Simulation of wind and solar energy generation over California with E3SM SCREAM regionally refined models at 3.25 km and 800 m resolutions

Jishi Zhang[1], Jean–Christophe Golaz[1], Matthew Vincent Signorotti[1], Hsiang–He Lee[1], Peter Bogenschutz[1], Minda Monteagudo[1], Paul Aaron Ullrich[1], Robert S. Arthur[1], Stephen Po–Chedley[1], Philip Cameron–smith[1], and Jean–Paul Watson[1]

[1]Lawrence Livermore National Laboratory, Livermore, California, United States

Correspondence: Jishi Zhang (zhang73@llnl.gov)

Abstract. This study produces wind and solar power generation estimates derived from the US Department of Energy's Simple Cloud-Resolving ~~Energy Exascale Earth System Model (~~E3SM ~~)~~ Atmosphere Model (SCREAM) ~~by leveraging regional mesh refinement over California (CARRM) simulations~~ Regionally refined Models (RRM) over California at 3.25 km and 800 m horizontal resolutions, using the Python wrapper of System Advisor Model (PySAM). The resulting wind and solar energy generation estimates are compared to monthly capacity factors reported to Energy Information Administration (EIA), High-Resolution Rapid Refresh (HRRR, 3 km resolution) forecast model, and E3SM North American regionally refined model (NARRM, 25 km resolution). We systematically assess the impacts of generation modeling assumptions, meteorological models, and horizontal resolution. Results show that resolution plays a dominant role for wind energy: increasing from 25 km to 3.25 km brings qualitative and quantitative improvements, most notably by resolving the phase error in the seasonal cycle found in coarser simulations. However, further refinement to 800 m offers minimal gains. SCREAM's performance for solar generation surpasses HRRR, likely due to more accurate surface radiation. The sensitivity of PySAM to system configuration, particularly for axis-tracking modeling in photovoltaics, is also highlighted. Overall, SCREAM-RRM shows strong potential for high-resolution energy assessments, with future progress depending on more in situ observations and clearer quantification of generation modeling uncertainties.

1 Introduction

Su Shi wrote in his *First Prose on the Red Cliff* (1082)[1]:

> "But as for the clear breeze over the river and the bright moon between the hills – sound to the ear, colour to the eye – free to take and never spent. They are the Creator's inexhaustible storehouse, and delights for you and me to share."

[1]Adapted from the translation by Graham (1965). Original Chinese text: "惟江上之清风，与山间之明月，耳得之而为声，目遇之而成色，取之无禁，用之不竭。是造物者之无尽藏也，而吾与子之所共适"苏轼«前赤壁赋»

This poetic reflection captures a truth that still resonates in today's energy landscape: wind and sunlight, gifts of nature, are invaluable, endlessly renewed, and shared by all. Yet their abundance comes with uncertainty. Unlike conventional fuels, these resources fluctuate with weather and climate. Harnessing them effectively is not a matter of extraction, but of prediction and adaptation. In response to this challenge, and amid long-term energy transition efforts and growing sustainability concerns over fossil fuels, renewable technologies like wind ~~and solar power~~ turbines and solar photovoltaic (PV) systems have advanced rapidly over the past decades, making intelligent use of these virtually inexhaustible forces for electricity generation and reflecting a design ethos that respects natural constraints while embracing engineering ingenuity. ~~Global renewable electricity generation – led by wind and solar power –~~

According to the California Energy Commission (CEC), California currently has approximately 87.8 GW of installed electric generation capacity distributed across more than 1,600 power plants statewide in 2023 (https://www.energy.ca.gov/data-reports/energy-almanac/california-electricity-data/2023-total-system-electric-generation, last accessed: February 20, 2026). Natural gas remains the dominant technology, accounting for about 39.7 GW (45%) of total nameplate capacity. Renewable energy resources contribute approximately 32.9 GW (37.5%) of installed capacity, including about 20.9 GW of solar PV capacity (24%) and 6.3 GW of wind capacity (7%). In terms of electricity generation rather than installed capacity, solar and wind energy supplied approximately 19.2% and 6.5%, respectively, of California's in-state electricity generation, together accounting for about one quarter of annual electricity demand on an energy basis. Global renewable electricity generation is expected to nearly double from 2023 to 2030, rising from 30% to 46% of global electricity supply and exceeding 17,000 TWh in total output (IEA, 2024, 2025).

The inherent variability of wind and solar resources – particularly their hourly fluctuations – presents significant challenges for system reliability and calls for robust planning and forecasting. The siting of wind and solar plants is fundamentally based on local assessments of wind and solar resources, most notably hub-height wind speed and surface solar radiation. These meteorological variables, together with modeling of technology-specific generation characteristics, inform the spatiotemporal distribution of wind and solar energy potential (e.g., Ryberg et al., 2019). These resources are inherently tied to local weather and climate conditions, featuring strong geographic heterogeneity and diurnal variability. They are also influenced by hourly variability and by rare weather events. Therefore, high spatio-temporal resolution modeling of meteorological state is critically needed across the energy sector – from resource assessment, to power plant development and grid planning (Rodrigo et al., 2017; Wang et al., 2018; Frank et al., 2020; McKenna et al., 2022; Davis et al., 2023; Chen et al., 2024; Arthur et al., 2025a). However, meteorological datasets that can be used directly to derive wind and solar power generation remain limited. Energy modeling requires at least hourly meteorological variables that are critical to power generation processes. For wind energy, this includes wind speeds at turbine hub height, while solar generation requires surface downwelling shortwave radiation and, ideally, its direct and diffuse components. Many widely used reanalyses such as ERA5 (European Centre for Medium-Range Weather Forecasts Reanalysis version 5; Hersbach et al., 2020) and MERRA2 (Modern-Era Retrospective Analysis for Research and Applications Version 2; Gelaro et al., 2017), provide only near-surface wind speed and total shortwave radiation or near-surface wind speed, requiring additional assumptions such as temporal interpolation, vertical extrapolation based on boundary-layer wind profiles, or radiation decomposition. In contrast, high-frequency

km-scale atmospheric simulations can directly provide physically consistent meteorological fields relevant to energy generation, making them a natural framework for evaluating wind and solar power potential.

Previously, Lee et al. (2025) evaluated wind and solar energy generation using the U.S. Department of Energy's Energy Exascale Earth System Model (E3SM) 25 km North American (NA) Regionally Refined Model (RRM) version 2 (hereafter referred to as "E3SM-25kmNARRM"; Tang et al., 2023). When evaluated against U.S. Energy Information Administration (EIA) monthly generation reports, they found that E3SM-25kmNARRM realistically simulates wind energy production across most of the continental U.S. However, E3SM-25kmNARRM fails to reproduce the observed seasonal cycle of wind energy production in the western U.S., particularly in California. This is perhaps unsurprising given that 25 km horizontal resolution is insufficient to resolve the region's complex topography, one of the most topographically diverse regions in the contiguous U.S.

Complex topography often ~~results in more~~ produces diverse microclimates and ~~stronger~~ strong spatial variability in meteorological conditions, ~~which demand~~ requiring high spatial resolution and accurate topographic representation in numerical models ~~. For instance, California displays pronounced geographic transitions: from complex coastal and mountainous zones, to an expansive inland valley, the Sierra Nevada range, and arid highland deserts.~~ (e.g., Zhang et al., 2024). These land–sea and topographic contrasts shape California's ~~diverse microclimates (e.g., Zhang et al., 2024). In summer, prevalent~~ heterogeneous wind and solar resources. For example, coastal stratocumulus clouds ~~reflect more shortwave radiation, limiting solar energy~~ frequently limit solar radiation near the coast~~. Wind energy is shaped by a cascade of meteorological processes across scales~~ , while wind energy is influenced by processes spanning multiple scales, from synoptic land–sea thermal gradients ~~and mesoscale pressure gradients,~~ to *km*-scale ~~topography-induced flows and sub-kilometer surface and vegetation heterogeneity (e.g., Davis et al., 2023; Arthur et al., 2025b). Accordingly, both wind and solar~~ terrain-induced flows (e.g., Davis et al., 2023; Arthur et al. . These characteristics suggest that high-resolution modeling may substantially improve renewable energy assessments in California~~stand to benefit from high-resolution modeling~~.

The Simple ~~Convective Resolving~~ Cloud-Resolving E3SM Atmosphere Model (SCREAM) is a ~~global~~ convection-permitting ~~model of~~ global atmosphere model within the E3SM project (Caldwell et al., 2021; Donahue et al., 2024). ~~For computational efficiency, Zhang et al. (2024) employed SCREAM~~ Using RRMs, SCREAM can perform kilometer-scale simulations over targeted regions while maintaining global coverage. Recent studies have applied SCREAM with a 3.25 km California ~~(CA) RRM to conduct 20-year climate simulations, and Bogenschutz et al. (2024) evaluated the 3.25 km CARRM in atmospheric river case studies and extended the analysis to an 800 m CARRM for higher-resolution comparison. At 3.25 km resolution, SCREAM-CARRM has been shown to reasonably capture atmospheric rivers, near-surface temperature,~~ RRM (CARRM) to investigate regional meteorology, including atmospheric rivers, coastal stratocumulus, and ~~the spatial distribution of precipitation and snowpack (Zhang et al., 2024)~~precipitation patterns (Zhang et al., 2024; Bogenschutz et al., 2024). However, these studies ~~did not directly assess~~ focused primarily on meteorological evaluation and did not assess the implications of SCREAM simulations for wind and solar energy generation~~potential~~.

Capacity factors (CFs) – defined as the ratio of actual energy generation to installed capacity – are commonly used to evaluate wind and solar ~~power~~ energy performance. A recent multi-model dataset from PLUSWIND (Millstein et al., 2023),

which provides plant-level CFs across the U.S., shows that CFs derived from the 3 km HRRR developed by NOAA (High-Resolution Rapid Refresh; Dowell et al., 2022) ~~modeled wind energy generation more closely matches~~ more closely match EIA-reported monthly generation ~~compared to the coarser~~ than those derived from coarser reanalysis products, including the 31 km ERA5 ~~(European Centre for Medium-Range Weather Forecasts Reanalysis version 5; Hersbach et al., 2020)~~ and 50 km MERRA2~~(Modern-Era Retrospective Analysis for Research and Applications Version 2; Gelaro et al., 2017)~~. PLUSWIND also raised two open questions: is HRRR's lower bias in regions such as California primarily due to its finer spatial resolution, or due to differences in meteorological modeling? And to what extent do generation modeling assumptions (such as wake loss representations) affect the estimated CFs? ~~Given Lee et al. (2025)'s findings~~ Both questions are insightful.

With respect to sensitivity to horizontal resolution, combined with the findings of Lee et al. (2025) that E3SM-25kmNARRM ~~cannot~~ fails to capture the observed seasonal cycle of ~~California wind energy , it is well worth investigating whether higher resolution~~ wind energy in California, we were motivated to examine whether higher-resolution SCREAM-CARRM simulations (i.e., 3.25 km ~~or~~ and 800 m) ~~simulations~~ reduce discrepancies with EIA wind and solar generation data. ~~Furthermore, juxtaposing our findings with PLUSWINDoffers~~ Regarding generation modeling, for all model datasets analyzed here we estimate CFs using the Python wrapper of the System Advisor Model (PySAM) (NREL). We note that our coupling of meteorological outputs to the PySAM wind generation model differs from several assumptions adopted in PLUSWIND, which provides an opportunity to ~~compare the relative accuracy of simulations with varied resolution, meteorological models, and energy generation modeling assumptions~~qualitatively assess the sensitivity of wind CFs to different generation model. In addition, following a similar approach to Lee et al. (2025), we configure different axis-tracking technologies in PySAM to examine their impacts on solar CFs.

To explore the sensitivity of modeled wind and solar generation to different factors, we conducted ~~three new~~ SCREAM-CARRM experiments of the current climate: two at 3.25 km resolution over California (hereafter referred to as "SCREAM-3kmCARRM") with different forcings, and one at 800 m (hereafter referred to as "SCREAM-800mCARRM"). These are analyzed alongside E3SM-25kmNARRM simulations and 3 km HRRR forecasts. ~~For all model datasets analyzed here, we estimate CFs using the Python wrapper of System Advisor Model (PySAM) (NREL). Together, these datasetsform a comparative framework through which we~~

By estimating CFs using PySAM and comparing all modeled results with EIA monthly data, we construct multiple angles of comparison. Our high-level study objectives are 1) whether km-scale SCREAM-CARRM climate data combined with PySAM constitutes an effective energy modeling framework, and 2) whether it can substantially reduce the pronounced wind energy biases identified in 25 km E3SM simulations over California, i.e., how important horizontal resolution is for wind and solar energy simulations. EIA provides the benchmark most directly relevant to applications; however, discrepancies relative to EIA reflect combined uncertainties associated with SCREAM-CARRM datasets, PySAM energy modeling, and EIA data quality. We therefore introduce additional evaluation targets to approximately assess the relative importance of these uncertainties. Specifically, we first assess whether PySAM is applied appropriately, then compare against HRRR at comparable resolution to evaluate whether SCREAM-3kmCARRM has similar performance, and finally examine the role of horizontal resolution. Accordingly, we address the following three questions:

1. *Do the use of different energy models affect wind CFs? How ~~much~~ do ~~generation modeling assumptions impact wind energy simulations~~different axis-tracking methods affect biases in modeled solar CFs?* This ~~is examined by comparing the HRRR-modeled wind~~question focuses on validating our use of PySAM in assessing whether the SCREAM-CARRM + PySAM framework is effective. For wind, we compare HRRR-derived CFs from PLUSWIND with those ~~computed using PySAMfrom the same HRRR data. Both sets are then benchmarked with~~ estimated using PySAM, with both evaluated against EIA monthly CFs~~to quantify discrepancies. The main sources of difference are turbine power curve selection, air density estimation, and wake loss assumptions.~~ . We also reproduce ~~comparison across~~ the PLUSWIND comparisons among ERA5, ~~MERRA2~~MERRA-2, and HRRR~~within PLUSWIND to provide broader context and assess discrepancies across reanalysis products.~~ , which helps assess the relative importance of horizontal resolution and generation modeling. For solar, we evaluate the impacts of fixed, single-axis, and dual-axis tracking assumptions in PySAM.

2. *How ~~much do meteorological models impact~~ different are wind and solar ~~energy~~CFs between SCREAM-3kmCARRM and HRRR at comparable horizontal resolution?* ~~This is evaluated by comparing the modeled CFs from HRRR 2018-2022 forecasts and~~ Using the same PySAM configuration, we assess the performance of SCREAM-CARRM by comparison with HRRR. Although these two meteorological models differ substantially in model structure, physical parameterizations, and simulation periods (HRRR forecasts over 2018–2022 vs. SCREAM-3kmCARRM ~~with the 2010s climatological forcings (hereafter referred to as "2010climo"), both calculated with PySAM. These two datasets have comparable horizontal resolution and are both forced with observed sea surface temperature (SST). HRRR assimilates high-frequency observations, so its weather evolution is expected to more closely match reality than the SCREAM-3kmCARRM simulation, which is driven by the~~ forced by observed climatology over 2005–~~2014. However, since our focus is on monthly climatology, the 2010climo forcing should produce climatological states from a similar time period as HRRR. Thus, discrepancies here primarily reflect structural differencesin the meteorological models, including dynamics, physical parameterizations, and data assimilation/nudging~~2014), they share a very similar horizontal resolution at the kilometer scale. We therefore ask whether, despite these differences, comparable horizontal resolution can dominate and lead to statistically similar wind and solar CFs.

3. *How much does horizontal resolution affect wind and solar energy?* ~~This is examined by comparing modeled~~ We compare wind and solar CFs from the 3.25 km and 800 m SCREAM-CARRM simulations, both driven by CMIP6 ~~future~~ forcings and evaluated ~~in~~ over the present-day period~~(2015–2020 water years; hereafter referred to as "present-day")~~, with CFs calculated using PySAM. These simulations ~~are forced with the same 1° E3SM SST and ice cover, use the same nudging strategy, and share the same SCREAM version 0 codebase. Therefore, resolution is~~ use identical model configurations, with horizontal resolution as the only difference~~. This setup provides~~ , providing a clean assessment of ~~horizontal~~ resolution sensitivity specific to SCREAM-CARRM. ~~The result may not generalize to all GCMs, as resolution sensitivity often depends on the model's physical parameterizations. In addition, we~~ We also compare against the E3SM-25kmNARRM ~~2005–2014 historical simulation which differs in several aspects, including resolution,~~

Table 1. Overview of SCREAM-CARRRMs, E3SM-25kmNARRM, HRRR simualtions providing acronyms used in this study, horizontal resolutions, dataset periods, descriptions, and references.

Simulation acronym	Resolution	Dataset Period	Simulation description	Reference[a]
SCREAM-3kmCARRM 2010climo	3.25 km	10-year	Observed SST and sea ice concentration representative of the 2010s	
HRRR	3 km	2018–2022	Forecasts with boundary conditions provided by the Rapid Refresh system and assimilates the NOAA Multi-Radar Multi-Sensor project	Dowell et al. (2022)
E3SM-25kmNARRM	25 km	10-year	Coupled atmosphere–ocean historical simulations	Tang et al. (2023)[b]
SCREAM-3kmCARRM present-day	3.25 km	5-year	Observed SST and sea ice concentration with boundary conditions provided by the present-day period of E3SMv1 SSP5-8.5 future projection	Zhang et al. (2024)[c]
SCREAM-800mCARRM present-day	800 m	5-year	Observed SST and sea ice concentration with boundary conditions provided by the present-day period of E3SMv1 SSP5-8.5 future projection	

[a] An empty or footnoted reference entry indicates that the simulation dataset was produced as part of this study.

[b] Rerun to get the additional outputs used for energy modeling.

[c] Rerun to use the same E3SMv1 boundary conditions as the SCREAM-800mCARRM present-day simulations.

~~physical parameterizations, and internal variability~~simulations to assess whether the km-scale simulations substantially reduce the wind generation biases identified at 25 km resolution.

Table 1 summarizes the acronyms for all simulations and datasets, along with their resolutions, ~~time~~ simulation durations/periods, descriptions, and references. Further details, including the development of the 3.25 km and 800 m SCREAM-CARRRMs, are provided in the Methods section.

The remainder of this paper is organized as follows. Section 2 describes the 3.25 km and 800 m SCREAM-CARRM configurations and other model outputs used to model wind and solar generation. Section 3 presents the simulation results and comparisons with observations. Section 4 discusses the implications and broader applicability of our findings.

2 Methods

This section describes the development and experimental design of the SCREAM-CARRMs used in this study, along with the other simulation datasets. Modeled CFs are computed by coupling meteorological outputs from SCREAM-CARRMs, E3SM-25kmNARRM, and HRRR with PySAM, and evaluated against monthly plant-level CFs from EIA ~~. Finally, to better diagnose CF discrepancies, we calculate CFs over all California grid points and analyze the spatial distribution of key input variables~~ as well as ~~the large-scale circulation that modulate them~~across the spatial distribution of California grid points.

2.1 Modeling strategy

2.2 ~~Introduction to SCREAM~~

2.1.1 Introduction to SCREAM

The 3.25 km and 800 m SCREAM-CARRMs are built upon SCREAM version 0 (SCREAMv0; Caldwell et al., 2021), developed under the E3SM project (Leung et al., 2020). SCREAM operates at a global horizontal resolution of 3.25 km, reaching convection-permitting scales without parameterizing deep convection. It uses 128 vertical layers on a hybrid coordinate, with vertical grid spacing increasing from 30 m near the surface to approximately 200 m at the boundary layer top, and a model top at roughly 40 km. The nonhydrostatic dynamical core is based on the HOMME (High-Order Methods Modeling Environment) framework (Taylor et al., 2020).

SCREAM's physics suite includes ~~SHOC (Simplified Higher-Order Closure; Bogenschutz and Krueger, 2013)~~ Simplified Higher-Order Closure (SHOC; Bogenschutz and Krueger, 2013) for unified cloud macrophysics, turbulence, and shallow convection; ~~P3 (~~Predicted Particle Properties (P3) microphysics (Morrison and Milbrandt, 2015); the ~~RTE+RRTMGP (~~Radiative Transfer for Energetics and Rapid Radiative Transfer Method for General circulation model applications – Parallel (RTE+RRTMGP) radiative transfer package (Pincus et al., 2019); and the ~~SPA (~~Simple Prescribed Aerosol (SPA) module for aerosol properties (Mahfouz et al., 2025). Surface conditions are based on the E3SM land model (Golaz et al., 2019) and the Los Alamos prescribed CICE4 and data ocean models (Hunke et al., 2008). Land surface fluxes are based on Monin-Obukhov similarity theory, while surface fluxes over ocean are computed using a bulk formulation.

SCREAM Version 1 (SCREAMv1), implemented in C++/Kokkos, has been publicly released and documented (Donahue et al., 2024). RRM capability was not yet available in SCREAMv1 at the time of our simulations, so this study used SCREAMv0. Although GPU resources were unavailable during this study, separate work using similar SCREAM-RRM configurations (Zhang et al., 2025) has demonstrated significant GPU-enabled performance gains.

2.1.2 CARRM configuration

The 3.25 km and 800 m CARRMs generally follow the configurations in Zhang et al. (2024) and Bogenschutz et al. (2024). The 3kmCARRM mesh spans mainland California, adjacent offshore areas, and part of western Nevada. Outside this domain, a 1° global resolution is applied, transitioning into the refined region via eighth-order refinement (Fig. 1a). The 800mCARRM mesh covers the same spatial extent but undergoes refinement to 800 m (Fig. 1b). Maximum Dinv-based element distortions are 3.02 for 3.25 km and 3.05 for 800 m meshes, indicating high-quality grids (Dinv is acceptable if below 4; it quantifies element shape distortion as the ratio of the maximum to minimum singular values obtained from an ~~SVD~~ singular value decomposition (SVD) of the element mapping. See https://acme-climate.atlassian.net/wiki/spaces/DOC/pages/872579110/Running+E3SM+on+New+Atmosphere+Grids, last accessed: June 25, 2025). The land model operates on the same grid as the atmosphere, while the data ocean component uses the ~~MPAS~~ Model for Prediction Across Scales (MPAS) 18-to-6 km mesh (oRRS18to6v3).

