https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

A physically constrained framework for reconstructing high-resolution solar irradiance from accumulated radiation in machine-learning weather forecasts

Zekai Dai[1], Man Wang[2], Yongwei Wang[1,3], Yuanjian Yang[1,3], Xindin Li[1], Xiaoying Jiang[1], Yajun Wang[1], and Jiangwei Ji[1]

[1]School of Atmospheric Physics, Nanjing University of Information Science and Technology, Nanjing, China
[2]Yunnan Institute of Meteorological Sciences, Kunming, Yunnan 650034, China
[3]China Meteorological Administration Xiong'an Atmospheric Boundary Layer Key Laboratory, Xiong'an New Area, Hebei 071700, China

Correspondence: Yongwei Wang (wyw@nuist.edu.cn)

Abstract. Machine-learning weather prediction (MLWP) supports global forecasting, but solar radiation is often provided on coarse grids as multihour accumulations, creating a scale mismatch with local applications. We present SolarFINE, a physically constrained hierarchical framework that converts $0.25°$ six-hour accumulated radiation and concurrent meteorological states to surface solar irradiance at $0.05°$ and 10-min resolution. It reconstructs hourly coarse-grid irradiance from solar geometry and meteorological states under a six-hour integral-consistency constraint, then constructs an hourly high-resolution spatial background and generates subhourly variability using conditional diffusion. The temporal stage was trained on ERA5, the high-resolution stage was supervised with a satellite radiation product, and SolarFINE-T was adapted to Artificial Intelligence Forecasting System (AIFS) inputs. Evaluation covered hourly reconstruction, high-resolution spatial structure, and AIFS-driven end-to-end station performance. In controlled ERA5 experiments, SolarFINE-T achieved a daytime hourly root-mean-square error (RMSE) of $40.37\ \mathrm{W\,m^{-2}}$, 29.9% lower than with solar-geometry allocation. At $0.05°$ and 10-min resolution, SolarFINE-J achieved an RMSE of $132.42\ \mathrm{W\,m^{-2}}$ and more closely reproduced the spatial power spectrum, spatial variability, and high-irradiance tail of the high-resolution reference product. In the independent station evaluation, RMSE skill scores relative to solar-geometry allocation were positive on forecast days 1–15, rising from 0.084 to 0.174. Improvements were largest and most consistent at intermediate clear-sky index. Adapting to AIFS inputs improved skill, whereas forcing hourly fields to match the original six-hour accumulation reduced it. The results indicate that MLWP-driven solar-irradiance downscaling requires exploiting atmospheric information beyond solar geometry while accounting for upstream forecast errors.

Keywords: machine-learning weather prediction; solar irradiance; spatiotemporal downscaling; physical constraints; conditional diffusion

1 Introduction

The growing share of solar power in electricity systems increases the need for irradiance forecasts that resolve local temporal and spatial variability (Appenzeller, 2025; Nijsse et al., 2023). Global horizontal irradiance (GHI) is governed by solar geome-

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

try and atmospheric conditions: solar elevation sets the basic diurnal cycle, while clouds, water vapor, and terrain modulate its spatial distribution and variability from hourly to minute scales. These variations affect photovoltaic output and motivate short-term irradiance forecasting (Chu et al., 2021); fluctuations in distributed photovoltaic generation also matter for grid stability (Smith et al., 2022). Meeting these needs requires methods that retain the large-scale weather background while recovering fine-scale irradiance variability from coarse-resolution forecast information.

Solar irradiance forecasting relies on global numerical weather prediction (NWP) for large-scale atmospheric information. These systems combine data assimilation, numerical integration, and physical parameterizations to represent atmospheric evolution and unresolved processes (Bauer et al., 2015). Machine-learning weather prediction (MLWP) provides an additional source of global forecasts. FourCastNet, Pangu-Weather, GraphCast, GenCast, FuXi, and Aurora learn from historical atmospheric data to generate global, multivariable forecasts (Pathak et al., 2022; Bi et al., 2023; Lam et al., 2023; Price et al., 2025; Chen et al., 2023; Bodnar et al., 2025). The Artificial Intelligence Forecasting System (AIFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF) also provides surface variables, including surface solar radiation downwards (SSRD) (Lang et al., 2024; Moldovan et al., 2026). Later versions, including FuXi-2.0 and Aurora 1.5, also predict accumulated surface solar radiation (Zhong et al., 2024; Weyn et al., 2026). However, global radiation forecasts on coarse grids and with multihour accumulation windows do not directly resolve local irradiance variability. Such accumulations do not specify the irradiance distribution within each window, while coarse grids do not resolve local subgrid variability. Connecting these forecasts with local GHI applications therefore requires recovering physically consistent local irradiance fields and their variability at hourly to minute scales from the coarse, accumulated outputs.

Dynamical downscaling offers one route to recovering fine spatial detail from global forecasts. Regional models such as WRF-Solar represent regional and mesoscale processes through enhanced radiation, aerosol, and cloud schemes (Jimenez et al., 2016), while higher horizontal resolution helps resolve local structures absent from coarse global models (Fallah et al., 2025). However, high-resolution dynamical downscaling remains computationally expensive, especially at long forecast horizons and for ensembles (Fallah et al., 2025; Mardani et al., 2025). Statistical and machine-learning methods recover high-resolution fields from relationships between large-scale atmospheric states and local conditions. These approaches include multilayer perceptrons, convolutional networks, U-Net architectures, and image super-resolution models (Vandal et al., 2017; Baño-Medina et al., 2020; González-Abad and Gutiérrez, 2025; Rampal et al., 2024; Lafferty and Sriver, 2023; Liang et al., 2021). Meteorological downscaling also needs to account for the characteristics of each variable. Precipitation involves intermittency, extreme-value tails, and terrain effects (Fuentes-Franco et al., 2025; Rampal et al., 2025; Wolvin et al., 2026); wind fields involve terrain, vector structure, and small-scale variability (Hu et al., 2023; Guo et al., 2025; Getter et al., 2024); and thermodynamic relationships can constrain temperature and humidity (Jiang et al., 2023). Terrain and land-surface information provide additional guidance for recovering local structure (Giroux et al., 2025; Liu et al., 2024; Chen et al., 2024; de Roda Husman et al., 2024). The training objective also affects the variability recovered: deterministic models trained with pointwise losses such as mean squared error (MSE) tend toward the conditional mean, smoothing subgrid variability and weakening variance and distribution tails (Stengel et al., 2020; Getter et al., 2024; Rampal et al., 2025).

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

Generative models offer another route to representing the fine-scale variability smoothed by deterministic predictions. They sample plausible high-resolution states conditioned on the large-scale weather state to describe spatial dependence, spectral structure, and distribution tails (Leinonen et al., 2021; Bischoff and Deck, 2024). Applications span forecast downscaling and climate reconstruction. Harris et al. (2022) generated spatially coherent hourly precipitation ensembles from global IFS forecasts and evaluated them against hourly radar accumulations, while other models have produced kilometer-scale, subhourly precipitation fields from radar or ERA5 inputs (Glawion et al., 2023, 2025). Diffusion models have emulated daily precipitation from regional convection-permitting simulations conditioned on coarse global climate fields (Addison et al., 2026), reconstructed regional climate (Ling et al., 2024), and performed kilometer-scale downscaling (Yi et al., 2026). AIFS-driven typhoon downscaling also illustrates generative refinement of forecast inputs (Niu et al., 2026). CorrDiff decomposes a high-resolution field into a deterministic background and a stochastic fine-scale residual, assessing the generated structure with spatial power spectra (Mardani et al., 2025). A latent diffusion model has similarly refined a U-Net background to downscale ERA5 temperature and wind to kilometer-scale regional model fields (Tomasi et al., 2025). Multivariate regional climate downscaling has generated temperature, precipitation, wind, and solar radiation, with evaluation of calibration, spatiotemporal structure, and extremes (Schillinger et al., 2026). Wind-field studies also indicate that generative models may recover high-wavenumber spectral energy and heavy-tailed velocity-gradient distributions even when improvements in pointwise error are limited (Stengel et al., 2020). These applications motivate evaluation of spatial structure and probabilistic behavior alongside pointwise error.

For high-resolution downscaling intended for weather forecasting, realistic generated structure is only one requirement. Fine-scale textures may appear visually plausible without being supported by the upstream weather state and may instead form algorithmically generated structures that are inconsistent with the conditioning information (Annau et al., 2023). To address this risk, previous studies have combined physically informed constraints with interpretability analyses and diagnostics of spatial structure, distributional behavior, and ensemble calibration to test whether generated fields remain consistent with their conditioning state (de Roda Husman et al., 2024; Harder et al., 2023; González-Abad et al., 2023; Leinonen et al., 2021; Harris et al., 2022; Mardani et al., 2025). For Earth system model precipitation, conditional diffusion has combined downscaling with bias correction while preserving spatial patterns above a selected correction scale (Aich et al., 2026). For forecast-driven applications, diffusion-based wind-speed downscaling has combined spatial refinement with bias correction and compared its outputs against the driving forecast and post-processing baselines (Springenberg et al., 2026). Nevertheless, downstream downscaling remains sensitive to the quality of its coarse inputs: differences between those inputs and the target data in spatiotemporal structure can reduce model performance (Glawion et al., 2025). Compared with precipitation and wind applications, such assessments of generated structure and probabilistic consistency remain relatively uncommon in GHI downscaling.

