¹ – occurs under the 9000 MW scenario. (c) Gives load duration curves for Zones A - F for each of the three scenarios over the winter cold spell spanning from January 22-26, 2014. No peaking capacity is displaced by solar PV generation, but a significant amount of baseload generation is displaced during mid-day if solar electricity is considered must-take.	28
2.7	Bar chart depicting a summary of peak load (<i>left</i>) and ramp rate (<i>right</i>) split by region, season, and scenario over the entire temporal period from 2010 - 2015. Summer encompasses June through August, and winter encompasses November through March each year. Note that peak loads and maximum observed ramping rates do not generally occur during the same hour. . . .	30
2.8	Scatter plot with points representing the mean value of the annual capacity factor, the summer peak production factor (SPPF), and the peak hour production factor (PHPF). Error bars show the range of each factor, which was calculated using electricity generated by a 2 MW solar farm. Single factors were calculated for each year between 2010-2015 at each of 32 NYS meteorological stations. Stations are sorted by zone.	33
2.9	(a) shows net load profiles for downstate Zones G - K under each of the three scenarios during the heat wave spanning from July 15-19, 2013. Solar generation clearly has a much smaller effect on net load downstate as far fewer suitable properties are available here. Peak load for this episode under the no solar scenario reached 21,705 MW. (b) depicts ramping rates in Zones G - K under each of the three scenarios over the same temporal period. Highest ramps remain during the morning ramp, and the highest ramp rate (1,520 MW h ⁻¹) occurred under the no solar scenario. (c) gives load duration curves for Zones G - K under each of the three scenarios during the heat wave spanning from July 15-19, 2013. Peak load is reduced by a modest amount when compared with the upstate zones.	42

2.10	(a) shows net load profiles for downstate Zones G - K under each of the three scenarios during a cold spell spanning from January 22-26, 2014. Peak loads as high as 14,086 MW were observed and are not reduced by solar PV generation. (b) depicts ramping rates in Zones G - K over the same temporal period. Only small changes in ramping rates are seen in this region, but the highest (1,117 MW h ⁻¹) occurs under the 9000 MW scenario. (c) gives load duration curves for Zones G - K for each of the three scenarios over a cold spell spanning from January 22-26, 2014. No peaking capacity or baseload capacity will be displaced by solar in this region under any of the considered scenarios.	45
3.1	a) on the left shows the locations of the north buoy and the south buoy, respectively. b) on the right shows one of the EOLOS FLS-200 buoys in operation (Photo provided by Ocean Tech Services Inc).	53
3.2	WRF domain configuration. The largest domain has a horizontal resolution of 12 km and is shown by the figure's bounding box covering the Northeastern US. Intermediate domain, d02, has a horizontal resolution of 4 km, and the smallest domain, d03, which we use for analysis, has a horizontal resolution of 1.33 km.	56
3.3	Mean bias of WRF ensemble members predicting wind speed (m s ⁻¹) averaged over all four seasons at the south (left) and north (right) buoys.	61
3.4	Mean bias of WRF ensemble members predicting wind speed (m s ⁻¹) averaged over each representative period in each of the four seasons at the north buoy.	62
3.5	Root mean squared error (RMSE) of WRF ensemble members predicting wind speed (m s ⁻¹) averaged over all four seasons at the south (left) and north (right) buoys.	63
3.6	Taylor diagrams depicting the standard deviations, correlation coefficient, and skill scores for horizontal wind speeds at 20 m, 100 m, and 200 m for each ensemble member. The left diagram shows data at the south buoy, and the right diagram shows data at the north buoy.	64
3.7	Log wind speed profiles separated by stability class at the south buoy for both the observations and each WRF ensemble member. The number of time steps that fall into each stability class is listed in the subtitle of each plot.	66
3.8	Log wind speed profiles separated by stability class at the north buoy for both the observations and each WRF ensemble member. The number of time steps that fall into each stability class is listed in the subtitle of each plot.	67

3.9	Rank histograms for 100 m horizontal wind speed at both the south (right plot) and north (left plot) LiDAR buoy locations. We clearly see evidence of both a negative bias and underdispersion at the south buoy but perhaps only underdispersion at the north buoy.	69
3.10	Mean bias of WRF ensemble members predicting wind speed (m s^{-1}) averaged over each representative period in each of the four seasons at the south buoy.	82
3.11	Root mean squared error (RMSE) of WRF ensemble members predicting wind speed (m s^{-1}) averaged over each representative period in each of the four seasons at the north buoy.	83
3.12	Root mean squared error (RMSE) of WRF ensemble members predicting wind speed (m s^{-1}) averaged over each representative period in each of the four seasons at the south buoy.	84
3.13	Mean low-level cloud fraction for each simulation period for the Lee 2017 member defined by an 850 hPa threshold. A cloud fraction equal to 1 corresponds to full cloudiness at every level below the threshold for every time step in the simulation. Buoy locations are denoted with a grey “x.”	85
3.14	Histograms showing stability by wind speed at the south buoy location for each of the five WRF ensemble members and for the LiDAR data (lower right subplot).	86
3.15	Histograms showing stability by wind speed at the north buoy location for each of the five WRF ensemble members and for the LiDAR data (lower right subplot).	87
3.16	Breakdown of stability conditions by hour of the day at the south buoy location for each of the five WRF ensemble members and for the LiDAR data (lower right subplot).	88
3.17	Breakdown of stability conditions by hour of the day at the north buoy location for each of the five WRF ensemble members and for the LiDAR data (lower right subplot).	89
4.1	Schematic diagram of the genetic algorithm used for determining a near optimal set of physics parameters for the WRF model based upon an arbitrary application-specific fitness function. The number of generations (cycles through the diagram) can be tuned by the user based upon size of the ensemble, available time, and computational resources.	99

4.2	Plot matrix with each of nine panels highlighting the monthly mean fitness of one physics option set and the results from the other eight plotted in grey. physics option sets selected by OptWRF are highlighted in teal and appear in the first five panels; whereas physics option sets recommended by NCAR in the WRF User Guide are highlighted in navy blue and occupy the final four panels. Each of the sets chosen by OptWRF outperformed each of the sets recommended by NCAR in all months of 2011 providing compelling evidence that these physics option sets produce better forecasts across seasons and diverse meteorological conditions.	106
4.3	Panels a and b show global horizontal irradiance (GHI) across the entire modeling domain in $\text{kWh m}^{-2} \text{ day}^{-1}$ for a one-day WRF simulation initialized on December 13, 2011 00 (UTC) and the ERA5 reanalysis. Similarly, panels c and d show the wind power density (WPD) also in $\text{kWh m}^{-2} \text{ day}^{-1}$. Data shown in b, d, f, and g is taken from the best-performing simulation produced by OptWRF. Errors in the GHI (panel f) and WPD (panel g) contribute to the overall model fitness (panel e).	107
4.4	SHAP values explain the impact that each random forest model feature (WRF parameterization scheme) has upon model error. Radiation parameterization schemes are listed along the left axis, and boxes are created from all the WRF model runs. Orange bars correspond to WRF models where the parameterization scheme was activated; whereas the blue bars correspond to those where the scheme was deactivated. Negative SHAP values correspond to those schemes that reduce model error (improving the forecast); whereas positive SHAP values correspond to those schemes that increase model error. Therefore, those schemes with orange bars centered furthest to the left produce the best WRF forecasts. The namelist options for all longwave and short-wave radiation schemes are included in Tables 4.5 and 4.6 in Section 4.6.3, respectively.	111
4.5	Planetary boundary layer parameterization schemes are listed along the left axis, and boxes are created from all the WRF model runs. Orange bars correspond to WRF models where the parameterization scheme was activated; whereas the blue bars correspond to those where the scheme was deactivated. Those schemes with orange bars centered furthest to the left produce the best WRF forecasts. The namelist options for all PBL schemes are included in Table 4.8 in Section 4.6.3.	112

4.6	The left panel shows a zoomed-in view depicting GHI error over New York State and parts of the Northeastern United States – the same quantity as plotted in Figure 4.3f. The largest errors in GHI occurred off the southern and eastern coasts of Long Island, NY. The right panel shows a zoomed-in view depicting WPD error over New York State and parts of the Northeastern United States – the same quantity as plotted in Figure 4.3g. The spit of land between Lake Erie and Lake Ontario as well as the White Mountains in Northern New Hampshire have the largest errors in WPD. However, the area off the coast of Long Island, NY also has moderately high WPD errors.	115
4.7	Scatter plots showing GHI error, WPD error, and fitness values on the top, middle, and bottom, respectively. Each individual point corresponds to a single (of more than 2000) distinct WRF simulation. GHI error clearly shows a seasonal trend attributable to the changing number of daylight hours throughout the year. This provides a visual justification for the correction of this error metric by daylight fraction. WPD error shows a greater spread during the non-Summer months, but no other obvious seasonal patterns exist. Notice that WPD error values are approximately four orders of magnitude higher than those for GHI error warranting a correction factor. Finally, overarching trends in both the GHI and WPD errors are represented in the fitness values. . .	137
4.8	SHAP values explain the impact that each random forest model feature (WRF parameterization scheme) has upon model error. Microphysics parameterization schemes are listed along the left axis, and boxes are created from all the WRF model runs. Orange bars correspond to WRF models where the parameterization scheme was activated; whereas the blue bars correspond to those where the scheme was deactivated. Negative SHAP values correspond to those schemes that reduce model error (improving the forecast); whereas positive SHAP values correspond to those schemes that increase model error. Therefore, those schemes with orange bars centered furthest to the left produce the best WRF forecasts.	138

4.9	SHAP values explain the impact that each random forest model feature (WRF parameterization scheme) has upon model error. Land surface parameterization schemes are listed along the left axis, and boxes are created from all the WRF model runs. Orange bars correspond to WRF models where the parameterization scheme was activated; whereas the blue bars correspond to those where the scheme was deactivated. Negative SHAP values correspond to those schemes that reduce model error (improving the forecast); whereas positive SHAP values correspond to those schemes that increase model error. Therefore, those schemes with orange bars centered furthest to the left produce the best WRF forecasts.	139
4.10	SHAP values explain the impact that each random forest model feature (WRF parameterization scheme) has upon model error. Cumulus parameterization schemes are listed along the left axis, and boxes are created from all the WRF model runs. Orange bars correspond to WRF models where the parameterization scheme was activated; whereas the blue bars correspond to those where the scheme was deactivated. Negative SHAP values correspond to those schemes that reduce model error (improving the forecast); whereas positive SHAP values correspond to those schemes that increase model error. Therefore, those schemes with orange bars centered furthest to the left produce the best WRF forecasts.	140
5.1	Flow chart depicting the main components of the OneMet framework. Blue blocks indicate third-party open-source tools, green blocks indicate internally developed tools, and gray blocks indicate files that are transferred among the various programs. Note that the update emissions block consists of three steps that are shown in the larger dotted rectangle on the left.	148
5.2	Map showing the locations of the four lease areas under study, Empire Wind 1, 2, Beacon Wind, and Sunrise Wind, off the coast of New York State. Map reproduced from NYSERDA [2].	150
5.3	Map showing the two WRF domains centered over the Northeastern US. The coarse 12 km domain covers the entire US but is later windowed by MCIP. The fine 4 km domain covers only the OTC region.	151

5.4	Generator profiles under the base case (solid purple lines) and with renewables case (dashed orange lines). The top plot shows the Allegany Generation Station, which acts as a baseload/load following resource remaining on most of the time, but it turns off more often in the renewables case. The bottom plot shows Vernon Blvd Unit 2, which acts more like a peaking unit but turns on less frequently during the renewables case.	154
5.5	The upper panel shows the mean percent difference in NO_x emissions from the base cast to the renewables case on August 6th, 2016 for the CMAQ-ready in-line point source emissions files prepared by SMOKE. Red cells indicate an increase in emissions, blue cells indicate a decrease in emissions, and gray cells indicate no change in emissions. The lower panel shows the mean percent difference in SO_2 emissions for the same day. Note that we window the domain over NYS as point source emissions should only change here.	158
5.6	The upper panel depicts the mean $\text{PM}_{2.5}$ concentration over the full 12 km domain for the entire CMAQ model run. Darker brown colors correspond to higher pollutant concentrations. The lower panel shows the mean absolute difference in $\text{PM}_{2.5}$ concentrations for the full model run. Blue corresponds to a decrease in concentration whereas red corresponds to an increase.	160
5.7	The upper panel depicts the mean $\text{PM}_{2.5}$ concentration over the 4 km domain for the entire CMAQ model run. Darker brown colors correspond to higher pollutant concentrations. The lower panel shows the mean absolute difference in $\text{PM}_{2.5}$ concentrations for the full model run. Blue corresponds to a decrease in concentration whereas red corresponds to an increase.	161
5.8	The upper panel depicts the mean O_3 concentration over the full 12 km domain for the entire CMAQ model run spanning August 6 - 13, 2016. Brighter yellow colors correspond to higher pollutant concentrations. The lower panel shows the mean absolute difference in O_3 concentrations for the full model run. Blue corresponds to a decrease in concentration whereas red corresponds to an increase.	163
5.9	The upper panel depicts the mean O_3 concentration over the 4 km domain for the entire CMAQ model run spanning August 6 - 13, 2016. Brighter yellow colors correspond to higher pollutant concentrations. The lower panel shows the mean absolute difference in O_3 concentrations for the full model run. Blue corresponds to a decrease in concentration whereas red corresponds to an increase.	164

CHAPTER 1

INTRODUCTION

At this point in history, society at large has recognized that global warming represents a clear and present danger to life as we know it. Despite a complicated patchwork of geopolitical priorities, the vast majority of the world came together to sign the Paris Accord in 2015. While not legally binding, this provided a good faith offering from nations of the world to begin to draw down fossil fuel emissions whilst leaving vast reserves in the ground unburned. Many U.S. states and cities gave teeth to their promises by penning legislation aimed at rapidly decarbonizing their industries, their buildings, and their transportation systems over the next 20 to 30 years. At the center of these laws and aspirations lies the largest single machine ever built – the electricity grid.

But, this system – connecting thousands of power plants to millions of businesses and residences using bundles of thin steel wires – remains chronically underfunded. Today, an average piece of U.S. grid infrastructure is forty years old even alongside consistent increases in spending on transmission and distribution infrastructure. Damage from extreme weather, shifting patterns in generation and consumption, and looming bankruptcies plague the players in the modern electricity business daily. Therefore, we require new paradigms to fund the grid and ensure its reliability. Furthermore, if we hope to use the grid as the mechanism by which to decarbonize our society, then we require tools that deepen our understanding of where and when emissions are produced and what actions can displace the resources that cause them.

Fundamentally, a decarbonized grid must harness Earth’s naturally occurring flows of energy, increasingly from the wind and sun, which requires not

only accurate weather data products but the ability to act on the information contained within these data. Figure 1.1 depicts the strong – and ever-growing – influence of the weather on the grid. Public and private sector entities already apply a patchwork of weather forecasts and reanalysis data for grid planning and operations. But such fragmented approaches fail to capture key linkages putting future grid reliability at risk. In response, this thesis documents open-source tools centered around meteorology to aid in renewable energy integration in a way that strategically reduces CO₂ and criteria air pollutants. My methods strive for internal consistency and can easily be adapted to new regions around the world.

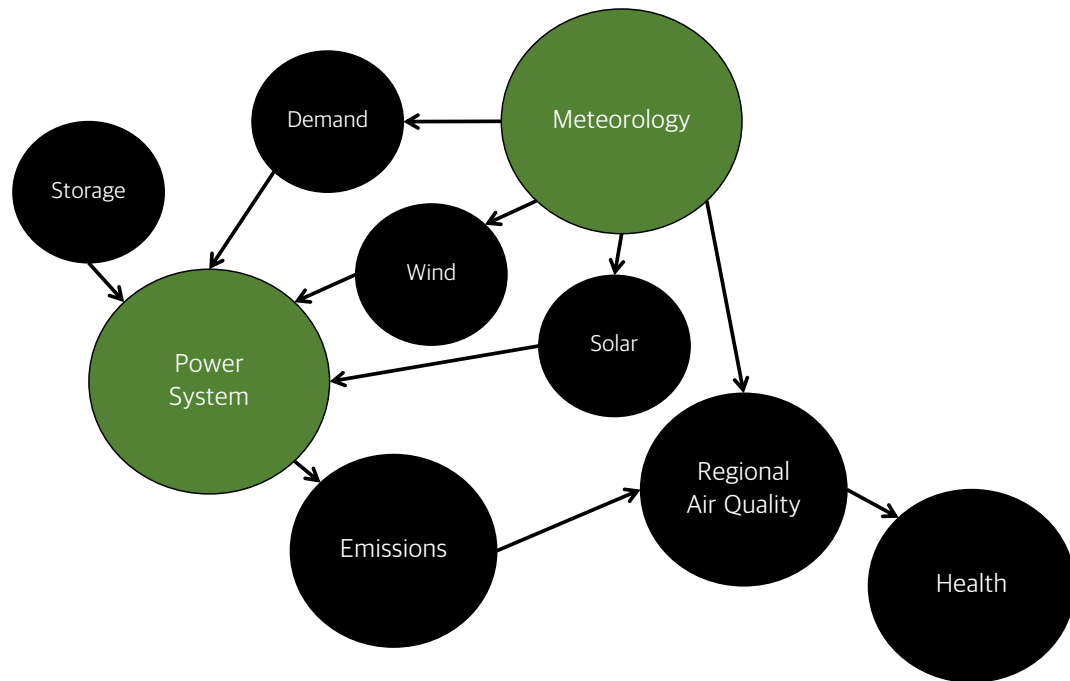


Figure 1.1: Diagram depicting the influence of meteorology upon the power system, air quality, and public health.

For much of this work, New York State (NYS) acts as a test-bed to demonstrate these tools due to several attractive features. First, electric power markets in NYS are deregulated and are operated independently from other states. This

means that NYS doesn't have to contend with as many complicated interstate dynamics during electricity planning. Rather, state policies can be directly implemented by the New York Independent System Operator (NYISO). In 2019, NYS passed the most ambitious state-level climate change policy to date, the Climate Leadership and Community Protection Act (Climate Act), which requires 100% zero-carbon electricity by 2040, 9 GW of offshore wind by 2035, and 10 GW of distributed solar by 2030. Finally, NYS is geographically and economically diverse with a majority of the population and energy demand concentrated Downstate surrounding New York City but with a majority of existing power plants and prime land for renewable development located Upstate. This configuration represents a classic "centralized" approach to electricity generation and distribution, which increases the risk of catastrophic failures when faced with extreme weather, cyber attacks, or simply deferred maintenance. Therefore, reliably providing enough clean power to Downstate NY offers an interesting challenge.

A brief summary of the four major projects featured in the coming Chapters is as follows:

In Chapter 2, I assess the implications of utility-scale solar development. After mapping available land across NYS, I designed development scenarios in line with NYS's renewable portfolio standards. Through this study, we learned that New York's species of "duck curve" will likely occur during sunny winter days. During such periods, flexibility resources will become highly valuable due to the potential for high ramps, but the wholesale cost of electricity will likely remain in negative territory during the midday with large quantities of zero marginal cost resources on the system.

Chapter 3 addresses offshore wind resources in the Northeastern U.S. While this region has the potential for vast wind farms, development remains nascent. I assess a multiphysics ensemble of WRF-simulated horizontal wind speeds in the New York Bight. This constitutes the most detailed analysis of WRF performance for offshore wind in the Northeast completed to date. For the first time, I had access to LiDAR data at multiple locations, which allowed me to comment on WRF biases in four dimensions. As expected, WRF wind speed biases increased with height, but biases also differed substantially between the two buoys even in estimates generated by the same member. This offers additional compelling evidence that extrapolating power generation from one location to another offshore causes errors. In other words, electricity planners and operators in the U.S. lack the observational data necessary to determine the spatiotemporal bias patterns in offshore wind, which are necessary to intimately understand resource coincidence and availability. I found that no single member performed the best in all stability classes but that the setup devised by Optis et al. [1] performed the best on average. Finally, I reported that the ensemble as a whole showed classical signs of underdispersion as well as a greater tendency to underpredict wind speeds offshore.

In Chapter 4, I take the first steps at breaking down the barriers to entry for customizing a WRF model for any application in any region called OptWRF. Generally, WRF setups are iteratively improved by a team of experts over a period of months to years putting in-house meteorological modeling out of reach for many governments and nonprofits across the world. OptWRF optimizes a WRF setup using a genetic algorithm. I showed that the setup found by OptWRF for the dual purpose of wind and solar downscaling outperformed all the setups recommended by NCAR in the WRF User's Guide for every month

of the year. The novel fitness function that I designed for combined wind and solar analysis featured a metric called wind power density, which allows for the direct comparison of errors in wind and solar resources. Such a metric can aid planners in investigating contingencies for periods of shortfall and excess generation alike. In other words, this combined metric offers a convenient way to characterize errors in a future variable renewable energy-dominated system. In addition to creating a database of viable WRF setups, I crafted a random-forest-based postprocessing algorithm to shed light on the effect that the parameterization of each major physical process has upon total model error.

Finally, in Chapter 5, I present an integrated framework – named OneMet – that uses consistent meteorology to forecast renewable energy generation, dispatch power plants or flexible resources to meet the remaining net load, and assess air quality. I argue that this internal consistency can offer more consistent bias patterns thereby making it easier to bias-correct using established statistical postprocessing techniques. Methods developed in Chapters 2-3 inform renewable energy development patterns, and their associated power generation profiles are derived from WRF. These profiles are then fed into an in-house representation of the New York State power system, and the resulting dispatch pattern dictates CO₂ and criteria pollutant emissions. Finally, updated emissions estimates drive a chemical transport model, which elucidates the air quality benefits – and trade-offs – associated with a particular renewable development scenario. Therefore, OneMet can inform government incentives for renewables seeking to encourage a pathway that will reap the greatest emissions and air quality benefits.

References

- [1] Mike Optis, Andrew Kumler, Joseph Brodie, and Travis Miles. Quantifying sensitivity in numerical weather prediction-modeled offshore wind speeds through an ensemble modeling approach. *Wind Energy*, 24:957–973, 2021.

CHAPTER 2

STRATEGIC PLANNING FOR UTILITY-SCALE SOLAR PHOTOVOLTAIC DEVELOPMENT – HISTORICAL EVENTS REVISITED

Abstract

Rapidly growing utility-scale solar photovoltaic (PV) holds promise for addressing energy and environmental challenges posed by high electricity demand days (HEDDs). We assessed the implications on strategic planning of future solar development in an emerging solar market, New York State (NYS) in the U.S., by synthesizing information on electrical infrastructure, tax assessment, geographical constraints and measured meteorological data. Considering three solar PV penetration scenarios (4500 MW, 6000 MW, and 9000 MW), we examined the impact of distributed utility-scale solar farms on peak demand reduction and ramping requirements during historical peak events. Our results reiterate that, across NYS, a wealth of low-value land exists to support utility-scale solar PV and that generation from these resources can reduce peak demand by up to 9.6% under the 9000 MW scenario. In addition, peak demand reduction displays locational and temporal dependency. Contingent upon local meteorology, a solar farm can reduce demand during the peak hour by anywhere between 10 and 74% of its rated capacity during summer HEDDs. However, the highest ramping requirements are more likely to occur during winter than summer. Furthermore, because developers cannot predict performance during the annual system peak, current capacity valuation methodologies for solar projects may not be adequate to promote a healthy competitive market for solar. Incorporating a broader spectrum of peak demand conditions into variable re-

source capacity valuation would improve strategic planning, not only in NYS, but across quickly growing solar markets worldwide.

2.1 Introduction

High electricity demand days (HEDDs) are usually driven by intensified usage of air conditioning during prolonged hot and humid summer periods. As the climate changes, regional heat waves are predicted to become more frequent [26], increasing the quantity of HEDDs and the magnitude of peak energy demand [31]. High energy costs, dangerous temperatures, and air pollution during HEDDs pose great challenges in maintaining grid reliability and protecting public health [40]. Recent rapid development of distributed energy resources (DERs) provides an economically viable opportunity to mitigate impacts of HEDDs.

DERs broadly include distributed generation (DG), energy efficiency, demand response, and energy storage. Driesen and Katiraei argue that restructuring the electrical system to accommodate higher penetrations of DERs can improve reliability [17]. Subsequent work showed the vast potential that renewable DERs possess in mitigation of greenhouse gas emissions [9]. However, siting numerous small generation resources further complicates power system planning, so an increasing amount of work has aimed to develop multi-objective optimization techniques for locating various types of DERs to attain the greatest overall benefit [10]. On the demand side, energy efficiency measures, dynamic pricing, and demand response programs have already been widely adopted to slash peak demand [40]. New flexible technologies, such as Vehicle-to-Grid

(V2G), may provide additional effective methods of reducing peak [37]. On the supply side, solar photovoltaic (PV) is the most promising form of renewable DG due to its comparatively low installed cost and because peak solar generation coincides with mid-day system peak demand [12].

The U.S. states that have dedicated carve-outs for solar within their renewable portfolio standards (RPS) (e.g. Arizona, Delaware, Nevada, Massachusetts, New Mexico, etc.) installed 722% more capacity between 1997 and 2009, on average, than those without [33]. However, since these carve-outs are often specified on a statewide basis, they fail to value the location-based benefits of solar. Similarly, traditional net metering incentives appear as a credit on a customer's monthly utility bill. Such a mechanism cannot elicit responses from behind-the-meter solar owners on time scales at which the day ahead or real time electricity markets operate. Furthermore, high system ramping rates resulting from aggressive solar PV penetrations may become an unintended consequence of high solar penetrations [8]. Under such scenarios, diurnal peak load tends to shift toward evening hours and many dispatchable generators remain idling during the day to avoid incurring exorbitant start-up costs during the evening ramp [16]. As such, steeper ramping requirements could lead to higher total emissions from thermal generators as emission rates deviate considerably during start-up, up-ramping, and part-load operation when compared with steady-state operation [36]. As articulated by [27], integration of solar and wind resources can only go so far given current system flexibility limitations - something that must be addressed in future legislation.

To fully realize potential benefits of DERs, solar PV in particular, on HEDDs requires improved quantification of several key impacts that DERs have on the

power system during the planning stage. First, benefits of DERs are locationally and temporally dependent due to existing power system network infrastructure and patterns of electricity consumption. Also, since DERs are frequently interconnected within power distribution systems, independent system operators (ISO) or regional transmission organizations (RTO) have neither visibility nor control over distributed generation, which makes it necessary to account for intermittency to ensure bulk power system reliability [7]. Finally, the impact of solar PV on system flexibility requirements should be assessed.

Various methodologies exist for siting intermittent resources and determining their impact. Previous studies have estimated solar PV resource potential at national scale in the US [18], China [21], and Australia [30]. One study looks specifically at environmental and health benefits associated with expanded solar development in the US [38]. Nikolakakis et al. designed an optimization model to determine the maximum penetration of intermittent resources in New York under a specific grid flexibility scenario assuming no transmission constraints [27]. Trade-offs between siting PV farms to maximize energy production or to minimize reserves were recently explored [35]. Pietzcker et al. used a coupled energy-economy-climate model to determine what role solar technologies might play in decarbonization of the power sector [29].

Specifically, GIS provides a useful aid in identifying and evaluating potential solar farm sites. One study, by Brewer et al., incorporates survey data to assess site suitability based upon public attitudes about solar farm locations [13]. Sanchez-Lozano uses a multi-criteria decision analysis (MCDA) technique to evaluate alternative locations for solar farm development [32]. While both of these studies offer developers valuable information on where they should site

individual solar farms, they provide little insight into long-term renewable integration planning. Recently, Yushchenko et al. treated a larger scale by conducting a MCDA assessment across West Africa to identify areas for grid-connected PV development [39]. However, they note that future estimates should include an economic evaluation to accompany an estimate of technical potential.

As such, the two main objectives in this study are: a) implement a GIS-based siting approach considering individual properties in a study region with non-ideal solar resources, and incorporate locational and temporal resource characteristics to estimate the impact of intermittent generation on peak load and ramping requirement during HEDDs; b) assess the current valuation method for installed solar capacity. The two objectives are closely connected because the peak load reduction potential is a critical component in valuating installed capacity. Using diverse data sources in a regional siting exercise can add to the discussion among academics, developers, and policy-makers about best practices for regional solar development and what impact different development trends may have upon electricity markets. Due to the capital-intensive nature of solar projects, how solar capacity is valued could determine if utility-scale solar projects remain viable investments. Furthermore, the energy value of solar decreases with increasing penetration of zero marginal cost resources [22]. Therefore, characterizing the extent to which solar farms mitigate peak demand and compensating them accordingly remains an open and vital task.

For this study, a bottom-up approach is defined as one that identifies appropriate locations for solar farms by assessing individual contiguous tracts of land of sufficient size, slope, and property class to support solar farm development. Such a method provides advantages over a top-down approach where an aver-

age capacity is specified over a study region. Specifically, using substation locations and geographic information for each site, a judgment can be made about which sites may be interconnected at relatively low cost. Finally, in knowing the exact geographic location of each potential solar farm, it becomes possible to target those sites for development which are located nearer to load pockets or are in locations where the solar resources – dictated by local meteorological conditions – are most favorable.

New York State (NYS) is taken as a testbed for demonstrating our method introduced above due to several attractive characteristics. NYS’s Clean Energy Standard (CES) mandates that 50% of all electricity be generated via renewable energy by 2030 – an estimated 70,500 GWh [7], which will provoke major investment in renewable energy infrastructure over the next decade. The New York State Public Service Commission (NYSPSC) has transitioned away from net metering to a new compensation mechanism for DERs known as the Value of Distributed Energy Resources (VDER) framework [41], which seeks to encompass locational, temporal, and environmental values provided by DERs. Therefore, results presented here can assist policy-makers with VDER implementation. Synthesizing granular datasets containing tax classification, substation location, and various geographic constraints, we estimated the spatial distribution of solar farms across NYS and modeled hourly solar output during HEDDs identified between 2010 and 2015. Then, a comprehensive evaluation of electricity system benefits provided by distributed solar farms, inclusive of intermittency considerations, was conducted during times when the system experiences the most stress. Finally, we recommend capacity valuation alternatives which can potentially create a mechanism to achieve state RPS targets and greenhouse gas reduction goals.

The remainder of this article is organized as follows: Section 2.2 describes our methodology. Net load profiles, peak demand reduction magnitudes, and ramping rate curves are presented and their implications discussed in Section 2.3. Finally, Section 2.4 draws conclusions as well as stipulates on future research aimed at improving utility-scale solar farm development.

2.2 Method

Figure 2.1 illustrates the overall method. We start with NYS-wide tax property data collected during assessment, and filter, then rank all properties based upon a set of criteria that affect the likelihood of those properties being developed into solar farms. Then, we model electricity outputs from the selected solar farms under three penetration scenarios using historical meteorological conditions and evaluate the impact on peak demand reduction and ramping requirements.

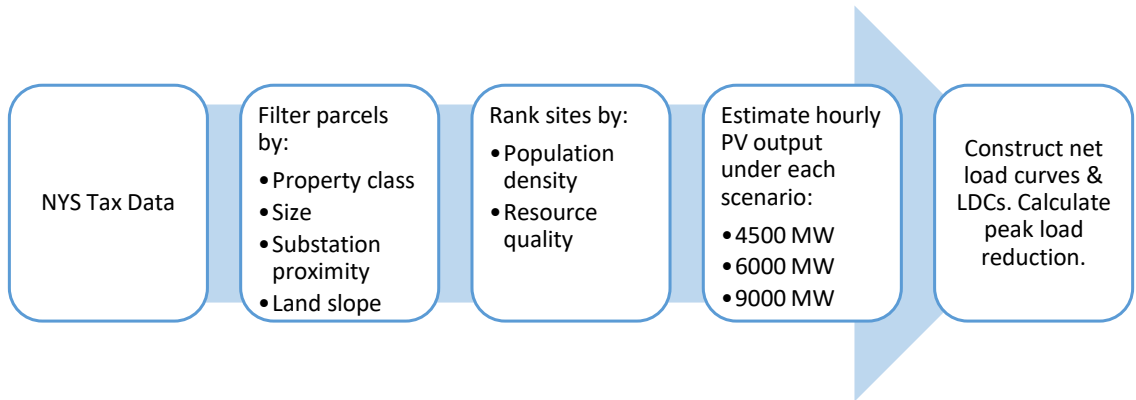


Figure 2.1: Flow chart depicting a bottom-up approach for utilizing tax data to draw conclusions about electricity system impacts associated with future solar PV development.

2.2.1 Utility-scale solar photovoltaic farm development

We utilized granular datasets including NYS-wide tax property classifications, electrical substation locations, and geographic constraints to estimate a spatial distribution of solar farms. Where solar farms are sited will determine whether or not they displace electricity generated by conventional generators and therefore whether or not they mitigate greenhouse gas emissions. As predicting precise locations encompasses numerous uncertainties, our focus lies on broad trends rather than the likelihood of development for any individual site.

Site selection process

Property classes from the NYS real property system suitable for utility-scale solar farms were identified and compared against those hosting 15 existing solar farms. Methods for filtering and exclusion of land are summarized in Figure 2.2. A 2015 statewide tax property dataset [3], obtained from the NYS Office of Information Technology Services, containing property boundaries and associated tax information for every taxable piece of land within NYS formed the basis for site selection. Figure 2.2a depicts property boundaries in a region near Buffalo, NY.

Table 2.1 in Section 2.5 lists property classes considered appropriate for future utility-scale solar farm development (listed as “Future” or “Both”), property classes currently hosting solar farms, and explanations of any discrepancies. Land that is fairly flat, clear of trees and buildings, and often vacant provides the simplest, most efficient, and most economical base for solar farm development. Property classes reflecting productive farmland were avoided due

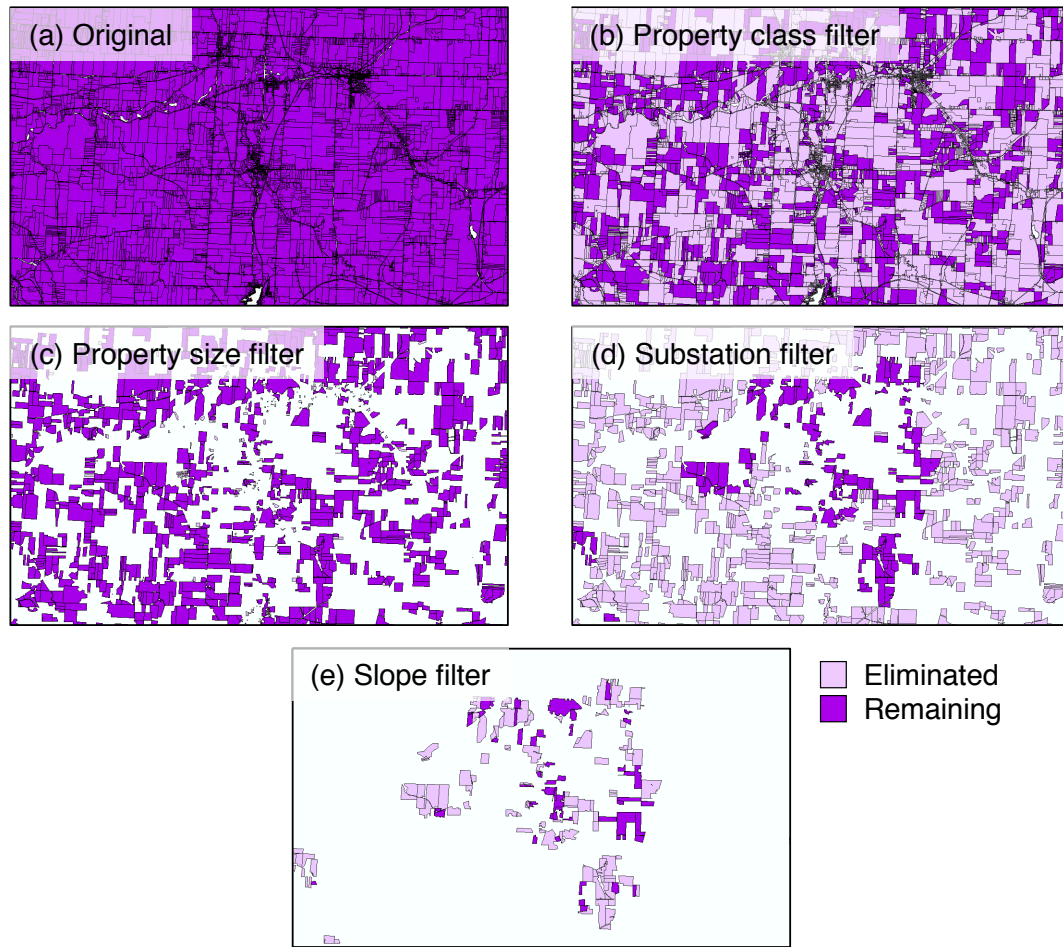


Figure 2.2: Properties remaining after each filter has been applied are colored in dark purple while light purple properties are those eliminated by the filter. (a) shows all original properties and their boundaries, (b) shows properties excluded due to unsuitable land classes, (c) shows which properties are larger than 10 acres, (d) depicts which properties are within 1-mile of a distribution level substation, and (e) shows all properties that have an average slope below 5%.

to anticipated future protection, but abandoned and vacant agricultural lands were considered. Remaining properties after property class filtering are shown in Figure 2.2b.

Properties were also filtered based on size, slope, and power system inter-connection cost. Only solar farms of 2 MW (nameplate capacity) or above – re-

quiring a minimum of 10 acres either on individual properties or collections of contiguous properties – were considered as this capacity represents a common threshold between utility-scale and commercial-scale solar. Since commercial-scale solar installations often occur on different property classes than utility-scale solar, this distinction is important. The 10-acre size for a 2-MW solar farm is selected based on the National Renewable Energy Laboratory’s (NREL) estimate for solar PV installation density 39 MW km^{-2} [28], which was confirmed by existing NYS solar farms. Land meeting both property class and size requirements are shown in Figure 2.2c.

Furthermore, properties were excluded if their average slope derived from the US National Elevation Dataset [34] (the finest resolution, 1/3 arc-second, was used for the current study) exceeded 5% – see Figure 2.2d – as additional costs incurred during initial development become prohibitive [4]. This fact was confirmed by examining existing NYS solar farms. Finally, properties were excluded if not located within a 1-mile radius of a distribution-level (up to 115 kV) substation as depicted in Figure 2.2d. Locations for distribution-level substations within New York were obtained from Transmission AtlasTM for the Eastern Interconnect region [2]. Euclidean distance to the nearest distribution level substation was calculated and used as a surrogate to represent power system interconnection costs, which are estimated by NREL at \$500,000/mile [14]. Although many studies assume no new interconnection lines will be installed for projects below 10 MW, local developer feedback suggested that capital costs for 2 MW farms become too great if the interconnection point is further than one mile away. Note that we do not consider if lines have required capacity or transfer capability to accommodate proposed solar farms.

A major shortcoming inherent to the filtering methodology is the exclusion of all land within New York City (NYC). This occurs because no NYC properties within property classes deemed appropriate for utility-scale solar development meet size requirements. However, as this methodology seeks to strategically deploy solar near load pockets, a solar PV capacity of 706 MW was assumed for NYC under each scenario. The estimate was obtained by linearly interpolating forecasted solar electricity production values reported for 2027 by the New York Independent System Operator (NYISO) [6] out to year 2030. Although this capacity counted toward the total statewide PV penetration, no solar generation values were calculated for Zone J (shown in Figure 2.3) as the model assumes utility-scale PV installations.

Site Ranking

With a focus on peak demand reduction, sites were ranked based on surrounding population density derived from 2016 census data – a proxy for electricity demand. Conversely, Yushchenko et al. considered a maximum distance from human settlements favorable for utility-scale PV installations to allow for future urban development [39]. Therefore, this study provides a perfect contrast and can aid policy-makers balancing urban development costs with energy transition infrastructure costs. This quantity shows greatest sensitivity to the chosen radius. For this study, we set this parameter to 5 miles around the centroid of each proposed site location. To consider solar resource quality, the average population density was multiplied by a solar resource drought factor obtained by dividing the annual average number of solar resource drought days at the nearest meteorological station by 365. Solar resource drought days are defined as

any day where daily average global horizontal irradiance (GHI) falls below 84 W m^{-2} [20].

Solar penetration scenarios

The effect of different solar PV penetrations was investigated under three scenarios, i.e., 4500 MW, 6000 MW, and 9000 MW. These scenarios are based on projections stemming from NYS' target of 3000 MW installed by 2023. The target capacities in the three scenarios were obtained through an adoption model projecting explosive growth in solar PV installations in the short term following by a leveling off of installations around 2025 as market saturation occurs [5]. Zonal allocation was achieved by calculating the average zonal peak load over the 40 highest historical electricity demand days and assigning a fractional capacity to each zone by dividing the zonal average peak load by the state peak load. The highest ranked sites in each zone were selected to meet zonal targets.

2.2.2 Hourly solar PV output

For each selected solar site, we used geographic coordinates at the centroid to locate the closest of 32 airport-based meteorological stations across NYS shown in Figure 2.3. Hourly solar GHI data were obtained from ASOS observations processed using an adapted Meyers and Dale model [11] and made available through the Northeast Regional Climate Center at Cornell University. GHI is the sum of direct (DNI) and diffuse (DHI) radiation received on a horizontal plane. Values for DNI were estimated using NREL's DISC model [25]. DISC is a quasi-physical model that predicts DNI from GHI recognizing that air mass is

the dominant parameter in the relationship between normal and global atmospheric transmittance.