The RRM workflow includes grid generation using SQuadGen v1.2.2 (Ullrich and Roesler, 2024), mapping/domain files, topography generation, land surface and land use data, and dry deposition files. Except for regenerating topography for 800mCARRM using higher-resolution digital elevation model (DEM) data, the workflow follows Zhang et al. (2024) and Bogenschutz et al. (2024). Specifically, the 800mCARRM uses the USGS GMTED2010 15-arc-second (500 m) DEM instead of the default GTOPO30 30-arc-second (1 km) DEM (Lauritzen et al., 2015). The 500 m GMTED2010 data were interpolated onto a global 800 m cubed-sphere grid to generate the 800mCARRM topography. The 3kmCARRM topography underwent 12 smoothing iterations, while 800mCARRM used six, following the E3SM Version 3 workflow (https://acme-climate.atlassian.net/wiki/spaces/DOC/pages/2712338924/V3+Topography+GLL+PG2+grids, last accessed: June 25, 2025). Final topographies are shown in Figs. 1c and 1d.

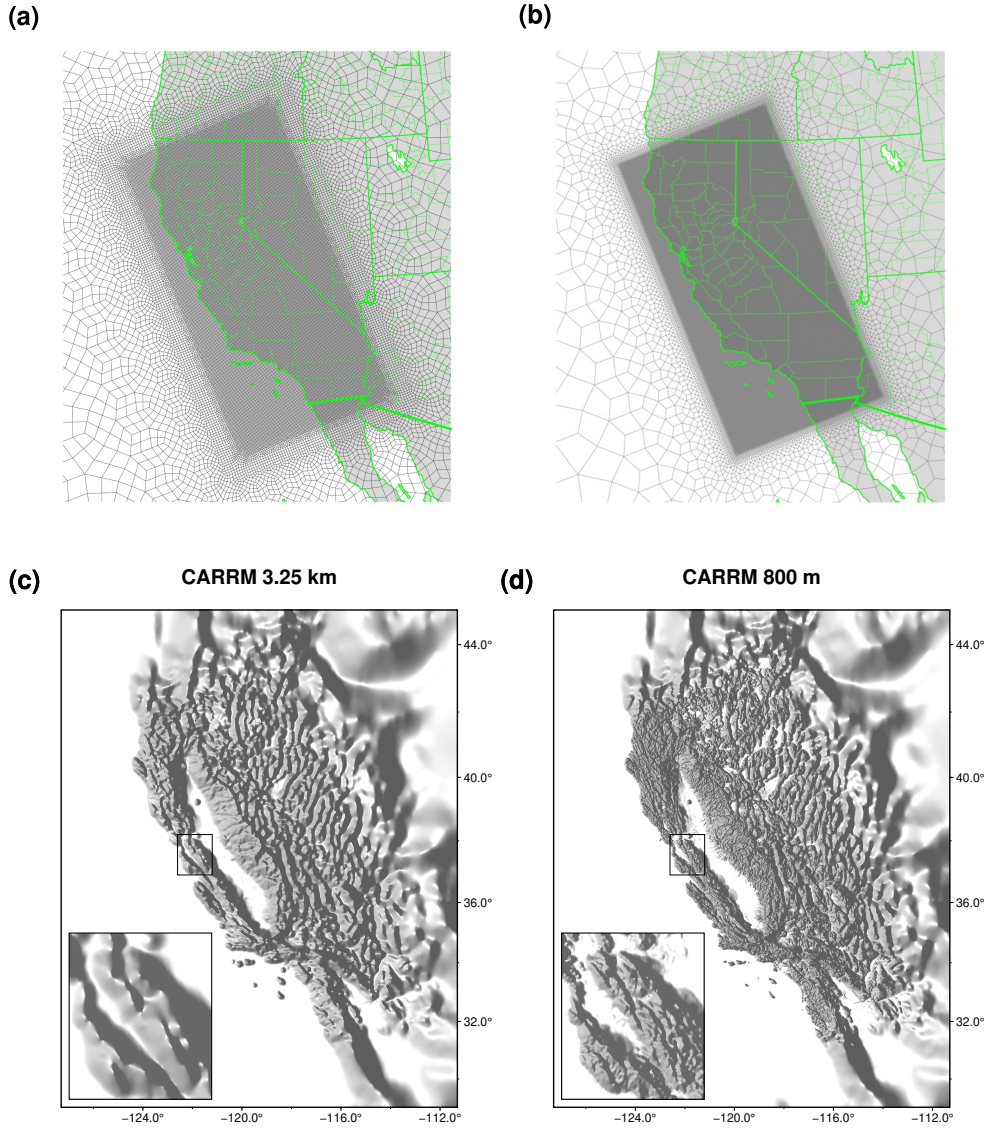

Figure 1. The grid of the (a) 3.25 km and (b) 800 m California regionally refined models. The topography of the (a) 3.25 km and (b) 800 m California regionally refined models, with the small rectangular box over part of the San Francisco Bay Area zoomed in.

Atmosphere initial conditions were generated with the HICCUP package (https://github.com/E3SM-Project/HICCUP, last accessed: June 25, 2025), which automates downloading and interpolating ERA5 pressure-level data. Vertical interpolation is performed by netCDF Operator (NCO) (Zender, 2008), while horizontal interpolation invokes various tools and algorithms through NCO. We employed NCO for linear-in-log(pressure) vertical interpolation and TempestRemap (Ullrich and Taylor, 2015; Ullrich et al., 2016) for high-order horizontal interpolation onto the RRM dynamical core grid. HICCUP also applies surface adjustments following Trenberth et al. (1993) to mitigate spurious gravity wave shocks associated with initial imbal-

ances. Land initial conditions were taken from a well spun-up (1850–2015) 1° E3SMv1 historical simulation at 00:00 UTC on 1 January 2015 (Golaz et al., 2019). Notably, this historical simulation serves as the precursor to the future simulation used to generate nudging data, ensuring internal consistency.

2.1.3 ~~Souce~~ Source of boundary conditions: 1° E3SMv1 simulations

Since SCREAM lacks a scale-aware deep convection scheme, SCREAM-RRM global simulations require nudging outside the high-resolution domain to constrain large-scale flow. Prescribed SST and sea ice concentration boundary conditions are also necessary due to SCREAM's atmosphere-land-only coupling. To ensure consistency across these boundary conditions, nudging variables and SST/sea ice concentration fields were obtained from a coarse-resolution E3SM simulation.

We conducted simulations with two different configurations to provide boundary conditions for the RRM simulations. The first is a rerun of a 1° fully coupled E3SMv1 simulation (Golaz et al., 2019), following the setup of Zhang et al. (2024), to produce 3-hourly fields of U, V, T, and Q (winds, temperature, and humidity) as well as monthly SST and sea ice concentration. Although this simulation ~~nominally~~ follows the SSP5-8.5 pathway (Zheng et al., 2022), our analysis focuses only on the nominal 2015–2020 water years (October–September accounting year used in hydrology) period, where radiative forcings and global climate statistics closely resemble those of the present day (e.g., Kriegler et al., 2017). We refer to this configuration as the "present-day" simulation. In parallel, we conducted a 10-year E3SMv1 simulation ~~with prescribed SST, sea ice concentration, and radiative forcing levels~~ using forcings representative of the 2010s, ~~derived~~ in which SST and sea ice are prescribed from the monthly climatology averaged over 2005–2014. ~~This~~ this type of experiment ~~simulations~~ simulates the atmosphere–land components interactively with prescribed ocean conditions. It is widely used for evaluating model performance against observational data. We refer to this configuration as the "2010climo" simulation.

Nudging data ~~were interpolated~~ are interpolated onto the RRM physics grid using NCO's ~~linear-in-log(pressure)~~ linear-in-log-pressure vertical interpolation and TempestRemap's high-order horizontal interpolation~~onto the RRM physics grid~~. U, V, T, and Q ~~Nudging with a relaxation timescale of 2d was applied using the~~ nudging with a 2-day relaxation timescale is applied following (Zhang et al., 2024). Regional nudging in E3SM ~~regional configuration with~~ employs a Heaviside window ~~to constrain the horizontal domain (Tang et al., 2019). The nudging strength was nearly uniform across vertical levels, with reduced strength in the lowest and highest model levels. This allows~~ that is zero within the refined region and transitions toward one in the surrounding coarse-resolution domain (e.g., Tang et al., 2019), allowing free-running conditions over California ~~while nudging~~ and applying nudging in the coarser-resolution domain~~with a two-day relaxation timescale (Zhang et al., 2024)~~. Vertically, the default Heaviside configuration is used, with weaker nudging at the lowest and highest model levels and full strength in between.

SSTs and sea ice concentrations were processed using the e3sm_to_cmip tool (https://github.com/E3SM-Project/e3sm_to_cmip, last accessed: June 25, 2025), converting MPAS outputs to 1° lat-lon time series. Missing land values were filled via Poisson equation relaxation before formatting into Data Ocean stream files (https://esmci.github.io/cime/versions/ufs_release_v1.1/html/data_models/data-ocean.html, last accessed: June 25, 2025).

Table 2. Column numbers, timesteps, and throughputs of SCREAMv0 global 3.25 km simulation, 3kmCARRM, and 800mCARRM.

Model	Column no.		Time step (s)		Throughput		
	Dynamics	Physics	Dynamics	Physics	Machine	Cores (Nodes)	Simulated Day Per Day
SCREAMv0	56,623,106	25,165,824	9.375	75	cori-knl	104448 (1536)	4-5
SCREAMv0-3kmCARRM	152,712	67,872	9.375	75	dane	1120 (10)	144
SCREAMv0-800mCARRM	146,976	587,904	2.34375	75	quartz	6480 (180)	12.5
					dane	6720 (60)	19.3

2.1.4 Timestep and computational cost

The 3kmCARRM grid (67,872 physics columns; Zhang et al., 2024) comprises only 0.27% of the global 3.25 km SCREAM physics columns (25,165,824 physics columns; Caldwell et al., 2021), while the 800mCARRM grid (587,904 columns; Bogenschutz et al., 2024) represents 2.3%, significantly reducing storage and computational costs compared to a global *km*-scale model.

All SCREAM-CARRM simulations were conducted on the Livermore Computing (LC) clusters Ruby, Dane, and Quartz. Ruby nodes feature 56 Intel(R) Xeon(R) Platinum 8276 CPUs @ 2.20 GHz; Dane nodes feature 112 Intel(R) Xeon(R) Platinum 8480+ CPUs @ 2.0 GHz; and Quartz nodes are equipped with 36 Intel(R) Xeon(R) E5-2695 v4 CPUs @ 2.10 GHz. For SCREAM-3kmCARRM 2010climo and present-day simulations, 10 Dane nodes were used. The SCREAM-800mCARRM present-day simulation, which incurred significantly higher computational costs, was executed in two phases due to machine and project transitions: 180 Quartz nodes were used from 1 October 2015 to 4 January 2017, and 60 Dane nodes from 4 January 2017 to 1 October 2020. All simulations used MPI-only processes, with no OpenMP threads.

The 3kmCARRM simulations advanced at a pace of approximately 144 simulated days per wall-clock day (including I/O overhead), running one simulation month per job submission. The 800mCARRM simulation was conducted in 10-day chunks, achieving throughputs of about 12.5 simulated days per day on 180 Quartz nodes, and about 19.3 simulated days per day on 60 Dane nodes. During the ~~5-yr~~ 5-year SCREAM-3kmCARRM present-day simulation, four "EOS (equation of state) bad state" failures occurred; five such failures occurred during the 10-year 2010climo simulation. These errors arise when the model's thermodynamic variables violate EOS, which governs the physical relationship among pressure, temperature, and density. No EOS failures were seen in the 800mCARRM simulation. Each EOS error was resolved by temporarily halving the dynamical core timestep for one job cycle before reverting to the default timestep. All simulation outputs have been archived.

Table 2 compares the column counts, timesteps, and throughputs between the 3.25 km and 800 m SCREAMv0-CARRMs and the global SCREAMv0 3.25 km simulation in Caldwell et al. (2021).

2.2 Evaluation strategy

This section introduces the evaluation strategy for modeled CFs. The plant-level EIA monthly data were first processed by reading raw tables, applying quality control, filtering valid plants, and aggregating them to model grids. Wind and solar CFs were then computed by coupling model outputs with EIA plant locations within our PySAM preprocessing pipeline. Additional simulation datasets and PLUSWIND-provided CFs beyond the SCREAM and E3SM simulations are also introduced. Finally, to better diagnose CF discrepancies and provide a broader perspective, we calculate CFs over all California grid points and analyze the spatial distribution of key input variables as well as the large-scale circulation that modulate them.

2.2.1 EIA monthly CF data

EIA provides plant-level electricity generation data through monthly generation records from EIA–923 (https://www.eia.gov/electricity/data/eia923/, last accessed: June 25, 2025) and generator-specific metadata for plants with nameplate capacities over 1 megawatt from EIA–860 (https://www.eia.gov/electricity/data/eia860/, last accessed: June 25, 2025).

The ~~eia~~EIA–processor package extracts and preprocesses EIA data to produce monthly wind and solar CFs for 2013–~~2023~~ 2024 (Lee et al., 2025). It reads generation values from the "Netgen" column in EIA–923 and retrieves plant attributes ("Plant Code", "Plant Name", "State", "County", "Latitude", "Longitude", "Nameplate Capacity", "Technology", and "Axis Tracking") from EIA–860. Initial quality control and duplicate removal were incorporated during EIA preprocessing.

The EIA data first underwent the following ~~steps of~~ plant-level quality control steps:

1. ~~Retained only~~ Only plants with a nameplate capacity of ~~no less than~~ at least 2.5 MW were included.
2. ~~Retained only~~ Only plants located in the state of California were included.
3. For solar generation, ~~retained only power plants with the technology type~~ only plants classified as "Solar Photovoltaic" ~~(which account~~ were retained (accounting for 97.3% of all solar facilities), ~~excluding those classified~~ while those categorized as "Solar Thermal without Energy Storage" were excluded. In this study, ~~when we refer to~~ the term *solar* ~~, we specifically mean~~ refers specifically to ~~*photovoltaic (*~~*PV (photovoltaic)* power.
4. For solar PV ~~, retained only plants with an axis-tracking entry~~ plants, only those with a reported array configuration in the EIA ~~data (three types are reported: fixed, single-axis, and dual-axis)~~dataset were retained, as indicated by at least one of the fields "Single-Axis Tracking?", "Dual-Axis Tracking?", or "Fixed Tilt?".
5. ~~Retained only plants with CF exceeding~~ Plants were excluded if CF was below 5% for ~~at least three months and with fewer~~ more than nine months or if more than 12 months of ~~missing data~~ data were missing.

These quality control steps resulted in 72 valid wind plants and 322 PV plants.

To align with the model resolution, the geographical locations (latitude and longitude) of individual plants reported in the EIA data were aggregated to the target model grid. For the 3kmCARRM grid, this process aggregated all valid wind plants

and PV plants into 32 and 222 model grid points, respectively. In the subsequent analysis, wind plants are clustered based on geographical location, while PV plants are clustered by axis-tracking type.

We adopted a gridcell-based evaluation framework instead of direct plant in situ evaluation for the following reasons: 1) It avoids artificially reducing the spread of the modeled climatological seasonal cycle. In plant-level evaluations, multiple plants located within the same model grid cell would be represented identically. This would substantially reduce the apparent monthly climatological spread and make modeled variability appear artificially smaller than that from observations. 2) Aggregating to the model grid ensures consistency between plant-level evaluations and assessments of spatial patterns across California, as both are conducted on the same grid. This facilitates a coherent transition between statewide spatial perspectives and local perspectives. In the seasonal-cycle analysis, we applied a weighted average based on plant nameplate capacity within each grid cell, so that the EIA mean line remains consistent across comparisons at different model resolutions.

For wind plants, because all quality control is applied at the plant level, and the clustering logic does not depend on plant-specific attributes, aggregating to different model grids does not affect the total nameplate capacity of each cluster. ~~However, for~~ For PV plants, ~~the axis-tracking type of each~~ we group samples by both model grid point ~~is calculated as the mean of fixed (0.0),~~ and axis-tracking type. For example, if a grid point G contains three single-axis ~~(1.0), and dual-axis (2.0) values of all plants within that grid point. During clustering, a model grid point is assigned to a specific~~ plants and one fixed-axis plant, two groups are formed: G_1 (single-axis) and G_0 (fixed). The seasonal cycles and evaluation metrics are then computed by clustering according to axis-tracking ~~type if the absolute difference between its averaged axis-tracking value and fixed (0.0), single-axis (1.0) , and dual-axis (2.0)is less than 0.1. Because the proportion of axis-tracking types within a grid point vary slightly with grid size,~~ type. This approach preserves the plant-level tracking type information reported in the EIA dataset, and the total nameplate capacity within each ~~PV cluster also vary slightly~~ cluster does not vary with model resolution. As will be shown in the results section, the axis-tracking type has a substantial influence on PV power generation.

Finally, we applied an "independent-plant ratio > 15%" threshold to all valid EIA plants. The independent-plant ratio is defined as the fraction of plants in a given year whose monthly CF time series are not perfectly correlated with those of other plants (identified using a Pearson correlation distance threshold of 10^{-6}). This was motivated by the fact that the EIA data contains many plants that only report annual values without providing monthly records, for which seasonal variability is scaled from other plants. The background of this issue ~~is a notable story (please refer to Appendix A for details)~~and the investigation that led to this finding are described in Appendix A. Applying this threshold removes years heavily affected by scaled plants, correcting the artificial reversal of the seasonal phase – compared to all models – at the Shasta wind plant and adjusting seasonal peaks in northern California, while having minimal impact elsewhere.

2.2.2 PLUSWIND-derived CFs

PLUSWIND (Millstein et al., 2023) provides hourly, plant-level generation estimates across the U.S. from 2018 to 2021, using meteorological inputs from three models: ERA5, MERRA2, and HRRR. They use hub-height wind speeds from HRRR, model-level wind speeds from ERA5, and near-surface wind speeds from MERRA2 to estimate wind power generation.

To improve plant geolocation, PLUSWIND computed wind plant centroids from the U.S. Wind Turbine Database (USWTDB; Rand et al., 2020) rather than relying on the self-reported EIA–860 coordinates. Each plant was matched to a representative wind turbine power curve based on key assumptions: 1) a cut-in speed of 3 m/s, 2) a cut-out speed of 25 m/s, and 3) a rated speed between 9-13 m/s, derived from the plant's average specific power, which was calculated from turbine diameter and capacity. Hub-height wind speeds were estimated differently across models: for ERA5, by vertical interpolation from model levels; for MERRA2, by applying a log wind profile using 2 m, 10 m, and 50 m wind speeds; and for HRRR, directly without additional correction. For repowered plants, data from the repowering year were excluded, and turbine characteristics were updated to the most recent configuration for all years. Air density correction for all models (ERA5, MERRA2, HRRR) was based on MERRA2 air density, interpolated to hourly resolution at hub height. Wake losses were estimated using a simple method: a constant 7% loss was applied when wind speeds fall below rated speed minus 0.5 m/s, tapering linearly to zero at rated speed plus 2 m/s.

2.2.3 Coupling with PySAM to estimate wind and solar CFs

All SCREAM-CARRM simulations output hourly meteorological variables, including surface pressure, 2 m temperature, 2 m relative humidity (used with temperature to derive the dew point), downward shortwave radiation (total and diffuse components), hub-height winds at 90 m, 10 m winds, and snow water equivalent. Together with static surface geopotential height, these variables serve as inputs for CF calculations through PySAM. HRRR provides the same set of hourly variables, except that hub-height wind is reported at 80 m instead of 90 m.

Developed by the National Renewable Energy Laboratory (NREL), the System Advisor Model (SAM) (https://sam.nrel.gov/, last accessed: June 25, 2025) provides a comprehensive set of algorithms and assumptions for calculating electricity generation and approximating CFs (i.e., actual generation divided by nameplate capacity). PySAM is the Python wrapper of SAM (https://github.com/NREL/pysam, last accessed: June 25, 2025).

In this study, we used our preprocessing ~~preprocessing~~ pipeline developed for PySAM (NREL) to calculate hourly generation outputs: wind power was estimated using *PySAM.Windpower.default("WindPowerNone")* and solar PV power using *PySAM.Pvsamv1.default("FlatPlatePVNone")*.

~~These calculations were based on several key assumptions: 1)~~ The default parameter set is used for *PySAM.Windpower.default("WindPow* including the use of a generic turbine power curve without plant- or manufacturer-specific corrections, ~~neglecting~~ the neglect of variations in turbine physical characteristics such as rotor diameter and rated capacity, ~~2)~~ and the use of ~~default wake effect model, 3) the use of~~ fixed 90 m hub-height wind inputs without turbine-specific height correction. The default wake calculation includes the simple wake model (https://samrepo.nrelcloud.org/help/wind_power.html, last accessed: February 20, ~~4)the use of a~~ 2026), together with an external wake loss of 1.1%. The simple wake model computes distances between downwind and crosswind turbines based on their relative positions within a wind farm and accounts for the effects of power coefficient, thrust coefficient, and turbulence intensity on wind speed. The default parameter set specifies relative x and y coordinates for 32 turbines within a wind farm. Because the primary focus of this study is model evaluation and the complexity of meteorological datasets, we chose to follow the default settings in PySAM to maintain focus.

The *PySAM.Pvsamv1.default("FlatPlatePVNone")* configuration is a detailed PV model that includes separate ~~models for~~ representations of the module, inverter, and cell temperature. The default use of nominal operating cell temperature (NOCT) model (Gilman et al., 2018a) relies on effective irradiance transmitted to the cell, air temperature, and near-surface wind speed. In addition to the default fixed-axis tracking configuration, single- and dual-axis tracking configurations are enabled by setting *SystemDesign.subarray1_track_mode*. The default option None is used for subarray shading, meaning no self-shading or backtracking is applied to avoid row shading.

Required inputs to PySAM include:

1. Locations (latitude, longitude) of power plants within California reported by EIA;

2. Hourly simulation outputs reformatted into a full-year (365-day calendar) 8,760-hour time series, with each variable stored in a separate file. Model grid points corresponding to plant locations were identified using the BallTree class from the sklearn.neighbors module (https://scikit-learn.org/1.6/, last accessed: June 25, 2025);

3. For solar PV, axis tracking types (fixed, single-axis, dual-axis) reported by EIA were assigned by setting *SystemDesign.subarray1_track_mode* to 0, 1, or 2, respectively, across all sub-arrays.