Existing studies have explored spatial super-resolution of GHI, diffusion-based cloud-field generation for solar forecasting, and kilometer-scale ensembles driven by machine-learning weather models (Cheng et al., 2024; Hatanaka et al., 2023; Liu et al., 2026). Station comparisons also show that a finer forecast grid need not improve GHI accuracy (Mayer et al., 2023). Yet few studies have examined the full chain of recovering hourly states from multihour accumulated radiation, generating subhourly high-resolution fields, and evaluating the system under upstream forecast errors. We therefore address three related but non-

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

interchangeable questions: How much hourly information can be recovered from accumulated radiation? Are the generated fine-scale structures statistically plausible? Do these capabilities translate into improvements against independent observations when the system is driven by forecast inputs?

To address these questions, we introduce SolarFINE, a physically constrained, hierarchical framework for spatiotemporal downscaling of accumulated radiation from MLWP systems. SolarFINE-T first combines six-hour accumulated radiation, concurrent meteorological states, and solar geometry to reconstruct hourly coarse-grid irradiance and meteorological fields, using a six-hour integral-consistency loss to regularize the temporal allocation during training. SolarFINE-J then uses these hourly fields in two sequential stages: a deterministic spatial stage that downscales hourly irradiance to $0.05°$, and a conditional diffusion stage that generates six 10-min irradiance fields per hour while enforcing nonnegativity and consistency with the hourly mean. SolarFINE-T was trained using the ERA5 reanalysis (Hersbach et al., 2020), whereas SolarFINE-J received high-resolution supervision from the Shortwave Downward Radiation over the East Asia–Pacific region data set (SWDR_EAP; Husi et al., 2021; Letu et al., 2022). Paired AIFS–ERA5 samples were used to adapt SolarFINE-T to AIFS forecast inputs. We evaluate the framework at three levels. Controlled ERA5 experiments assess hourly reconstruction; SWDR_EAP serves as the high-resolution reference for evaluating the errors, spatial structure, and probabilistic characteristics of the $0.05°$, 10-min outputs; and the complete AIFS-driven system is evaluated against independent ground observations for end-to-end hourly GHI forecasting. Hourly station comparisons assess the forecast-driven chain at the hourly scale; intrahour generation is evaluated separately using the native 10-min outputs.

2 Methods

2.1 Problem Formulation

Given six-hour accumulated SSRD and concurrent meteorological states from reanalysis or MLWP, this study seeks to learn a mapping across temporal and spatial scales that transforms radiation information on a coarse $0.25°$ grid into surface solar irradiance fields at $0.05°$ and 10-min resolution. SolarFINE-T first reconstructs hourly irradiance and meteorological states on the coarse grid; SolarFINE-J then uses these time-matched outputs for spatial refinement and intrahour generation. Controlled experiments use ERA5 reanalysis as input to evaluate temporal reconstruction and high-resolution generation without upstream forecast errors. Forecast-domain experiments instead use AIFS forecasts to evaluate the end-to-end performance of the complete SolarFINE system in the presence of forecast errors and differences in the input distribution. ERA5 hourly fields and SWDR_EAP provide the respective references for temporal reconstruction and high-resolution generation, while end-to-end station evaluation is conducted at the hourly scale.

The study domain covers the middle and lower reaches of the Jinsha River basin and most of Yunnan Province, extending from $21.0°\mathrm{N}$ to $29.75°\mathrm{N}$ and from $97.5°\mathrm{E}$ to $106.25°\mathrm{E}$ (Figure 1). The region lies in the transition from the Tibetan Plateau to the Yunnan–Guizhou Plateau and is jointly influenced by the East Asian monsoon, the South Asian monsoon, and circulation along the plateau margins. During the rainy season, warm and moist airflows interact with complex terrain to influence cloud formation and evolution, producing pronounced spatial heterogeneity and temporal variability in surface shortwave radiation.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

Surface irradiance is generally higher during the dry season, although local terrain and weather conditions can still produce marked irradiance differences. The region therefore provides a challenging environment for evaluating the conversion of coarse-resolution accumulated radiation into high-resolution irradiance fields over complex terrain. We further use surface radiation observations within the study area to independently evaluate AIFS-driven SolarFINE outputs at station locations.

We use the following notation. A six-hour window is indexed by q, the hourly steps within each window by $h = 1, \ldots, 6$, and the 10-min frames within each hour by $t = 1, \ldots, 6$. Coarse-grid fields at $0.25°$, or low resolution (LR), are denoted by the superscript LR, whereas high-resolution (HR) fields at $0.05°$ are denoted by HR. Accumulated solar radiation energy is denoted by E, with units of $\mathrm{J\,m^{-2}}$, and mean irradiance over a specified interval is denoted by I, with units of $\mathrm{W\,m^{-2}}$. Accumulated radiation from each data source is first converted to accumulated energy over the corresponding time window according to its native temporal definition and is then converted to mean irradiance over the required interval for model input or evaluation. The specific treatment of accumulated radiation from AIFS and ERA5 is described in Section 2.2.

2.2 Data

SolarFINE comprises two sequentially connected submodels: SolarFINE-T reconstructs hourly coarse-grid states from 6 h accumulated radiation, and SolarFINE-J performs high-resolution spatial and subhourly downscaling conditioned on the reconstructed hourly states (Section 2.3). We use five categories of data: AIFS forecasts, ERA5 reanalysis, the SWDR_EAP high-resolution radiation product, SRTM topography and solar-geometry information, and surface radiation observations. ERA5 data from 2016–2019 are used for model training and validation, and 2020 is reserved for independent controlled testing. AIFS forecasts for 2023 are paired with the corresponding ERA5 data to adapt SolarFINE-T to forecast inputs, whereas the 2024 AIFS forecasts and surface observations are used only for independent end-to-end evaluation and do not enter model training.

2.2.1 AIFS Forecasts

Forecast inputs are generated by AIFS, a data-driven global weather prediction model that learns the spatiotemporal evolution of the atmosphere to produce multivariable medium-range forecasts (Lang et al., 2024; Moldovan et al., 2026).

To ensure that forecast-input adaptation and independent evaluation use the same upstream model, we use the publicly released AIFS Single v1.0 checkpoint from ECMWF (Pinnington et al., 2025; ECMWF, 2025) with anemoi-inference to regenerate retrospective forecasts for 2023–2024, hereafter referred to as the AIFS forecasts. All experiments use the same model weights, variable configuration, and inference procedure, and each retrospective forecast is initialized from the corresponding ERA5 atmospheric state. In this study, forecasts are initialized daily at 00 UTC, extend to a maximum lead time of 15 d, and provide output every 6 h on a $0.25°$ horizontal grid. The 2023 forecasts are paired with ERA5 to adapt SolarFINE-T to forecast inputs. The 2024 forecasts are independent of model training and input adaptation and are used only as upstream inputs for the end-to-end station evaluation. SolarFINE-J remains frozen during forecast-input adaptation. AIFS provides SSRD as independent, nonoverlapping 6 h accumulations; each accumulation is used directly as E_q^{LR} and converted to mean irradiance without differencing adjacent forecast lead times.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