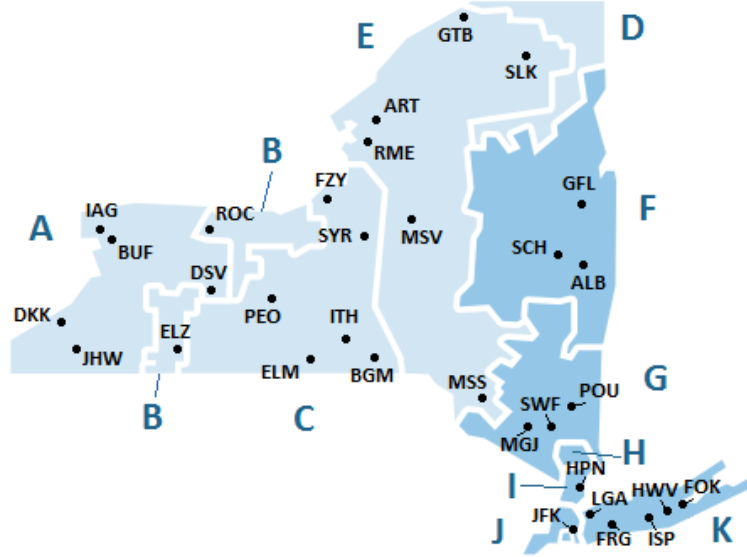


Figure 2.3: NYCA load zones depicted with letters A - K. Here, Zones A - F are referred to as "upstate zones" and Zones G - K as "downstate zones". Weather station locations are shown and identified by their three character abbreviations.

Using ambient temperature, wind speed, GHI, and DNI at the closest weather station, we employed NREL's solar PV System Advisor Model (PVSAM), a robust model for estimating electricity production from grid-connected solar PV arrays [19], to determine AC energy output at each site. PVSAM evolved and incorporated functionality from Sandia National Laboratory's photovoltaic array performance model [23], which was later improved by De Soto et al. [15], as well as a grid-connected inverter model [24]. This model accounts for multiple loss assumptions, reported in Table 2.2 in Section 2.5, in its estimate of electricity generation. Azimuth angles of 180° clockwise from due north and array tilts equal to latitude were assumed. Each farm was modeled in 2 - 3 MW blocks composed of CS6X-315P modules manufactured by Canadian Solar connected to a single SC2200-US 385V inverter manufactured by SMA Solar Technologies. This hardware pair was chosen because it appears in multiple

projects approved for construction within the NY-Sun database [1], which contains information about all operational solar projects across the state as well as those in the interconnection queue.

2.2.3 Net load profiles, peak demand reduction, and ramping rates

The NYISO operates NYS's power grid (also known as New York Control Area, NYCA) and wholesale electricity markets. There are eleven NYCA load zones (A - K) as marked in Figure 2.3. Net load profiles and ramping rates were calculated at a regional level bisecting NYS into upstate Zones A - F (shown in light blue in Figure 2.3) and downstate Zones G - K (shown in dark blue in Figure 2.3). Zones were grouped in this manner due to the presence of the Central-East transmission constraint within NYCA. When this constraint is active, no additional power can be transferred across the interface between upstate and downstate zones, thereby making it impossible for zones downstate to take advantage of resource flexibility within upstate zones. For example, if cooling load during a HEDD in NYC increases by 5 MW when the Central-East constraint is active, a resource within Zones G - K with an adequate response rate must be dispatched to meet this load even if there is excess electricity being generated by solar upstate that could otherwise balance this demand increase. Therefore, as a system operator, the NYISO must ensure that sufficient flexible ramping capability exists independently within each of these two regions to balance the bulk power system for all reasonable fluctuations in net load.

Since wholesale electricity prices are set at the sub-zonal level, peak demand

represents a more local issue. However, because most end-use customers pay zonal electricity prices, we chose to study peak reduction at the zonal level as well.

Net load profiles by NYISO zone

In order to analyze the impact of different solar PV penetration scenarios on NYISO zonal and regional load conditions, we calculated the hourly net load profiles by load zone under each scenario. Historical load data are publicly available through NYISO. Generation by each solar farm developed under a scenario was estimated using the methodology described in Section 2.2.2. Electricity produced by all solar farms within each load zone was then summed together and subtracted from the zonal load over the same period to obtain the net load, as shown in Equation 2.1,

$$\text{Zonal Net Load} = \text{NYISO Load} - \sum_{i=1}^N (\text{Solar Farm Generation})_i \quad (2.1)$$

where N is the number of solar farms selected for development within the load zone under the chosen PV scenario.

Zonal load duration curves and peak demand reduction

Net load profiles are used to compute load duration curves (LDCs) and peak demand reduction. LDCs are created by sorting net load from highest to lowest, depicting what fraction of the time various generation capacities are required. LDCs allow for long-term planning of what category of generation capacity – baseload, load following, or peaking – should be developed to satisfy forecasted demand changes.

Two distinct methods exist for defining peak demand reduction. A simple method takes total solar generation during a NYISO historical peak load hour. The other method is to define peak reduction as the difference, in MW, between the maximum of the LDCs constructed from the NYISO historical load and modeled net load during any given day, which accounts for shifting of peak load toward evening hours when solar resources provide less electricity. This ‘delayed peak’ is what the NYISO must accommodate during future system operations. The first method usually reports higher numerical values in peak demand reduction, but the second method was used because it provides a better metric to assess the peak demand reduction value of solar resources.

Ramping rate curves

Ramping rates were calculated by subtracting the net load during the previous hour from the net load during the current hour as shown in Equation 2.2.

$$\text{Ramping Rate} = \frac{(\text{Net Load})_t - (\text{Net Load})_{t-1}}{1 \text{ hour}} \quad (2.2)$$

Ramping rates were calculated (in MW/h) and plotted over a temporal period matching that for net load.

2.3 Results and Discussion

2.3.1 Spatial distribution of projected solar farm sites

As illustrated in Figure 2.4, different spatial distributions of solar PV farm sites resulted from each capacity penetration scenario. In this study, spatial distribu-

tion specifically refers to the organization of solar farms across the state. The 678 highest-ranked sites fulfill the 4500 MW capacity scenario. The number of sites under the 6000 MW and 9000 MW capacity scenarios are 920 and 1320, respectively.

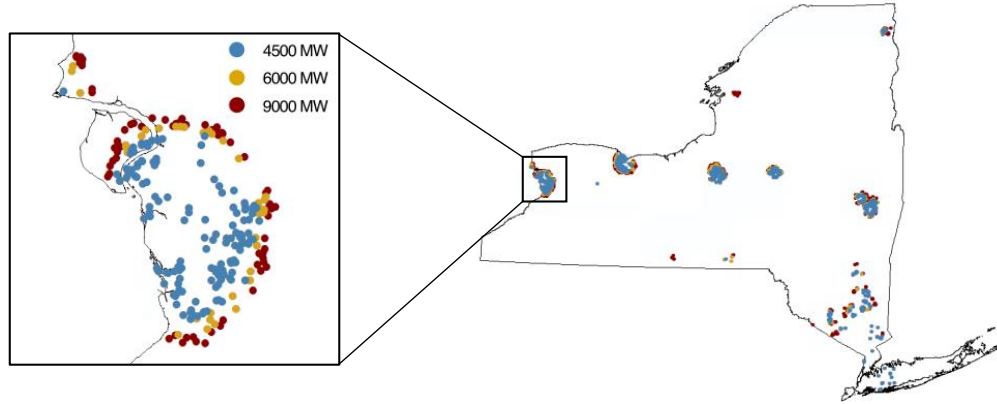


Figure 2.4: Solar farm distribution across NYS. Blue points correspond to sites selected under only the 4500 MW scenario. Gold points correspond to sites added moving to the 6000 MW scenario, and dark red points correspond to sites added under the 9000 MW scenario.

The map clearly shows that a sufficient quantity of low-value land exists across the state where solar PV could be deployed – a feature consistent with other GIS studies: Brewer et al. reported that between 43-71% of all land within California counties chosen for the study was suitable for solar PV development [13], and Yushchenko et al. reported a total PV generation potential of up to 686,686 TWh/year for West Africa [39]. Ranking sites primarily by population density resulted in many sites clustered tightly together, which could cause local over-generation decreasing the locational marginal price (LMP) of electricity at the sub-zonal level. However, such sub-zonal level effects are tempered when net load is aggregated at the zonal level. As mentioned previously, although wholesale prices are set at the sub-zonal level, most end-use customers pay the

zonal price justifying this aggregation. Regardless, future work must address sub-zonal level analysis of distribution systems with high solar PV penetrations. It should be noted that the distributions depicted depend on the filtering and ranking criteria presented in Section 2.2.1. Our study is focused on demonstrating an integrated approach inclusive of market values, not on predicting precisely where solar farms will be developed.

2.3.2 Implications for system flexibility and peak load reduction

Peak load requirements during summer

Table 2.4 in Section 2.5 lists the 40 highest electricity demand days, their peak load, and GHI values during the peak load hour on each day. Historically, the highest loads have occurred in early evening hours from June through September. To illustrate the effect that increased solar development will have upon system peak, we selected an extreme summer episode running from July 15-19, 2013, which overlaps with one of longest historical heat waves ever recorded – lasting for seven days. On July 19, 2013, NYS as well as NYC hit its highest historical electricity demand. For this episode, the LDC depicted in Figure 2.5c clearly shows that solar provides a valuable peak mitigation service for the upstate region (Zones A - F) during the summer season. Similar trends for the downstate region (Zones G - K) are observed in Figure 2.9c in Section 2.5. Under the 9000 MW penetration scenario, solar causes slight and infrequent displacement of baseload resources during mid-day in the upstate zones. Ramping rates, shown in Figures 2.5b and 2.9b, are lower during morning hours but in-

crease in the late afternoon for the upstate zones where the majority of solar resources are installed.

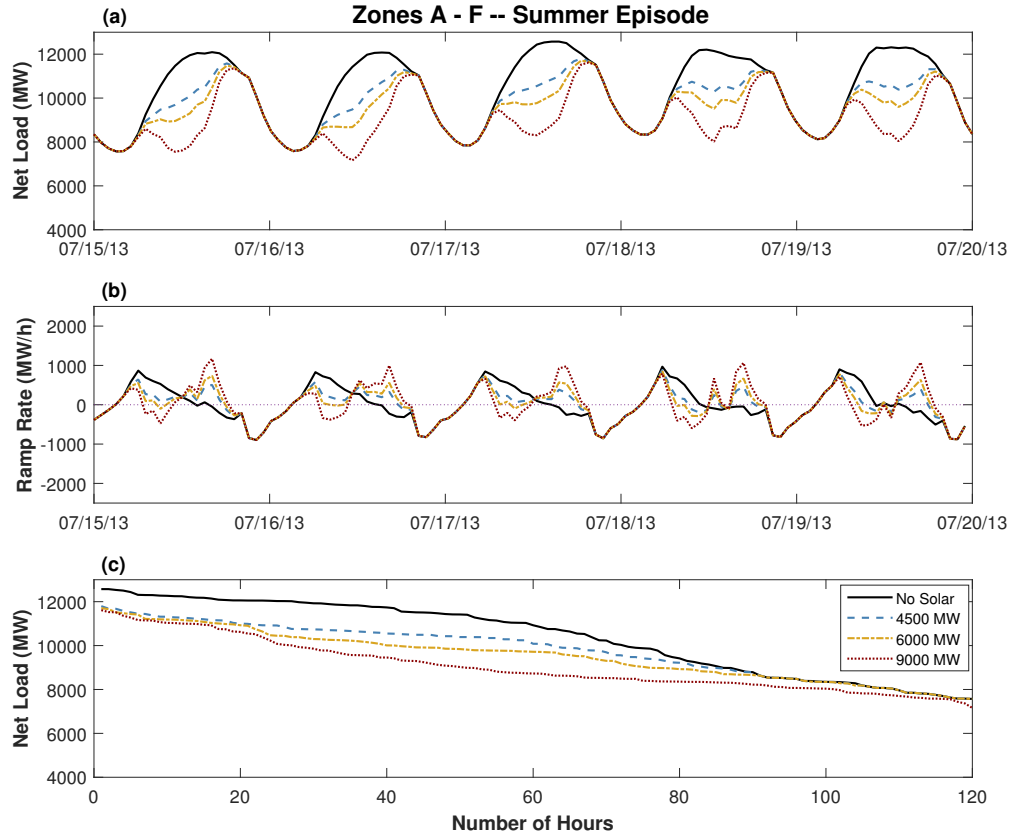


Figure 2.5: (a) Shows net load profiles for upstate Zones A - F under each of the three scenarios during the heat wave spanning from July 15-19, 2013. Peak load for this episode under the no solar scenario reached 12,573 MW. (b) Depicts ramping rates for Zones A - F over the same temporal period. Solar farms mitigate a significant portion of peak load without causing significant cycling of thermal power plants. The highest ramping rate calculated ($1,171 \text{ MW h}^{-1}$) occurred under the 9000 MW scenario. (c) Gives load duration curves for Zones A - F under each of the three scenarios during the heat wave spanning from July 15-19, 2013. These confirm that peak load is reduced and shows that little to no baseload capacity is displaced by solar generation from the baseline scenario.

Figure 2.7 provides a summary of historically observed peak loads and maximum ramping rates as well as those modeled for each scenario for both upstate

and downstate regions in NYS. Over 40 historical HEDDs (listed in Table 2.4), utility-scale solar PV resources reduce peak demand between 0.7-7.0%, 0.7-8.2%, and 0.7-9.6% among different zones under the 4500 MW, 6000 MW, and 9000 MW scenarios, respectively. Mean peak demand reduction values within each zone under each scenario are summarized in Table 2.3 in Section 2.5. Large deviations from the mean occur because output from a solar farm during a single hour is highly sensitive to meteorological conditions. For example, a sudden increase in local cloud coverage leads to a subsequent drop in solar PV output.

Ramping rates are physically constrained by an electricity network's existing infrastructure, and peak demand reduction does not guarantee a reduction in ramping rates. Although solar developers may seek to install solar PV in regions with favorable resources or high LMP to increase their revenue stream, system planners need to ensure that local power systems can accommodate changes in ramping rates. Depicted in Figure 2.7, the maximum ramping rate in the upstate region (Zones A - F) reached $1,096 \text{ MW h}^{-1}$ historically and could reach $1,694 \text{ MW h}^{-1}$ under the 9000 MW scenario. This ramping rate exceeds proven ramping capability by 55%. Therefore, utilities and operators should evaluate currently available ramping capability and estimate an optimal solar PV capacity – or upper limit – within each region given existing grid infrastructure. Utilizing current operational paradigms, the NYS electricity system can comfortably accommodate 4500 MW of additional solar during summer months, but increased flexibility becomes necessary before solar PV capacity reaches 6000 MW.

Ramping requirements during winter

During winter months, diurnal peak electricity demand falls outside the generation window for solar PV. As a result, solar generation provides no peak reduction but still affects ramping requirements. Therefore, it is prudent to analyze the impact of solar generation during the winter season. As in summer, we illustrate this impact via an extreme winter episode.

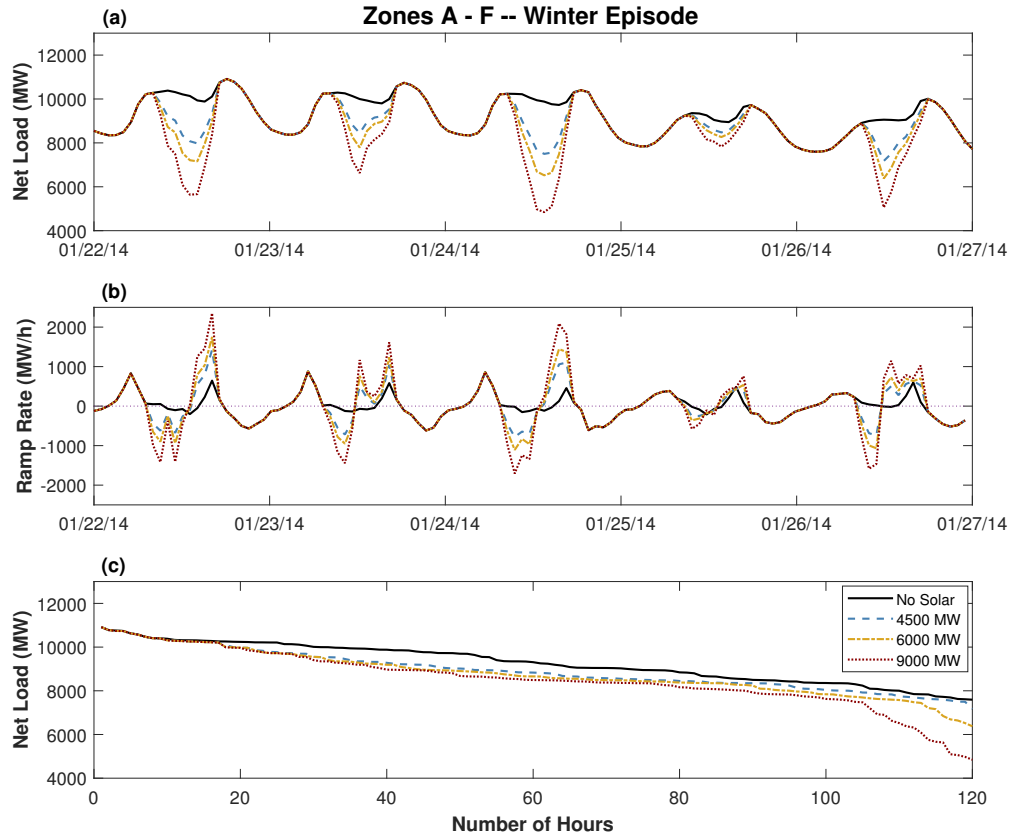


Figure 2.6: (a) Shows net load profiles for upstate Zones A - F under each of the three scenarios over a cold spell spanning from January 22-26, 2014. Peak loads as high as 10,914 MW were observed and are not reduced by solar PV generation. (b) Depicts ramping rates for Zones A - F over the same temporal period. The most drastic increases in required ramp rate occurs to meet the evening peak load, which has more than doubled on some days. The highest ramping rate – 2,356 MW h⁻¹ – occurs under the 9000 MW scenario. (c) Gives load duration curves for Zones A - F for each of the three scenarios over the winter cold spell spanning from January 22-26, 2014. No peaking capacity is displaced by solar PV generation, but a significant amount of baseload generation is displaced during mid-day if solar electricity is considered must-take.

A winter peak episode spanning from January 22-26, 2014 occurred during a cold spell in the wake of a moderate snow storm. Most places across NYS experienced temperatures remaining in the teens for the entire week. These freezing temperatures were accompanied by sunshine on several days which

can easily be seen in the net load curves plotted in Figure 2.6a for the upstate region, and Figure 2.10a for the downstate region. Daytime solar production is particularly evident on January 24. Accompanying strong solar production is an equally strong up-ramp as peak load hits after the sun sets as shown clearly in Figure 2.6b. As expected, solar generation does not aid in peak reduction during the winter as is evident from common maximum values in LDCs depicted in Figures 2.6c and 2.10c. However, solar generation under the 9000 MW scenario often displaces upstate baseload generators around mid-day during winter.

With solar displaying deep displacement of baseload power generation during the mid-day, load-following generators would need to be cycled in order to avoid any solar curtailment during this episode. As noted previously, this can lead to higher overall emissions if generators remain idle. Under the 9000 MW scenario, depicted in Figure 2.7, the maximum up-ramping rate in the upstate region (Zones A - F) could reach $2,754 \text{ MW h}^{-1}$ – a 151% increase from the maximum historical ramping rate observed within the region ($1,096 \text{ MW h}^{-1}$). Consequently, ramping requirements in the upstate region during winter represent the greatest concern for future system planning. However, as a greater fraction of the heating and transportation sectors become electrified, winter load patterns may shift. Conjointly, introduction of new technologies such as a flexible ramping product in the NYISO wholesale market, electrical energy storage, and fast-response natural gas facilities will aid in accommodating increased ramping requirements on the supply-side to a certain degree.

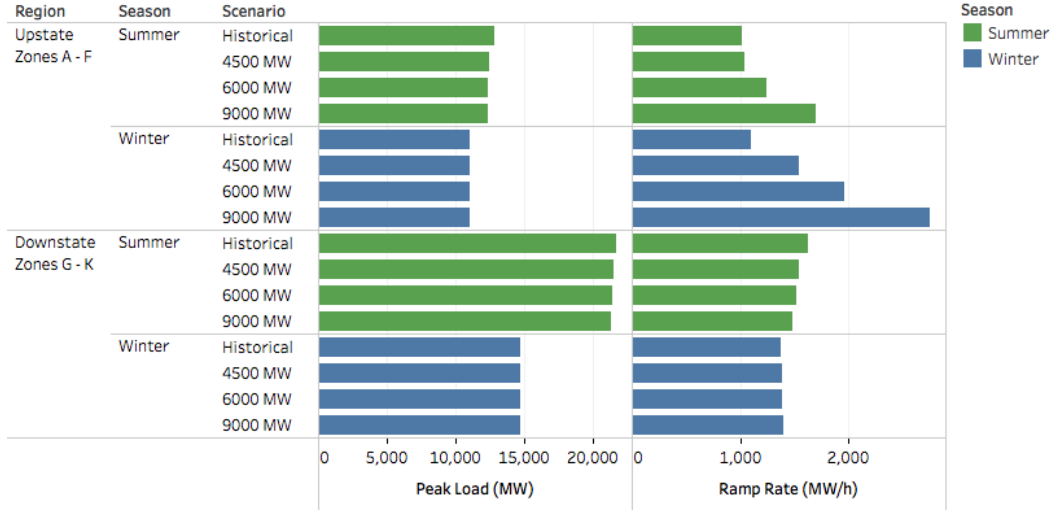


Figure 2.7: Bar chart depicting a summary of peak load (*left*) and ramp rate (*right*) split by region, season, and scenario over the entire temporal period from 2010 - 2015. Summer encompasses June through August, and winter encompasses November through March each year. Note that peak loads and maximum observed ramping rates do not generally occur during the same hour.

2.3.3 Implications for installed capacity valuation

In the draft VDER proposed by NYSDPS, there are three compensation options for the installed capacity of a DER project [41], which are directly associated with peak demand reduction and critical to the project’s financial viability. These options are: 1) a small \$/kWh credit for every kWh injected to the grid, 2) a higher \$/kWh value for each kWh injected to the grid between 2 and 6pm from June 1 through August 31 (i.e., summer peak hours), or 3) a lump sum payment in terms of \$/kW injected during the system peak hour (i.e., the one hour each year when the NYISO system peak occurs). The first option is proportional to the annual capacity factor. For the second option, we define the summer peak production factor (SPPF), as shown in Equation 2.3, to facilitate comparisons,

$$\text{SPPF}_i = \frac{\sum_{h \in H} (\text{Energy Generation})_h}{(\text{Nameplate Capacity}) \cdot (460 \text{ Hours})}, \quad (2.3)$$

where h represents each individual hour within the set H of 460 summer peak hours occurring between 2-6 pm, inclusive, spanning June 1 through August 31 during a single year i . For the third option, we define the peak hour production factor (PHPF), shown in Equation 2.4 below:

$$\text{PHPF}_i = \frac{\text{Energy Generation}_{ph}}{(\text{Nameplate Capacity}) \cdot (1 \text{ Hour})}, \quad (2.4)$$

where ph represents the single hour for a given year i when the NYISO system peak occurred.

Figure 2.8 shows the mean values (as points) and ranges (as error bars) of annual capacity factors, SPPFs, and PHPFs calculated for each year between 2010 and 2015 at all NYS meteorological stations (locations shown in Figure 2.3). Mean values of SPPFs at individual sites are within a narrow range between 0.24 and 0.33 – much higher than the corresponding annual capacity factors ranging from 0.11 to 0.16. Thus, on average, a solar farm in NYS contributes between 24-33% of its capacity toward summer peak demand reduction. Mean SPPFs (inclusive only of peak hours) are 65-171% higher than the average annual capacity factors (inclusive of all hours), confirming that favorable solar resources often coincide with peak load hours. PHPF values differ substantially among different years at any given location and across different stations ranging from 0.10 to 0.74. Such variations remain unsurprising as cloud cover drastically decreases output during unlucky hours.

Our analysis implies several key challenges in developing sound valuation methodologies for solar projects from a system planning perspective. Ideally, a competitive market for solar development should incentivize projects to perform well during system peak hours thereby displacing power produced from the most expensive fossil-fueled generators. As the annual capacity factors and

SPPFs (corresponding to the first and second options described above) experience smaller fluctuations from year to year, these metrics provide solar developers with desired financial consistency. However, neither of these two options truly capture the role that solar resources play in meeting system peak demand – the underlying purpose of the installed capacity market. The merit of the traditionally required third option is its contingency on the single hour of the year when demand peaks. Therefore, the effect of resource intermittency is implicitly considered as evinced by the significant variability observed in PHPFs. Unfortunately, since the hour of the system-wide peak is unpredictable beforehand and long-term solar forecasting remains wildly inaccurate, option three proves too volatile to act as an effective market signal to solar developers.

As such, a combination of two options (options 1 and 3, or options 2 and 3) may be more effective than simply exempting DER generators completely from valuation based on system peak performance – as currently proposed by VDER. However, it is important to bear in mind that adding additional complexity may erect a barrier hindering further solar development. The overall approach presented in this chapter can assist system planners and solar developers with strategic resource siting. In this way, clear incentives may be designed that ensure adequate system capacity, alleviate congestion, and provide developers with predictable revenue streams.

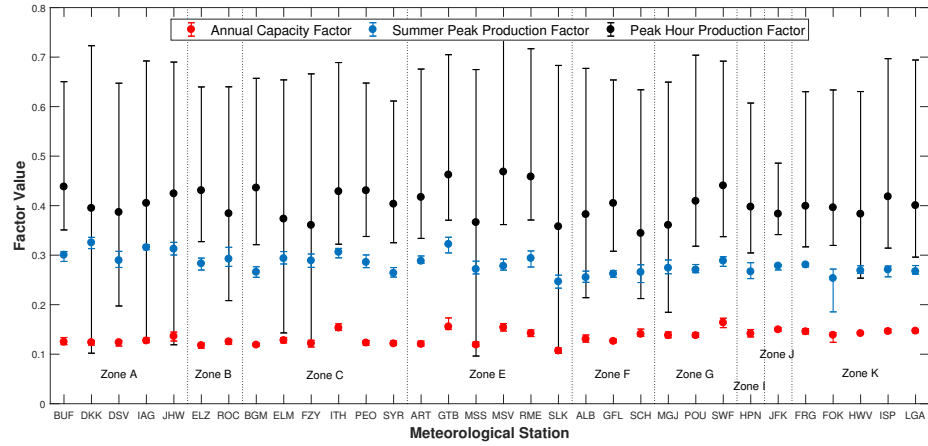


Figure 2.8: Scatter plot with points representing the mean value of the annual capacity factor, the summer peak production factor (SPPF), and the peak hour production factor (PHPF). Error bars show the range of each factor, which was calculated using electricity generated by a 2 MW solar farm. Single factors were calculated for each year between 2010-2015 at each of 32 NYS meteorological stations. Stations are sorted by zone.

2.4 Conclusions

In this chapter, we assembled a set of comprehensive and spatially granular datasets, combined with meteorological and electricity system data, to assess the impact of utility-scale solar on power systems. The bottom-up approach we adopted estimated the spatial distribution of solar farms across NYS in GIS using tax classification data, electricity network infrastructure data, and census data to determine the effect of solar farm development on peak load and ramping rate within different regions using real-world meteorological data from historical events.

We showed that on average, a solar farm in NYS generates between 24-33% of its nameplate capacity during peak demand hours and confirmed that favor-

able solar resources are generally correlated with high energy demand hours. In addition, peak demand reduction displays locational and temporal dependency. During the system peak hours between 2010 and 2015, a solar farm generated anywhere between 10 and 74% of capacity. Such variability in performance during system peak creates uncertainty in a solar farm's installed capacity payment thereby making it more difficult to obtain project financing. Therefore, an alternative installed capacity compensation mechanism that values all summer peak hours in addition to the system peak and sends a more predictable market signal to developers should be established.

Our results clearly demonstrated the importance of evaluating ramping requirements alongside peak demand reduction. While solar generation often reduces peak load during summertime, it tends to displace baseload resources during wintertime. Examining historical summer peak and winter peak episodes indicated that maximum ramping rates will likely take place during wintertime under high solar penetration scenarios. As such, system flexibility constraints during wintertime should be assessed.

Due to the highly uncertain nature of local public support surrounding solar farms, a shortcoming of this and related research is its inability to consider the social science aspect of solar siting. Clearly, public opinion must be characterized fully before a truly accurate spatial distribution of solar farms can be constructed. Further, substation location was the only electrical infrastructure data used in this methodology. Future work will consider more detailed treatment of electrical infrastructure, imports and exports to and from adjacent electricity markets, and the role that electrical energy storage may play in enhancing flexibility of grid-connected solar PV. Finally, future work will also

encompass micro-meteorological conditions by supplementing meteorological data obtained from weather stations with that obtained via numerical weather prediction.

Although results presented here use NYS data, the approach can be generalized to aid in developing or assessing compensation methodologies for the demand reduction value and installed capacity value of solar projects in other deregulated electricity markets.

References

- [1] Solar Electric Programs Reported by NYSERDA: Beginning 2000. Technical report, New York State Energy Research and Development Authority, Albany, NY.
- [2] Transmission Atlas™, 2013.
- [3] Public NYS Tax Parcels. Technical report, New York State Department of Taxation and Finance, Albany, NY, 2015.
- [4] Capital Cost Estimates for Utility Scale Electricity Generating Plants. Technical report, U.S. Energy Information Administration, Washington, D.C., 2016.
- [5] Solar Impact on Grid Operations - An Initial Assessment. Technical report, The New York Independent System Operator, Rensselaer, NY, 2016.
- [6] 2017 Load & Capacity Data "Gold Book". Technical report, New York Independent System Operator, 2017.

- [7] Power Trends 2017. Technical report, The New York Independent System Operator, Rensselaer, NY, 2017.
- [8] Jamshid Aghaei and Mohammad-Iman Alizadeh. Demand response in smart electricity grids equipped with renewable energy sources: A review. *Renewable and Sustainable Energy Reviews*, 18:64–72, 2013.
- [9] Mudathir Funsho Akorede, Hashim Hizam, and Edris Pouresmaeil. Distributed energy resources and benefits to the environment. *Renewable and Sustainable Energy Reviews*, 14(2):724–734, 2010.
- [10] Arturo Alarcon-Rodriguez, Graham Ault, and Stuart Galloway. Multi-objective planning of distributed energy resources: A review of the state-of-the-art. *Renewable and Sustainable Energy Reviews*, 14:1353–1366, 2010.
- [11] Brian N Belcher and Arthur T DeGaetano. Integration of asos weather data into model-derived solar radiation. *ASHRAE Transactions*, 111(1), 2005.
- [12] Severin Borenstein. The Market Value and Cost of Solar Photovoltaic Electricity Production. Technical report, University of California Energy Institute, Berkeley, CA, 2008.
- [13] Justin Brewer, Daniel P Ames, David Solan, Randy Lee, and Juliet Carlisle. Using GIS analytics and social preference data to evaluate utility-scale solar power site suitability. *Renewable Energy*, 81:825–836, 2015.
- [14] Donald Chung, Carolyn Davidson, Ran Fu, Kristen Ardani, and Robert Margolis. U.S. Photovoltaic Prices and Cost Breakdowns: Q1 2015 Benchmarks for Residential, Commercial, and Utility-Scale Systems. Technical report, National Renewable Energy Laboratory, Golden, CO, 2015.

- [15] W De Soto, S A Klein, and W A Beckman. Improvement and validation of a model for photovoltaic array performance. *Solar Energy*, 80:78–88, 2006.
- [16] Paul Denholm and Maureen Hand. Grid flexibility and storage required to achieve very high penetration of variable renewable electricity. *Energy Policy*, 39:1817–1830, 2011.
- [17] J. Driesen and F. Katiraei. Design for distributed energy resources. *IEEE Power and Energy Magazine*, 6(3):30–40, 2008.
- [18] Vasilis Fthenakis and Hyung Chul Kim. Land use and electricity generation: A life-cycle analysis. *Renewable and Sustainable Energy Reviews*, 13:1465–1474, 2009.
- [19] P Gilman. SAM Photovoltaic Model Technical Reference. Technical report, National Renewable Energy Laboratory, Golden, CO, 2015.
- [20] Jiajun Gu, Jeffrey A Sward, Ye Lin Kim, Jessica Spaccio, Arthur T. DeGaetano, and K. Max Zhang. Resources Climatology: An analysis of renewable resources droughts and implications on distributed energy systems resiliency. 2019.
- [21] Gang He and Daniel M Kammen. Where, when and how much solar is available? A provincial-scale solar resource assessment for China. *Renewable Energy*, 85:74–82, 2016.
- [22] Lion Hirth. Market value of solar power: Is photovoltaics cost-competitive? *IET Renewable Power Generation*, 9(1):37–45, 2014.
- [23] D L King, W E Boyson, and J A Kratochvill. Photovoltaic Array Performance Model. Technical report, Sandia National Laboratories, Albuquerque, NM, 2004.

- [24] David L King, Sigifredo Gonzalez, Gary M Galbraith, and William E Boyson. Performance Model for Grid-Connected Photovoltaic Inverters. Technical report, Sandia National Laboratories, 2007.
- [25] Eugene L. Maxwell. A Quasi-Physical Model for Converting Hourly Global Horizontal to Direct Normal Insolation. Technical report, Solar Energy Research Institute, Golden, CO, 1987.
- [26] Gerald A. Meehl and Claudia Tebaldi. More Intense, More Frequent, and Longer Lasting Heat Waves in the 21st Century. *Science*, 305(2004).
- [27] Thomas Nikolakakis and Vasilis Fthenakis. The optimum mix of electricity from wind-and solar-sources in conventional power systems: Evaluating the case for New York State. *Energy Policy*, 39:6972–6980, 2011.
- [28] Sean Ong, Clinton Campbell, Paul Denholm, Robert Margolis, and Garvin Heath. Land-Use Requirements for Solar Power Plants in the United States. Technical report, National Renewable Energy Laboratory, Golden, CO, 2013.
- [29] Robert Carl Pietzcker, Daniel Stetter, Susanne Manger, and Gunnar Luderer. Using the sun to decarbonize the power sector: The economic potential of photovoltaics and concentrating solar power. *Applied Energy*, 135:704–720, 2014.
- [30] Abhnil A. Prasad, Robert A. Taylor, and Merlinde Kay. Assessment of solar and wind resource synergy in Australia. *Applied Energy*, 190:354–367, 2017.
- [31] Cynthia Rosenzweig, William Solecki, Arthur DeGaetano, Megan O’Grady, Susan Hassol, and P. Grahborn. Responding to Climate Change in New York State: the ClimAID Integrated Assessment for Effective Climate

Change Adaptation. Synthesis Report. Technical report, The New York State Energy Research and Development Authority, Albany, NY, 2011.

- [32] Juan M Sánchez-Lozano, Jerónimo Teruel-Solano, Pedro L Soto-Elvira, and M Socorro García-Cascales. Geographical Information Systems (GIS) and Multi-Criteria Decision Making (MCDM) methods for the evaluation of solar farms locations: Case study in south-eastern Spain. *Renewable and Sustainable Energy Reviews*, 24:544–556, 2013.
- [33] Andrea Sarzynski, Jeremy Larrieu, and Gireesh Shrimali. The impact of state financial incentives on market deployment of solar technology. *Energy Policy*, 46:550–557, 2012.
- [34] Larry J Sugarbaker, Diane F Eldridge, Allyson L Jason, Vicki Lukas, David L Saghy, Jason M Stoker, and Diana R Thunen. Status of the 3D Elevation Program, 2015 Open-File Report 2016-1196. Technical report, 2017.
- [35] Michaelangelo D Tabone, Christoph Goebel, and Duncan S Callaway. The effect of PV siting on power system flexibility needs. *Solar Energy*, 139:776–786, 2016.
- [36] Tereje Gjengedal. Emission Constrained Unit-Commitment (ECUC). *IEEE Transactions on Energy Conversion*, 11(1996).
- [37] Corey D White and K Max Zhang. Using vehicle-to-grid technology for frequency regulation and peak-load reduction. *Journal of Power Sources*, 196:3972–3980, 2011.
- [38] Ryan Wiser, Dev Millstein, Trieu Mai, Jordan Macknick, Alberta Carpenter, Stuart Cohen, Wesley Cole, Bethany Frew, and Garvin Heath. The environ-

mental and public health benefits of achieving high penetrations of solar energy in the United States. *Energy*, 113:472–486, 2016.

- [39] Alisa Yushchenko, Andrea de Bono, Bruno Chatenoux, Martin Kumar Patel, and Nicolas Ray. GIS-based assessment of photovoltaic (PV) and concentrated solar power (CSP) generation potential in West Africa. *Renewable and Sustainable Energy Reviews*, 81:2088–2103, jan 2018.
- [40] Xiyue Zhang and K. Max Zhang. Demand response, behind-the-meter generation and air quality. *Environmental Science and Technology*, 49(3):1260–1267, 2015.
- [41] Audrey Zibelman, Gregg C. Sayre, and Diane X. Burman. Order on Net Energy Metering Transition, Phase One of Value of Distributed Energy Resources, and Related Matters, 2017.

2.5 Supporting Information

Table 2.1 shows which property classes where solar farms have been developed in the past as well as those were deemed adequate to support utility-scale solar farms in the future. Less than 1% of properties within the tax dataset have no assigned property class. A limitation of this dataset is that only one property class can be assigned to each property, so only the code for the property class that takes up a majority of the land is reported. Detailed information about all NYS property codes can be obtained from the New York State Department of Taxation and Finance.

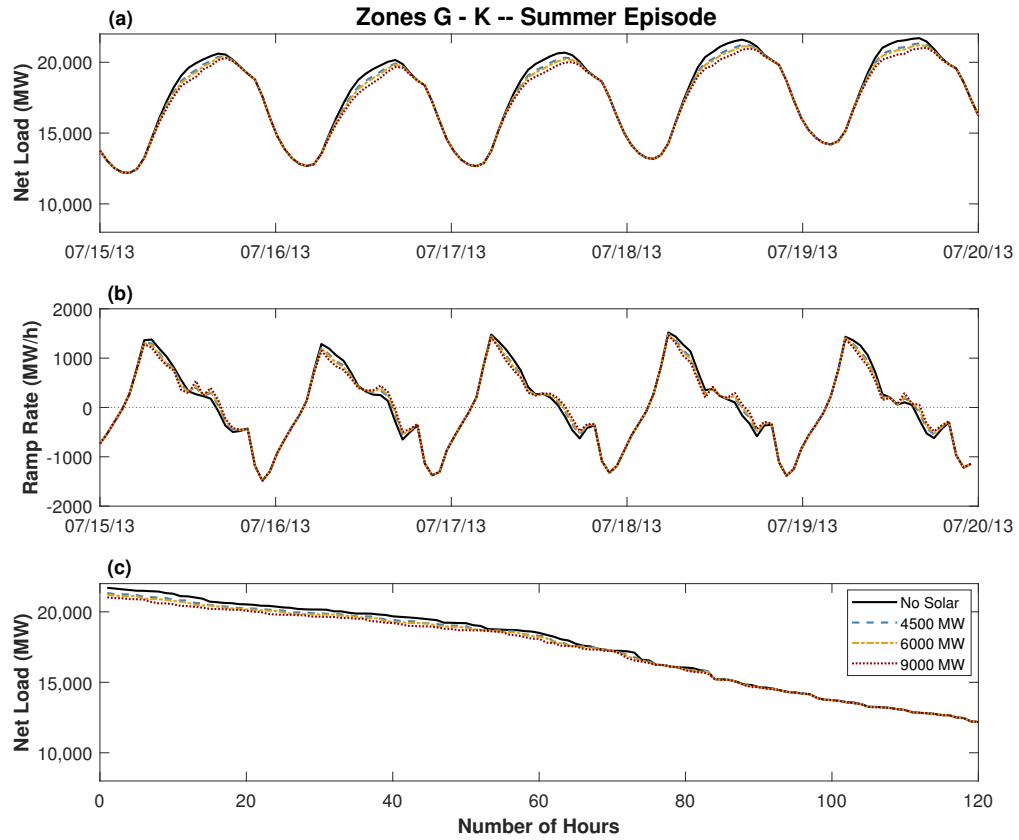


Figure 2.9: (a) shows net load profiles for downstate Zones G - K under each of the three scenarios during the heat wave spanning from July 15-19, 2013. Solar generation clearly has a much smaller effect on net load downstate as far fewer suitable properties are available here. Peak load for this episode under the no solar scenario reached 21,705 MW. (b) depicts ramping rates in Zones G - K under each of the three scenarios over the same temporal period. Highest ramps remain during the morning ramp, and the highest ramp rate ($1,520 \text{ MW h}^{-1}$) occurred under the no solar scenario. (c) gives load duration curves for Zones G - K under each of the three scenarios during the heat wave spanning from July 15-19, 2013. Peak load is reduced by a modest amount when compared with the upstate zones.

Table 2.1: The first two columns give property class codes and descriptions for land either on which existing solar farms are located or that assumed suitable for future utility-scale solar PV farm development. The third column identifies if a property class has seen solar development, is assumed to see solar development in the future, or both. Lands classified under all codes without “Future” or “Both” listed in column three were excluded. Column four provides a justification for discrepancies between historically observed and future solar development for a property class.