Note that the EIA data are reported after curtailment, whereas our generation modeling do not include this effect. As a reference, Lee et al. (2025) applied a curtailment correction of 2% for wind and 3% for solar CFs over California.

Executing our PySAM preprocessing pipeline generated hourly CFs for each power plant across all SCREAM-CARRM simulations, HRRR, and E3SM-25kmNARRM.

2.2.4 HRRR

Inspired by the PLUSWIND dataset, we evaluate the HRRR (Dowell et al., 2022) as an additional source of meteorological data to drive PySAM, given that its horizontal and boundary-layer vertical resolutions are comparable to SCREAM-CARRMs. HRRR assimilated conventional observations and radar reflectivity at each forecast cycle and remains the highest-resolution historical forecast dataset over North America.

We used HRRR +1 hour forecasts from 2018 to 2022, substituting +2 or +3 hour analysis when +1 hour data are unavailable. Despite model biases and PySAM assumption errors, HRRR is expected to perform well due to its frequent data assimilation. In fact, HRRR is widely used in the wind energy industry (Shaw et al., 2019). Thus, HRRR provides another reference for assessing SCREAM-RRM performance in electricity generation. Among the three SCREAM-CARRM simulations, the 3kmCARRM 2010climo experiment design aligns most closely with HRRR, offering comparable resolution and similar climatology. Key differences arise from model structure, data assimilation, and internal variability (i.e., 2010climo being a 10-year simulation with the 2010s SST ~~, versus~~ vs. HRRR as a 5-year simulation driven by evolving SST from 2018 to 2022).

We also compare HRRR-modeled wind CFs from PLUSWIND with our PySAM-derived results to isolate sensitivities to wind power modeling assumptions. These sensitivities are further compared with sensitivities to model structure (HRRR vs. SCREAM-3kmCARRM 2010climo), and to resolution (HRRR vs. ERA5/MERRA2 in PLUSWIND; SCREAM-3kmCARRM vs. SCREAM-800mCARRM vs. E3SM-25kmNARRM).

2.2.5 Statewide CFs, meteorological drivers, and large-scale circulation

To better understand the differences in CFs derived from different simulations, we extend our analysis beyond EIA plant locations and estimate CFs at all grid points across California using PySAM. This provides a comprehensive spatial view of wind and solar resource availability within the state and enables direct comparison with local topography and meteorological drivers. We present the spatial distributions of variables directly used in PySAM calculations over California including: 2 m temperature~~and dew point~~, 90 m hub-height winds, and total downwelling solar flux at surface, alongside global-scale atmospheric circulation patterns. ERA5 is used as a reference to evaluate biases in both the regional meteorological fields and the large-scale circulation. However, some ERA5 variables (such as hub-height wind speeds) are less reliable for evaluating *km*-scale models as they are strongly affected by topography features (e.g., Millstein et al., 2022, 2023).

3 Results

This section begins by discussing some of the data limitations and clarifies our interpretation of the comparison results given these obstacles. We also outline the structure and rationale of our evaluation groups. We then present the results sequentially, organized under two main categories: wind and solar.

3.1 Uncertainty in EIA monthly data and interpretation of discrepancies

A fundamental issue with the monthly EIA data is that the procedures used for collecting and processing are not publicly known, particularly how missing data, outages, curtailments, and obvious anomalies are handled during temporal averaging (see similar discussion in Ryberg et al. (2019)). Another limitation is the issue of non-independent samples related to scaled plants, as discussed in "Methods – EIA monthly CF data" section and Appendix A. Therefore, we cannot assert that generation estimates from model outputs are inherently erroneous, nor can we determine the true degree of error. The difference between model-simulated and EIA-reported CFs should be viewed not as a conventional *bias*, but rather as a *discrepancy*.

~~The~~ In principle, the most accurate model evaluation data would ideally come from on-site meteorological towers or remote sensing systems, such as lidar or meteorological masts deployed by wind plant developers, which provide high-quality wind speed observations. However, such data are expensive, technically challenging to obtain, and not publicly available (Jourdier et al., 2023; Pronk et al., 2022). ~~A more practical proxy observation is high-frequency (e.g., hourly) generation data. A few utilities release these as part of their market-facing datasets (Davidson and Millstein, 2022). While still affected by conversion assumptionsfrom wind to power, hourly data contain~~ Widely available meteorological station data are typically limited to near-surface (10-m) observations, which are insufficient for direct wind power evaluation. Previous studies have shown that interannual variability in wind power generation is only weakly correlated with observed near-surface wind speed (Millstein et al., 2022), motivating PLUSWIND to evaluate HRRR using EIA generation data rather than surface wind observations. Extrapolating 10-m winds to hub height requires power-law assumptions, whose exponents depend on atmospheric stability and surface roughness and are known to perform poorly under stable conditions (Frank et al., 2020). This is particularly relevant

in California, where wind power generation tends to be stronger at night (Millstein et al., 2023) when atmospheric stability is higher.

An alternative and more direct observational benchmark would be high-frequency (e.g., hourly) plant-level generation data, which retain diurnal and weather-scale variability and can better capture generation losses ~~due to~~ associated with disruptive meteorological or hydrological events (Davidson and Millstein, 2022). Unfortunately, ~~we lack access to such hourly, plant-level data for California. Therefore, we caution readers against interpreting~~ such data are not publicly available for California at the plant level. Monthly CFs, while inevitably smoothing sub-hourly and synoptic-scale variability, therefore remain the only feasible observational constraint for statewide evaluation.

Given these limitations, the use of monthly EIA data in this study reflects a trade-off between observational fidelity and data availability. We therefore caution against attributing discrepancies between modeled generation and ~~monthly EIA values as purely resulting from model inaccuracies~~ EIA monthly CFs solely to deficiencies in either the atmospheric simulations or the power generation modeling. Instead, these discrepancies should be interpreted in light of the combined uncertainties inherent in both the observational data and the meteorological ~~whether in the meteorological simulation or generation process.~~ generation modeling modeling framework. This is why we also include comparisons with PLUSWIND and HRRR to help assess the relative importance of different uncertainty sources.

3.2 Evaluation ~~group~~ design and ~~key questions~~structure

~~To answer~~ The Results section is organized to address the three guiding questions introduced in the Introduction ~~, our analysis is organized around three comparisonsdesigned to isolate key sources of uncertainty~~through a set of controlled comparisons. All CFs are calculated using a consistent PySAM configuration unless otherwise noted. Because discrepancies relative to EIA reflect combined uncertainties from meteorological datasets, power-generation modeling, and EIA data quality, additional benchmarks are introduced to help contextualize the relative importance of these uncertainty sources.

For both wind and solar energy, the analysis follows a common structure in general:

1. *Generation* ~~*modeling assumptions:HRRR with PLUSWIND vs. HRRR with PySAM. Outside of this section, all HRRR-based CF calculations will utilize our results from PySAM*~~ *model dependence (Q1).* For wind, we compare CFs derived from HRRR using PLUSWIND and PySAM to assess whether generation model choices materially affect CF estimates. For solar PV, we quantify the sensitivity of modeled CFs to axis-tracking assumptions within PySAM. These tests serve as a validation of PySAM before inter-comparing meteorological datasets.

2. *Meteorological* ~~*model structure:*~~*dataset dependence at comparable resolution (Q2).* We compare SCREAM-3kmCARRM (2010climo~~vs. HRRR. Both have similar horizontal resolution and are forced with observed SSTs,~~) against HRRR, both of which provide km-scale meteorological fields for PySAM but differ in ~~physics, data assimilation, and internal variability.~~ model framework. This comparison quantifies the magnitude of CF differences attributable to the meteorological model when horizontal resolution is not the primary differentiator.

3. ~~*Horizontal resolution:* E3SM-25kmNARRM vs. SCREAM-3kmCARRM~~ *Horizontal resolution sensitivity (Q3).* We compare present-day ~~vs. SCREAM-800mCARRM present-day. The two~~ SCREAM-CARRM simulations ~~share identical physics, forcing, and nudging; resolution is the only difference~~at 3.25 km with the E3SM-25kmNARRM simulation to assess whether km-scale configurations substantially reduce the pronounced wind CF biases identified at 25 km resolution over California. Sensitivity to further refinement of resolution is then examined by comparing the 3.25 km and 800 m CARRM simulations.

~~Each following subsection presents results by energy type (wind or solar), with each category further broken down by the evaluation questions~~Within each comparison group, results are presented separately for wind and solar. For each ~~group, we~~ energy type, we first show multi-year monthly ~~CFs,~~ CF climatology aggregated by plant ~~cluster (geographic for wind; axis-tracking for solar PV). After the grouped comparisons, we present additional illustrative figures, including spatial distributions of statewide CFs, key meteorological input variables used in PySAM(e.g., hub-height wind speeds, 2 m temperature and dew point, total downward shortwave radiation), large-scale patterns, and diurnal cycles. Each main category (wind or solar)~~ clusters, followed by diagnostics that help interpret CF differences, including diurnal cycle of simulated CFs, statewide spatial patterns of CFs and key meteorological driver variables used by PySAM. Each energy-type section concludes with a summary ~~figure showing the~~ metric (mean absolute error~~to synthesize key findings~~ , MAE) to synthesize the main findings across meteorological datasets.

3.3 Wind

3.3.1 ~~How much do generation modeling assumptions impact~~ Do the ~~wind~~ use of different energy models affect wind CFs?

Figure 2 shows our selected groupings for wind plants. We clustered geographically due to strong dependence of wind simulation on topographic features and surface type. Wind resources at each location are driven by multiscale atmospheric dynamics. Large-scale circulation, synoptic patterns, and geopotential gradients drive synoptic flows that are influenced by coastal proximity. At smaller sclaes, topography and land-surface heterogeneity govern mesoscale to microscale effects, such as moutain-valley circulations and buoyancy-driven flows. Additionally, winds are influenced by heterogeneous surface roughness that includes different vegetation types (e.g., grassland, forest canopy), urban areas, and water surfaces.

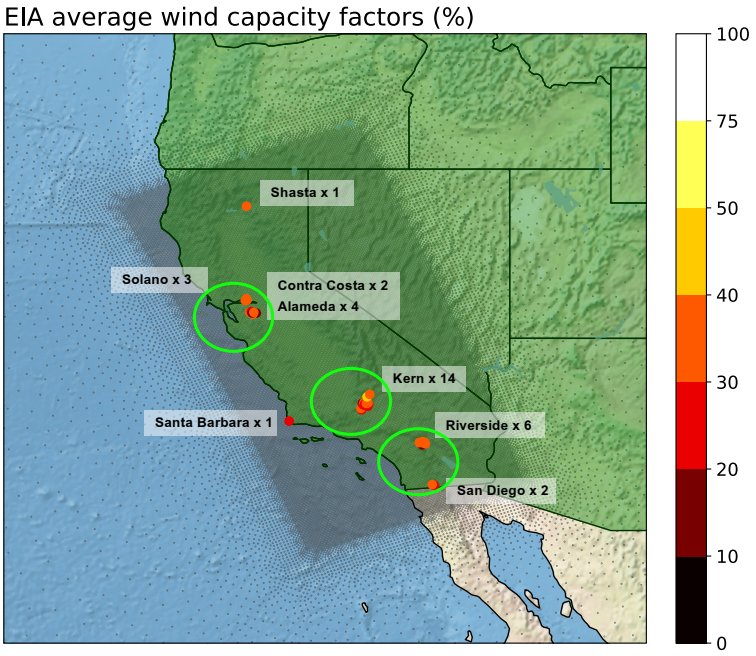

Figure 2. Clusters of wind plants across California, grouped by geographic proximity. Each cluster is marked with a green ellipse and annotated with the county name and number of plants. EIA plants are aggregated to the 3kmCARRM model grid; each filled dot represents the mean wind CF (%) from valid monthly EIA records between 2013 and ~~2023.~~2024.

We define three plant groups in California: 1) "NorCal", including 3 grid points (counted based on the 3kmCARRM model grid) in Solano County, 3 in Alameda County, and 2 in Contra Costa County, with a total installed capacity of 928.1 MW, 2) "Kern" grid points as a standalone category, given its large sample size (12 grid points) and a total installed capacity of 3,322.5 MW, and 3) "SoCal", including 4 grid points in Riverside County and 2 in San Diego County, with a total installed capacity of 696.8 MW. Each cluster is weighted by the nameplate capacity summed across all plants assigned to each model grid point, ensuring integral conservation.

The northernmost part of California contains only one power plant – Shasta. The Shasta plant is part of the Hatchet Ridge Wind Project, which consists of 44 turbines with a total nameplate capacity of 101.2 MW. In our analysis, we found that Shasta exhibited a distinct wind CF seasonal pattern: it peaks in summer, in phase with the rest of California in the EIA monthly records, but opposite to all modeled patterns. The analysis of Shasta confirmed the scaled-plants issue in the EIA data (see "Methods – EIA monthly CF data" section and Appendix A for details). By excluding years with too few independent samples, the artificial seasonal cycle at Shasta was corrected (Fig. A3a). However, given the small sample size ($N = 1$) relative to other regions, we exclude ~~this~~ the Shasta site from further analysis. The only other single-sample case occurs in Santa Barbara, which passed quality control only for 2024.

Figure 3 presents a comparison of modeled CFs for different generation modeling, including PLUSWIND and PySAM. In PLUSWIND, ERA5 (yellow) yields the lowest CFs, followed by MERRA2 (orange), with both staying below EIA val-

ues throughout the year and across all geographic clusters (see more discussions in Dowell et al., 2022). HRRR (red) in PLUSWIND gives the highest CFs, typically exceeding EIA, especially in Kern and SoCal (by 10–25%). PLUSWIND computes three different generation outputs for each model: raw estimates (dotted), density-adjusted (dashed), and density + loss adjusted (solid). The impact of these assumptions are secondary to the choice of meteorological model, highlighting that the generation modeling assumptions in PLUSWIND are a minor source of uncertainty.

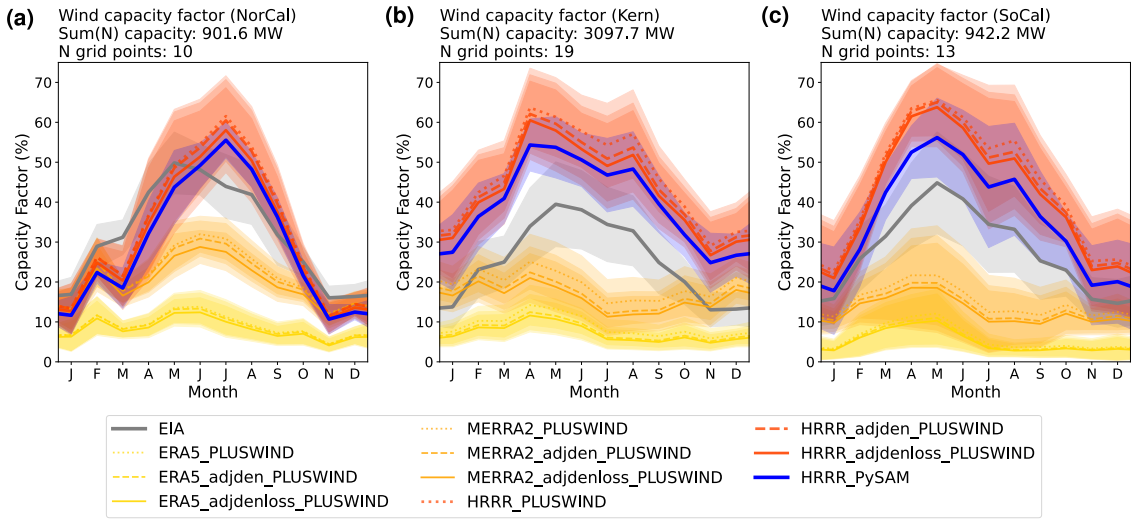

Figure 3. Seasonal variation in wind CFs (%) estimated using different generation modeling options, including PLUSWIND and PySAM. Each subplot corresponds to a cluster in Fig. 2. Different meteorological datasets in PLUSWIND are color-coded: ERA5 (yellow), MERRA2 (orange), and HRRR (red). For each dataset, different generation modeling assumptions by PLUSWIND are shown using line styles: default (dotted), air-density adjusted (dashed), and both air-density and loss adjusted (solid). PySAM-estimated generation from HRRR is shown in solid blue, and EIA data is shown in gray. The bold line shows the multi-year climatology, with shading indicating $\pm$ 1 standard deviation.

PySAM-modeled HRRR estimates (solid blue) show similar seasonal patterns to those from PLUSWIND (solid red) but are overall lower by 2-8%, and generally closer to EIA. This suggests that the effect of modeling assumptions (turbine power curves, density estimation, wake losses) between PLUSWIND and PySAM is larger than the effect of internal corrections within PLUSWIND, but smaller than the effect of varying meteorological model ~~inputs~~datasets.

Note that we first tested two plant coordinates when inputting HRRR data into PySAM: one based on EIA-reported plant locations (default), and the other based on PLUSWIND-provided coordinates derived from USWTDB centroids. The resulting CF differences were minimal (not shown). This is likely due to plant aggregation on the model grid. For HRRR and CARRMs, this grid is *km*-scale, which remains considerably coarser than the footprint of an individual plant. Therefore, spatial uncertainty from plant location has negligible impact on modeled generation.

PLUSWIND uses more realistic turbine power curves compared to the uniform default curve in PySAM's *default("WindPowerNone")* configuration. While PySAM supports more detailed, plant-specific wind simulations (e.g., customized turbine parameters and variable hub heights), we used the default configuration to ensure consistency and simplicity across datasets. In addition,

PLUSWIND applies MERRA2 air density correction and simple empirical losses of about 7% to all model datasets. In contrast, PySAM uses near-surface temperature and humidity from each model to derive air density directly, and applies different loss assumptions. This might help explain why PLUSWIND estimates are different from PySAM.

In summary, generation modeling assumptions, especially air density and wake loss estimation, do influence CF outcomes. The differences between PLUSWIND and PySAM are more substantial than the variations introduced by internal assumptions within PLUSWIND itself. In general, PySAM's estimates for California wind generation align more closely with EIA data. Nevertheless, the impact of generation assumptions is minor relative to the much larger discrepancy driven by meteorological inputs from different models.

3.3.2 How ~~much do meteorological models impact the~~ different are wind ~~energy~~CFs between SCREAM-3kmCARRM and HRRR at comparable horizontal resolution?

Figure 4 presents the second evaluation group comparing wind power generation estimates from SCREAM-3kmCARRM 2010climo and HRRR, both processed using PySAM. HRRR, assimilating high-frequency observations, effectively functions as a *km*-scale reanalysis product, with internal variability closely aligned with historical conditions (2018–2022). The 10-year SCREAM-3kmCARRM simulation is forced by a repeating 2010s climatology for SST and sea ice concentration, resulting in weaker interannual variability.

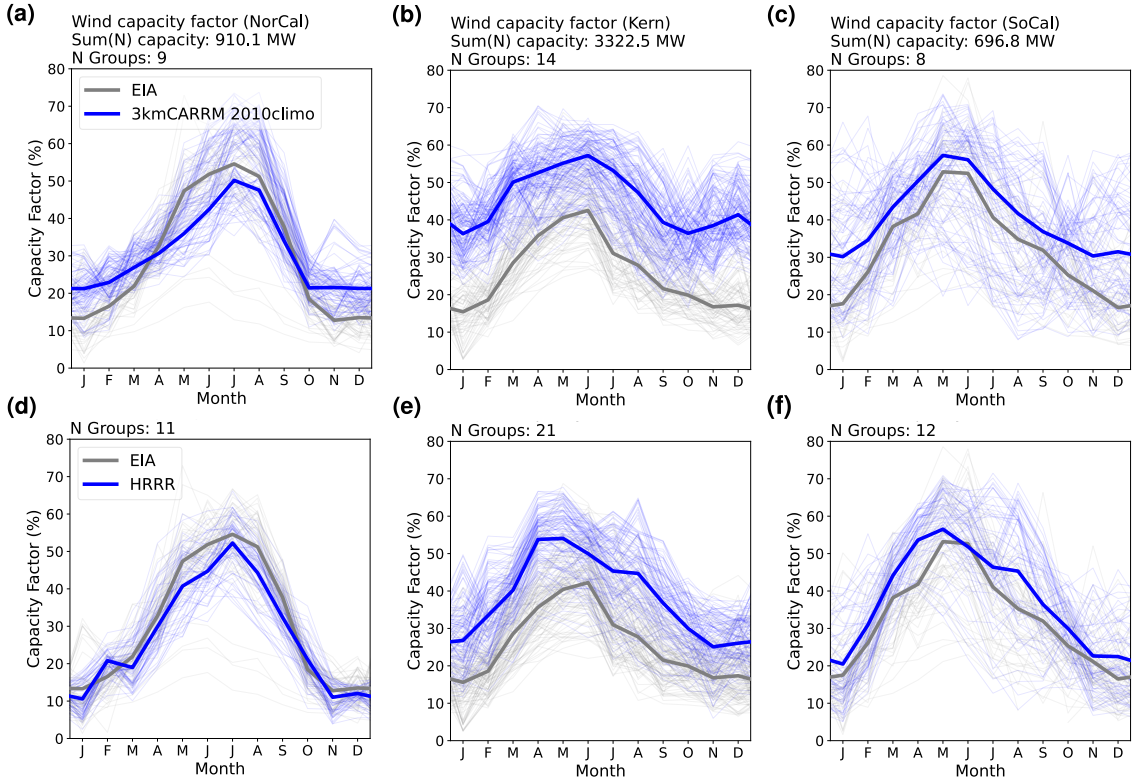

Figure 4. Seasonal variation in wind CFs (%) estimated using PySAM from SCREAM-3kmCARRM 2010climo (top) and HRRR (bottom). Each subplot corresponds to a cluster in Fig. 2. Simulations are shown in blue, and EIA records in gray. Light lines represent individual years; bold lines represent the multi-year climatology.