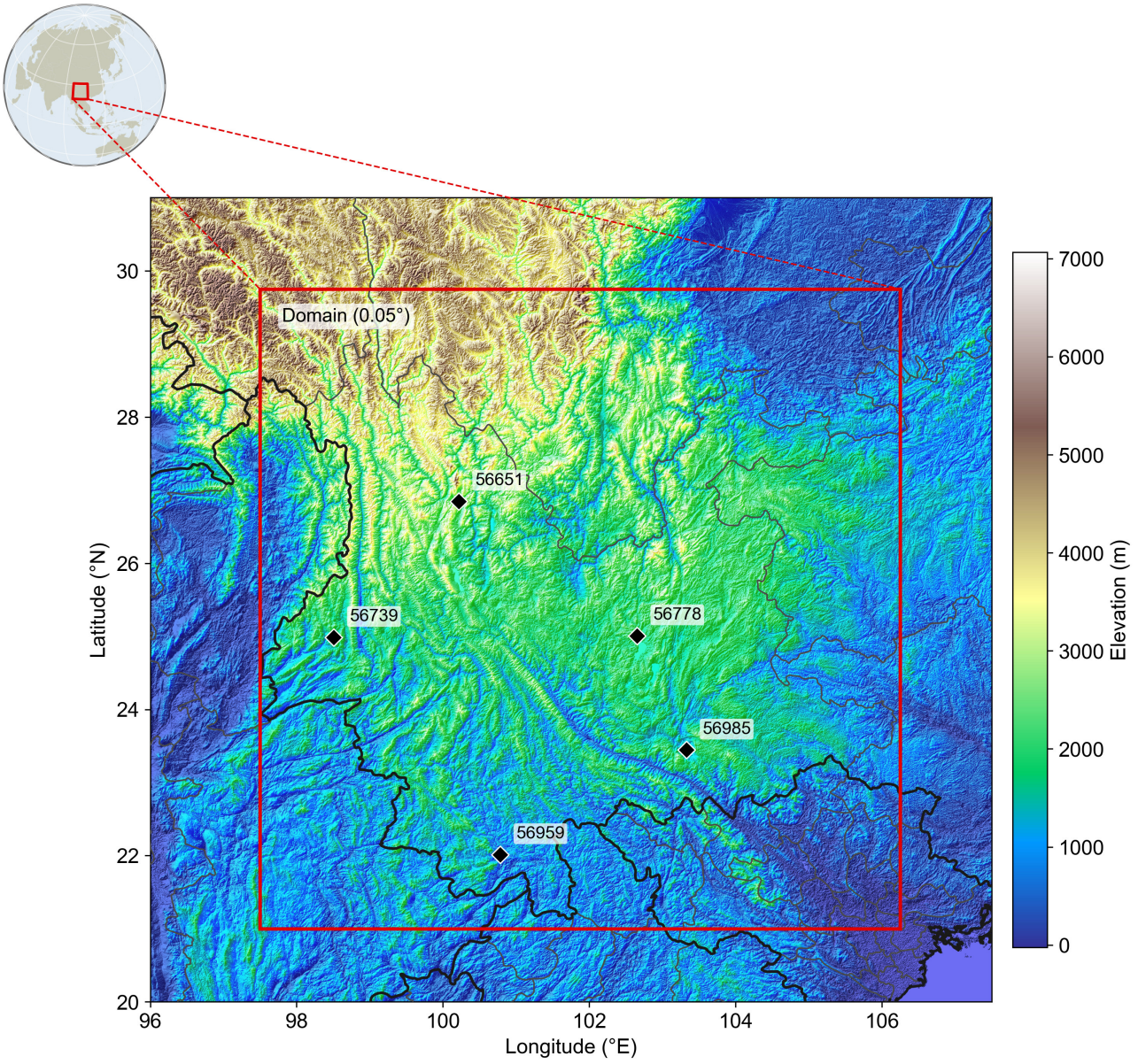


Figure 1. Study area and surface radiation observation stations. The red box denotes the SolarFINE model domain, and the black diamonds mark the surface radiation observation stations. Gray and black solid lines indicate provincial administrative boundaries and national borders, respectively. The inset shows the geographical location of the study area.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.



2.2.2 ERA5 Reanalysis

ERA5 is a global hourly reanalysis produced by ECMWF; we use data on its 0.25° grid (Hersbach et al., 2020). ERA5 serves two purposes in SolarFINE. In the controlled experiments, it provides both the low-resolution inputs and the corresponding hourly targets, allowing temporal reconstruction to be evaluated without upstream forecast errors. In 2023, it is paired with AIFS forecasts to provide supervisory targets for adapting SolarFINE-T to forecast inputs.

SolarFINE-T uses 20 dynamic variables: SSRD, 2 m temperature, mean sea level pressure, total column water vapor, total cloud cover, high-, medium-, and low-level cloud cover, and temperature, specific humidity, and the two horizontal wind components at 200, 500, and 850 hPa. These variables describe the radiation background, thermal state, moisture, clouds, and large-scale circulation that influence surface shortwave radiation. The corresponding hourly variables serve both as supervisory targets for SolarFINE-T and as dynamic conditions for SolarFINE-J.

The shortwave radiation accumulations from ERA5 are converted to hourly accumulated energy according to their native temporal definition and divided by 3600 s to obtain hourly mean irradiance. For the 6 h input windows of SolarFINE-T, the six corresponding hourly radiation-energy fields are summed to form a 6 h accumulation. The accumulated input and hourly targets in the controlled ERA5 experiments therefore come from the same source, allowing the reconstruction of hourly states from 6 h accumulated information to be evaluated separately.

ERA5 samples from 2016–2019 are used for model training and validation, with the specific split described in Section 2.4. The full year of 2020 is reserved as an independent controlled test set. ERA5 data from 2023 are used only as supervisory targets for forecast-input adaptation and do not enter the 2024 end-to-end evaluation.

2.2.3 SWDR_EAP High-Resolution Radiation Reference Product

High-resolution supervision and controlled evaluation use the Shortwave Downward Radiation over the East Asia–Pacific region (SWDR_EAP) product (Husi et al., 2021). This product is retrieved primarily from Himawari-8/AHI observations and incorporates MODIS and surface observations. It covers 60°S–60°N and 80°E–160°W from 2016 to 2020, with a temporal resolution of 10 min and a spatial resolution of 0.05°. Letu et al. (2022) describe the retrieval and validation and report hourly and daily root-mean-square errors (RMSEs) of 104.9 and 31.5 $\mathrm{W\,m^{-2}}$, respectively.

SWDR_EAP serves as the high-resolution reference product for spatial and subhourly supervision of SolarFINE-J and for the controlled high-resolution evaluation in 2020; it is not used to train the temporal reconstruction in SolarFINE-T. For target hour h, the six consecutive 10-min fields are denoted by $\{Y_{h,t}^{\mathrm{HR}}\}_{t=1}^{6}$, with each frame defined on a 176×176 high-resolution grid. Hourly results are evaluated using the arithmetic mean of the six frames.

2.2.4 Topographic and Solar-Geometry Conditions

Static topographic conditions are obtained from the Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM). The original elevation data at approximately 90 m resolution are spatially aggregated or resampled to the 0.05° SolarFINE target grid to represent the high-resolution topographic background.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

Solar-geometry variables are calculated analytically at each target grid point and valid time. The cosine of the solar zenith angle, $\cos\theta_z$, represents the basic geometric constraint on the diurnal cycle, and top-of-atmosphere solar irradiance is calculated from the solar constant and solar zenith angle. Theoretical clear-sky GHI is estimated using the Haurwitz model (Haurwitz, 1945) and used as an auxiliary condition for SolarFINE-J. The clear-sky index used to stratify sky conditions is defined as $k_c = I_{\mathrm{actual}}/I_{\mathrm{clear}}$ and is calculated only when $\cos\theta_z > 0.05$. Actual GHI is taken from SWDR_EAP in the controlled high-resolution evaluation and from surface observations in the station evaluation. The index describes irradiance relative to the Haurwitz clear-sky estimate and is influenced by clouds, water vapor, aerosols, and clear-sky model errors. High, intermediate, and low k_c therefore describe different irradiance conditions.

2.2.5 Surface Radiation Observations

The independent station evaluation uses hourly surface solar radiation observations for 2024 provided by the Yunnan Provincial Meteorological Bureau. The observed variable is total shortwave radiation reaching the surface. Stations 56651, 56739, 56778, 56959, and 56985 are used, and their locations are shown in Figure 1. These observations are used to evaluate the hourly GHI produced by the complete AIFS-driven SolarFINE system and are excluded from model training, validation, and the 2023 forecast-input adaptation.

Before evaluation, the observations are time aligned, screened for anomalous values and missing samples, and paired with the nearest 0.05° model grid point and the corresponding model valid time. In the station order listed above, the missing-data rates are 0.05%, 0.00%, 0.59%, 0.62%, and 0.20%, respectively. The station records contain hourly accumulated radiation energy, which is divided by 3600 s to obtain mean irradiance for each hour in $\mathrm{W\,m^{-2}}$. All methods are evaluated using the same station–valid-time pairs.