Code	Description	Existing/ Future	Justification
140	Truck Crops – Not Muck-lands	Existing	Future protection expected
240	Rural Residence with Acreage	Both	N/A
241	Primary residential, also used in agricultural production	Existing	Future protection expected
311	Residential Vacant Land	Future	Can be repurposed with PV
320	Rural	Future	Catch all category
321	Abandoned Agricultural Land	Future	Solar already exists on productive farm land
322	Residential Vacant Land Over 10 Acres	Both	N/A
323	Other Rural Vacant Lands	Both	N/A
330	Vacant Land Located in Commercial Areas	Both	N/A
340	Vacant Land Located in Industrial Areas	Future	Similar to 340
613	Colleges and Universities	Existing	Does not generally have enough land for PV development
652	Office Building	Existing	Does not generally have enough land for PV development
714	Light Industrial Manufacturing and Processing	Both	N/A
720	Mining and Quarrying	Future	Can be repurposed with PV after retirement
852	Landfills and Dumps	Both	N/A
877	Electric Power Generating Facility – Other Fuel	Existing	Only classified as this after construction

Table 2.2: Loss factors assigned within NREL’s System Advisor Model PV Module

Description	Loss
Average soiling	5.0%
Module mismatch	2.0%
Diodes and connections	0.5%
DC wiring	2.0%
Tracking error	0.0%
Nameplate	0.0%
DC power optimizer loss	0.0%
AC wiring	1.0%

Table 2.3: Mean and standard deviation of peak demand reduction values for each load zone. Mean taken over the peak demand reduction on the 40 highest electricity demand days, which explains the high standard deviation values (i.e., cloud cover and therefore solar irradiance on each high electricity demand day shows a great deal of variability). All values are rounded to the nearest MW. Values in zones H,I, and K do not change because so few sites were selected that all available sites in these zones were developed under each scenario. No sites in zone J were selected using the current methodology.

Load Zone	4500 MW	6000 MW	9000 MW
Zone A	179(±54)	199(±57)	227(±63)
Zone B	135(±37)	157(±42)	185(±42)
Zone C	155(±56)	180(±61)	203(±66)
Zone D	15(±9)	15(±9)	16(±10)
Zone E	70(±37)	80(±42)	92(±45)
Zone F	103(±57)	121(±61)	138(±63)
Zone G	118(±58)	144(±71)	173(±81)
Zone H	14(±7)	14(±7)	14(±7)
Zone I	3(±1)	3(±1)	3(±1)
Zone J	0(±0)	0(±0)	0(±0)
Zone K	38(±11)	38(±11)	38(±11)

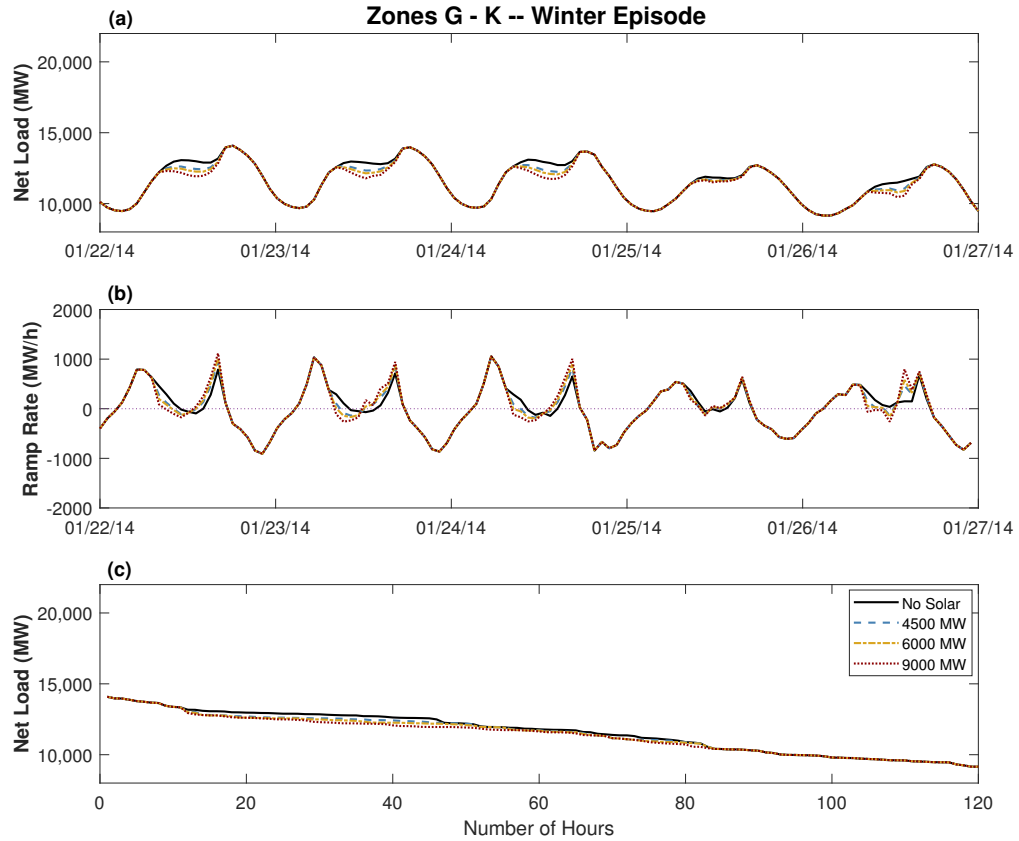


Figure 2.10: (a) shows net load profiles for downstate Zones G - K under each of the three scenarios during a cold spell spanning from January 22-26, 2014. Peak loads as high as 14,086 MW were observed and are not reduced by solar PV generation. (b) depicts ramping rates in Zones G - K over the same temporal period. Only small changes in ramping rates are seen in this region, but the highest ($1,117 \text{ MW h}^{-1}$) occurs under the 9000 MW scenario. (c) gives load duration curves for Zones G - K for each of the three scenarios over a cold spell spanning from January 22-26, 2014. No peaking capacity or baseload capacity will be displaced by solar in this region under any of the considered scenarios.

Table 2.4: Highest historical electricity demand days in NYS.

Rank	Date	Hour
1	July 19, 2013	17
2	July 22, 2011	16
3	July 21, 2011	17
4	July 6, 2010	17
5	July 18, 2013	17
6	July 17, 2013	17
7	July 7, 2010	16
8	July 15, 2013	17
9	July 17, 2012	17
10	July 16, 2013	17
11	July 18, 2012	14
12	June 21, 2012	15
13	July 12, 2011	15
14	Sept 2, 2010	16
15	Sept 1, 2010	17
16	July 8, 2010	17
17	June 20, 2012	17
18	Aug 31, 2010	17
19	July 20, 2011	17
20	Sept 11, 2013	17
21	July 29, 2015	17
22	Aug 5, 2010	17
23	Sept 8, 2015	17
24	June 29, 2012	17
25	July 16, 2012	17
26	Aug 4, 2010	17
27	June 9, 2011	14
28	July 11, 2011	17
29	June 8, 2011	17
30	July 20, 2015	16
31	July 6, 2012	17
32	July 19, 2011	18
33	Aug 17, 2015	17
34	July 5, 2012	16
35	July 16, 2010	17
36	July 28, 2015	17
37	July 23, 2013	17
38	Aug 1, 2011	16
39	July 8, 2013	17
40	June 28, 2010	14

CHAPTER 3

**SPATIAL BIASES REVEALED BY LIDAR IN A MULTIPHYSICS WRF
ENSEMBLE DESIGNED FOR OFFSHORE WIND**

Abstract

Numerical weather predictions (NWP) have become essential in offshore wind energy planning and operations. Thus, rigorous assessments of NWP model performance are critical to integrating offshore wind power into existing power systems. Taking advantage of two LiDAR buoys launched off the coast of New York in 2019, we assess the performance of a multiphysics Weather Research and Forecast (WRF) model ensemble with a 1.33-km spatial resolution for estimating the power system impacts associated with New York's offshore wind target. Our work is the first to report WRF horizontal wind speed biases not only at multiple heights above sea level but at two locations while still considering all seasons. WRF tends to overpredict wind speeds during spring and summer and underpredict wind speeds during winter. However, the patterns in wind speed biases differ substantially between the two buoys offering compelling evidence against spatially uniform biases, which impacts the performance of numerous bias correction methods frequently used to post-process WRF data. Therefore, additional measurements of wind speeds throughout the lower atmosphere are necessary to fully characterize bias patterns. With the recent goal set by the U.S. to install 30 GW of offshore wind by 2030 – largely along the East Coast, mispredictions carry important policy implications. Absent accurate offshore wind uncertainty forecasts, power system operators throughout the Eastern Interconnection will be forced to dispatch their most expensive and likely high emitting

power plants to compensate for periods of underperformance.

3.1 Introduction

Despite staggering global contractions across many industries, solar and wind developers set records in 2020 [28]. The United States was no exception adding 19.2 GW of solar [44] and 14.2 GW of wind [13], which accounted for 81% of new capacity. However, even with renewable portfolio standards in place and access to technologies proven in other parts of the world, the United States remains behind when it comes to offshore wind. Two proof-of-concept projects – the Block Island Wind Farm totaling 30 MW and the Coastal Virginia Offshore Pilot project totaling 12 MW – account for all the offshore wind connected to the US power grid. Still, projects at various stages of the offshore wind development pipeline have swelled to 35.3 GW [38]. Given the multi-year regulatory and industrial timelines associated with these projects, governments and non-profit agencies need robust assessment tools today in order to analyze the impacts on energy and social systems as this transition happens.

As observational wind speed data above the surface level are difficult to come by, numerical weather predictions (NWP) provide vital insight into boundary layer meteorology and wind power prediction. In particular, the Weather Research and Forecasting (WRF) model [45] is ideally suited for conducting everything from short-range forecasts [51, 6] to regional wind resource assessments [11, 32]. Draxl et al. used WRF to create the WIND Toolkit for the entire United States [11]. WRF’s flexible framework can be attributed to the staggering number of schemes available for parameterizing subgrid-scale pro-

cesses, i.e., physical phenomena that are not directly resolved by the model at the user-specified grid spacing. For wind energy analyses, the choice of planetary boundary layer (PBL) scheme likely has the highest impact on model results as it controls how profiles are represented in the lowest part of the atmosphere. Specifically, turbulence is mostly a subgrid-scale process in mesoscale models, and PBL schemes seek to capture the turbulent fluxes of heat, moisture, and momentum that occur in the lower troposphere [9]. Previous work has looked at how well different PBL schemes simulate winds under different atmospheric stability conditions [12], conducted a WRF physics sensitivity analysis for winds over complex terrain [14], created ensembles to improve wind speed prediction near turbine hub height [10], and reported the value of downscaling various global datasets with WRF for regional wind resource assessments [16]. Fitch et al. introduced a wind farm parameterization (WFP) that represents an individual wind turbine with a turbulent kinetic energy source and momentum sink, which captures turbine-turbine interactions, as well as the minor effect that wind farms have upon mesoscale meteorology [15], which has subsequently been evaluated for both onshore [33] and offshore [29] applications.

The number of studies applying WRF to offshore wind has exploded in recent years compelling Banta et al. to suggest that offshore wind energy can provide a backdrop for making improvements to NWP models[4]. Broadly, these studies fall into two main categories a) assessment, sensitivity, or validation of offshore winds and b) longer-term wind resource assessments (e.g., regional wind climatologies); often, the wind resource assessments offer some validation thereby covering both categories. Gryning and Floors compare WRF forecasts and downscalings to a LiDAR buoy in the North Sea [22]. Giannakopoulou and Nhili assess four PBL schemes and two reanalysis products, ERA-Interim and

NCEP FNL, using data from a 100 m mast in the North Sea [17]. A similar validation was performed for an ERA-Interim downscaling using a larger domain in the same region [5]. Kikuchi et al. use measurements from an offshore mast in Japan to nudge WRF improving the modeled wind speeds [30]. Comparing WRF to LiDAR measurements, Goit et al. found that WRF tends to overestimate wind speeds at higher altitudes off the coast of Japan [21]. Several studies also validate offshore wind produced by WRF using surface-level wind speed measurements at buoys extrapolated to wind turbine hub heights [8]. Offshore wind resource assessments were conducted using WRF for the North Sea [26, 35], the North and Baltic Seas [23], Europe [25], Alaska [32], Northwest India [31], China [34, 26], Japan [46], and Chile [36].

In the United States, Archer et al. expressed the need for more offshore observations and uncertainty characterization [3] later estimating offshore wind forecast errors at wind turbine hub height by analyzing the forecast errors at 23 onshore wind farms across the U.S. East Coast [1]. Separately, Archer et al. analyzed data collected at a meteorological mast during the 2000s in Nantucket Sound and reported that unstable atmospheric conditions dominate [2]. Pichugina et al. stress the importance of high-resolution LiDAR measurements by comparing two forecast models – the NCEP Rapid Refresh (RAP) and the North American Mesoscale Forecast System Rapid Refresh (NAMRR) – each run over the Gulf of Maine with and without ingesting data from ship and land-based LiDAR measurements [41].

While many WRF studies consider offshore wind or the Northeastern U.S., few look at the intersection of these, and none have validated WRF wind speeds using continuously available LiDAR data. Turbines rely on winds in the low-

est part of the atmosphere where turbulent mixing creates complicated flow patterns not only in three-dimensional space but across time as well. Therefore, a complete model validation should ideally assess each of these dimensions. Of course, previous studies have made use of available data often relying on land-based measurements, surface wind speed measurements, or reanalysis data products, which cannot capture the complicated spatiotemporally varying behavior of wind speeds aloft. Here, we validate a multiphysics ensemble of WRF models, each member (i.e., WRF simulation with a distinct set of physics parameterizations) of which has been used for wind studies previously, by comparing them to data from the two LiDAR buoys off the coast of New York. The New York State Energy Research and Development Authority (NY-SERDA) funded the LiDAR buoys, which went online in August and September 2019, respectively [18, 19]. These buoys give us insight into the trends in WRF horizontal wind speed biases in four dimensions improving on previous data-limited validation work.

3.2 Method

This section covers two major topics. First, we explain the WRF model, the multiphysics ensemble created for this study, and how we chose representative simulation periods. Then, we cover model validation – both from the perspective of individual members and from the ensemble as a whole.

3.2.1 Offshore Wind Observational Data

In this study, we leverage a unique public offshore wind dataset that provides wind speed observations up to 200 m. This data comes from two EOLOS FLS-200 buoys that NYSERDA contracted DNV GL to deploy in the New York-New Jersey Bight. The locations of these two buoys are shown in panel a) of Figure 3.1. For the remainder of this chapter, we will refer to the buoy further to the north and east as the “north buoy” and the buoy further to the south and west as the “south buoy.” The image in panel b) on the right in Figure 3.1 shows what these buoys look like floating on the ocean. Wind speed measurements are available every 20 m up to 200 m above sea level (a.s.l.). Before deployment, each buoy was tested to ensure that mean wind speeds did not deviate by more than 1% from benchmark observations and that at least 90% of measurements had an absolute wind speed difference of within 5% and less than 0.5 m s^{-1} from the benchmark observations [18, 19]. DNV GL reported a systematic uncertainty on the order of 2% for the floating LiDAR buoys for all heights that they compared from wind speeds between $4 - 16 \text{ m s}^{-1}$. To the best of our knowledge, this is the longest-running offshore wind dataset providing observations at multiple heights and multiple locations within the same region anywhere in the United States.

3.2.2 WRF Ensemble Setup

WRF’s flexibility poses both challenges and opportunities. A staggering number of possible customizations make WRF the go-to tool for forecasting and downscaling studies alike across myriad diverse regions and applications.

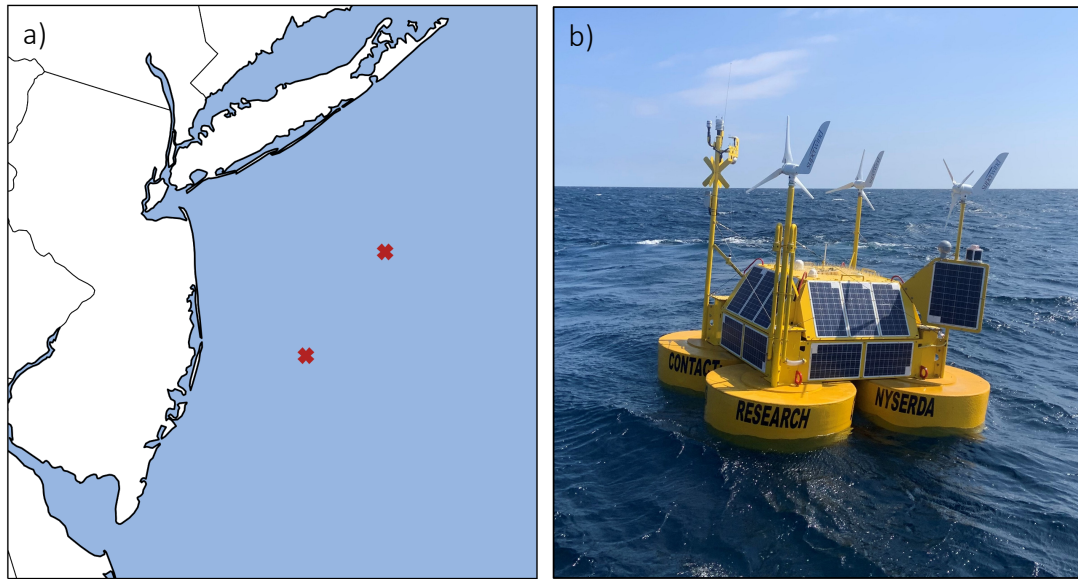


Figure 3.1: a) on the left shows the locations of the north buoy and the south buoy, respectively. b) on the right shows one of the EOLOS FLS-200 buoys in operation (Photo provided by Ocean Tech Services Inc).

However, tracking WRF's performance over a set of studies considering the same region or application becomes difficult as model setups are sometimes institutionalized and gradually improved over multiple projects. This is especially prevalent in the case where institutional models are passed down from an earlier study.

Here, we create an ensemble using WRF physics setups that already appear in the WRF offshore wind literature. Five major physical processes cannot be resolved by the WRF model and are instead parameterized. These include microphysics, radiation (separate schemes handle longwave (LW) and shortwave (SW) radiation), the land surface (LSM), the PBL, and cumulus clouds. Note that WRF run with a grid spacing below about 4 km begins to resolve clouds directly, so we turn off the cumulus scheme below this grid spacing threshold. A separate surface layer scheme parameterizes the fluxes exchanged between the land surface and the PBL, but the choice of the surface layer scheme is heav-

ily constrained by the PBL scheme. Several more minor parameterizations are also available in WRF (e.g., lake physics and urban canopy), but we omit these because they should have little to no effect on winds offshore, and therefore, it is likely that these were left at the WRF defaults in previous studies.

Table 3.1 provides information about parameterization schemes used in the five WRF setups from the literature adopted for this work. Each of these setups was chosen for a specific reason. Optis et al. [40] and Veron et al. [50] are the only studies that use WRF for offshore wind in the Northeastern US. The domain configuration used by Optis et al. is similar to the one we use in this work except that our finest domain has a horizontal resolution of 1.33 km rather than 3 km. We included the Lee and Lundquist setup [33] because, although they conducted their validation over the continental US, their work is currently the most comprehensive evaluation of the WRF WFP, which can easily be implemented offshore. The final two setups were both taken from Draxl et al. [12]. The difference between these two setups lies in the choice of PBL and surface layer parameterization schemes. Specifically, Draxl et al. found that the setup with the YSU PBL scheme (we call this setup Draxl 2014a) outperforms the rest when the atmosphere is unstable at Høvsøre, and the setup with the MYJ scheme (we call this setup Draxl 2014b) performs the best during stable and very stable conditions. Given that previous work reported a high prevalence of unstable atmospheric conditions off the coast of Massachusetts [2], we wanted to investigate if these trends in PBL scheme performance hold in the Northeastern US as well.

Table 3.1: WRF Model Setups

	Optis 2021 [40]	Veron 2018 [50]	Lee 2017 [33]	Draxl 2014a [12]	Draxl 2014b [12]
Microphysics	Thompson	Thompson	Thompson Aerosol	Thompson	Thompson
LW Radiation	RRTMG	RRTM	RRTMG	RRTM	RRTM
SW Radiation	Goddard	Dudhia	RRTMG	Dudhia	Dudhia
LSM	Noah	Noah	Noah	Noah	Noah
PBL	MYNN2	MYJ	MYNN2	YSU	MYJ
Cumulus	None	BJM	Kain-Fritsch	Kain-Fritsch	Kain-Fritsch
Surface Layer	Revised MM5	Eta Similarity	Revised MM5	Revised MM5	Eta Similarity

WRF model domain and boundary condition data

The modeling domain, shown in Figure 3.2, consists of three two-way nested domains with horizontal resolutions of 12 km, 4 km, and 1.33 km, respectively. The largest domain, represented by the bounding box of Figure 3.2, is centered over the Northeast United States. The intermediate domain (d02) is outlined in white, and the finest domain (d03), which we use for all analysis, is outlined in red. All three domains have 36 levels in the vertical direction, and the pressure at the model top is 5000 Pa. The eta levels were set manually (Table 3.4 in Section 3.5.1 gives the exact eta levels) to ensure that 6 vertical levels represented approximately the lowest 200 m above ground level. This should better accommodate the WRF WFP, which performs better when a greater number of cells exist lower in the modeling domain [33]. We used a 45 second time step in the coarse domain for all simulations, and for the finest domain (d03), we changed the output data resolution to 10 minutes to match that of the validation data (see Section 3.2.3). Refer to Section 3.5.2 for the remaining constant WRF namelist parameters. ERA5 data provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) was used as boundary condition data for each of the five members [27].

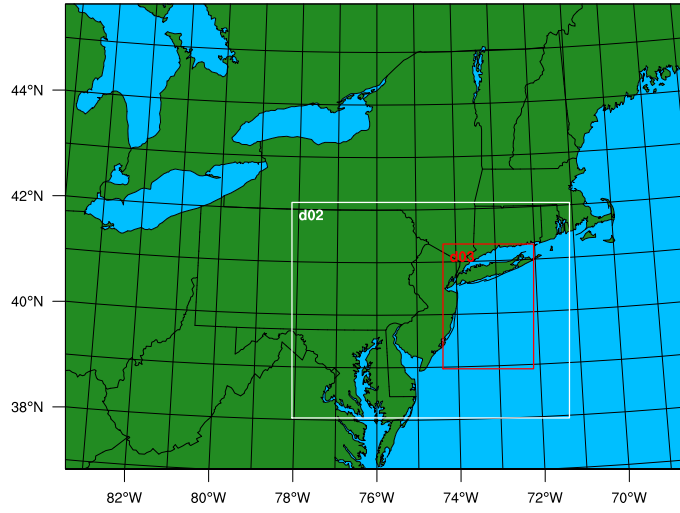


Figure 3.2: WRF domain configuration. The largest domain has a horizontal resolution of 12 km and is shown by the figure’s bounding box covering the Northeastern US. Intermediate domain, d02, has a horizontal resolution of 4 km, and the smallest domain, d03, which we use for analysis, has a horizontal resolution of 1.33 km.

Simulation periods

Given the computational cost of running WRF simulations at 1.33 km resolution, we selected four “representative weeks” in 2020 and ran each ensemble member for a continuous period of 8 days initialized at 00:00 UTC. The first day is used as a spin-up period and omitted from the validation data. We chose a single representative week in each season by borrowing an optimization technique from energy-economy models aimed at integrating intermittent renewables [42]. This is accomplished by minimizing the error between the duration curves for the entire time series and for the representative period. To obtain the duration curve from a given time series we sort the values in the time series in descending order. In this case, we used the 100 m wind speed at one of the buoys for all of 2020 to select a single 7-day representative period in each season. Here, we specified the seasons as January - March, April - June, July -

September, and October - December. We added the additional condition that no more than 10% of the buoy time series could be missing during a representative period, and we used the time series from the south buoy to select the representative period for the first three seasons, and the time series from the north buoy to select the Winter period. This was done because the south buoy had a greater quantity of missing data. However, as most of the data at the south buoy was missing from mid-October through December 2020 we used the north buoy data for the winter instead. The representative weeks are shown in Table 3.2.

Table 3.2: Representative Weeks for Simulations (all 2020)

Winter	Spring	Summer	Fall
Feb 5 - 11	Jun 3 - 9	Jul 1 - 7	Nov 26 - Dec 2

3.2.3 WRF Ensemble Validation

For this study, we bisect ensemble validation by first assessing the performance of each ensemble member separately before looking at the performance of the five-member ensemble as a whole. Both use the same observational data source for validation – from the two LiDAR buoys.

Individual ensemble member performance

We use several common verification statistics to judge how well each ensemble member performed based on data from both LiDAR buoys. We calculate the mean bias in horizontal wind speed at each height for each member by first determining the absolute error for each variable in each 10-minute interval and then calculating the mean over each representative period. We also compute

the root mean square error (RMSE) for each variable. Finally, we include the wind profile error (WPE), introduced by Draxl et al. [12], which we compute by averaging the RMSE errors at each height available from the buoy data between 40 m and 200 m. In each case, the WRF data was interpolated to match the buoy heights. To visually communicate the relative performance of each ensemble member with respect to the observations, we use a subset of the wind speeds to create a Taylor Diagram. The statistics displayed on a Taylor diagram allow for a quick visualization of how much of the RMSE can be attributed to poor correlation and difference in variance between the model and the observations [47].

Performance based on atmospheric stability

Previous work by Draxl et al. shows that different WRF PBL parameterizations perform better under different atmospheric stability regimes [12]. Furthermore, they found that no single PBL parameterization performs the best under all atmospheric stability regimes. As such, we determine the atmospheric stability regime using the Obukhov Length, L , which compares mechanical effects to buoyancy [39]. In other words, in a neutral atmosphere where buoyancy is absent, L is infinite. Equation 3.1 defines the Obukhov length

$$L = -\frac{u_*^3}{\kappa \frac{g}{T_0} \frac{H_0}{\rho c_p}}, \quad (3.1)$$

where u_* is the friction velocity, $\kappa = 0.4$ is the von Kármán constant, g is the acceleration due to gravity, T_0 is the temperature near the surface, H_0 is the surface heat flux (upward defined as positive), and ρ and c_p are the density and specific heat of air, respectively.

Since no direct measurements of the surface heat flux are available, we estimate the value of L using an iterative technique first presented by Van Vijk et al. [48] but also used in more recent studies [37, 2]. Different studies often adopt different thresholds for L to denote the stability regimes. Here we adopt the thresholds used by Archer et al., which define $-100 < L < -5$ as very unstable, $-500 < L < -100$ as unstable, $|L| > 500$ as neutral, $100 < L < 500$ as stable, and $5 < L < 100$ as very stable [2].

Overall ensemble performance

Energy system planners could benefit from ensemble forecasts. In fact, probabilistic wind forecasts are already working their way into power systems operations within some regions of the US [7]. Therefore, we also validate the column of wind speeds for our WRF ensemble using verification rank histograms. Rank histograms have long been a tool for assessing the quality of ensemble forecasts and diagnosing errors in the mean and spread of an ensemble [24]. The rank histogram seeks to verify the probabilities provided by an ensemble for a given variable. Each of the n ensemble members gives one estimate of this variable, and if these estimates are pooled into a vector and sorted in ascending order, the observation should have an equal chance of occupying each rank between 1 and $n + 1$. A rank of 1 represents the case where all the estimates from all the ensemble members are higher than the observation. Continuing this process, for each point in a sample – e.g., for each time-step in the WRF simulation – should then result in a uniform histogram created with the rank values. In other words, the observations occupy each rank the same number of times in an ideal ensemble. Biases show up in rank histograms as overpopulation of the highest or lowest

ranks of the histogram, and dispersion errors produce concave or convex histograms. We create a rank histogram at each buoy for horizontal wind speeds at each vertical level.

3.3 Results and Discussion

3.3.1 Ensemble Member Validation

To our knowledge, this is the first WRF validation study that has had access to data from multiple offshore buoys, which allows us to assess how biases in horizontal wind speeds vary not only across time but across space as well. Biases of the five WRF ensemble members are shown in Figure 3.3. From these plots, we note that a slight positive bias in horizontal wind speed exists for all of the members at nearly every level, and biases tend to be greatest in the lower to middle levels. However, no member overestimates mean wind speeds by more than 0.95 and 0.77 m s^{-1} at the south and north buoy, respectively. The Draxl 2014a, Veron 2018, and Optis 2021 members have the lowest biases overall, but Draxl 2014a and Veron 2018 underestimate wind speeds above 140 m whereas Optis 2021 overestimates wind speeds at the first several levels. Since we observe a similar overestimation from Lee 2017, we partially attribute this to the MYNN2 PBL scheme used by both of these two members.

These mean biases over the full simulation period cannot tell the full story. When biases are plotted by season, as shown in Figure 3.4 for the north buoy, more trends in the biases appear. Overall, the ensemble performs best during the Winter, with a maximum absolute bias of 0.42 m s^{-1} . All members show a

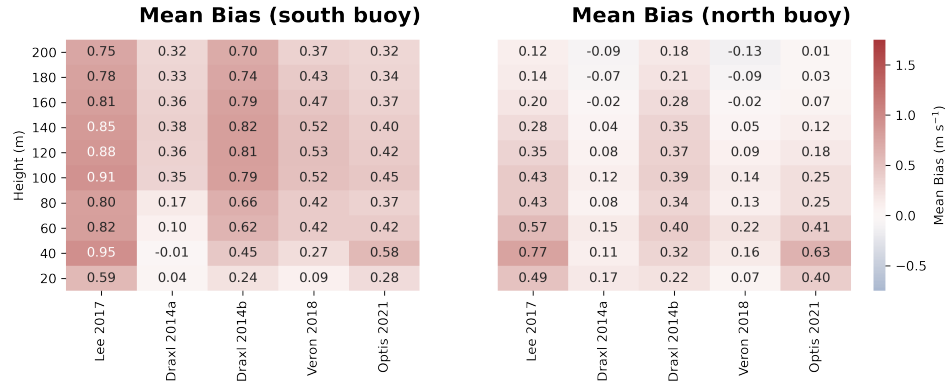


Figure 3.3: Mean bias of WRF ensemble members predicting wind speed (m s^{-1}) averaged over all four seasons at the south (left) and north (right) buoys.

negative bias above 120 m, and Draxl 2014a, Draxl 2014b, and Veron 2018 all have negative biases for the entire column. We attribute this to a prevalence of higher wind speeds during the winter period, which WRF tends to underpredict. Of course, this means that any wind energy analyses informed by WRF-simulated winds, in turn, will underpredict wind power by a power of three greater than the wind speed underprediction. In the case where an underprediction of wind speeds fails to capture winds that exceed the cut-out speed of wind turbines, available wind power may be drastically overestimated.

Most members performed worse during the Summer and much worse during the Spring simulation periods, with positive biases reaching 1.72 m s^{-1} for Lee 2017. Still, during those seasons each member showed similar bias trends (i.e., each member generally overpredicted wind speeds). Overpredicting wind speeds during the Summer poses a particular challenge to operators grappling with peak electricity demand. Therefore, positive biases in wind power prediction would exacerbate already expensive electricity costs on high electricity demand days. Biases during the Fall period proved the most unique varying in both magnitude and direction across ensemble members. The Veron 2018

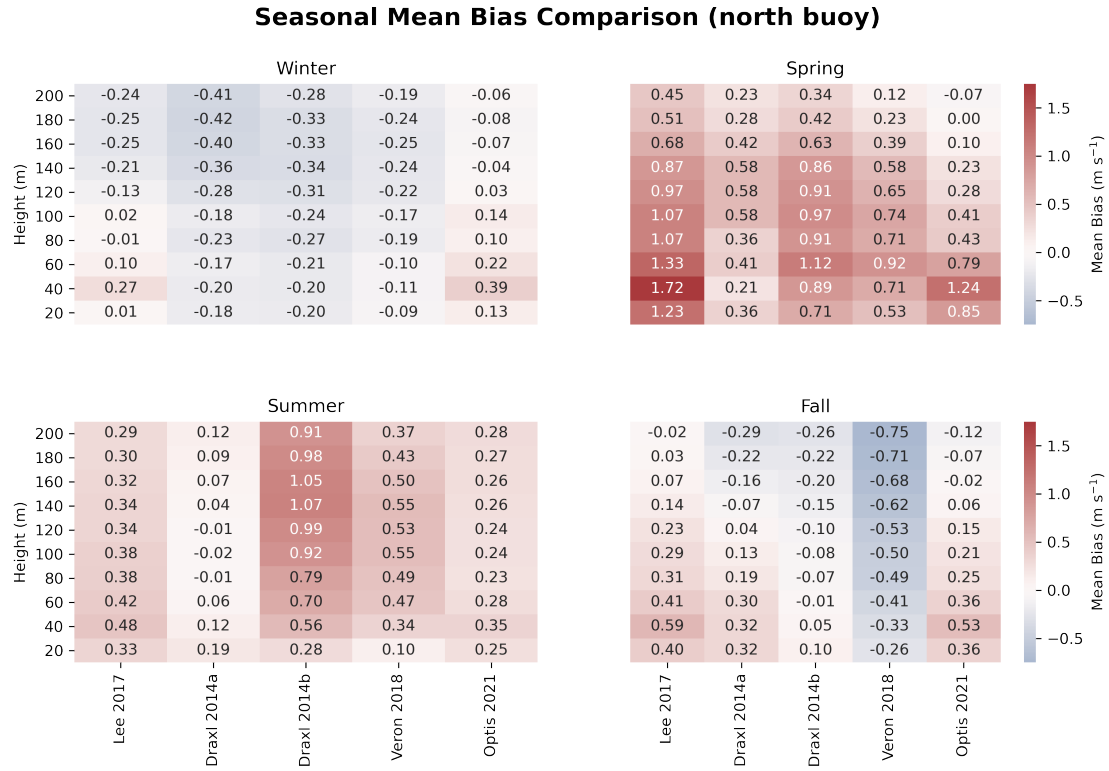


Figure 3.4: Mean bias of WRF ensemble members predicting wind speed (m s^{-1}) averaged over each representative period in each of the four seasons at the north buoy.

member deviated substantially from the other members underpredicting wind speeds – sometimes significantly – at all levels. We can directly attribute this deviation to the BJM cumulus scheme as this is the only difference between the Veron 2018 and Draxl 2014b member, which produced much smaller biases during the Fall period. These results show that researchers should proceed with caution when selecting a WRF setup for a wind energy analysis if the season in which the setup was validated does not match that of their application. Since the south buoy was missing data for the fall representative period, we present the seasonal breakdown in biases in Section 3.5.3 in Figure 3.10. Mostly, similar bias trends up the column exist at the south buoy for the other three seasons, but the magnitudes of the differences are larger reiterating the importance of

localized offshore wind measurements for model validation exercises.

The mean RMSEs (shown in Figure 3.5) tell a similar story to that of the biases varying across all members at all levels from 2.04 to 3.47 m s^{-1} and from 1.90 to 2.77 m s^{-1} at the south and north buoys, respectively. Unlike the mean biases, however, the greatest RMSE tended to occur at higher levels (above 100 m) – likely due to higher winds at these levels. Looking at each individual season (shown in Figures 3.12 and 3.11 in Section 3.5.4), we observe again the highest RMSE values higher in the atmosphere during the Spring and Summer months.

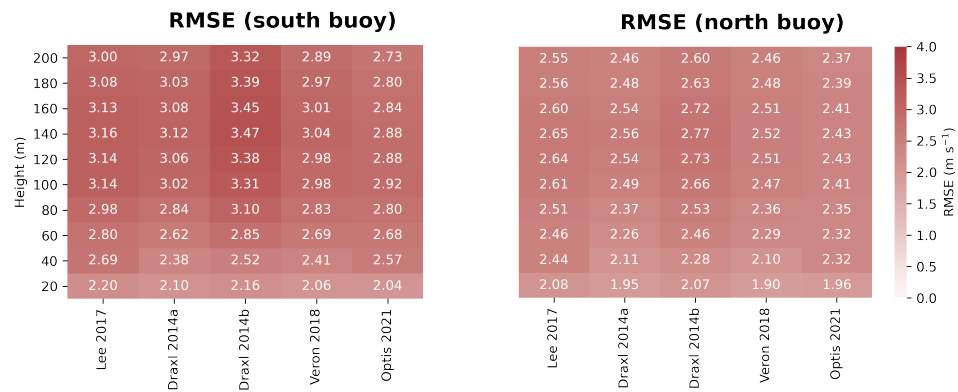


Figure 3.5: Root mean squared error (RMSE) of WRF ensemble members predicting wind speed (m s^{-1}) averaged over all four seasons at the south (left) and north (right) buoys.

The WPE provides a single performance metric that summarizes errors spanning a wind turbine’s rotor area. The Optis 2021 member showed the lowest WPE at both the south and north buoys (see the bolded values in Table 3.3). Similar to the bias and RMSE values, higher WPEs were calculated for the south buoy.

Taylor diagrams offer a convenient depiction of model performance with respect to observations. Specifically, they communicate the standard deviation,

Table 3.3: Mean Wind Profile Error Values (m s^{-1})

	Optis 2021	Veron 2018	Lee 2017	Draxl 2014a	Draxl 2014b
South buoy	2.78	2.85	3.00	2.87	3.15
North buoy	2.38	2.39	2.56	2.41	2.59

correlation with observations, and the skill associated with a set of modeled data. We plot the horizontal wind speeds at 20 m, 100 m, and 200 m for each ensemble member on Taylor diagrams using different colored markers at the south and north buoys (shown in Figure 3.6). Distance from the origin indicates the standard deviation of the variable with a dotted line marking a standard deviation equal to one – matching that of the observations. The zenith angle corresponds to how well correlated a variable is with observations; a zenith angle of 90° indicates a perfect correlation. Finally, isopleths indicate the skill score.

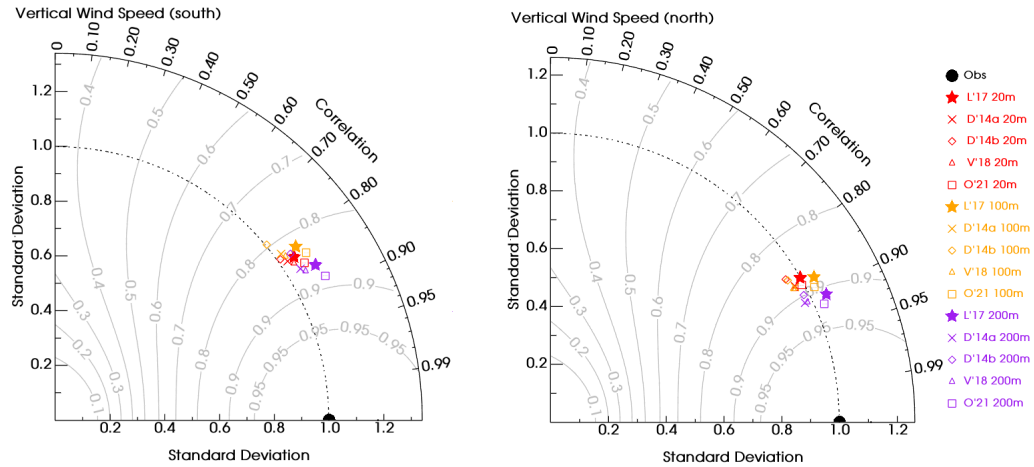


Figure 3.6: Taylor diagrams depicting the standard deviations, correlation coefficient, and skill scores for horizontal wind speeds at 20 m, 100 m, and 200 m for each ensemble member. The left diagram shows data at the south buoy, and the right diagram shows data at the north buoy.

Looking at the Taylor diagrams, we can immediately make several major ob-

servations. First, the Optis 2021 member has the highest skill across each height at both buoys, and the Draxl 2014b member has the lowest. Also, all members have higher skill in downscaling 200 m wind speeds than 100 m and 20 m wind speeds, but the standard deviation in 20 m wind speeds is captured better than those at 100 m or 200 m due to lower wind speeds. Finally, the clustering patterns of variables at both the south and north buoys are similar, but the standard deviations are higher and the correlations are lower at the south buoy leading to lower skill scores. This offers additional evidence to our previous point that the precise location of the offshore wind validation data matters by showing that model skill at two locations within the same offshore region deviate from one another significantly.

3.3.2 Effect of Atmospheric Stability on Wind Speed Profiles

Previous work by Draxl et al. showed that WRF performance varies substantially based on atmospheric stability conditions [12]. We plotted the log wind speed profiles at the south (Figure 3.7) and north (Figure 3.8) buoys. Note that, unlike the bias and RMSE presented previously, we did not interpolate WRF winds to the heights available at the buoys, but we simply plotted the wind speed profiles from the WRF ensemble up to 200 m. At the south buoy, all the ensemble members capture the vertical wind speed profile accurately in very stable, stable, and neutral conditions, although they show a slight tendency to overpredict the profile, which is consistent with our earlier discussion on bias and RMSE. Lee 2017 and Optis 2021 have the most difficulty reproducing a neutral profile because they overpredict wind speeds at the second vertical level (near 30 m). All the members have a more difficult time reproducing the wind

speed profile in stable and very stable atmospheres. For stable atmospheres, none of the members capture the bending back of the wind speed profile observed at the southern buoy. In very stable conditions, WRF fails to predict both the magnitude and the shape of the wind speed profile. We partially attribute this to the fact that very stable conditions are often characterized by lower wind speeds, which WRF tends to overpredict. As other studies have remarked, researchers should characterize atmospheric stability as it clearly affects WRF model performance.

