https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

CERFRES v1.0: a physics-based modeling framework for farm-level renewable energy generation in Europe

Emil Bruvik[1,*] and Asgeir Sorteberg[1]

[1]Geophysical Institute, University of Bergen (UiB), Bergen, Norway

Correspondence: Emil Bruvik (emil.bruvik@uib.no)

Abstract. We present CERFRES (CERRA-derived Farm-level Renewable Energy Systems generation for Europe), a physics-based modeling framework for simulating hourly renewable energy generation from utility-scale and distributed wind and solar installations across Europe. The framework combines generator-level metadata with the high-resolution Copernicus European Regional ReAnalysis (CERRA, 5.5 × 5.5 km) to convert local meteorological conditions into per-plant power output via physically motivated models for wind speed extrapolation, multi-turbine power aggregation, and photovoltaic performance. The resulting time series span the 1995–2025 period at the individual asset level, alongside national and bidding-zone aggregations, covering 8,203 wind farms and 22,190 solar PV installations. Because historical generation records at individual plant level are rarely disclosed by transmission system operators or plant owners, a validated open-access modeling framework capable of reproducing sub-national variability provides critical infrastructure for power-system research. Validation against ENTSO-E reported actual generation demonstrates strong temporal agreement across Europe; solar PV modeling achieves Pearson correlation coefficients exceeding 0.95 in most major national markets. A case-study farm-level validation against 61 Norwegian onshore wind farms for January 2023 yields correlation coefficients of 0.78–0.93 at individual plant level, indicating that the 5.5 km meteorological forcing can resolve sub-national variability. Benchmarking against EMHIRES and Renewables Ninja, two widely used open-access European renewable generation datasets, shows country-mean normalized RMSE reductions for solar PV of 31% and 37%, with corresponding normalized MAE reductions of 34% and 39%, over the commonly covered countries.

1 Introduction

The growing share of variable renewable energy in European electricity systems creates demand for physically consistent, spatially explicit generation models that can reproduce farm-level output at high resolution. Accurate representation of spatial and temporal generation variability is critical for power-system analyses involving transmission constraints, flexibility dispatch, hydro–wind coordination, and long-term capacity planning. Existing approaches often rely on coarse-resolution meteorological inputs or administrative-level aggregations that obscure the sub-national variability essential for such applications. In this paper, we present CERFRES v1.0, a physics-based modeling framework for simulating hourly renewable energy generation at the individual plant level across Europe. The framework covers utility-scale and distributed solar photovoltaic (PV) generation, as well as onshore, offshore (bottom-fixed), and floating offshore wind generation.

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

Several datasets have been developed to support large-scale renewable energy modeling. The EMHIRES dataset provides hourly wind and solar generation time series at a regional level across Europe, based on CM-SAF SARAH data (5 km) for solar (Gonzalez Aparicio et al., 2017) and MERRA data (approximately 70×60 km) for wind (Gonzalez Aparicio et al., 2016). The dataset is derived from meteorological inputs using spatial downscaling methods to produce generation profiles at the NUTS-2 level, capturing localized resource variability. A report by Schmitz and Després (2022) provides a global dataset of hourly wind and solar time series from 2004 to 2018 calculated with the MERRA-2 data using different clustering methods.

Similarly, the RE-Europe dataset offers an integrated, continent-scale collection of renewable generation and demand time series designed for future European power system studies (Jensen and Pinson, 2017). Their work combines a transmission network model with scalable 3-year weather-driven renewable forecasts and conventional generation data to analyze varying penetration scenarios. More recently, open modeling frameworks such as PyPSA-Eur have combined weather-based renewable availability with transmission system representations to enable reproducible system optimization studies (Hoersch et al., 2018). These efforts have advanced the accessibility and consistency of renewable energy modeling data.

A fundamental obstacle for renewable energy modeling is the limited availability of historical generation data at the individual plant level. While aggregated national or bidding-zone generation is reported through platforms such as the ENTSO-E Transparency Platform, records for individual wind and solar farms are rarely disclosed publicly. This reflects both commercial sensitivity and the governance structure of national transmission system operators (TSOs), which are not uniformly required to publish plant-level metered output. As a result, validated models that can reproduce hourly generation at the individual asset level are essential for applications that cannot rely on measured plant-level data. Such applications can include long-term scenario analysis, power flow studies in networks not yet constructed, and sensitivity analyses over weather years predating the current fleet. The farm-level validation performed in this study, using plant-level records from the Norwegian NVE wind register and sub-national data from Denmark, is unusual in the literature precisely because such data are so difficult to obtain.

Despite this progress, important gaps remain for applications requiring both high temporal resolution and detailed spatial representation. Many existing datasets rely on spatial aggregation at administrative or transmission-zone levels, which can smooth out localized variability that is critical when assessing congestion, curtailment, or the interaction between renewable generation and flexible resources such as hydropower. Wind speed variability is often significantly underestimated when using coarse-resolution data such as MERRA. Statistical downscaling techniques, such as those presented in (González-Aparicio et al., 2017), have demonstrated the importance of high-resolution reanalysis data for power system modeling. In addition, differences in reanalysis products, interpolation methods, height extrapolation, and technology assumptions have been observed to cause systematic biases when datasets are used directly in operational or market-oriented models. Furthermore, no existing open framework combines geo-located farm-level metadata with a high-resolution regional reanalysis and a physics-based conversion chain validated at both national and individual plant scales.

The present study addresses these limitations by developing high-resolution wind and solar photovoltaic generation time series for Europe based on the Copernicus European Regional ReAnalysis (CERRA) (Ridal et al., 2024). Meteorological variables are processed at the grid-cell level and converted into power production using physically motivated models for wind speed extrapolation, turbine power aggregation, and photovoltaic performance. Spatial interpolation is applied to account for

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

the mismatch between reanalysis grid points and the actual locations of renewable energy installations. A location-specific correction factor is introduced to reconcile modeled output with reported generation at the bidding-zone level.

The resulting modeling framework is designed to support power system analyses that are sensitive to transmission constraints and flexibility requirements, such as market simulations, hydro–wind coordination studies, and offshore wind integration assessments. By combining transparent methodologies, open meteorological data, and validation against observed generation, this work provides a reproducible and system-oriented alternative to existing regional datasets. The model is intended both as a standalone research product and as an input to large-scale energy system models requiring realistic representations of renewable variability.

2 Data sources

2.1 Generator metadata

Information on installed renewable generation capacity is required to convert meteorological inputs into power generation time series. Generator-level metadata used in this study are obtained from the Global Energy Monitor (GEM) solar and wind tracker databases (Global Energy Monitor, 2026a, b). This database provides geospatial and technical information for individual installations or aggregated projects, including location (latitude/longitude), technology type, commissioning year, and nameplate capacity.

The level of technological detail available in GEM varies across countries and technologies. In particular, plant-specific technical parameters such as turbine specifications or PV module characteristics are often incomplete or unavailable. Consequently, the modeling framework does not attempt to represent installation-specific configurations. Instead, generators are categorized by technology class (onshore wind, offshore wind, solar PV, solar thermal) and commissioning period, and are assigned representative technology parameters consistent with typical European installations.

For solar PV, the dataset includes both utility-scale installations and distributed systems. Utility-scale projects are reported at the facility level for installations larger than 1 MW, while distributed PV systems below 1 MW are represented as nationally aggregated capacities. Installations are modeled as fixed-tilt systems with representative tilt and azimuth angles derived from latitude-dependent relationships, with single-axis tracking activated for large modern installations. These assumptions are applied uniformly across all locations to ensure methodological consistency and comparability between regions. For wind power, hub heights and power curves are parameterized as a function of commissioning year and capacity class, reflecting the historical trend toward larger turbines and increased hub heights (see Table 1).

Two scenario variants are released. The fixed reference scenario applies the full 2025 fleet to every year in the meteorological record. The as-built scenario instead includes only installations whose reported commissioning year is finite and not later than the simulated year, after filtering the source records by a country-specific set of status classes treated as active in the workflow. In practice, this means that the historical fleet is reconstructed from the available metadata rather than from a complete month-by-month operating history. Installations with missing commissioning year are excluded from the as-built subset. The as-built product should therefore be interpreted as a best-effort metadata-constrained reconstruction of the historical fleet, with the

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

largest uncertainty expected in early years and in countries where smaller or distributed assets are incompletely represented in the underlying registries.

To assess the quality of the GEM metadata, a comparison was performed for Norway using the Norwegian Water Resources and Energy Directorate (NVE) wind farm registry as ground truth (NVE, 2026). The NVE registry lists 64 operating onshore wind farms with a combined installed capacity of 5082 MW, representing a near-complete national inventory. Restricting the GEM wind tracker to onshore entries with operating status and a valid commissioning year yields 61 farms totalling 4891 MW, an aggregate underestimate of 191 MW (-3.8%) relative to NVE. Among matched farms, 66% show exact or near-exact capacity agreement ($|\Delta C| \leq 1$ MW). The largest discrepancies – Sørmarkfjellet (-117 MW), Smøla (-106 MW), and Midtfjellet (-97 MW) are multi-phase expansions recorded at original rather than updated capacity.

2.2 Weather data sources and processing

Meteorological forcing is derived from the Copernicus European Regional ReAnalysis (CERRA), which is a regional reanalysis product over Europe provided at high spatial resolution (5.5 $\times$ 5.5 km) (Ridal et al., 2024). CERRA covers the period from mid-1984 to present, making it suitable for both historical validation and future scenario analysis. The wind power model uses CERRA wind speeds at 30, 50, 75, 100, 150, and 200 m to represent the vertical wind profile and estimate hub-height conditions. PV generation is modeled using 10-m wind speed, 2-m temperature, and surface solar radiation. CERRA has been extensively validated against observations and found to outperform ERA5 in regions with complex terrain (Pelosi, 2023).

2.3 Generation data for validation

Validation of the modeled renewable generation time series is conducted using historical generation data from the European Network of Transmission System Operators for Electricity (ENTSO-E) Transparency Platform (ENTSO-E, 2026). This platform provides aggregated generation by fuel type for each European bidding zone. These data serve as ground truth for temporal validation and as the basis for the per-country bias-correction factors described in Sect. 3.5. However, the platform has well-documented data quality limitations. Coverage is entirely absent for several countries (Albania, Ukraine, Georgia), substantially incomplete for others (United Kingdom before 2021, Republic of Ireland, Croatia), and inconsistencies between the platform and alternative sources such as Eurostat and national TSO websites are common (Hirth et al., 2018). These inconsistencies partly reflect the platform's governance structure, in which primary data owners submit data through intermediary providers with limited quality enforcement. These limitations imply that bias-correction factors derived from ENTSO-E actuals may themselves carry errors in countries with known reporting inconsistencies, and validation metrics should be interpreted with this caveat in mind.

For the localized validation, two additional observational datasets are used. First, hourly production records for 61 individual Norwegian onshore wind farms are taken from the NVE windpower database (NVE, 2026). These historical profiles are matched to CERFRES farm-level simulations by installation coordinates and are used as a plant-level case study for January 2023. Second, hourly solar PV generation for the Danish administrative regions Midtjylland and Sjælland is obtained from the Energi Data Service "Production per Region per Hour" dataset (Energinet, 2026). These observations are compared with CER-

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

FRES output aggregated to the corresponding NUTS-2 regions over a two-week period in June 2023. Because such localized validation data are only available for a small subset of regions and technologies, they are used as targeted case studies rather than as Europe-wide evaluation datasets.

3 Model description

The modeling approach consists of several key steps, as outlined in Figure 1: (1) spatial aggregation of generator metadata to the reanalysis grid, (2) conversion of meteorological variables to power generation using technology-specific models, (3) spatial interpolation to account for the mismatch between reanalysis grid points and actual installation locations, and (4) application of location-specific correction factors to scale modeled output with observed generation at the bidding-zone level.

3.1 Spatial aggregation of generator metadata

The CERRA reanalysis data is provided on a fixed grid, whereas wind and solar power plants are located at specific coordinates that rarely align perfectly with the grid nodes. To estimate meteorological variables at the exact location of each power plant, an Inverse Distance Weighting (IDW) interpolation method is used. The interpolated value $u(\mathbf{x})$ at the power plant location $\mathbf{x}$ is calculated as a weighted average of the values u_i observed at the nearest grid points:

$$u(\mathbf{x}) = \begin{cases} u_i & \text{if } d(\mathbf{x}, \mathbf{x}_i) = 0 \\ \frac{\sum_{i=1}^{N} w_i(\mathbf{x}) u_i}{\sum_{i=1}^{N} w_i(\mathbf{x})} & \text{if } d(\mathbf{x}, \mathbf{x}_i) > 0 \end{cases} \tag{1}$$

The weights w_i are inversely proportional to the distance d_i between the plant and the grid node i, raised to a power parameter p:

$$w_i(\mathbf{x}) = \frac{1}{d(\mathbf{x}, \mathbf{x}_i)^p} \tag{2}$$

The interpolation used considers the $N = 4$ nearest grid neighbors surrounding the point of interest. A power parameter of 2 is used.

