



Genetic algorithm selection of the weather research and forecasting model physics to support wind and solar energy integration

J.A. Sward^a, T.R. Ault^b, K.M. Zhang^{a,*}

^a Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, NY, 14853, United States

^b Earth and Atmospheric Sciences, Cornell University, Ithaca, NY, 14853, United States

ARTICLE INFO

Article history:

Received 24 December 2021

Received in revised form

10 May 2022

Accepted 22 May 2022

Available online 26 May 2022

Keywords:

Solar energy

Wind energy

WRF model

Genetic algorithm

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 that 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.

© 2022 Elsevier Ltd. All rights reserved.

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 [1]. This increase comprises nearly 80% of all growth in electricity production, dwarfing that of fossil fuels. 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 (NWPs) are becoming indispensable ingredients in the cocktail of electricity systems operations and planning [2–4].

One NWP tool in particular—the Weather Research and Forecasting (WRF) model [5]—is the favored tool for wind and solar across scales. For example, WRF underlies operational forecasting

systems tuned for both wind [6] and solar [7], provided the forecast system to create NREL's wind tool kit [8], and helped uncover the costs of uncoordinated wind farm development [9]. 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 even at 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 [10]. 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.

* Corresponding author.

E-mail address: kz33@cornell.edu (K.M. Zhang).

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. Refs. [11–18]). Stergiou et al. even use multi-criteria decision analysis to aid in assessing sensitivity to model physics [19]. 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) [20]. 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” [21]. 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 configuration 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 configurations, and the “fittest” configurations are then hybridized with each other to produce a new generations of configurations. 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 [22,23]. 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 [24]. In a brief study, Oana and Spataru are the only researchers who have used a GA for initial WRF parameter selection [25]. They report promising early results for humidity and temperature 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 configuration 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 configuration would immediately halve the computational cost. Among those concerned with wind energy, researchers have introduced and benchmarked a wind farm parameterization [26–29] and assessed WRF’s sensitivity to different spatial resolutions, different boundary condition data sources, different PBL schemes [30–32], parameters within PBL schemes, or some combination of these [33–36]. Others assessed how well a single WRF physics option set performs before conducting a wind resource analysis [37]. Some of these analyses culminated in public meteorological data products to aid in wind power integration (e.g., NREL’s WIND Toolkit [8]).

Many studies also customize WRF for solar energy. WRF-SOLAR provides a WRF physics option set for solar energy forecasting [7], and urban WRF-SOLAR extends this by adding an urban canopy model and building energy model [38]. 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 [2,39] and another evaluated the performance of different shortwave parameterization schemes in predicting GHI [40]. 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 the sets recommended as a starting point by NCAR in the WRF User’s Guide.

The remainder of this paper is organized as follows: in Section 2 will discuss our methodology for developing and applying the genetic algorithm to WRF model setup, in Section 4 we discuss our results with an emphasis on the influence of physics options on WRF output and applying this method to operationalize the WRF model in an arbitrary location for an arbitrary application, and finally, Section 5 provides a conclusion and final recommendations for applying the OptWRF methodology.

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 2.1, as each setup represents an individual in the OptWRF GA. The following explanations in Section 2.2 provide specifics of the GA algorithm operators and control parameters that produce better WRF forecasts through successive generations.

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 [5]. Microphysics, radiation (separate schemes handle longwave and shortwave radiation), planetary boundary layer (PBL), land surface, and convection (at a horizontal grid spacing above ~4 km [5]) 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.

2.1.1. Model parameterizations

During WRF model initialization, choices for parameterizing various unresolvable 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 [41]. See Section Appendix A.3.1 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 [42]. 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 [22,23], 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.

2.1.2. WRF model domain, boundary condition data, and simulation length

The focus of this work is upon the parameterizations that govern how physical processes within WRF are represented. Therefore, the majority of the WRF namelist options – particularly those governing the model domain and dynamics – are kept constant in each configuration as are the boundary conditions. The modeling domain is centered over the Northeast United States and has a 12 km 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 A.1 in Appendix A.1 gives the exact eta levels). The time step used for all simulations was set to 45 s 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 Appendix A.2. ERA-Interim data created by the European Centre for Medium-Range Weather Forecasts (ECMWF) was used as boundary condition data for all downscaling [43]. Each simulation runs for 24 h and is initialized at 00:00 UTC on a random day in 2011. We analyze all the data from each simulation rather than treating the first several hours as a spinup period.

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. Fig. 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.

2.2.1. 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

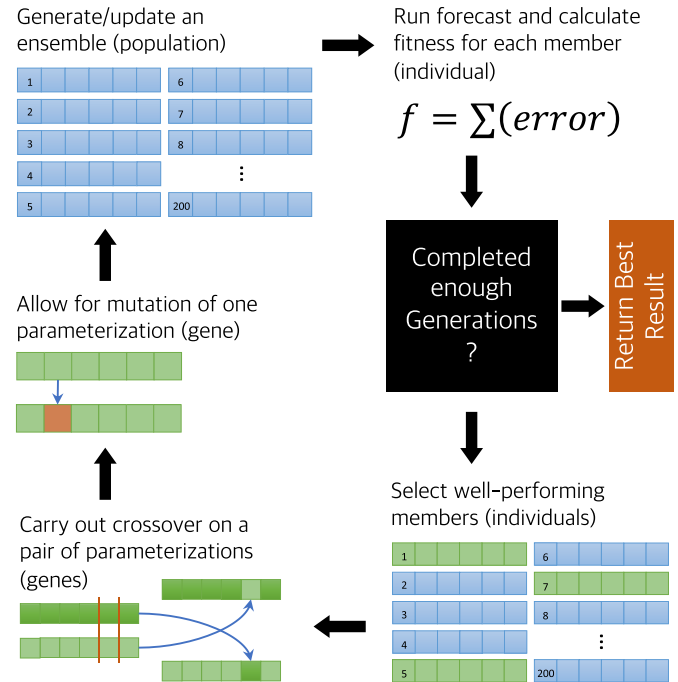


Fig. 1. Schematic diagram of the genetic algorithm used for determining a near optimal set of physics parameters for the WRF model based upon an arbitrary application-specific fitness function. The number of generations (cycles through the diagram) can be tuned by the user based upon size of the ensemble, available time, and computational resources.