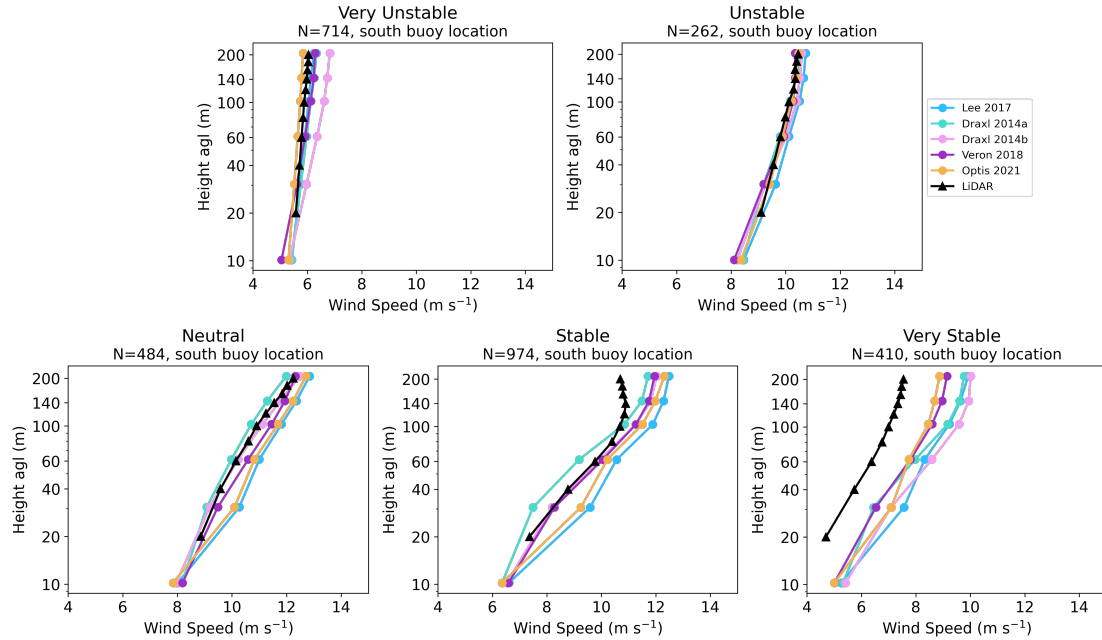


Figure 3.7: Log wind speed profiles separated by stability class at the south buoy for both the observations and each WRF ensemble member. The number of time steps that fall into each stability class is listed in the subtitle of each plot.

Interestingly, we see different patterns at the north buoy, which are shown in Figure 3.8. All the members underpredict the wind speeds in unstable conditions, and they also fail to capture the elbow at 100 m that shifts the profile toward higher wind speeds. Also, all members capture the shape of the stable profile better at the north buoy although the Lee and Optis 2021 profiles show

the same overprediction at the second vertical level.

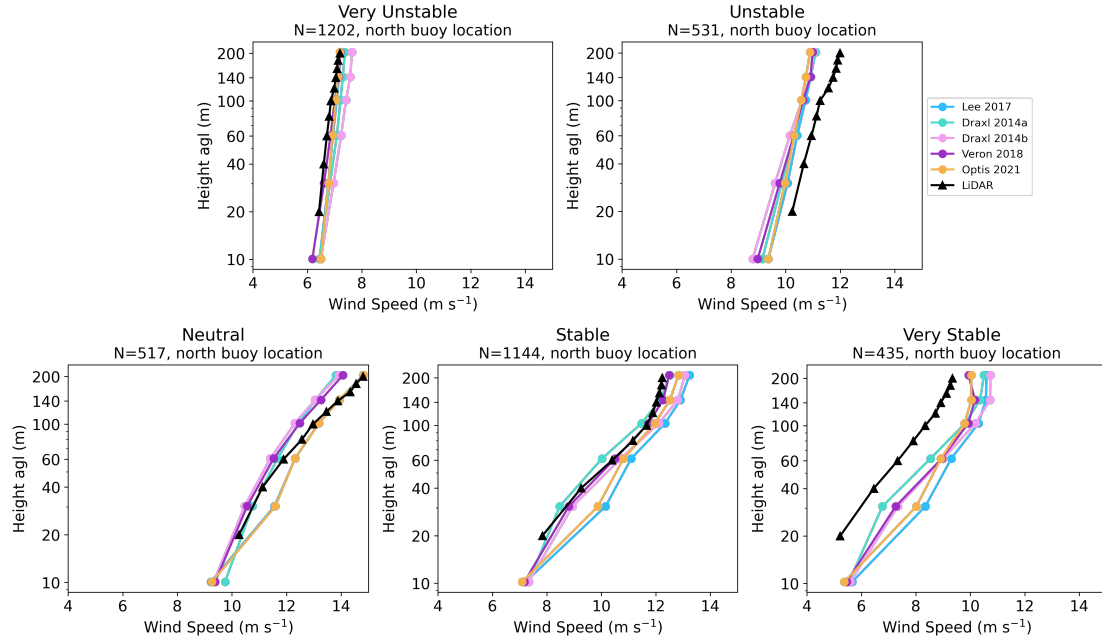


Figure 3.8: Log wind speed profiles separated by stability class at the north buoy for both the observations and each WRF ensemble member. The number of time steps that fall into each stability class is listed in the subtitle of each plot.

3.3.3 Performance of the WRF Ensemble for Offshore Wind Studies

As is clear from the validation exercises we performed on each ensemble member, no single member presents itself as the best setup for predicting wind speeds at all heights in all atmospheric stability conditions. As such, the full multiphysics ensemble could be used to produce a more complete picture of offshore winds across all heights and conditions. We created rank histograms to assess the bias and spread of the ensemble as a whole. A rank histogram is plotted for 100 m wind speeds at both locations (shown in Figure 3.9). For wind speeds at the south buoy shown on the right, the ensemble appears to be quite

underdispersed and negatively biased as well – the observation falls into the first rank about 55% of the time. In a perfect ensemble, we would expect that an observation falls into each rank about 16.7% of the time, which corresponds to the dotted line in the histograms. This means that all the ensemble members produce wind speed downscalings higher than those observed at the south buoy with a frequency of 55%. Such an ensemble could cause problems if used by a utility for a long-term planning exercise. Looking at the rank histogram for the north buoy, shown on the left in Figure 3.9, we observe a lower probability that the observation falls into the first rank – only about 35%. This rank histogram suggests that the ensemble is underdispersed at the north buoy because ranks one and six are both high rather than skewed substantially toward higher or lower ranks. Therefore, using the ensemble mean to carry out a wind resource assessment at the surrounding the north buoy location is more appropriate and should offer an improvement over using a single ensemble member to carry out such an analysis. Of course, without additional measurement data, we have no way of knowing how far this behavior extends around the buoy.

Numerous methods exist for the postprocessing of ensemble forecasts that can correct errors in the ensemble mean and/or spread [49]. Two of the most common techniques are ensemble model output statistics (EMOS) [20] and Bayesian model averaging (BMA) [43]. Either of these can provide probabilistic forecasts across the entire simulation domain using a combination of previous model output and observations as training data. Naturally, this training data only reflects the bias patterns associated with the locations where observations exist, so either uniform bias patterns or good coverage of observations produces the best forecasts. Here, we found that biases differ substantially between the north and south buoy locations calling into question the accuracy of offshore

wind probabilistic forecasts produced using common ensemble postprocessing techniques. Unfortunately, absent additional observational data, we cannot further assess these post-processing techniques. Therefore, future work should continue to assess offshore wind ensemble forecasting methods as more data become available.

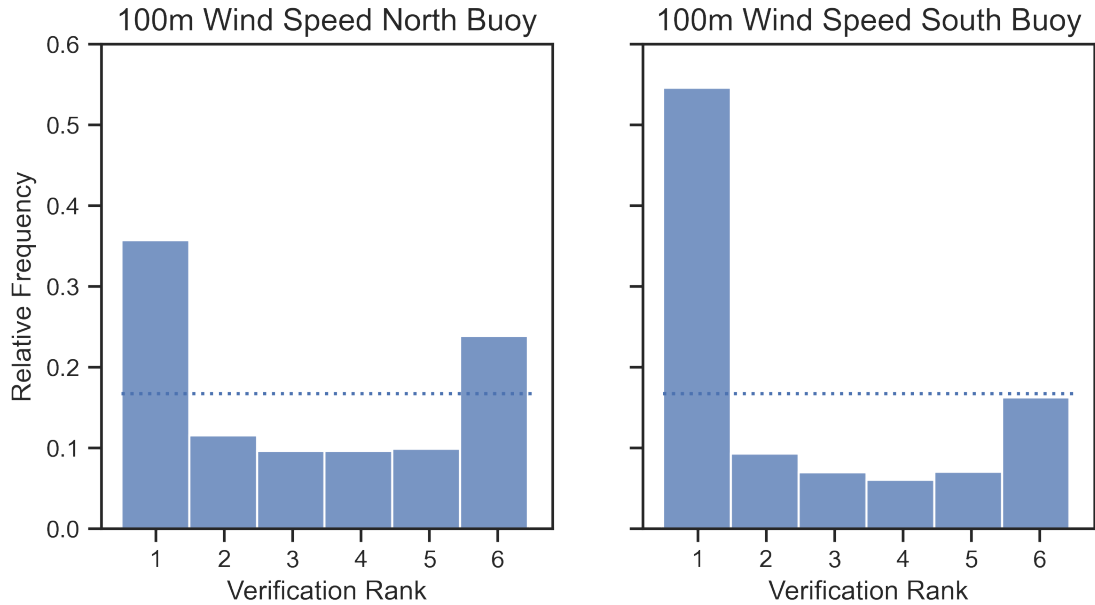


Figure 3.9: Rank histograms for 100 m horizontal wind speed at both the south (right plot) and north (left plot) LiDAR buoy locations. We clearly see evidence of both a negative bias and underdispersion at the south buoy but perhaps only underdispersion at the north buoy.

3.4 Conclusion

In this study, we validated a multiphysics WRF ensemble using data from two floating LiDAR buoys located in the New York Bight. We found that biases in horizontal wind speeds generally increase with height, and the Draxl 2014a, Veron 2018, and Optis 2021 ensemble members have the lowest biases making them the best choices for long-term offshore wind analyses in the Northeast

US. Similarly, the highest RMSE values occur at or above 100 m. However, the member biases vary significantly among different seasons with overpredictions in the Spring and Summer and underpredictions in the Winter and Fall, so researchers should be careful adopting a WRF model setup if the season used for validation does not match that of their application. Furthermore, we note that similar trends are seen at the north and south buoys, but the magnitude of biases and RMSE values is higher at the south buoy. This highlights the importance of collecting offshore wind data at multiple locations even in the same region, which echoes many other offshore wind studies. This would also mitigate the effects that missing data has on validation analysis. That being said, all other studies in this area (e.g., [17, 12, 2]) have only looked at wind speed observations from multiple heights at a single location, so the extent to which WRF wind speed biases vary across the Northeastern US remains a question for further research.

We found that the Optis 2021 setup had the highest model skill score and the Draxl 2014b setup had the lowest across multiple heights and geographic locations meaning that if a single WRF model setup must be used for an offshore wind analysis in the Northeastern US, the Optis 2021 setup is the best choice. However, researchers should note that this setup often overpredicts wind speeds, especially below 80 m above sea level, which could be linked to the MYNN2 PBL scheme. Additionally, other setups (e.g., Draxl 2014b and Veron 2018) perform better than the Optis 2021 setup for stable conditions, which is consistent with previous work by Draxl et al. We found that the Draxl 2014b setup performs well for stable conditions in the Northeast US as well. As such, if the study period is concerned with predominantly stable atmospheric conditions, then this setup is more appropriate.

Finally, we determined that the multiphysics ensemble is underdispersed at both the north and south buoy locations, and additionally, the ensemble is negatively biased at the south buoy. Statistical correction of this multiphysics ensemble is left for future work but is complicated by the lack of observational data and nonuniform bias patterns offshore. Still, some advantage may be gained by using the ensemble mean over a single model setup for the region surrounding the north buoy.

References

- [1] C.L. Archer, H.P. Simão, W. Kempton, W.B. Powell, and M.J. Dvorak. The challenge of integrating offshore wind power in the U.S. electric grid. Part I: Wind forecast error. *Renewable Energy*, 103:346–360, apr 2017.
- [2] Cristina L. Archer, Brian A. Colle, Dana L. Veron, Fabrice Veron, and Matthew J. Sienkiewicz. On the predominance of unstable atmospheric conditions in the marine boundary layer offshore of the U.S. northeastern coast. *Journal of Geophysical Research: Atmospheres*, 121(15):8869–8885, aug 2016.
- [3] Cristina L Archer, Brian A Colle, Luca Delle Monache, Michael J Dvorak, Julie Lundquist, Bruce H Bailey, Philippe Beaucage, Matthew J Churchfield, Anna C Fitch, Branko Kosovic, Patrick J Moriarty, Hugo Simao, J A M Stevens, Dana Veron, and John Zack. Meteorology for Coastal/Offshore Wind Energy in the United States Recommendations and Research Needs for the Next 10 Years. *Bulletin of the American Meteorological Society*, 94:515–519, 2014.

- [4] Robert M. Banta, Yelena L. Pichugina, W. Alan Brewer, Eric P. James, Joseph B. Olson, Stanley G. Benjamin, Jacob R. Carley, Laura Bianco, Irina V. Djalalova, James M. Wilczak, R. Michael Hardesty, Joel Cline, and Melinda C. Marquis. Evaluating and improving NWP forecast models for the future: How the Needs of Offshore Wind Energy Can Point the Way. *Bulletin of the American Meteorological Society*, 99(6):1155–1176, jun 2018.
- [5] Idar Barstad. Offshore validation of a 3 km ERA-Interim downscaling-WRF model’s performance on static stability. *Wind Energy*, 19:515–526, 2016.
- [6] Stanley G. Benjamin, Stephen S. Weygandt, John M. Brown, Ming Hu, Curtis R. Alexander, Tatiana G. Smirnova, Joseph B. Olson, Eric P. James, David C. Dowell, Georg A. Grell, Haidao Lin, Steven E. Peckham, Tracy Lorraine Smith, William R. Moninger, Jaymes S. Kenyon, and Geoffrey S. Manikin. A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. *Monthly Weather Review*, 144(4):1669–1694, apr 2016.
- [7] Ricardo J Bessa, Corinna Möhrle, Vanessa Fundel, Malte Siefert, Jethro Browell, Sebastian Haglund El Gaidi, Bri-Mathias Hodge, Umit Cali, and George Kariniotakis. Towards Improved Understanding of the Applicability of Uncertainty Forecasts in the Electric Power Industry. *Energies*, 10(9):1402–undefined, 2017.
- [8] D Carvalho, A Rocha, M G Omez-Gesteira, and Silva Santos. Offshore winds and wind energy production estimates derived from ASCAT, OSCAT, numerical weather prediction models and buoys e A comparative study for the Iberian Peninsula Atlantic coast. *Renewable Energy*, 102:433–444, 2017.

- [9] Ariel E Cohen, Steven M Cavallo, Michael C Coniglio, and Harold E Brooks. A Review of Planetary Boundary Layer Parameterization Schemes and Their Sensitivity in Simulating Southeastern U.S. Cold Season Severe Weather Environments. *Weather and Forecasting*, 30(3):591–612., 2015.
- [10] Adam J. Deppe, William A. Gallus, and Eugene S. Takle. A WRF ensemble for improved wind speed forecasts at turbine height. *Weather and Forecasting*, 28(1):212–228, 2013.
- [11] Caroline Draxl, Andrew Clifton, Bri-Mathias Hodge, and Jim Mccaa. The Wind Integration National Dataset (WIND) Toolkit. *Applied Energy*, 151:355–366, 2015.
- [12] Caroline Draxl, Andrea N. Hahmann, Alfredo Pena, and Gregor Giebel. Evaluating winds and vertical wind shear from Weather Research and Forecasting model forecasts using seven planetary boundary layer schemes. *Wind Energy*, 17:39–55, 2014.
- [13] EIA. Preliminary Monthly Electric Generator Inventory. Technical report, U.S. Energy Information Administration (EIA), Washington D.C., 2021.
- [14] Sergio Fernández-González, María Luisa Martín, Eduardo García-Ortega, Andrés Merino, Jesús Lorenzana, José Luis Sánchez, Francisco Valero, and Javier Sanz Rodrigo. Sensitivity analysis of the WRF model: Wind-resource assessment for complex terrain. *Journal of Applied Meteorology and Climatology*, 57(3):733–753, mar 2018.
- [15] Anna C Fitch, Joseph B Olson, Julie K Lundquist, Jimy Dudhia, Alok K Gupta, John Michalakes, and Idar Barstad. Local and Mesoscale Impacts

- of Wind Farms as Parameterized in a Mesoscale NWP Model. *Monthly Weather Review*, 140:3017–3038, 2012.
- [16] M García-Díez, J Fernández, D San-Martín, S Herrera, and J M Gutiérrez. Assessing and Improving the Local Added Value of WRF for Wind Downscaling. *Journal of Applied Meteorology and Climatology*, 54:1556–1568, 2015.
 - [17] Evangelia-Maria Giannakopoulou and Regis Nhili. WRF Model Methodology for Offshore Wind Energy Applications. *Advances in Meteorology*, 2014.
 - [18] DNV GL. Independent analysis and reporting of ZX Lidars performance verification executed by Zephir Ltd. at Pershore test site, including IEC compliant validation analysis. Technical report, DNV GL, Bærum, 2019.
 - [19] DNV GL. Independent analysis and reporting of ZX Lidars performance verification executed by ZX Lidars at the UK Remote Sensing Test Site. Technical report, DNV GL, Bærum, 2019.
 - [20] Tilmann Gneiting, Adrian E. Raftery, Anton H. Westveld, and Tom Goldman. Calibrated probabilistic forecasting using ensemble model output statistics and minimum crps estimation. *Monthly Weather Review*, 133:1098–1118, 5 2005.
 - [21] Jay Prakash Goit, Atsushi Yamaguchi, and Takeshi Ishihara. Measurement and prediction of wind fields at an offshore site by scanning doppler LiDAR and WRF. *Atmosphere*, 11(5), 2020.
 - [22] Sven Erik Gryning and Rogier Floors. Investigating predictability of offshore winds using a mesoscale model driven by forecast and reanalysis data. *Meteorologische Zeitschrift*, 29(2):117–130, 2020.

- [23] Andrea N. Hahmann, Claire L. Vincent, Alfredo Peña, Julia Lange, and Charlotte B. Hasager. Wind climate estimation using WRF model output: Method and model sensitivities over the sea. *International Journal of Climatology*, 35(12):3422–3439, oct 2015.
- [24] T. M. Hamill. Interpretation of rank histograms for verifying ensemble forecasts. *Monthly Weather Review*, 129(3):550–560, mar 2001.
- [25] Charlotte B Hasager, Andrea N Hahmann, Tobias Ahsbahs, Ioanna Karagali, Tija Sile, Merete Badger, and Jakob Mann. Europe’s offshore winds assessed with synthetic aperture radar, ASCAT and WRF. *Wind Energ. Sci*, 5:375–390, 2020.
- [26] Charlotte Bay Hasager, Poul Astrup, Rong Zhu, Rui Chang, Merete Badger, and Andrea N. Hahmann. Quarter-century offshore winds from SSM/I and WRF in the North Sea and South China Sea. *Remote Sensing*, 8(9), 2016.
- [27] Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, Adrian Simmons, Cornel Soci, Saleh Abdalla, Xavier Abellan, Gianpaolo Balsamo, Peter Bechtold, Gionata Biavati, Jean Bidlot, Massimo Bonavita, Giovanna De Chiara, Per Dahlgren, Dick Dee, Michail Diamantakis, Rossana Dragani, Johannes Flemming, Richard Forbes, Manuel Fuentes, Alan Geer, Leo Haimberger, Sean Healy, Robin J. Hogan, Elías Hólm, Marta Janisková, Sarah Keeley, Patrick Laloyaux, Philippe Lopez, Cristina Lupu, Gabor Radnoti, Patricia de Rosnay, Iryna Rozum, Freja Vamborg, Sebastien Villaume, and Jean Noël Thépaut. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, jul 2020.

- [28] IRENA. Renewable capacity statistics 2021. Technical report, International Renewable Energy Agency (IRENA), Abu Dhabi, 2021.
- [29] Pedro A. Jiménez, Jorge Navarro, Ana M. Palomares, and Jimmy Dudhia. Mesoscale modeling of offshore wind turbine wakes at the wind farm resolving scale: A composite-based analysis with the Weather Research and Forecasting model over Horns Rev. *Wind Energy*, 18(3):559–566, mar 2015.
- [30] Yuka Kikuchi, Masato Fukushima, and Takeshi Ishihara. Assessment of a coastal offshore wind climate by means of mesoscale model simulations considering high-resolution land use and sea surface temperature data sets. *Atmosphere*, 11(4):1–16, 2020.
- [31] Rohan Kumar, Tim Stallard, and Peter K Stansby. Large-scale offshore wind energy installation in northwest India: Assessment of wind resource using Weather Research and Forecasting and levelized cost of energy. *Wind Energy*, 24:174–192, 2020.
- [32] Jared A. Lee, Paula Doubrawa, Lulin Xue, Andrew J. Newman, Caroline Draxl, and George Scott. Wind resource assessment for Alaska’s offshore regions: Validation of a 14-year high-resolution WRF data set. *Energies*, 12(14), 2019.
- [33] Joseph C. Y. Lee and Julie K. Lundquist. Evaluation of the wind farm parameterization in the Weather Research and Forecasting model (version 3.8.1) with meteorological and turbine power data. *Geoscientific Model Development Discussions*, (1993):1–31, 2017.
- [34] Yichao Liu, Daoyi Chen, Sunwei Li, and P W Chan. Discerning the spatial

- variations in offshore wind resources along the coast of China via dynamic downscaling. *Energy*, 160:582–596, 2018.
- [35] Torge Lorenz and Idar Barstad. A dynamical downscaling of ERA-Interim in the North Sea using WRF with a 3 km grid-for wind resource applications. *Wind Energy*, 19(10):1945–1959, oct 2016.
- [36] Cristian Mattar and Dager Borvarán. Offshore wind power simulation by using WRF in the central coast of Chile. *Renewable Energy*, 94:22–31, aug 2016.
- [37] M. Motta, R. J. Barthelmie, and P. Vølund. The influence of non-logarithmic wind speed profiles on potential power output at danish offshore sites. *Wind Energy*, 8(2):219–236, apr 2005.
- [38] Walter Musial, Paul Spitsen, Philipp Beiter, Patrick Duffy, Melinda Marquis, Aubryn Cooperman, Rob Hammond, and Matt Shields. Offshore wind market report: 2021 edition, 2021.
- [39] A. M. Obukhov. Turbulence in an atmosphere with a non-uniform temperature. *Boundary-Layer Meteorology*, 2(1):7–29, mar 1971.
- [40] Mike Optis, Andrew Kumler, Joseph Brodie, and Travis Miles. Quantifying sensitivity in numerical weather prediction-modeled offshore wind speeds through an ensemble modeling approach. *Wind Energy*, 24:957–973, 2021.
- [41] Yelena L. Pichugina, Robert M. Banta, Joseph B. Olson, Jacob R. Carley, Melinda C. Marquis, Alan Brewer, James M. Wilczak, Irina Djalalova, Laura Bianco, Eric P. James, Stanley G. Benjamin, and Joel Cline. Assessment of NWP forecast models in simulating offshore winds through

the lower boundary layer by measurements from a ship-based scanning doppler lidar. *Monthly Weather Review*, 145(10):4277–4301, oct 2017.

- [42] Kris Poncelet, Hanspeter Hoschle, Erik Delarue, Ana Virag, and William Drhaeseleer. Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems. *IEEE Transactions on Power Systems*, 32(3):1936–1948, 2017.
- [43] Adrian E. Raftery, Tilmann Gneiting, Fadoua Balabdaoui, and Michael Polakowski. Using bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review*, 133:1155–1174, 5 2005.
- [44] SEIA. U.S. Solar Market Insight Report. Technical report, Solar Energy Industries Association (SEIA), Washington D.C., 2021.
- [45] William C Skamarock and Joseph B Klemp. A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *Journal of Computational Physics*, 227:3465–3485, 2008.
- [46] Yuko Takeyama, Teruo Ohsawa, | Jun Tanemoto, Susumu Shimada, | Katsutoshi Kozai, and Tetsuya Kogaki. A comparison between Advanced Scatterometer and Weather Research and Forecasting wind speeds for the Japanese offshore wind resource map. *Wind Energy*, 23:1596–1609, 2020.
- [47] Karl E. Taylor. Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research Atmospheres*, 106(D7):7183–7192, apr 2001.
- [48] A J M Van Wijk, A C M Beimaars, A A M Holtslag, and W C Turkenburg. EVALUATION OF STABILITY CORRECTIONS IN WIND SPEED PROFILES OVER THE NORTH SEA. Technical report, 1990.

- [49] Stéphane Vannitsem, Daniel S. Wilks, and Jakob W. Messner, editors. *Statistical Postprocessing of Ensemble Forecasts*. Elsevier, 1st edition, 2019.
- [50] Dana E. Veron, Joseph F. Brodie, Yosef A. Shirazi, and Justin R. Gilchrist. Modeling the electrical grid impact of wind ramp-up forecasting error offshore in the Mid-Atlantic region. *Journal of Renewable and Sustainable Energy*, 10(1):013308, jan 2018.
- [51] Jing Zhao, Zhen-Hai Guo, Zhong-Yue Su, Zhi-Yuan Zhao, Xia Xiao, and Feng Liu. An improved multi-step forecasting model based on WRF ensembles and creative fuzzy systems for wind speed. *Applied Energy*, 162:808–826, 2016.

3.5 Supporting Information

3.5.1 WRF Eta Levels

The custom eta levels used in the model are shown in Table 3.4

Table 3.4: WRF Eta Levels

1.000,	0.9975,	0.995,	0.990,	0.985,	0.980,
0.970,	0.960,	0.950,	0.940,	0.930,	0.920,
0.910,	0.900,	0.880,	0.860,	0.840,	0.820,
0.800,	0.770,	0.740,	0.700,	0.650,	0.600,
0.550,	0.500,	0.450,	0.400,	0.350,	0.300,
0.250,	0.200,	0.150,	0.100,	0.050,	0.000

3.5.2 WRF Namelist Parameters

Constant WRF namelist parameters are provided in Table 3.5. Any parameters not listed in the table were left at their WRFv4.0 default values.

Table 3.5: Constant Namelist Options

interval_seconds	10800
history_interval	60, 60, 10
time_step	30
max_dom	3
e_we	111, 148, 136
e_sn	83, 115, 136
e_vert	36, 36, 36
num_metgrid_levels	38
num_metgrid_soil_levels	4
dx	12000, 4000, 1333.333
dy	12000, 4000, 1333.333
grid_id	1, 2, 3
parent_id	1, 1, 2
i_parent_start	1, 41, 81
j_parent_start	1, 10, 39
parent_grid_ratio	1, 3, 3
parent_time_step_ratio	1, 3, 3
swint_opt	1
radt	15, 5, 1
bldt	0, 0, 0
cudt	0, 0, 0
surface_input_source	1
num_soil_layers	4
num_land_cat	21
damp_opt	0

3.5.3 Bias Comparison at the South Buoy

Seasonal biases at the South Buoy are shown in Figure 3.10 (see Figure 3.4 in the main text for those at the North Buoy). While most members show similar bias trends going up the column, i.e., the largest biases occur between 100 - 180 m, the magnitudes of the biases at the south buoy are substantially higher than those seen at the north buoy.

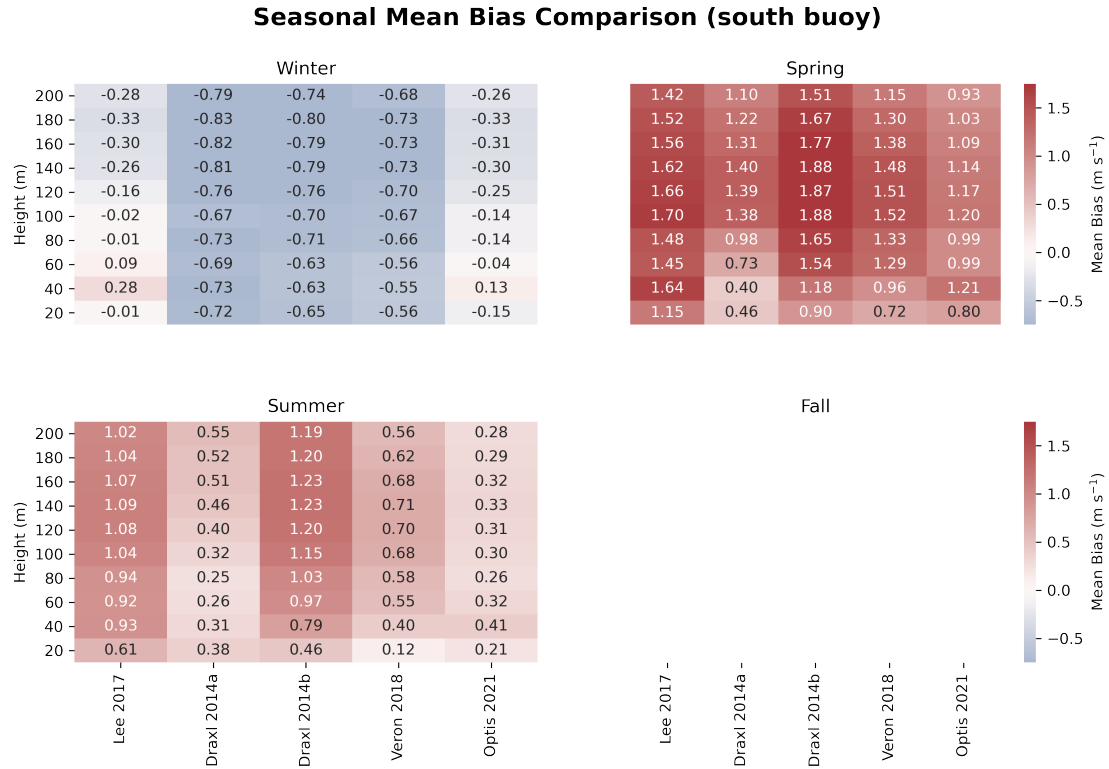


Figure 3.10: Mean bias of WRF ensemble members predicting wind speed (m s^{-1}) averaged over each representative period in each of the four seasons at the south buoy.

3.5.4 RMSE Comparison at Both Buoys

A similar story follows from the seasonal breakdown in RMSE shown in Figure 3.11 and Figure 3.10 for the north and south buoys, respectively. The notable deviation in the trend occurs for the Lee 2017 and Optis 2021 members during the winter where RMSE values at the south buoy are lower than those at the north buoy.

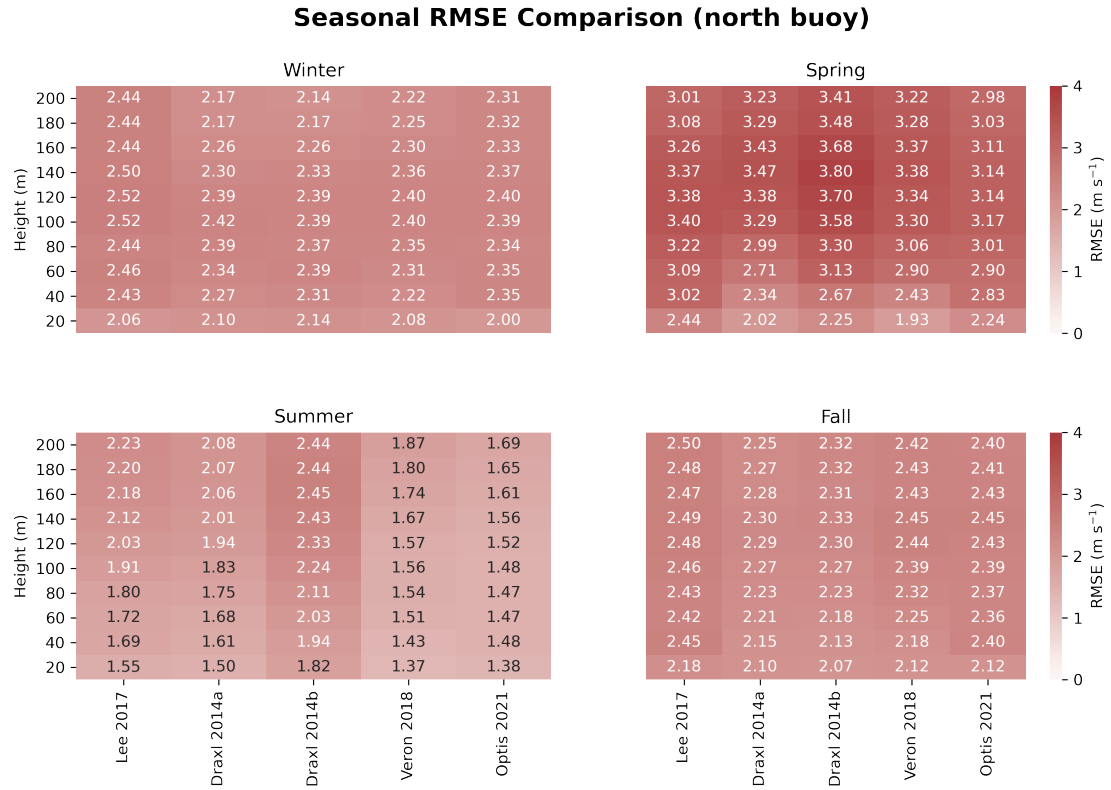


Figure 3.11: Root mean squared error (RMSE) of WRF ensemble members predicting wind speed (m s^{-1}) averaged over each representative period in each of the four seasons at the north buoy.

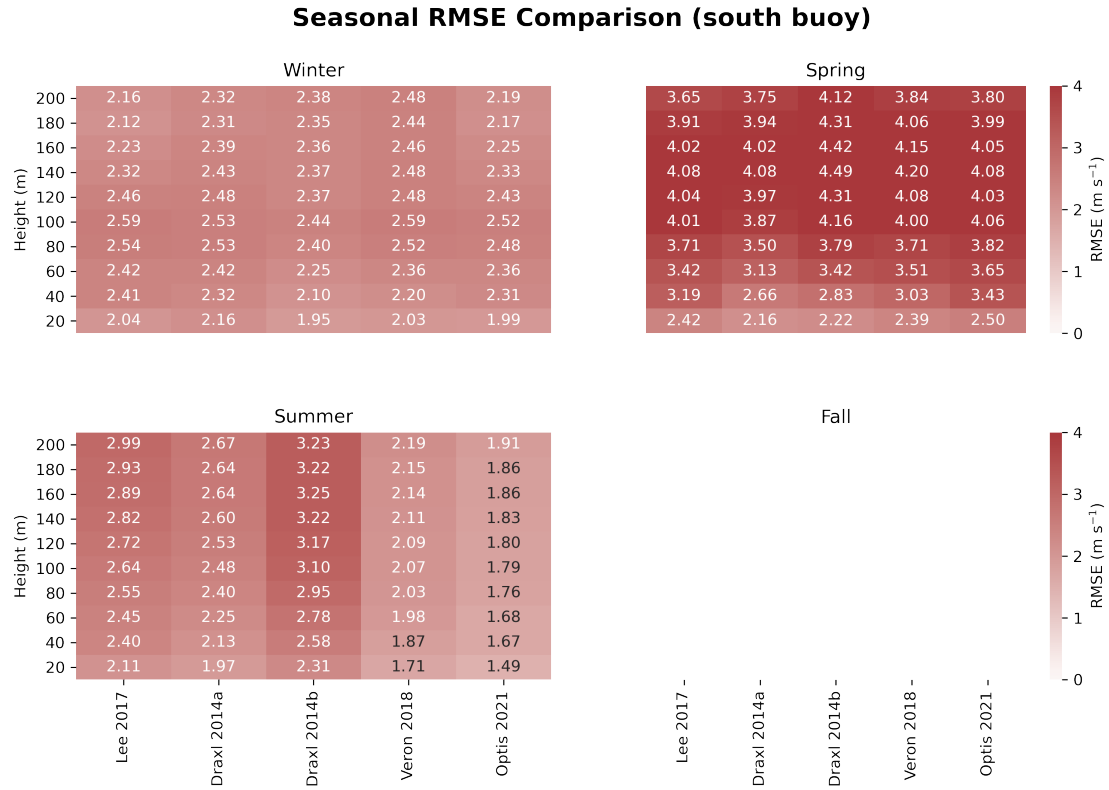


Figure 3.12: Root mean squared error (RMSE) of WRF ensemble members predicting wind speed (m s^{-1}) averaged over each representative period in each of the four seasons at the south buoy.

3.5.5 Potential Errors from Orographic Waves

We initially hypothesized that poor performance at the south buoy could be attributed to orographic waves emanating from the coast of New Jersey. With the south buoy located closer to land, perhaps such forcing from the land affected the south buoy to a greater degree than the north buoy, and if the wavelength of these orographic waves was mispredicted by WRF, this could potentially explain the errors in wind speeds up to 200 m. However, upon inspection of low-level clouds offshore, which we define with a threshold of 850 hPa, we observed no overarching orographic cloud patterns. Figure 3.13 shows these mean low-

level clouds for each simulation week for the Lee 2017 member.

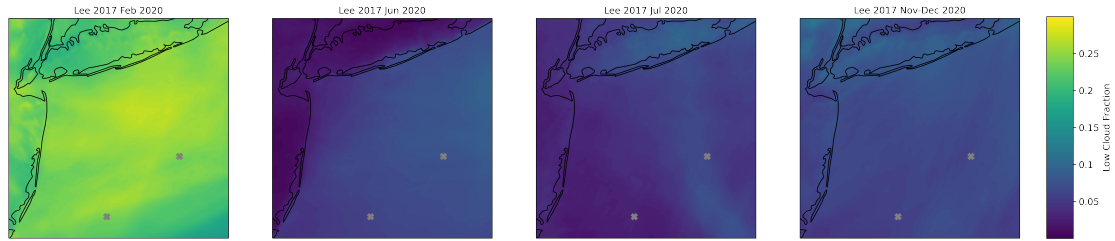


Figure 3.13: Mean low-level cloud fraction for each simulation period for the Lee 2017 member defined by an 850 hPa threshold. A cloud fraction equal to 1 corresponds to full cloudiness at every level below the threshold for every time step in the simulation. Buoy locations are denoted with a grey “x.”

3.5.6 Stability by Wind Speed

Stability class overlaid on a histogram of wind speed distributions are shown in Figure 3.14 and Figure 3.15 for the south and north buoys, respectively. Trends are largely similar, but WRF tends to overpredict very stable conditions in all members, and each member represents the top of the wind speed distribution slightly differently.

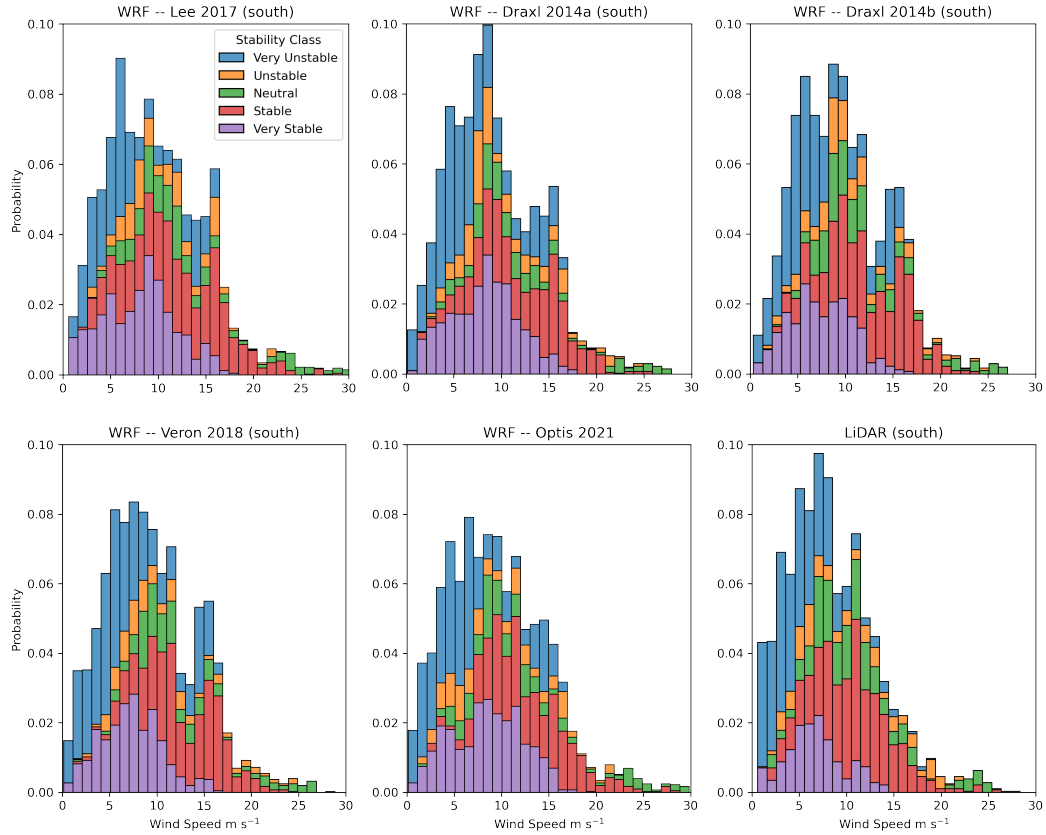


Figure 3.14: Histograms showing stability by wind speed at the south buoy location for each of the five WRF ensemble members and for the LiDAR data (lower right subplot).