Overall, HRRR and SCREAM-3kmCARRM yield similar results, with differences substantially smaller than those between HRRR and MERRA2/ERA5 in PLUSWIND (Fig. 3). Considering the structural differences between HRRR and SCREAM – including dynamics, physics parameterizations, and data assimilation/nudging – this suggests that horizontal resolution, rather than model structural details, exerts a first-order control on wind energy simulations. Note that the nominal resolution of 3.25 km refers to SCREAM's dynamical core, while its physical grid has a resolution of ~4.5 km, slightly coarser than HRRR. While this could lead to marginal degradation in results, wind simulations at *km*-scale are likely dominated by resolved processes, with limited sensitivity to turbulence scheme differences. This is consistent with findings from Altamont Pass, where two planetary boundary layer schemes in Weather Research and Forecasting model produced comparable performance in simulating wind conditions and CFs at similar resolution (Arthur et al., 2025b).

In terms of discrepancies between modeled CFs and EIA records across plant clusters, the largest occurs in Kern County. Note that we did not add a curtailment correction to the PySAM outputs; given the 2% curtailment applied by Lee et al. (2025) to California wind energy, this is unlikely to explain the majority of the discrepancies. On multi-year average, HRRR aligns more closely with EIA in this region, with an annual discrepancy of 13%. In contrast, SCREAM-3kmCARRM's discrepancy

in Kern increases from 15% in summer to 25% in winter. A similar seasonal pattern is observed in ~~NorCal~~SoCal, where 3kmCARRM's discrepancy grows from 10% to 15% between summer and winter. In the same region, HRRR's overestimation declines from 10% in summer and fall to only 3% in winter. For ~~SoCal~~NorCal, both models captured the ~~July~~ summer peak well. During winter, 3kmCARRM tends to overestimate by ~~8~~10%, while HRRR underestimates by 4%. SoCal shows the smallest discrepancy (<5%) among all clusters, especially during spring and fall.

In summary, the two fundamentally different meteorological models ~~—~~(SCREAM-3kmCARRM and HRRR~~—~~) produce overall similar wind generation estimates when using comparable horizontal resolution, with HRRR results closer to EIA values. Given HRRR's nature as a reanalysis-like product, its lower discrepancy might be expected. ~~The contrast between this section and the previous suggests suggests that horizontal resolution is a first-order control on wind power simulation, despite large differences in model structure and internal variability.~~ Additionally, we find that 3kmCARRM consistently overestimates winter wind generation relative to summer, whereas HRRR does not exhibit this pattern. This may indicate a systematic wintertime wind bias in SCREAM-CARRMs.

3.3.3 How much does horizontal resolution affect ~~the~~ wind energy?

Figure 5 presents California wind ~~power~~ energy estimates from E3SM-25kmNARRM, as well as the 3.25 km and 800 m SCREAM-CARRM present-day simulations, all processed using PySAM. The only difference between the two SCREAM-CARRM simulations is horizontal resolution. SCREAM-CARRMs and E3SM-25kmNARRM differ in several aspects, including resolution, physical parameterizations, and internal variability.

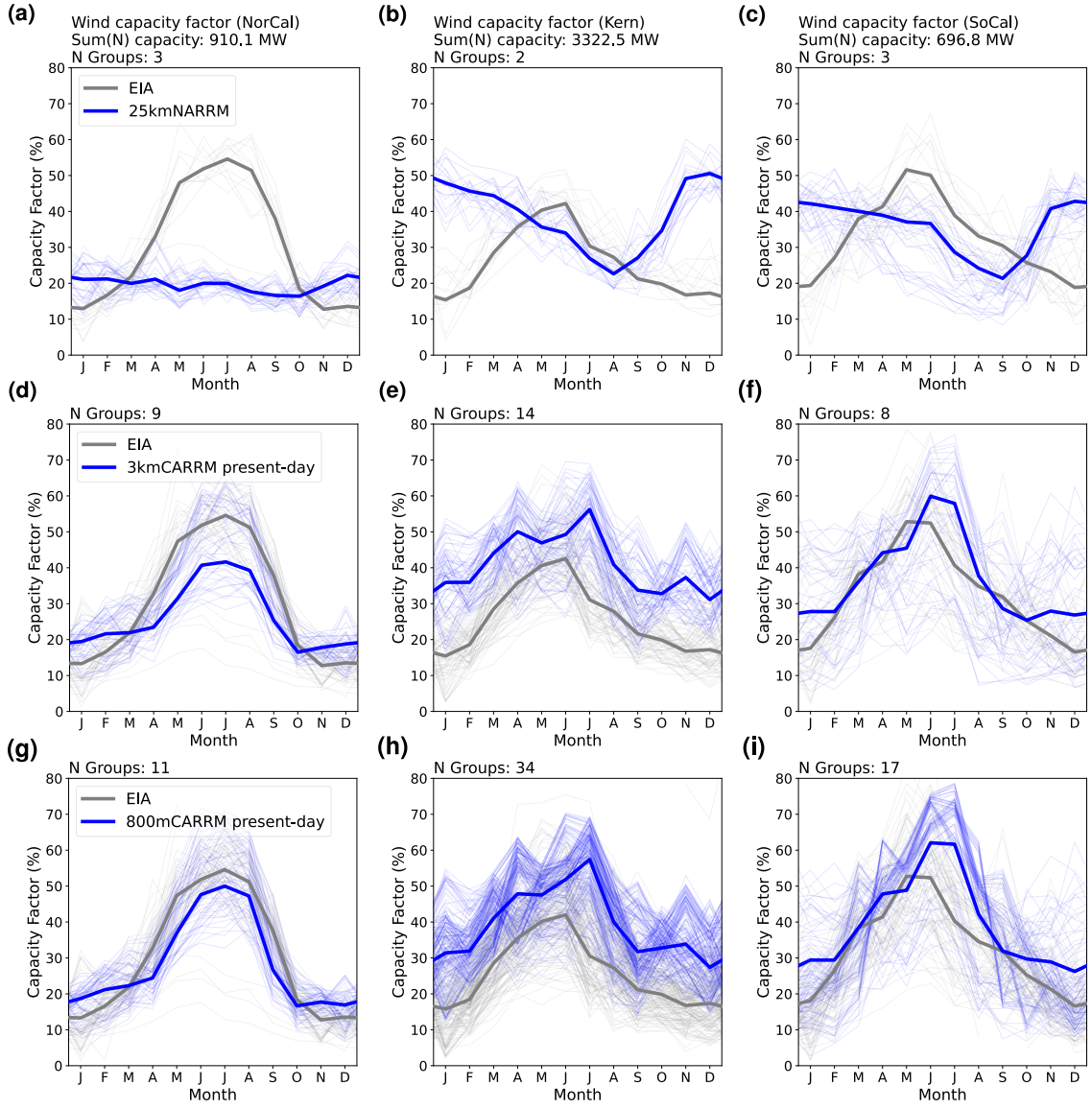

Figure 5. As in Fig. 4, but comparing ~~the E3SM-25kmNARRM 2005~~(a)–~~2014 simulation~~ (c) E3SM-25kmNARRM simulations with ~~the~~ (d)–(f) 3.25 km and (g)–(i) 800 m SCREAM-CARRM present-day simulations.

Strikingly, the E3SM-25kmNARRM simulation shows either a near-reversal or a strong damping of seasonal variability in wind generation across California. This pattern was first identified by Lee et al. (2025). In contrast, the SCREAM-3kmCARRM simulation captures the EIA seasonal phase much more accurately, with clear improvements particularly in seasonal timing. Since all EIA processing, PySAM configuration, and data filtering were held constant in our analysis, the differences seen here can be attributed solely to the meteorological models. Resolution plays a far more dominant role than large-scale circulation, as shown in the next section, where similar large-scale patterns still lead to substantial differences due to resolution.

For the 3.25 km vs. 800 m SCREAM-CARRM simulations, the modeled wind CFs differ only slightly. In Kern and SoCal, wintertime generation from SCREAM-800mCARRM is closer to EIA, with discrepancies reduced by 5% compared to the 15–20% overestimation from SCREAM-3kmCARRM. However, during summer (particularly in July) overestimation increases with resolution across NorCal, Kern, and SoCal. This suggests that a 4× increase in resolution yields only modest sensitivity in SCREAM wind generation estimates, with no consistent improvement in the match to EIA, as quantified in the subsequent mean absolute error plot (Fig. 9). Note that both simulations are based on present-day forcing, not observed SST and sea ice concentration, so their discrepancies relative to EIA are larger than those in Fig. 4. Therefore, the discrepancy magnitude in this group reflects not only resolution-related biases. It also captures mismatches between the 1° E3SM present-day simulation (2015–2020 water years under CMIP6 SSP5-8.5 forcing) and either the observed 2010s climatology or the actual interannual variability from the empirical years within 2013–~~2023.~~ 2024.

In summary, moving from 25 km to 3.25 km results in a qualitative leap, especially for the seasonal cycle, while the resolution sensitivity from 3.25 km to 800 m appears minor overall in SCREAM. The contrast between this section and and the previous suggests that horizontal resolution is a first-order control on wind energy simulation, despite large differences in model structure and internal variability.

3.3.4 ~~Statewide~~ Diurnal cycle of simulated wind ~~CFs, meteorological factors, and large-Scale circulation~~power generation

Figure 6 shows the diurnal cycle in simulated wind CFs. Note that EIA data are not included, as hourly EIA records are not available. Each panel displays average seasonal (horizontal, month) and diurnal (vertical, time PST) variation, grouped by geographic cluster.

First, all simulations exhibit highly consistent diurnal cycles, with higher values at night and lower during the day. E3SM-25kmNARRM appears highly anomalous, mainly due to the previously discussed seasonal phase reversal and significantly underestimated summer generation. Nevertheless, its diurnal timing remains relatively consistent with the other simulations. All models simulated peak values occur at approximately 21–24 PST, while minimum values occur near 12–13 PST. Second, in the two comparisons with similar internal variability, diurnal cycle differences are substantially smaller than seasonal cycle differences. The seasonal cycles in Figure 6 represent an alternate visualization of those in Figs. 3–5.

Overall, the inter-model differences in the diurnal cycle of wind power generation are smaller than the differences between regions (i.e., between northern and southern regions). The wind generation peak period in NorCal, located in north-central California, is slightly shorter than the peak duration observed in the southern regions of Kern and SoCal.

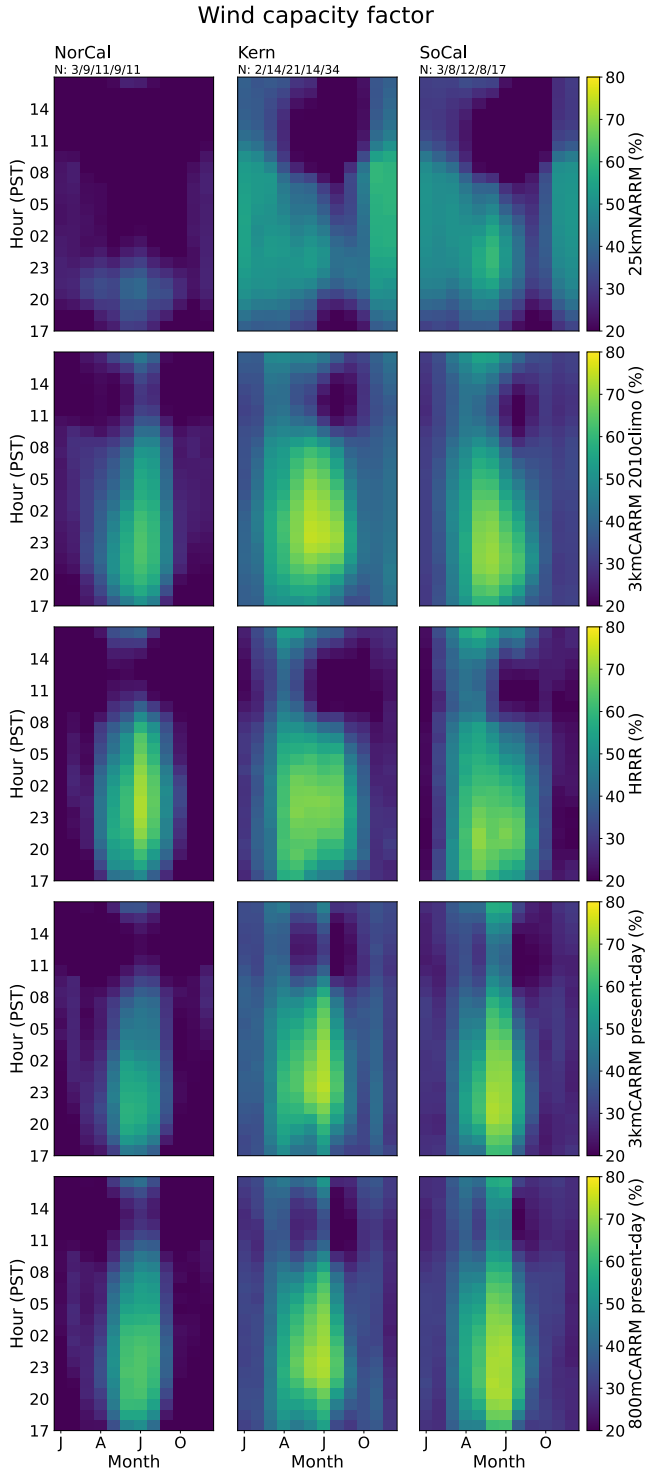

Figure 6. Diurnal and seasonal variation in wind CFs (%) estimated using PySAM. From top to bottom: E3SM-25kmNARRM, SCREAM-3kmCARRM 2010climo, HRRR, SCREAM-3kmCARRM present-day, and SCREAM-800m CARRM present-day simulations, respectively. Each subplot corresponds to a cluster in Fig. 2.

3.3.5 Statewide wind CFs and meteorological factors

Figure 7 shows the spatial distribution of wind energy resources throughout California. Overall, the best wind resources are found near mountainous regions – such as the coastal ranges near the Bay Area, the southern Sierra Nevada, and the eastern highlands of California – where wind CFs generally reach around 50%. Spatial variability is most pronounced in summer, with the lowest values in areas such as the Central Valley, which maintains the lowest CF year-round, while the highest-value regions reach their seasonal peak. This figure highlights an intuitive but crucial point: wind CF is strongly dependent on local topographic features, and existing wind plants are largely situated in the most resource-rich regions.

Figure 7 also shows the month of peak CF in California. First, all *km*-scale models predict broadly consistent seasonal peaks in wind generation across the state: generally higher in winter and lower in summer in mountainous regions, and the opposite pattern in the coast, Central Valley and the desert highlands on the leeward side of Sierra Nevada. Most existing wind farms exhibit summer peaks, whereas northern California (including the Shasta plant) generally peaks in winter, consistent with the EIA plant-level assessment (Figs. 3–5, A3a). Although present-day simulations predict a smaller extent of summer peaks in the Central Valley than the 2010climo simulation and HRRR forecasts, these differences are unlikely to affect planners, given the area's overall weak wind resource. The seasonal diversity of peak wind generation across California may be advantageous for planners in balancing the supply and demand of different renewable resources, especially given that solar resources are systematically greater in summer. Notably, seasonal peak reversals can occur over short distances, such as winter peaks along the coastline and summer peaks in the adjacent Coast Ranges.

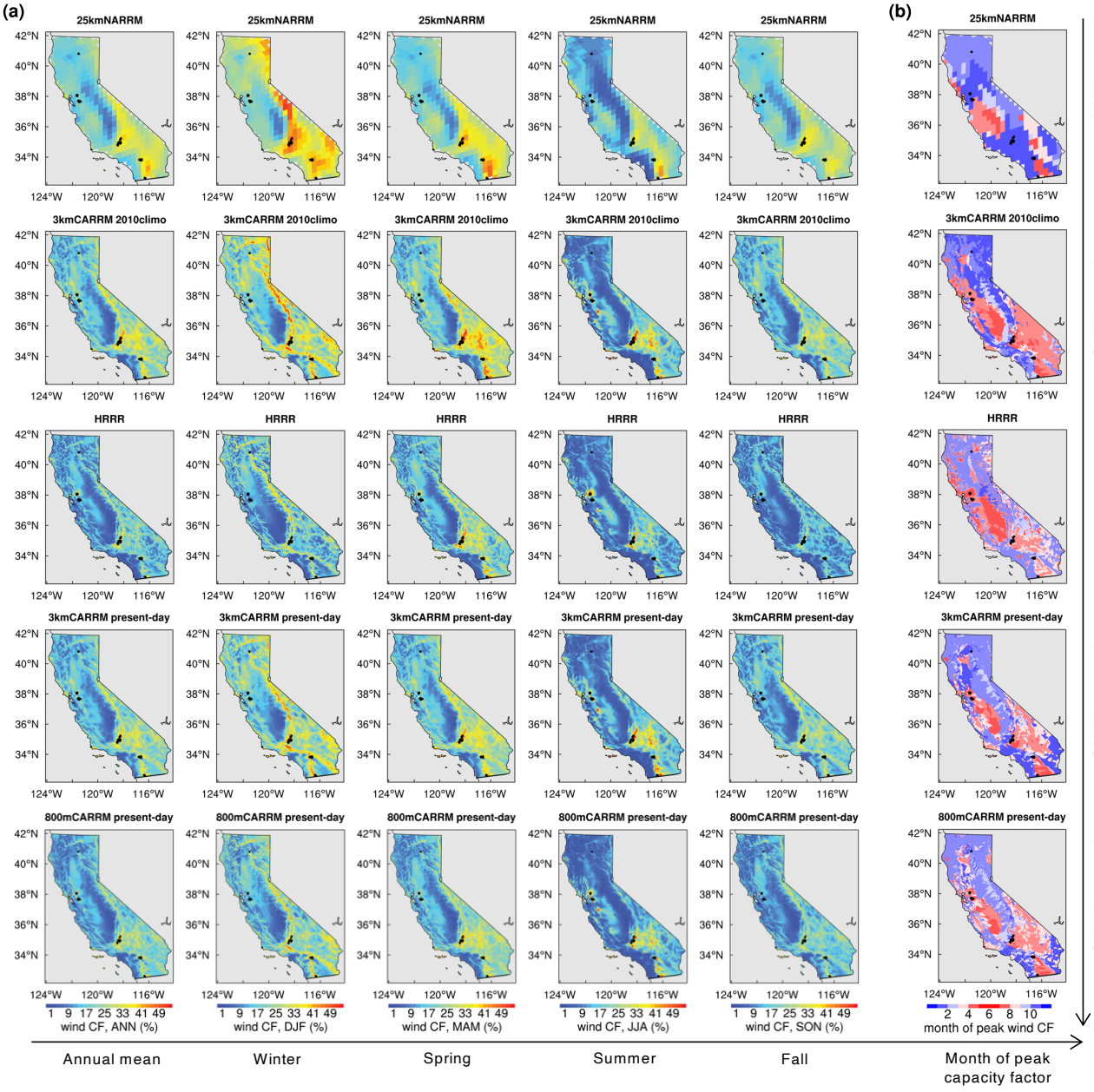

Figure 7. Statewide spatial distribution of (a) wind CFs (%) for California ~~,~~ estimated using PySAM, ~~along with~~ and (b) the month of peak CF. From left to right: annual mean (ANN), winter (DJF), spring (MAM), summer (JJA), and fall (SON) wind CFs, and the corresponding month of peak CF. From top to bottom: E3SM-25kmNARRM, SCREAM-3kmCARRM 2010climo, HRRR, SCREAM-3kmCARRM present-day, and SCREAM-800mCARRM present-day simulations. Valid EIA plant locations are indicated by small black dots on the map.

Consistent with earlier findings from the EIA, E3SM-25kmNARRM stands out as an outlier, especially in winter when its wind CF is significantly higher than that of other models. Given the strong dependence of wind resources on local topography, resolution and the model's ability to resolve topography clearly play a role. When zooming into an evaluation region such as Kern County, the higher-resolution models exhibit pronounced spatial heterogeneity within a single county, whereas the 25 km model lacks this spatial detail (Figs. 1c–1d).

To better understand the winter time discrepancy, Figure 8 focuses on key variables directly influencing winter wind CF, particularly the hub-height winds, as well as large-scale circulation. Notably, E3SM-25kmNARRM not only simulates stronger zonal winds over California but also exhibits stronger northerly winds over Southern California. Its winter blocking ridge is both stronger and located farther offshore relative to ERA5, consistent with its hub-height wind bias (i.e., excessive westerly and northerly winds). Note that the ridge shape in ~~3kmCARRM-2010climo~~ 3kmCARRM 2010climo resembles that in 25km-NARRM, although it extends further over land (Fig. 8). Nevertheless, 3kmCARRM produces much more realistic wind CFs. Another notable feature is the pronounced ridge difference between 1) the group consisting of 2010climo and 25kmNARRM (both ultimately reflecting the 2010–2014 climate) and 2) the two present-day simulations (reflecting the nominal 2015–2020 present-day climate under CMIP6 SSP5-8.5 forcing). The differences in wind CFs, however, are still primarily driven by the resolution differences across the groups. Together, these comparisons suggest that large-scale circulation plays a relatively minor role in wind simulation, while resolution and topographic representation are the dominant factors.

For the SCREAM-CARRM simulations, both present-day simulations exhibit a stronger ridge than ERA5, while the 2010climo simulation is closer in ridge strength. However, somewhat unexpectedly, the present-day simulations better capture the ridge's position and structure.