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

2.2.2. 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) to judge model performance based on wind speed. WPD captures the amount of power that could be extracted from the wind per square meter of wind turbine sweep area, $WPD = \frac{P_{wind}}{A} = \frac{1}{2} \rho v^3$, where P_{wind} is the power extracted by a wind turbine, A is the rotor sweep area of the turbine, ρ is the air density, and v is the 100 m wind speed. Performance based on solar energy was judged by model-calculated GHI. The fitness function shown in Equation (1) calculates the accumulated errors between a WRF model configuration and the ERA5 reanalysis dataset [44] in GHI (denoted by the two terms within the first set of absolute value brackets) and WPD (denoted by the two terms

within the second set of absolute value brackets) across each hourly time step, t , and across all model grid cells, g , as shown below.

$$fitness = \sum_t \sum_g \left(\left(D_{frac} |GHI_{WRF} - GHI_{ERA5}| + C_{wind} \left| \left(\frac{P_{wind}}{A} \right)_{WRF} - \left(\frac{P_{wind}}{A} \right)_{ERA5} \right| \right) \right)_g \quad (1)$$

Since GHI varies substantially across seasons, we multiplied the GHI errors by a day length fraction $D_{frac} = D/D_L$, where D is the day length calculated using the algorithm provided in [Appendix A.5](#) [70], and D_L corresponds to the longest day of the year. While WPD and GHI both have units of $W m^{-2}$, a scalar correction factor, $C_{wind} = (E_{GHI}/D_{frac})/E_{WPD}$, was employed to ensure that accumulated errors in each quantity during the year 2011 were the same. C_{wind} was derived by running one simulation for each month of 2011 using the same parameterization schemes, where E_{GHI} corresponds to the mean daily GHI error, E_{WPD} to the mean daily WPD error, and D_{frac} to the mean daily day length fraction. 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. Note that this fitness function can change to accommodate a variety of different applications or tuned to show improvement over an existing configuration. Namely, a user might choose to normalize GHI and WPD errors with respect to those produced by an existing model setup, leading the fitness function to represent the improvement (or degradation) in performance with respect to the existing setup.

2.2.3. Population initialization and tournament selection

A population within the GA is composed of WRF model configurations 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 [Appendix A](#) 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 2011. 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 between 50 and 200 members, whose performance will be discussed in Section 4.

Following population initialization, the fitness of each member of the population was calculated by running each WRF model configuration. All the configurations in a generation were submitted to a cluster simultaneously using Python's concurrent futures module. Each configuration was allocated a maximum of 6 h on eight cores for a maximum of 48 core hours per simulation – the vast majority of configurations finished in that amount of time.

Those configurations that failed to finish within the allotted 6 h were terminated and assigned a large fitness value. Likewise, those

configurations 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 error, so lower fitness values are desirable. In this way, the algorithm discovers and records which physics scheme combinations produce errors.

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.

2.2.4. 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.

2.2.5. The mutation operator

After the offspring population has been filled, mutation provides a mechanism to introduce additional genes into individuals that their parent configurations 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 configuration. Researchers previously delved into both the best operators to carry out mutation [46] and the optimal mutation probability [47]. After mutation, the final version of the offspring population is complete. The cycle (in [Fig. 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 configurations extracted from the final population.

Although OptWRF cannot guarantee that every member in the

Table 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%

offspring population will result in a viable WRF simulation, elitism ensures that the top third of the population is transferred into the next generation. This, in combination with the disproportionately poor fitness value that failed members receive leads OptWRF to filter failed members out over time.

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 3.3). The top five performing physics option sets are shown in Table 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 1.

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 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 (Fig. 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.

3.2. Wind and solar trends

Simulations are judged based upon the fitness function (Equation (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 [48]. 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 (Fig. 3 a)–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.

The same method was applied to produce WPD plots for both the best-performing WRF simulation and the ERA5 reanalysis (Fig. 3 c)–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 (Fig. 3 e)–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 A1 in the Supplemental Information.

3.3. Effects of physics parameterizations

In order to determine the absolute effect that each physics parameterization scheme has upon the fitness value, we post-processed the OptWRF results using a random forest regressor [49]. 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