3.5.7 Stability by Hour

In this section, stability class is overlaid on the histogram showing the hour of the day at the south (Figure 3.16) and north (Figure 3.17) buoys. Of course, this histogram is uniform unless some data is missing (e.g., the case for the LiDAR buoys). Again, all the WRF members tend to overpredict very stable conditions especially during the night. Also, several members tend to overpredict unstable

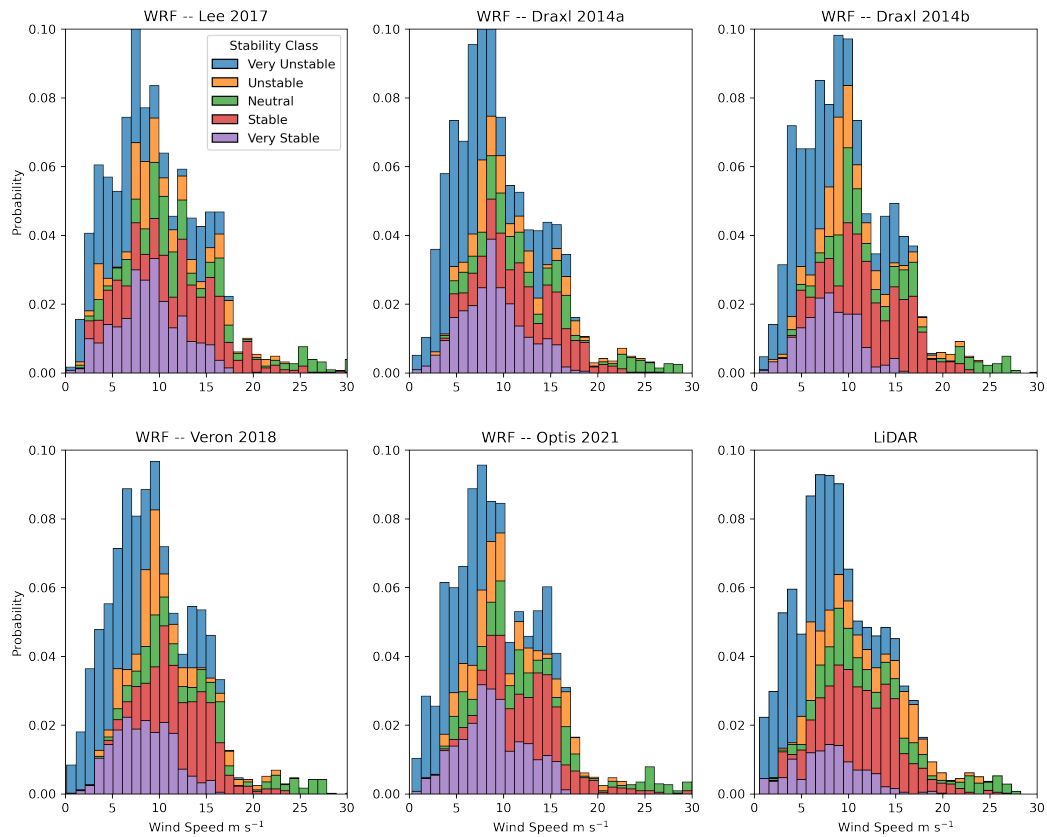


Figure 3.15: Histograms showing stability by wind speed at the north buoy location for each of the five WRF ensemble members and for the LiDAR data (lower right subplot).

conditions near sunset at the south buoy. Unstable and very unstable conditions occur more often at the south buoy, which could be a contributing factor to the increased member biases observed here.

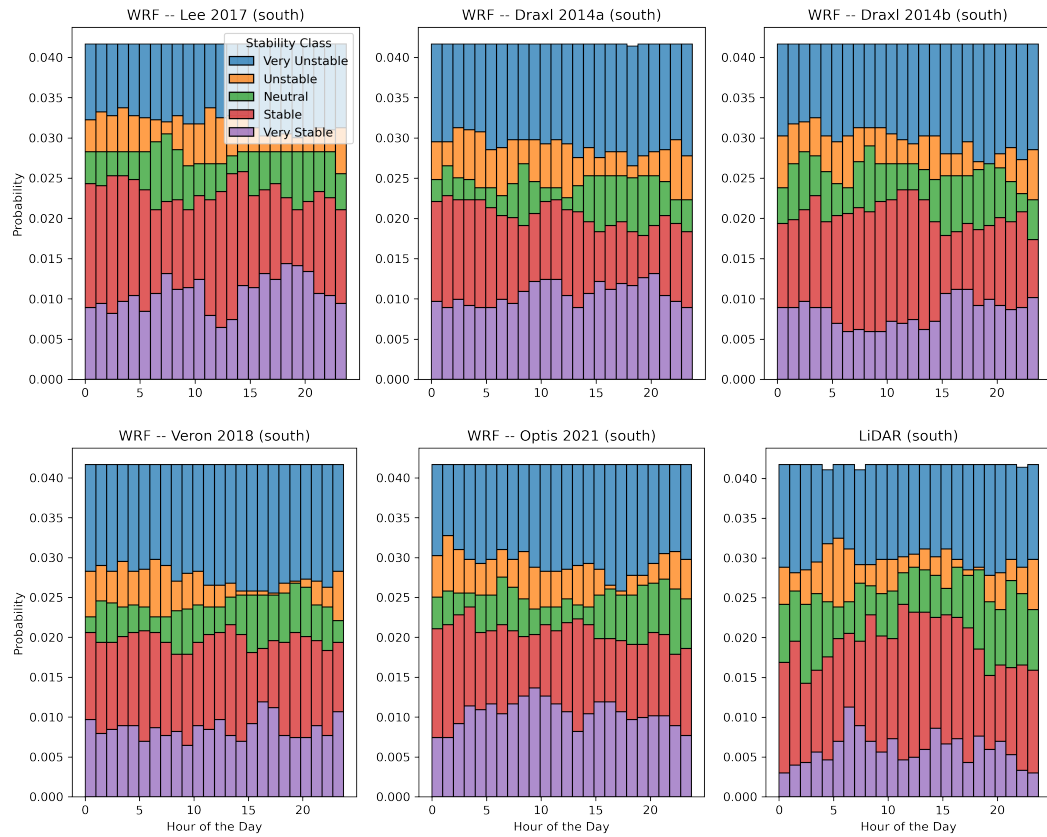


Figure 3.16: Breakdown of stability conditions by hour of the day at the south buoy location for each of the five WRF ensemble members and for the LiDAR data (lower right subplot).



Figure 3.17: Breakdown of stability conditions by hour of the day at the north buoy location for each of the five WRF ensemble members and for the LiDAR data (lower right subplot).

CHAPTER 4

GENETIC ALGORITHM SELECTION OF THE WEATHER RESEARCH AND FORECASTING MODEL PHYSICS TO SUPPORT WIND AND SOLAR ENERGY INTEGRATION

Abstract

To make a future run by renewable energy possible, we must design our power system to seamlessly collect, store, and transport the Earth's naturally occurring flows of energy – namely the sun and the wind. Such a future will require accurate representations of wind and solar resources and their associated variability permeate power systems planning and operational tools. Practically speaking, we must merge weather and power systems modeling. Although many meteorological phenomena that affect wind and solar power production are well-studied in isolation, no coordinated effort has sought to improve medium- and long-term power systems planning using numerical weather prediction (NWP) models. One modern open-source NWP tool – the weather research and forecasting (WRF) model – offers the complexity and flexibility required to integrate weather prediction with a power systems model in any region. However, there are over one million distinct ways to set up WRF. Here, we present a methodology for optimizing the WRF model physics for forecasting wind power density and solar irradiance using a genetic algorithm. The top five setups created by our algorithm outperform all of the recommended setups. Using the simulation results, we train a random forest model to identify which WRF parameters contribute to the lowest forecast errors and produce plots depicting the performance of key physics options to guide energy researchers in quickly setting up

an accurate WRF model.

4.1 Introduction

Propelled by the Paris Climate Agreement in 2016 and the urgent need to reduce greenhouse gas emissions, the world’s capacity to produce electricity from renewable energy grew by 184 GW in 2019 alone. This increase comprises nearly 80% of all growth in electricity production, dwarfing that of fossil fuels [1]. The vast majority of global new additions harvest energy from wind and solar accounting for 118 GW (64%) and 61 GW (33%), respectively. However, electricity generated by these sources is both intermittent and variable, as we cannot control where the wind blows nor when the sun will shine. Accordingly, numerical weather predictions (NWP) are becoming indispensable ingredients in the cocktail of electricity systems operations and planning [43, 42, 35].

One NWP tool in particular—the Weather Research and Forecasting (WRF) model [60]—is the favored tool for wind and solar across scales. For example, WRF underlies operational forecasting systems tuned for both wind [70] and solar [31], provided the forecast system to create NREL’s wind tool kit [11], and helped uncover the costs of uncoordinated wind farm development [39]. Developed under the guidance of the National Center for Atmospheric Research (NCAR), WRF is a mesoscale, non-hydrostatic NWP modeling system that produces forecasts by numerically integrating the dynamical equations of fluid flow – the Euler equations. However, since turbulence exists across many scales in the atmosphere, WRF cannot explicitly resolve all of the dynamics using a grid spacing on the kilometer scale; thus, sub-grid scale processes must

be parameterized – i.e., specified using empirical relations or simplified physical models [53]. Microphysics, radiation, the planetary boundary layer (PBL), the land surface, and cumulus clouds are all parameterized. A number of parameterization schemes for each process have been designed and painstakingly tailored to a broad range of applications. Operationalizing the WRF model from scratch for use in a different region or for a different application is no small task. There are in excess of one million different combinations of the parameterization schemes that dictate how sub-grid scale processes are represented within the WRF model framework. To make matters worse, while some parameterization schemes are compatible with others, many combinations cause WRF to fail.

In more recent releases of WRF, NCAR has included example “physics option sets,” but warns that these are merely meant as starting points for testing the model for a given application (WRF ARW User’s Guide v4.2). In the absence of an existing WRF physics option set, new users are left to scour the literature in hopes of finding works with similar applications or in regions to their own. Rarely do these articles document in detail the lengthy process by which they arrived at their final setup for WRF, and fewer still compare distinctly different physics option sets. This complexity puts operationalizing a WRF model out of reach for many potential users.

Testing all plausible combinations of WRF parameterization schemes remains impossible. Accordingly, studies typically conduct sensitivity analyses with one or more schemes (e.g., [2, 5, 33, 21, 16, 49, 57, 51]). Stergiou et al. even use multi-criteria decision analysis to aid in assessing sensitivity to model physics [62]. To reduce systematic bias from poor parameterization scheme

choices, we propose a comprehensive methodology for operationalizing a WRF model that works in any region for any application of interest.

Recently, researchers have adopted algorithms that mimic evolution to “breed” optimal solutions to computational problems. Collectively, these optimization strategies are called “evolutionary algorithms” (EAs) [68]. A subset of EAs, called “genetic algorithms” (GAs), act as a digital-analog to natural selection by encoding variables as “genes” which combine to form an “individual” [47]. Each individual achieves a fitness score based upon how well it performs with respect to a user-defined objective. Applying this approach, we develop and evaluate a GA for breeding an optimal WRF model to predict wind and solar energy integration in the Northeastern US. The parameterization schemes comprising WRF are analogous to the genetic material comprising DNA; changes to the WRF code affect its macroscale behavior just as changes to the genetic material of an organism will affect its macroscale structure. These genes are mixed and matched between sequential “generations” of models, and the “fittest” models are then hybridized with each other to produce a new generations of models. We will refer to this approach as “OptWRF” for the remainder of this manuscript.

EAs and GAs have aided in WRF parameter tuning before, but usually to optimize parameters within a scheme [28, 8]. Diaz-Isaac et al. used a GA to select a subset of a 45-member multi-physics ensemble based on the flatness of the rank histogram and found that they could improve the representation of model error variances with few ensemble members [10]. In a brief study, Oana and Spataru are the only researchers who have used a GA for initial WRF parameter selection [50]. They report promising early results for humidity and temper-

ature forecasts but note that further study is necessary to assess the utility of using a GA to aid in WRF model setup.

Although many studies design a WRF model for wind or solar resources separately, no studies have sought to optimize – or even assess – WRF’s ability to downscale reanalysis data for the dual purpose of wind and solar. Such a model would immediately halve the computational cost. Among those concerned with wind energy, researchers have introduced and benchmarked a wind farm parameterization [15, 14, 32, 34] and assessed WRF’s sensitivity to different spatial resolutions, different boundary condition data sources, different PBL schemes [20, 12, 52], parameters within PBL schemes, or some combination of these [22, 7, 17, 67]. Others assessed how well a single WRF physics option set performs before conducting a wind resource analysis [44]. Some of these analyses culminated in public meteorological data products to aid in wind power integration (e.g., NREL’s WIND Toolkit [11]).

Many studies also customize WRF for solar energy. WRF-SOLAR provides a WRF physics option set for solar energy forecasting [31], and urban WRF-SOLAR extends this by adding an urban canopy model and building energy model [19]. Two recent studies assessed how well the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS) and the North American Model (NAM) Global Forecast System (GFS) performed for short-term solar forecasting [43, 37] and another evaluated the performance of different shortwave parameterization schemes in predicting GHI [69]. In this work, we explore the feasibility and computational cost of choosing an initial physics option set for combined wind-solar analyses using a GA. Further, we aim to uncover – for the first time – the influence that each physics

category and individual parameterization scheme has upon the final model output. We provide the first database enabling new users to quickly assess which physics option sets work best for wind and solar analyses in the Northeastern US and how they compare to those recommended by NCAR.

The remainder of this chapter is organized as follows: in Section 4.2 will discuss our methodology for developing and applying the genetic algorithm to WRF model setup, in Section 4.4 we discuss our results with an emphasis on the influence of physics options on WRF output and applying this method to operationalize a WRF model in an arbitrary location for an arbitrary application, and finally, Section 4.5 provides a conclusion and final recommendations for applying the OptWRF methodology.

4.2 Method

This method section is divided into two major subsections. We begin with a description of the WRF model and associated setup options in Section 4.2.1, as each model represents an individual in the OptWRF GA. The following explanations in Section 4.2.2 provide specifics of the GA algorithm operators and control parameters that produce better WRF forecasts through successive generations.

4.2.1 The Weather Research and Forecasting Model

WRF produces mesoscale forecasts by numerically integrating the dynamical equations of fluid flow, and parameterizes those processes that cannot be resolved at the user-defined grid spacing [60]. Microphysics, radiation (separate

schemes handle longwave and shortwave radiation), planetary boundary layer (PBL), land surface, and convection (at a horizontal grid spacing above ~4km [60]) are parameterized. Through parameterization, WRF attempts to capture all key couplings that affect the earth system at the mesoscale resulting in an accurate overall forecast or downscaling of wind and solar variables.

Model Parameterizations

During WRF model initialization, choices for parameterizing various unresolved processes must be specified. One scheme is chosen for each of the six major parameterizations – microphysics, long and shortwave radiation, PBL, land surface, and cumulus – that govern how WRF runs. Each unique combination of parameter choices, therefore, constitutes a new physics option set for WRF. Note that the surface layer (i.e., the interface layer between the PBL and the land surface) is parameterized separately in WRF, but as the scheme choice is heavily constrained by the choice of PBL physics, we chose to select the surface layer scheme based upon the PBL scheme. If multiple choices exist, we use the revised MM5 Monin-Obukhov scheme [29]. See Section 4.6.3 for a full discussion about dependencies among different parameterization options and schemes. Parameterization schemes will become the building blocks – the genes – upon which our genetic algorithm is constructed.

Different methods or schemes for parameterization have been developed over the years sometimes capturing underlying physics more accurately and other times offering a speed up when compared with previous parameterization techniques. For this work, we use parameterizations available in WRF version 4.2.1 [59]. Parameters within each scheme can also be tuned to further customize

the model, which increases the number of possible unique WRF “instances” by orders of magnitude. Since previous research has already dealt with optimization of parameters within individual schemes [28, 8], we omit this added complexity here and all internal scheme parameters are left at their default values. In all cases, when one parameterization option is swapped for another, nonlinear coupling among different parts of the model can lead to drastically different model results that often cannot be readily diagnosed.

WRF model domain and boundary condition data

The focus of this work is upon the parameterizations that govern how physical processes within WRF are represented. Therefore, the majority of the WRF namelists options – particularly those governing the model domain and dynamics – are kept constant for each model run as are the boundary conditions. The modeling domain is centered over the Northeast United States and has a 12km horizontal resolution with 36 levels in the vertical direction. The pressure at the model top is 5000 Pa, and the eta levels were set manually (Table 4.2 in Section 4.6.1 gives the exact eta levels). The time step used for all simulations was set to 45 seconds to reduce the likelihood of Courant-Friedrichs-Lewy (CFL) errors as we made every effort to accommodate the broadest array of physics parameterizations. A copy of our constant WRF namelist parameters is included in Section 4.6.2. ERA-Interim data created by the European Centre for Medium-Range Weather Forecasts (ECMWF) was used as boundary condition data for all downscaling [9].

4.2.2 Genetic Algorithm

We have employed a simple genetic algorithm (GA) to identify a WRF physics option set that performs better than out-of-the-box recommendations from NCAR. It's important to note that a GA does not guarantee an optimal solution, but improves the solution space through an iterative evolutionary process. Each WRF physics option set combined with a specific run date represents an individual within this population. GAs are constructed from a number of operators that govern how the algorithm functions: fitness, selection, crossover, and mutation. Figure 4.1 depicts a general layout for the algorithm. These processes are repeated once per generation until a pre-specified number of generations are reached. Each operator will be described in greater detail in the sections that follow.

Physics schemes as genes; WRF simulations as individuals

Like any real population, the space of possible individuals depends indelibly upon the gene pool. Genes represent the basic building block upon which a population is built, and for this case, physics schemes created for each of the six major physical processes parameterized within the model are each encoded as one of six genes. These processes are microphysics with 25 schemes, longwave radiation with 8 schemes, shortwave radiation with 8 schemes, land surface with 6 schemes, PBL with 11 schemes, and cumulus with 14 schemes (See Section 4.6 for a complete list of parameterization schemes). To set up a WRF model run, exactly one parameterization scheme from each category must be chosen. Combinatorially, this allows for 1,478,400 distinct possible WRF model setups. Taking model run date into account, that number swells to 539,616,000. Each

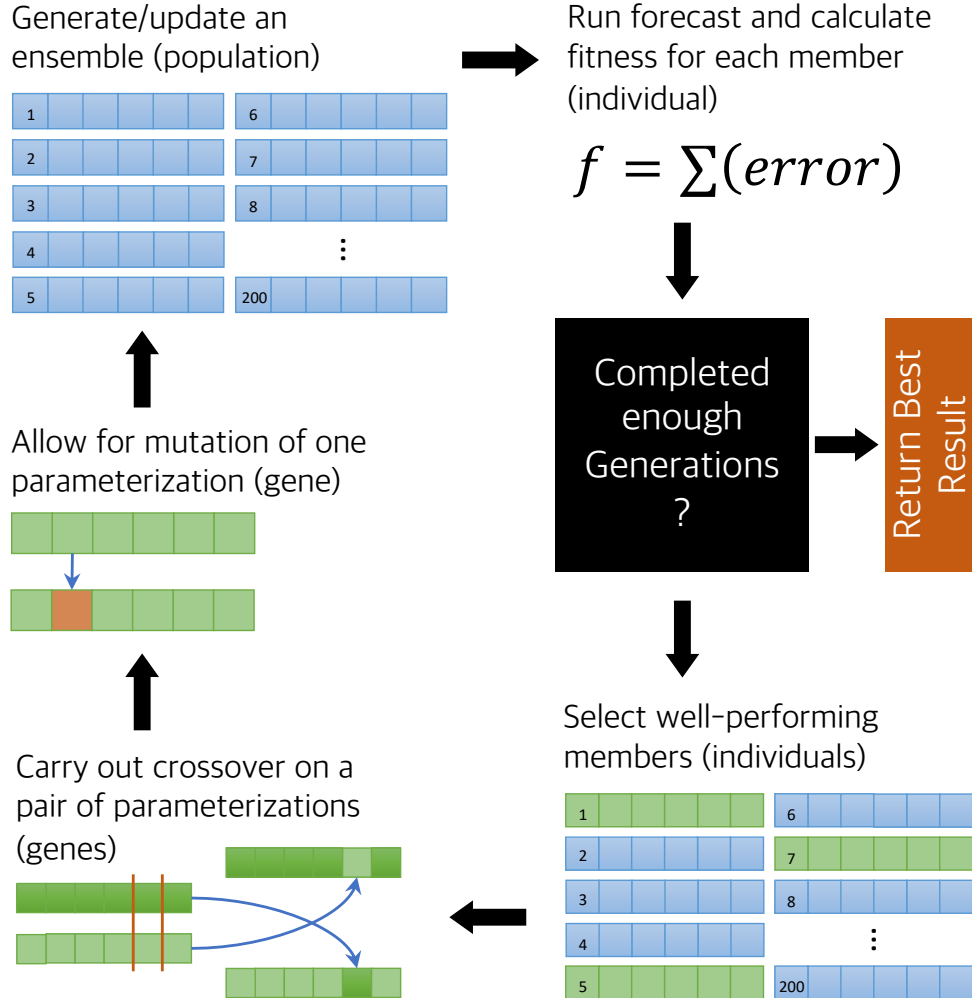


Figure 4.1: Schematic diagram of the genetic algorithm used for determining a near optimal set of physics parameters for the WRF model based upon an arbitrary application-specific fitness function. The number of generations (cycles through the diagram) can be tuned by the user based upon size of the ensemble, available time, and computational resources.

of these distinct model setups corresponds to a single individual that together makes up a population for the genetic algorithm.

The fitness function

In order to judge how well each individual within the population performs, a metric for judgment must be specified. For the application of medium- to long-term power systems planning, we care about WRF's ability to downscale meteorological variables affecting wind and solar energy production. We created a metric dubbed the wind power density (WPD), which varies with the cube of the turbine hub height (100m) wind speed, to judge model performance based on wind speed. Performance based on solar energy was judged by model-calculated GHI. The fitness function shown in Equation 4.1 calculates the accumulated errors between a WRF model realization and the ERA5 reanalysis dataset [25] in WPD and GHI across time (t) and across all model grid cells (g) as shown below. While WPD and GHI both have units of W m^{-2} , a correction factor was employed to ensure that accumulated errors in each quantity during the year 2011 were the same. Reanalysis products blend observations with past short-range forecasts providing the most complete picture of global weather in existence making them an ideal choice for testing the skill of a gridded model.

fitness

$$\begin{aligned} &= \sum_t \sum_g \left(\left(D_{frac} |GHI_{WRF} - GHI_{ERA5}| \right. \right. \\ &\quad \left. \left. + C_{wind} \left| \left(\frac{P_{wind}}{A} \right)_{WRF} - \left(\frac{P_{wind}}{A} \right)_{ERA5} \right| \right) \right) \end{aligned} \quad (4.1)$$

Population initialization and tournament selection

A population within the GA is composed of WRF models with different physics and/or run dates. During population initialization, the namelist files governing how each model will run are written. First, physics schemes for the six major parameterization options are selected randomly from the options available (See Section 4.6 for the complete list). Throughout testing, we uncovered numerous incompatibilities while attempting to run the WRF model a certain way. When possible, these physics scheme incompatibilities were incorporated into the initialization function signaling the operator to avoid these combinations. The run date is selected randomly within the year. This year was selected for convenience, and a comparison of performance for different years was beyond the scope and computational resources available for this work, but the choice to select start dates within a single year was made deliberately to ensure that seasonal variability was taken into account without having to consider meteorological variability on longer time scales. Experiments were run using several population sizes, containing 50 - 200, whose performance will be discussed in Section 4.4.

Following population initialization, the fitness of each member of the population was calculated by running each WRF model. All the models in a generation were submitted to a cluster simultaneously using Python's `concurrent.futures` module. Each model was allocated a maximum of six hours on eight cores for a maximum of 48 core hours per simulation – the vast majority of models finished in that amount of time. Those models that failed to finish within the allotted six hours were terminated and assigned a large fitness value. Likewise, those models that failed due to physics scheme incompatibilities or

other WRF model vicissitudes were assigned a high fitness value. Recall that the fitness value represents the accumulated model error, so lower fitness values are desirable.

With fitness values calculated, the selection operator identifies individuals within the existing population and constructs a mating population from which offspring will be conceived. Tournament selection creates a mating population that has both strength and diversity. In tournament selection, 10% of the existing population is selected randomly, and the individual with the highest fitness in this group is placed in the mating population. Tournament selection is continued iteratively until the mating population is half the size of the existing population. Previous research has investigated the sensitivity of output with respect to many of these GA control parameters [45], but such analysis was beyond the scope of this work. The randomization introduced by tournament selection is designed to keep the GA from converging to a sub-optimal local minimum prematurely by ensuring that the population contains individuals carrying genes with a greater variety of fitness values than just the ones at the top, which we carry through to future generations regardless.

The crossover operator

An offspring population is formed via two separate mechanisms – crossover and elitism. Elitism simply takes a prescribed number of individuals – one-third of the population in this study – and places them in the offspring population unchanged. The crossover operator is responsible for filling the remaining two-thirds of the offspring population. The crossover operator randomly selects two individuals from the mating population and gives them a 50% chance to

form two offspring. When crossover does happen, a single gene (i.e., a physics scheme) is selected at random, and swapped between the two individuals creating two new offspring.

The mutation operator

After the offspring population has been filled, mutation provides a mechanism to introduce additional genes into individuals that their parent models did not contain. The mutation operator is applied to each offspring in the population, and the probability of a mutation occurring is equal to one over the population size. In other words, on average, one member of each offspring population will experience a mutation. When a mutation occurs, a single gene within an individual is randomly changed – one physics scheme is swapped for another in one model. Researchers previously delved into both the best operators to carry out mutation [6] and the optimal mutation probability [23]. After mutation, the final version of the offspring population is complete. The cycle (in Figure 4.1) begins again and this population becomes the parent population for the next generation. Only once the prescribed number of generations have elapsed are the best models extracted from the final population.

4.3 Results

By running OptWRF with a population size of 200 over 10 generations, 2000 simulations were initialized. Of those, 1823 distinct simulations – considering the date of model initialization and physics option set together – ran successfully. The balance of the 2000 total simulations can be accounted for either

via the 8% (16 simulations) that are passed into each successive generation unchanged (a total of 144 simulations) or failed for one of many possible reasons along the way. We also retained some additional simulations run during algorithm testing to bolster the data used in the analysis of WRF model behavior based upon physics option sets (see Section 4.3.3). The top five performing physics option sets are shown in Table 4.1. Note that the best performing physics option set (NSSL 1 moment microphysics, RRTMG longwave and shortwave radiation, Pleim-Xiu land surface model, BouLac PBL, and old Kain-Fritsch cumulus) appears three times in the top seven demonstrating its ability to produce low error values when WRF is initialized on different days. All the top-performing physics option sets found by OptWRF offer an improvement to the WRF User Guide-Recommended physics option sets as shown in Table 4.1.

Table 4.1: Highest Performing OptWRF and WRF User Guide-Recommended Setups. Improvements are reported with respect to the highest performing setup recommended in the WRF user guide, which is shown directly below the midrule. Plus signs indicate a performance improvement, and minus signs indicate a performance deterioration.

Date	Microphysics	LW Rad	SW Rad	LSM	PBL	Cumulus	Fitness	Improvement
OptWRF Setups								
Dec 13	NSSL 1 Moment	RRTMG	RRTMG	Pleim-Xiu	BouLac	Old KF	7858.	+16.0%
Jan 14	NSSL 1 Moment	RRTMG	RRTMG	Pleim-Xiu	BouLac	Old KF	7859.	+16.0%
Dec 13	Thompson	RRTMG	Dudhia	Pleim-Xiu	BouLac	KF	8207.	+12.2%
Dec 02	Sbu-Ylin	RRTMG	RRTMG	Pleim-Xiu	BouLac	Old KF	8480.	+9.3%
Jan 14	NSSL 1 Moment	RRTMG	Dudhia	Pleim-Xiu	BouLac	Old KF	8488.	+9.2%
Dec 02	NSSL 1 Moment	RRTMG	RRTMG	Pleim-Xiu	BouLac	Old KF	8519.	+8.9%
Jan 14	NSSL 1 Moment	RRTMG	CAM	Pleim-Xiu	BouLac	KF	8601.	+8.0%
WRF User Guide-Recommended Setups								
Dec 13	Thompson	RRTMG	RRTMG	Noah	MYJ	Tied TKE	9351.	N/A
Dec 13	WSM6	CAM	CAM	Noah	YSU	KF	16071.	-71.9%
Dec 13	Thompson	RRTMG	RRTMG	Noah	YSU	Grell-Freitas	16117.	-72.4%
Dec 13	WSM5	RRTMG	Goddard	Noah	MYJ	KF	16704.	-78.6%

4.3.1 Annual Comparisons

We designed OptWRF to cover the greatest possible number of physics option sets initialized on random days throughout a single year. Of course, there is an inherent trade-off between coverage and computational cost. Running simulations over longer time periods, or with multiple time periods dispersed throughout the year, would have given us more insight into how skilled a particular physics option set was at downscaling a diverse set of meteorological conditions, but we would have been able to investigate fewer physics option sets with the same computational resource. Since we sought to determine how different parameterization options affected wind and solar forecast errors, higher coverage of physics option sets better matched the goals of this work.

Still, we wanted greater confidence that physics option sets ultimately selected by OptWRF outperformed the WRF User Guide recommendations across a variety of meteorological conditions. We ran the top five physics option sets selected by OptWRF and four from the WRF User Guide (Table 4.1) for the entire year of 2011 in a series of single-day simulations similar to the original experiment. Monthly mean fitness values for each of the five physics option sets found by OptWRF outperformed each of the four from the WRF User Guide (Figure 4.2). Note that several of the simulations initialized in July and August for the NCAR 1 physics option set failed to run, so these monthly means were generated using data from fewer simulations. However, it is unlikely this affected the overall trend presented here.

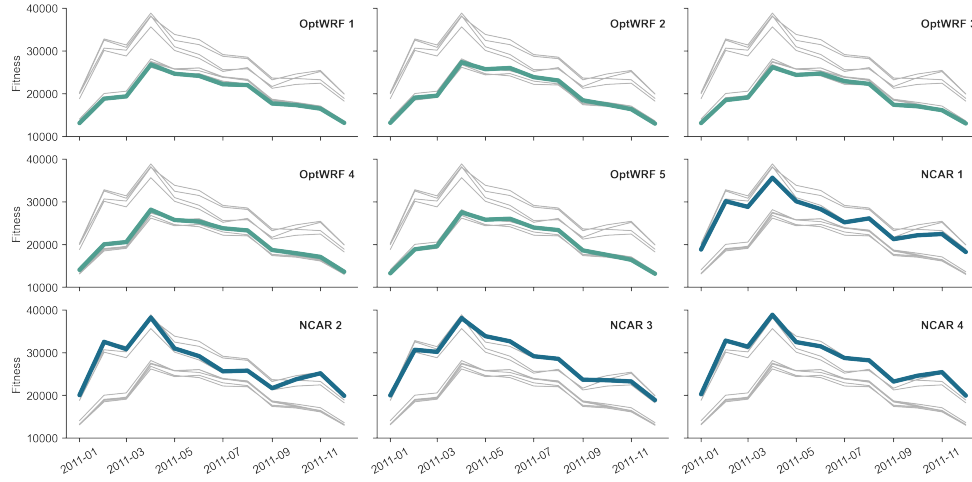


Figure 4.2: Plot matrix with each of nine panels highlighting the monthly mean fitness of one physics option set and the results from the other eight plotted in grey. physics option sets selected by OptWRF are highlighted in teal and appear in the first five panels; whereas physics option sets recommended by NCAR in the WRF User Guide are highlighted in navy blue and occupy the final four panels. Each of the sets chosen by OptWRF outperformed each of the sets recommended by NCAR in all months of 2011 providing compelling evidence that these physics option sets produce better forecasts across seasons and diverse meteorological conditions.

4.3.2 Wind and Solar Trends

Simulations are judged based upon the fitness function (Equation 4.1). For this application, we designed the fitness function to deliberately consider wind and solar resources equally on an annual basis as we would like the forecast skill to be equal for both wind and solar energy. We chose GHI to benchmark solar energy as this depends on both the diffuse and direct components of solar radiation – although both are shortwave [41]. We integrated the GHI for one single-day simulation period and plotted it across the entire domain for both the best-performing WRF simulation and the ERA5 reanalysis (Figure 4.3a-b). This integration was performed using the hourly values for GHI in kW m^{-2} , which when summed over a day, result in $\text{kWh m}^{-2} \text{ day}^{-1}$. Clouds, moisture,

and aerosol creation or advection are largely responsible for the sharp GHI gradients shown in such single-day snapshots.

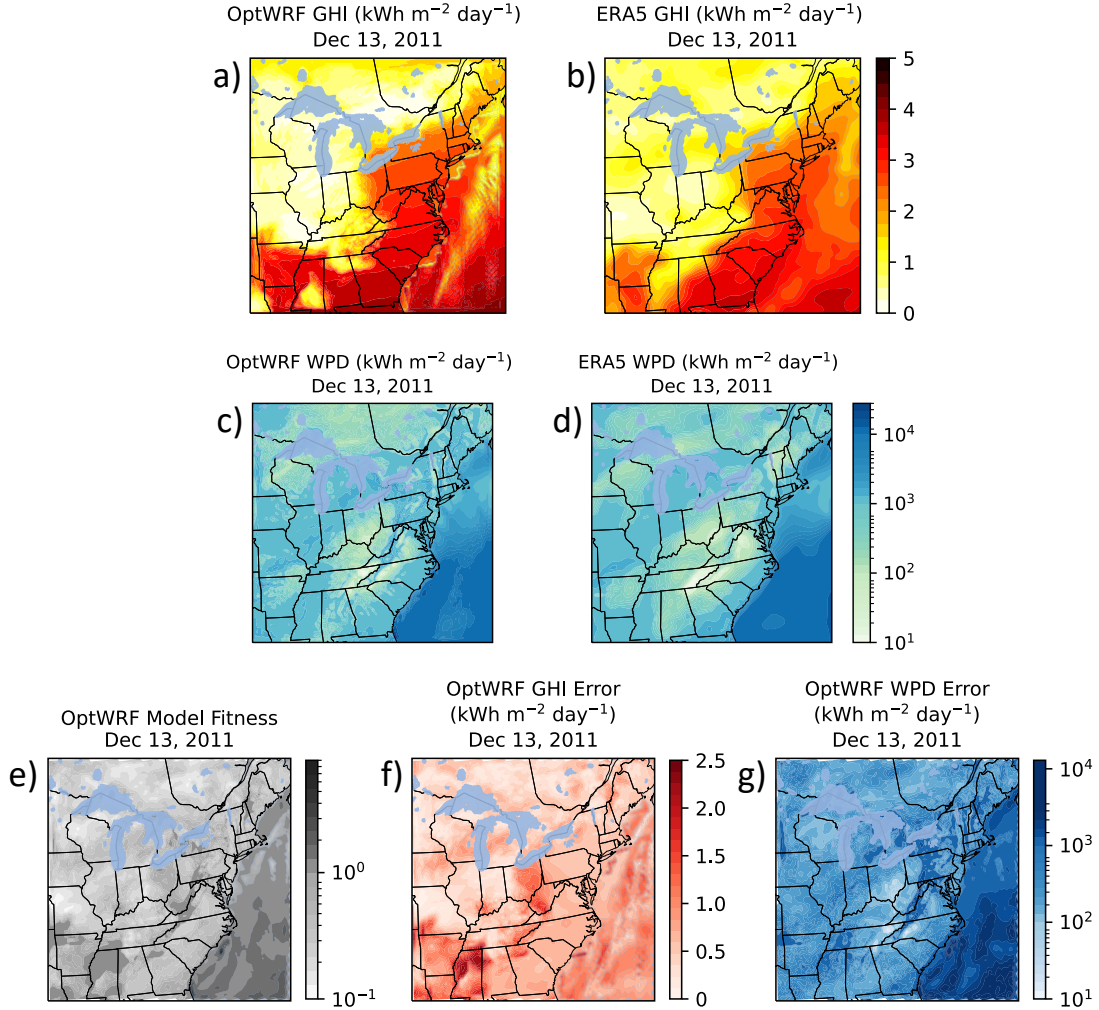


Figure 4.3: Panels a and b show global horizontal irradiance (GHI) across the entire modeling domain in $\text{kWh m}^{-2} \text{ day}^{-1}$ for a one-day WRF simulation initialized on December 13, 2011 00 (UTC) and the ERA5 reanalysis. Similarly, panels c and d show the wind power density (WPD) also in $\text{kWh m}^{-2} \text{ day}^{-1}$. Data shown in b, d, f, and g is taken from the best-performing simulation produced by OptWRF. Errors in the GHI (panel f) and WPD (panel g) contribute to the overall model fitness (panel e).

The same method was applied to produce WPD plots for both the best-performing WRF simulation and the ERA5 reanalysis (Figure 4.3c-d). Since

wind power depends on the cube of the wind speed, the WPD is plotted on a log scale.

Visualizing the fitness is simply a matter of combining the errors incurred by a WRF simulation with respect to the ERA5 reanalysis, controlling for day length in the GHI error, and correcting the WPD error so the two error metrics are of the same order of magnitude. Fitness varies across the domain with the errors in GHI and WPD (Figure 4.3e-g), and errors are the highest where meteorological extremes are poorly captured. A panel of scatter plots showing GHI error, WPD error, and fitness values for each of the almost 2000 simulations is shown in Figure 4.7 in the Supplemental Information.

4.3.3 Effects of Physics Parameterizations

In order to determine the absolute effect that each physics parameterization scheme has upon the fitness value, we made use of a random forest regressor [4]. This method was chosen because the physics parameterization options are non-ordinal categorical variables; so principle component analysis cannot shed much light on the effect that each physics option has upon the model output. Random forest allowed us to one-hot encode each parameterization scheme as a feature so each could be considered separately. Although the random forest regressor assigns a weight (importance) to each feature, these weights are relative, and their shortcomings are well documented [63]. To determine the importance of each parameterization scheme, we employed the SHapley Additive exPlanation (SHAP) Python Package [38]. SHAP provides a high-speed exact algorithm for explaining the output of tree-based machine learning mod-

els and has been shown to correctly rank the contributions of predictor variables to model output [24]. The SHAP values for each radiation and each planetary boundary layer parameterization scheme in each model are shown in box plots in Figure 4.4 and Figure 4.5, respectively. When a scheme is turned on (i.e., used in a WRF forecast), an orange box is used, and when a scheme is turned off (i.e., not used in a WRF forecast), a blue box is used. Negative SHAP values correspond to those schemes – either turned on or off – that push the fitness (error) value lower while positive SHAP values correspond to those physics options that push the fitness (error) value higher. In other words, when an orange box appears to the left of the zero SHAP value line, the scheme improves the WRF forecast skill when used.

Looking at the radiation schemes shown in Figure 4.4, several perform quite poorly while only a couple contribute to lower fitness (error) in a majority of cases. Clearly, shortwave FLG, longwave Held-Suarez, and shortwave New Goddard are linked to much higher errors in GHI and WPD as all non-outlier SHAP values are positive. WRF users investigating wind and solar integration should avoid using these schemes. Unfortunately, no schemes can be linked to an overwhelming reduction in error; so we can offer no clear choice of radiation schemes. The shortwave RRTMG scheme is the only one linked to a modest reduction in forecast errors in nearly all cases. This is consistent with the best physics option sets selected by OptWRF (see Table 4.1). Longwave RRTMG and longwave FLG are both linked to lower errors in the majority of cases and represent the best choices for longwave radiation. All remaining schemes contribute to higher errors in the majority of cases and therefore users should carry out further model tuning exercises if they want to use one of these schemes for wind and solar integration.

Turning now to the performance of the planetary boundary layer schemes shown in Figure 4.5, only a couple contribute to higher error most of the time and several perform well. Only the MYNN3 scheme is linked to higher errors in GHI and WPD in almost all cases, but ACM2, MYNN2, TEMF, and YSU are all linked to higher errors a majority of the time. Therefore, users should avoid these schemes for wind and solar integration if at all possible or carry out further model tuning if the use of these schemes is warranted for some ancillary reason. For example, this raises an issue for the WRF wind farm parameterization [15], which is currently only compatible with the MYNN2 scheme. We will engage this point further in Section 4.4. Encouragingly, the BouLac scheme appears to drive down GHI and WPD errors substantially in all cases that it is used making it an easy choice for a PBL scheme. GBM, MYJ, QNSE, Shin-Hong, and UW schemes are also all linked to lower errors for the majority of models in which they were activated. Section 4.6.7 presents SHAP values plots for the remaining parameterization schemes (microphysics, land surface, and cumulus).

4.4 Discussion

As with any modeling approach, there are caveats that affect the applicability of this approach, which warrant further discussion. Foremost, we want to stress the importance of the choice of domain and fitness function. Then, we discuss why certain physics options sets may have outperformed others.