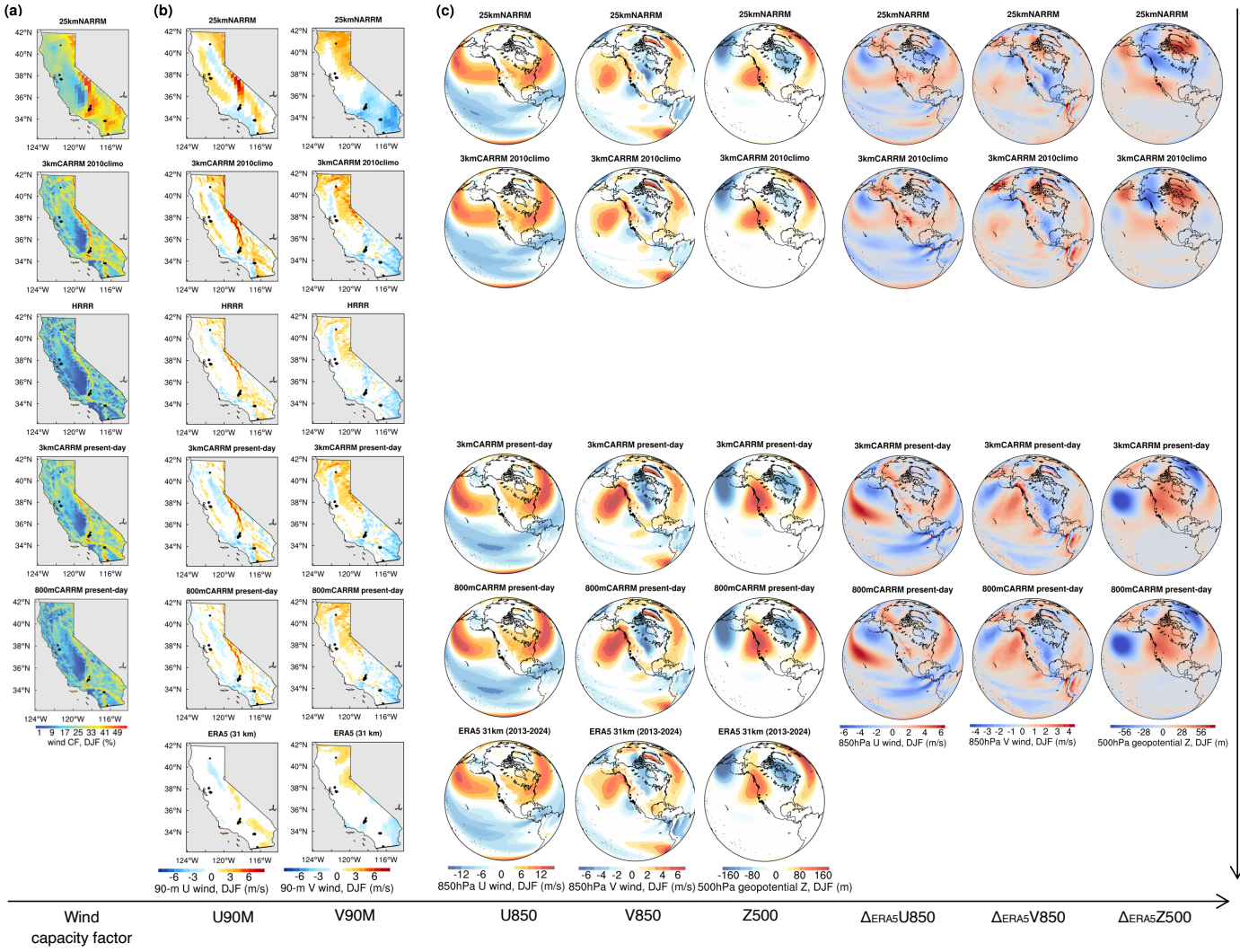

Figure 8. Direct meteorological drivers of wind ~~CF and large-scale circulation~~CFs. From left to right: (a) wind CFs~~, 90 m~~ ; (~~hub-height~~b) 90 m zonal winds (U90M) ~~,~~and 90 m meridional winds (V90M), representing hub-height winds; and (c) 850 hPa zonal winds (U850), 850 hPa meridional winds (V850), and 500 hPa geopotential height with the zonal mean removed (Z500), ~~and~~ together with the differences between each model-simulated U850, V850, and Z500 and those from ERA5~~, respectively~~. All quantities are wintertime averages. From top to bottom: E3SM-25kmNARRM, SCREAM-3kmCARRM 2010climo, HRRR, SCREAM-3kmCARRM present-day, SCREAM-800mCARRM present-day simulations, and ERA5. ~~As~~ Because ERA5 provides wind output only at 100 m height, the 100 m winds are used as ~~the~~ hub-height winds for ERA5. Global U850, V850, and Z500 are not available from HRRR because it is a regional model and does not cover the global domain.

3.3.6 ~~Diurnal cycle of simulated wind power generation~~

Figure 6 shows the diurnal cycle in simulated wind CFs. Note that EIA data are not included, as hourly EIA records are not available. Each panel displays average seasonal (horizontal, month) and diurnal (vertical, time PST) variation, grouped by geographic cluster.

Diurnal and seasonal variation in wind CFs (%) estimated using PySAM. From top to bottom: E3SM-25kmNARRM, SCREAM-3kmCARRM 2010climo, HRRR, SCREAM-3kmCARRM present-day, and SCREAM-800m CARRM present-day simulations, respectively. Each subplot corresponds to a cluster in Fig. 2.

First, all simulations exhibit highly consistent diurnal cycles, with higher values at night and lower during the day. E3SM-25kmNARRM appears highly anomalous, mainly due to the previously discussed seasonal phase reversal and significantly underestimated summer generation. Nevertheless, its diurnal timing remains relatively consistent with the other simulations. All models simulated peak values occur at approximately 21–24 PST, while minimum values occur near 12–13 PST. Second, in the two comparisons with similar internal variability, diurnal cycle differences are substantially smaller than seasonal cycle differences. The seasonal cycles in Figure 6 represent an alternate visualization of those in Figs. 3–5.

3.3.6 Overall mean absolute error of wind energy generation

Figure 9 summarizes the overall discrepancy in simulated wind CFs using mean absolute error (MAE, black bar) as the metric. For the NorCal and Kern plants, HRRR shows the lowest MAE, while SCREAM-3kmCARRM 2010climo yields the highest. The SCREAM-800mCARRM simulation produces slightly lower MAE than SCREAM-3kmCARRM in the present-day group. In SoCal, the lowest MAE is found in E3SM-25kmNARRM and the SCREAM-3kmCARRM present-day simulation. All the MAE differences remain within 8%, with the vast majority under 2%, indicating that the choice of meteorological model has a relatively small overall impact on wind power generation estimates. However, we caution against overinterpreting the MAE plots in isolation, since MAE only reflects the magnitude of at the annual scale across all plants, while compensating errors across different seasons cannot be captured in this figure.

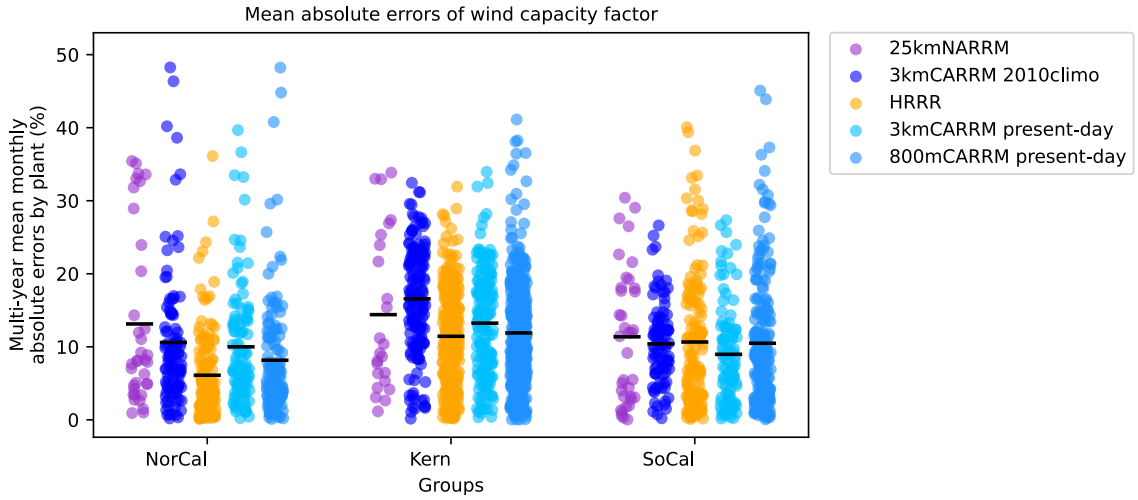

Figure 9. Multi-year mean monthly absolute errors of wind CFs by plant (%) from PySAM simulations compared to monthly EIA data: E3SM-25kmNARRM (darkorchid), SCREAM-3kmCARRM 2010climo (blue), HRRR (orange), SCREAM-3kmCARRM present-day (deepskyblue), and SCREAM-800m CARRM present-day (dodgerblue). Each dot represents the absolute difference between the multi-year mean monthly CF and the corresponding EIA value for each plant. The black horizontal bar shows the mean absolute error (MAE) for each group, averaged over all months and plants.

3.4 Solar

For solar energy, we did not apply the geographic clustering. Instead, we grouped all plants by tracker configuration: single-axis, fixed-axis, and dual-axis. Unlike wind, solar is more uniformly distributed and hence no natural classification emerges. Fig. 10 shows the PV plants after aggregating the EIA plants to the 3.25 km CARRM grid.

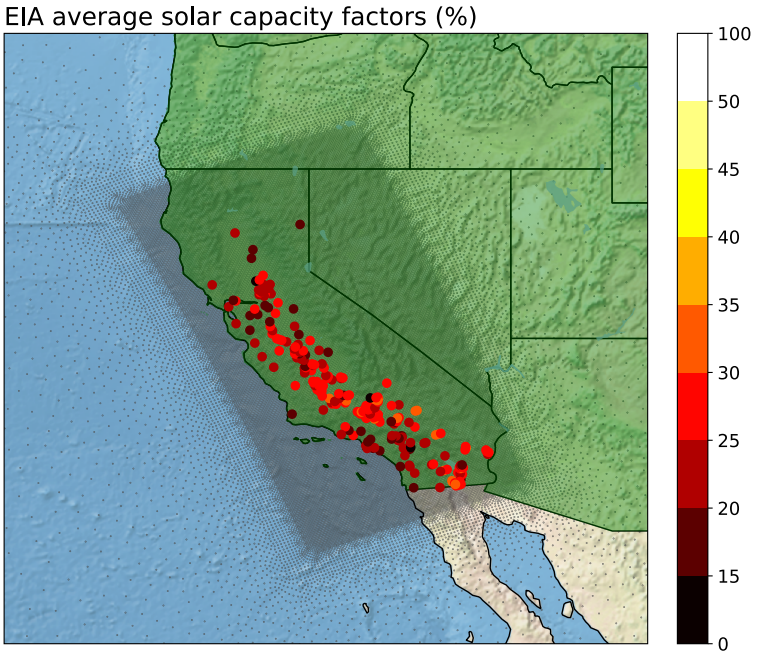

Figure 10. Geographic distribution of solar plants in California. EIA plants are aggregated to the 3kmCARRM model grid; each filled dot represents the mean solar CF (%) from valid monthly EIA records between 2013 and ~~2023.~~2024.

As outlined in the Methods section, PV systems are classified by tracker configuration into three categories based on the EIA records: fixed-axis, single-axis, and dual-axis. Fixed-axis arrays are expected to deliver the lowest energy yield, whereas single- and dual-axis trackers continuously adjust panel orientation along one or two axes to follow the trajectory of the sun, keeping the modules closer to normal incidence and thus systematically producing more electricity than fixed-axis systems.

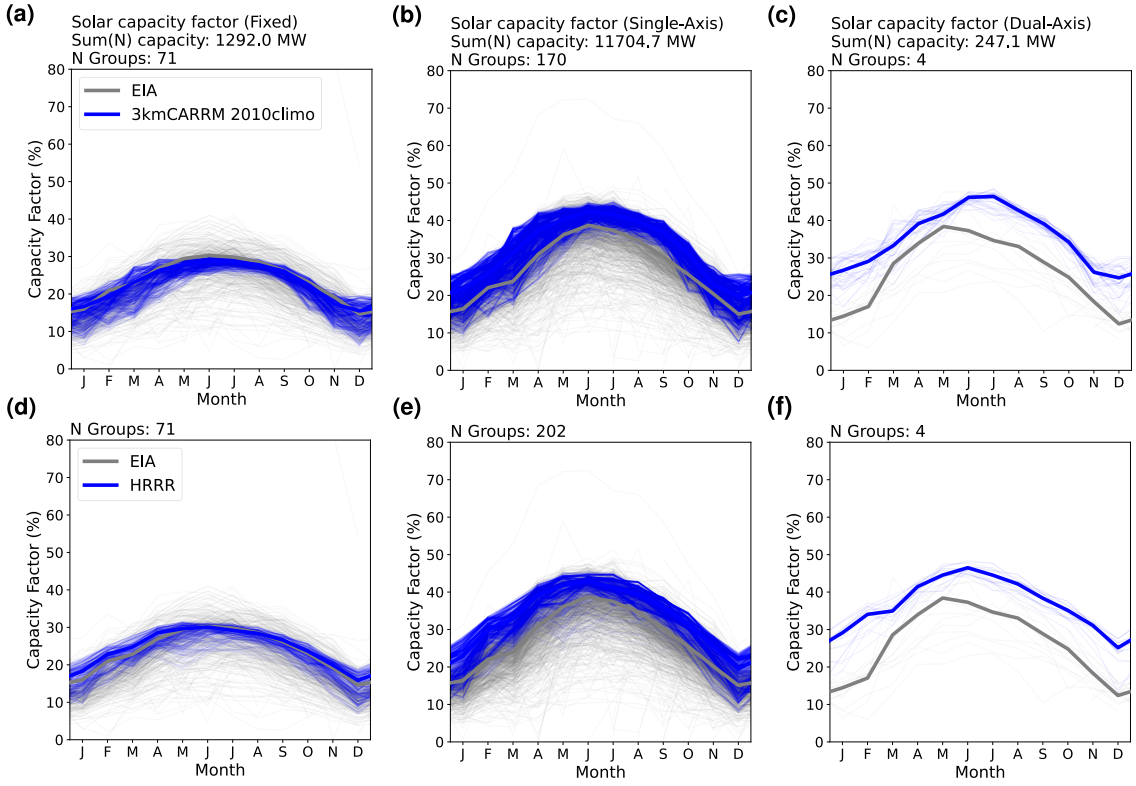

Figure 11. As in Fig. 4, but for solar CFs grouped by tracking type: (a, d) fixed-axis, (b, e) single-axis, and (c, f) dual-axis.

3.4.1 How ~~much~~ do ~~meteorological models impact the~~ different axis-tracking methods affect biases in modeled solar ~~energy~~CFs?

Figure 11 shows the solar CFs in California as estimated by PySAM, using SCREAM-3kmCARRM 2010climo and HRRR simulations. Compared to wind, the seasonal cycle of solar CFs is much simpler, with a clear summer peak and winter ~~low. Even in the~~ minimum. Even for dual-axis ~~case~~systems, despite having only three samples, the seasonal pattern remains smooth. This likely reflects the relatively stable solar resource in California's sunny climate~~. In cloudier regions , where solar variabilityis higher~~; in cloudier regions with higher solar variability, such smoothness may not hold.

~~As in Fig. 4, but for solar CFs, grouped by tracking type: fixed-axis, single-axis, and dual-axis.~~

Across the three tracker classes, the overestimation of ~~the~~ simulated solar CFs relative to EIA records ~~grows~~ increases with tracker complexity. For fixed-axis arrays, both SCREAM-3kmCARRM 2010climo and HRRR reproduce the observations well: on average, 3kmCARRM deviates by ~~<~~ less than 3% in every month. For single-axis systems, the overestimation in both models ~~stays~~ remains below 6%, ~~and~~ while for dual-axis systems it ~~remains~~ stays below 13%.

This monotonic increase suggests that the tracker algorithms implemented in PySAM are somewhat idealized~~; in~~ . Solar tracking systems are designed to increase energy production by maintaining the panel orientation close to perpendicular

to the incoming solar radiation. In principle, dual-axis trackers should therefore achieve higher efficiency than single-axis systems. However, this pattern is not evident in the EIA generation records. One possible reason is that dual-axis systems are mechanically more complex, and the additional costs and operational constraints associated with the control and driving mechanisms may offset part of the theoretical energy gain; such information, however, is not reported in the EIA-860 dataset. In addition, the benefit of multi-axis tracking may be limited in hot climates, where the increased incident irradiance can raise module temperatures and reduce photovoltaic output due to thermal losses (El Hammoumi et al., 2022).

In the SAM PV model, tracker orientation follows geometric tracking algorithms that determine panel tilt and azimuth from solar position and array geometry: for single-axis systems, the panel azimuth tracks the solar azimuth while the tilt remains fixed relative to the rotation axis, whereas for dual-axis systems both tilt and azimuth follow the solar zenith and azimuth angles, respectively (Gilman et al., 2018a). In practice, mechanical ~~or control limitations likely cause~~ constraints, tracking control strategies, and stow conditions can lead to slower or less precise tracking, reducing the ~~actual gains~~ realized energy gains relative to these idealized simulations.

3.4.2 How different are solar CFs between SCREAM-3kmCARRM and HRRR at comparable horizontal resolution?

Figure 11 also compares solar CFs derived from SCREAM-3kmCARRM 2010climo and HRRR simulations. Overall, the two models produce broadly similar seasonal cycles and magnitudes of solar CFs across the three tracker categories.

Some systematic differences nevertheless emerge. Notably, SCREAM-3kmCARRM 2010climo performs remarkably well in reproducing EIA-recorded solar power generation, even outperforming the high-frequency-assimilated HRRR ~~.~~simulation in some months. For fixed-axis systems, the positive bias in HRRR is more pronounced before July; SCREAM-3kmCARRM overestimates the EIA values by about 0-1%, whereas HRRR overestimates them by roughly 1–2%. However, for single-axis systems, the positive bias in SCREAM-3kmCARRM during autumn is slightly more evident than in HRRR, whereas before June the bias is smaller than in HRRR.

Agreement in energy output alone does not necessarily imply accurate meteorological modeling, ~~as~~ since uncertainties in generation modeling (such as the representation of axis tracking) are difficult to quantify and the reliability of EIA monthly records is not well characterized. Nevertheless, ~~we will show in a later section that SCREAM also outperforms~~ differences in the simulated meteorological drivers of solar power generation likely contribute to the CF discrepancies between the models. Surface downwelling shortwave radiation is the primary determinant of solar CF, and differences in cloud representation can therefore lead to systematic CF biases. As shown in Section 3.4.4, SCREAM exhibits smaller biases relative to ERA5 than HRRR in simulating surface solar radiation~~, further supporting its potential value for solar resource assessment.~~ .

In addition, although this section focuses on CF evaluation, we examined the corresponding meteorological variables in ERA5 over the same period as HRRR (2018–2022) and found spatial patterns very similar to the 2013–2024 climatology (shown in Section 3.4.5). This suggests that HRRR solar CFs averaged over 2018–2022 would likely be comparable to those averaged over the EIA record period, implying that the differences identified here primarily reflect differences between meteorological models rather than sampling effects. However, given the many structural differences between the two models, this evaluation quantifies the magnitude of differences in simulated solar CFs but cannot isolate specific causal mechanisms.

3.4.3 How much does horizontal resolution affect ~~the~~ solar energy?

Figure 12 shows comparisons that help to understand the impact of resolution on solar energy simulations. ~~As with the previous evaluation group, CFs from plants with tracking~~

First, the dominant influence of axis-tracking configuration remains evident here: plants equipped with tracking systems (single- or dual-axis) ~~tend to~~ generally produce more solar energy than those without tracking. ~~On the other hand~~In contrast, the influence of spatial resolution ~~(25 km vs.~~ appears relatively small in this context. For fixed-axis systems, all three simulations show a slight underestimation relative to the EIA records, with 25kmNARRM exhibiting a smaller bias than the 3.25 km ~~vs.~~ and 800 m ~~) appears minor in this context. However, in~~ CARRM simulations. For single- dual-axis systems all simulations show an overestimation, and among single-axis systems the overestimation is most pronounced in 25kmNARRM. As will be shown in the next section on the diurnal cycle, the solar CF simulated by 25kmNARRM reaches particularly high daily peak values, where it is not only higher than in the 3.25 km and 800 m CARRM simulations but even exceeds that of HRRR.

Overall, for solar CFs, the differences between the 25 km and km-scale simulations are less pronounced than those observed for wind CFs. Horizontal resolution therefore does not appear to be the dominant factor controlling inter-model differences, although several notable discrepancies remain. Nevertheless, in cloudier climates, resolution may play a larger role by ~~better~~ more accurately capturing sub-grid cloud structures that affect irradiance.

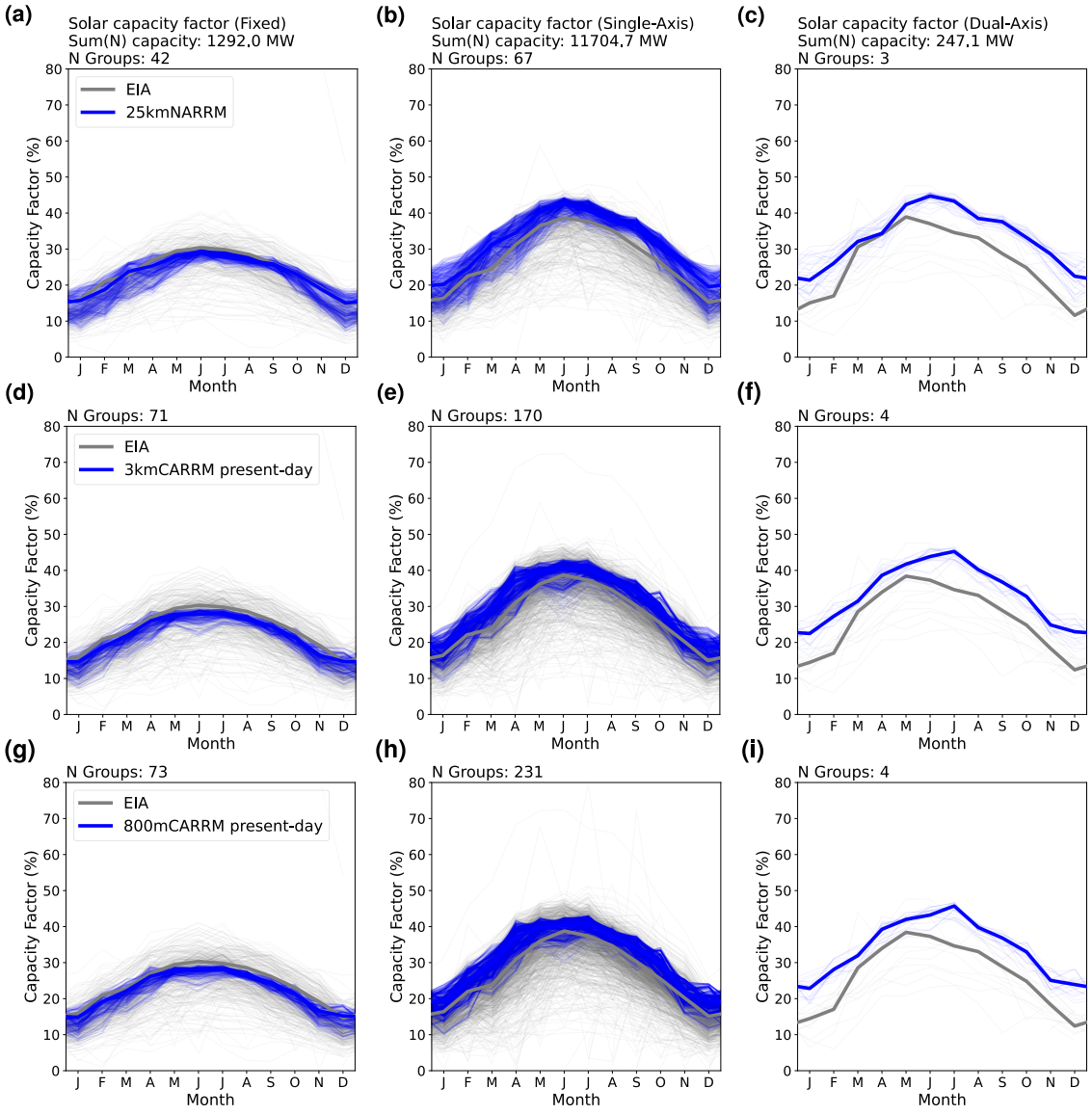

Figure 12. As in Fig. 11 but ~~for for~~ comparing (a)–(c) E3SM-25kmNARRM simulations with (d)–(f) 3.25 km ~~vs.~~ and (g)–(i) 800 m ~~SCREAM-CARRMs~~ SCREAM-CARRM present-day simulations.