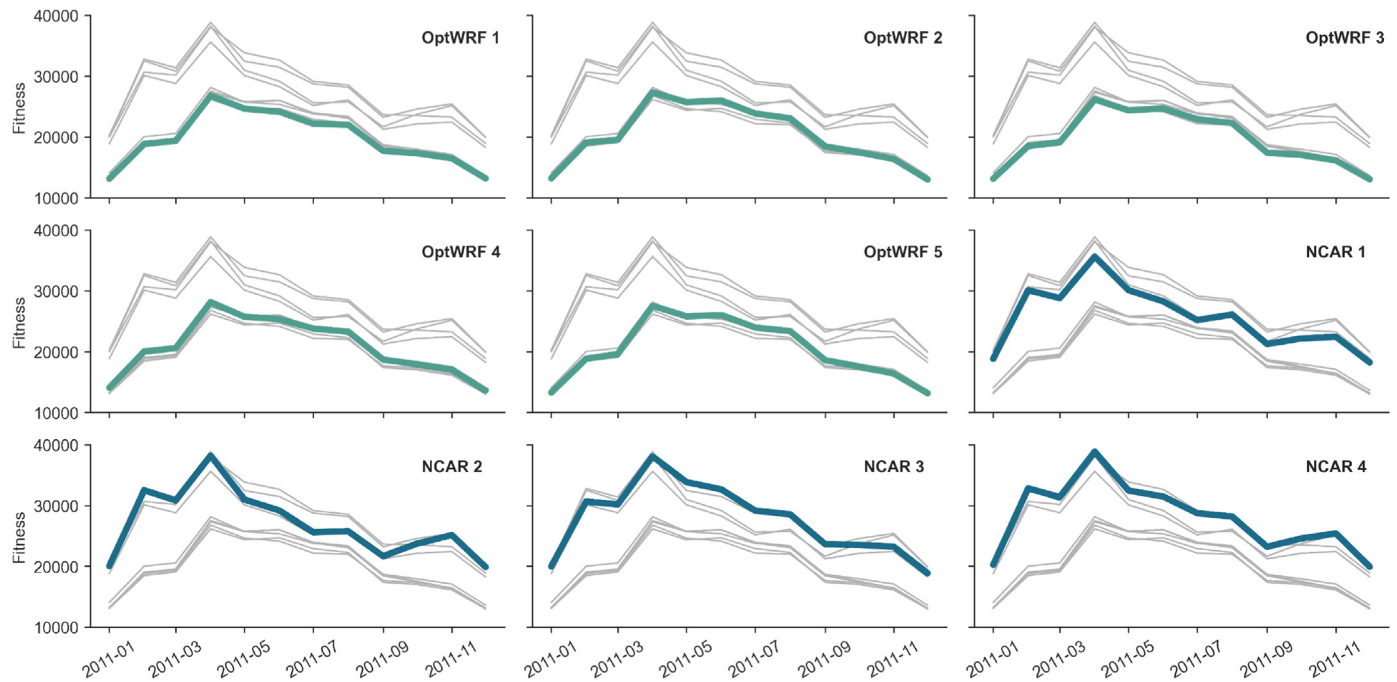


Fig. 2. Plot matrix with each of nine panels highlighting the monthly mean fitness of one physics option set and the results from the other eight plotted in grey. Physics option sets selected by OptWRF are highlighted in teal and appear in the first five panels; whereas physics option sets recommended 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.

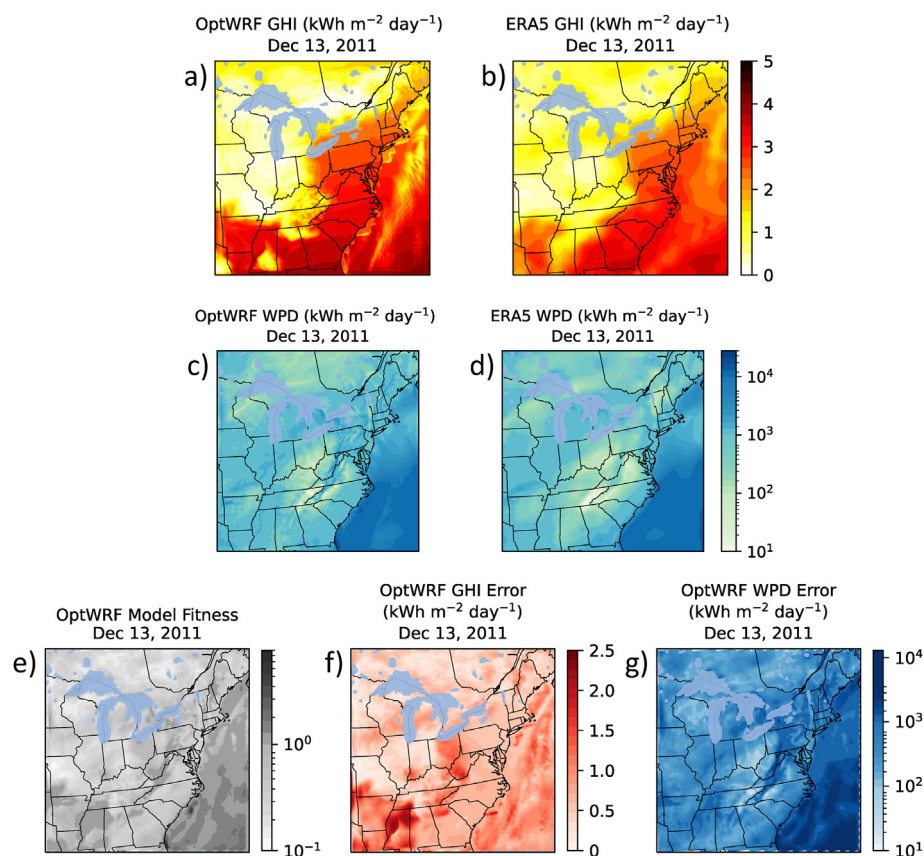


Fig. 3. Panels a) and b) show global horizontal irradiance (GHI) across the entire modeling domain in $\text{kWh m}^{-2} \text{day}^{-1}$ for a one-day WRF simulation initialized on December 13, 2011 00 (UTC) and the ERA5 reanalysis. Similarly, panels c) and d) show the wind power density (WPD) also 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)).