3.2 Spatial aggregation to power system nodes

Energy system models typically operate on a reduced spatial representation in which generators and loads are connected to a finite set of network nodes, such as bidding zones or transmission substations. To ensure compatibility with such models, the high-resolution renewable generation time series are aggregated from grid-cell level to predefined power system nodes. Each wind and solar installation is assigned to a corresponding bidding zone based on its geographic location for a direct comparison with ENTSO-E data. Meteorological variables are first interpolated to the installation coordinates using the four nearest neighbors. Then, they are converted to power output at the grid-cell level, and finally aggregated across all installations within the same bidding zone. As we have no more detailed information about the specific production of individual sites, this approach preserves spatial heterogeneity in weather-driven variability while producing time series compatible with zonal

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

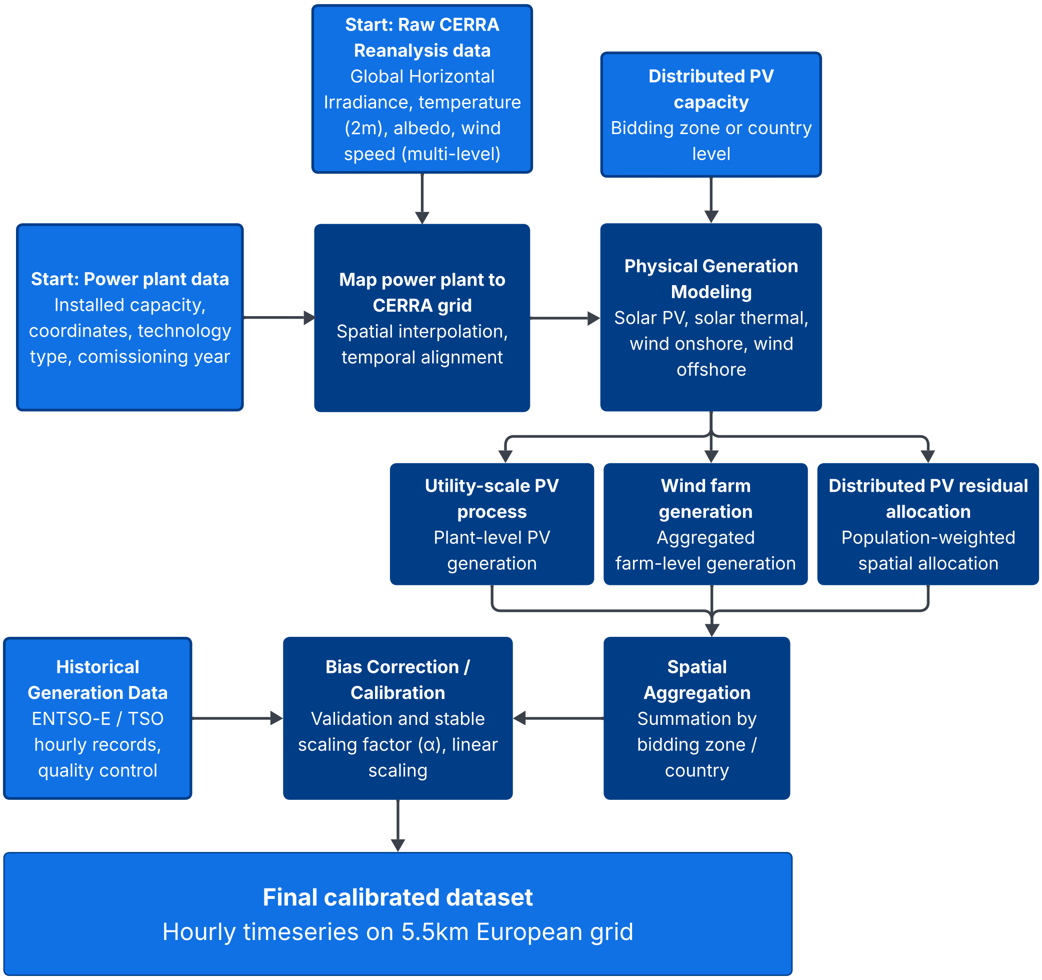


Figure 1. Summary of the procedure used to calculate renewable power generation time series.

155 market and network models. The zonal market compatibility is especially important for energy system modeling and balancing net power demand changes.

Formally, the aggregated renewable generation $P_z(t)$ for bidding zone z at time t is computed as

$$P_z(t) = \sum_{i \in \mathcal{I}_z} P_i(t), \tag{3}$$

where $\mathcal{I}_z$ denotes the set of wind or solar installations located within zone z, and $P_i(t)$ is the estimated power output of

160 installation i at time t.

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

3.3 Calculation of PV generation

The estimation of PV energy production is performed using the PVlib library, an open-source software toolbox for PV performance simulation (Holmgren et al., 2018). PVlib uses a sequential chain of physical models to convert meteorological reanalysis data into electrical output. First, the weather variables, including global horizontal irradiance (GHI), 10-m wind speed, 2-m temperature, and albedo, are extracted for each site using 4-neighbor inverse distance weighting interpolation or direct grid-cell extraction. Missing records, which occur only as rare, isolated hours in the CERRA archive, are set to zero across all weather variables; for irradiance this conservatively yields zero generation in the affected hour, while the impact of the corresponding zero-filled temperature and wind speed on the temperature-dependent conversion is negligible at aggregate scales. A minimum irradiance threshold ($I_{\min} = 10\ \mathrm{W\,m^{-2}}$) is additionally applied to filter out low, non-physical positive values.

Next, the solar geometry and irradiance components are determined. Because CERRA irradiance is interval-accumulated, the solar position is evaluated at the mid-interval ($t - 30$ minutes) to avoid systematic phase errors between the reported irradiance and incidence-angle calculations. The GHI is then decomposed into direct normal irradiance (DNI) and diffuse horizontal irradiance (DHI) using the Erbs model (Erbs et al., 1982).

Once the solar geometry is established, the irradiance on the plane of the PV array (POA) is computed. The array orientation dictates the incidence angle of the solar radiation. For a fixed azimuth angle (a_p) and a fixed tilt angle (t), the plane incidence angle θ is:

$$\cos(\theta) = \sin(h)\cos(t) + \cos(h)\sin(t)\cos(a_p - a_s) \tag{4}$$

where h is the sun altitude and a_s is the sun azimuth angle. The direct and diffuse plane irradiance ($I_{dir,p}$ and $I_{dif,p}$) can then be computed from the global direct and diffuse horizontal irradiance (I_{dir} and I_{dif}) by:

$$I_{dir,p} = I_{dir}\frac{\cos(\theta)}{\sin(h)} \tag{5}$$

$$I_{dif,p} = I_{dif}\left(\frac{1+\cos(t)}{2}\right) + (I_{dir} + I_{dif})\alpha\left(\frac{1-\cos(t)}{2}\right) \tag{6}$$

where α is the surface albedo. In the default model, fixed-tilt orientation is implemented using a latitude-based heuristic ($t = |\phi| - 5°$) and a south-facing azimuth in the Northern Hemisphere. Alternatively, the model activates single-axis tracking for large modern plants (capacity > 50 MW and commissioned ≥ 2019) to continually adjust the arrays along a tracking axis, minimizing the incidence angle throughout the day.

Finally, the power output from the panel is calculated from the in-plane irradiance. Angle-of-incidence losses are represented using the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE IAM) (Souka and Safwat, 1966), with the incidence angle modifier transmission model set to 0.05 for the direct and diffuse components. The cell temperature is modeled using the Sandia PV Array Performance Model (SAPM) (King et al., 2004) temperature models

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

based on ambient conditions ("open-rack glass-glass" coefficients for utility-scale systems and "close-mount glass-glass" for smaller distributed configurations). The DC electrical output is then computed using the National Renewable Energy Laboratory's PVWatts (Version 5) DC power model (Dobos, 2014), with a temperature coefficient of $\gamma_{pdc} = -0.004\,°\mathrm{C}^{-1}$. Additional system losses are dominated by inverter conversion, which converts the DC output into AC power. This conversion is handled by the PVWatts inverter model using a nominal inverter efficiency of 96 % and a DC/AC provisioning ratio of 1.25 to model inverter loading and clipping limits. A sensitivity analysis on these generator characteristics indicated that the reproduction of historical time series is not significantly impacted (at a bidding-zone level) by the exact coefficients used.

3.3.1 Distributed PV generation

In addition to utility-scale solar farms, a significant share of European PV capacity is installed in small, distributed systems such as residential rooftop installations and commercial building-integrated arrays. These installations are not individually geo-referenced in the main generator registry and must therefore be treated separately.

Distributed PV capacity is taken from the Global Solar Power Tracker (GSPT) distributed installations file (Global Energy Monitor, 2026a). This file records national or sub-national aggregates of small-scale PV capacity that are not resolved to individual site coordinates. For each country, the total distributed capacity C_{dist} is extracted as the sum of all operationally active entries in this dataset.

Because the main GSPT file covers only installations that can be geo-located, the combined nameplate capacity of geo-located farms (C_{farm}) is typically lower than the total national PV capacity. Distributed capacity is therefore added as a scaling supplement to the geo-located total. When processing is performed at bidding-zone level, the distributed capacity is apportioned to each zone in proportion to the zone's share of the country's total geo-located capacity:

$$C_{dist,z} = C_{dist} \cdot \frac{C_{farm,z}}{C_{farm,country}}, \tag{7}$$

where $C_{farm,z}$ is the geo-located capacity in zone z and $C_{farm,country}$ is the country total. The effective aggregate capacity for the zone then becomes $C_{farm,z} + C_{dist,z}$, and the modeled generation time series is scaled by the resulting capacity ratio:

$$\alpha_{scale,z} = \frac{C_{farm,z} + C_{dist,z}}{C_{farm,z}}. \tag{8}$$

The power time series used for all distributed capacity in a zone is identical in shape to that derived from the geo-located farm simulations, because the same CERRA irradiance and temperature fields, PV performance model, and system configuration assumptions apply. The capacity scaling factor $\alpha_{scale,z}$ is calculated multiplicatively after the bias-correction step.

The country-aggregated or zone-aggregated distributed PV generation is disaggregated back onto the 5.5 × 5.5 km CERRA grid using spatial weights derived from a population distribution raster (WorldPop, 2026). Cells with higher population density are assigned larger fractions of the distributed generation. Formally, the distributed generation at grid cell (y,x) is:

$$P_{dist}(t,y,x) = P_{dist,total}(t) \cdot w_{pop}(y,x), \tag{9}$$

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

where $w_{pop}(y,x)$ is the population-weighted fraction of the country (or zone) total allocated to cell (y,x). The weights are normalized so that their sum over all grid cells equals 1. When population raster data are unavailable for a given country, the weights fall back to the distribution of geo-located farm capacity across grid cells.

3.4 Calculation of wind generation

Wind speeds are extracted at fixed height intervals from the CERRA reanalysis dataset (30, 50, 75, 100, 150 and 200 m). To estimate wind speeds at the specific hub height of each turbine, the two CERRA levels immediately bracketing the hub height are first identified, and the power law profile is fitted to this pair of levels to derive a local shear exponent. The hub-height wind speed is then obtained by applying this locally fitted profile:

$$w(z) = w_{lower}\left(\frac{z}{z_{lower}}\right)^{\alpha_{local}} \tag{10}$$

where z_{lower} is the lower bracketing level, w_{lower} is the corresponding wind speed, and α_{local} is estimated from the two bracketing levels as:

$$\alpha_{local} = \frac{\ln(w_{upper}/w_{lower})}{\ln(z_{upper}/z_{lower})} \tag{11}$$

Here z_{upper} and w_{upper} are the height and wind speed at the upper bracketing level. The power conversion is performed assuming standard air density (1.225 kg m^{-3}) for simplicity.

3.4.1 Single-turbine power output

The single-turbine approach applies a specific manufacturer (or reference) power curve $P_{single}(v)$ directly to the wind speed at hub height. This relationship is typically defined piecewise:

$$P_{single}(v) = \begin{cases} 0 & v < v_{cut\text{-}in} \\ \frac{1}{2}C_p\rho A v^3 & v_{cut\text{-}in} \leq v < v_{rated} \\ P_{rated} & v_{rated} \leq v < v_{cut\text{-}out} \\ 0 & v \geq v_{cut\text{-}out} \end{cases} \tag{12}$$

where v is the wind speed, ρ is air density, A is the rotor swept area, and C_p is the power coefficient. This curve preserves the cubic response to wind speed fluctuations, capturing the sharp ramp rates and peak output potential.

3.4.2 Multi-turbine power curve

Since the aim is to calculate the power production for a farm of turbines rather than that of a single turbine, an aggregate power curve representing multiple turbines is employed to reduce the variability or "spikes" in the power production time series. This accounts for the fact that turbines within a large farm do not experience the exact same wind speed simultaneously due to spatial distribution (Nørgaard and Holttinen, 2004).

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

The multi-turbine curve is derived by calculating the expected value of the single-turbine power curve, assuming the local wind speeds across the farm follow a Gaussian distribution centered on the mean farm wind speed (Staffell and Pfenninger, 2016). The smoothed power output per turbine, $P_{multi}(w)$, for a mean farm wind speed w is calculated as the convolution of the single-turbine power curve with the Gaussian probability density function:

$$P_{multi}(w) = \int_{0}^{\infty} P_{single}(v) \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(v-w)^2}{2\sigma^2}\right) dv \tag{13}$$

Here, v is the local wind speed experienced by an individual turbine, w is the mean wind speed across the farm (derived from the reanalysis grid), and σ is the standard deviation controlling the degree of spatial wind speed dispersion. In the numerical implementation, this integral is approximated as a sum over discrete wind speed bins using a fixed bin width.