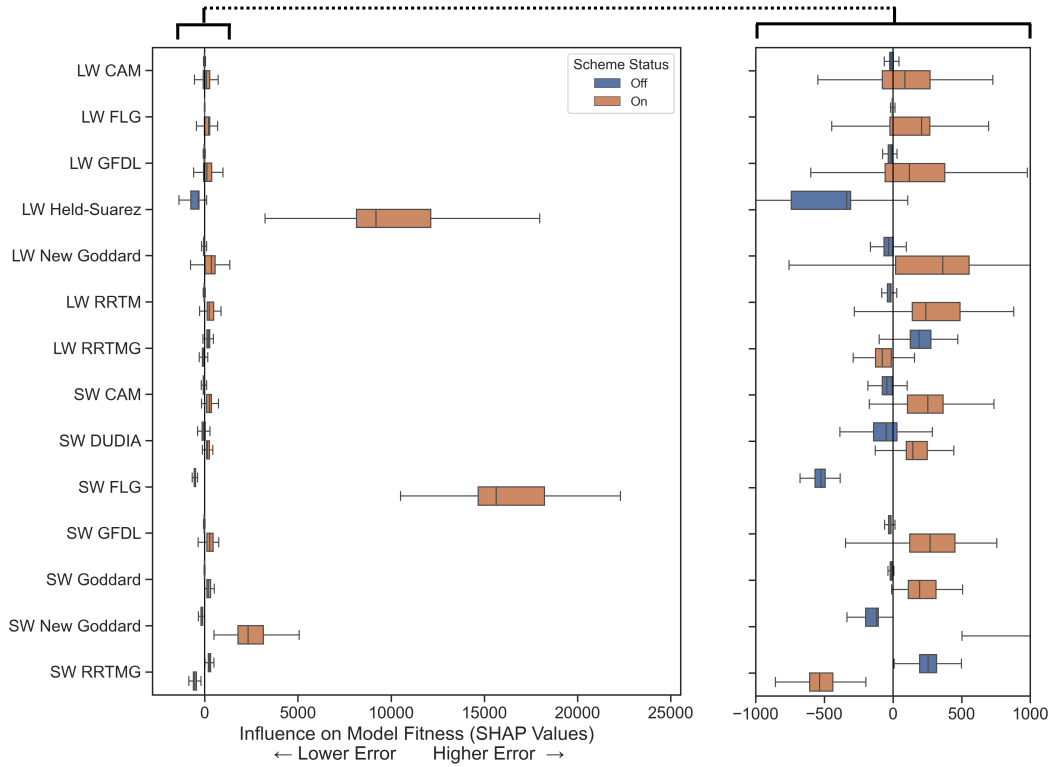


Figure 4.4: SHAP values explain the impact that each random forest model feature (WRF parameterization scheme) has upon model error. Radiation parameterization schemes are listed along the left axis, and boxes are created from all the WRF model runs. Orange bars correspond to WRF models where the parameterization scheme was activated; whereas the blue bars correspond to those where the scheme was deactivated. Negative SHAP values correspond to those schemes that reduce model error (improving the forecast); whereas positive SHAP values correspond to those schemes that increase model error. Therefore, those schemes with orange bars centered furthest to the left produce the best WRF forecasts. The namelist options for all longwave and shortwave radiation schemes are included in Tables 4.5 and 4.6 in Section 4.6.3, respectively.

4.4.1 Domain and Fitness Function

All the results presented here are inextricably linked to how well WRF can predict GHI and WPD across the entire domain covering most of the eastern United States. We chose this domain to overlap with the Ozone Transport Commission's domain and its placement enables researchers studying one of the three

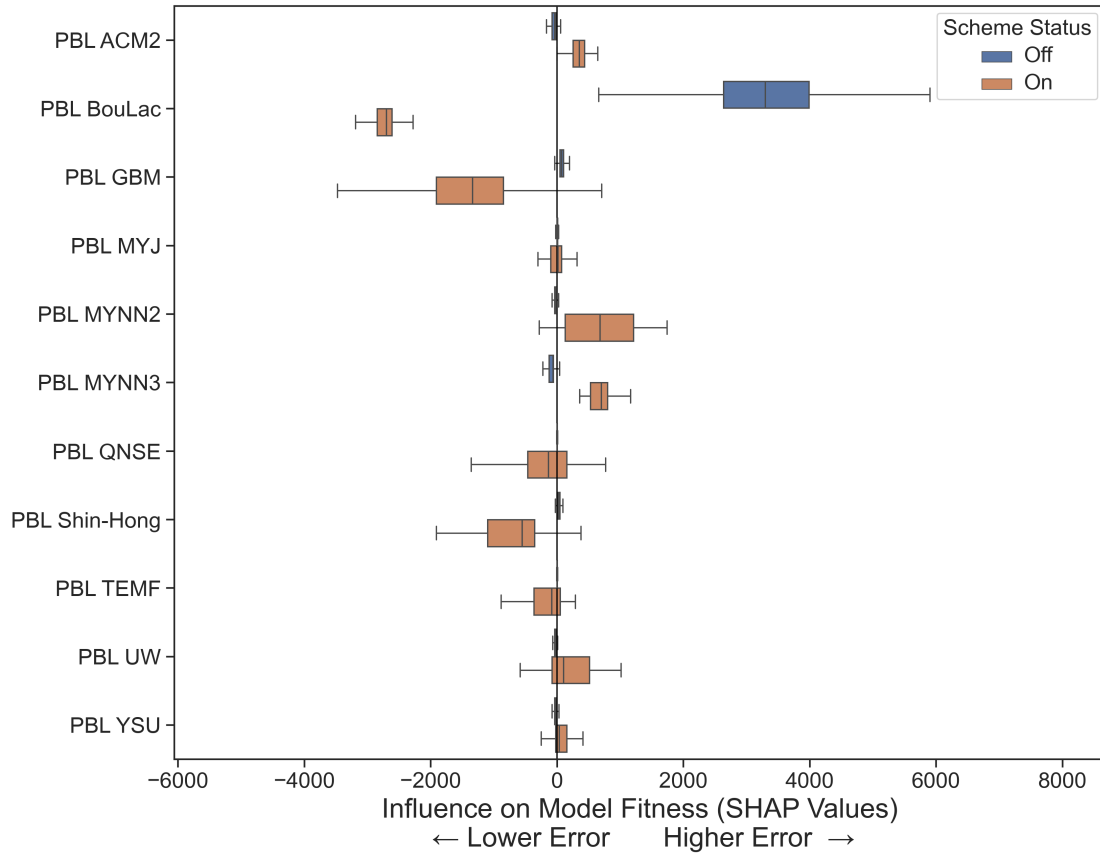


Figure 4.5: Planetary boundary layer parameterization schemes are listed along the left axis, and boxes are created from all the WRF model runs. Orange bars correspond to WRF models where the parameterization scheme was activated; whereas the blue bars correspond to those where the scheme was deactivated. Those schemes with orange bars centered furthest to the left produce the best WRF forecasts. The namelist options for all PBL schemes are included in Table 4.8 in Section 4.6.3.

major eastern regional transmission organizations – PJM Interconnection, New York ISO, and New England ISO – direct insight into the best-performing WRF physics option set across the entire region. However, the fitness function judges all errors across this domain equally; so we caution users who study primarily a small subset within this domain and encourage them to undertake further benchmarking. For example, the best performing model produced by the GA incurred high errors in WPD off the coast of North Carolina (see Figure 4.3g).

As such, this model setup would almost certainly not be the best choice for simulating the power output from an offshore wind farm there.

On a related note, our approach applies to a variety of spatial and temporal scales. For this work, we selected a 12-km domain across the eastern United States and ran many one-day simulations within 2011 to balance computational resources with coverage of many different physics option sets. Users could easily modify this setup in several ways (i) change the location, spatial extent, or grid spacing of the domain, (ii) include a nested domain in which to judge errors, and (iii) extend the time period for each simulation or the time period from which model start dates are drawn. Keep in mind that, barring a reduction in the spatial extent of a domain or an increase in grid spacing, each of these modifications will increase the computational cost of running OptWRF. For some studies, using a single year as we have here is appropriate. For example, short- to medium-term power systems capacity planning exercises generally utilize historical peak demand for a single season or year. On the other hand, long-term capacity planning exercises (e.g., planning wind and solar facility deployments through 2050) should incorporate simulations from a longer time period to capture variability at larger scales (e.g., interannual cycles, ENSO, and interdecadal cycles) that will affect how much energy a wind or solar facility will produce over its lifetime.

4.4.2 Best Physics Option Sets

We turn now to a discussion of why these particular physics option sets (refer back to Table 4.1) may have produced the best downscaling of GHI and WPD.

To frame this discussion, we will zoom in on a portion of our domain, New York State, to discuss the spatial pattern in the GHI errors (shown on the left in Figure 4.6) and WPD errors (shown on the right in Figure 4.6). Errors in GHI appear uniform across much of NYS with the largest errors occurring off the coast, to the South and East, of Long Island. We attribute these to clouds that are resolved in the WRF downscaling but not in the ERA5 reanalysis. However, since these features sit offshore, their presence represents no concern for forecasting solar resources over land thereby potentially making this model more attractive in practice. An extension of this work could consider only the GHI over land. Modestly high GHI error values occur near the Great Lakes and near the White Mountains in Northern New Hampshire. We stipulate that these errors may stem from the added physical complexity associated with the transition from lake to land and the wake of a mountainous region, respectively. In other words, the microphysics scheme likely omits some of this complexity and mispredicts the spatial or temporal extent of clouds in these areas.

Interestingly, we found that the NSSL single-moment scheme (see [40] for the original NSSL two-moment scheme) selected by OptWRF was linked to higher errors in a majority of cases, shown in Figure 4.8. Since the same is true for the other two microphysics schemes used in the best-performing models (Thompson [65] and Sbu-Ylin [36]), the microphysics scheme does not appear to strongly influence GHI and WPD error values. This stands in contrast to the radiation schemes (Refer back to Figure 4.4) where both the longwave and shortwave RRTMG [27] schemes are linked to lower errors. Longwave RRTMG is based on the correlated-k method [48] that generates required k-distributions and optical depths using a line-by-line radiative transfer model drawing data from the HITRAN database [58]. We found that this new “look-up-table” pa-

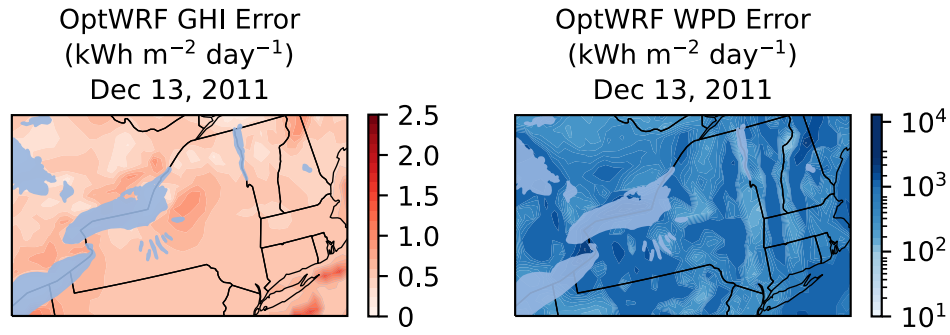


Figure 4.6: The left panel shows a zoomed-in view depicting GHI error over New York State and parts of the Northeastern United States – the same quantity as plotted in Figure 4.3f. The largest errors in GHI occurred off the southern and eastern coasts of Long Island, NY. The right panel shows a zoomed-in view depicting WPD error over New York State and parts of the Northeastern United States – the same quantity as plotted in Figure 4.3g. The spit of land between Lake Erie and Lake Ontario as well as the White Mountains in Northern New Hampshire have the largest errors in WPD. However, the area off the coast of Long Island, NY also has moderately high WPD errors.

parameterization method [53] outperforms the longwave schemes based upon the older broadband emissivity method [61] (e.g., see [13]).

Turning to errors in WPD shown in the right panel of Figure 4.6, we observe some expected trends. The highest errors occur between Lake Erie and Lake Ontario and near the White Mountains in Northern New Hampshire. Since the spit of land between the two Great Lakes is approximately 36km wide, it can be spanned by three WRF grid cells (12km) and is on the same order as the ERA5 grid (~30km). Therefore, it's not surprising that winds deviate here as the benchmark data represents this feature so coarsely. For the White Mountain region, winds likely deviate due to the complex terrain. Most PBL parameterizations – with the possible exception of the YSU scheme [26] – poorly represent unresolved orographical features existing in mountainous terrain and remains an open area of research [29, 30, 22]. Draxl et al. reported that the choice of PBL schemes depends on atmospheric stability [12]. Unfortunately, the BouLac

scheme [3] selected by OptWRF was not part of their study. Therefore, the performance of this scheme under different stability regimes remains a subject for future work. Of course, the PBL scheme exchanges heat and moisture fluxes with the surface layer scheme but these fluxes are ultimately determined by the physics represented in the land surface model. Therefore, we want to highlight that the Pleim-Xiu land surface model [66], which was developed in conjunction with the ACM2 PBL scheme [55, 56] and the Pleim-Xiu surface layer scheme [54], was selected in isolation of its companion schemes by OptWRF. Clearly, future work should continue to explore hidden synergies among WRF’s many parameterization schemes.

4.5 Conclusion

We developed a methodology that utilizes a genetic algorithm to aid in the setup and benchmarking of a numerical weather prediction model (WRF) with hundreds of thousands of potential physics option sets. This method can be applied over any region, and the fitness function can be tailored to a specific application (or variable) to find a more optimal model physics option set. The benefits of the method are that it allows users to a) find a model setup in the absence of an inherited model or similar application existing within the literature, b) easily discover sensitivities to different physics option sets and model run dates that would not otherwise be apparent, and c) provides users with multiple options – allowing them to choose a quicker sub-optimal model. Of course, running a GA with WRF models as the individual computational unit is inherently computationally expensive; so not all researchers will have access to computational resources necessary to carry out such an expensive exercise. We have therefore

stored all the models run throughout the course of this work in an SQL database and will be included with the Supplemental Information.

The dataset produced by this work will aid future researchers in medium to long-term power systems planning. The sheer number of meteorological data products creates a high barrier to entry for the atmospheric sciences. This creates a problem for energy system planners who historically conducted long-term planning exercises with simple models to predict electricity demand growth and help them to decide where to invest in new power plants and lines that would support the system several decades into the future. The proliferation of renewable energy demands a much more complex approach to system planning integrated intimately with meteorological modeling. Weather now not only loosely dictates how much electricity people use but is responsible for exactly how much renewable energy will be supplied. Since many of the parameterization options were developed before the rapid deployment of renewable energy, there is no reason to believe that the prediction of renewable energy production was even a passing consideration for the designers of these physics schemes. Our results provide an exhaustive view of how well each physics scheme operates when judged on wind and solar energy prediction – power systems researchers can use these directly to quickly set up a WRF modeling effort.

Open-source software tools and publicly accessible datasets are vital to both the meteorological modeling community and the energy systems community. We have put all the code used in this work – namely an OptWRF python package – on a public GitHub site in the hopes that others can make use of this methodology producing and disseminating datasets in different geographic re-

gions or for different applications. The meteorological modeling community is one containing a wealth of freely available data modeling products; the WRF model is a prime example. The power systems community has some work to do, but a similar approach will help more regions across the world develop renewable energy and decarbonize their electricity systems faster.

To be sure, there are some areas for improvement in tailoring a GA to effectively set up a WRF modeling effort. GA control parameters are important tools to help guide the algorithm toward the best solution, and while we adopted the population size and elite percentage from previous work [46], other recommendations were omitted to reduce complexity. For example, Mills et al. found that an adaptive mutation rate could improve GA efficiency. The domain, physics, and dynamics of the WRF model itself are also exceedingly complex. Future work could investigate GA solution sensitivity to horizontal and vertical grid spacing specified in each WRF model – an incredibly computationally expensive task as it would require repeating the work described here at a variety of different grid resolutions. We also have not investigated the sensitivity of the GA solution to the geographic location where the WRF models are run. We expect that different parameterization options will perform better in different geographic areas, but there is little way to know how well the model setups reported in Section 4.3 may apply to an arbitrary location across the globe. The best physics option sets reported here are probably most relevant to climates, latitudes, and topographical regimes similar to that of the Northeastern United States, but the OptWRF approach could be deployed anywhere (e.g., in the tropics where the relative contributions of solar radiative heating and the Coriolis force are markedly different). Finally, some work has been done to optimize the control parameters within an individual parameterization option to improve the

forecast skill (e.g., [28, 8]). This work could be integrated into the current framework as a second step after the model physics has been selected during the first. Each of these future directions would require staggering computational resources; so we echo again the importance of making such WRF performance analytics available to the broader user community who may lack necessary resources.

Not all environmental nonprofits, local governments, community colleges, and the like have the knowledge or budget to painstakingly tailor a forecasting system to fit their needs over a period of years. However, as developers of wind and solar float ever-cheaper bids, planners across institutions require tools to make sense of an increasingly interconnected system. As the weather will dictate both electricity supply and demand, WRF is the indispensable open-source tool that can help all planners decarbonize their systems. This work takes the first major step toward making location and application forecasts using WRF more accessible.

References

- [1] Global Trends in Renewable Energy Investment. Technical report, Frankfurt School - UN Environment Programme Centre, Frankfurt, Germany, 2020.
- [2] Rafael Borge, Vassil Alexandrov, Juan José Del Vas, Julio Lumbreras, and Encarnacion Rodríguez. A comprehensive sensitivity analysis of the WRF model for air quality applications over the Iberian Peninsula. *Atmospheric Environment*, 42:8560–8574 Contents, 2008.

- [3] P Bougeault and P Lacarrere. Parameterization of Orography-Induced Turbulence in a Mesobeta-Scale Model. *Monthly Weather Review*, 117:1872–1890, 1989.
- [4] Leo Breiman. Random forests. *Machine Learning*, 45:5–32, 2001.
- [5] Melissa S Bukovsky and David J Karoly. Precipitation Simulations Using WRF as a Nested Regional Climate Model. *Journal of Applied Meteorology and Climatology*, 48(10):2152–2159, 2009.
- [6] Razvan Cazacu. Comparative Study between the Improved Implementation of 3 Classic Mutation Operators for Genetic Algorithms. *Procedia Engineering*, 181:634–640, 2017.
- [7] Xsitaaz T Chadee, Naresh R Seegobin, and Ricardo M Clarke. Optimizing the Weather Research and Forecasting (WRF) Model for Mapping the Near-Surface Wind Resources over the Southernmost Caribbean Islands of Trinidad and Tobago. *Energies*, 10(931), 2017.
- [8] Sandeep Chinta and C. Balaji. Calibration of WRF model parameters using multiobjective adaptive surrogate model-based optimization to improve the prediction of the Indian summer monsoon. *Climate Dynamics*, 55(3-4):631–650, 2020.
- [9] D P Dee, S M Uppala, A J Simmons, P Berrisford, P Poli, S Kobayashi, U Andrae, M A Balmaseda, G Balsamo, P Bauer, P Bechtold, A C M Beljaars, L van de Berg, J Bidlot, N Bormann, C Delsol, R Dragani, M Fuentes, A J Geer, L Haimberger, S B Healy, H Hersbach, L Isaksen, P Kållberg, M Köhler, M Matricardi, A P McNally, B M Monge-Sanz, J-j Morcrette, B-k Park, C Peubey, P de Rosnay, C Tavorato, J-n Thépaut, F Vitart, Beljaars

- Acm, van L de Berg, Morcrette J-j, Park B-k, and de P Rosnay. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society Q. J. R. Meteorol. Soc.*, 137:553–597, 2011.
- [10] Liza I. Díaz-Isaac, Thomas Lauvaux, Marc Bocquet, and Kenneth J. Davis. Calibration of a multi-physics ensemble for estimating the uncertainty of a greenhouse gas atmospheric transport model. *Atmospheric Chemistry and Physics*, 19(8):5695–5718, apr 2019.
- [11] Caroline Draxl, Andrew Clifton, Bri-Mathias Hodge, and Jim Mccaa. The Wind Integration National Dataset (WIND) Toolkit. *Applied Energy*, 151:355–366, 2015.
- [12] Caroline Draxl, Andrea N. Hahmann, Alfredo Pena, and Gregor Giebel. Evaluating winds and vertical wind shear from Weather Research and Forecasting model forecasts using seven planetary boundary layer schemes. *Wind Energy*, 17:39–55, 2014.
- [13] Jimy Dudhia. Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model. *Journal of the Atmospheric Sciences*, 42(20):3077–3107, 1989.
- [14] Anna C. Fitch. Notes on using the mesoscale wind farm parameterization of Fitch et al. (2012) in WRF. *Wind Energy*, 19:1757–1758, 2015.
- [15] Anna C Fitch, Joseph B Olson, Julie K Lundquist, Jimy Dudhia, Alok K Gupta, John Michalakes, and Idar Barstad. Local and Mesoscale Impacts of Wind Farms as Parameterized in a Mesoscale NWP Model. *Monthly Weather Review*, 140:3017–3038, 2012.

- [16] Emmanouil Flaounas, Sophie Bastin, and Serge Janicot. Regional climate modelling of the 2006 West African monsoon: sensitivity to convection and planetary boundary layer parameterisation using WRF. *Climate Dynamics*, 36:1083–1105, 2011.
- [17] R. Floors, A. N. Hahmann, and A. Peña. Evaluating Mesoscale Simulations of the Coastal Flow Using Lidar Measurements. *Journal of Geophysical Research: Atmospheres*, 123(5):2718–2736, mar 2018.
- [18] William C Forsythe, Edward J Rykiel, Randal S Stahl, Hsin-I Wu, and Robert M Schoolfield. A model comparison for daylength as a function of latitude and day of year. Technical report, 1995.
- [19] Harold Gamarro, Jorge E. Gonzalez, and Luis E. Ortiz. On the Assessment of a Numerical Weather Prediction Model for Solar Photovoltaic Power Forecasts in Cities. *Journal of Energy Resources Technology*, 141, 2019.
- [20] Evangelia-Maria Giannakopoulou and Regis Nhili. WRF Model Methodology for Offshore Wind Energy Applications. *Advances in Meteorology*, 2014.
- [21] Robert C. Gilliam and Jonathan E. Pleim. Performance assessment of new land surface and planetary boundary layer physics in the WRF-ARW. *Journal of Applied Meteorology and Climatology*, 49(4):760–774, apr 2010.
- [22] J J Gómez-Navarro, C C Raible, and S Dierer. Sensitivity of the WRF model to PBL parametrisations and nesting techniques: evaluation of wind storms over complex terrain. *Geosci. Model Dev*, 8:3349–3363, 2015.
- [23] R.N. Greenwell, J.E. Angus, and M. Finck. Optimal mutation probability for genetic algorithms. *Mathematical and Computer Modelling*, 21(8):1–11, apr 1995.

- [24] Jiajun Gu, Bo Yang, Michael Brauer, and K. Max Zhang. Enhancing the evaluation and interpretability of data-driven air quality models. *Atmospheric Environment*, 246, 2021.
- [25] Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Di-nand Schepers, Adrian Simmons, Cornel Soci, Saleh Abdalla, Xavier Abellan, Gianpaolo Balsamo, Peter Bechtold, Gionata Biavati, Jean Bidlot, Massimo Bonavita, Giovanna De Chiara, Per Dahlgren, Dick Dee, Michail Diamantakis, Rossana Dragani, Johannes Flemming, Richard Forbes, Manuel Fuentes, Alan Geer, Leo Haimberger, Sean Healy, Robin J. Hogan, Elías Hólm, Marta Janisková, Sarah Keeley, Patrick Laloyaux, Philippe Lopez, Cristina Lupu, Gabor Radnoti, Patricia de Rosnay, Iryna Rozum, Freja Vamborg, Sebastien Villaume, and Jean Noël Thépaut. The ERA5 global re-analysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, jul 2020.
- [26] Song You Hong, Yign Noh, and Jimmy Dudhia. A new vertical diffusion package with an explicit treatment of entrainment processes. *Monthly Weather Review*, 134(9):2318–2341, sep 2006.
- [27] Michael J. Iacono, Jennifer S. Delamere, Eli J. Mlawer, Mark W. Shephard, Shepard A. Clough, and William D. Collins. Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research*, 113, jul 2008.
- [28] Hisham Ihshaish, Ana Cortés, and Miquel A. Senar. Parallel multi-level genetic ensemble for numerical weather prediction enhancement. *Procedia Computer Science*, 9:276–285, 2012.

- [29] Pedro A. Jiménez and Jimmy Dudhia. Improving the representation of resolved and unresolved topographic effects on surface wind in the wrf model. *Journal of Applied Meteorology and Climatology*, 51(2):300–316, feb 2012.
- [30] Pedro A. Jiménez and Jimmy Dudhia. On the ability of the WRF model to reproduce the surface wind direction over complex terrain. *Journal of Applied Meteorology and Climatology*, 52(7):1610–1617, jul 2013.
- [31] Pedro A Jimenez, Joshua P Hacker, Ellen Haupt, Jose A Ruiz-Arias, Chris A Gueymard, Gregory Thompson, Trude Eidhammer, and Aijun Deng. WRF-Solar is the first NWP model specifically designed to meet the growing demand for specialized numerical forecast products for solar power applications. WRF-SOLAR Description and Clear-Sky Assessment of an Augmented NWP Model for Solar Power Prediction. *Bulletin of the American Meteorological Society*, 97:1249–1264, 2015.
- [32] Pedro A. Jiménez, Jorge Navarro, Ana M. Palomares, and Jimmy Dudhia. Mesoscale modeling of offshore wind turbine wakes at the wind farm resolving scale: A composite-based analysis with the Weather Research and Forecasting model over Horns Rev. *Wind Energy*, 18(3):559–566, mar 2015.
- [33] Jiming Jin, Norman L Miller, and Nicole Schlegel. Sensitivity Study of Four Land Surface Schemes in the WRF Model. *Advances in Meteorology*, 2010.
- [34] Joseph C. Y. Lee and Julie K. Lundquist. Evaluation of the wind farm parameterization in the Weather Research and Forecasting model (version 3.8.1) with meteorological and turbine power data. *Geoscientific Model Development Discussions*, (1993):1–31, 2017.

- [35] Francisco J.L. Lima, Fernando R. Martins, Enio B. Pereira, Elke Lorenz, and Detlev Heinemann. Forecast for surface solar irradiance at the Brazilian Northeastern region using NWP model and artificial neural networks. *Renewable Energy*, 87:807–818, mar 2016.
- [36] Yanluan Lin and Brian A. Colle. A new bulk microphysical scheme that includes riming intensity and temperature-dependent ice characteristics. *Monthly Weather Review*, 139(3):1013–1035, mar 2011.
- [37] Francis M. Lopes, Hugo G. Silva, Rui Salgado, Afonso Cavaco, Paulo Canhoto, and Manuel Collares-Pereira. Short-term forecasts of GHI and DNI for solar energy systems operation: assessment of the ECMWF integrated forecasting system in southern Portugal. *Solar Energy*, 170:14–30, aug 2018.
- [38] Scott M Lundberg, Gabriel Erion, Hugh Chen, Alex DeGrave, Jordan M Prutkin, Bala Nair, Ronit Katz, Jonathan Himmelfarb, Nisha Bansal, and Su-In Lee. From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, 2(1):56–67, 2020.
- [39] J. K. Lundquist, K. K. DuVivier, D. Kaffine, and J. M. Tomaszewski. Costs and consequences of wind turbine wake effects arising from uncoordinated wind energy development. *Nature Energy* 2018 4:1, 4:26–34, 11 2018.
- [40] Edward R. Mansell, Conrad L. Ziegler, and Eric C. Bruning. Simulated electrification of a small thunderstorm with two-moment bulk microphysics. *Journal of the Atmospheric Sciences*, 67(1):171–194, jan 2010.
- [41] Amaya Martinez-Gracia, Inmaculada Arauzo, and Javier Uche. Solar energy availability. In *Solar Hydrogen Production: Processes, Systems and Technologies*, pages 113–149. Elsevier, aug 2019.

- [42] Patrick Mathiesen, Craig Collier, and Jan Kleissl. A high-resolution, cloud-assimilating numerical weather prediction model for solar irradiance forecasting. 2013.
- [43] Patrick Mathiesen and Jan Kleissl. Evaluation of numerical weather prediction for intra-day solar forecasting in the continental United States. *Solar Energy*, 85:967–977, 2011.
- [44] Cristian Mattar and Dager Borvarán. Offshore wind power simulation by using WRF in the central coast of Chile. *Renewable Energy*, 94:22–31, aug 2016.
- [45] K. L. Mills, J. J. Filliben, and A. L. Haines. Determining relative importance and effective settings for genetic algorithm control parameters. *Evolutionary Computation*, 23(2):309–342, 2015.
- [46] K L Mills, J J Filliben, A L Haines, and Warren Rogers Associates. Determining Relative Importance and Effective Settings for Genetic Algorithm Control Parameters. Technical report.
- [47] Seyedali Mirjalili. *Genetic Algorithm*, pages 43–55. Springer International Publishing, Cham, 2019.
- [48] Eli J. Mlawer, Steven J. Taubman, Patrick D. Brown, Michael J. Iacono, and Shepard A. Clough. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research Atmospheres*, 102(14):16663–16682, jul 1997.
- [49] P A Mooney, F J Mulligan, and R Fealy. Evaluation of the Sensitivity of the Weather Research and Forecasting Model to Parameterization Schemes for

Regional Climates of Europe over the Period 1990-95. *Journal of Climate*, 26:1002–1017, 2013.

- [50] Liviu Oana and Adrian Spataru. Use of genetic algorithms in numerical weather prediction. In *Proceedings - 18th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, SYNASC 2016*, pages 456–461, 2016.
- [51] Linlin Pan, Yubao Liu, Jason C Knievel, Luca Delle Monache, and Gregory Roux. Evaluations of WRF Sensitivities in Surface Simulations with an Ensemble Prediction System. *Atmosphere*, 9(106), 2018.
- [52] Mohammadreza Mohammadpour Penchah, Hossein Malakooti, and Mohammad Satkin. Evaluation of planetary boundary layer simulations for wind resource study in east of Iran. *Renewable Energy*, 111:1–10, 2017.
- [53] Roger A Pielke. *Mesoscale Meteorological Modeling*, volume 98. Academic Press, New York, 3 edition, 2013.
- [54] Jonathan E. Pleim. A simple, efficient solution of flux-profile relationships in the atmospheric surface layer. *Journal of Applied Meteorology and Climatology*, 45(2):341–347, 2006.
- [55] Jonathan E. Pleim. A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: Model description and testing. *Journal of Applied Meteorology and Climatology*, 46(9):1383–1395, 2007.
- [56] Jonathan E. Pleim. A combined local and nonlocal closure model for the atmospheric boundary layer. Part II: Application and evaluation in a mesoscale meteorological model. *Journal of Applied Meteorology and Climatology*, 46(9):1396–1409, 2007.

- [57] Jiping Quan, Zhenhua Di, Qingyun Duan, Wei Gong, Chen Wang, Yanjun Gan, Aizhong Ye, and Chiyuan Miao. An evaluation of parametric sensitivities of different meteorological variables simulated by the WRF model. *Quarterly Journal of the Royal Meteorological Society Q. J. R. Meteorol. Soc.*, 142:2925–2934, 2016.
- [58] L S Rothman, I E Gordon, A Barbe, D Chris Benner, P F Bernath, M Birk, V Boudon, L R Brown, A Campargue, J.-P Champion, K Chance, L H Coudert, V Dana, V M Devi, S Fally, J.-M Flaud, R R Gamache, A Goldman, D Jacquemart, I Kleiner, N Lacome, W J Lafferty, J.-Y Mandin, S T Massie, S N Mikhailenko, C E Miller, N Moazzen-Ahmadi, O V Naumenko, A V Nikitin, J Orphal, V I Perevalov, A Perrin, A Predoi-Cross, C P Rinsland, M Rotger, M A H Smith, K Sung, S A Tashkun, J Tennyson, R A Toth, A C Vandaele, and J Vander Auwera. The HITRAN 2008 molecular spectroscopic database. *Journal of Quantitative Spectroscopy & Radiative Transfer*, 110:533–572, 2009.
- [59] W.C. Skamarock, J.B. Klemp, Jimy Dudhia, D.O. Gill, Liu Zhiqun, Judith Berner, Wei Wang, J.G. Powers, Michael G Duda, Dale M Barker, and Xiang-Yu Huang. A Description of the Advanced Research WRF Model Version 4. Technical report, NCAR, Boulder, CO, 2019.
- [60] William C Skamarock and Joseph B Klemp. A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *Journal of Computational Physics*, 227:3465–3485, 2008.
- [61] G. L. Stephens. The parameterization of radiation for numerical weather prediction and climate models. *Monthly Weather Review*, 112(4):826–867, apr 1984.

- [62] Ioannis Stergiou, Efthimios Tagaris, and Rafaella-Eleni P. Sotiropoulou. Sensitivity Assessment of WRF Parameterizations over Europe. In *2nd International Electronic Conference on Atmospheric Sciences*, 2017.
- [63] Carolin Strobl, Anne Laure Boulesteix, Achim Zeileis, and Torsten Hothorn. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, 8(1), dec 2007.
- [64] Jeffrey A. Sward. OptWRF, 2020.
- [65] Gregory Thompson, Paul R. Field, Roy M. Rasmussen, and William D. Hall. Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Monthly Weather Review*, 136(12):5095–5115, dec 2008.
- [66] A. Xiu and J. E. Pleim. Development of a land surface model. Part I: Application in a mesoscale meteorological model. *Journal of Applied Meteorology*, 40(2):192–209, feb 2001.
- [67] Ben Yang, Larry K Berg, Yun Qian, Chen Wang, Zhangshuan Hou, Ying Liu, Hyeyum Hailey Shin, Songyou Hong, and Mikhail Pekour. Parametric and Structural Sensitivities of Turbine-Height Wind Speeds in the Boundary Layer Parameterizations in the Weather Research and Forecasting Model. *Journal of Geophysical Research: Atmospheres*, 124:5951–5969, 2019.
- [68] Xinjie Yu and Mitsuo Gen. *Introduction to Evolutionary Algorithms*. Springer, 2010.
- [69] Melina-Maria Zempila, Theodore M Giannaros, Alkiviadis Bais, Dimitris Melas, and Andreas Kazantzidis. Evaluation of WRF shortwave radiation

parameterizations in predicting Global Horizontal Irradiance in Greece. *Renewable Energy*, 86:831–840, 2016.

- [70] Jing Zhao, Zhen-Hai Guo, Zhong-Yue Su, Zhi-Yuan Zhao, Xia Xiao, and Feng Liu. An improved multi-step forecasting model based on wrf ensembles and creative fuzzy systems for wind speed. *Applied Energy*, 162:808–826, 2016.

4.6 Supporting Information

4.6.1 WRF Eta Levels

The custom eta levels used in the model are shown in Table 4.2. Note that these were set manually to accommodate the use of the WRF wind farm parameterization (WFP) [15], which performs better when a greater number of cells exist lower in the modeling domain. For those applications that omit the WFP, default levels will suffice.

Table 4.2: WRF Eta Levels

1.000,	0.9975,	0.995,	0.990,	0.985,	0.980,
0.970,	0.960,	0.950,	0.940,	0.930,	0.920,
0.910,	0.900,	0.880,	0.860,	0.840,	0.820,
0.800,	0.770,	0.740,	0.700,	0.650,	0.600,
0.550,	0.500,	0.450,	0.400,	0.350,	0.300,
0.250,	0.200,	0.150,	0.100,	0.050,	0.000

4.6.2 WRF Namelist Parameters

Constant WRF namelist parameters are provided in Table 4.3. Any parameters not listed in that table were left at their WRFv4.0 default values.

4.6.3 WRF Physics Parameterization Options

All the physics parameterizations options available within WRF version 4.2.1 are given below along with their respective namelist identifiers.

Table 4.3: Constant Namelist Options

interval_seconds	10800
history_interval	60
time_step	45
max_dom	1
e_we	192
e_sn	192
e_vert	36
num_metgrid_levels	38
num_metgrid_soil_levels	4
dx	12000
dy	12000
swint_opt	1
radt	15
bldt	0
cudt	0
surface_input_source	1
num_soil_layers	4
num_land_cat	21
damp_opt	0

Dependencies among parameterizations

Given their diverse development origins, it is often impossible to tell whether or not a given scheme is even compatible with an arbitrary set of other physics options unless that set-up is reported in previous literature. Some parameterization dependencies are also documented in the WRF User’s Guide. During this study, we discovered that, generally, well over 50% of randomly chosen physics option sets contain incompatibilities not explicitly reported in the User’s Guide.

4.6.4 Notes on parallelism and computing

WRF runs best compiled for distributed memory parallelism in a high-performance computing (HPC) system such as NCAR Cheyenne supercom-

Table 4.4: WRF Microphysics (MP) Parameterization Options

Scheme Name	Namelist Option
Kressler	1
Lin	2
WSM 3	3
WSM 5	4
Eta	5
WSM 6	6
Goddard	7
Thompson	8
Milbrandt 2-moment	9
Morrison 2-moment	10
CAM 5.1	11
Sbu-Ylin	13
WDM 5	14
Ferrier	15
WDM 6	16
NSSL 2-moment	17
NSSL 2-moment w/ CCN	18
NSSL 1-moment	19
NSSL 1-moment LFO	21
NSSL 1-moment no hail	22
Thompson Aerosol	28
Huji (fast)	30
Huji (full)	32
P3	50
P3 2-moment	51

Table 4.5: WRF Longwave Radiation (LW) Parameterization Options

Scheme Name	Namelist Option
RRTM	1
CAM	3
RRTMG	4
New Goddard	5
FLG	7
RRTMG (fast)	24
Held-Suarez	31
GFDL	99

Table 4.6: WRF Shortwave (SW) Radiation Parameterization Options

Scheme Name	Namelist Option
Dudhia	1
Goddard	2
CAM	3
RRTMG	4
New Goddard	5
FLG	7
RRTMG (fast)	24
GFDL	99

Table 4.7: WRF Land Surface (LSM) Parameterization Options

Scheme Name	Namelist Option
5-layer	1
Noah	2
RUC	3
Noah-mp	4
CLM4	5
Pleim-Xiu	7

Table 4.8: WRF PBL Parameterization Options

Scheme Name	Namelist Option
YSU	1
MYJ	2
QNSE	4
MYNN 2	5
MYNN 3	6
ACM 2	7
BouLac	8
UW	9
TEMF	10
Shin-Hong	11
GBM	12

puter [59]. Within this environment, new users can copy pre-compiled WRF executables, which substantially lowers the barrier to entry as an arbitrary number of complexities can arise when compiling WRF on a different system. WRF's workflow is also quite constrained especially when interacting with an HPC

Table 4.9: WRF Cumulus Parameterization Options

Scheme Name	Namelist Option
Kain-Fritsch	1
BJM	2
Grell-Freitas	3
GFS SAS	4
Grell-3D	5
TiedTKE	6
Zhang-McFarlane	7
Modified Kain-Fritsch	10
Multi Kain-Fritsch	11
New GFS SAS	14
New TiedTKE	16
Grell-Devenyi	93
HWRF GFS SAS	84
Old Kain-Fritsch	99

scheduler. For example, the WRF preprocessing system (WPS) is generally compiled in serial and only needs to be run once for each domain; whereas both the Real and WRF executables are compiled in parallel and must be run each time any of the namelist options change. In this study, we submitted multiple Real and WRF parallel jobs to the scheduler at the same time from different threads using the python *concurrent.futures* module. However, as some WRF helper scripts (e.g. `link_grib.csh`) presume that users will run WRF from the current directory, modifications were necessary. All python code and modified WRF scripts are available in our public OptWRF GitHub repository [64].

4.6.5 Day Length Model

Equation 4.2, reproduced from Forsythe et al. [18], provides a method for estimating day length based upon latitude and day of the year.

$$D = 24 - \frac{24}{\pi} \arccos \left(\frac{\sin \frac{p\pi}{180} + \sin \frac{L\pi}{180} \sin \phi}{\cos \frac{L\pi}{180} \cos \phi} \right), \quad (4.2)$$

where,

$$\theta = 0.2163108 + 2 \arctan (0.9671396 \tan 0.00860 \times (J - 186)),$$

and,

$$\phi = \arcsin (0.39795 \cos \theta).$$

D is the day length (inclusive of twilight), L is latitude, J is Julian date, and p is the daylight coefficient, in degrees. We retained the value of 0.8333 for p from Forsythe et al., and both θ and ϕ are in radians.

4.6.6 Error and Fitness Scatter Plots

Below is a panel of scatter plots showing GHI error, WPD error, and fitness values for each of the almost 2000 OptWRF simulation which comprises our database. Note the seasonal nature of the GHI error values – higher in the summer – but that this seasonality is removed by the day length model. A difference in fitness values still appears with seasonality but this can be at least partially attributed to the larger variability in wind during the winter and spring.