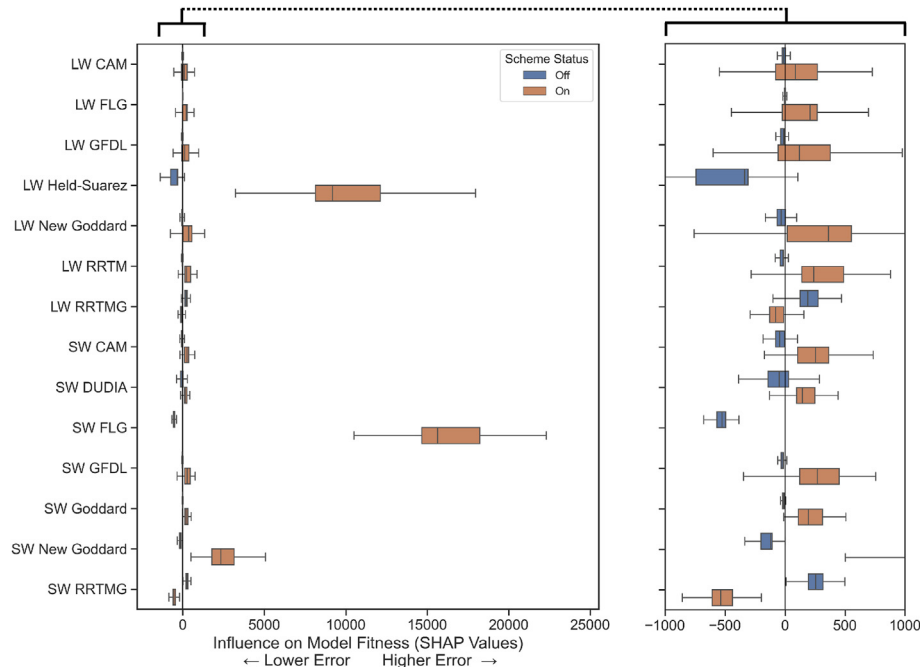


Fig. 4. SHAP values explain the impact that each random forest model feature (WRF parameterization scheme) has upon model error. Radiation parameterization schemes are listed along the left axis, and boxes are created from all the WRF model runs. Orange bars correspond to WRF configurations where the 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 A.4 and A.5 in Appendix A.3, respectively.

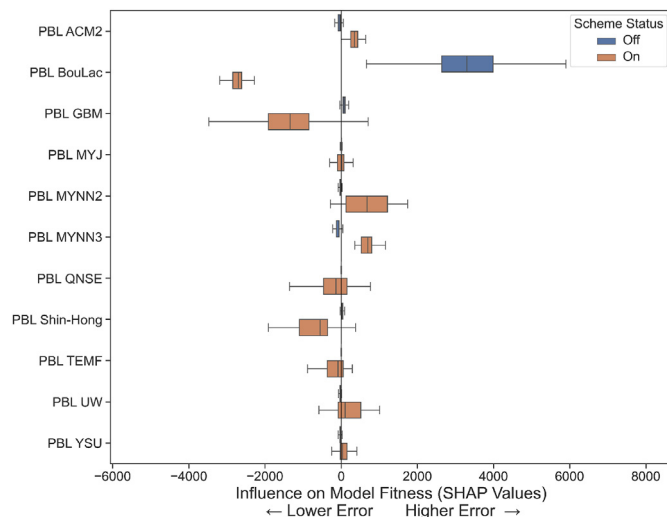


Fig. 5. Planetary boundary layer parameterization schemes are listed along the left axis, and boxes are created from all the WRF model runs. Orange bars correspond to WRF configurations where the parameterization scheme was activated; whereas the blue bars correspond to those where the scheme was deactivated. 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 A.7 in Appendix A.3.

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 [50]. To determine the importance of each parameterization scheme, we employed the SHapley Additive exPlanation (SHAP) Python Package [51]. SHAP provides a high-speed exact algorithm

for explaining the output of tree-based machine learning models and has been shown to correctly rank the contributions of predictor variables to model output [52]. The SHAP values for each radiation and each planetary boundary layer parameterization scheme in each setup are shown in box plots in Fig. 4 and Fig. 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 Fig. 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 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 Fig. 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 in the Northeast United States 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 [26], which is currently only compatible with the MYNN2 scheme. We will engage this point further in Section 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 configurations in which they were activated. Appendix A.7 presents SHAP values plots for the remaining parameterization schemes (microphysics, land surface, and cumulus).

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.

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 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 configuration produced by the GA incurred high errors in WPD off the coast of North Carolina (see Fig. 3 g)). As such, this model setup would almost certainly not be the best choice for simulating the power output from an offshore wind farm there. In other words, the fitness function can be tuned to a user's specific application, and different error metrics may be more appropriate for other applications. For example, a researcher studying power system reliability might care deeply about overpredicting renewable generation. In such a case, a non-absolute error metric should underlie the fitness function. Finally, the choice of data products used for boundary conditions and comparison will impact the model fitness. This work coincided with the transition between ERA-Interim and ERA5 at ECMWF, so both datasets made their way into our work. However, with ERA5 fully available, we recommend that users seeking to apply our method simply use ERA5 for both boundary conditions and comparison, which could more clearly elucidate errors associated with physical parameterization.

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,

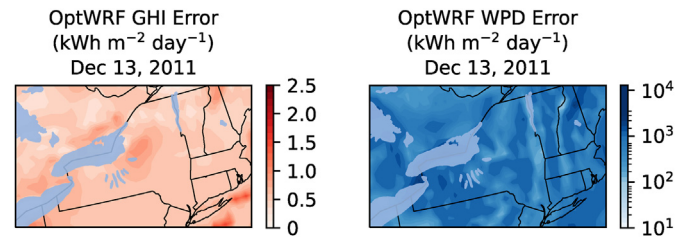


Fig. 6. The left panel shows a zoomed-in view depicting GHI error over New York State and parts of the Northeastern United States – the same quantity as plotted in Fig. 3 f). The largest errors in GHI occurred off the southern and eastern coasts of Long Island, NY. The right panel shows a zoomed-in view depicting WPD error over New York State and parts of the Northeastern United States – the same quantity as plotted in Fig. 3 g). 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.