The standard deviation σ is modeled dynamically as a function of the mean farm wind speed w:

$$\sigma = 0.6 + 0.1w \tag{14}$$

Sensitivity analyses on the coefficients in Eq. (14) showed a significant impact, with increasing coefficient values leading to a wider probability distribution, resulting in a higher degree of smoothing and underestimation of peak power.

Due to the lack of plant details, installations are mapped to specific reference curves (e.g., NREL, DTU, IEA) based on their commissioning year and installation type. Table 1 details the technology assumptions used in the model. The representative models are based on the available turbine models in the wind turbine power curve archive by NREL (National Renewable Energy Laboratory, 2025).

3.5 Bias correction

Although the uncalibrated time series reproduce the temporal variability of renewable generation well, their absolute energy volumes can deviate systematically from reported production at the country or bidding-zone level. These discrepancies originate primarily in the generator metadata and in physical processes deliberately excluded from the conversion chain, such as incomplete or outdated capacity records, missing technology detail, unmodeled curtailment, scheduled maintenance, and residual biases in the meteorological forcing. In energy system modeling, both the national renewable fleet and electricity demand are commonly aggregated into a single node, requiring annual generation to be consistent with observations. The purpose of the bias correction is to remove these aggregate, zone-level offsets while leaving the weather-driven temporal profile intact. The factor α_z is defined as the ratio of total observed to total modeled energy over the calibration period:

$$\alpha_z = \frac{\sum_t P_{actual,z}(t)}{\sum_t P_{modeled,z}(t)} \tag{15}$$

where $P_{actual,z}(t)$ is the observed generation for zone z at time t and $P_{modeled,z}(t)$ the corresponding uncorrected modeled generation. Separate factors are derived for each renewable energy technology and each bidding zone, and are applied multi-

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

Table 1. Mapping of commissioning years to representative wind turbine technologies and hub heights.

Installation type	Commissioning year	Representative model	Hub height (m)
Onshore	≤ 2000	Vestas V47 (660 kW)	65
Onshore	2001–2005	DOE GE 1.5 MW (77 m)	80
Onshore	2006–2010	DOE GE 1.5 MW (77 m)	90
Onshore	2011–2015	2017 COE Market Average 2.3 MW (113 m)	100
Onshore	2016–2018	IEA Reference 3.4 MW (130 m)	125
Onshore	2019–2021	2020 ATB NREL Reference 4 MW (150 m)	135
Onshore	2022–2025	2023 NREL Bespoke 6 MW (170 m)	150
Offshore floating	≤ 2020	2020 ATB NREL Reference 5.5 MW (175 m)	110
Offshore floating	2021–2025	2023 NREL Bespoke 8.3 MW (196 m)	130
Offshore (other)	≤ 2005	DOE GE 1.5 MW (77 m)	70
Offshore (other)	2006–2010	2017 COE Market Average 2.3 MW (113 m)	90
Offshore (other)	2011–2015	2020 ATB NREL Reference 4 MW (150 m)	100
Offshore (other)	2016–2019	2020 ATB NREL Reference 5.5 MW (175 m)	110
Offshore (other)	2020–2025	2023 NREL Bespoke 8.3 MW (196 m)	140

plicatively to the entire time series of that zone and technology:

$$P_{corrected,z}(t) = \alpha_z \cdot P_{modeled,z}(t). \tag{16}$$

Because α_z is a single constant per zone and technology, it rescales the aggregate energy to match reported production without changing the shape of the time series; the correlation, the timing of ramps, and the relative variability are all preserved.

The scope of this correction should be emphasized. The factor is a property of the zonal aggregate: it absorbs fleet-wide effects such as capacity underrepresentation, systematic bias, and average curtailment, and is meaningful only when the national or zonal fleet is treated as a single node in energy-system or market studies. It is not intended to be applied to individual farm-level series, as plant-specific outages, wake losses, and curtailment cannot be captured by a zone-average ratio. The uncorrected series should therefore be used for asset-level analysis, with the correction reserved for nodal, system-level studies.

4 Results and discussion

This section evaluates the performance of CERFRES against historical generation records across multiple spatial scales, from national aggregates to individual renewable-energy assets. Temporal accuracy is first evaluated for wind and solar PV at the country and bidding-zone levels, followed by sub-national and farm-level assessments using Norwegian and Danish observations. The section concludes with a benchmark comparison against two open-source datasets, EMHIRES and Renewables Ninja.

The spatial structure of the modeled output is illustrated over northern Europe in Figure 2 for a representative week in July 2025.

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

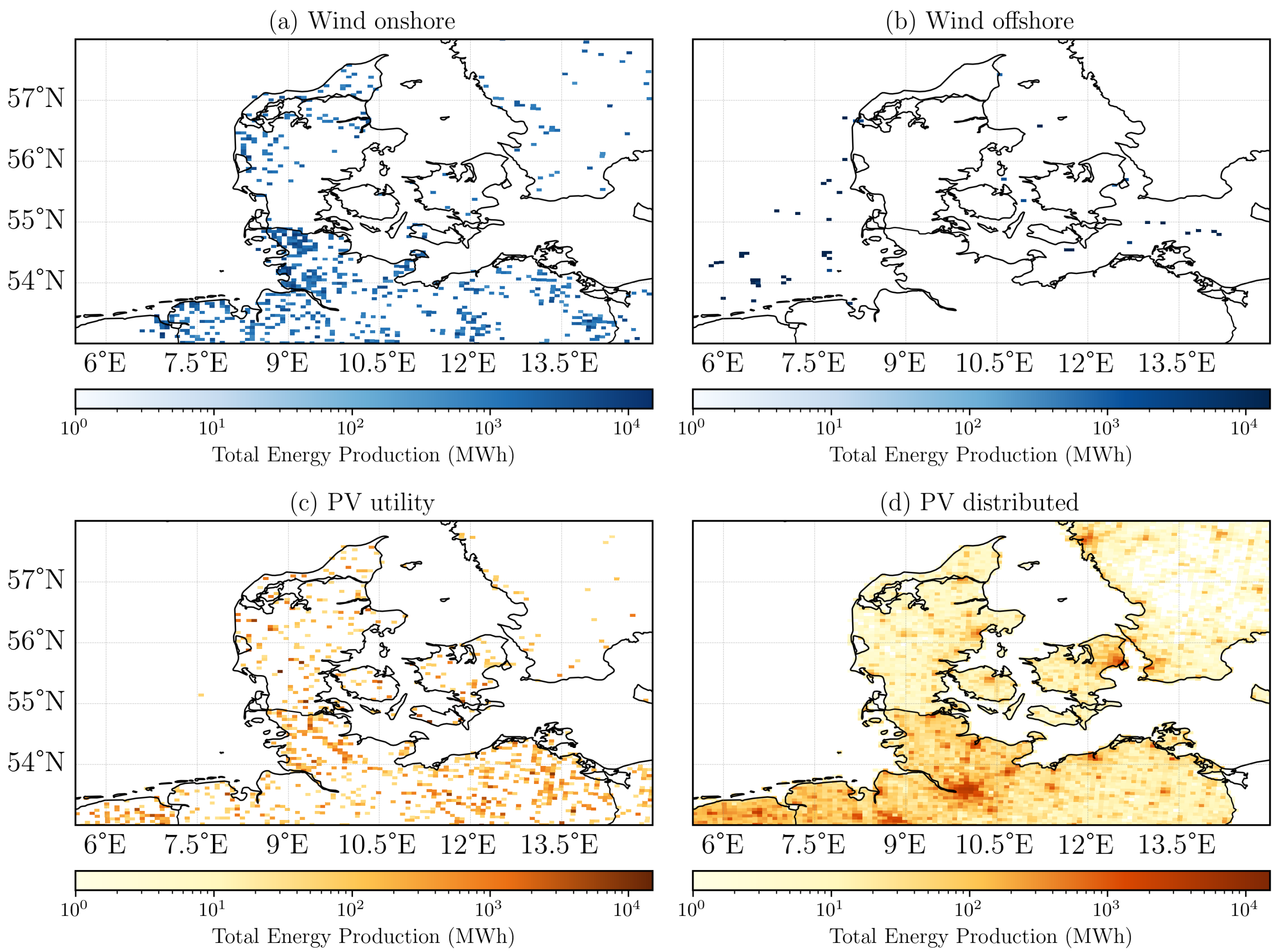


Figure 2. Total wind and solar power generation in northern Europe on the CERRA 5.5 × 5.5 km grid for 1–7 July 2025.

4.1 Performance metrics

The accuracy of the processed dataset is assessed by comparing it with nationally aggregated generation profiles from countries and bidding zones where historic data is available. The validation compares the raw, uncalibrated modeled generation (P_{mod}) with the ENTSO-E time series (P_{act}). All metrics are computed on absolute generation time series aggregated at the country or bidding zone level. This isolates the structural and meteorological alignment errors from capacity-related magnitude errors. The assessment focuses primarily on the temporal accuracy of the model, specifically its ability to capture both everyday diurnal patterns and meteorologically driven ramps. Absolute generation levels are influenced mainly by capacity assumptions and bias correction, whereas temporal metrics provide a more direct measure of the quality of the meteorological forcing and modeling approach.

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

The statistical indicators used to benchmark modeled against measured generation are the normalized mean absolute error (NMAE), the normalized root mean square error (NRMSE), the Pearson correlation coefficient (r), and the fractional bias (FB). Normalizing error metrics by installed capacity C_z allows direct comparison across bidding zones with very different system sizes.

NMAE measures the average absolute deviation between modeled and observed generation, expressed as a fraction of installed capacity:

$$\mathrm{NMAE} = \frac{1}{N \cdot C_z} \sum_{t=1}^{N} |P_{act}(t) - P_{mod}(t)| \tag{17}$$

NRMSE penalizes larger errors more strongly and is similarly normalized:

$$\mathrm{NRMSE} = \frac{1}{C_z} \sqrt{\frac{1}{N} \sum_{t=1}^{N} (P_{act}(t) - P_{mod}(t))^2} \tag{18}$$

Low NRMSE indicates that the model correctly estimates both the timing and amplitude of power fluctuations relative to the system size.

The Pearson correlation coefficient (r) quantifies the linear relationship between modeled and observed generation and primarily reflects the model's ability to capture temporal variability:

$$r = \frac{\sum_{t=1}^{N} (P_{mod}(t) - \bar{P}_{mod})(P_{act}(t) - \bar{P}_{act})}{\sqrt{\sum_{t=1}^{N} (P_{mod}(t) - \bar{P}_{mod})^2} \sqrt{\sum_{t=1}^{N} (P_{act}(t) - \bar{P}_{act})^2}} \tag{19}$$

where $\bar{P}$ denotes the mean generation over the period. A value close to 1 indicates excellent temporal alignment.

FB measures systematic over- or underestimation, with positive values indicating overestimation of actual generation by the model and negative values indicating underestimation:

$$\mathrm{FB} = \frac{2 \cdot \sum_{t=1}^{N} (P_{mod}(t) - P_{act}(t))}{\sum_{t=1}^{N} (P_{act}(t) + P_{mod}(t))} \tag{20}$$

FB is bounded between -2 and $+2$, with values close to zero indicating negligible systematic bias.

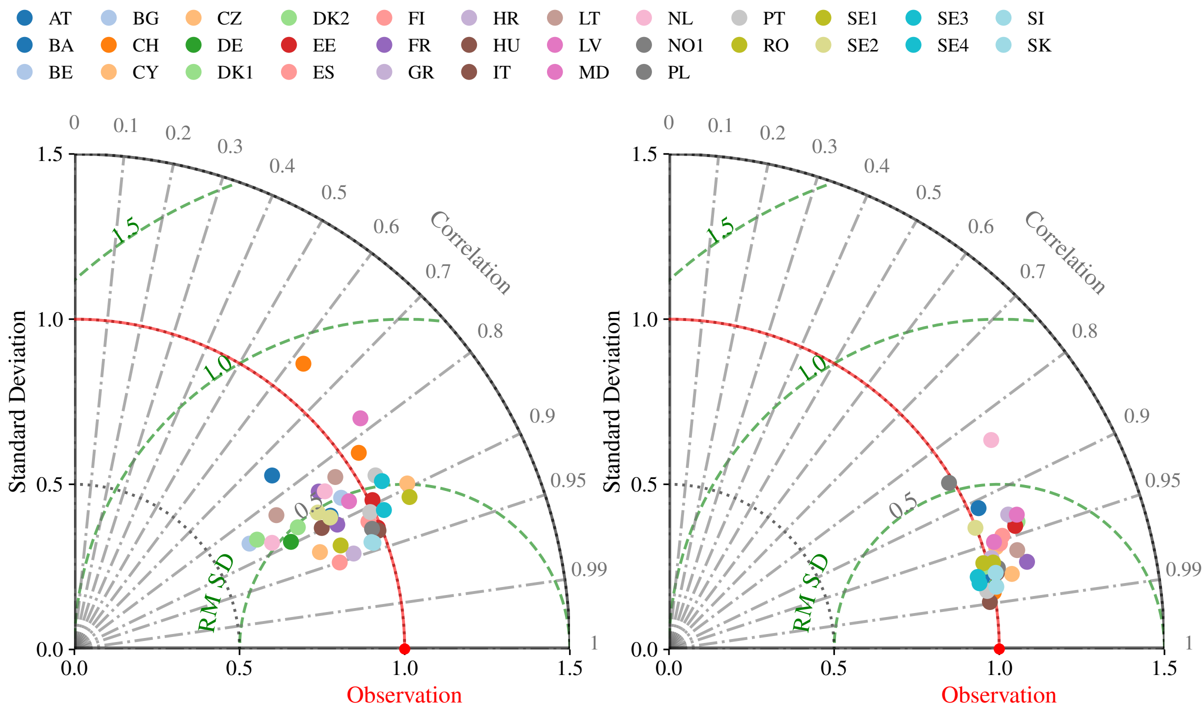
To evaluate the dispersion of the modeled and observed time series around the mean, the standard deviation is also evaluated. It is defined as:

$$\sigma_P = \sqrt{\frac{1}{N} \sum_{t=1}^{N} \left(P(t) - \bar{P}\right)^2} \tag{21}$$

where $\bar{P}$ denotes the mean generation over the period.