2.3 Model

SolarFINE sequentially transforms 6 h accumulated radiation on a coarse 0.25° grid into hourly states at the same resolution, hourly irradiance fields at 0.05°, and 10-min irradiance fields at 0.05°. For a 6 h window q, these transformations are written as

$$\left(E_q^{\mathrm{LR}}, \mathbf{X}_q^{\mathrm{LR}}, \mathbf{\Gamma}_q\right) \xrightarrow{T_\phi} \left(\{\hat{I}_{q,h}^{\mathrm{LR}}\}_{h=1}^{6}, \{\hat{\mathbf{X}}_{q,h}^{\mathrm{LR}}\}_{h=1}^{6}\right) \xrightarrow{H_{\theta,\psi}(\cdot;D)} \{\hat{Y}_{q,h,t}^{\mathrm{HR}}\}_{h=1,\ldots,6;t=1,\ldots,6}. \tag{1}$$

Here, E_q^{LR} denotes 6 h accumulated radiation, $\mathbf{X}_q^{\mathrm{LR}}$ is the coarse-grid meteorological state, $\mathbf{\Gamma}_q$ contains analytical conditions such as solar geometry and clear-sky irradiance, and D denotes the high-resolution topographic condition. SolarFINE comprises two sequentially connected submodels. SolarFINE-T reconstructs hourly coarse-grid irradiance and meteorological states from the 6 h accumulated inputs, whereas SolarFINE-J uses the reconstructed hourly states to generate irradiance fields at 0.05° and 10-min resolution. Specifically, $H_{\theta,\psi} = D_\psi \circ S_\theta$, where S_θ is the hourly spatial conditional generator and D_ψ is the subhourly conditional diffusion generator. The overall architecture is shown in Figure 2.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

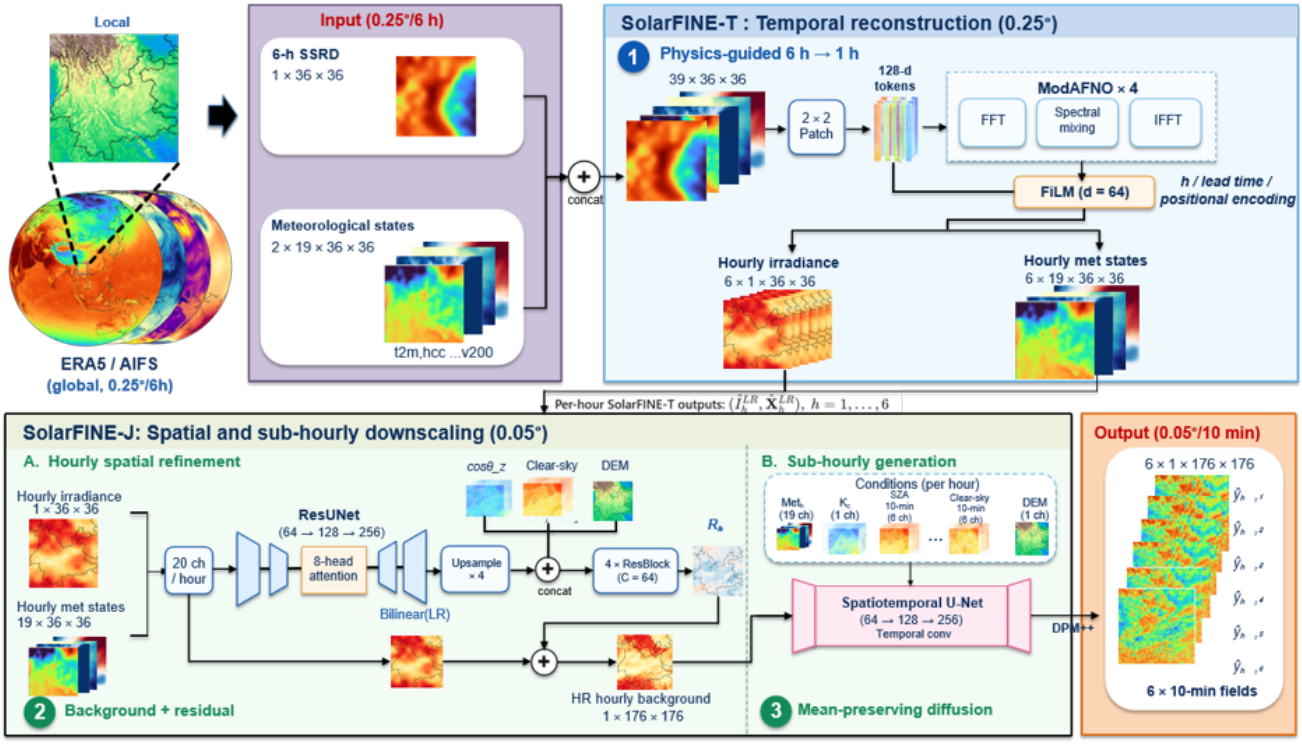


Figure 2. SolarFINE framework for spatiotemporal downscaling. SolarFINE-T reconstructs hourly coarse-grid fields from 6 h accumulated radiation and concurrent meteorological states. SolarFINE-J then generates high-resolution hourly backgrounds and subhourly irradiance fields. The successive stages incorporate solar geometry, 6 h integral consistency, nonnegativity, and hourly-mean consistency.

2.3.1 SolarFINE-T

SolarFINE-T transforms 6 h accumulated radiation and concurrent meteorological states into six coarse-grid states at hourly intervals. Its inputs include the nonradiative meteorological variables at the beginning and end of the window, 6 h accumulated SSRD, and the cosine of the solar zenith angle, $\cos\theta_z$, for each of the six target hours. Top-of-atmosphere (TOA) solar irradiance is supplied as an analytical condition to the radiation decoder. A shared decoder produces hourly irradiance and the same set of meteorological variables used on input for all six target hours; all reconstructed meteorological channels are passed to SolarFINE-J.

The 6 h accumulation constrains the total radiation within the window, whereas its hourly allocation also depends on the concurrent meteorological state and solar elevation at each target hour. To map the same accumulated input to six distinct hourly states, SolarFINE-T uses a modulated Adaptive Fourier Neural Operator (ModAFNO) to represent large-scale spatial relationships on the coarse grid. The target-hour position encoding and forecast lead time provide the conditioning variables for feature-wise linear modulation (FiLM). The network configuration is described in Appendix A.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

Hourly irradiance is parameterized using a bounded effective shortwave transmission factor. The model first predicts this factor $\hat{\tau}_{q,h}$ for each target hour and then multiplies it by the corresponding TOA irradiance:

$$\hat{I}^{\mathrm{LR}}_{q,h} = \hat{\tau}_{q,h} \odot I^{\mathrm{TOA}}_{q,h}, \qquad \hat{\tau}_{q,h} \in [0,1]. \tag{2}$$

Here, $I^{\mathrm{TOA}}_{q,h} = 1361 \max(\cos\theta_z, 0)$, with units of $\mathrm{W\,m^{-2}}$. The bounded factor $\hat{\tau}_{q,h}$ represents effective shortwave transmission in the model. The parameterization incorporates solar geometry and nonnegativity and uses this TOA reference as an upper bound on model output. The fixed solar constant and target-time solar geometry approximate the hourly mean irradiance, without explicitly accounting for seasonal Earth–Sun distance variations or local cloud enhancement. The TOA reference therefore provides an approximate upper bound on model output.

SolarFINE-T jointly optimizes the hourly meteorological states, hourly irradiance, and 6 h integral consistency. The total loss is

$$\mathcal{L}_T = \mathcal{L}_{T,\mathrm{met}} + \mathcal{L}_{T,\mathrm{rad}} + \lambda_E \mathcal{L}_{T,E},$$

where $\mathcal{L}_{T,\mathrm{met}}$ and $\mathcal{L}_{T,\mathrm{rad}}$ are the losses for the hourly meteorological fields and irradiance, respectively. The integral-consistency term is

$$\mathcal{L}_{T,E} = \ell\left(3600 \sum_{h=1}^{6} \hat{I}^{\mathrm{LR}}_{q,h}, E^{\mathrm{truth}}_{q}\right), \qquad E^{\mathrm{truth}}_{q} = 3600 \sum_{h=1}^{6} I^{\mathrm{truth}}_{q,h}. \tag{3}$$

This term constrains the integrated energy of the six reconstructed hourly irradiance fields to remain consistent with the corresponding 6 h accumulated target. It is normalized by the standard deviation of accumulated radiation in the training samples, and we set $\lambda_E = 0.15$. Training settings and the treatment of cross-scale consistency are provided in Appendix B.

2.3.2 SolarFINE-J

SolarFINE-J further transforms the hourly coarse-grid states into high-resolution spatial fields and subhourly irradiance sequences. It uses a two-stage structure that combines deterministic hourly spatial downscaling with conditional diffusion. The first stage constructs a high-resolution spatial background for each hour, and the second stage generates six 10-min irradiance fields conditioned on that background.

To retain the large-scale radiation background while recovering local differences associated with weather conditions, solar geometry, and terrain, the hourly spatial stage uses residual downscaling. Its dynamic inputs are the current-hour coarse-grid irradiance and meteorological states produced by SolarFINE-T. The cosine of the solar zenith angle, Haurwitz clear-sky irradiance, and the DEM are included as additional conditions. This stage is implemented using a ResUNet with a global attention module at the bottleneck. Starting from the bilinearly interpolated coarse-grid irradiance field $B^{\mathrm{bilinear}}_{q,h}$, the hourly high-resolution spatial background is given by

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

$$\hat{M}^{\mathrm{HR}}_{q,h} = \max\left(0, B^{\mathrm{bilinear}}_{q,h} + R_{q,h}\right), \tag{4}$$

where $R_{q,h}$ is the high-resolution spatial correction; nonnegative clipping is applied at inference.