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.

4.2. Best physics option sets

We turn now to a discussion of why these particular physics option sets (refer back to Table 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 Fig. 6) and WPD errors (shown on the right in Fig. 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 (see maps in Fig. 3a)–b)). However, since these features sit offshore, their presence represents no concern for forecasting solar resources over land thereby potentially making this setup 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, challenges remain in representing unresolvable quantities (e.g., remaining subgrid scale clouds) in regions where terrain features change dramatically within a small number of grid cells. At 12 km, WRF cannot yet resolve convective clouds, which are a feature often produced by flow over mountainous terrain. Therefore, we would expect substantially lower errors running WRF at a higher resolution (e.g., 4 km), but that in turn, would increase the computational and output data burden of OptWRF.

Interestingly, we found that the NSSL single-moment scheme (see Ref. [53] for the original NSSL two-moment scheme) selected by OptWRF was linked to higher errors in a majority of cases, shown in Figure A.2. Since the same is true for the other two microphysics schemes used in the best-performing configurations

(Thompson [54] and Sbu-Ylin [55]), 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 Fig. 4) where both the longwave and shortwave RRTMG [56] schemes are linked to lower errors. Longwave RRTMG is based on the correlated-k method [57] 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” parameterization method [10] outperforms the longwave schemes based upon the older broadband emissivity method [59] (e.g., see Ref. [60]). Of course, there are other sources of error in addition to those related to the parameterizations (e.g., terrain, initial and boundary conditions, etc.). SHAP values may offer more insight if we were able to first estimate the error contributions from various sources and then observe the impact that a particular scheme had upon the portion of the error contributed by the parameterizations.

Turning to errors in WPD shown in the right panel of Fig. 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 36 km wide, it can be spanned by three WRF grid cells (12 km) and is on the same order as the ERA5 grid (~30 km). 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 [61] – poorly represent unresolved orographical features existing in mountainous terrain and remains an open area of research [33,41,62]. Draxl et al. reported that the choice of PBL schemes depends on atmospheric stability [31]. Unfortunately, the BouLac scheme [63] 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 [64], which was developed in conjunction with the ACM2 PBL scheme [65,66] and the Pleim-Xiu surface layer scheme [67], was selected in isolation of its companion schemes by OptWRF. Since we should not judge the overall performance of a scheme based on such limited simulation data, future work should continue to explore hidden synergies among WRF's many parameterization schemes.

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 setup 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 setup. Of course, running a GA with a WRF model 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 configurations 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.

Additional nuance exists when selecting boundary condition data for use in different power systems applications. We selected the ERA5 reanalysis because we presented an example of how OptWRF might be used to inform medium-to long-term power systems planning. Such planners can make use of historical data to extrapolate future trends. However, researchers aiming to implement OptWRF for operational wind and solar forecasting must initialize their simulations using data available forecast data that is updated in real time. Two common data products matching that fulfill this requirement are the Global Forecast System (GFS) provided by the National Centers for Environmental Prediction (NCEP), and the high resolution 10-day forecast (HRES) produced by ECMWF.

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 regions or for different applications. In this repository, we include fitness functions that consider wind and solar in isolation and remark that these use cases may make SHAP values easier to interpret for certain parameterizations (e.g., radiation schemes in a solar only case). 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 [68], 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 setup – 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 setups 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 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. Refs. [22,23]). This work could be integrated into the current framework as a second step after the model physics has been selected during the first. Similarly, WRF contains options with the ability to change the behavior of advection, diffusion, and damping. While many of these options have recommendations, they too could be integrated into this framework to fully encompass all options available to WRF users. 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.

Author statement

Jeffrey Sward: Conceptualization, Methodology, Investigation, Writing – original draft, Reviewing and Editing. **Toby Ault:** Conceptualization, Methodology, Investigation, Writing-Reviewing. **K. Max Zhang:** Conceptualization, Methodology, Investigation, Writing-Reviewing and Editing, Supervision and Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to acknowledge the support from the New York State Energy Research and Development Authority (NYSERDA). Jeffrey A. Sward's work was supported by the National Science Foundation Graduate Research Fellowship under grant number DGE-1650441.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.energy.2022.124367>.