At country and bidding-zone scales, geographic aggregation substantially reduces plant-level noise and local metadata errors. Therefore, a high correlation (r) combined with moderate residual RMSE typically indicates correct event timing but imperfect scaling. On the other hand, lower correlation points toward unresolved sub-grid meteorology or metadata mismatches.

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026



a) Wind

b) Solar

Figure 3. Taylor diagrams for hourly (a) wind and (b) solar energy generation in 2025.

4.2 Validation on country-aggregated level

To broadly compare the timing and magnitude across Europe for both solar and wind, Figure 3 visualizes modeled vs actual time series using a Taylor diagram. The standard deviation is represented by the radial distance from the origin (red line), correlation is indicated by the azimuthal angle (gray lines), and the centered root-mean-square difference as the Euclidean distance between the model point and the reference observation point (green line). Areas with no or negligible installed capacity are excluded from the diagram, as the metrics become less meaningful when the observed generation is close to zero.

Table 2 summarizes validation metrics for nationally aggregated wind generation in 2025. NRMSE and NMAE values are normalized by the modeled installed generating capacity (status="operating"). Most regions show strong temporal agreement, with Pearson correlation coefficients ranging from 0.75 to 0.95, indicating that large-scale weather systems are well captured. However, the wind energy model shows greater variability in performance compared to solar generation due to the higher spatial heterogeneity and sensitivity of wind to the local terrain.

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

Table 2. Hourly wind generation validation metrics (uncalibrated) by area, 2025. Countries with insufficient historical ENTSO-E generation data are excluded (United Kingdom, Georgia, Moldova, Slovenia, Slovakia).

Area	Bias (FB)	NMAE	NRMSE	r
Austria (AT)	0.027	0.095	0.131	0.887
Bosnia and Herz. (BA)	1.205	0.252	0.348	0.746
Belgium (BE)	0.381	0.104	0.128	0.864
Bulgaria (BG)	0.269	0.107	0.145	0.866
Switzerland (CH)	−0.211	0.122	0.171	0.783
Cyprus (CY)	−0.010	0.082	0.122	0.615
Czech Republic (CZ)	−0.546	0.206	0.275	0.867
Germany (DE)	0.295	0.100	0.127	0.917
DK1	0.040	0.035	0.047	0.888
DK2	0.619	0.051	0.059	0.852
Estonia (EE)	0.681	0.217	0.262	0.855
Spain (ES)	0.178	0.049	0.071	0.886
Finland (FI)	0.327	0.113	0.146	0.929
France (FR)	0.436	0.119	0.138	0.936
Greece (GR)	0.304	0.084	0.112	0.918
Croatia (HR)	−0.080	0.070	0.096	0.903
Hungary (HU)	0.279	0.103	0.144	0.845
Ireland (IE)	0.480	0.187	0.221	0.933
Italy (IT)	−0.296	0.015	0.021	0.931
Lithuania (LT)	0.209	0.077	0.102	0.893
Luxembourg (LU)	−0.327	0.115	0.162	0.841
Latvia (LV)	0.575	0.008	0.011	0.765
Montenegro (ME)	0.067	0.054	0.077	0.880
North Macedonia (MK)	0.326	0.111	0.175	0.774
Netherlands (NL)	0.845	0.304	0.339	0.887
NO1	−0.232	0.233	0.313	0.843
NO2	0.147	0.082	0.111	0.921
NO3	0.012	0.064	0.088	0.862
NO4	0.221	0.086	0.115	0.905
Poland (PL)	0.269	0.097	0.122	0.928
Portugal (PT)	−0.170	0.042	0.056	0.910
Romania (RO)	0.452	0.077	0.104	0.870
Serbia (RS)	0.313	0.121	0.175	0.893
SE1	0.153	0.056	0.078	0.911
SE2	0.390	0.115	0.157	0.891
SE3	0.182	0.157	0.208	0.934
SE4	0.116	0.042	0.057	0.935

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

Table 3. Hourly PV generation validation metrics (uncalibrated) by area, 2025. Countries with insufficient historical ENTSO-E generation data are excluded (Ireland, Luxembourg, Montenegro, NO3, SE2).

Area	Bias (FB)	NMAE	NRMSE	r
Austria (AT)	0.607	0.070	0.131	0.958
Bosnia and Herz. (BA)	1.353	0.140	0.251	0.907
Belgium (BE)	−1.592	1.149	2.121	0.968
Bulgaria (BG)	−0.155	0.062	0.103	0.930
Switzerland (CH)	0.753	0.096	0.179	0.920
Cyprus (CY)	−0.373	0.121	0.219	0.923
Czech Republic (CZ)	0.430	0.055	0.111	0.974
Germany (DE)	0.372	0.038	0.077	0.978
DK1	0.168	0.037	0.084	0.936
DK2	0.106	0.024	0.048	0.966
Estonia (EE)	−0.058	0.033	0.068	0.940
Spain (ES)	0.299	0.076	0.124	0.938
Finland (FI)	−0.082	0.029	0.056	0.951
France (FR)	0.321	0.052	0.098	0.970
Greece (GR)	0.000	0.026	0.056	0.951
Croatia (HR)	−0.314	0.087	0.157	0.909
Hungary (HU)	0.385	0.052	0.104	0.972
Italy (IT)	−0.633	0.062	0.106	0.986
Lithuania (LT)	0.543	0.053	0.116	0.951
Latvia (LV)	−0.132	0.045	0.089	0.913
Moldova (MD)	0.381	0.078	0.161	0.806
Netherlands (NL)	1.935	0.131	0.246	0.845
NO1	0.313	0.065	0.137	0.825
Poland (PL)	−1.196	0.124	0.213	0.963
Portugal (PT)	0.373	0.059	0.102	0.979
Romania (RO)	0.608	0.076	0.143	0.953
SE1	−0.247	0.053	0.119	0.925
SE3	−0.053	0.027	0.054	0.961
SE4	0.834	0.047	0.091	0.972
Slovenia (SI)	−1.426	0.813	1.616	0.844
Slovakia (SK)	0.699	0.081	0.155	0.980

For certain countries with complex topography or low overall aggregate wind production (e.g., North Macedonia (MK), Bosnia and Herzegovina (BA), or Switzerland (CH)), correlation is lower (typically < 0.8), while RMSE values remain small due to the limited installed capacity. This indicates that even at the 5.5 km resolution, the model still struggles to resolve localized wind channeling and orographic effects, which are critical for accurate generation estimates in mountainous regions.

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

Validation against ENTSO-E records shows consistently high temporal accuracy across Europe. This performance reflects both the physically based PV modeling framework and the high-resolution (5.5 km) CERRA irradiance fields, which capture cloud-driven variability.

Table 3 shows that the simulated hourly PV time series reproduce observed generation well across most countries and bidding zones, with r exceeding 0.95 in most major markets. This strong performance is most evident in countries with large and geographically distributed PV fleets, where spatial aggregation benefits from averaging over heterogeneous irradiance conditions, reducing local errors. Italy (IT), Portugal (PT), Germany (DE), and Greece (GR) illustrate this pattern particularly well, with r values of 0.986, 0.979, 0.978, and 0.951, respectively. As for wind, RMSE and MAE scale with installed capacity and should therefore be interpreted relative to system size. The MAE and RMSE are largely proportional to the cumulative installed PV capacity in the respective zones, indicating that residual errors scale consistently with system size rather than reflecting large structural mismatches.

A small number of areas in Table 3 stand out as pronounced outliers and warrant explicit comment. Belgium (FB $= -1.59$, NMAE $= 1.15$), Slovenia (FB $= -1.43$), and Poland (FB $= -1.20$) all combine strongly negative fractional biases with high correlations ($r = 0.97$, 0.84, and 0.96, respectively). This pattern of correct temporal structure but severely deficient magnitude arises from the generator metadata rather than from the meteorological forcing or the conversion chain: in these countries the geo-located and distributed capacity captured by GEM falls well short of the TSO-registered fleet (for Belgium, 4.9 GW modeled against 14.1 GW registered; cf. Table B1 in Appendix B). Because the error metrics are normalized by the modeled capacity, NMAE and NRMSE can exceed unity where the missing capacity is large, as for Belgium. These cases illustrate why the multiplicative bias correction of Sect. 3.5 is an integral part of the released product rather than an optional refinement.

In Figure 3 (right), PV simulations group together close to the observation reference point, indicating high correlation and relatively low spread in normalized standard deviation compared to wind. This reflects more consistent reproduction of variability across regions. The time series comparisons in Figure 4 support this interpretation, showing that the model reproduces both the diurnal cycle and short-term variability driven by cloud cover. However, the uncalibrated PV series exhibits a systematic tendency to overestimate observed generation, visible as a persistent positive offset in the error panels across all shown regions. These positive biases can likely be attributed to unmodeled system losses such as soiling, shading, and degradation, and the absence of curtailment in the model. The systematic differences between modeled and observed energy volumes motivate the use of the bias-correction factors defined in Sect. 3.5. Details of their derivation, seasonal structure, and inter-annual stability are provided in Appendix C, and the associated country-level capacity discrepancies are summarized in Appendix B.

4.3 Sub-national and farm-level validation

To assess model accuracy below the national and bidding zone scales examined above, we compare CERFRES output against two sub-national observational datasets. Hourly production records for 61 individual Norwegian onshore wind farms from the NVE (NVE, 2026) windpower database are matched to corresponding CERFRES per-farm simulations by coordinates. This database also provides wind farm-level information such as number of operating turbines, and turbine-level information such as

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

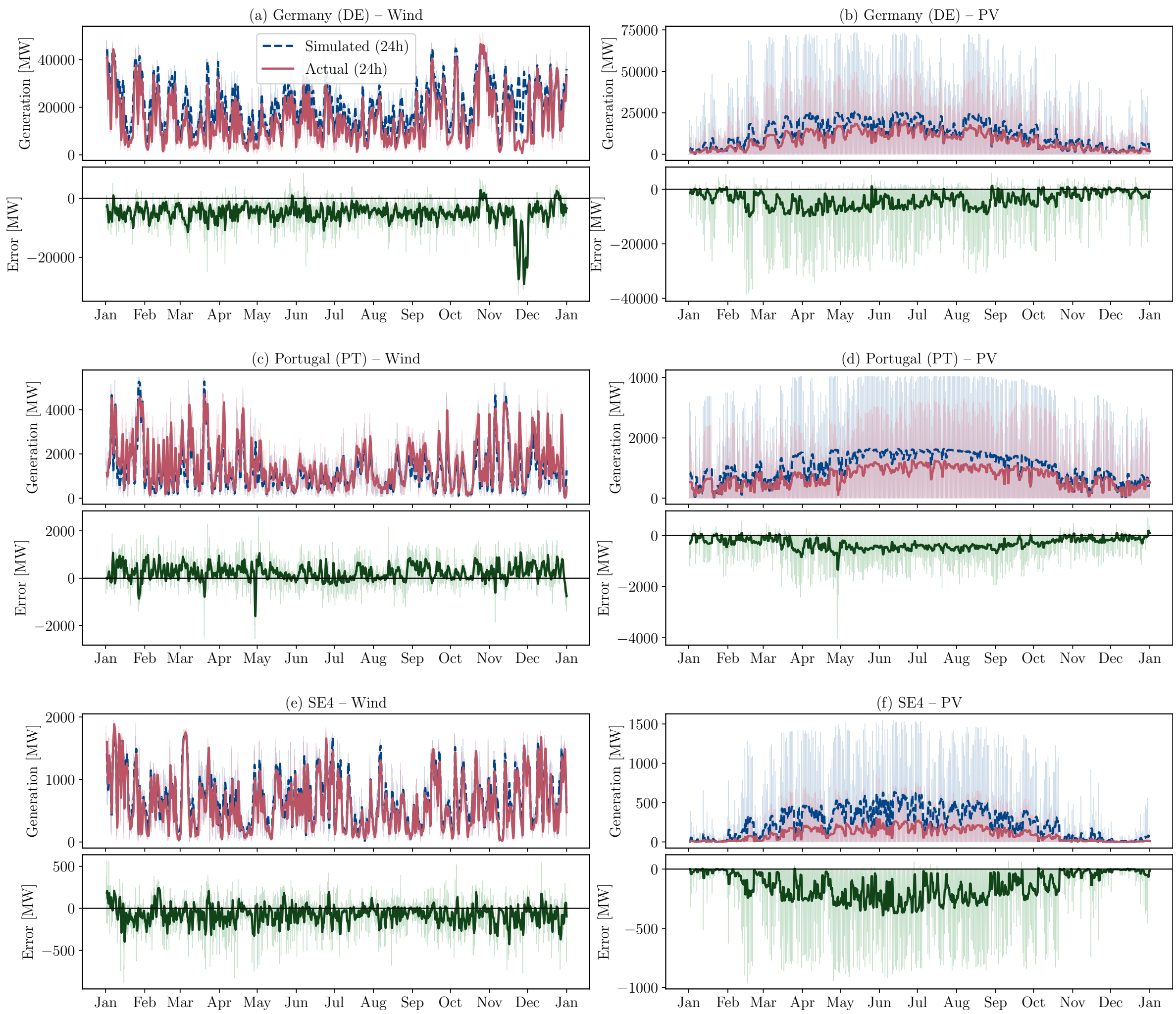


Figure 4. Comparison of simulated and observed renewable generation time series for wind (left column) and PV (right column) aggregated across the indicated countries/bidding zones, 2025.

hub height, rotor diameter and rated power. In addition, hourly solar PV generation for two representative Danish administrative regions (Midtjylland and Sjælland) from the Energi Data Service "Production per Region per Hour" dataset (Energinet, 2026) is compared against CERFRES output aggregated to the corresponding Nomenclature of Territorial Units for Statistics level 2 (NUTS-2) regions.