3.4.4 ~~Statewide~~ Diurnal cycle of simulated solar ~~CFs, meteorological factors, and large-Scale circulation~~power generation

735 The simulated diurnal cycle of solar CFs is shown in Fig. 13.

First, the diurnal pattern of PV generation and its seasonal variation appear highly idealized, reflecting the diurnal evolution of solar azimuth angle and the seasonal evolution of solar zenith angle. Both the diurnal cycle and its seasonal dependence show

little inter-model or regional variation. For example, the peak solar CF consistently occurs in the afternoon, and the timing of this peak remains nearly unchanged throughout the year across all months.

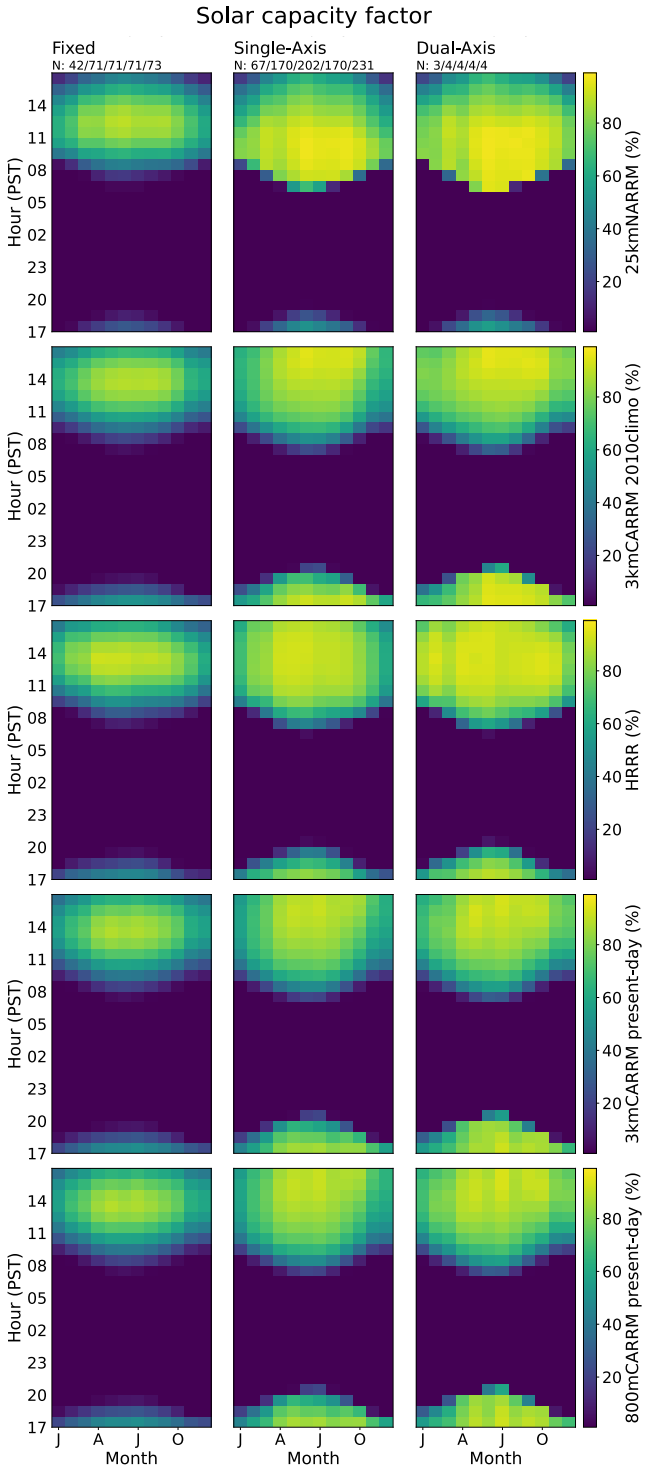

Figure 13. Diurnal and seasonal variation in solar CFs (%) estimated using PySAM. From top to bottom: E3SM-25kmNARRM, SCREAM-3kmCARRM 2010climo, HRRR, SCREAM-3kmCARRM present-day, and SCREAM-800m CARRM present-day simulations, respectively. Each subplot corresponds to clusters of fixed, single- and dual- axis tracking.

In terms of magnitude, the km-scale simulations exhibit differences similar to those seen in the seasonal analysis: HRRR generally produces higher solar CFs than all SCREAM-CARRM simulations. The 25kmNARRM simulation stands out in this figure, with extremely high solar CF peaks (note that the bright yellow in the colorbar approaches 100%). Its peak values are not only higher than those of the km-scale SCREAM-CARRMs but even exceed the peak values simulated by HRRR. In addition, for the single-axis and dual-axis categories in 25kmNARRM, PV generation increases very abruptly from near zero to peak values, and the duration of high output is shorter than in the other simulations. For example, in 25kmNARRM during May–July, generation rises from nearly zero at around 06 PST to near peak levels by about 08 PST, remains high until around noon, and then begins to decrease. In contrast, in all CARRM simulations the increase after 07 PST is much more gradual, reaching peak values around 14-15 PST, followed by a gradual decline until about 20 PST. The diurnal amplitude of HRRR during the midday-to-evening decline is very similar to that of the CARRMs, while the morning ramp-up is neither as steep as in 25kmNARRM nor as gradual as in the CARRMs, lying between the two. The HRRR peak hour also occurs slightly earlier than in the CARRM simulations. Overall, however, the diurnal patterns among the km-scale simulations remain similar in general.

This section provides an additional perspective to the seasonal analysis and again shows that HRRR and 25kmNARRM tend to produce higher solar CFs than the CARRM simulations, although several additional details emerge. In particular, the diurnal cycle in 25kmNARRM is the steepest, especially during the morning ramp-up toward the peak. However, when averaged over the full day, the differences appear less pronounced in the seasonal analysis than in the peak comparisons, because the rapid increase is followed by a similarly rapid decline, resulting in a relatively short duration of high output. The peak timing also differs, with 25kmNARRM reaching peak generation much earlier than the km-scale simulations. Rapid ramps in PV generation pose significant challenges for power system operation, making accurate simulation of the diurnal cycle particularly important. Unfortunately, the EIA dataset does not provide hourly solar generation records, and there are no hourly proxy observations of surface solar radiation that could help determine which simulation is most accurate.

3.4.5 Statewide solar CFs and meteorological factors

Figure 14 shows the spatial distribution of solar CFs across California. Due to the state's predominance of clear-sky conditions, the spatial heterogeneity of solar CFs is much smaller than that of wind CFs, generally decreasing from south to north in latitudinal bands. In winter, the northernmost part of California has the lowest CFs, around 9%, while in summer, most regions reach 25–30%.

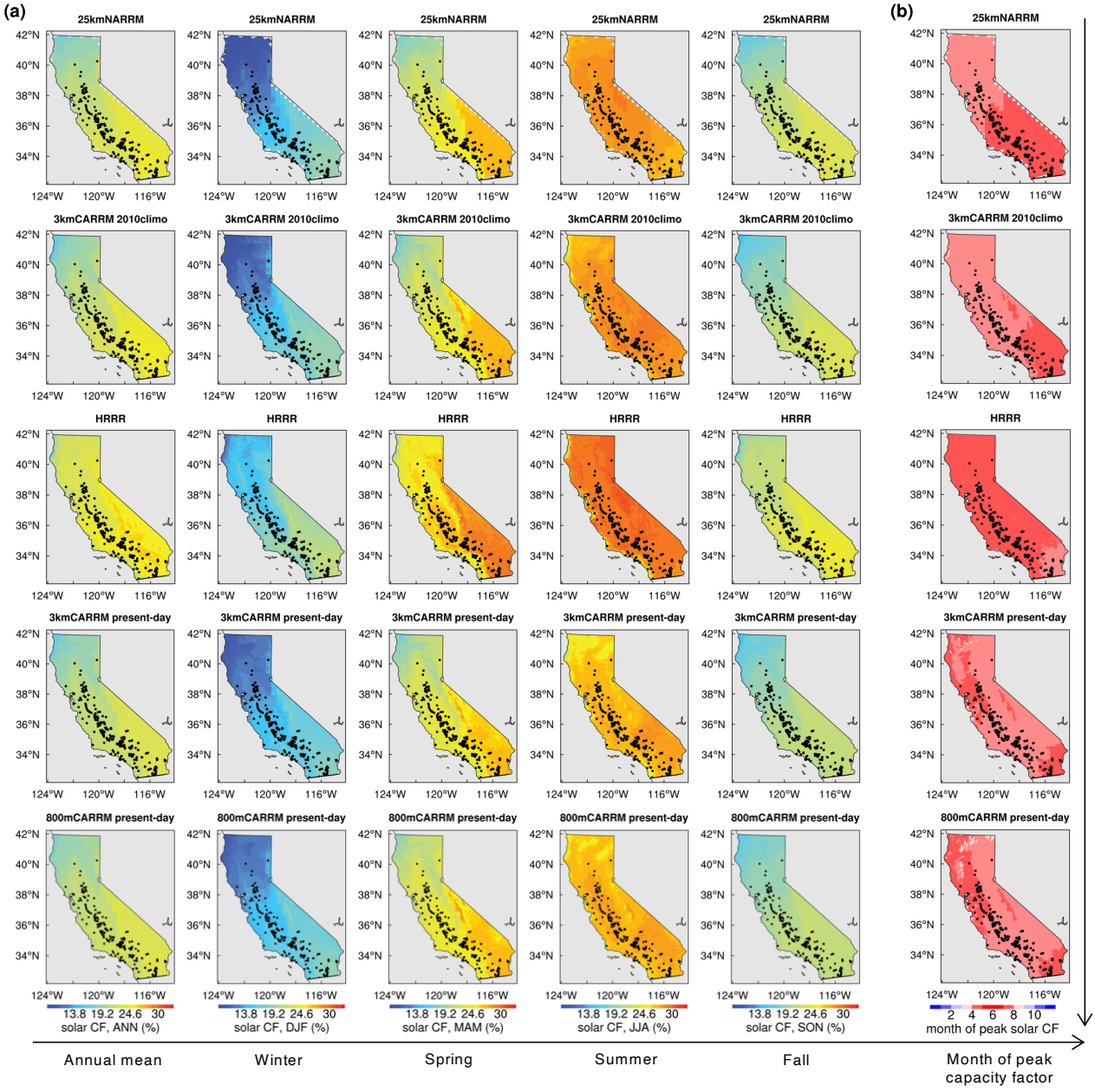

Figure 14. As in Fig. 7, but for solar CFs using the fixed-axis tracking assumption.

Among all simulations, HRRR stands out as an outlier, consistently producing higher solar CFs than all SCREAM/E3SM RRM simulations evaluated here (Fig. 14). In particular, HRRR exhibits the largest positive bias relative to the EIA records. Moreover, HRRR simulates ~~more shortwave radiation reaching the ground (Figure~~larger amounts of surface downwelling shortwave radiation (Fig. 15). This likely explains why HRRR deviates more from EIA ~~records than SCREAM-CARRMs~~observations

than the SCREAM-CARRM simulations, as surface shortwave radiation ~~and~~ is the dominant meteorological driver of PV generation. It determines the amount of solar energy collected by PV modules and, together with near-surface ~~air temperature are the major meteorological drivers of solar CF. The former determines the available solar energy , while the latter affects PV performance andcell temperature. Radiation~~ air temperature, influences cell temperature and PV performance (e.g., Mayer and Gróf, 2021;) . Surface radiation is highly sensitive to cloud conditions, ~~which are subject to large uncertainties and depend on temperature and humidity. HRRR shows lower humidity than SCREAM-CARRMs, as indicated by the dew point, which is consistent with its higher shortwave radiation. Cell temperature , which affects the electrical characteristics and efficiency of PV modules, is mainly driven by ambient temperature and insolation; some detailed PV models also take dew point and wind speed into account (Mayer and Gróf, 2021). Temperature simulations are fairly~~ and clouds remain one of the largest sources of uncertainty in atmospheric modeling because they depend on multiple interacting physical processes. On the other hand, temperature simulations are relatively consistent across all models~~and closely match ERA5~~.

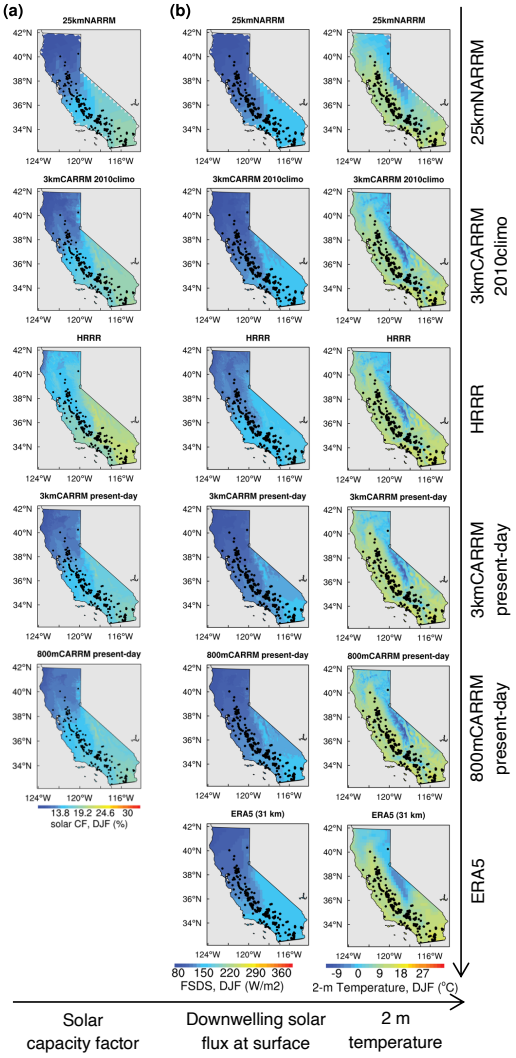

Figure 15. Direct meteorological drivers of solar CF~~and large-scale circulation~~. From left to right: (a) solar CFs, (b) total downwelling solar flux at surface (FSDS) ~~, 2 m temperature,~~ and 2 m ~~dew point~~temperature.

3.4.6 Overall mean absolute error of solar energy generation

Figure 16 summarizes the overall discrepancies in simulated solar CFs using MAE. Fixed-axis systems show the smallest MAE (all below 4%), while single-axis tracking ranges from 4–6%, and dual-axis systems exhibit overestimation of 8–12%. This pattern suggests a possible systematic effect: as tracking complexity increases, so do the discrepancies. One plausible explanation is that PySAM's tracking algorithms may idealize real-world performance, especially for dual-axis systems that rarely achieve perfect sun-following in practice. Notably, SCREAM-CARRMs show smaller discrepancies than HRRR across all tracking types.

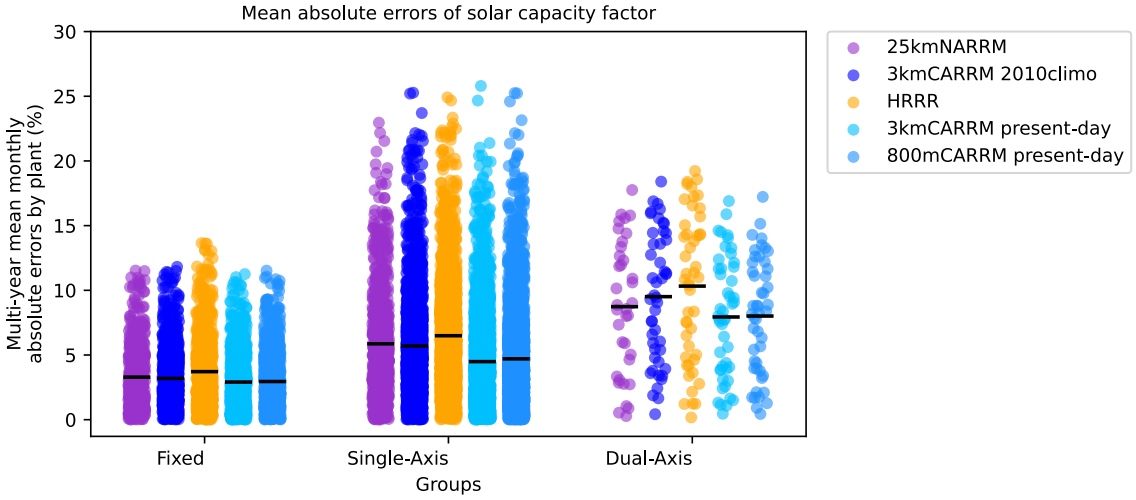

Figure 16. As in Fig. 10 but for for solar CFs.

4 Conclusions

This study evaluates wind and solar power generation derived via PySAM from SCREAM-CARRMs at 3.25 km and 800 m resolutions, through comparisons with EIA records, PLUSWIND, and HRRR. The following summarizes our responses to the three central questions raised in the Results:

1. *How much do generation modeling assumptions impact the wind energy?* For wind, differences between PySAM and PLUSWIND are larger than those arising from internal assumptions within PLUSWIND, but still smaller than discrepancies caused by differing meteorological models. For solar, however, modeling assumptions play a bigger role – especially for systems with axis tracking. As seen in Figure 16, errors increase with tracking complexity, suggesting that PySAM's idealized tracking assumptions may not reflect the mechanical or operational limitations of real-world systems.

2. *How much do meteorological models impact the wind and solar energy?* Despite being based on two fundamentally different meteorological models, both SCREAM-3kmCARRM 2010climo and HRRR yield similar wind and solar ~~power outputs~~ generation when processed with PySAM. The MAE differs by 1-5% for wind and 1-2% for solar. While the data-assimilated HRRR differs less from EIA data for wind, SCREAM-3kmCARRM aligns more closely for solar ~~power~~generation.

3. *How much does horizontal resolution affect the wind ~~and solar~~ energy?* The transition from E3SM-25kmNARRM to SCREAM-3kmCARRM results in a performance leap for wind energy, particularly in capturing the seasonal cycle, highlighting the importance of improved horizontal resolution and better representation of topography. In contrast, the difference between the 3.25 km and 800 m SCREAM-CARRM present-day simulations (and E3SM-3kmCARRM and HRRR from bullet 2) is minimal, with MAE differences below 2%, indicating limited resolution sensitivity at these finer

scales in SCREAM. Solar PV generation shows much lower resolution sensitivity across E3SM-25kmNARRM, 3.25 km SCREAM-CARRM, and 800 m SCREAM-CARRM than the differences between SCREAM and HRRR.

It should be noted that MAE alone cannot fully reveal model performances. Seasonal patterns offer more insight, especially for those interested in meteorological model biases. Wind power discrepancies are particularly intriguing: in the absence of direct hub-height wind observations and hourly EIA data, we cannot precisely quantify the source of EIA–model differences without more data. Although absolute biases remain unknown, comparisons across the three questions provide a sense of the relative influence of different factors on modeled generation.

The ~~comparison between SCREAM-3kmCARRM and HRRR suggests that structural differences in the meteorological models (such as dynamics, physical parameterizations, and data assimilation/nudging techniques) have less impact on generation estimates than horizontal resolution. The~~ importance of resolution for wind energy is supported by findings from PLUSWIND and by our comparison between 25kmNARRM, 3kmCARRM and 800mCARRM – where the seasonal phase mismatch previously identified in NARRM is no longer present, underscoring the improvements gained through *km*-scale modeling. This has critical implications for power system planning, as seasonal variability plays a key role in balancing supply and demand, and its uncertainty cannot be effectively mitigated by short-duration storage technologies such as batteries (e.g., Staadecker et al., 2024). The conclusion that km-scale simulations perform substantially better than O(10 km) simulations for wind power generation is likely applicable to other regions with complex terrain, as topography is one of the most dominant lower boundary conditions controlling winds. However, increasing resolution from 3.25 km to 800 m adds little value for SCREAM's wind ~~power~~generation, in particular, the seasonal pattern shows that while the 800 m simulation reduces wintertime overestimation, it amplifies summertime overestimation. ~~The~~

Regarding the lack of large sensitivity between 3.25 km and 800 m SCREAM-RRMs, on the one hand, likely depends on SCREAM's turbulence scheme and should not necessarily be generalized to other models.