The hourly background does not uniquely determine cloud-driven variability within the hour. Conditional diffusion represents the subhourly variability that remains unresolved given the hourly weather and radiation background and generates multiple possible sequences of six 10-min irradiance fields. The conditioning variables include the hourly spatial background, the dynamic meteorological states produced by SolarFINE-T, the clear-sky index k_c, the solar zenith angle and clear-sky irradiance at each of the six target 10-min times, and the DEM. When used as a model condition, k_c is calculated using the hourly spatial background from SolarFINE-J as the numerator and theoretical clear-sky GHI at the same location and hour as the denominator. By contrast, the sky-condition categories used in the controlled and station evaluations are calculated using SWDR_EAP and station-observed GHI, respectively, as the numerator. The denoiser is a spatiotemporal U-Net that combines two-dimensional residual blocks for spatial features with temporal convolutions across the six frames. It is trained using a 1000-step cosine noise schedule and a v-prediction objective. Inference uses DPM-Solver++(2M) SDE with 20 sampling steps, distinct from the training noise schedule. The denoised six-frame variation is mapped back to nonnegative irradiance while preserving the given hourly mean, as described in Eqs. (5)–(6). The network configuration is provided in Appendix A.

The conditional diffusion stage uses the hourly high-resolution irradiance background and weather, solar-geometry, and terrain conditions to generate six consecutive 10-min irradiance fields at 0.05° resolution, representing possible within-hour variations while retaining the prescribed hourly mean. To represent intrahour variability separately from the hourly irradiance amplitude, the diffusion model operates on the mean-centered logarithms of the six reference irradiance fields. This transformation removes the six-frame mean in log space; it does not subtract the predicted hourly background. For the six reference fields within the same hour, the clean diffusion variable is defined as

$$z_t = \log(Y_t + 1) - \frac{1}{6}\sum_{j=1}^{6} \log(Y_j + 1). \tag{5}$$

Let z denote the stack of these six fields and k index the diffusion noise level, distinct from the 10-min frame index t. During training, Gaussian noise ϵ is added through $z^{(k)} = \alpha_k z + \sigma_k \epsilon$, where $\alpha_k^2 + \sigma_k^2 = 1$ and the coefficients follow the cosine noise schedule. The denoiser predicts the velocity target $v^{(k)} = \alpha_k \epsilon - \sigma_k z$ from $z^{(k)}$, the noise level, and the conditioning information described above. The diffusion loss is the mean squared error between predicted and target v, averaged over the six frames and spatial pixels and then over valid daytime training samples. Thus, Eq. (5) specifies the clean intrahour representation, whereas v specifies the denoising supervision. The hourly irradiance amplitude is supplied separately during the constrained reconstruction in Eq. (6).

The reconstruction step converts the sampled six-frame variations into irradiance while preserving the hourly background supplied by the spatial stage. Irradiance is measured in $\mathrm{W\,m^{-2}}$, so the additive constant 1 in Eq. (5) corresponds to $1\ \mathrm{W\,m^{-2}}$. At each spatial pixel and for each ensemble member, the six sampled log weights are first centered by subtracting their six-

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.



frame mean, clipped to $[-6, 6]$, and exponentiated to give positive weights w_t. Each weight is divided by the mean of all six weights and then multiplied by the hourly background M, giving $Y_t = Mw_t/[(w_1 + \cdots + w_6)/6]$. The normalized weights therefore have a six-frame mean of one. This mapping guarantees nonnegative irradiance and preserves the prescribed hourly mean:

$$\frac{1}{6}\sum_{t=1}^{6} \hat{Y}_{q,h,t}^{\mathrm{HR},(m)} = \hat{M}_{q,h}^{\mathrm{HR}}, \tag{6}$$

where m denotes a diffusion ensemble member. This constraint prevents the stochastic subhourly variations from changing the mean irradiance of the hourly spatial background. Samples below the clear-sky irradiance threshold skip stochastic sampling; all six frames then take the same nonnegative hourly background value, becoming zero where that background is zero.

2.4 Training Strategy

We trained SolarFINE-T using ERA5 samples from 2016–2019, with 6 h accumulated radiation and concurrent meteorological states as inputs and the corresponding hourly irradiance and meteorological states as supervisory targets. Data from 2016–2018 and days 1–15 of each month in 2019 were used for training, whereas days 16 through the end of each month in 2019 were used for validation. The full year of 2020 was reserved as an independent test set. The parameters of SolarFINE-T were frozen after training.

Using the frozen SolarFINE-T, we generated hourly states for the 2016–2019 samples and used them as dynamic inputs to SolarFINE-J, with the corresponding high-resolution SWDR_EAP irradiance fields as supervisory targets. The hourly spatial generator was first trained against hourly mean SWDR_EAP fields and then frozen. The conditional diffusion model was subsequently trained against the six 10-min SWDR_EAP irradiance fields within each hour. SolarFINE-J used the same training and validation periods as SolarFINE-T. Network configurations are provided in Appendix A, and the optimizer, learning rate, batch size, and number of training epochs are reported in Appendix B.

ERA5 reanalysis and AIFS forecasts differ systematically in their variable distributions and forecast-error characteristics. To adapt the hourly reconstruction to forecast inputs, we further adapted SolarFINE-T using AIFS forecasts throughout 2023 paired with the corresponding ERA5 data. The split was based on forecast initialization dates: forecasts initialized on days 5, 12, 19, and 26 of each month were reserved for validation, and those initialized on the remaining dates were used for adaptation training. SolarFINE-J remained frozen throughout this stage. The ERA5-pretrained model without forecast-input adaptation is denoted SolarFINE-noTune.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

2.5 Inference and Evaluation Design

2.5.1 Inference, Evaluation Framework, and Samples

AIFS forecasts from 2024 are used for end-to-end inference. Each AIFS input is first transformed by SolarFINE-T into hourly coarse-grid states and then passed to the frozen SolarFINE-J to generate irradiance fields at 0.05° and 10-min resolution. Before evaluation, outputs and references are time matched and screened for missing data. Daytime is defined by $\cos\theta_z > 0.05$; station evaluation and sky-condition stratification use this criterion. Overall high-resolution RMSE uses all spatial pixels in the six-frame test windows; aggregation RMSE is calculated from the same sample pool at the temporal and spatial scales described in Section 2.5.3. We also diagnose nonzero nighttime values, negative irradiance, 6 h integral consistency, and hourly-mean consistency.

We evaluate SolarFINE at three successive levels. Level 1 evaluates the reconstruction of hourly irradiance from 6 h accumulated radiation under controlled ERA5 conditions. Level 2 uses SWDR_EAP as the high-resolution reference product to evaluate pointwise error, spatial structure, distributional characteristics, and probabilistic behavior at 0.05° and 10-min resolution. Level 3 uses AIFS forecasts as inputs and evaluates hourly GHI against independent surface observations. Under the hourly-mean constraint in Eq. (6), averaging the six generated 10-min fields recovers the shared hourly spatial background for every ensemble member. The station evaluation therefore assesses the hourly reconstruction and spatial background under forecast inputs, while Level 2 assesses the generated intrahour fields and their conditional probabilistic behavior.

Table 1. Hierarchical evaluation framework for SolarFINE.

Level	Input	Reference	Evaluation target
Level 1: Hourly reconstruction	ERA5 6 h accumulated radiation and meteorological fields	ERA5 hourly irradiance	Ability to reconstruct hourly irradiance from 6 h accumulated information
Level 2: High-resolution downscaling	Hourly states reconstructed by SolarFINE-T	SWDR_EAP fields at 0.05° and 10-min resolution	Pointwise error, spatial structure, distributions, and probabilistic characteristics of the high-resolution fields
Level 3: End-to-end forecasting	AIFS forecasts	Surface radiation observations	Hourly station-level GHI performance under forecast inputs

All reported 95% confidence intervals are estimated using 2000 block-bootstrap resamples. Calendar dates form the blocks for the controlled evaluations, and forecast initialization dates form the blocks for the AIFS-driven station evaluation.

2.5.2 Hourly Reconstruction Evaluation

Level 1 uses the independent ERA5 test set from 2020. The solar zenith angle (SZA) allocation distributes each 6 h accumulated radiation value among the six target hours in proportion to their TOA irradiance while preserving the original 6 h total. ConvLSTM serves as the temporal deep-learning baseline, and the original ModAFNO provides a spectral-architecture

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

comparison. Evaluation metrics include hourly RMSE, the residual in 6 h integrated radiation, and errors stratified by cloud cover and SZA.