Figure 5 presents hourly generation for January 2023 at four Norwegian onshore wind farms (Storheia, Storøy, Stigafjellet, and Lista), selected to represent various capacities and wind conditions across Norway. At all four sites, the simulated series closely reproduce the observed large-scale atmospheric variability, including calm intervals during high-pressure blocking

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

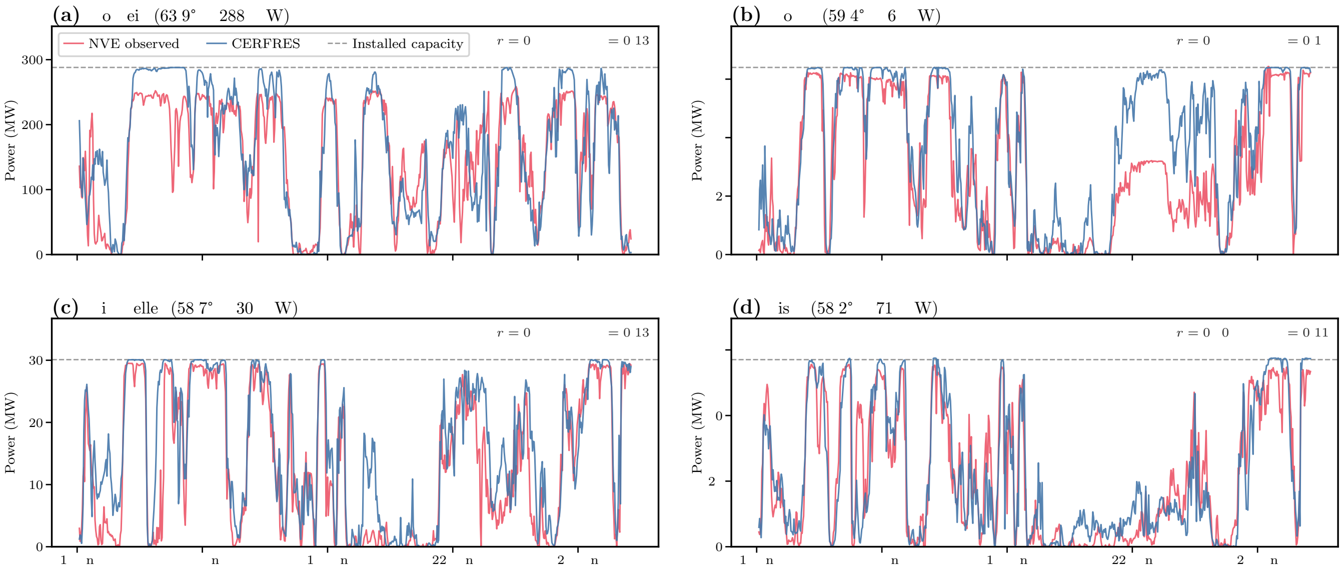


Figure 5. Farm-level wind generation comparison between CERFRES (blue) and NVE observed (rose) for four Norwegian onshore wind farms during January 2023.

events and sharp ramps during the passage of extratropical cyclones. Across the full NVE inventory of 61 farms, Pearson correlations for January 2023 lie between 0.78 and 0.93, and the NMAE remains below 0.18 at each of the four farms shown.

In terms of discrepancies, two issues specifically stand out, since both originate in the representation of installed capacity within the modeling framework rather than in the meteorological forcing. First, during sustained high-wind periods, CERFRES reaches the registered capacity, whereas the observed output of the larger wind farms typically reached a maximum of approximately 10–15% below this value (Figure 5a, first week of January). The power conversion applies an uncorrected reference power curve (Table 1) to every registered megawatt. As a result, intra-array wake losses, transformer losses, and routine turbine unavailability, which in combination typically limit attainable farm output to 85–95% of installed capacity, are excluded from the uncalibrated series. This is intentional, as the model aims to represent the idealized performance of each turbine under standard conditions, which is useful for energy system planners. Second, the divergence at Storøy during 22–29 January (Figure 5b), where CERFRES sustains output near the 6.4 MW installed capacity while the observed generation plateaus at almost exactly 3.2 MW. Storøy wind farm consists of only two Enercon E-115 turbines rated at 3.2 MW each, so the observed plateau corresponds exactly to one turbine operating at rated power while the other was out of service. Because CERFRES treats the capacity recorded in the generator metadata as fully available at every time step and carries no representation of unit-level outages, such an event lies outside what the model can reproduce.

Figure 6 aggregates the farm-level comparison to monthly capacity factors for six farms spanning nearly the full latitudinal range of the Norwegian onshore fleet. Every site shows the characteristic winter maximum and summer minimum of the northern European wind climate. The model also reproduces the geographic differences in the amplitude of this cycle, ranging

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

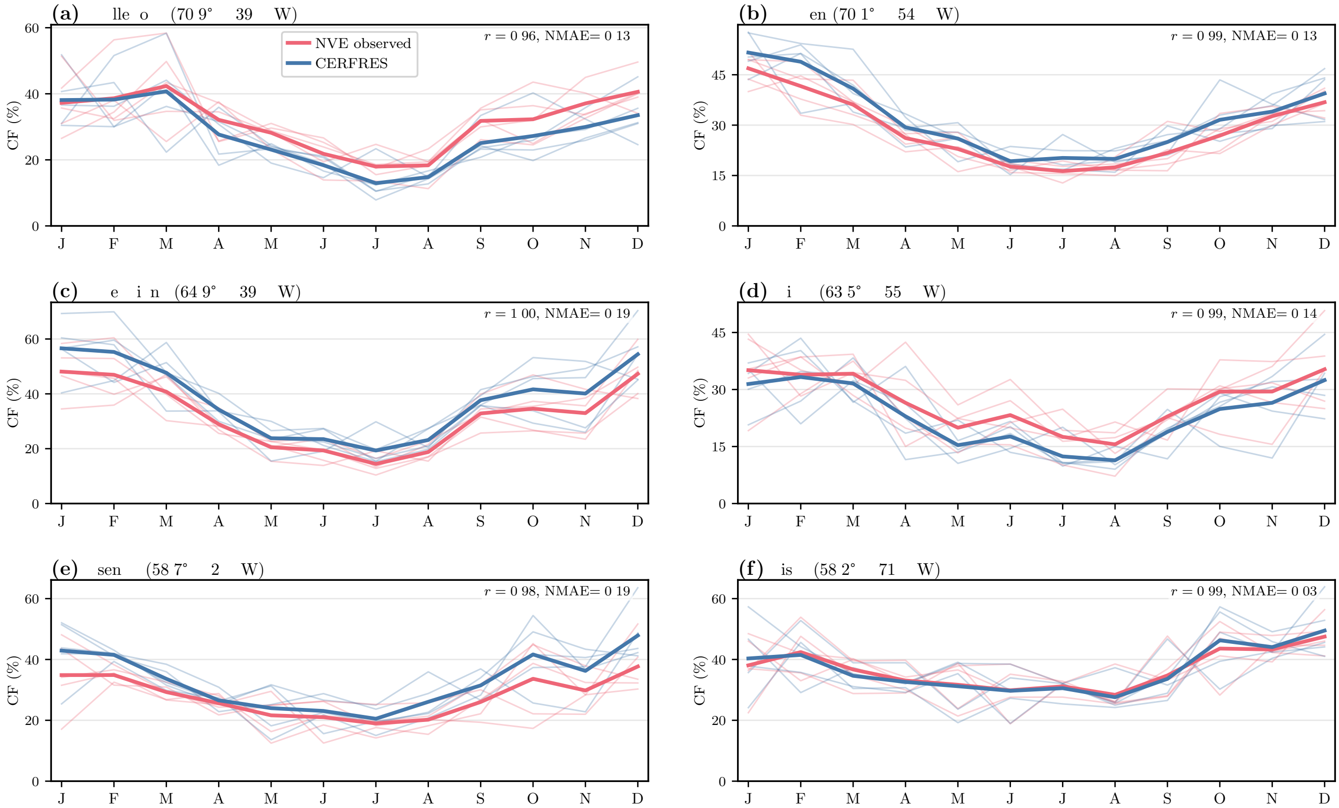


Figure 6. Monthly wind capacity factor (CF) for six Norwegian onshore wind farms arranged N–S (71°N–58°N) using the CERFRES simulations and NVE observations. Thin lines show individual years (2015, 2017, 2019, 2021, 2023); bold lines show the multi-year mean.

from northern sites where capacity factors decrease from nearly 50% in winter to below 20% in July, to the much flatter seasonal pattern observed at Lista on the southwestern coast. Figure 6 further indicates that the model tracks the inter-annual variation, which reaches 10–15 percentage points about the climatological mean in winter, on a year-by-year basis rather than merely in the multi-year average.

Figure 7 compares CERFRES with observed generation in Midtjylland and Sjælland during 1–14 June 2023, together with the corresponding CERRA surface irradiance averaged over each NUTS-2 region. CERFRES reproduces the diurnal solar cycle with high fidelity in both regions (Pearson $r = 0.964$ for Midtjylland and $r = 0.981$ for Sjælland), and simulated power closely follows the CERRA irradiance forcing throughout the period. The remaining discrepancies therefore originate primarily from the meteorological forcing rather than the photovoltaic conversion routine. While the observed generation exhibits short-lived reductions caused by localized cloud cover, these events are not resolved by the smoother CERRA irradiance field and are consequently absent from the simulations. This indicates that the temporal and spatial resolution of the radiation forcing, rather than the power-conversion methodology, is the main limitation in capturing short-term solar variability.

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

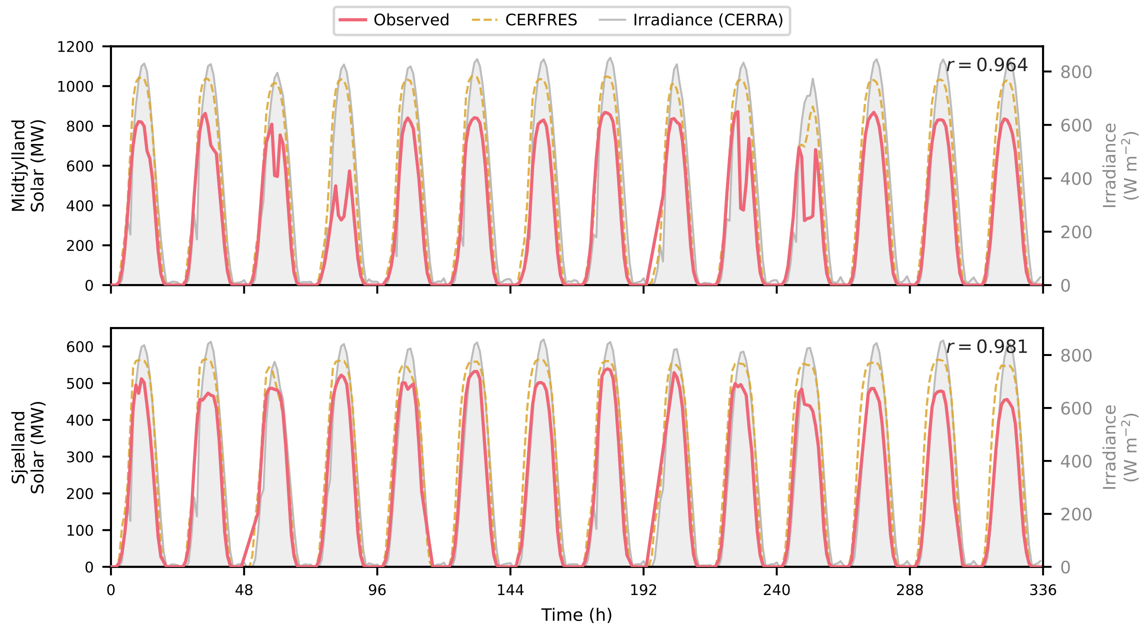


Figure 7. Sub-national solar PV validation for two Danish regions, Midtjylland (top) and Sjælland (bottom), over a two-week window in June 2023 (1–14 June).