~~The more pronounced wintertime overestimation of wind power by SCREAMcompared to HRRR may indicate a systematic seasonal bias in SCREAM-CARRMs~~SCREAM's ~~wind simulations. This can be partially explained by the weaker high-pressure blocking ridges in the 2010climo simulation, which tend to favor an overestimated frequency of atmospheric rivers and associated cold-front winds. The limited added value from increased resolution (~~s SHOC turbulence scheme is scale-aware and scale-insensitive, meaning that a larger fraction of turbulent kinetic energy is explicitly resolved as resolution increases, while the representation of clouds and thermodynamics does not change substantially with resolution (Bogenschutz et al., 2023). In contrast, turbulence schemes lacking this property may artificially exaggerate the contribution of subgrid transport at 800 ~~m vs. 3.25 km) is consistent with the weak sensitivity of SCREAM-CARRMs' sub-grid scale transport of turbulence in the 5 km to 800~~ m ~~regime, as reported in Bogenschutz et al. (2023), and could~~ resolution. On the other hand, the muted sensitivity may also be attributed in part to model errors ~~in~~ associated with the turbulence gray zone (e.g., Wyngaard, 2004; Chow et al., 2019; Honnert et al., 2020)~~(also see discussions in Zhang et al., 2025)~~; see also discussions in Zhang et al. (2025). Specifically, horizontal turbulent mixing (neglected in most PBL schemes including SHOC) may become non-negligible at 800 m, especially in complex terrain where three-dimensional effects matter (e.g., Juliano et al., 2022; Arthur et al., 2025b).

The comparison between SCREAM-3kmCARRM and HRRR suggests that structural differences in the meteorological models have less impact on generation estimates than horizontal resolution. The more pronounced wintertime overestimation of wind generation by SCREAM compared to HRRR may indicate a systematic seasonal bias in SCREAM-CARRMs' wind simulations. This can be partially explained by the weaker high-pressure blocking ridges in the 2010climo simulation, which tend to favor an overestimated frequency of atmospheric rivers and associated cold-front winds.

For generation modeling, PySAM applied to HRRR yields wind generation estimates closer to EIA and with less overestimation than PLUSWIND, suggesting that loss estimation may be the dominant factor. Although PLUSWIND's turbine power curve derived from plant-level turbine diameter and capacity is theoretically more accurate than PySAM's default curve, its assumed ~7% loss factor (Millstein et al., 2023) appears to be underestimated, based on the comparison between HRRR+PySAM and HRRR+PLUSWIND. In addition, we provide PySAM with each model's own hourly temperature, humidity, and pressure fields to calculate air density – ensuring consistency with the wind inputs. In contrast, PLUSWIND uses the same MERRA2-based air density estimates even when driven by high-resolution wind fields like HRRR, potentially introducing a resolution mismatch.

For ~~solar, SCREAM shows smaller discrepancies from EIA than HRRR. This~~ the comparison between solar CFs simulated by 3kmCARRM and 25kmNARRM, the conclusion that PV power generation is relatively insensitive to horizontal resolution may partly reflect California's predominantly clear-sky conditions, ~~where~~ under which model biases are less amplified~~. The better performance of~~, and may therefore be limited to regions where mesoscale cloud-precipitating processes are not prevalent. Although California contains extensive mountain ranges, precipitation is primarily associated with large-scale systems (atmospheric rivers) and is relatively infrequent overall, particularly during summer when solar generation is at its peak. Local orographically forced precipitation and mesoscale convective systems are largely absent.

The smaller solar CF discrepancies between SCREAM-CARRMs ~~in simulating solar CF is~~ and EIA, relative to HRRR, are consistent with their smaller bias in surface-reaching shortwave radiation~~. It highlights~~ , highlighting SCREAM's potential applicability in regions with similar climates. ~~As discussed earlier,~~ Sensitivity to different meteorological models may also be broadly applicable, since the representation of clouds (especially low clouds) remains one of the largest sources of uncertainty in atmospheric models.

For PV generation modeling, axis-tracking type remains the dominant source of variation, with discrepancies increasing alongside tracking complexity. This suggests that PySAM's multi-axis tracking assumptions may be overly idealized compared to real-world system behavior. The sensitivity of PV generation to axis tracking arises from the geometric design of PV panels and should therefore be robust, whereas it remains unclear whether the overestimation of multi-axis tracking found in PySAM applies to other power generation models.

Several limitations of this study are worth noting. First, although the EIA monthly data represent the best available "observations" they are not automatically collected, and the quality-control procedures used to derive monthly means are unclear, and imputation introduces an additional concern. As a result, we cannot disentangle the impacts of data processing or operational losses in practice, such as outages and curtailment, nor can we fully avoid the effects of imputation, which would otherwise substantially reduce the effective sample size and damp the spread of interannual variability. Second, none of the simulations conducted

here are weather hindcasts, i.e., the simulated timeseries do not correspond one-to-one with historical observations. Thus, our evaluation is limited to climatological statistics rather than specific time series or historical events. Finally, we did not conduct an extensive sensitivity analysis and instead chose to follow the default assumptions of PySAM, as detailed plant-level optimization of a energy model is not the primary focus of this work.

Overall, this work demonstrates the importance of high-resolution simulations for wind and solar estimates and that SCREAM-RRM is a viable tool for such energy resource assessments. With improved observational constraints and efforts to better quantify generation modeling uncertainties, future work can further understand SCREAM's biases and apply SCREAM to long-term power generation prediction and optimization.

Code and data availability. The SCREAM 100 m San Francisco Bay Area Regionally Refined Model 0.0 version code, model, and analysis data (including the lists of valid EIA wind and PV plants in California used in our analysis) can be found at Zhang (2025) (https://doi.org/10.5281/zenodo.16809290, last access: 11 August 2025). The SCREAM 3.25 km and 800 m CARRMs source code is also available on GitHub at: https://github.com/jsbamboo/E3SM/tree/jzhang/update20231221_bogensch_CA_32xRRM (for SCREAMv0, last access: 11 August 2025) and an archive (CApySAM-CA3kmCA800m-SCREAMv0-v0.0; https://github.com/jsbamboo/E3SM/releases/tag/CApySAM-CA3kmCA800m-SCREAMv0-v0.0, last access: 11 August 2025).

Appendix A: Applying an independent-plant ratio threshold to EIA monthly data

During our analysis, we encountered a puzzling phenomenon: for all simulations, the seasonal cycle of the Shasta CF (2013–~~2023~~ 2024 climatology) was completely out of phase with the multi-year mean seasonal cycle recorded in the EIA data. This plant was so unusual because no other plant in the *km*-scale simulations exhibited such an off-seasonal-phase issue. In the absence of nearby wind observations, it was difficult to determine whether all models were wrong or whether the EIA record was wrong.

A key insight came from an anonymous reviewer of Lee et al. (2025), who cautioned:

> "EIA gathers monthly data from a subset of plants from each state, creates monthly weights for that state, and then, for the rest of the plants in the state, distributes reported annual generation to each month based on the state-level weights... You can observe this in the data if you take the ratio of July to January (or any two months) across a set of plants in a state. Some of the plants will have independent ratios, but many plants will have the exact same July to January ratio. For those plants with non-independent monthly profiles, EIA only gathers annual generation records."

In fact, the EIA-923 documentation (https://www.eia.gov/electricity/monthly/pdf/technotes.pdf, last accessed: 8 August 2025) briefly describes the *imputation* method on pages 3 and 15:

> "Imputation: For monthly data, if the reported values appeared to be in error and the data issue could not be resolved with the respondent, or if the facility was a nonrespondent, a regression methodology is used to impute

for the facility. The same procedure is used to estimate ("predict") data for facilities not in the monthly sample. The regression methodology relies on other data to make estimates for erroneous or missing responses."

"Imputation: For select survey data elements collected monthly, regression prediction, or imputation, is done for missing data, including non-sampled units and any non-respondents. For data collected annually, imputation is performed for non-respondents."

To assess the degree of imputation in the annually collected data, and in particular to understand the mysterious behavior of Shasta, we quantified the severity of scaling for each plant and each year.

For the valid EIA plants selected by the aforementioned uality control criteria, we computed the Pearson correlation distance (i.e., 1 minus the Pearson correlation coefficient) for monthly data between every pair of plants for each year. The yearly distances were plotted as scatterplots (Fig. A1). If the distance value was less than 10^{-6} (i.e., close to 0, accounting for truncation error), then the plant pair was deemed *non-independent*. Vertical lines in Fig. A1 indicate that a plant i has an identical correlation coefficient with multiple other plants ($j = 1, \ldots, N$), implying that the j plants are *non-independent* and plant i is *independent*. If in a given year, there were no plant pairs with zero distance and no vertical lines, the year was classified as a *fully-independent year* , meaning no scaling existed among any plants. Taking wind as an example: in 2017, all plants are non-independent (all scatter points are bright green with no blue vertical lines); in 2013, most plants are independent, but some are non-independent (clusters with $distance = 0$ appear larger in size and are more blue in color); in 2023 and 2024, there are no blue vertical lines and no clusters with $distance = 0$ for wind CFs (all scatter points have the same size), indicating that all plants in that year are independent.

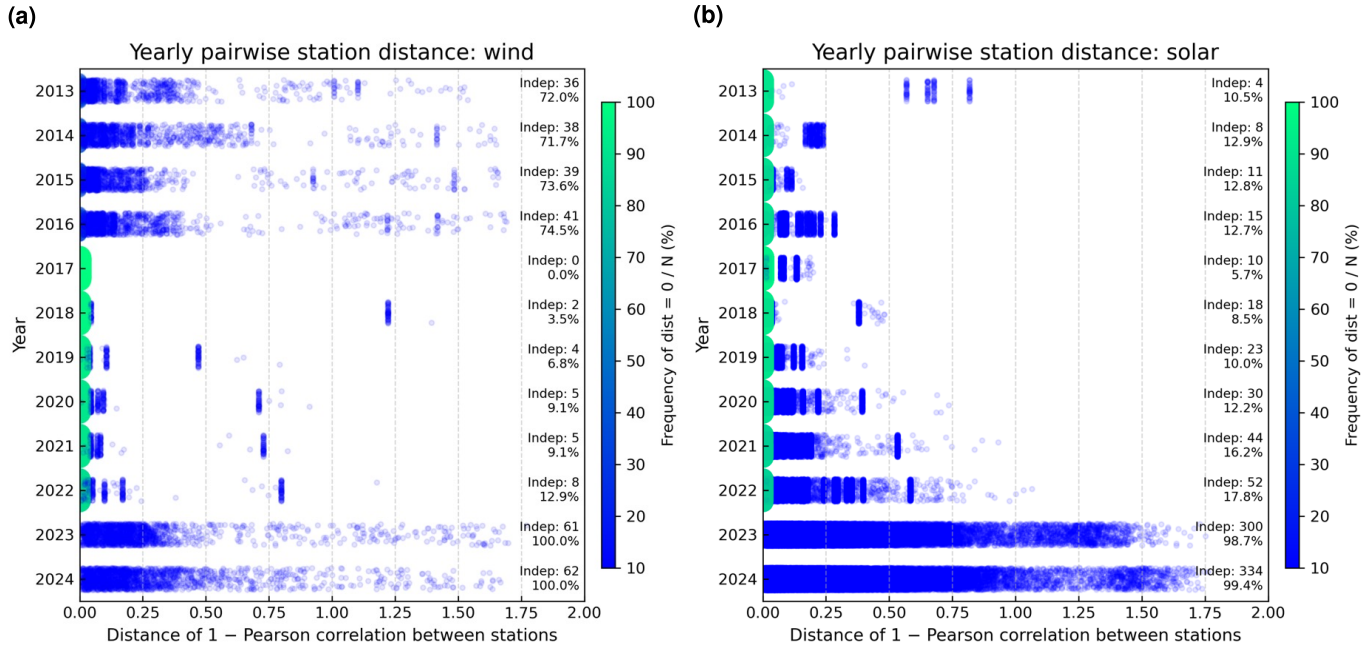

Figure A1. Scatter plots for assessing plant independence for (a) wind and (b) solar. Each point represents the Pearson correlation distance for a given year and plant pair, with the value on the x-axis and the year on the y-axis. Distances closer to zero (considering a truncation error of 10^{-6}) are shown in brighter green, and the sizes of the corresponding scatter points are set larger; conversely, larger distances are shown in blue with a constant size. A vertical jitter of ± 0.25 is applied to the y-axis for each year. This visualization step highlights independent samples: plants that have a fixed Pearson correlation distance with multiple other plants, which appear as short blue vertical lines in the figure. The number of independent plants and their proportion relative to all plants in that year are quantified and displayed as text to the right of each subplot for each year.

From Fig. A1, it can be seen that 2023 ~~was the only fully-independent year~~ and 2024 were the only fully independent years between 2013 ~~–2023. (Note: as of 9 August 2025 21:43 PST, the EIA has not yet released the~~ and 2024 ~~EIA–860 form)~~for wind CFs, and solar CFs in these two years were also nearly fully independent. The degree of scaling (i.e., the proportion of plants whose seasonal cycles are scaled from other plants) varies from year to year: for wind, in 2017–2022, the vast majority of plants (>90%) had artificial seasonal cycles, whereas for solar, a large fraction (>80%) of plants had artificial seasonal cycles starting as early as 2013. To visualize this more directly, Fig. A2 shows Shasta's independence status for each year from 2013–~~2023~~2024: red for independent years, blue for non-independent years, and thin gray lines for other plants.

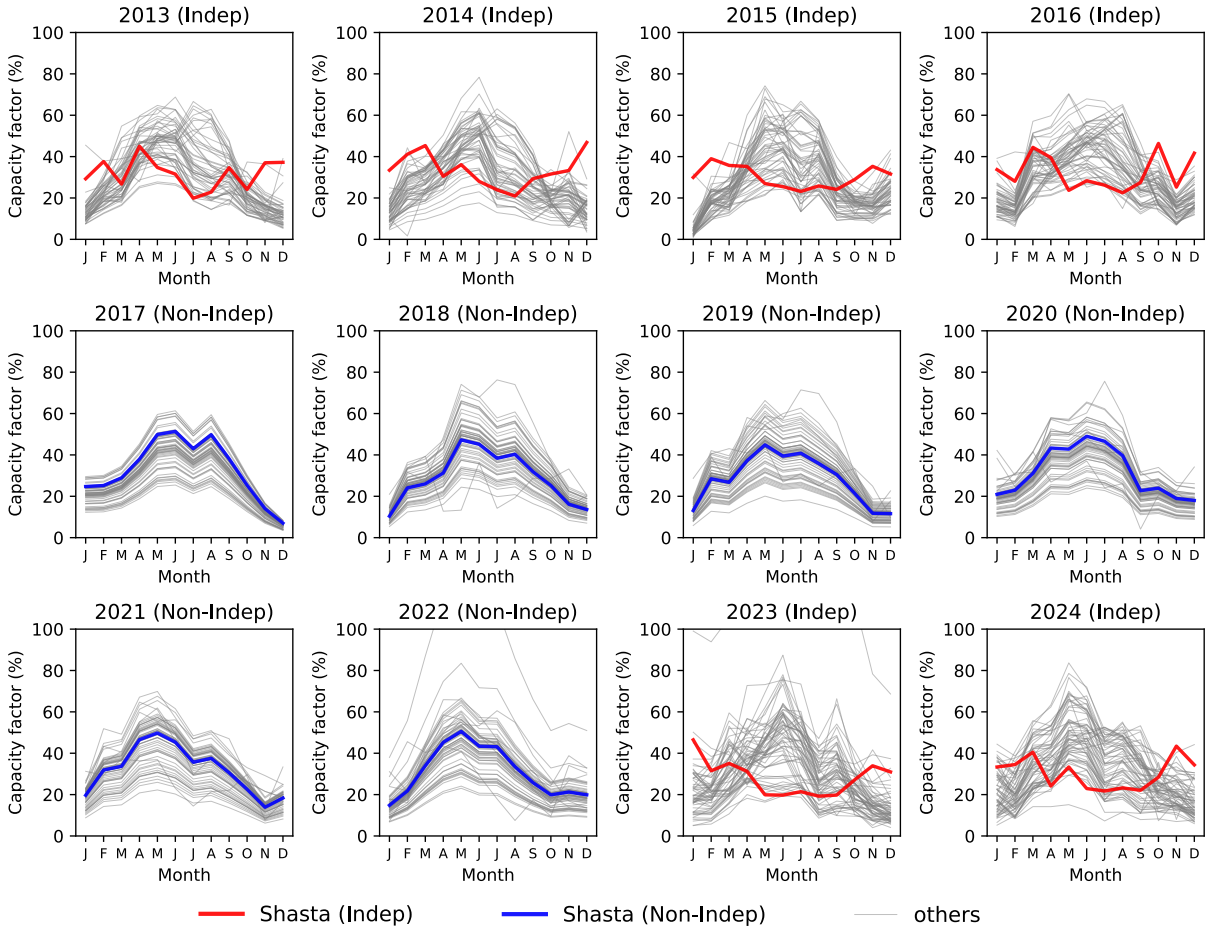

Figure A2. Determination of whether the Shasta plant is independent in each year based on the Pearson correlation distance for each plant pair in that year. When Shasta is independent, it is shown as a thick red line; when non-independent, as a thick blue line. All other valid plants are shown as thin gray lines.

Based on this analysis, we conclude that removing years when scaled plants are too dominant is logically more correct for seasonal pattern analysis. However, removing those years would reduce the sample size for annual-mean discrepancy analysis. Therefore, we set an "independent-plant ratio" threshold of $N_{independent} > 15\%$, which leads to the exclusion of the following years: wind: {2017, 2018, 2019, 2020, 2021, 2022}, solar: {2013, 2015, 2017, 2018, 2019}. Figure A3 compares the seasonal cycle of wind and solar CFs before and after applying the "independent-plant ratio > 15%" threshold. However, this distance-correlation filter cannot be applied to PLUSWIND, as it would remove all available years (2018–2021).

Comparing results with and without this threshold shows that the largest impact is on the Shasta wind plant (Fig. A3a): without the threshold, Shasta exhibits an exactly out-of-phase seasonal cycle between all model simulations and the EIA record; with the threshold applied, this artificial behavior disappears entirely, since 2017–2022 are exactly the years when the Shasta wind CF was scaled from other plants (Fig. A2). This correction also affects northern California, shifting the wind

CF peak month from May to July (Fig. A3a), but has little effect on wind in other regions of California or on any PV plants (Fig. A3b).

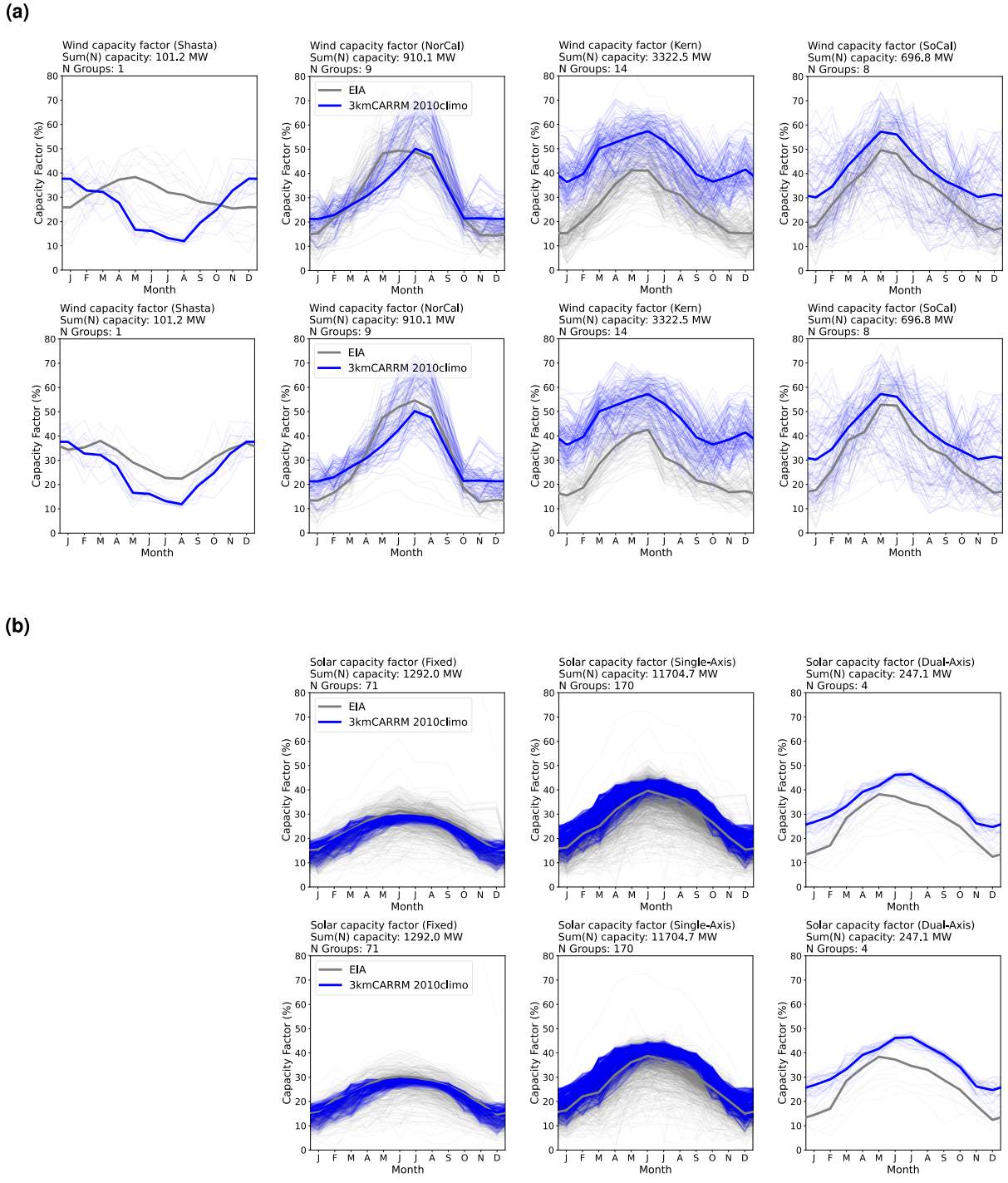

Figure A3. Seasonal variation in the (a) wind and (b) solar CFs (%) estimated using PySAM from SCREAM-3kmCARRM 2010climo, before (upper) and after (bottom) applying the "independent-plant ratio > 15%" threshold. The years before filtering are 2013–~~2023~~2024, while the remaining valid years after filtering are {2013, 2014, 2015, 2016, 2023, 2024} for wind and {2014, 2016, 2020, 2021, 2022, 2023, 2024} for solar, respectively. Each ~~subplot~~ column corresponds to a cluster ~~in Fig. 2~~for wind or solar CFs, with Shasta also included here. Simulations are shown in blue, and EIA records in gray. Light lines represent individual years; bold lines represent the multi-year climatology.