2.5.3 High-Resolution and Probabilistic Evaluation

Level 2 uses the common test samples from 2020. Overall RMSE in Figure 4 and aggregation RMSE in Figure 5 pool all spatial pixels across the six 10-min frames in each of the 3,878 windows, without an additional pixel-level solar-zenith mask. The clear-sky index used for sky-condition stratification is still calculated from daytime pixels. SolarFINE-J is compared with bilinear interpolation, DeepSD-Solar, and SwinIR-Solar. These three deterministic baselines first produce hourly high-resolution fields, which are allocated to six 10-min fields using solar geometry and clear-sky irradiance weights. A deterministic configuration with the same hourly spatial generator but without diffusion is included as a control for the subhourly generation mechanism. RMSE at 0.05° and 10-min resolution is the primary deterministic metric. SolarFINE-J uses the 20-member ensemble mean for deterministic evaluation and the full ensemble for the continuous ranked probability score (CRPS). Spatial spectra and marginal distributions are examined separately for the ensemble mean and a representative member (member 0). Radially averaged spatial power spectra, marginal irradiance distributions, spatial standard-deviation error, 99th-percentile (P99) bias, and RMSE after temporal and spatial aggregation are also used to evaluate the statistical structure of the high-resolution fields. Sky conditions are divided into high-, medium-, and low-k_c categories using k_c calculated from SWDR_EAP. Terrain complexity is divided into terciles using the local standard deviation of the DEM within a 9×9 neighborhood.

To distinguish errors in the shared hourly background from those in the generated intrahour evolution, we apply an orthogonal MSE decomposition to the 10-min ensemble mean. Let $\widetilde{Y}^{\mathrm{HR}}_{q,h,t}$ denote the ensemble-mean prediction, $Y^{\mathrm{HR}}_{q,h,t}$ the corresponding reference, and $e_{q,h,t} = \widetilde{Y}^{\mathrm{HR}}_{q,h,t} - Y^{\mathrm{HR}}_{q,h,t}$ their error. Defining the hourly-mean error $\bar{e}_{q,h} = \frac{1}{6}\sum_{t=1}^{6} e_{q,h,t}$ and the intrahour-anomaly error $e'_{q,h,t} = e_{q,h,t} - \bar{e}_{q,h}$ gives

$$\underbrace{\frac{1}{6}\sum_{t=1}^{6} e^2_{q,h,t}}_{\mathrm{MSE}_{\mathrm{total}}} = \underbrace{\bar{e}^2_{q,h}}_{\mathrm{MSE}_{\mathrm{hourly\ mean}}} + \underbrace{\frac{1}{6}\sum_{t=1}^{6}\left(e'_{q,h,t}\right)^2}_{\mathrm{MSE}_{\mathrm{intrahour\ anomaly}}} . \tag{7}$$

Because the anomaly errors sum to zero within each hour, the cross term vanishes and the two components add exactly to the total MSE. The components are first calculated for each hour and grid point and then averaged over the same valid samples and spatial weights.

2.5.4 Independent Station Evaluation

Level 3 drives the complete SolarFINE system with AIFS forecasts from 2024 and matches the resulting hourly forecasts with concurrent surface radiation observations. AIFS-SZA distributes each AIFS 6 h radiation accumulation among the target hours using TOA irradiance weights and serves as the primary physical baseline. The complete SolarFINE system is also compared with SolarFINE-noTune and SolarFINE-C. SolarFINE-C uses the adapted SolarFINE-T and the same SolarFINE-J. At each coarse-grid point within a 6 h window, it multiplies the six hourly irradiance values from SolarFINE-T by a common factor,

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

$c_q = E_q^{\mathrm{AIFS}}/(3600\sum_{j=1}^{6}\hat{I}_{q,j}^{\mathrm{LR}})$, so that their integrated total equals the original AIFS 6 h accumulation while their relative hourly allocation is preserved. The rescaled hourly irradiance and the corresponding meteorological states are then passed to SolarFINE-J. This configuration is used to examine the sensitivity of the model results to the treatment of accumulated-radiation totals.

Forecast lead times are grouped by forecast day: D1 covers +7 to +24 h after initialization, each subsequent forecast day comprises the next consecutive 24 h window, and D15 covers +337 to +360 h.

The primary station-evaluation metric is daytime hourly GHI RMSE. Sky conditions are stratified using k_c calculated from station-observed GHI. The RMSE skill score relative to AIFS-SZA is defined as

$$SS = 1 - \frac{\mathrm{RMSE}_{\mathrm{model}}}{\mathrm{RMSE}_{\mathrm{AIFS-SZA}}}, \tag{8}$$

where positive values indicate lower RMSE than AIFS-SZA. Errors in 6 h accumulated radiation are also examined to diagnose differences in accumulated radiation totals. Results are further stratified by forecast lead time, month, and sky condition.

2.5.5 Ablation and Input Attribution

Temporal-downscaling ablations follow the Level 1 protocol and sequentially test the SZA input, transmittance parameterization, and radiation-loss design. Input ablations for SolarFINE-J remove the DEM, the hourly meteorological states, or both the meteorological states and DEM to assess how different conditioning variables influence spatial downscaling. The complete diffusion configuration is also compared with deterministic subhourly allocation to evaluate the contribution of conditional diffusion to intrahour variability and probabilistic behavior.

Integrated Gradients (IG) is used to analyze the sensitivity of the hourly spatial generator in SolarFINE-J to its input variables. For input $\mathbf{x}$, reference input $\mathbf{x}'$, and scalar output function F, the attribution for input i is defined as

$$\mathrm{IG}_i(\mathbf{x}) = (x_i - x_i')\int_0^1 \frac{\partial F[\mathbf{x}' + \alpha(\mathbf{x}-\mathbf{x}')]}{\partial x_i}\,\mathrm{d}\alpha. \tag{9}$$

The attribution target is the high-resolution residual $R_{q,h}$ produced by the spatial generator and aggregated over a fixed region. Absolute attributions are summarized separately for each input-variable group. The dry and rainy seasons use the same input standardization, and their SZA ranges are matched to reduce the influence of seasonal differences in solar elevation. IG and the input-ablation experiments are used jointly to examine the dependence of the spatial generator on radiation, meteorological states, solar conditions, and terrain. The variable groups and seasonal comparison procedure are described in Appendix B.

Use of generative AI tools. ChatGPT and Codex (OpenAI) were used to assist with English translation, editing and restructuring of author-provided manuscript drafts, and drafting research code. All AI-assisted text and code were reviewed by the authors before use. The authors take responsibility for the final manuscript, code, analyses, and conclusions.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

3 Results

We evaluate SolarFINE at three successive levels: temporal reconstruction, high-resolution spatial and probabilistic fidelity, and end-to-end forecast performance against station observations.

3.1 Temporal Downscaling From 6-h Accumulations to Hourly Irradiance

On the independent ERA5 test set for 2020, SolarFINE-T yielded a daytime hourly RMSE of 40.37 $\mathrm{W\,m^{-2}}$ (95% CI [39.33, 41.36]), compared with 57.60 $\mathrm{W\,m^{-2}}$ for SZA allocation (95% CI [56.50, 58.70]), 54.88 $\mathrm{W\,m^{-2}}$ for ConvLSTM, and 46.20 $\mathrm{W\,m^{-2}}$ for ModAFNO (Figure 3). Relative to SZA allocation, SolarFINE-T reduced RMSE by 17.23 $\mathrm{W\,m^{-2}}$, corresponding to a reduction of 29.9%. This comparison indicates that concurrent meteorological states provide information for hourly reconstruction beyond that available from solar geometry alone.

SolarFINE-T yielded lower RMSE than all three baselines at every hourly position within the 6 h window (Figure 3a). Errors from SZA allocation and ConvLSTM exhibited a pronounced U-shaped pattern, decreasing from approximately 64–68 $\mathrm{W\,m^{-2}}$ at the beginning and end of the window to about 45–50 $\mathrm{W\,m^{-2}}$ near its center. The corresponding RMSE for ModAFNO was approximately 50–51 $\mathrm{W\,m^{-2}}$ at the window boundaries, whereas SolarFINE-T remained within approximately 39–42 $\mathrm{W\,m^{-2}}$ across all six positions. The differences among the methods were therefore largest near the window boundaries, where SolarFINE-T retained its lower error.

The 6 h integrated-radiation residual provides a complementary measure of reconstruction performance (Figure 3b). SZA allocation has zero residual by construction because it directly preserves the input accumulation. The residuals from SolarFINE-T were concentrated near zero and had a narrower distribution than those from ConvLSTM and ModAFNO. Thus, the lower hourly reconstruction error of SolarFINE-T was not accompanied by a marked increase in the 6 h integrated-radiation residual.