4.4 Benchmark evaluation against open European datasets

For comparison, CERFRES is benchmarked against two widely used datasets: EMHIRES (1986–2015) (Gonzalez Aparicio et al., 2016, 2017) and Renewables Ninja (2000–2018) (Staffell and Pfenninger, 2016; Pfenninger and Staffell, 2016; Schmitz and Després, 2022). All three datasets are evaluated against ENTSO-E hourly reported generation as a common reference year within the overlapping coverage. The comparison is performed at national level, since EMHIRES and Ninja provide country-aggregated time series, while CERFRES aggregates farm-level simulations. This enables a direct comparison of final generation profiles despite differences in underlying modeling approaches.

The comparison highlights two primary sources of performance differences. The first is the choice of meteorological forcing data. EMHIRES uses MERRA-1 for wind and CM-SAF SARAH for solar PV, whereas Renewables Ninja uses MERRA-2 for wind and CM-SAF SARAH for solar PV. Second, they differ in spatial modeling. EMHIRES and Ninja apply power conversion to aggregated meteorological fields, whereas CERFRES models individual farms before aggregation. In homogeneous systems, differences mainly reflect meteorological input quality, while in heterogeneous or complex terrain they also capture the impact of spatial resolution.

A third source of differences lies in generator metadata quality and completeness. Both EMHIRES and Renewables Ninja rely on the thewindpower.net database (Pierrot, 2015), where hub heights were unknown for 62% of farms in Ninja (Staffell and Pfenninger, 2016) and had to be estimated for 61% in EMHIRES (Gonzalez Aparicio et al., 2016); turbine models were simi-

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

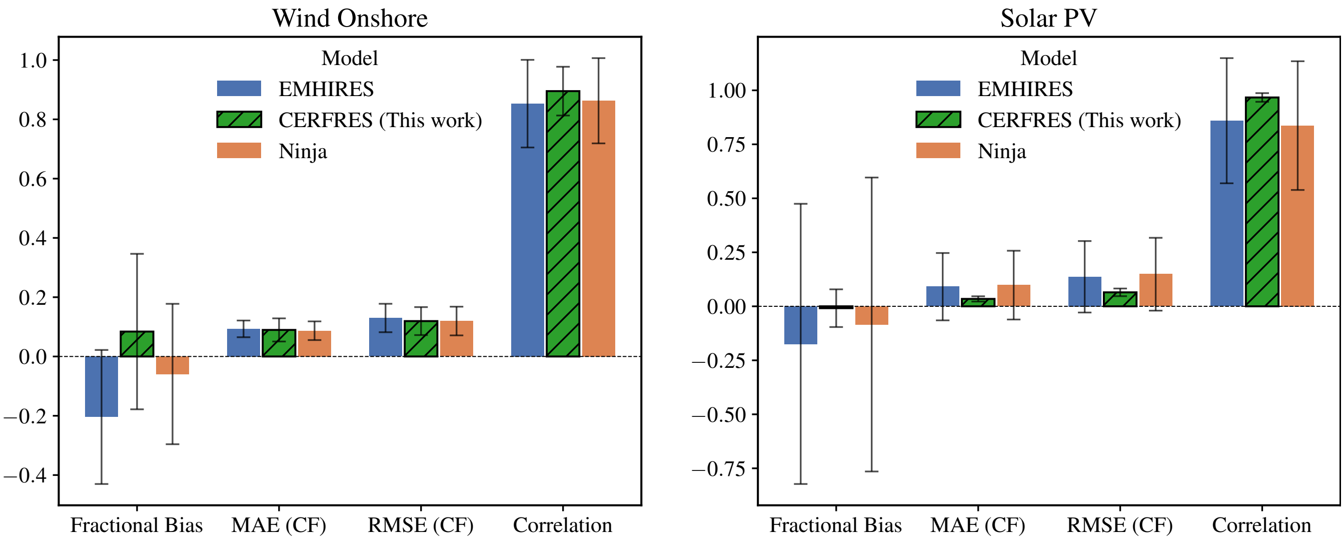


Figure 8. Mean validation metrics across all common countries for wind onshore (left) and solar PV (right) for the year 2015. Error bars show the standard deviation.

larly inferred for a substantial fraction of installations. For solar PV, neither EMHIRES nor Renewables Ninja uses geo-located farm databases. EMHIRES assumes a uniform 30° south-facing tilt for all European installations and distributed capacity across NUTS-2 regions using solar potential weights rather than actual farm locations (Gonzalez Aparicio et al., 2017). Renewables Ninja simulates PV output using site-level panel metadata where available. CERFRES uses the Global Solar and Wind Power Trackers (Global Energy Monitor, 2026a, b), which provide geo-located farm-level metadata and allow explicit spatial modeling, but introduce their own uncertainties through capacity recording practices, commissioning year errors, and incomplete coverage of distributed installations, as discussed in Sect. 2.1. These metadata differences contribute to systematic biases that are difficult to disentangle from meteorological effects in the benchmark comparison.

Tables 4 and 5 compare the three datasets country by country for wind onshore and solar PV, respectively, across FB, MAE, RMSE, and r. Figure 8 complements these tables by summarizing the country-mean values for each dataset.

The benchmark reveals a consistent overall pattern: CERFRES tends to improve amplitude-related performance most clearly where higher-resolution meteorological forcing matters, while the benefits of farm-level spatial modeling are more country dependent. For wind, the strongest improvements over EMHIRES and Renewables Ninja appear in countries with heterogeneous wind regimes or complex terrain, where coarse meteorological inputs and national averaging smooth out local variability. This is particularly evident in Italy, Bulgaria, and Finland, where CERFRES significantly reduces bias and error relative to EMHIRES. The findings confirm that higher-resolution CERRA data better captures terrain effects and other sub-grid variability that are smoothed in MERRA-2-based products.

For solar PV, differences are driven primarily by systematic bias rather than temporal structure. EMHIRES shows a consistent negative FB across most countries, indicating a tendency to underestimate annual PV output. CERFRES reduces this systematic

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

Table 4. Wind onshore benchmark metrics (capacity-factor units) against ENTSO-E reported generation for 2015. E = EMHIRES, N = Renewables Ninja, and C = CERFRES. Countries for which a dataset provides no coverage are marked "–". Bold values indicate the best-performing dataset for a given metric and country.

	FB			NMAE			NRMSE			Correlation		
Country	E	N	C	E	N	C	E	N	C	E	N	C
AT	**0.068**	–	0.199	0.100	–	**0.089**	0.149	–	**0.123**	0.859	–	**0.907**
BE	**0.292**	0.445	0.788	**0.074**	0.101	0.227	**0.107**	0.127	0.276	**0.962**	0.952	0.822
BG	-0.466	-0.176	**0.117**	0.140	0.099	**0.080**	0.193	0.140	**0.112**	0.749	0.851	**0.913**
CH	**-0.059**	-0.171	-0.195	0.136	0.107	**0.106**	0.201	0.157	**0.145**	0.554	0.605	**0.716**
CY	-0.708	**-0.270**	-0.398	0.125	0.103	**0.095**	0.180	0.143	**0.141**	0.424	0.611	**0.657**
CZ	**-0.009**	-0.060	0.088	0.083	**0.057**	0.078	0.113	**0.078**	0.106	**0.926**	0.925	0.896
DE	-0.144	**-0.137**	0.268	0.048	**0.046**	0.087	**0.058**	0.066	0.108	**0.982**	0.976	0.946
EE	-0.167	**-0.120**	0.241	0.072	**0.060**	0.095	0.093	**0.084**	0.122	0.940	**0.948**	**0.948**
ES	-0.192	-0.140	**-0.112**	0.084	0.060	**0.045**	0.102	0.077	**0.058**	0.924	0.948	**0.966**
FI	-0.266	0.348	**-0.102**	0.102	0.148	**0.055**	0.138	0.180	**0.072**	0.816	0.796	**0.947**
FR	-0.225	**-0.108**	0.171	0.066	**0.048**	0.062	0.077	**0.066**	0.079	**0.972**	0.968	0.968
HU	-0.351	**-0.034**	0.040	0.099	0.089	**0.081**	0.143	0.121	**0.117**	0.875	**0.898**	0.887
IE	-0.148	**-0.075**	0.181	0.066	**0.062**	0.089	0.087	**0.084**	0.117	**0.966**	0.955	0.951
IT	-0.581	-0.374	**-0.048**	0.119	0.087	**0.042**	0.148	0.119	**0.058**	0.878	0.918	**0.956**
LT	-0.093	**0.055**	0.149	0.066	0.070	**0.065**	**0.092**	**0.092**	**0.092**	0.926	0.926	**0.938**
LU	-0.312	0.379	**0.119**	0.066	0.097	**0.065**	**0.093**	0.123	0.094	**0.920**	0.910	0.903
LV	-0.452	**0.091**	0.217	0.112	**0.083**	0.092	0.159	**0.106**	0.128	0.921	**0.922**	0.918
NL	**0.086**	0.109	0.354	0.076	**0.075**	0.141	**0.104**	**0.104**	0.185	**0.900**	0.892	0.784
PL	**-0.116**	-0.133	0.122	0.057	0.060	**0.055**	0.075	0.089	**0.071**	0.959	0.953	**0.966**
PT	-0.324	**-0.151**	-0.434	0.104	**0.100**	0.114	0.140	**0.134**	0.147	0.884	0.871	**0.923**
RO	-0.359	-0.179	**-0.167**	0.118	0.100	**0.096**	0.163	0.139	**0.136**	0.837	**0.876**	0.866
SI	**-0.043**	-0.537	–	**0.147**	0.173	–	**0.258**	0.282	–	**0.553**	0.419	–
UK	**-0.126**	–	0.251	**0.085**	–	0.111	**0.112**	–	0.138	0.892	–	**0.917**

bias and yields results that are more symmetrically distributed around zero, while preserving similarly high, and in several cases improved r values. Averaged over the countries covered jointly in Table 5, CERFRES reduces the country-mean NRMSE by 31% relative to EMHIRES and 37% relative to Renewables Ninja, with corresponding NMAE reductions of 34% and 39%. This suggests that the main improvements for PV come from the irradiance forcing and meteorological representation rather than from large changes in temporal structure, which is already reproduced well by all three datasets.

Since Renewables Ninja and CERFRES rely on similar reanalysis inputs, differences more directly demonstrate how meteorology is translated into national generation. In major wind markets such as Spain, Italy, and Finland, CERFRES achieves lower MAE and RMSE than Ninja together with higher correlation, suggesting that farm-level aggregation reduces smoothing and improves representation of variability. The improvement is less uniform in systems dominated by concentrated offshore

Table 5. Solar PV benchmark metrics (capacity-factor units) vs. ENTSO-E actuals, year 2015.

	FB			NMAE			NRMSE			Correlation		
Country	E	N	C	E	N	C	E	N	C	E	N	C
AT	**-0.332**	–	–	**0.068**	–	–	**0.121**	–	–	**0.930**	–	–
BE	-0.227	-0.149	**-0.039**	0.039	0.041	**0.036**	0.075	0.082	**0.073**	**0.973**	0.951	0.953
BG	-0.209	-0.164	**0.085**	0.044	0.047	**0.036**	0.082	0.090	**0.075**	**0.973**	0.964	0.969
CH	0.167	0.314	**0.132**	0.047	0.062	**0.035**	0.093	0.124	**0.080**	0.894	0.850	**0.932**
CZ	-0.204	-0.081	**0.051**	0.034	0.038	**0.032**	**0.063**	0.075	0.067	**0.984**	0.955	0.965
DE	-0.412	-0.181	**-0.045**	0.055	0.040	**0.028**	0.102	0.079	**0.051**	**0.987**	0.964	0.977
EE	-1.596	**-1.513**	–	0.709	**0.689**	–	0.783	**0.765**	–	0.139	**0.172**	–
ES	-0.380	-0.408	**-0.190**	0.096	0.098	**0.067**	0.140	0.144	**0.098**	0.939	0.940	**0.956**
FR	-0.388	-0.324	**-0.057**	0.067	0.056	**0.024**	0.117	0.099	**0.044**	0.980	**0.988**	**0.988**
IT	-0.321	-0.234	**-0.003**	0.060	0.047	**0.021**	0.105	0.089	**0.039**	0.978	0.986	**0.991**
LT	-0.351	**-0.151**	–	0.057	**0.047**	–	0.105	**0.086**	–	0.913	**0.928**	–
LU	**-0.072**	0.079	–	**0.029**	0.047	–	**0.060**	0.096	–	**0.964**	0.899	–
NL	**1.992**	1.993	–	**0.115**	0.125	–	**0.214**	0.229	–	−0.000	−0.005	–
PT	-0.100	-0.328	**0.037**	0.037	0.072	**0.031**	0.066	0.124	**0.060**	0.983	0.975	**0.984**
RO	-0.248	-0.259	**-0.074**	0.048	0.055	**0.028**	0.090	0.104	**0.051**	**0.988**	0.972	0.984
SI	**0.066**	0.231	–	**0.028**	0.038	–	**0.057**	0.077	–	**0.959**	0.943	–
SK	-0.225	**-0.172**	–	**0.056**	0.068	–	**0.098**	0.116	–	**0.956**	0.906	–
UK	-0.296	–	**0.009**	0.047	–	**0.040**	0.091	–	**0.074**	0.929	–	**0.939**

E = EMHIRES, N = Renewables Ninja, C = CERFRES. Bold values indicate the best-performing dataset for that metric and country.

or near-shore production. Belgium and the Netherlands are notable cases where CERFRES produces a larger positive FB and higher RMSE than the benchmark datasets. These results may reflect sensitivity to power-curve assumptions and incomplete representation of offshore or distributed capacity in the underlying dataset.