4.6.7 Shapley Value Plots for Microphysics, Land Surface, and Cumulus Schemes

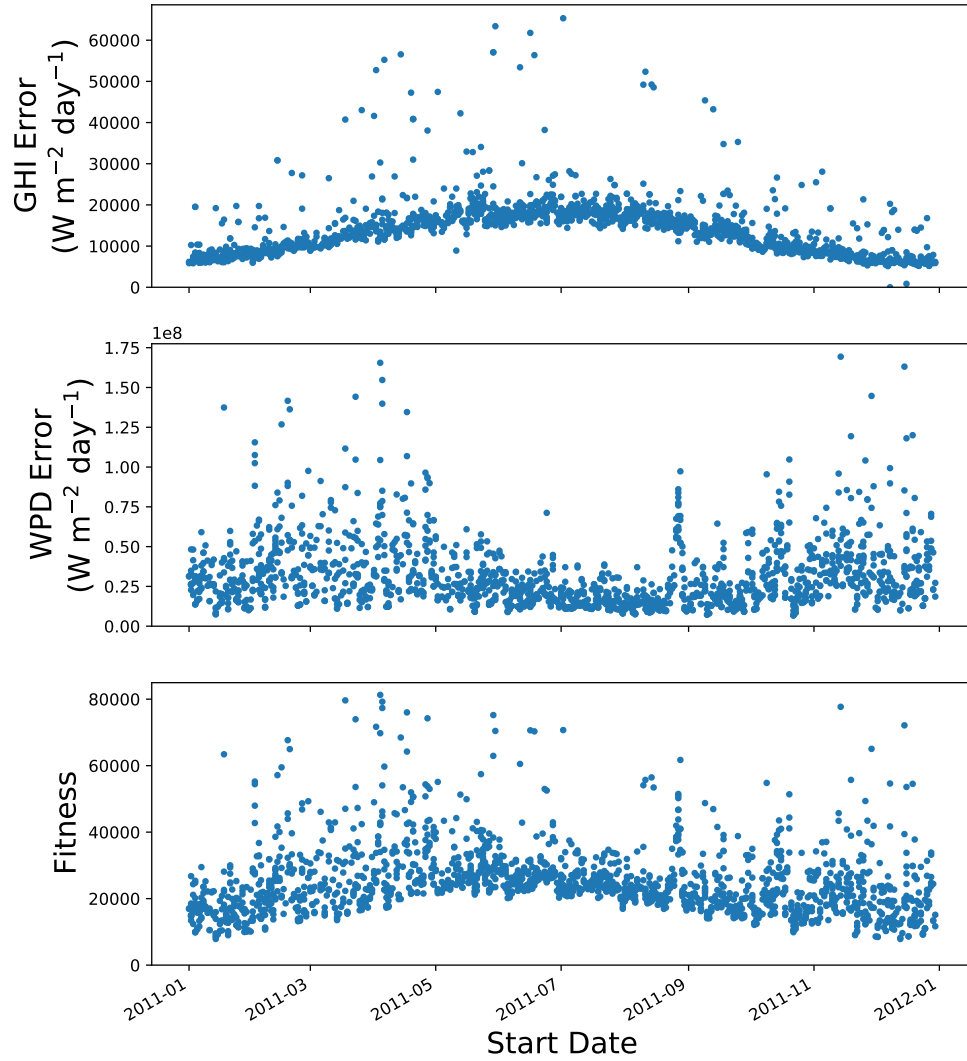


Figure 4.7: Scatter plots showing GHI error, WPD error, and fitness values on the top, middle, and bottom, respectively. Each individual point corresponds to a single (of more than 2000) distinct WRF simulation. GHI error clearly shows a seasonal trend attributable to the changing number of daylight hours throughout the year. This provides a visual justification for the correction of this error metric by daylight fraction. WPD error shows a greater spread during the non-Summer months, but no other obvious seasonal patterns exist. Notice that WPD error values are approximately four orders of magnitude higher than those for GHI error warranting a correction factor. Finally, overarching trends in both the GHI and WPD errors are represented in the fitness values.

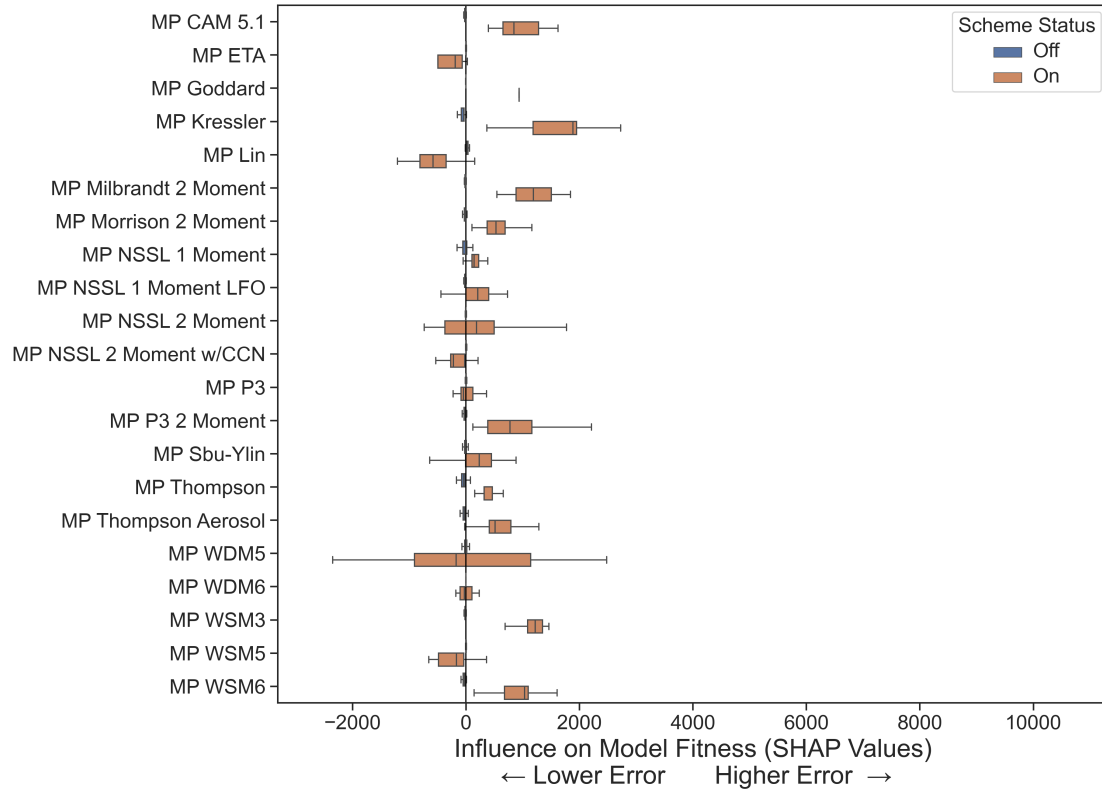


Figure 4.8: SHAP values explain the impact that each random forest model feature (WRF parameterization scheme) has upon model error. Microphysics parameterization schemes are listed along the left axis, and boxes are created from all the WRF model runs. Orange bars correspond to WRF models where the parameterization scheme was activated; whereas the blue bars correspond to those where the scheme was deactivated. Negative SHAP values correspond to those schemes that reduce model error (improving the forecast); whereas positive SHAP values correspond to those schemes that increase model error. Therefore, those schemes with orange bars centered furthest to the left produce the best WRF forecasts.

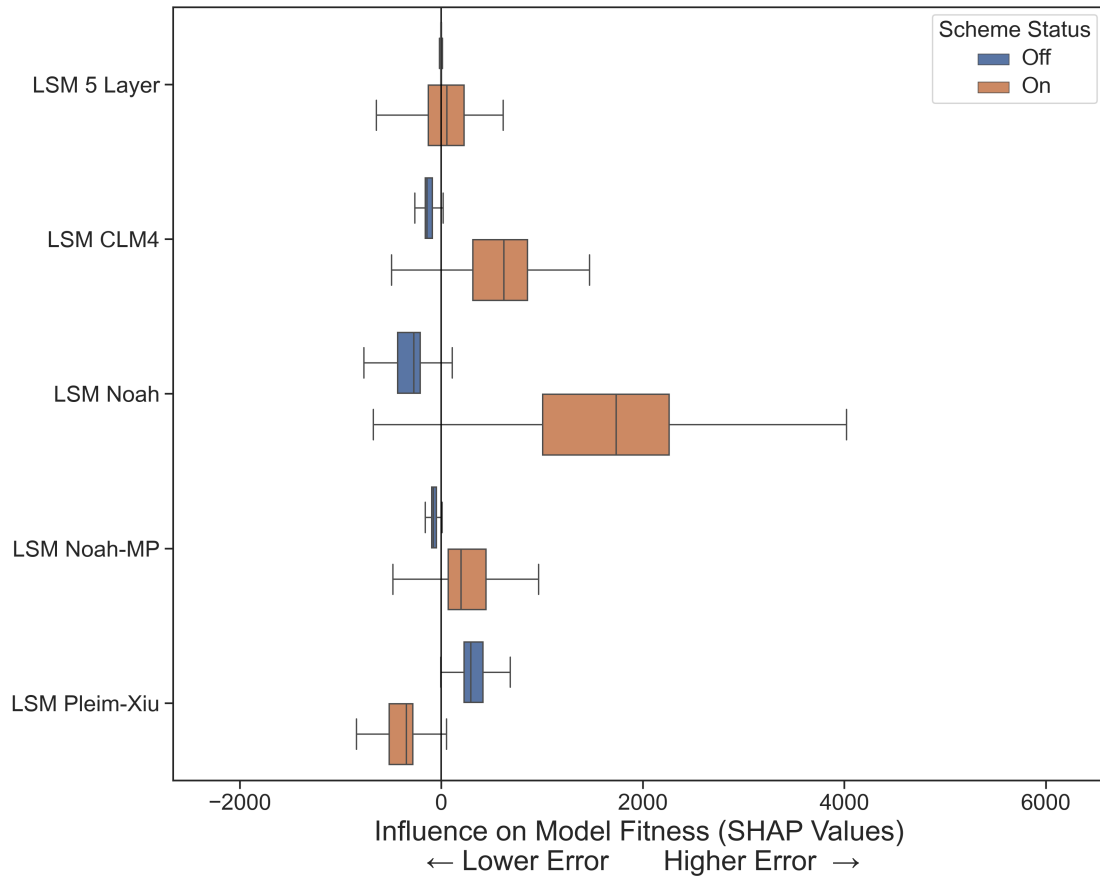


Figure 4.9: SHAP values explain the impact that each random forest model feature (WRF parameterization scheme) has upon model error. Land surface parameterization schemes are listed along the left axis, and boxes are created from all the WRF model runs. Orange bars correspond to WRF models where the parameterization scheme was activated; whereas the blue bars correspond to those where the scheme was deactivated. Negative SHAP values correspond to those schemes that reduce model error (improving the forecast); whereas positive SHAP values correspond to those schemes that increase model error. Therefore, those schemes with orange bars centered furthest to the left produce the best WRF forecasts.

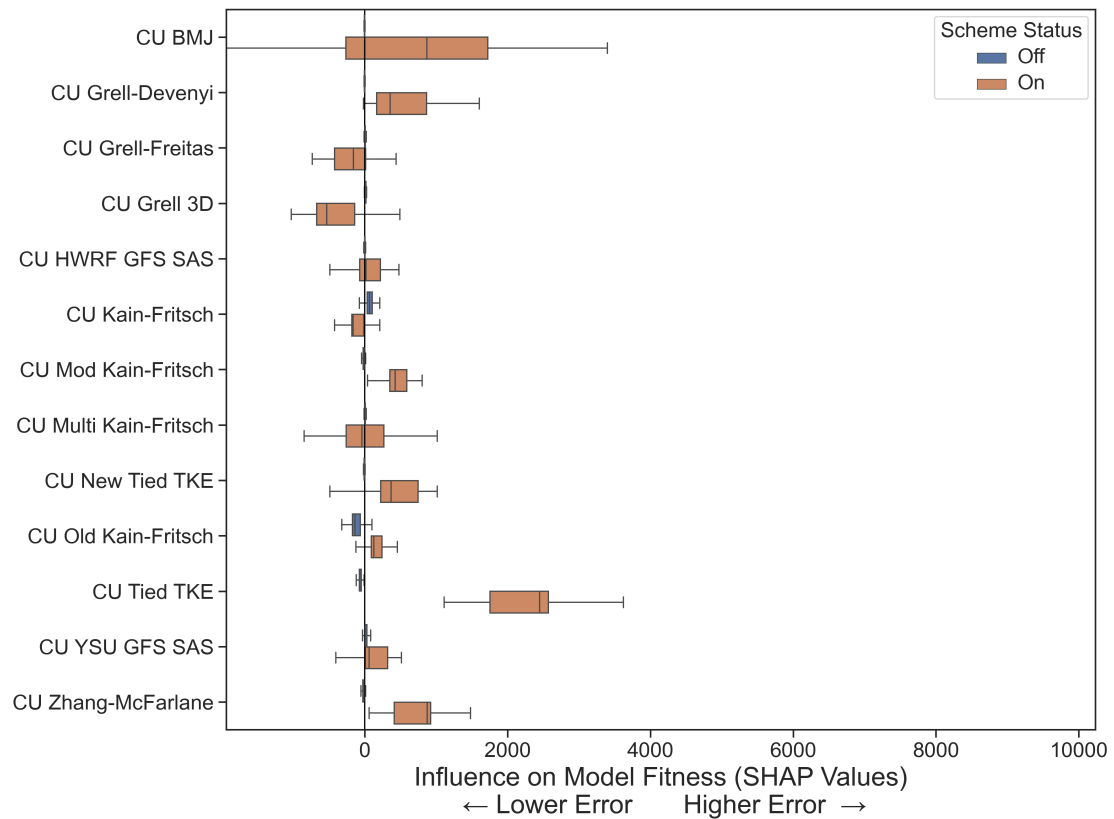


Figure 4.10: SHAP values explain the impact that each random forest model feature (WRF parameterization scheme) has upon model error. Cumulus parameterization schemes are listed along the left axis, and boxes are created from all the WRF model runs. Orange bars correspond to WRF models where the parameterization scheme was activated; whereas the blue bars correspond to those where the scheme was deactivated. Negative SHAP values correspond to those schemes that reduce model error (improving the forecast); whereas positive SHAP values correspond to those schemes that increase model error. Therefore, those schemes with orange bars centered furthest to the left produce the best WRF forecasts.

CHAPTER 5

AN INTERNALLY CONSISTENT FRAMEWORK FOR DETERMINING
THE AIR QUALITY CO-BENEFITS OF WIND AND SOLAR
DEVELOPMENT

Abstract

Wind and solar enjoy broad support and therefore offer a politically feasible way to combat climate change in the absence of policies that directly regulate carbon emissions. However, wind and solar provide society with additional benefits by displacing electricity generated by power plants that cause harmful air pollution. These “co-benefits” can be quantified by linking a numerical weather prediction tool, the Weather Research and Forecasting (WRF) model for example, with a chemical transport model, such as the Community Multiscale Air Quality (CMAQ) model. Historically, weather modeling has taken a back seat in these studies. But in a future where our electricity comes largely from wind and solar, the weather plays a central role in the reliability of all systems. Therefore, we propose a framework for determining the air quality co-benefits of future wind and solar development that relies on the same WRF downscaling to produce estimates for wind and solar generation as well as provide meteorological inputs for CMAQ. As an example, we determine the changes in ozone and $PM_{2.5}$ concentration across the Northeastern United States associated with expected near-term wind and solar development in New York State (NYS).

5.1 Introduction

How about that weather? This question often escapes our lips when we find ourselves searching for the correct words to fill that awkward space between us and a stranger. But awkwardly enough, the same question captures our uncertainty about the future smooth operation of post-industrial society. Governments across the world have pledged to wean their economies off fossil fuels, and most have hopped on the electric train – that is, decarbonize the economy by electrifying everything and generating clean electricity using renewable energy. In this future, the weather is promoted from a subject for small talk to the lifeblood of society. This means we require tools built around the weather, characterizing its uncertainty, and devising contingencies for when it behaves unexpectedly.

Of course, a fundamental challenge that stymies decarbonization is that the mechanisms by which economic growth is achieved – and the underlying structure of the economy – are at odds with efforts to mitigate climate change. This conflict exists because unregulated markets cannot decarbonize society at a rate necessary to avoid climate catastrophe [31]. Such markets fail to value negative externalities associated with fossil-based energy production and consumption. Existing research theorizes alternative economic structures (e.g. circular economy [15] and degrowth [20]), but societal reorganizations of this magnitude tend to happen on generational time scales. To avoid the worst effects of climate change this century, we require a method for “correcting” markets, which would improve the value proposition of renewable energy and put a price on emissions and environmental degradation. In the United States, we have unfortunately remained unsuccessful at passing any national-scale legislation that

would put a price on carbon, and nearly all progress in regulating emissions has worked by expanding the Clean Air Act. Many of us remember all too well the failures of the Waxman-Markey cap and trade bill in 2009 followed by the Clean Power Plan in 2015. Therefore, an alternative way to aid the market in transitioning the US to a clean economy is by valuing the co-benefits of renewable energy (e.g., air quality, water quality, public health, etc.).

Here, we focus on the air quality improvements associated with future renewable energy development. The justification for this is three-fold. First, renewable energy development has been driven both by market forces as well as government incentives, and while the levelized cost of energy for wind and solar often beats that of conventional resources, these zero marginal cost resources tend to cannibalize their own profits as more of them appear within the same electrical system. Second, the Clean Air Act is arguably the most successful piece of environmental legislation of all time in the US and can be utilized to further climate goals given the current state of congressional gridlock. Finally, the supply of renewable energy and air quality both depend intimately on the weather, so a framework that investigates them together offers greater internal consistency.

Existing literature has sought to quantify the air quality benefits associated with an array of climate and clean energy scenarios. A substantial number of studies tackle this problem using purely data-driven methods including spatial autocorrelation [35], quantile regression [21], and the Environmental Kuznets Curve (EKC) [5]. While these methods offer computationally efficient methods for teasing out the relationship between air pollution and renewable energy, they do not attempt to capture the physics driving air pollution, and therefore,

cannot elucidate changes in air pollution patterns that may result from a future policy action. Such studies must employ a chemical transport model (CTM) and devise a methodology for estimating expected emissions changes under a given policy. A recent review by Gallagher and Holloway remarks on the sparsity of studies occupying this interdisciplinary space and highlights the disconnect between research focusing on energy and climate as opposed to air quality and health [13].

Still, over the past decade, several multidisciplinary studies have quantified emissions changes based on an existing inventory and used a CTM to determine the commensurate change in pollutant concentrations over a region or country. Three main methods exist to develop a new emissions inventory for a sector of interest. The first simply scales an existing emissions inventory by making assumptions about the emissions response to a given action. For example, Gammarra et al. use published emissions factors and projected changes in the Spanish energy system coupled to the CHIMERE model to investigate the air quality benefit of increased electric vehicle (EV) use and biofuel blending [14]. Also, Peng et al. investigated the benefit of sending electricity generated with a mix of coal and renewables vs. coal only via proposed transmission lines in China using WRF-CHEM [27]. The second approach utilizes an energy-economy model to forecast changes in the generation mix and corresponding emissions over the long term. In this category, Plachinski et al. study increased efficiency and renewable energy in a 2024 policy scenario for Wisconsin using the MyPower electricity sector model [28]. MyPower produces annual load duration curves for each power plant in Wisconsin that were used to estimate emissions changes before CMAQ provided air quality concentration changes across the entire great lakes region in response to Wisconsin's policy. Schmid et al. used the TIMES

PanEU model to predict changes in the energy mix and emissions across the European Union, but here, a reduced-form air quality model, EcoSense, provided external pollutant cost factors for a subsequent run of TIMES PanEU [30]. Similar studies have been conducted using different region-specific modeling frameworks in both the US [10] and Australia [24]. Pan et al. considered several combinations of automation, electrification, and ride-sharing in the US transportation sector using the IEA's Mobility Model (MoMo). MoMo outputs were used to adjust the 2011 EPA national emissions inventory, which was then fed into CMAQ [26]. Finally, Abel et al. looked at the co-benefits of energy efficiency using EPA's AVERT tool [4].

However, energy-economy models usually include only crude representations of the power system, which seek to balance the total electricity supply and demand across a large region for a set of representative days. This approach does not take into account existing transmission constraints or fully characterize the availability of renewable energy. Therefore, the third method addresses these shortcomings by using an economic dispatch model to inform changes in power plant emissions. Looking to determine the impact that plug-in EVs and additional wind energy development would have on air quality in Southern California, Razeghi et al. coupled the Spatially and Temporally Resolved Energy and Environment Tool (STREET) to the University of California Irvine – California Institute of Technology (UCI-CIT) atmospheric chemistry and transport model [29]. Interestingly enough, they found that, while EVs will generally benefit urban air quality, vehicle charging profiles have little impact on air quality. Others have used commercial models PROSYM from ABB and GridView from Hitachi to estimate emissions changes associated with renewable energy development or energy-efficiency [6, 3].

While three studies exist that successfully couple dispatch models, one focuses on a relatively small region in California, and the remaining two rely on proprietary representations of the power system. Analyses focusing on small regions tend to ignore the broader regional air quality impacts of local policies as pointed out by Gallagher [13]. Proprietary models not only increase the barrier to entry for groups with limited resources, but modeling assumptions are often difficult if not impossible to track down thereby limiting reproducibility. In this work, we develop a framework, which we refer to as OneMet, for assessing the air-quality co-benefits of renewable energy development and climate policy in New York State (NYS) that centers around WRF as a versatile tool to predict renewable energy supply and provide inputs to CMAQ. We build this framework using only open-source models and data resources – including a representation of the NYS power system created in-house and a machine learning-based electric generating unit (EGU) emission estimation tool. Recognizing that air quality is an inherently regional issue, we adopt the CMAQ domain used by the Ozone Transport Commission (OTC). In this way, we supplement the existing methodology that Northeastern states already use to devise their implementation plans.

The remainder of this chapter is organized as follows Section 5.2 explains the details of OneMet, Section 5.3 presents results covering a 8-day high electricity demand period in 2016, Section 5.4 engages in a discussion on the performance and limitations of OneMet as well as opportunities for extending the framework to other policies and regions. Finally, Section 5.5 summarizes our conclusions.

5.2 Method

Challenges abound in determining the air quality co-benefits associated with variable renewable energy development. Numerous physical processes and man-made systems overlap adding layers of complexity. We designed OneMet (depicted in Figure 5.1) to manage these complicated components and deliver the air quality benefits at a spatial resolution of 4 km across the Ozone Transport Commission’s high-resolution domain, which focuses on the Northeast Corridor. The framework consists of five major steps 1) determining future wind and solar development patterns, 2) deriving wind and solar generation profiles using WRF, 3) dispatching power plants, 4) estimating power plant emissions based on dispatch patterns, and 5) calculating air pollutant concentrations using CMAQ. In the subsections that follow, we will expand upon each of these components in greater detail while focusing on renewable energy targets in NYS to demonstrate the advantages of OneMet.

5.2.1 Future Wind and Solar Development

Exactly where and when new wind and solar facilities come online depends on a combination of the value proposition for developers, federal to local incentives, permitting processes, and public perceptions surrounding individual projects. We have discussed the complexity of these issues in the context of utility-scale solar siting in previous work [33, 32], and we adopt the same methodology for determining the locations of individual solar facilities presented in Sward et al. [33]. Briefly, we filter land from a NYS tax dataset based on class, size, slope, and distance to a distribution-level substation, rank the remaining sites by popula-

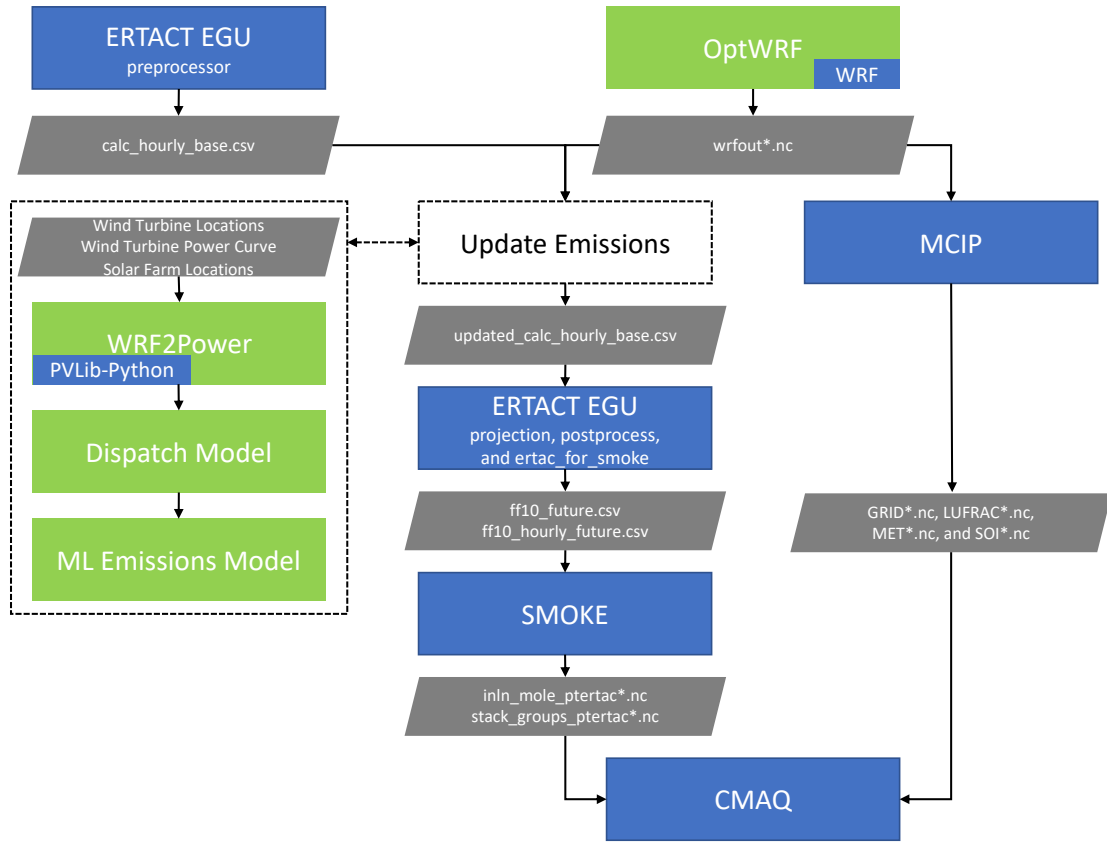


Figure 5.1: Flow chart depicting the main components of the OneMet framework. Blue blocks indicate third-party open-source tools, green blocks indicate internally developed tools, and gray blocks indicate files that are transferred among the various programs. Note that the update emissions block consists of three steps that are shown in the larger dotted rectangle on the left.

tion density and resource quality, and select sites up to a certain policy-relevant level. For this study, we chose a relatively modest solar installation number of 2700 MW to reflect existing solar policy goals in NYS while accounting for projects that are already in the development pipeline. This step simply provides locations and nameplate capacities of future solar developments and, therefore, can easily be swapped for an alternative methodology in a different region.

Given that offshore wind permitting and development timelines span several years, we chose to model four lease areas for which the New York State

Energy Research and Development Authority (NYSERDA) has already signed contracts shown in Figure 5.2. Empire Wind 1 and Sunrise Wind won the 2018 solicitations, which we model with nameplate capacities of 816 MW and 880 MW, respectively. Empire Wind 2 and Beacon Wind won NYSEDA's 2020 solicitations, which we model with nameplate capacities of 1264 MW and 1232 MW, respectively. In total, this sums to 4192 MW of offshore wind – almost half of NYS's 9 GW target for 2035. These nameplate capacities were obtained by placing 8 MW turbines up to the number of turbines that most closely matches the nameplate capacity listed in the corresponding NYSEDA solicitation. In determining exact turbine locations, we assumed that the turbines were separated by five turbine diameters on both major axes. The remaining offshore lease area shown in Figure 5.2, South Fork Wind Farm, is much smaller than the others, so we chose to omit this site. Like that used to determine future solar facilities, the method used to determine individual turbine locations is flexible.

5.2.2 WRF for Wind, Solar, and CMAQ

Meteorological modeling is the common thread connecting variable renewable energy and air quality researchers. However, while energy researchers have adopted ever more sophisticated methods for improving forecasts (e.g., ensemble forecasting [22], statistical post-processing [34], etc.), meteorological modeling remains on the back burner in most air quality studies. This must change particularly for those studies seeking to quantify the air quality co-benefits associated with renewable energy. In such studies, common base-year meteorology oversimplifies the compounding weather-dependent effects of renewable energy supply on air quality. Therefore, we propose using the same WRF model

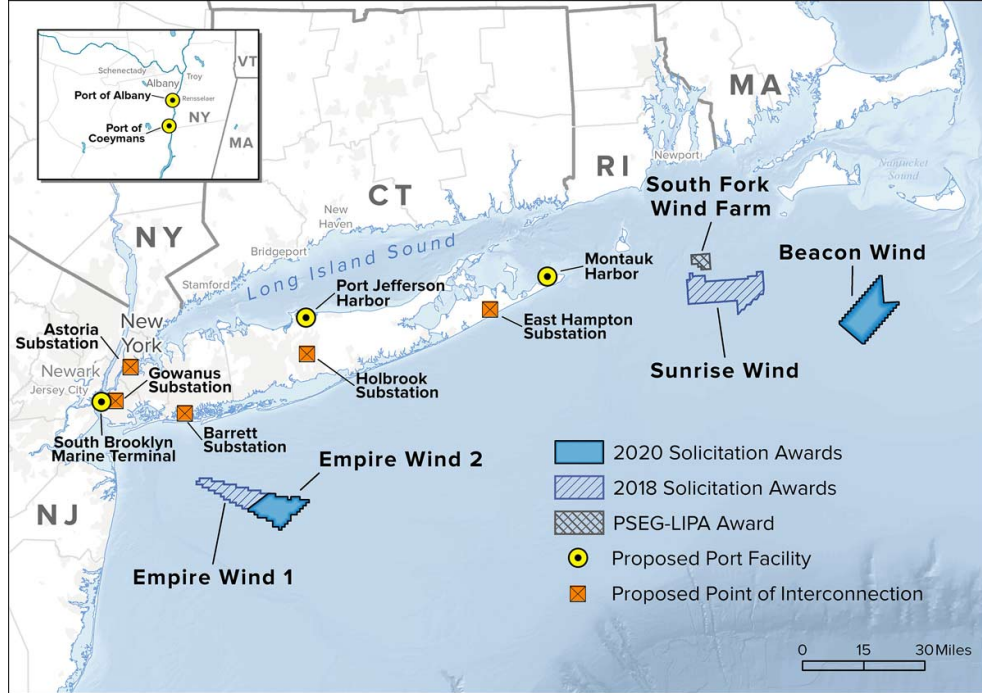


Figure 5.2: Map showing the locations of the four lease areas under study, Empire Wind 1, 2, Beacon Wind, and Sunrise Wind, off the coast of New York State. Map reproduced from NYSERDA [2].

output to estimate renewable energy generation and run CMAQ.

Setting up and benchmarking WRF for an arbitrary application can be a daunting task, and in Chapter 4, we propose a method for optimizing the WRF model for any application in any region, which we referred to as OptWRF. However, the computational burden of OptWRF is substantial for nested domains, so we adopted the best-performing member from the multiphysics WRF ensemble that we investigated in Chapter 3 – the setup proposed by Optis et al. [25]. We feel justified in this choice not only because this setup performed well when benchmarked against NYSERDA’s two offshore wind lidar buoys making it the obvious singular choice for investigating wind energy in the Northeastern United States, but because its ability to predict offshore wind speeds at multiple vertical levels should translate to a solid prediction of the synoptic-

scale winds – i.e., the same winds that drive interstate pollutant transport. The Optis setup uses Thompson microphysics, RRTMG longwave radiation, Goddard shortwave radiation, the Noah land surface model, the MYNN2 planetary boundary layer scheme, no cumulus scheme, and the revised MM5 surface layer scheme. As shown in the diagram (Figure 5.1), we leverage the OptWRF code base to run WRF, which is available on GitHub at [jeffreysward/met4ene](https://github.com/jeffreysward/met4ene).

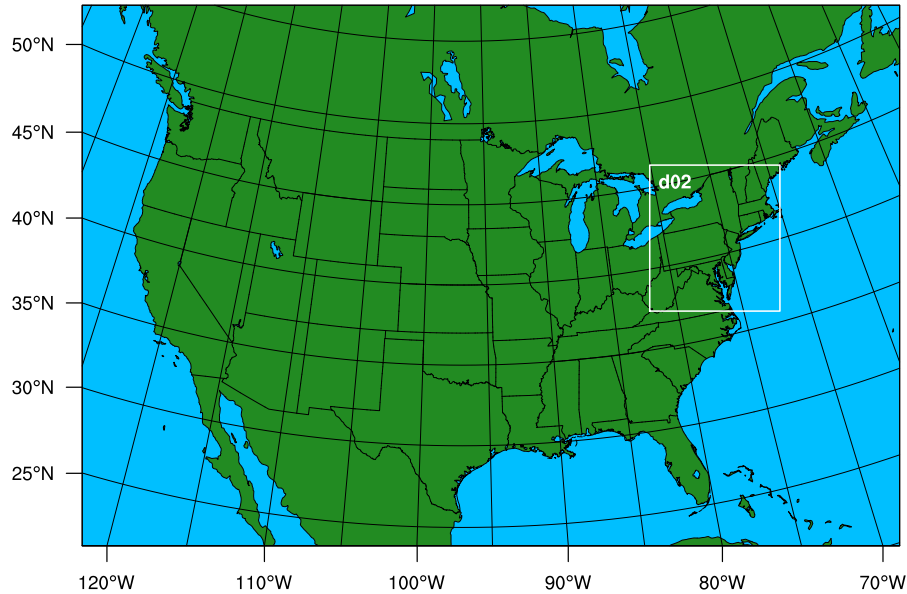


Figure 5.3: Map showing the two WRF domains centered over the Northeastern US. The coarse 12 km domain covers the entire US but is later windowed by MCIP. The fine 4 km domain covers only the OTC region.

Our WRF domain matches the one used by OTC states to conduct air pollutant analyses, shown in Figure 5.3. This domain consists of two two-way nested domains with spatial resolutions of 12 km and 4 km, respectively, uses a Lambert Conformal map projection, and contains 36 vertical layers up to a pressure level of 5000 hPa. Additional details are given in the WRF namelist included in Section 5.6.1. We also edited the WRF Registry to have WRF write the surface downward direct normal irradiance and diffuse irradiance to the output file.

We ran WRF for 11 days starting on August 5, 2016, 00:00 (UTC), and used the ERA5 reanalysis [18] for initial and boundary conditions. In other words, we dynamically downscaled ERA5 over the Northeastern US.

Upon completion, WRF outputs were processed in two separate ways – one produced wind and solar generation profiles – the left pathway in Figure 5.1 – while the other provided CMAQ-ready meteorology files – the right pathway in Figure 5.1. Hourly wind generation profiles for each turbine were created using a published power curve for an 8 MW reference wind turbine [9] and WRF horizontal wind speeds interpolated to 100m. Since we do not account for turbine-turbine interactions, this represents an upper limit for wind power production. However, this WRF setup is compatible with the wind farm parameterization [12], which can offer improved wind power profiles. Hourly solar generation profiles are created with the help of a customized version of `PVLib-Python` [19] that can ingest WRF meteorological data. We use the default PVWatts photovoltaic module and inverter parameters, the Sandia Array Performance Model with open rack glass-polymer configuration, and assume modules have a fixed tilt equal to the site’s latitude. We chose the default PVWatts representations here for clarity and consistency, but `PVLib-Python` includes numerous ways to represent individual solar facilities more realistically if such information is available. Our code allowing users to generate wind and solar profiles using WRF output data is available in the `jeffreysward/wrf2power` repository on GitHub.

CMAQ-ready meteorological inputs were created using the Meteorology-Chemistry Interface Processor (MCIP) version 5.3.3. We processed only subsets of the WRF domain setting $X0 = 141$, $Y0 = 15$, $NCOLS = 273$, and $NROWS = 246$

for the 12 km domain, and $X0 = 87$, $Y0 = 9$, $NCOLS = 126$, and $NROWS = 156$ for the 4 km domain.

5.2.3 Generator Unit Commitment and Economic Dispatch

Earlier, we alluded to the challenge associated with estimating future power plant operational profiles. To accomplish this realistically requires information about the power system's topology. However, a majority of US grid infrastructure is classified as critical infrastructure information, and even if it existed in the public domain, the grid is too complex to model exhaustively for planning purposes. Therefore, reduced form models are necessary to make the problem tractable. Some such models appear in the literature (e.g., Buonocore et al. [6] and Abel et al. [3]). Unfortunately, both of these studies included industrial partners with proprietary grid modeling software meaning that the underlying topology and model assumptions remain hidden from the research community.

To combat this challenge, Liu et al. recently published an open-source representation of the NYS electric grid [23], which we adopt here. Briefly, this representation is a WARD-type equivalent of the Northeastern Power Coordinating Council (NPCC) 140-bus system with a reduced number of busses outside of NYS. To calculate operational profiles for each generator in NYS, we ran a DC-Optimal Power Flow (OPF) using MATPOWER [36] using this network topology. We assumed that both Empire Wind facilities would inject power into New York Independent System Operator (NYISO) load zone J, whereas Sunrise and Beacon wind would inject power into NYISO zone K. Individual solar facilities were added to the system at the bus nearest to their geographic locations. Both

wind and solar are modeled with zero marginal cost meaning that they rank at the top of the dispatch stack. Sample dispatch patterns for two units are shown in Figure 5.4. As evinced from these plots, DC-OPF was run independently hour-to-hour. Future work will incorporate generator ramping constraints by running a multiperiod DC-OPF.

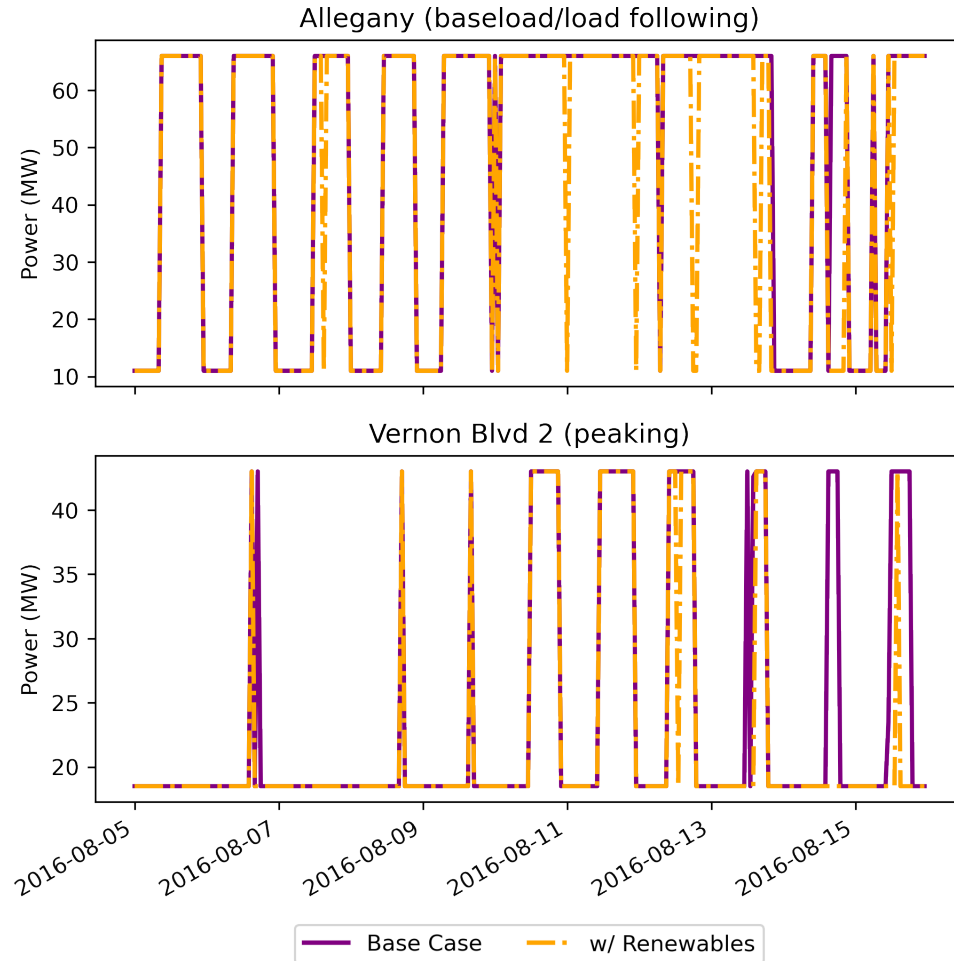


Figure 5.4: Generator profiles under the base case (solid purple lines) and with renewables case (dashed orange lines). The top plot shows the Allegany Generation Station, which acts as a baseload/load following resource remaining on most of the time, but it turns off more often in the renewables case. The bottom plot shows Vernon Blvd Unit 2, which acts more like a peaking unit but turns on less frequently during the renewables case.

5.2.4 Generator Emissions Estimation

While studies often directly scale emissions using new generation profiles, we take a more nuanced approach that seeks to capture the non-linear relationship between load and EGU emissions. Namely, EGU emissions deviate from expected values during start-up, ramping, and part-load operation [16]. Therefore, we apply a reduced model for predicting NO_x , SO_2 , and CO_2 emission rates developed by Gu et al. [17]. This approach makes use of EPA’s continuous emissions monitoring systems (CEMS) data reported by each EGU greater than 25 MW to build separate machine learning models for each pollutant. The model for NO_x directly estimates the emission rate, but the models for SO_2 and CO_2 predict the heat input and use equations from Appendices D and G to Part 75 of Title 40 of the Code of Federal Regulations to calculate the emission rates of SO_2 and CO_2 , respectively. Note that this approach for SO_2 is only appropriate for gas-fired units and therefore suffices in NYS where no EGUs burn coal. With this step, we have concluded the portion of the method depicted within the dashed black box in the diagram shown in Figure 5.1.

From here, preparing CMAQ-ready emissions files remains a multi-step process but is aided by the Eastern Regional Technical Advisory Committee (ERTAC) EGU Tool [11]. This tool offers an alternative to ICF’s Integrated Planning Model (IPM) for EGU emissions projection and formatting. Here we simply use it to fill values for the remaining criteria pollutants and for data formatting. Explicitly, we run the ERTAC EGU Tool preprocessing step, which produces a `calc_hourly_base.csv` file, update this file with emissions calculated from the machine learning model discussed earlier, perform a quality assurance check that no negative emissions values appear, and run the remain-

ing ERTAC EGU Tool steps. We used the CONUS 16.0 inputs for the ERTAC EGU Tool, which can be obtained from the Mid-Atlantic Regional Air Management Association (MARAMA). ERTAC EGU outputs were then processed using version 4.7 of the Sparse Matrix Operator Kernel Emissions (SMOKE) Modeling System [7]. We note two important caveats with the implementation of SMOKE. First, we reverted back one version from the current release of SMOKE (version 4.8) due to numerous compile time and runtime errors that we believe stemmed from the GNU compilers (version 9.3.1) running alongside older hardware and software on our system. Second, courtesy of NYDEC, we corrected a source-code error in SMOKE that resulted in a one hour time-shift of EGU emissions inventories prepared using the ERTAC EGU Tool. Finally, we have reached the final blue block, CMAQ, shown in Figure 5.1.