References

- [1] Global trends in renewable energy Investment, Tech. Rep. Frankfurt, Germany: Frankfurt School - UN Environment Programme Centre; 2020.
- [2] Mathiesen P, Kleissl J. Evaluation of numerical weather prediction for intraday solar forecasting in the continental United States. *Solar Energy* 2011;85: 967–77. <https://doi.org/10.1016/j.solener.2011.02.013>.
- [3] P. Mathiesen, C. Collier, J. Kleissl, A high-resolution, cloud-assimilating numerical weather prediction model for solar irradiance forecasting doi: 10.1016/j.solener.2013.02.018.
- [4] Lima FJ, Martins FR, Pereira EB, Lorenz E, Heinemann D. Forecast for surface solar irradiance at the Brazilian Northeastern region using NWP model and artificial neural networks. *Renewable Energy* 2016;87:807–18. <https://doi.org/10.1016/j.renene.2015.11.005>.
- [5] Skamarock WC, Klemp JB. A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *J Comput Phys* 2008;227: 3465–85. <https://doi.org/10.1016/j.jcp.2007.01.037>.
- [6] Zhao J, Guo Z-H, Su Z-Y, Zhao Z-Y, Xiao X, Liu F. An improved multi-step forecasting model based on wrf ensembles and creative fuzzy systems for wind speed. *Appl Energy* 2016;162:808–26. <https://doi.org/10.1016/j.apenergy.2015.10.145>. URL <https://doi.org/10.1016/j.apenergy.2015.10.145>.
- [7] Jimenez PA, Hacker JP, Haupt E, Ruiz-Arias JA, Gueymard CA, Thompson G, Eidhammer T, Deng A. 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. *Bull Am Meteorol Soc* 2015;97:1249–64. <https://doi.org/10.1175/BAMS-D-14-00279.1>.
- [8] Draxl C, Clifton A, Hodge B-M, Mccaa J. The wind integration national dataset (WIND) Toolkit. *Appl Energy* 2015;151:355–66. <https://doi.org/10.1016/j.apenergy.2015.03.121>.
- [9] Lundquist JK, DuVivier KK, Kaffine D, Tomaszewski JM. Costs and consequences of wind turbine wake effects arising from uncoordinated wind energy development. *Nat Energy* 2018;4(1 4):26–34. <https://doi.org/10.1038/s41560-018-0281-2>. 2018, <https://www.nature.com/articles/s41560-018-0281-2>.
- [10] Pielke RA. *Mesoscale meteorological modeling*. third ed., vol. 98. New York: Academic Press; 2013.
- [11] Borge R, Alexandrov V, José Del Vas J, Lumbreras J, Rodríguez E. A comprehensive sensitivity analysis of the WRF model for air quality applications over the Iberian Peninsula. *Atmos Environ* 2008;42:8560–74. <https://doi.org/10.1016/j.atmosenv.2008.08.032>. Contents.
- [12] Bukovsky MS, Karoly DJ. Precipitation simulations using WRF as a nested regional climate model. *J Appl Meteorol Climatol* 2009;48(10):2152–9. <https://doi.org/10.1175/2009JAMC2186.1>.
- [13] J. Jin, N. L. Miller, N. Schlegel, Sensitivity study of four land surface schemes in the WRF model, *Adv Meteorol*:10.1155/2010/167436.
- [14] Gilliam RC, Pleim JE. Performance assessment of new land surface and planetary boundary layer physics in the WRF-ARW. *J Appl Meteorol Climatol* 2010;49(4):760–74. <https://doi.org/10.1175/2009JAMC2126.1>.
- [15] Flaounas E, Bastin S, Janicot S. Regional climate modelling of the 2006 West African monsoon: sensitivity to convection and planetary boundary layer parameterisation using WRF. *Clim Dynam* 2011;36:1083–105. <https://doi.org/10.1007/s00382-010-0785-3>.
- [16] Mooney PA, Mulligan FJ, Fealy R. Evaluation of the sensitivity of the weather research and forecasting model to parameterization schemes for regional climates of Europe over the period 1990–95. *J Clim* 2013;26:1002–17. <https://doi.org/10.1175/JCLI-D-11-00676.1>.
- [17] Quan J, Di Z, Duan Q, Gong W, Wang C, Gan Y, Ye A, Miao C. 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* 2016;142:2925–34. <https://doi.org/10.1002/qj.2885>.
- [18] L. Pan, Y. Liu, J. C. Kniviel, L. Delle Monache, G. Roux, Evaluations of WRF sensitivities in surface simulations with an ensemble prediction system, *Atmosphere* 9 (106). doi:10.3390/atmos9030106.
- [19] Stergiou I, Tagaris E, Sotiropoulou R-EP. Sensitivity Assessment of WRF parameterizations over Europe. 2nd International Electronic Conference on Atmospheric Sciences; 2017. <https://doi.org/10.3390/ecas2017-04138>.
- [20] Yu X, Gen M. *Introduction to evolutionary algorithms*. Springer; 2010.
- [21] Mirjalili S. Genetic algorithm. Cham: Springer International Publishing; 2019. p. 43–55. https://doi.org/10.1007/978-3-319-93025-1_4.
- [22] Ishaish H, Cortés A, Senar MA. Parallel multi-level genetic ensemble for numerical weather prediction enhancement. *Procedia Comput Sci* 2012;9: 276–85. <https://doi.org/10.1016/j.procs.2012.04.029>.
- [23] Chinta S, Balaji C. Calibration of WRF model parameters using multiobjective adaptive surrogate model-based optimization to improve the prediction of the Indian summer monsoon. *Clim Dynam* 2020;55(3–4):631–50. <https://doi.org/10.1007/s00382-020-05288-1>.
- [24] Díaz-Isaac LI, Lauvaux T, Bocquet M, Davis KJ. Calibration of a multi-physics ensemble for estimating the uncertainty of a greenhouse gas atmospheric transport model. *Atmos Chem Phys* 2019;19(8):5695–718. <https://doi.org/10.5194/acp-19-5695-2019>.