The results should also be interpreted in light of the benchmark year. The comparison is restricted to 2015 because it lies within the overlapping historical coverage of all three datasets. It therefore provides the fairest common-year assessment, even though CERFRES bias-correction factors are derived from a broader 2015–2025 calibration period. For that reason, this section should be read as a benchmark of the raw modeling frameworks rather than of the final bias-corrected CERFRES product. Overall, the comparison indicates that both improved meteorological forcing and farm-level spatial modeling contribute to better performance, but in different ways. The gain from higher-resolution meteorology is the more systematic one, especially for reducing amplitude bias and aggregate error.

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

5 Conclusions

This paper presented CERFRES v1.0, a physics-based modeling framework for generating high-resolution, farm-level wind and solar power production time series across Europe using the CERRA reanalysis. The framework introduces a modular conversion pipeline combining inverse-distance-weighted meteorological interpolation to plant locations, a Gaussian-smoothed multi-turbine power aggregation scheme, a full photovoltaic modeling chain, and monthly bias-correction against observed bidding-zone generation. This approach addresses the absence of openly available historical records at individual plant level by providing a reproducible, validated alternative at farm and sub-national scales.

Validation against ENTSO-E reported generation shows that the model captures the temporal behavior of solar PV particularly well and represents wind variability robustly across most countries and bidding zones. PV results show consistently high agreement, with correlation values above 0.90 in most major markets, supporting the conclusion that the combination of high-resolution irradiance data and geolocated generator metadata accurately reproduces both the diurnal cycle and short-term weather-driven fluctuations. Wind performance is more heterogeneous, but remains strong in regions with large installed capacity and less complex terrain.

For the comparison between actual and simulated wind farm-level time series, the conclusion that model accuracy is particularly sensitive to the quality of generator metadata is supported by farms with turbines undergoing maintenance or curtailment, which show larger discrepancies than those with more stable operation. In the Norwegian January 2023 case study, the results also indicate that the CERRA meteorological forcing sufficiently resolves sub-national wind variability, as shown by high Pearson correlation coefficients at individual farm level. For solar PV, the Danish two-week regional case study suggests that the model captures the diurnal cycle and cloud-driven variability well, with remaining discrepancies primarily attributable to localized cloud cover rather than errors in the photovoltaic conversion routine.

Benchmarking against EMHIRES and Renewables Ninja indicates that replacing the coarse MERRA-2 meteorological forcing with the higher-resolution CERRA reanalysis is the primary driver of improvement. For solar PV, CERFRES achieves country-mean NRMSE reductions of 31% and 37% relative to EMHIRES and Renewables Ninja, respectively, with corresponding NMAE reductions of 34% and 39%, together with a significant reduction in the systematic negative bias that is typical of the EMHIRES product. For wind, improvements are most pronounced in countries with heterogeneous meteorological conditions or complex terrain, where higher-resolution forcing better captures sub-grid variability. Farm-level spatial modeling provides additional benefits, especially in systems with geographically dispersed generation, where national aggregation would otherwise smooth important spatial differences.

Code and data availability. All results presented in this paper were produced using version 1.1.0 of the CERFRES output dataset, which is publicly archived on Zenodo (https://doi.org/10.5281/zenodo.19762446) (Bruvik and Sorteberg, 2026a) under the CC BY 4.0 license (https://creativecommons.org/licenses/by/4.0/). The label "CERFRES v1.0" in the paper title refers to the version of the modeling framework (software), which is separately archived on Zenodo (https://doi.org/10.5281/zenodo.20772916) (Bruvik, 2026) and corresponds to the GitHub release `v1.0` in the public CERFRES repository (https://github.com/EmilBruvik/CERFRES/releases/tag/v1.0). The source code is

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

distributed under the MIT license (https://opensource.org/license/mit). The framework is implemented in Python (version 3.10 or later) and depends only on openly available libraries. Quality flags and licensing constraints derived from upstream sources are documented in Table D1. A subset of the Global Solar Power Tracker records used as model input carry a CC BY-NC 4.0 license inherited from the TransitionZero Solar Asset Mapper (Table D1). These records are used only to inform the location, capacity, and technology of individual PV installations within the conversion model. The released CERFRES output does not reproduce this source metadata but instead reports simulated power aggregated either to the independently licensed CERRA reanalysis grid or to national and bidding-zone areas, without exposing per-installation capacity, coordinates, or identifiers. The released dataset therefore does not redistribute NC-licensed material and remains compatible with its CC BY 4.0 license. Because several validation inputs originate from continuously updated live platforms with no fixed historical snapshot, the exact data extracts used to produce the validation results (Figs. 5, 6, 7; Tables 2, 3) are archived as version 1.1.1 of the same Zenodo record (https://doi.org/10.5281/zenodo.21479427) (Bruvik and Sorteberg, 2026b), which adds these extracts to the unchanged output data of version 1.1.0. These include the February 2026 Global Energy Monitor wind tracker snapshot and the accompanying aggregated distributed-PV capacity file from the Global Solar Power Tracker (country-level aggregates only, with no individual-installation records; `CC BY 4.0`), the Energi Data Service regional hourly solar and wind production records for the Danish regions (2021–2025, `CC BY 4.0`), the NVE hourly wind production records (2002–2025) for named Norwegian onshore wind farms, and the monthly ENTSO-E Aggregated Generation per Type exports (2014–2025) used as the observed reference for validation and bias correction. The ENTSO-E Transparency Platform and NVE do not assign a standard open license to bulk data extracts. The archived files reproduce only the source extracts used to compute the reported metrics, consistent with ENTSO-E's public disclosure obligation under EU Regulation 543/2013 and with NVE's stated terms of reuse. The EMHIRES and Renewables Ninja benchmark datasets used in Table 4 and Table 5 are not re-archived, as both are static, independently versioned datasets available from their original sources (Gonzalez Aparicio et al., 2016, 2017; Staffell and Pfenninger, 2016; Pfenninger and Staffell, 2016).

Author contributions. Emil Bruvik: modeling framework, data processing, validation, and manuscript preparation. Asgeir Sorteberg: conceptualization, supervision, and writing – review and editing.

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

520 ## Appendix A: Overview of related open meteorological and generation datasets

Table A1. Comparison of different open-access meteorological databases and their corresponding renewable generation datasets (solar and wind).

Name	Database	Temp. Res.	Spatial Res.	Years	Region	License
Renewables Ninja	NASA MERRA-2, CM-SAF SARAH	1 h	∼50 km / 5 km	2000–2018	Global	CC BY-NC 4.0
EMHIRES	MERRA-1	1 h	70×60 km	1986–2015	EU-28	EU reuse
Open power system data	ENTSO-E	15 min	Country/Zone	2015–2020	Europe	CC BY 4.0
ERA5 Reanalysis	ERA5	1 h	31 km	1950–2020	Europe	CC BY 4.0
MERRA2 Derived	MERRA-2	1 h	∼50 km	1980–2018	Europe	CC BY 4.0
C3S	ECMWF ERA5	1 h	31 km	1979–2022	Europe	Copernicus 1.0
NREL NSRDB	NREL PSM, NOAA GOES, NIC IMS, NASA MODIS, MERRA-2	5 min–1 h	Various	1998–2021	Global	CC BY 3.0
PECD (2021 update)	ERA5, GWA version 2	1 h	31 km	1982–2019	Europe	CC BY 4.0
CERFRES	**Copernicus CERRA**	**1 h**	**5.5 km**	**1995–2025**	**Europe**	**CC BY 4.0**

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

Appendix B: Aggregate generation capacity and capacity factors

Table B1 compares TSO-registered installed capacity (wind + solar) and the corresponding annual capacity factors with their modeled counterparts for 2025. The modeled capacity is derived from the GEM fleet after status and commissioning-year filtering, and the capacity factors normalize annual generation by the respective capacity figure. The comparison helps distinguish errors arising from fleet representation from those related to the meteorological-to-power conversion, and underlies the correction-factor analysis in Appendix C.

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

Table B1. Installed generation capacity (wind + solar) and capacity factors (CF) by country/zone, 2025.

	Installed generation capacity (MW)		Capacity factor (–)	
Area	TSO	Modeled	TSO	Modeled
Austria (AT)	11 315.0	11 355.4	0.143	0.195
Belgium (BE)	14 104.1	4871.1	0.182	0.453
Bulgaria (BG)	5279.0	4137.0	0.165	0.196
Croatia (HR)	1501.0	1366.6	0.275	0.270
Cyprus (CY)	692.9	567.9	0.161	0.144
Czech Republic (CZ)	3890.0	5853.8	0.156	0.149
DK1	8451.0	7020.3	0.239	0.306
DK2	2785.0	4206.4	0.238	0.274
Estonia (EE)	1259.0	1733.6	0.211	0.236
Finland (FI)	6645.0	9901.8	0.393	0.360
France (FR)	40 499.0	55 160.0	0.214	0.234
Germany (DE)	144 897.2	158 283.7	0.160	0.203
Greece (GR)	11 765.0	15 700.5	0.213	0.187
Hungary (HU)	4354.2	7892.6	0.192	0.155
Italy (IT)	17 608.0	42 449.9	0.327	0.082
Latvia (LV)	439.0	685.3	0.228	0.158
Lithuania (LT)	2453.0	4110.6	0.255	0.211
Moldova (MD)	192.0	165.7	0.154	0.162
Montenegro (ME)	118.0	134.9	0.288	0.293
Netherlands (NL)	39 674.0	38 559.1	0.063	0.244
NO1	414.8	219.4	0.324	0.488
NO2	1447.4	1480.2	0.299	0.339
NO3	2120.0	2811.6	0.293	0.224
NO4	1160.0	1328.0	0.306	0.333
Poland (PL)	24 191.6	22 232.1	0.197	0.176
Portugal (PT)	7168.6	11 396.5	0.312	0.203
Romania (RO)	4251.0	9097.0	0.231	0.181
SE1	2728.7	7971.2	0.290	0.116
SE2	5968.0	12 115.4	0.310	0.226
SE3	3657.7	6021.2	0.368	0.262
SE4	2196.7	7462.1	0.317	0.122
Serbia (RS)	674.8	803.3	0.228	0.305
Spain (ES)	54 026.6	72 371.1	0.205	0.195
Switzerland (CH)	4034.0	7539.7	0.151	0.176

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

Appendix C: Bias-correction factors

The bias-correction factors are intended for use when a national or zonal fleet is represented as a single node rather than for rescaling individual farm-level series. Historical ENTSO-E generation data are available from 2015 onward for per-type generation, allowing direct estimation of monthly correction factors for 2015–2025. For earlier years (1995–2014), calibration is not possible due to missing observations. Instead, fallback correction factors are constructed for each country and technology (PV and wind) using recency-weighted averages of post-2015 monthly factors. Each calendar month is treated separately, such that, for example, January values are corrected using historical January factors. This preserves the seasonal structure of the correction. This approach assumes that systematic biases remain stationary over time, which may not hold during periods of rapid capacity expansion. As a result, pre-2015 corrections carry higher uncertainty.

Figure C1 shows the evolution of bias-correction factors for France, Spain, and Germany. Values above 1 indicate underestimation by the raw model, while values below 1 indicate overestimation. All three countries exhibit clear seasonal patterns, with PV factors typically increasing in summer and decreasing in winter, while wind shows the opposite tendency. This highlights the importance of monthly calibration for correcting recurrent intra-annual biases.

PV correction factors show comparatively low stochastic variability, with most monthly values remaining close to the 3-month rolling mean. The spread also decreases slightly over time, suggesting improved consistency between generator metadata and modeled PV output. These patterns are consistent with unresolved influences such as curtailment, outages, and residual metadata uncertainty.