RMSE increased with cloud cover for all methods (Figure 3c). SolarFINE-T yielded the lowest error in every cloud-cover bin, although its absolute error under moderate and high cloud cover remained greater than under low cloud cover. Rapid cloud-related changes in irradiance therefore remain a major challenge for hourly reconstruction. When stratified by SZA, SolarFINE-T also maintained lower RMSE across all bins (Figure 3d). Absolute errors were smaller for all methods at large SZA, consistent with the lower irradiance magnitude under these conditions.

SolarFINE-T also reconstructs hourly meteorological fields for SolarFINE-J. Among the 19 accompanying variables (Table B2), cloud fractions, specific humidity, and most wind fields had lower RMSE than linear interpolation, while sea-level pressure and some temperature fields had higher RMSE. ModAFNO had lower RMSE than SolarFINE-T for all 19 meteorological variables. The main temporal reconstruction result here is the improvement in irradiance.

The progressive ablation experiments show that individual design choices affected error magnitude and temporal correlation differently (Table 2). Adding SZA as an input reduced RMSE from 45.05 $\mathrm{W\,m^{-2}}$ for M0 to 42.75 $\mathrm{W\,m^{-2}}$. Introducing the transmittance parameterization did not further reduce RMSE, which increased slightly to 43.46 $\mathrm{W\,m^{-2}}$, whereas the Pearson correlation coefficient increased from 0.924 to 0.961. This component therefore improved agreement in the hourly variations rather than reducing the overall error magnitude. With the hourly irradiance loss included (M3), RMSE and MAE decreased to

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

430 41.18 and 25.52 $\mathrm{W\,m^{-2}}$, respectively. Adding the 6 h integral-consistency soft constraint to M3 gave the complete SolarFINE-T configuration, which yielded both the lowest RMSE, 40.37 $\mathrm{W\,m^{-2}}$, and the highest Pearson correlation coefficient, 0.972.

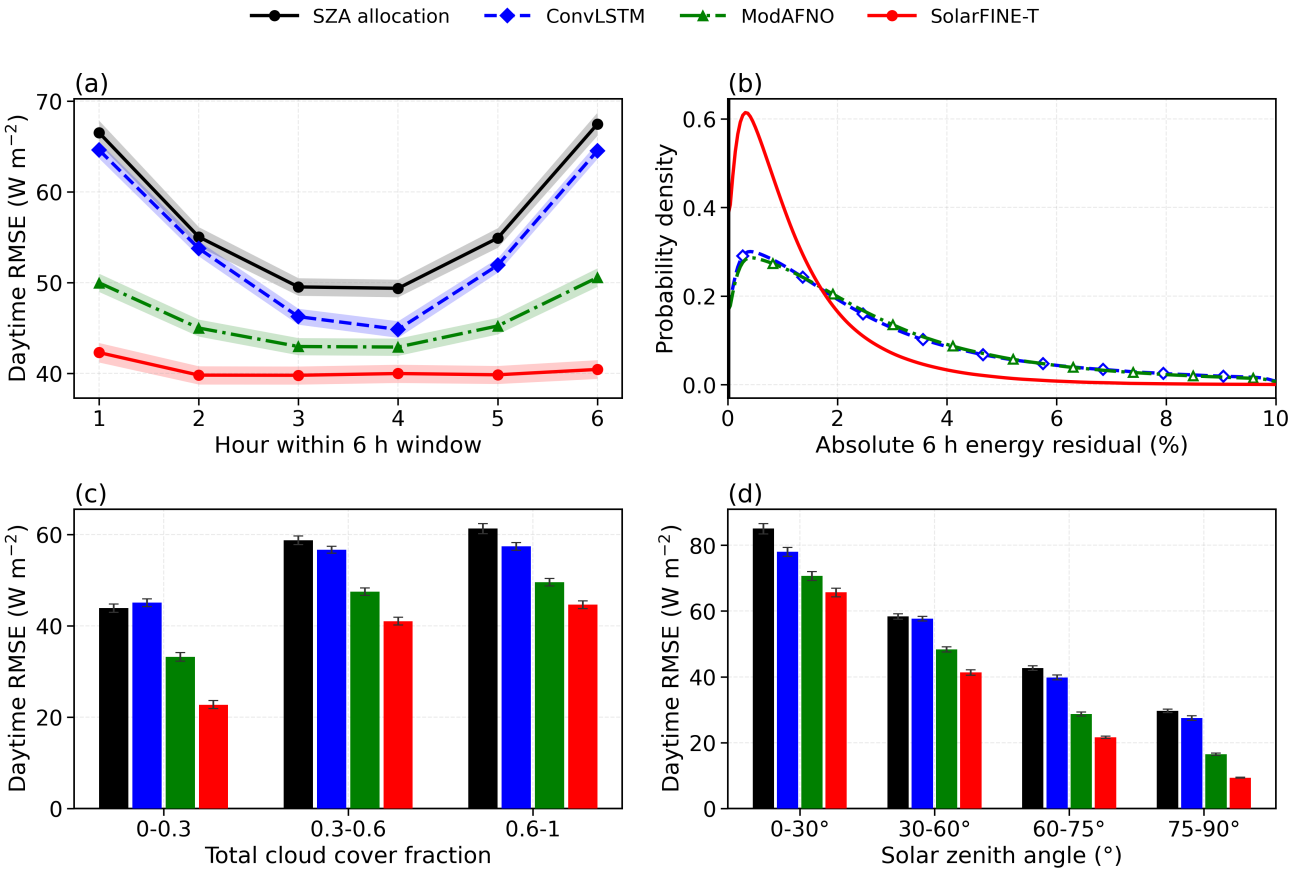


Figure 3. Hourly reconstruction performance of SolarFINE-T. (a) RMSE at each hourly position within the 6 h window. (b) Distribution of the absolute relative residual in 6 h integrated radiation (%); SZA allocation is a point mass at zero. (c) RMSE stratified by cloud cover. (d) RMSE stratified by solar zenith angle. Shading and error bars denote 95% confidence intervals.

3.2 Errors and Spatial Structure of High-Resolution Irradiance

Using SWDR_EAP for 2020 as the reference, we compare SolarFINE-J with bilinear interpolation, DeepSD-Solar, and SwinIR-Solar at 0.05° and 10-min resolution. Each deterministic baseline first produces an hourly high-resolution field, which
435 is then allocated to six 10-min fields using solar-geometry and clear-sky irradiance weights. SolarFINE-J is evaluated using its 20-member ensemble mean. The 10-min RMSE in Table 3 is evaluated frame by frame on the 0.05° grid; hourly RMSE is calculated after taking the arithmetic mean of the six prediction and reference fields separately. Spatial standard-deviation

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table 2. Progressive ablation experiments for SolarFINE-T.

Experiment	Configuration	RMSE	MAE	Pearson r	RMSE reduction
		$\mathrm{W\,m^{-2}}$			relative to SZA (%)
SZA	Solar-geometry allocation	57.60	39.76	0.935	—
M0	Base model	45.05	29.74	0.920	21.8
M1	M0 + solar zenith angle	42.75	27.79	0.924	25.8
M2	M1 + transmittance parameterization	43.46	28.65	0.961	24.5
M3	M2 + hourly irradiance loss	41.18	25.52	0.967	28.5
SolarFINE-T	M3 + integral-consistency loss	40.37	25.18	0.972	29.9

error measures differences in spatial variability from SWDR_EAP, and P99 bias measures the bias in the high-irradiance tail; values closer to zero indicate smaller bias. All metrics in Table 3 are in $\mathrm{W\,m^{-2}}$.

At the native $0.05°$ and 10-min resolution, SolarFINE-J yielded an RMSE of 132.42 $\mathrm{W\,m^{-2}}$, compared with 167.32 $\mathrm{W\,m^{-2}}$ for bilinear interpolation, 153.66 $\mathrm{W\,m^{-2}}$ for DeepSD-Solar, and 136.34 $\mathrm{W\,m^{-2}}$ for SwinIR-Solar (Table 3). This corresponds to an RMSE reduction of 20.9% relative to bilinear interpolation but only 2.9% relative to SwinIR-Solar. SolarFINE-J therefore showed a clear improvement over simple interpolation, whereas its native-resolution pointwise RMSE was close to that of the stronger learning-based super-resolution baseline.

Larger differences emerged in the spatial statistics. SolarFINE-J had a spatial standard-deviation error of 26.68 $\mathrm{W\,m^{-2}}$, lower than those of the other three methods, and its P99 bias of -36.31 $\mathrm{W\,m^{-2}}$ was closest to zero. After the six 10-min fields were aggregated to the hourly scale, its RMSE decreased to 113.23 $\mathrm{W\,m^{-2}}$, remaining the lowest among the four methods. Compared with SwinIR-Solar, the main differences were therefore found in spatial variability, the high-irradiance tail, and error after aggregation, rather than in a large reduction in native-resolution pointwise RMSE.