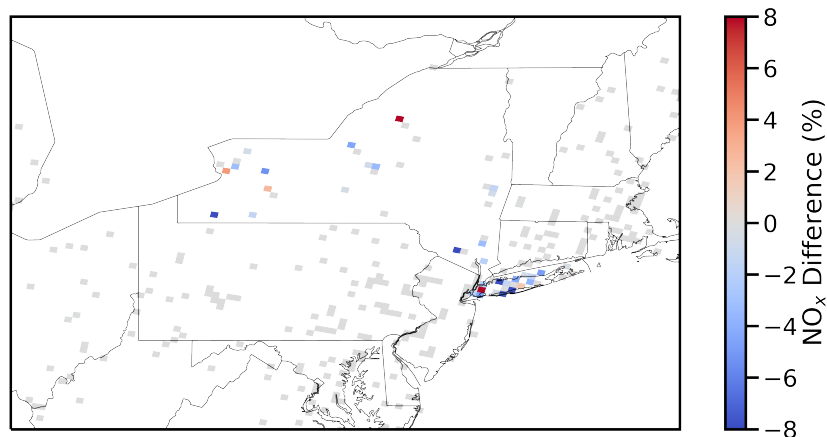
5.2.5 The Community Multiscale Air Quality Model

We run two simulations using version 5.3.3 of the Community Multiscale Air Quality (CMAQ) Model [8] with the same meteorological inputs. The first is a base case with EGU emissions produced using a default run of the ERTAC EGU Tool and the second captures the emissions changes associated with the additional wind and solar installations. The remaining emissions needed to run CMAQ were produced from version 1 of the 2016 EPA emissions modeling platform and were prepared by NYDEC. CMAQ is run for both the 12 km and 4 km domains with 35 layers in the vertical direction for eight days from August 6 - 13, 2016. We have wrapped much of the process shown in Figure 5.1 into the `jeffreysward/cmaqpy` repository available on GitHub.

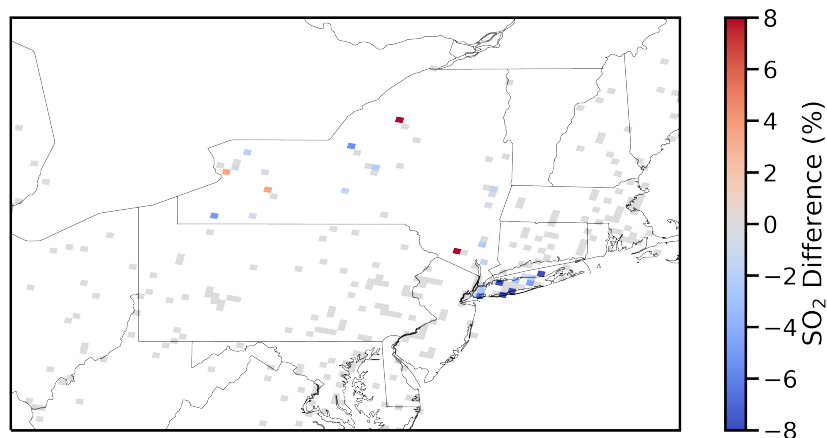
5.3 Results

From the CTM results, we can determine the air quality co-benefits across the Northeastern US associated with NYS renewable energy development goals. For this study, we modeled changes in the NO_x and SO_2 emissions for all major power plants in NYS. Maps showing the percent difference in the emissions of these pollutants are given in Figure 5.5a and 5.5b, respectively. Since CMAQ calculates point source plume rise within the model (“in-line”), the `inlineto2d` CMAQ utility is required to visualize these emissions on the modeling grid. Grid cells colored blue indicate a reduction in point source emissions and vice versa for red cells. Gray cells have the same emissions under both scenarios, and white cells contain no point source emissions. We only modeled changes for generators in NYS, so we expect all grid cells outside NYS to remain gray or white.

Notice that the largest emissions changes occur near New York City (NYC) and on Long Island. The reason for this is two-fold. First, the majority of people in NYS reside in or around NYC, so a numerous fast-response power generation resources are located there. Also, this is where the new offshore wind resources will interconnect with the grid displacing emissions from fossil generators. Emissions change in both directions (increases and decreases) with respect to the base scenario. A modest decrease in NO_x emissions can be seen at most grid points in Figure 5.5a, but two grid cells (i.e., one in NYC and one in Northern NYS) show substantial increases. Similarly, most point sources experience a decrease in average SO_2 emissions in Figure 5.5b, but two grid points (e.g., one in lower Hudson Valley and one in Northern NYS) show substantial increases.



(a)



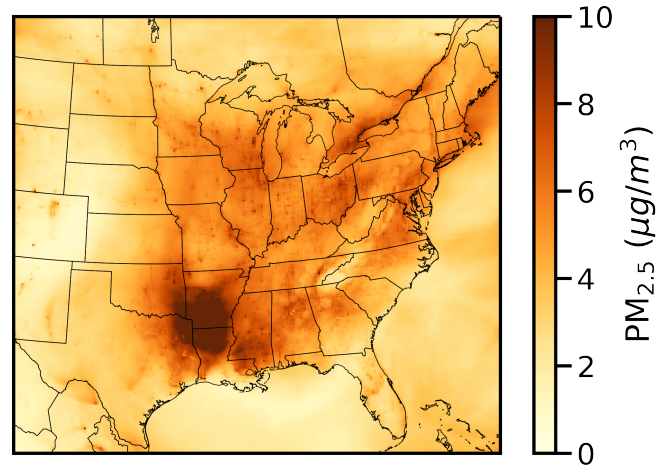
(b)

Figure 5.5: The upper panel shows the mean percent difference in NO_x emissions from the base cast to the renewables case on August 6th, 2016 for the CMAQ-ready in-line point source emissions files prepared by SMOKE. Red cells indicate an increase in emissions, blue cells indicate a decrease in emissions, and gray cells indicate no change in emissions. The lower panel shows the mean percent difference in SO_2 emissions for the same day. Note that we window the domain over NYS as point source emissions should only change here.

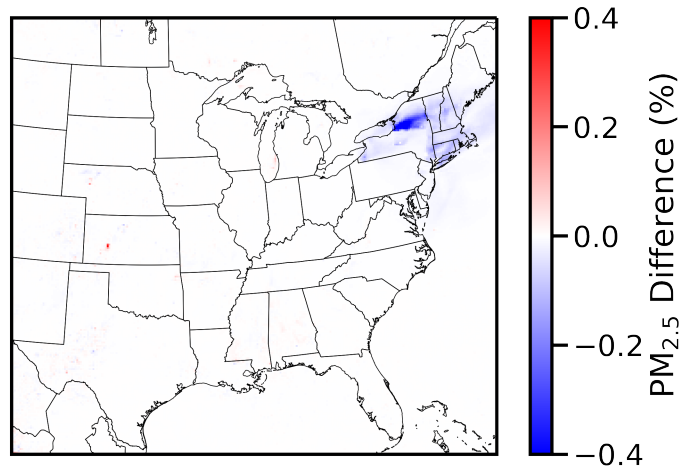
While CMAQ provides concentration fields for myriad pollutants, we present results for $\text{PM}_{2.5}$ (Figure 5.6) and O_3 (Figure 5.8). The coarse, 12 km, domain extends far beyond our focus area of New York State. This is partially because pollution does not recognize state lines, but rather travels wherever the weather dictates, and partially to avoid placing a domain boundary over steep

geographical features where large vertical gradients can cause nonphysical phenomena to appear in the results. The highest $\text{PM}_{2.5}$ concentrations in Figure 5.6a occur in the South Central US, where extractive industry and biomass burning can cause high particulate concentrations. However, heavy rain fell across parts of Southeastern Louisiana from August 11 - 13, dropping up to 30 inches of rain. We see this manifest as reduced $\text{PM}_{2.5}$ concentrations along the Louisiana Gulf Coast up toward cancer alley where numerous refineries often contribute to poor air quality. Concentrations surrounding NYS are as expected with lower concentrations occurring over the Adirondack Mountains and higher concentrations downwind of industry-heavy Toronto, Canada. Looking at the difference in $\text{PM}_{2.5}$ concentrations shown in Figure 5.6b, the largest reductions occur downwind of Watertown, NY with modest reductions visible across much of NYS and New England. This indicates that the additional renewable energy development displaced energy from gas and oil units during this period. However, we also note that the background concentration of $\text{PM}_{2.5}$ is quite low across most of NYS, so the corresponding changes in the absolute pollutant concentrations are small.

Zooming in, the 4 km domain covers much of the Northeast Corridor. $\text{PM}_{2.5}$ concentrations have fallen here over the past couple of decades as the region has worked to phase out coal and heavier fuel oils. The effect of these trends is visible in the mean $\text{PM}_{2.5}$ concentration plot shown in Figure 5.7a, where most concentrations fall within the $2 - 6 \mu\text{g}/\text{m}^{-3}$ range with only isolated points reaching $10 \mu\text{g}/\text{m}^{-3}$. From Figure 5.7b, we observe that the greatest improvements in this area occur downwind of the lower Hudson Valley and Long Island. Some modest increases are also visible highlighting the importance of quantifying the air quality impacts associated with complicated changes in power plant dis-



(a)

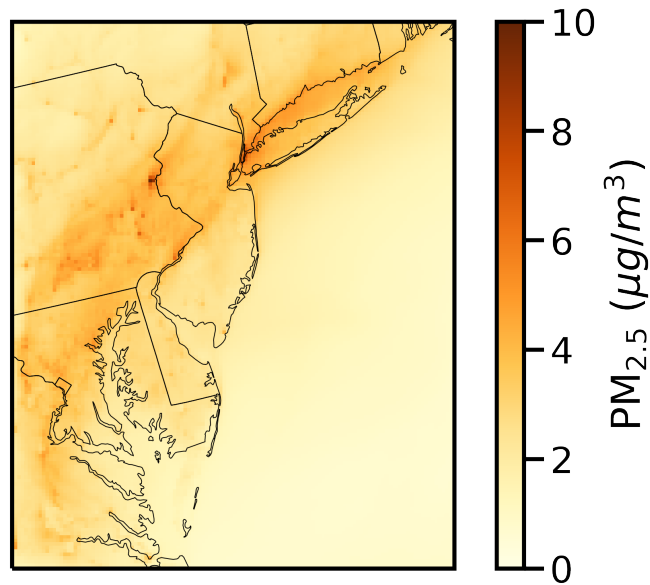


(b)

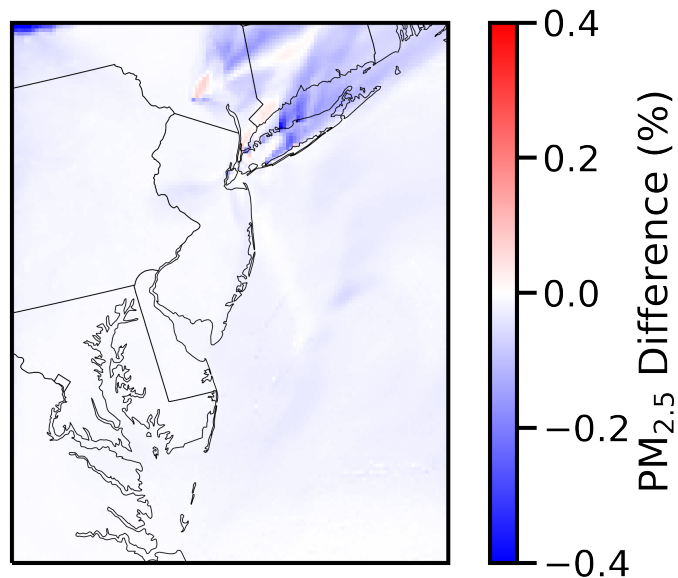
Figure 5.6: The upper panel depicts the mean $PM_{2.5}$ concentration over the full 12 km domain for the entire CMAQ model run. Darker brown colors correspond to higher pollutant concentrations. The lower panel shows the mean absolute difference in $PM_{2.5}$ concentrations for the full model run. Blue corresponds to a decrease in concentration whereas red corresponds to an increase.

patch patterns. In other words, renewable energy development might not mean better air for everyone at all times – even if it improves air quality on average. The relationships are convoluted.

The highest ozone values occur over the Northern Great Plains (as seen in Figure 5.8a) and can be partially attributed to NO_x that comes with heavy sum-



(a)



(b)

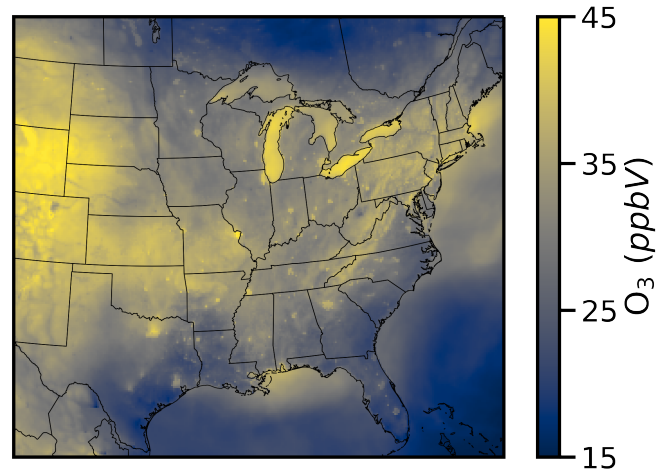
Figure 5.7: The upper panel depicts the mean $PM_{2.5}$ concentration over the 4 km domain for the entire CMAQ model run. Darker brown colors correspond to higher pollutant concentrations. The lower panel shows the mean absolute difference in $PM_{2.5}$ concentrations for the full model run. Blue corresponds to a decrease in concentration whereas red corresponds to an increase.

mer thunderstorm activity. Higher values also occur over the Great Lakes and in industry or metropolitan adjacent coastal regions where precursors are ad-

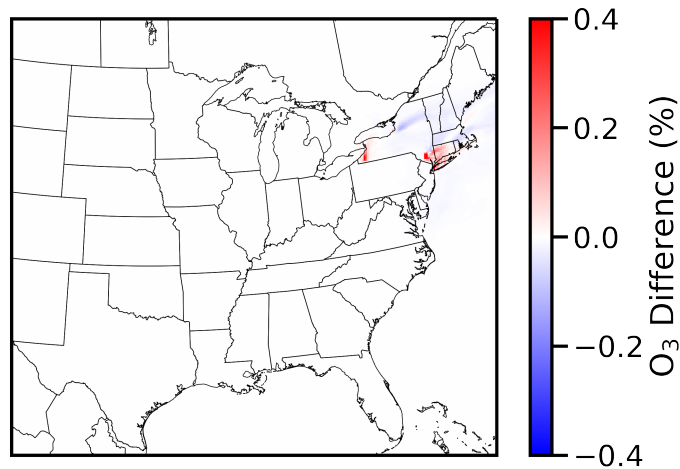
vected offshore into favorable conditions for ozone formation. Changes between the renewable development scenario and the base case show more variation for NO_x than $\text{PM}_{2.5}$, which follows from the complex relationship between NO_x emissions, volatile organic compound (VOC) emissions, and O_3 pollution. Often during the summer in the Northeast where natural vegetation sources contribute to VOC concentrations, a decrease in NO_x emissions causes a decrease in O_3 concentrations. This appears to be the case downwind of Watertown, NY shown in Figure 5.8b.

On the other hand, lower Hudson Valley, NYC, and Long Island all see a substantial increase in O_3 concentrations. There are multiple possible explanations. Less vegetation and high vehicle NO_x emissions could mean that O_3 formation is VOC-limited in these areas. In such a case, reductions in NO_x emissions can cause an increase in O_3 concentrations. This phenomenon is often called the “weekend effect” as it is most often observed on weekends when vehicle NO_x emissions subside. But, it is also possible that the absolute NO_x emissions in some grid cells are driving these changes but are poorly represented in Figure 5.5a because it compares only relative differences at each grid cell.

Results from the higher resolution, 4 km, domain help to elucidate these patterns. Notice that the increase in O_3 concentrations in the lower Hudson Valley shown in Figure 5.9b matches the location with the lowest absolute O_3 concentration in the domain shown in Figure 5.9a. In other words, the renewable energy development scenario has higher ozone concentrations in this area, but since the area has lower baseline concentrations, the increase might not pose any additional issues for air quality compliance. We also note that this episode coincides with the summer peak electrical demand day in NYS, which occurred



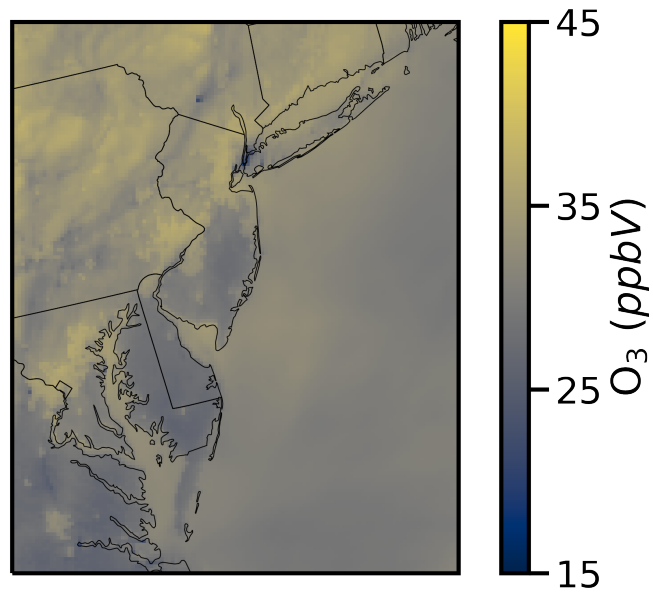
(a)



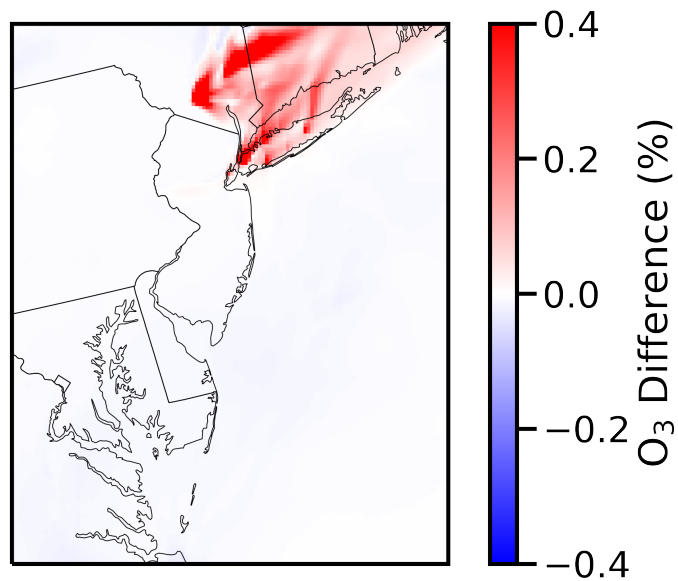
(b)

Figure 5.8: The upper panel depicts the mean O_3 concentration over the full 12 km domain for the entire CMAQ model run spanning August 6 - 13, 2016. Brighter yellow colors correspond to higher pollutant concentrations. The lower panel shows the mean absolute difference in O_3 concentrations for the full model run. Blue corresponds to a decrease in concentration whereas red corresponds to an increase.

on August 11, 2016 [1]. Summer peaks in NYS are driven primarily by cooling load – meaning that NYS experienced unusually hot and humid conditions during this period. Both renewable energy performance and air quality would have been affected by these conditions to some degree, so future work should characterize impacts over several ozone seasons before reporting the full air quality



(a)



(b)

Figure 5.9: The upper panel depicts the mean O_3 concentration over the 4 km domain for the entire CMAQ model run spanning August 6 - 13, 2016. Brighter yellow colors correspond to higher pollutant concentrations. The lower panel shows the mean absolute difference in O_3 concentrations for the full model run. Blue corresponds to a decrease in concentration whereas red corresponds to an increase.

co-benefits of a renewable energy development scenario.

5.4 Discussion

While we have taken steps to explain and simplify methodological steps in the OneMet framework, it remains a complex cross-disciplinary modeling setup. Several components comprising OneMet are flexible enough to be directly swapped with a similar method of a researcher's choosing. Here, we discuss the broader applicability of OneMet along with any caveats.

5.4.1 Power Plant Dispatch and Emissions Modeling

Of course, the exact method by which emission changes are modeled will decide the resultant the air quality co-benefits produced by the framework. We chose to estimate air quality changes using a publicly available reduced-form machine learning model built using EPA CAMD data and an open-source representation of the NYS power system. Such a combination balances accuracy with transparency – two ingredients essential in tools aiming to quicken the pace of electricity sector decarbonization. While several private-sector companies have created impressive dispatch models, none of these companies can hope to contribute to electricity sector planning across every corner of the system. Therefore, planners should exercise caution in using proprietary models, and better yet, incentives for developing open-source energy planning data tools and resources should be offered at the federal, state, and local levels. During the winter storm in Texas in 2021, we witnessed again that existing markets and regu-

lations cannot adequately incentivize power system reliability in the face of low probability events. As far back as the California Electricity Crisis, lower than expected renewable resource availability played a role. With extreme weather and the number of inverter-based resources connecting to the grid increasing, we cannot even afford subpar forecasts – let alone another Enron. The best antidote to this remains clear modeling assumptions backed up by a reliable data network.

Two important modifications would improve the accuracy of the dispatch and emissions models for use in any arbitrary power system. First, the dispatch model employed here considers each time period independently. Therefore, no interhourly information is taken into account meaning that we cannot capture generator ramping capabilities and the emissions impacts associated with different ramping patterns. Future work lies in expanding the economic dispatch model to a multi-period multi-area paradigm. Moving on to emissions estimates, recall that we predicted SO_2 by first estimating a generator's heat input before calculating SO_2 based on the equation provided in federal regulations. However, this equation for the SO_2 emission rate from Appendix D of Part 75 to 40 CFR applies only to gas-fired units. This calculation is appropriate in NYS where all coal units have been retired, but a separate model that directly estimates the SO_2 emission rate would be necessary for systems where coal units still operate regularly. As many states are interested in phasing out coal in the coming years, characterizing the amount of coal-fired electricity that may be displaced by renewables would almost certainly produce high air-quality co-benefits of renewables when compared with the more complicated benefits associated with changes in ozone precursors.

5.4.2 WRF Meteorology for Air Quality

Here, we adopted a WRF setup from a successful offshore wind simulation study. But, different WRF setups perform differently in different regions and for different applications as we discuss in detail in Chapter 4. Given enough computational resources, we could devise a fitness function considering WRF wind and solar forecast errors as well CMAQ air quality concentration errors and find a better setup for WRF. A simpler first step could add observational nudging to the WRF modeling process and determine which initial and boundary condition datasets result in the best performance. Finally, as we touch on in Chapter 3, the present and future of forecasting are probabilistic. As such, the OneMet framework should eventually evolve to ingest probabilistic wind and solar profiles and produce probabilistic air quality concentrations. Each of these additions adds a substantial computational burden, so sensitivity analyses should explore if additional complexity results in a commensurate increase in understanding.

5.5 Conclusion

We developed and demonstrated an open-source framework – which we call OneMet – for quantifying the air quality co-benefits associated with wind and solar development. Related studies within the existing literature conduct separate analyses to determine renewable energy generation profiles and pollutant concentrations over the same region. But, both depend primarily on underlying meteorological conditions. Therefore, our framework derives wind and solar estimates from the same WRF output that drives CMAQ’s chemical trans-

port model. Such internal consistency reduces potential sources of bias thereby making it simpler to characterize and correct those that remain.

To demonstrate OneMet's performance, we quantified co-benefits associated with NYS's near-term renewable energy development policy goals. Specifically, we modeled an additional 4192 MW of offshore wind and 2700 MW of solar for 8 days in August 2016. We showed reductions in PM_{2.5} concentrations up to 0.5% in parts of NYS with more modest improvements across New England. Changes in O₃ were mixed with modest decreases in some Upstate NYS regions, but larger increases near 0.5% in the lower Hudson Valley, NYC, and Long Island. We want to stress that these results represent an 8-day snapshot with uncommon meteorological conditions and that fully characterizing the co-benefits of renewable energy will require a substantially longer simulation period. However, the main focus of this work was to present the integrated OneMet framework, so we leave additional simulations for future study.

Several additions would improve the OneMet framework. Critically, future work will expand the dispatch model to capture multi-period multi-area interactions. This will elucidate necessary ramping requirements and their associated air quality implications. Also, it will make it possible to study the impact of policy – or indeed a lack thereof – in neighboring electricity markets. Additional work should also better characterize the uncertainty in air-quality co-benefits. In other words, how much do WRF and CMAQ modeling assumptions affect the overall result? Therefore, future work seeks to transform this framework into one that can provide probabilistic pollutant concentrations informed by probabilistic wind and solar forecasts.

Finally, we advocate for open-source renewable energy planning tools and

data where modeling assumptions are exposed for discussion, verification, and sensitivity analyses. Only with such an approach can we hope to collaborate across sectors to achieve a zero-carbon power sector in the future.

References

- [1] 2017 Load & Capacity Data "Gold Book". Technical report, New York Independent System Operator, 2017.
- [2] New York State Offshore Wind – 2020 Offshore Wind Solicitation Awards. Technical report, New York State Energy Research and Development Authority, 2022.
- [3] David Abel, Tracey Holloway, Monica Harkey, Arber Rrushaj, Greg Brinkman, Phillip Duran, Mark Janssen, and Paul Denholm. Potential air quality benefits from increased solar photovoltaic electricity generation in the eastern united states. *Atmospheric Environment*, 175:65–74, 2 2018.
- [4] David W Abel, Tracey Holloway, Javier Martínez-Santos, Monica Harkey, Madankui Tao, Cassandra Kubes, and Sara Hayes. Air quality-related health benefits of energy efficiency in the united states. 2019.
- [5] Agustin Alvarez-Herranz, Daniel Balsalobre-Lorente, Muhammad Shahbaz, and José María Cantos. Energy innovation and renewable energy consumption in the correction of air pollution levels. *Energy Policy*, 105:386–397, 6 2017.
- [6] Jonathan J Buonocore, Patrick Luckow, Gregory Norris, John D Spengler, Bruce Biewald, Jeremy Fisher, and Jonathan I Levy. Health and climate

benefits of different energy-efficiency and renewable energy choices. *NATURE CLIMATE CHANGE* 1, 6, 2016.

- [7] Community Modeling and Analysis System Center. Sparse Matrix Operator Kernel Emissions (SMOKE) Modeling System Version 4.7, 2019.
- [8] Community Modeling and Analysis System Center. Community Multi-scale Air Quality (CMAQ) Model Version 5.3.3, 2021.
- [9] Cian Desmond, Jimmy Murphy, Lindert Blonk, and Wouter Haans. Description of an 8 mw reference wind turbine. *Journal of Physics: Conference Series*, 753:092013, 2016.
- [10] Emil G Dimanchev, Sergey Paltsev, Mei Yuan, Daniel Rothenberg, Christopher W Tessum, Julian D Marshall, and Noelle Selin. Health co-benefits of sub-national renewable energy policy in the us. 2019.
- [11] Eastern Regional Technical Advisory Committee. ERTAC EGU Projection Tool Version 2.1.1, 2018.
- [12] Anna C Fitch, Joseph B Olson, Julie K Lundquist, Jimy Dudhia, Alok K Gupta, John Michalakes, and Idar Barstad. Local and Mesoscale Impacts of Wind Farms as Parameterized in a Mesoscale NWP Model. *Monthly Weather Review*, 140:3017–3038, 2012.
- [13] Ciaran L. Gallagher and Tracey Holloway. Integrating air quality and public health benefits in u.s. decarbonization strategies. *Frontiers in Public Health*, 8:520, 11 2020.
- [14] Ana R. Gamarra, Yolanda Lechón, Marta G. Vivanco, Mark Richard Theobald, Carmen Lago, Eugenio Sánchez, José Luis Santiago, Juan Luis

Garrido, Fernando Martín, Victoria Gil, and Alejandro Rodríguez-Sánchez. Avoided mortality associated with improved air quality from an increase in renewable energy in the spanish transport sector: Use of biofuels and the adoption of the electric car. *Atmosphere* 2021, Vol. 12, Page 1603, 12:1603, 11 2021.

- [15] Martin Geissdoerfer, Paulo Savaget, Nancy M.P. Bocken, and Erik Jan Hultink. The circular economy – a new sustainability paradigm? *Journal of Cleaner Production*, 143:757–768, 2 2017.
- [16] Terje Gjengedal. Emission constrained unit-commitment (ecuc). *IEEE Transactions on Energy Conversion*, 11:132–138, 1996.
- [17] Jiajun Gu, Jeffrey A. Sward, Huilin Zhang, Ye Jiang, and K. Max Zhang. Predicting power plant emissions using machine learning.
- [18] Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, Adrian Simmons, Cornel Soci, Saleh Abdalla, Xavier Abellan, Gianpaolo Balsamo, Peter Bechtold, Gionata Biavati, Jean Bidlot, Massimo Bonavita, Giovanna De Chiara, Per Dahlgren, Dick Dee, Michail Diamantakis, Rossana Dragani, Johannes Flemming, Richard Forbes, Manuel Fuentes, Alan Geer, Leo Haimberger, Sean Healy, Robin J. Hogan, Elías Hólm, Marta Janisková, Sarah Keeley, Patrick Laloyaux, Philippe Lopez, Cristina Lupu, Gabor Radnoti, Patricia de Rosnay, Iryna Rozum, Freja Vamborg, Sebastien Villaume, and Jean Noël Thépaut. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, jul 2020.
- [19] William F. Holmgren, Clifford W. Hansen, and Mark A. Mikofski. pvl

- python: a python package for modeling solar energy systems. *Journal of Open Source Software*, 3:884, 9 2018.
- [20] Giorgos Kallis, Christian Kerschner, and Joan Martinez-Alier. The economics of degrowth. *Ecological Economics*, 84:172–180, 12 2012.
- [21] Matheus Koengkan, José Alberto Fuinhas, and Nuno Silva. Exploring the capacity of renewable energy consumption to reduce outdoor air pollution death rate in latin america and the caribbean region. *Environmental Science and Pollution Research*, 28:1656–1674, 1 2021.
- [22] M. Leutbecher and T. N. Palmer. Ensemble forecasting. *Journal of Computational Physics*, 227:3515–3539, 3 2008.
- [23] M. Vivienne Liu, Bo Yuan, Zongjie Wang, Jeffrey A. Sward, K. Max Zhang, and C. Lindsay Anderson. An open source representation for the nys electric grid to support power grid and market transition studies, 2021.
- [24] Mandana Mazaheri, Yvonne Scorgie, Richard A. Broome, Geoffrey G. Morgan, Bin Jalaludin, and Matthew L. Riley. Monetising air pollution benefits of clean energy requires locally specific information. *Energies 2021, Vol. 14, Page 7622*, 14:7622, 11 2021.
- [25] Mike Optis, Andrew Kumler, Joseph Brodie, and Travis Miles. Quantifying sensitivity in numerical weather prediction-modeled offshore wind speeds through an ensemble modeling approach. *Wind Energy*, 24:957–973, 2021.
- [26] Shuai Pan, Lewis M. Fulton, Anirban Roy, Jia Jung, Yunsoo Choi, and H. Oliver Gao. Shared use of electric autonomous vehicles: Air quality and health impacts of future mobility in the united states. *Renewable and Sustainable Energy Reviews*, 149, 10 2021.

- [27] Wei Peng, Jiahai Yuan, Yu Zhao, Meiyun Lin, Qiang Zhang, David G. Victor, and Denise L. Mauzerall. Air quality and climate benefits of long-distance electricity transmission in china. *Environmental Research Letters*, 12:064012, 6 2017.
- [28] Steven D. Plachinski, Tracey Holloway, Paul J. Meier, Gregory F. Nemet, Arber Rrushaj, Jacob T. Oberman, Phillip L. Duran, and Caitlin L. Voigt. Quantifying the emissions and air quality co-benefits of lower-carbon electricity production. *Atmospheric Environment*, 94:180–191, 9 2014.
- [29] Ghazal Razeghi, Marc Carreras-Sospedra, Tim Brown, Jack Brouwer, Donald Dabdub, and Scott Samuelsen. Episodic air quality impacts of plug-in electric vehicles. *Atmospheric Environment*, 137:90–100, 7 2016.
- [30] Dorothea Schmid, Pinar Korkmaz, Markus Blesl, Ulrich Fahl, and Rainer Friedrich. Analyzing transformation pathways to a sustainable european energy system—internalization of health damage costs caused by air pollution. *Energy Strategy Reviews*, 26:100417, 11 2019.
- [31] Servaas Storm. Capitalism and climate change: Can the invisible hand adjust the natural thermostat? *Development and Change*, 40:1011–1038, 11 2009.
- [32] Jeffrey A. Sward, Roberta S. Nilson, Venkatesh V. Katkar, Richard C. Stedman, David L. Kay, Jennifer E. Ifft, and K. Max Zhang. Integrating social considerations in multicriteria decision analysis for utility-scale solar photovoltaic siting. *Applied Energy*, 288:116543, 4 2021.
- [33] Jeffrey A. Sward, Jackson Siff, Jiajun Gu, and K. Max Zhang. Strategic

planning for utility-scale solar photovoltaic development – historical peak events revisited. *Applied Energy*, 250:1292–1301, 9 2019.

- [34] Stéphane Vannitsem, Daniel S. Wilks, and Jakob W. Messner, editors. *Statistical Postprocessing of Ensemble Forecasts*. Elsevier, 1st edition, 2019.
- [35] Yongfeng Zhu, Zilong Wang, Jie Yang, and Lingling Zhu. Does renewable energy technological innovation control china’s air pollution? a spatial analysis. *Journal of Cleaner Production*, 250, 3 2020.
- [36] Ray Daniel Zimmerman, Carlos Edmundo Murillo-Sánchez, and Robert John Thomas. Matpower: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on Power Systems*, 26:12–19, 2 2011.

5.6 Supplemental Information

5.6.1 WRF Namelist Parameters

Constant WRF namelist parameters are provided in Table 5.1. Any parameters not listed in that table were left at their WRFv4.0 default values.

Table 5.1: Constant Namelist Options

interval_seconds	10800
history_interval	60, 60
time_step	45
max_dom	2
e_we	472, 226
e_sn	312, 253
e_vert	35, 35
num_metgrid_levels	38
num_metgrid_soil_levels	4
dx	12000, 4000
dy	12000, 4000
swint_opt	1
radt	15, 5
bldt	0, 0
cudt	0, 0
surface_input_source	1
num_soil_layers	4
num_land_cat	21
damp_opt	0

CHAPTER 6

MAJOR CONTRIBUTIONS AND FUTURE WORK

Broadly speaking, in this thesis, I have taken two well-known community models, assessed and customized them for renewable energy development in the Northeastern United States, and used them in a novel framework for investigating the power system caveats and air quality benefits of evolving into a decarbonized society. But with large modeling effort, perfect is the enemy of the good. As such, this chapter aims to concisely describe my contributions from the preceding chapters, revisit assumptions, and conceptualize future paths for the OneMet framework.

In Chapter 2, I conducted a spatial analysis to aid in strategic long-term planning for utility-scale solar development. Through this analysis, I examined land in New York State on a parcel by parcel basis to determine whether it might realistically host a solar farm. From this study, it was clear that the most attractive land from a solar developer's standpoint was located Upstate. Indeed subsequent work by Venkatesh V. Katkar showed that farmland accounted for the lion's share of land where solar would be the easiest to develop. Considering three different scenarios describing PV build-out across the state, I determined that these new resources could substantially offset peak demand during summer high electricity demand days. PV offers less value during the winter as the sun is down during the winter peak, and frequent cloudiness substantially reduces capacity factors. However, on sunny winter days, which have lower midday demand than the summer, PV can substantially increase system-wide flexibility requirements. As such, a mechanism to shift load to the middle of the day during the winter or store this infrequent winter sun would mitigate

operational challenges associated with utility scale-PV development.

Of course, as much as researchers and planners attempt to map out expected future development, energy systems have and always been part of the fabric of local communities. Historically, these groups have focused primarily on salient technical, economic, or legal issues while blatantly overlooking the inhabitants of these communities. This omission has led to poor support or outright opposition stalling many otherwise viable solar projects. Therefore, a great deal of future work should develop guidelines for a more community-centric approach to spatial modeling algorithms. Furthermore, traditional "decide-announce-defend" development models inspire opposition, particularly within well-resourced communities. As such, more interactive and responsive siting processes should be devised along with associated metrics for tracking this goal. While utility-scale solar farms may not pose the existential health threats that fossil fuel plants or waste incinerators do, this should not suggest that their siting is without equity implications. Therefore, future work should continue to question the environmental justice implications associated with all energy development regardless of how clean and green – actually or ostensibly – the development claims to be.

In Chapter 3, I assessed the performance of multiple physical setups of the Weather Research and Forecasting (WRF) model across different seasons using data from LiDAR buoys at multiple locations offshore for the first time. Performance varies substantially not only among the different setups but between the two buoys for the same setup. Such spatial nonuniformity in biases calls into question the accuracy of offshore wind probabilistic forecasts produced using common ensemble postprocessing techniques as it remains impossible to truly

characterize the uncertainty in offshore wind forecasts.

Given the Biden Administration’s goal to develop 30 GW of offshore wind by 2030, future work must develop tools to characterize this resource – and its associated variability – in greater detail. Local ensemble forecasts could inform ISOs, DSOs, and smart cities about likely offshore wind generation levels, which offers the closest thing we have to baseload power in the variable renewable energy world. Ensemble postprocessing techniques would improve these predictions but require substantial observational data. With enough wind power scattered along the east coast to rival New York State’s summer peak load, a network of instruments – ideally LiDAR – should supply data in near real-time. While the private sector will, be instrumental in carrying out the physical construction of these wind farms, the data cannot remain locked within the private sector. These data must flow into government planning models and proprietary forecaster codes alike. Only with such open data practices can we hope to supply the necessary tools for operating the power system under a majority variable renewables paradigm.

In Chapter 4, I presented a methodology for optimizing the WRF model setup in any region for any application using a genetic algorithm. I demonstrate this approach, which I called OptWRF, for wind and solar resources in the Northeastern U.S. As computer resources become cheaper and more readily available, this method could be applied by governments or non-profit organizations who want to conduct wind and solar resource assessments or produce their own operational forecasts. Today, such a task requires iterations of arduous WRF model tuning.

Numerous improvements could make OptWRF more effective and acces-

sible. While the code base is currently open-source a substantial amount of additional work would be required to make the code base easily portable to a variety of systems. Improving the modularity would also make it easier to plug and play with different WRF domains, boundary condition data sources, and fitness functions. Further work could include optimization of a wider set of parameters including parameters within each physics scheme. Through this, it would be possible to characterize the sensitivity of the wind and solar estimates to each of these wider parameters – something that we’ve already done for the major parameterization options. Finally, additional effort could be spent making OptWRF run more efficiently by changing the behavior or the structure of the underlying genetic algorithm.

Finally, in Chapter 5, I construct an internally consistent framework for quantifying the air quality co-benefits associated with specific wind and solar development scenarios. This framework – dubbed OneMet – is designed to help researchers assess which scenarios provide the greatest reduction in air pollutant concentrations thereby building a more compelling case for additional renewable energy development. Collaborative work on an open-source power system representation for New York State by Liu et al. [2] and emissions modeling by Gu et al. [1] represent critical components of OneMet. OneMet offers internal consistency by deriving wind and solar generation profiles from the same WRF downscaling that provides inputs to CMAQ. No other framework determines the air quality co-benefits of renewable energy using such an internally consistent open-source method.

Several important changes could improve OneMet’s accuracy and impact. First, improvements in both the dispatch model and the emissions model would

provide better accuracy in estimating the emissions changes caused by new renewable energy developments. Notably, the dispatch model currently runs a simple DC optimal power flow algorithm, which cannot capture hour-to-hour dependencies including generator ramping constraints. Second, SMOKE and CMAQ rely on a roundabout set of scripts, which could be more cleanly encapsulated within the existing `CMAQPY` python package that I created alongside this project. Finally, a great deal of work – at least a dissertation or two – could focus on adapting and expanding this framework to quantify the health co-benefits of renewable energy development, particularly in disadvantaged communities. The test simulation included in Chapter 5 indicates that renewable energy development will not result in improved air quality in all locations at all times. As with all development, trade-offs exist. Legislation in several states, New York included, already requires a specific percentage of clean energy spending to realize benefits in disadvantaged communities, but the tools required to characterize and achieve this currently do not exist. OneMet represents a solid foundation.

References

- [1] Jiajun Gu, Jeffrey A. Sward, Huilin Zhang, Ye Jiang, and K. Max Zhang. Predicting power plant emissions using machine learning.
- [2] M. Vivienne Liu, Bo Yuan, Zongjie Wang, Jeffrey A. Sward, K. Max Zhang, and C. Lindsay Anderson. An open source representation for the nys electric grid to support power grid and market transition studies, 2021.