Author contributions. JZ performed the 3.25 km and 800 m SCREAM-CARRM simulations. JCG reran the 2005–2014 E3SM-25kmNARRM to generate additional outputs for energy generation modeling. JZ and JCG developed the analysis plan, with significant input from regular group meetings with MVS, HHL, MM, PAU, RSA, SPC, and JPW, including code design concepts from previous projects and public data sharing. MVS developed and maintained the PySAM preprocessing pipeline for generation modeling. PB, JZ, and PCS secured computing-resource funding for the SCREAM-800mCARRM simulations; JZ and PB designed the SCREAM-CARRM climate-length experiments. JPW was responsible for overall funding acquisition. JZ defined the paper scope and prepared the first draft. All co-authors contributed to the manuscript.

Competing interests. At least one of the (co-)authors is a member of the editorial board of Geoscientific Model Development. The authors declare that they have no other competing interests.

Acknowledgements. The development of SCREAM is supported by the Energy Exascale Earth System Model (E3SM) project (https://e3sm.org/), funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research. The authors thank Samuel Liner for downloading and initially preprocessing the HRRR data. This work is supported by LLNL LDRD project [25-SI-007] "Framework for Optimal Critical-Infrastructure Solutions and Decision Support", and LLNL LDRD project [22-SI-008] "Climate Resilience for National Security". LLNL Institutional Computing Grand Challenge program provides the computing support for this work. Work at LLNL was performed under the auspices of the U.S. DOE by LLNL under contract (grant no. DE-AC52-07NA27344; IM release: LLNL-JRNL-2008844).

References

Arthur, R. S., Golaz, J.-C., Lee, H.-H., Wert, J., Signorotti, M., and Watson, J.-P.: High-resolution climate model datasets for energy infrastructure planning in a renewable-dependent future, Journal of Renewable and Sustainable Energy, 17, 2025a.

Arthur, R. S., Rybchuk, A., Juliano, T. W., Rios, G., Wharton, S., Lundquist, J. K., and Fast, J. D.: Evaluating mesoscale model predictions of diurnal speedup events in the Altamont Pass Wind Resource Area of California, Wind Energ. Sci., 10, 1187–1209, https://doi.org/10.5194/wes-10-1187-2025, wES, 2025b.

Bogenschutz, P., Zhang, J., Tang, Q., and Cameron-Smith, P.: Atmospheric River Induced Precipitation in California as Simulated by the Regionally Refined Simplified Convective Resolving E3SM Atmosphere Model, Geoscientific Model Development, 2024.

Bogenschutz, P. A. and Krueger, S. K.: A simplified PDF parameterization of subgrid-scale clouds and turbulence for cloud-resolving models, Journal of Advances in Modeling Earth Systems, 5, 195–211, https://doi.org/10.1002/jame.20018, 2013.

Bogenschutz, P. A., Eldred, C., and Caldwell, P. M.: Horizontal Resolution Sensitivity of the Simple Convection-Permitting E3SM Atmosphere Model in a Doubly-Periodic Configuration, Journal of Advances in Modeling Earth Systems, 15, https://doi.org/10.1029/2022ms003466, 2023.

Caldwell, P. M., Terai, C. R., Hillman, B., Keen, N. D., Bogenschutz, P., Lin, W., Beydoun, H., Taylor, M., Bertagna, L., Bradley, A. M., Clevenger, T. C., Donahue, A. S., Eldred, C., Foucar, J., Golaz, J. C., Guba, O., Jacob, R., Johnson, J., Krishna, J., Liu, W., Pressel, K., Salinger, A. G., Singh, B., Steyer, A., Ullrich, P., Wu, D., Yuan, X., Shpund, J., Ma, H. Y., and Zender, C. S.: Convection-Permitting Simulations With the E3SM Global Atmosphere Model, Journal of Advances in Modeling Earth Systems, 13, https://doi.org/10.1029/2021ms002544, 2021.

Chen, S. Y., Goergen, K., Franssen, H. J. H., Winkler, C., Poll, S., Wahabou, Y. H. Z., Linssen, J., Vereecken, H., Stolten, D., and Heinrichs, H.: Higher Onshore Wind Energy Potentials Revealed by Kilometer-Scale Atmospheric Modeling, Geophysical Research Letters, 51, https://doi.org/ARTN e2024GL110122 10.1029/2024GL110122, 2024.

Chow, F., Schär, C., Ban, N., Lundquist, K., Schlemmer, L., and Shi, X.: Crossing Multiple Gray Zones in the Transition from Mesoscale to Microscale Simulation over Complex Terrain, Atmosphere, 10, https://doi.org/10.3390/atmos10050274, 2019.

Davidson, M. R. and Millstein, D.: Limitations of reanalysis data for wind power applications, Wind Energy, 25, 1646–1653, 2022.

Davis, N. N., Badger, J., Hahmann, A. N., Hansen, B. O., Mortensen, N. G., Kelly, M., Larsén, X. G., Olsen, B. T., Floors, R., Lizcano, G., Casso, P., Lacave, O., Bosch, A., Bauwens, I., Knight, O. J., Potter van Loon, A., Fox, R., Parvanyan, T., Krohn Hansen, S. B., Heathfield, D., Onninen, M., and Drummond, R.: The Global Wind Atlas: A High-Resolution Dataset of Climatologies and Associated Web-Based Application, Bulletin of the American Meteorological Society, 104, E1507–E1525, https://doi.org/https://doi.org/10.1175/BAMS-D-21-0075.1, 2023.

Donahue, A. S., Caldwell, P. M., Bertagna, L., Beydoun, H., Bogenschutz, P. A., Bradley, A., Clevenger, T. C., Foucar, J., Golaz, C., and Guba, O.: To exascale and beyond—The Simple Cloud-Resolving E3SM Atmosphere Model (SCREAM), a performance portable global atmosphere model for cloud-resolving scales, Journal of Advances in Modeling Earth Systems, 16, e2024MS004 314, 2024.

Dowell, D. C., Alexander, C. R., James, E. P., Weygandt, S. S., Benjamin, S. G., Manikin, G. S., Blake, B. T., Brown, J. M., Olson, J. B., Hu, M., Smirnova, T. G., Ladwig, T., Kenyon, J. S., Ahmadov, R., Turner, D. D., Duda, J. D., and Alcott, T. I.: The High-Resolution Rapid Refresh (HRRR): An Hourly Updating Convection-Allowing Forecast Model. Part I: Motivation and System Description, Weather and Forecasting, 37, 1371–1395, https://doi.org/10.1175/waf-d-21-0151.1, 2022.

El Hammoumi, A., Chtita, S., Motahhir, S., and El Ghzizal, A.: Solar PV energy: From material to use, and the most commonly used techniques to maximize the power output of PV systems: A focus on solar trackers and floating solar panels, Energy Reports, 8, 11 992–12 010, 2022.

Frank, C. W., Pospichal, B., Wahl, S., Keller, J. D., Hense, A., and Crewell, S.: The added value of high resolution regional reanalyses for wind power applications, Renewable Energy, 148, 1094–1109, https://doi.org/10.1016/j.renene.2019.09.138, 2020.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., and Reichle, R.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), Journal of climate, 30, 5419–5454, 2017.

Gilman, P., Dobos, A., DiOrio, N., Freeman, J., Janzou, S., and Ryberg, D.: SAM photovoltaic model technical reference update, NREL: Golden, CO, USA, 2018a.

Gilman, P., Dobos, A., DiOrio, N., Freeman, J., Janzou, S., and Ryberg, D. S.: SAM photovoltaic model technical reference update, Report, https://docs.nlr.gov/docs/fy18osti/67399.pdf, 2018b.

Golaz, J. C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., Abeshu, G., Anantharaj, V., Asay-Davis, X. S., Bader, D. C., Baldwin, S. A., Bisht, G., Bogenschutz, P. A., Branstetter, M., Brunke, M. A., Brus, S. R., Burrows, S. M., Cameron-Smith, P. J., Donahue, A. S., Deakin, M., Easter, R. C., Evans, K. J., Feng, Y., Flanner, M., Foucar, J. G., Fyke, J. G., Griffin, B. M., Hannay, C., Harrop, B. E., Hoffman, M. J., Hunke, E. C., Jacob, R. L., Jacobsen, D. W., Jeffery, N., Jones, P. W., Keen, N. D., Klein, S. A., Larson, V. E., Leung, L. R., Li, H. Y., Lin, W. Y., Lipscomb, W. H., Ma, P. L., Mahajan, S., Maltrud, M. E., Mametjanov, A., McClean, J. L., McCoy, R. B., Neale, R. B., Price, S. F., Qian, Y., Rasch, P. J., Eyre, J. E. J. R., Riley, W. J., Ringler, T. D., Roberts, A. F., Roesler, E. L., Salinger, A. G., Shaheen, Z., Shi, X. Y., Singh, B., Tang, J. Y., Taylor, M. A., Thornton, P. E., Turner, A. K., Veneziani, M., Wan, H., Wang, H. L., Wang, S. L., Williams, D. N., Wolfram, P. J., Worley, P. H., Xie, S. C., Yang, Y., Yoon, J. H., Zelinka, M. D., Zender, C. S., Zeng, X. B., Zhang, C. Z., Zhang, K., Zhang, Y., Zheng, X., Zhou, T., and Zhu, Q.: The DOE E3SM Coupled Model Version 1: Overview and Evaluation at Standard Resolution, Journal of Advances in Modeling Earth Systems, 11, 2089–2129, https://doi.org/10.1029/2018ms001603, 2019.

Graham, A. C.: Anthology of Chinese Literature: Volume I: From Early Times to the Fourteenth Century, Grove Press, New York, 1965.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., and Schepers, D.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, 2020.

Honnert, R., Efstathiou, G. A., Beare, R. J., Ito, J., Lock, A., Neggers, R., Plant, R. S., Shin, H. H., Tomassini, L., and Zhou, B. W.: The Atmospheric Boundary Layer and the "Gray Zone" of Turbulence: A Critical Review, Journal of Geophysical Research-Atmospheres, 125, https://doi.org/ARTN e2019JD030317 10.1029/2019JD030317, 2020.

Hunke, E., Lipscomb, W., Turner, A., Jeffery, N., and Elliott, S.: CICE: The Los Alamos sea ice model, documentation and software, Report, version 4.0, Tech. Rep. LA-CC-06-012, Los Alamos National Laboratory, 2008.

IEA: Renewables 2024: Analysis and Forecast to 2030, Tech. rep., International Energy Agency, Paris, https://www.iea.org/reports/renewables-2024, licence: CC BY 4.0, 2024.

IEA: Global Energy Review 2025, Tech. rep., International Energy Agency, Paris, https://www.iea.org/reports/global-energy-review-2025, licence: CC BY 4.0, 2025.

Jourdier, B., Diaz, C., and Dubus, L.: Evaluation of CERRA for wind energy applications, Report, Copernicus Meetings, 2023.

Juliano, T. W., Jiménez, P. A., Eghdami, M., Haupt, S. E., and Martilli, A.: "Gray Zone" Simulations Using a Three-Dimensional Planetary Boundary Layer Parameterization in the Weather Research and Forecasting Model, Monthly Weather Review, 150, 1585–1619, https://doi.org/10.1175/Mwr-D-21-0164.1, 2022.

Kriegler, E., Bauer, N., Popp, A., Humpenöder, F., Leimbach, M., Strefler, J., Baumstark, L., Bodirsky, B. L., Hilaire, J., Klein, D., Mouratiadou, I., Weindl, I., Bertram, C., Dietrich, J.-P., Luderer, G., Pehl, M., Pietzcker, R., Piontek, F., Lotze-Campen, H., Biewald, A., Bonsch, M., Giannousakis, A., Kreidenweis, U., Müller, C., Rolinski, S., Schultes, A., Schwanitz, J., Stevanovic, M., Calvin, K., Emmerling, J., Fujimori, S., and Edenhofer, O.: Fossil-fueled development (SSP5): An energy and resource intensive scenario for the 21st century, Global Environmental Change, 42, 297–315, https://doi.org/https://doi.org/10.1016/j.gloenvcha.2016.05.015, 2017.

Lauritzen, P. H., Bacmeister, J. T., Callaghan, P. F., and Taylor, M. A.: NCAR_Topo (v1.0): NCAR global model topography generation software for unstructured grids, Geoscientific Model Development, 8, 3975–3986, https://doi.org/10.5194/gmd-8-3975-2015, 2015.

Lee, H.-H., Arthur, R., Golaz, J.-C., Signorotti, M., Wert, J., and Watson, J.-P.: Using Multiple High-Resolution Datasets to Benchmark the Energy Exascale Earth System Model (E3SM) for Renewable Resource Assessment, AIP Advances, under review, 2025.

Leung, L. R., Bader, D. C., Taylor, M. A., and McCoy, R. B.: An Introduction to the E3SM Special Collection: Goals, Science Drivers, Development, and Analysis, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/ARTN e2019MS001821 10.1029/2019MS001821, 2020.

Mahfouz, N., Beydoun, H., Mülmenstädt, J., Keen, N., Varble, A. C., Bertagna, L., Bogenschutz, P., Bradley, A., Christensen, M. W., and Clevenger, T. C.: Prescribing the aerosol effective radiative forcing in the Simple Cloud-Resolving E3SM Atmosphere Model v1, EGUsphere, 2025, 1–25, 2025.

Mayer, M. J. and Gróf, G.: Extensive comparison of physical models for photovoltaic power forecasting, Applied Energy, 283, 116 239, 2021.

McKenna, R., Pfenninger, S., Heinrichs, H., Schmidt, J., Staffell, I., Bauer, C., Gruber, K., Hahmann, A. N., Jansen, M., and Klingler, M.: High-resolution large-scale onshore wind energy assessments: A review of potential definitions, methodologies and future research needs, Renewable Energy, 182, 659–684, 2022.

Millstein, D., Bolinger, M., and Wiser, R.: What can surface wind observations tell us about interannual variation in wind energy output?, Wind Energy, 25, 1142–1150, https://doi.org/10.1002/we.2717, 2022.

Millstein, D., Jeong, S., Ancell, A., and Wiser, R.: A database of hourly wind speed and modeled generation for US wind plants based on three meteorological models, Scientific Data, 10, https://doi.org/ARTN 883 10.1038/s41597-023-02804-w, 2023.

Morrison, H. and Milbrandt, J. A.: Parameterization of Cloud Microphysics Based on the Prediction of Bulk Ice Particle Properties. Part I: Scheme Description and Idealized Tests, Journal of the Atmospheric Sciences, 72, 287–311, https://doi.org/10.1175/jas-d-14-0065.1, 2015.

NREL: PySAM Version 7.0.0, https://github.com/NREL/pysam, accessed: Aug 08, 2025.

Pincus, R., Mlawer, E. J., and Delamere, J. S.: Balancing Accuracy, Efficiency, and Flexibility in Radiation Calculations for Dynamical Models, Journal of Advances in Modeling Earth Systems, 11, 3074–3089, https://doi.org/https://doi.org/10.1029/2019MS001621, 2019.

Pronk, V., Bodini, N., Optis, M., Lundquist, J. K., Moriarty, P., Draxl, C., Purkayastha, A., and Young, E.: Can reanalysis products outperform mesoscale numerical weather prediction models in modeling the wind resource in simple terrain?, Wind Energy Science, 7, 487–504, https://doi.org/10.5194/wes-7-487-2022, 2022.

Rand, J. T., Kramer, L. A., Garrity, C. P., Hoen, B., Diffendorfer, J. E., Hunt, H. E., and Spears, M.: A continuously updated, geospatially rectified database of utility-scale wind turbines in the United States, Scientific Data, 7, https://doi.org/10.1038/s41597-020-0353-6, 2020.

Rodrigo, J. S., Arroyo, R. A. C., Moriarty, P., Churchfield, M., Kosovic, B., Réthoré, P. E., Hansen, K. S., Hahmann, A., Mirocha, J. D., and Rife, D.: Mesoscale to microscale wind farm flow modeling and evaluation, Wiley Interdisciplinary Reviews-Energy and Environment, 6, https://doi.org/ARTN e214 10.1002/wene.214, 2017.

Ryberg, D. S., Caglayan, D. G., Schmitt, S., Linssen, J., Stolten, D., and Robinius, M.: The future of European onshore wind energy potential: Detailed distribution and simulation of advanced turbine designs, Energy, 182, 1222–1238, https://doi.org/10.1016/j.energy.2019.06.052, 2019.

Shaw, W. J., Berg, L. K., Cline, J., Draxl, C., Djalalova, I., Grimit, E. P., Lundquist, J. K., Marquis, M., McCaa, J., Olson, J. B., Sivaraman, C., Sharp, J., and Wilczak, J. M.: The Second Wind Forecast Improvement Project (WFIP2): General Overview, Bulletin of the American Meteorological Society, 100, 1687–1699, https://doi.org/10.1175/Bams-D-18-0036.1, 2019.

Staadecker, M., Szinai, J., Sánchez-Pérez, P. A., Kurtz, S., and Hidalgo-Gonzalez, P.: The value of long-duration energy storage under various grid conditions in a zero-emissions future, Nature Communications, 15, https://doi.org/ARTN 9501 10.1038/s41467-024-53274-6, 2024.

Tang, Q., Klein, S. A., Xie, S. C., Lin, W. Y., Golaz, J. C., Roesler, E. L., Taylor, M. A., Rasch, P. J., Bader, D. C., Berg, L. K., Caldwell, P., Giangrande, S. E., Neale, R. B., Qian, Y., Riihimaki, L. D., Zender, C. S., Zhang, Y. Y., and Zheng, X.: Regionally refined test bed in E3SM atmosphere model version 1 (EAMv1) and applications for high-resolution modeling, Geoscientific Model Development, 12, 2679–2706, https://doi.org/10.5194/gmd-12-2679-2019, 2019.

Tang, Q., Golaz, J.-C., Van Roekel, L. P., Taylor, M. A., Lin, W., Hillman, B. R., Ullrich, P. A., Bradley, A. M., Guba, O., Wolfe, J. D., Zhou, T., Zhang, K., Zheng, X., Zhang, Y., Zhang, M., Wu, M., Wang, H., Tao, C., Singh, B., Rhoades, A. M., Qin, Y., Li, H.-Y., Feng, Y., Zhang, Y., Zhang, C., Zender, C. S., Xie, S., Roesler, E. L., Roberts, A. F., Mametjanov, A., Maltrud, M. E., Keen, N. D., Jacob, R. L., Jablonowski, C., Hughes, O. K., Forsyth, R. M., Di Vittorio, A. V., Caldwell, P. M., Bisht, G., McCoy, R. B., Leung, L. R., and Bader, D. C.: The fully coupled regionally refined model of E3SM version 2: overview of the atmosphere, land, and river results, Geoscientific Model Development, 16, 3953–3995, https://doi.org/10.5194/gmd-16-3953-2023, 2023.

Taylor, M. A., Guba, O., Steyer, A., Ullrich, P. A., Hall, D. M., and Eldrid, C.: An Energy Consistent Discretization of the Nonhydrostatic Equations in Primitive Variables, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/ARTN e2019MS001783 10.1029/2019MS001783, 2020.

Trenberth, K. E., Berry, J. C., and Buja, L. E.: Vertical interpolation and truncation of model-coordinate data, National Center for Atmospheric Research, Climate and Global Dynamics Division, https://doi.org/https://doi.org/10.5065/D6HX19NH, 1993.

Ullrich, P. and Roesler, E.: ClimateGlobalChange/squadgen: v1.2.2 (v1.2.2), https://doi.org/https://doi.org/10.5281/zenodo.13241731, 2024.

Ullrich, P. A. and Taylor, M. A.: Arbitrary-order conservative and consistent remapping and a theory of linear maps: Part I, Monthly Weather Review, 143, 2419–2440, 2015.

Ullrich, P. A., Devendran, D., and Johansen, H.: Arbitrary-order conservative and consistent remapping and a theory of linear maps: Part II, Monthly Weather Review, 144, 1529–1549, 2016.

Wang, M. N., Ullrich, P., and Millstein, D.: The future of wind energy in California: Future projections with the Variable-Resolution CESM, Renewable Energy, 127, 242–257, https://doi.org/10.1016/j.renene.2018.04.031, 2018.

Wyngaard, J. C.: Toward numerical modeling in the "terra incognita", Journal of the Atmospheric Sciences, 61, 1816–1826, https://doi.org/Doi 10.1175/1520-0469(2004)061<1816:Tnmitt>2.0.Co;2, 2004.

Zender, C. S.: Analysis of self-describing gridded geoscience data with netCDF Operators (NCO), Environmental Modelling & Software, 23, 1338–1342, 2008.

Zhang, J.: Code, model, and analysis data for simulation of wind and solar energy generation over California with E3SM SCREAM regionally refined models at 3.25 km and 800 m resolutions [Data set]., https://doi.org/10.5281/zenodo.16809290, 2025.

Zhang, J., Bogenschutz, P., Tang, Q., Cameron-smith, P., and Zhang, C.: Leveraging regional mesh refinement to simulate future climate projections for California using the Simplified Convection-Permitting E3SM Atmosphere Model Version 0, Geoscientific Model Development, 17, 3687–3731, 2024.

Zhang, J., Bogenschutz, P., Taylor, M., and Cameron-Smith, P.: Pushing the Simplified Convection Permitting E3SM Atmosphere Model to 100 m by leveraging regional mesh refinement over the Bay Area of California, Geoscientific Model Development, 2025.

Zheng, X., Li, Q., Zhou, T., Tang, Q., Van Roekel, L. P., Golaz, J. C., Wang, H. L., and Cameron-Smith, P.: Description of historical and future projection simulations by the global coupled E3SMv1.0 model as used in CMIP6, Geoscientific Model Development, 15, 3941–3967, https://doi.org/10.5194/gmd-15-3941-2022, 2022.