- [25] Oana L, Spataru A. Use of genetic algorithms in numerical weather prediction. In: Proceedings - 18th International Symposium on Symbolic and numeric algorithms for Scientific computing. SYNASC 2016; 2016. p. 456–61. <https://doi.org/10.1109/SYNASC.2016.075>.
- [26] Fitch AC, Olson JB, Lundquist JK, Dudhia J, Gupta AK, Michalak J, Barstad I. Local and mesoscale impacts of wind farms as parameterized in a mesoscale NWP model. *Mon Weather Rev* 2012;140:3017–38. <https://doi.org/10.1175/MWR-D-11-00352.1>.
- [27] Fitch AC. Notes on using the mesoscale wind farm parameterization of Fitch et al. (2012). In: WRF, wind energy, vol. 19; 2015. p. 1757–8. <https://doi.org/10.1002/we>.
- [28] Jiménez PA, Navarro J, Palomares AM, Dudhia J. 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* 2015;18(3):559–66. <https://doi.org/10.1002/we.1708>.
- [29] Lee JCY, Lundquist JK. 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); 2017. p. 1–31. <https://doi.org/10.5194/gmd-2017-128>.
- [30] E.-M. Giannakopoulou, R. Nhili, WRF model methodology for offshore wind energy applications, *Adv Meteorol*:10.1155/2014/319819.
- [31] Draxl C, Hahmann AN, Peña A, Giebel G. Evaluating winds and vertical wind shear from Weather Research and Forecasting model forecasts using seven planetary boundary layer schemes. *Wind Energy* 2014;17:39–55. <https://doi.org/10.1002/we>. arXiv:arXiv:1006.4405v1.
- [32] Penchah MM, Malakooti H, Satkin M. Evaluation of planetary boundary layer simulations for wind resource study in east of Iran. *Renewable Energy* 2017;111:1–10. <https://doi.org/10.1016/j.renene.2017.03.040>.
- [33] Gómez-Navarro JJ, Raible CC, Dier S. Sensitivity of the WRF model to PBL parameterisations and nesting techniques: evaluation of wind storms over complex terrain. *Geosci. Model Dev* 2015;8:3349–63. <https://doi.org/10.5194/gmd-8-3349-2015>.
- [34] X. T. Chadee, N. R. Seegobin, R. 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). doi:10.3390/en10070931.
- [35] Floors R, Hahmann AN, Peña A. Evaluating mesoscale simulations of the coastal flow using lidar measurements. *J Geophys Res Atmos* 2018;123(5): 2718–36. <https://doi.org/10.1002/2017JD027504>.
- [36] Yang B, Berg LK, Qian Y, Wang C, Hou Z, Liu Y, Shin HH, Hong S, Pekour M. Parametric and structural sensitivities of turbine-height wind speeds in the boundary layer parameterizations in the weather research and forecasting model. *J Geophys Res Atmos* 2019;124:5951–69. <https://doi.org/10.1029/2018JD029691>.
- [37] Mattar C, Borvarán D. Offshore wind power simulation by using WRF in the central coast of Chile. *Renewable Energy* 2016;94:22–31. <https://doi.org/10.1016/j.renene.2016.03.005>.
- [38] H. Gamarro, J. E. Gonzalez, L. E. Ortiz, On the assessment of a numerical weather prediction model for solar photovoltaic power forecasts in cities, *J Energy Resour Technol* 141.
- [39] Lopes FM, Silva HG, Salgado R, Cavaco A, Canhoto P, Collares-Pereira M. Short-term forecasts of GHI and DNI for solar energy systems operation: assessment of the ECMWF integrated forecasting system in southern Portugal. *Solar Energy* 2018;170:14–30. <https://doi.org/10.1016/j.solener.2018.05.039>.
- [40] Zempila M-M, Giannaros TM, Bais A, Melas D, Kazantzidis A. Evaluation of WRF shortwave radiation parameterizations in predicting Global Horizontal Irradiance in Greece. *Renewable Energy* 2016;86:831–40. <https://doi.org/10.1016/j.renene.2015.08.057>.
- [41] Jiménez PA, Dudhia J. Improving the representation of resolved and unresolved topographic effects on surface wind in the wrf model. *J Appl Meteorol Climatol* 2012;51(2):300–16. <https://doi.org/10.1175/JAMC-D-11-084.1>.
- [42] Skamarock W, Klemp J, Dudhia J, Gill D, Zhiquan L, Berner J, Wang W, Powers J, Duda MG, Barker DM, Huang X-Y. A description of the Advanced research WRF model version 4. Tech. Rep.. Boulder, CO: NCAR; 2019.
- [43] Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmasada MA, Balsamo G, Bauer P, Bechtold P, M Beljaars AC, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Isaksen I, Källberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette J-j, Park B-k, Peubey C, de Rosnay P, Tavolato C, Thépaut J-n, Vitart F, Acn B, de Berg vL, M J-j, P B-k, Rosnay dP. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society Q. J. R. Meteorol. Soc* 2011;137:553–97. <https://doi.org/10.1002/qj.828>.
- [44] Hersbach H, Bell B, Berrisford P, Hirasara S, Horányi A, Muñoz-Sabater J, Nicolas J, Peubey C, Radu R, Schepers D, Simmons A, Soci C, Abdalla S, Abellan X, Balsamo G, Bechtold P, Biavati G, Bidlot J, Bonavita M, De Chiara G, Dahlgren P, Dee D, Diamantakis M, Dragani R, Flemming J, Forbes R, Fuentes M, Geer A, Haimberger L, Healy S, Hogan RJ, Hólm E, Janisková M, Keeley S, Laloyaux P, Lopez P, Lupu C, Radnoti G, de Rosnay P, Rozum I, Vamborg F, Villaume S, Thépaut JN. The ERA5 global reanalysis. *Q J R Meteorol Soc* 2020;146(730):1999–2049. <https://doi.org/10.1002/qj.3803>.
- [45] Mills KL, Filliben JJ, Haines AL. Determining relative importance and effective settings for genetic algorithm control parameters. *Evol Comput* 2015;23(2): 309–42. https://doi.org/10.1162/EVCO_a_00137.