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

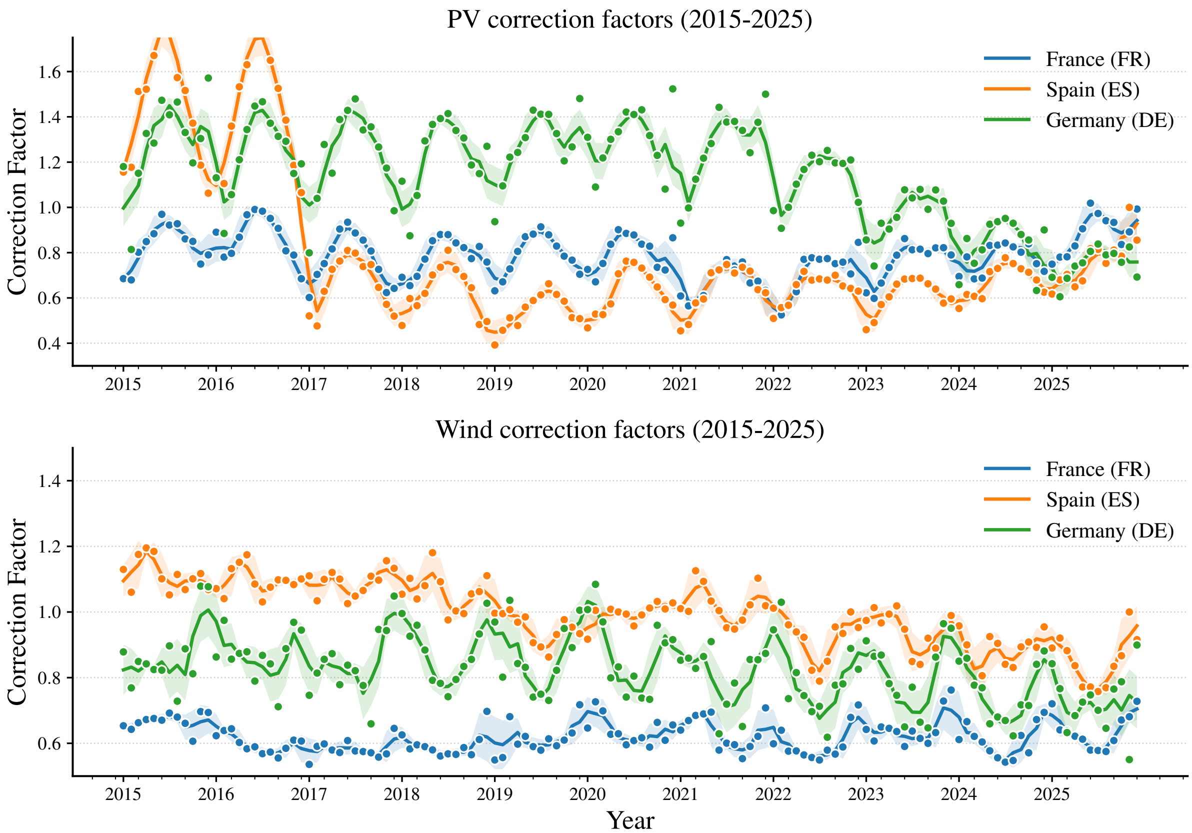


Figure C1. Evolution of PV and wind country-aggregated bias-correction factors for the indicated countries. Solid lines represent a 3-month rolling mean and points represent monthly values.

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

Appendix D: Dataset uncertainty and recommended use

D1 Uncertainty and product scope

The released CERFRES time series are weather-conditioned estimates derived from representative physical conversion models and incomplete asset metadata, and should not be interpreted as direct reconstructions of metered plant output. Three main uncertainty sources apply at all spatial scales.

First, the quality of the final product depends on the completeness of the underlying generator inventories, which consists of: commissioning year, installation type, turbine specifications, PV panel configuration, and distributed generation coverage. These gaps are largest in early years and in countries where smaller or distributed assets are incompletely represented in the source registries.

Second, unresolved physical effects, including curtailment, wake losses, outages, maintenance, and market-driven dispatch, lie outside the physical core model and are only partially absorbed by the monthly bias-correction factors.

Third, correction-factor reliability varies by area and period. For 2015–2025, factors are estimated directly from ENTSO-E actuals. For pre-2015 years, where historical actual generation is unavailable, factors are constructed from recency-weighted averages of post-2015 monthly values for the same area and calendar month. Where no historical factors exist, the factor defaults to 1.0. Users should treat pre-2015 time series with correspondingly higher uncertainty, particularly during periods of rapid capacity expansion when the stationarity assumption underlying the fallback strategy is least likely to hold.

These limitations are strongest at the individual asset level and decrease significantly after aggregation to bidding zone and national scales, where the dataset has been validated and is most reliable.

D2 Recommended and non-recommended use

CERFRES output has been validated at national, bidding-zone, and individual farm level, making it suitable for applications requiring physically consistent, weather-driven renewable generation profiles across these scales. Examples include power-system balancing analyses, hydro–wind coordination studies, transmission-constrained resource assessments, and scenario comparisons between historically evolving fleets and a fixed reference fleet.

The dataset is not recommended for asset-level commercial settlement, forensic reconstruction of specific outage or curtailment events, or plant-level claims of exact metered production. While farm-level time series are provided, systematic validation at this spatial scale has been performed only for Norwegian onshore wind farms; accuracy for individual assets in other countries or technologies has not been established due to lack of validation data and should not be assumed.

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

Table D1. Provenance and licensing of input data sources.

Source		Fields used	Role & license
Name	Provider		
CERRA single-level	Copernicus CDS	Surface solar radiation, 10-m wind speed, 2-m temperature, albedo	PV generation modeling. `CC BY 4.0`
CERRA multi-level	Copernicus CDS	Wind speeds at 30, 50, 75, 100, 150, 200 m	Wind generation modeling. `CC BY 4.0`
ENTSO-E Transparency Platform	ENTSO-E	Actual generation by production type, area code, timestamps	Validation and bias-correction. Freely accessible; no formal open license assigned. See https://transparency.entsoe.eu for current terms. The aggregated hourly series used for validation are archived alongside CERFRES (see Code and data availability).
Global Solar Power Tracker	Global Energy Monitor	Coordinates, capacity, status, start year, technology fields	Utility-scale and distributed PV metadata. `CC BY 4.0` (records with TZ ID: `CC BY-NC 4.0`). NC-licensed records are used as model input but are not reproduced in the released output; see Code and data availability.
Global Wind Power Tracker	Global Energy Monitor	Coordinates, capacity, status, start year, installation type	Wind metadata and offshore classification. `CC BY 4.0`
WorldPop population raster	WorldPop	Population-weighted grid	Distributed PV spatial downscaling. `CC BY 4.0`
National registries	Various	Capacity, location, commissioning year	Metadata supplementation. Varies by country; verify before redistribution. The matched NVE farm-level comparison extract used in the sub-national validation is archived alongside CERFRES (see Code and data availability).

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Parts of the processing code were developed with the assistance of Claude Code (Anthropic), an AI-based programming tool. Claude was also used to support literature search and to refine and improve portions of the manuscript text. All AI-assisted outputs were reviewed and verified by the authors.

Financial support. This research has been supported by the University of Bergen.

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

References

Bruvik, E.: CERFRES v1.0 Software, https://doi.org/10.5281/zenodo.20772916, 2026.

Bruvik, E. and Sorteberg, A.: CERFRES: High-Resolution European Renewable Energy Generation Dataset (v1.1.0), https://doi.org/10.5281/zenodo.19762446, 2026a.

Bruvik, E. and Sorteberg, A.: CERFRES: High-Resolution European Renewable Energy Generation Dataset (v1.1.1), https://doi.org/10.5281/zenodo.21479427, 2026b.

Dobos, A. P.: PVWatts Version 5 Manual, Tech. Rep. NREL/TP-6A20-62641, National Renewable Energy Laboratory, Golden, CO, https://www.nrel.gov/docs/fy14osti/62641.pdf, 2014.

Energinet: Production per Region per Hour, https://www.energidataservice.dk/dso-electricity/ProductionRegionHour, hourly electricity production per Danish region, divided by hydro, wind, solar and thermal generation, 2026.

ENTSO-E: ENTSO-E Transparency Platform, https://www.entsoe.eu/data/transparency-platform/, 2026.

Erbs, D., Klein, S., and Duffie, J.: Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation, Solar energy, 28, 293–302, 1982.

Global Energy Monitor: Global Solar Power Tracker, https://globalenergymonitor.org/projects/global-solar-power-tracker/, february 2026 release. Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0). Some records include data from TransitionZero Solar Asset Mapper (CC BY-NC 4.0)., 2026a.

Global Energy Monitor: Global Wind Power Tracker, https://globalenergymonitor.org/projects/global-wind-power-tracker/, february 2026 release. Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)., 2026b.

Gonzalez Aparicio, I., Zucker, A., Careri, F., Monforti, F., Huld, T., and Badger, J.: EMHIRES dataset. Part I: Wind power generation. European meteorological derived high resolution RES generation time series for present and future scenarios, Tech. Rep. EUR 28171 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-79-63193-1, ISSN 1831-9424, https://doi.org/10.2790/831549, 2016.

Gonzalez Aparicio, I., Huld, T., Careri, F., Monforti, F., and Zucker, A.: EMHIRES dataset. Part II: Solar power generation. European meteorological derived high resolution RES generation time series for present and future scenarios, using the PVGIS model, Tech. Rep. EUR 28629 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-79-69285-7, ISSN 1831-9424, https://doi.org/10.2760/044693, 2017.

González-Aparicio, I., Monforti, F., Volker, P., Zucker, A., Careri, F., Huld, T., and Badger, J.: Simulating European wind power generation applying statistical downscaling to reanalysis data, Applied Energy, 199, 155–168, https://doi.org/https://doi.org/10.1016/j.apenergy.2017.04.066, 2017.

Hirth, L., Mühlenpfordt, J., and Bulkeley, M.: The ENTSO-E Transparency Platform – A review of Europe's most ambitious electricity data platform, Applied Energy, 225, 1054–1067, https://doi.org/https://doi.org/10.1016/j.apenergy.2018.04.048, 2018.

Hoersch, J., Hofmann, F., Schlachtberger, D., and Brown, T.: PyPSA-Eur: An open optimisation model of the European transmission system, Energy Strategy Reviews, 22, 207–215, https://doi.org/10.1016/j.esr.2018.08.012, 2018.

Holmgren, W. F., Hansen, C. W., and Mikofski, M. A.: pvlib python: A python package for modeling solar energy systems, Journal of Open Source Software, 3, 884, https://doi.org/10.21105/joss.00884, 2018.

Jensen, T. V. and Pinson, P.: RE-Europe, a large-scale dataset for modeling a highly renewable European electricity system, Scientific Data, 4, 170 175, https://doi.org/10.1038/sdata.2017.175, 2017.

https://doi.org/10.5194/egusphere-2026-3654
Preprint. Discussion started: 24 July 2026

King, D. L., Boyson, W. E., and Kratochvil, J. A.: Photovoltaic Array Performance Model, Tech. Rep. SAND2004-3535, Sandia National Laboratories, Albuquerque, NM, https://doi.org/10.2172/919131, 2004.

National Renewable Energy Laboratory: turbine-models: Python package, https://pypi.org/project/turbine-models/, accessed: 2026-02-23, 2025.

Nørgaard, P. and Holttinen, H.: A multi-turbine power curve approach, in: Proceedings of the Nordic Wind Power Conference, vol. 1, pp. 1–5, Chalmers University of Technology Göteborg, Sweden, 2004.

NVE: Data for utbygde vindkraftverk i Norge, https://www.nve.no/energi/energisystem/vindkraft-paa-land/data-for-utbygde-vindkraftverk-i-norge/, accessed: 2026. Dataset published 18 August 2022, last updated 29 January 2026, 2026.

Pelosi, A.: Performance of the Copernicus European Regional Reanalysis (CERRA) dataset as proxy of ground-based agrometeorological data, Agricultural Water Management, 289, 108 556, 2023.

Pfenninger, S. and Staffell, I.: Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data, Energy, 114, 1251–1265, https://doi.org/https://doi.org/10.1016/j.energy.2016.08.060, 2016.

Pierrot, M.: The Wind Power: Wind Energy Market Intelligence, https://www.thewindpower.net, accessed: 2026, 2015.

Ridal, M., Bazile, E., Le Moigne, P., Randriamampianina, R., Schimanke, S., Andrae, U., Berggren, L., Brousseau, P., Dahlgren, P., Edvinsson, L., et al.: CERRA, the Copernicus European regional reanalysis system, Quarterly Journal of the Royal Meteorological Society, 150, 3385–3411, 2024.

Schmitz, A. and Després, J.: Global temporal power data collection: electricity load and power generation from solar and wind, JRC Publications Repository, https://doi.org/10.2760/182749 (online), 2022.

Souka, A. F. and Safwat, H. H.: Determination of the optimum orientations for the double-exposure flat-plate collector and its reflectors, Solar Energy, 10, 170–174, https://doi.org/10.1016/0038-092x(66)90004-1, 1966.

Staffell, I. and Pfenninger, S.: Using bias-corrected reanalysis to simulate current and future wind power output, Energy, 114, 1224–1239, https://doi.org/https://doi.org/10.1016/j.energy.2016.08.068, 2016.

WorldPop: Global High Resolution Population Denominators Project, https://www.worldpop.org, 100 m resolution gridded population dataset. Accessed 2026-03-11, 2026.