- [46] Cazacu R. Comparative study between the improved implementation of 3 classic mutation operators for genetic algorithms. *Procedia Eng* 2017;181: 634–40. <https://doi.org/10.1016/j.proeng.2017.02.444>.
- [47] Greenwell R, Angus J, Finck M. Optimal mutation probability for genetic algorithms. *Math Comput Model* 1995;21(8):1–11. [https://doi.org/10.1016/0895-7177\(95\)00035-Z](https://doi.org/10.1016/0895-7177(95)00035-Z).
- [48] Martínez-Gracia A, Arauzo I, Uche J. Solar energy availability. In: *Solar Hydrogen production: processes, systems and Technologies*. Elsevier; 2019. p. 113–49. <https://doi.org/10.1016/B978-0-12-814853-2.00005-9>.
- [49] Breiman L. Random forests. *Mach Learn* 2001;45:5–32. <https://doi.org/10.1201/9780429469275-8>.
- [50] C. Strobl, A. L. Boulesteix, A. Zeileis, T. Hothorn, Bias in random forest variable importance measures: illustrations, sources and a solution, *BMC Bioinf* 8 (1). doi:10.1186/1471-2105-8-25.
- [51] Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, Katz R, Himmelfarb J, Bansal N, Lee S-I. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell* 2020;2(1):56–67. <https://doi.org/10.1038/s42256-019-0138-9>.
- [52] J. Gu, B. Yang, M. Brauer, K. M. Zhang, Enhancing the evaluation and interpretability of data-driven air quality models, *Atmos Environ* 246. doi:<https://doi.org/10.1016/j.atmosenv.2020.118125>.
- [53] Mansell ER, Ziegler CL, Bruning EC. Simulated electrification of a small thunderstorm with two-moment bulk microphysics. *J Atmos Sci* 2010;67(1): 171–94. <https://doi.org/10.1175/2009JAS2965.1>.
- [54] Thompson G, Field PR, Rasmussen RM, Hall WD. Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: implementation of a new snow parameterization. *Mon Weather Rev* 2008;136(12): 5095–115. <https://doi.org/10.1175/2008MWR2387.1>.
- [55] Lin Y, Colle BA. A new bulk microphysical scheme that includes riming intensity and temperature-dependent ice characteristics. *Mon Weather Rev* 2011;139(3):1013–35. <https://doi.org/10.1175/2010MWR3293.1>.
- [56] M. J. Iacono, J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, W. D. Collins, Radiative forcing by long-lived greenhouse gases: calculations with the AER radiative transfer models, *J Geophys Res* 113. doi:10.1029/2008JD009944.
- [57] Mlawer EJ, Taubman SJ, Brown PD, Iacono MJ, Clough SA. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J Geophys Res Atmos* 1997;102(14):16663–82. <https://doi.org/10.1029/97jd00237>.
- [58] Rothman LS, Gordon IE, Barbe A, Chris Benner D, Bernath PF, Birk M, Boudon V, Brown LR, Campargue A, Champion J-P, Chance K, Coudert LH, Dana V, Devi VM, Fally S, Flaud J-M, Gamache RR, Goldman A, Jacquemart D, Kleiner I, Lacombe N, Lafferty WJ, Mandin J-Y, Massie ST, Mikhailenko SN, Miller CE, Moazzen-Ahmadi N, Naumenko OV, Nikitin AV, Orphal J, Perevalov VI, Perrin A, Predoi-Cross A, Rinsland CP, Rotger M, Smith MAH, Sung K, Tashkun SA, Tennyson J, Toth RA, Vandaele AC, Auwera JV. The HITRAN 2008 molecular spectroscopic database. *J Quant Spectrosc Radiat Transfer* 2009;110:533–72. <https://doi.org/10.1016/j.jqsrt.2009.02.013>.
- [59] Stephens GL. The parameterization of radiation for numerical weather prediction and climate models. *Mon Weather Rev* 1984;112(4):826–67. [https://doi.org/10.1175/1520-0493\(1984\)112<0826:TPORFN>2.0.CO;2](https://doi.org/10.1175/1520-0493(1984)112<0826:TPORFN>2.0.CO;2).
- [60] Dudhia J. Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *J Atmos Sci* 1989;42(2):3077–107.
- [61] Hong SY, Noh Y, Dudhia J. A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon Weather Rev* 2006;134(9): 2318–41. <https://doi.org/10.1175/MWR3199.1>.
- [62] Jiménez PA, Dudhia J. On the ability of the WRF model to reproduce the surface wind direction over complex terrain. *J Appl Meteorol Climatol* 2013;52(7):1610–7. <https://doi.org/10.1175/JAMC-D-12-0266.1>.
- [63] Bougeault P, Lacarrere P. Parameterization of orography-induced turbulence in a mesobeta-scale model. *Mon Weather Rev* 1989;117:1872–90.
- [64] Xiu A, Pleim JE. Development of a land surface model. Part I: application in a mesoscale meteorological model. *J Appl Meteorol* 2001;40(2):192–209. [https://doi.org/10.1175/1520-0450\(2001\)040<0192:DOALSM>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<0192:DOALSM>2.0.CO;2).
- [65] Pleim JE. A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: model description and testing. *J Appl Meteorol Climatol* 2007;46(9):1383–95. <https://doi.org/10.1175/JAM2539.1>.
- [66] Pleim JE. A combined local and nonlocal closure model for the atmospheric boundary layer. Part II: application and evaluation in a mesoscale meteorological model. *J Appl Meteorol Climatol* 2007;46(9):1396–409. <https://doi.org/10.1175/JAM2534.1>.
- [67] Pleim JE. A simple, efficient solution of flux-profile relationships in the atmospheric surface layer. *J Appl Meteorol Climatol* 2006;45(2):341–7. <https://doi.org/10.1175/JAM2339.1>.
- [68] K. L. Mills, J. J. Filliben, A. L. Haines, W. R. Associates, Determining relative importance and effective settings for genetic algorithm control parameters, [Tech. rep].
- [70] Forsythe WC, Rykiel EJ, Stahl RS, Wu H-I, Schoolfield RM. A model comparison for daylength as a function of latitude and day of year. 1995. Tech. rep.