Figure 4a compares the high-resolution fields under different sky conditions. At high k_c, all methods reproduced the large-scale irradiance gradient, with the main differences occurring in local gradients and the magnitude of spatial variability. The differences among the methods were more apparent at intermediate k_c. Bilinear interpolation produced visibly smoothed fields, whereas DeepSD-Solar and SwinIR-Solar recovered some local variability and SolarFINE-J retained more small-scale fluctuations. At low k_c, the local spatial structure became more fragmented and the differences among the methods were more complex.

These differences are more clearly resolved by the spatial power spectra (Figure 4b). Spectral energy was similar among the methods at longer wavelengths. As wavelength decreased, spectral energy declined more rapidly for the deterministic baselines, whereas SolarFINE-J remained closer to SWDR_EAP. The marginal irradiance distribution produced by SolarFINE-J was also closer to the reference product in the high-value range (Figure 4c), consistent with its smaller negative P99 bias in Table 3. Pointwise error and statistical structure therefore provide complementary measures of model performance.

When further stratified by sky condition and terrain complexity, SolarFINE-J yielded lower RMSE than bilinear interpolation under all nine combinations (Figure 4d). At high k_c, the RMSE reductions were 39.0%, 31.1%, and 24.3% for low,

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

moderate, and high terrain complexity, respectively, decreasing as terrain complexity increased. The corresponding reductions were 21.1%, 17.3%, and 17.8% at intermediate k_c, and 15.5%, 14.1%, and 17.9% at low k_c. A clear monotonic relationship with terrain complexity was therefore evident only at high k_c.

Table 3. Performance comparison between SolarFINE-J and high-resolution downscaling baselines.

Method	10-min RMSE	Hourly RMSE	Spatial-SD error	P99 bias	RMSE 95% CI
Bilinear-Solar	167.32	153.26	41.64	-78.65	163.71–170.90
DeepSD-Solar	153.66	138.37	54.32	-92.47	150.60–156.88
SwinIR-Solar	136.34	119.02	35.69	-65.44	133.11–139.47
SolarFINE-J	132.42	113.23	26.68	-36.31	129.31–135.51

The effects of temporal and spatial aggregation are summarized in Figure 5. Relative to bilinear interpolation, the RMSE reduction increased from 20.9% at the 10-min scale (95% CI [19.8, 22.0]) to 26.1% at the hourly scale (95% CI [24.9, 27.3]). It also increased from 20.9% at 0.05° resolution (95% CI [19.8, 21.9]) to 30.5% at 0.25° resolution (95% CI [29.1, 31.9]). Relative to SwinIR-Solar, temporal aggregation increased the RMSE reduction from 2.9% (95% CI [2.2, 3.5]) to 4.9% (95% CI [4.0, 5.7]), while spatial aggregation increased it to 6.5% (95% CI [5.4, 7.5]).

The relative improvement from SolarFINE-J persisted after aggregation and became larger at coarser temporal and spatial scales. Figure 5 shows that the RMSE difference between models depends on the evaluation scale.

3.3 Subhourly Generation and Ensemble Calibration

To distinguish errors in the shared hourly background from those in the generated intrahour evolution, we use an exact orthogonal decomposition of the 10-min ensemble-mean MSE into additive hourly-mean-error and intrahour-anomaly-error components (Figure 6a).

The hourly-mean component accounted for most of the total MSE at low and high k_c (Figure 6a). In the $k_c = 0.6$–0.8 bin, the intrahour-anomaly error reached 12,340 $\mathrm{W^2\,m^{-4}}$, accounting for approximately 47% of the total MSE. Intrahour variability therefore became an important component of the total error under these conditions. Across all k_c bins, the diffusion ensemble yielded lower CRPS than deterministic subhourly allocation for both the full irradiance field and the intrahour anomalies (Figure 6b). Conditional generation thus improved the probabilistic score of the subhourly distribution for a given hourly background.

The reduction in CRPS was not accompanied by fully matched ensemble spread. The full-field spread–skill points were predominantly underdispersive (Figure 6c), with a spread–skill ratio of 0.60 for the 20-member ensemble (Figure B1). After removing the hourly mean, the spread–skill ratio for the intrahour anomalies was 1.16. As the ensemble size increased, the full-field ratio changed little, whereas the anomaly ratio continued to increase.

This difference is consistent with the generation strategy of SolarFINE-J. All diffusion members share the same hourly spatial background, while stochastic variation is confined to the six 10-min fields within each hour. The standard deviation

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

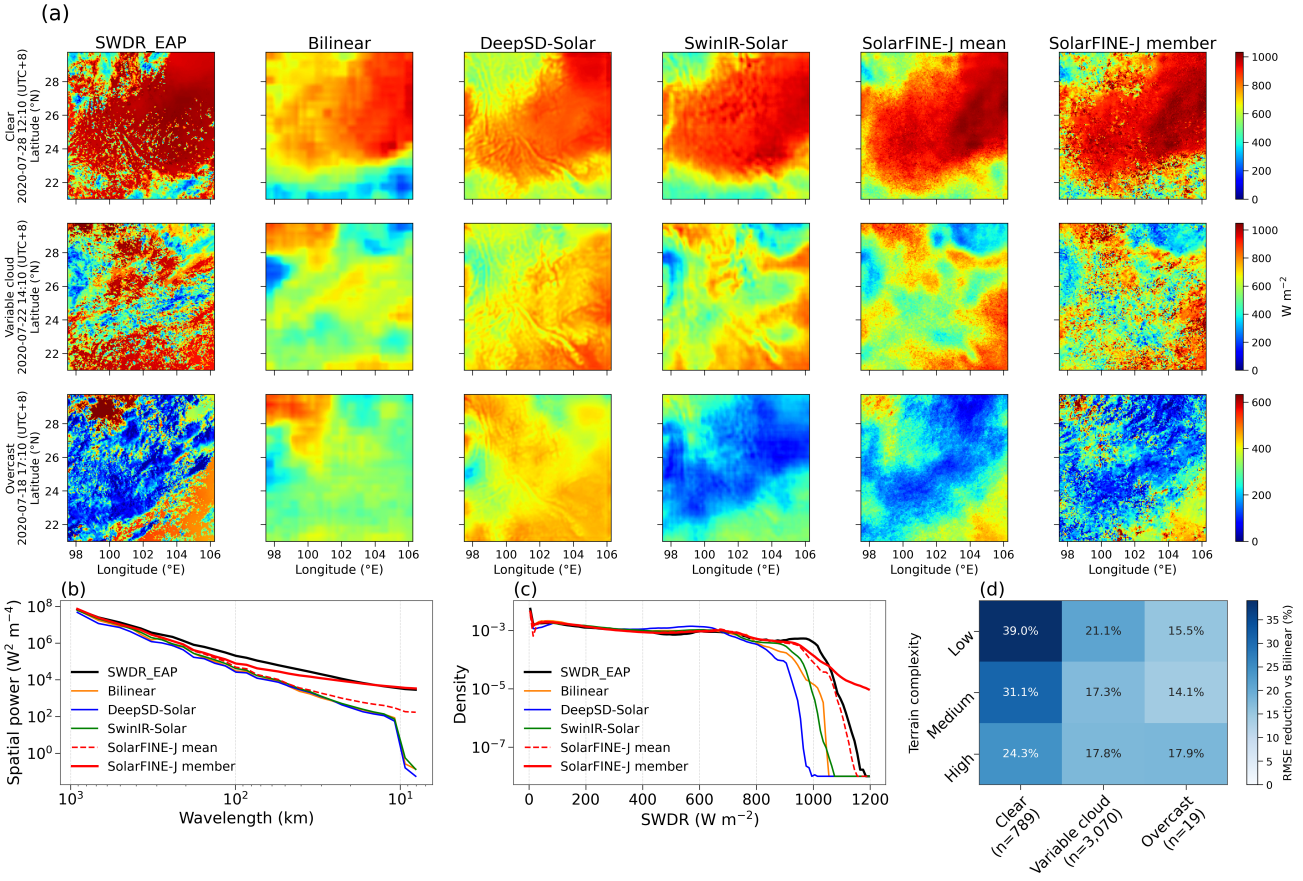


Figure 4. High-resolution downscaling performance of SolarFINE-J. (a) High-resolution irradiance fields under different clear-sky-index conditions; the SolarFINE-J columns show the ensemble mean and one representative member. High, intermediate, and low clear-sky-index conditions use $k_c > 0.8$, $0.3 \leq k_c \leq 0.8$, and $k_c < 0.3$, respectively. (b) Radially averaged spatial power spectra. (c) Probability density distributions of irradiance. (d) RMSE reductions relative to bilinear interpolation across sky conditions and levels of terrain complexity. SWDR_EAP serves as the high-resolution reference product. The horizontal axis in (d) reports the number of windows in each sky-condition group.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

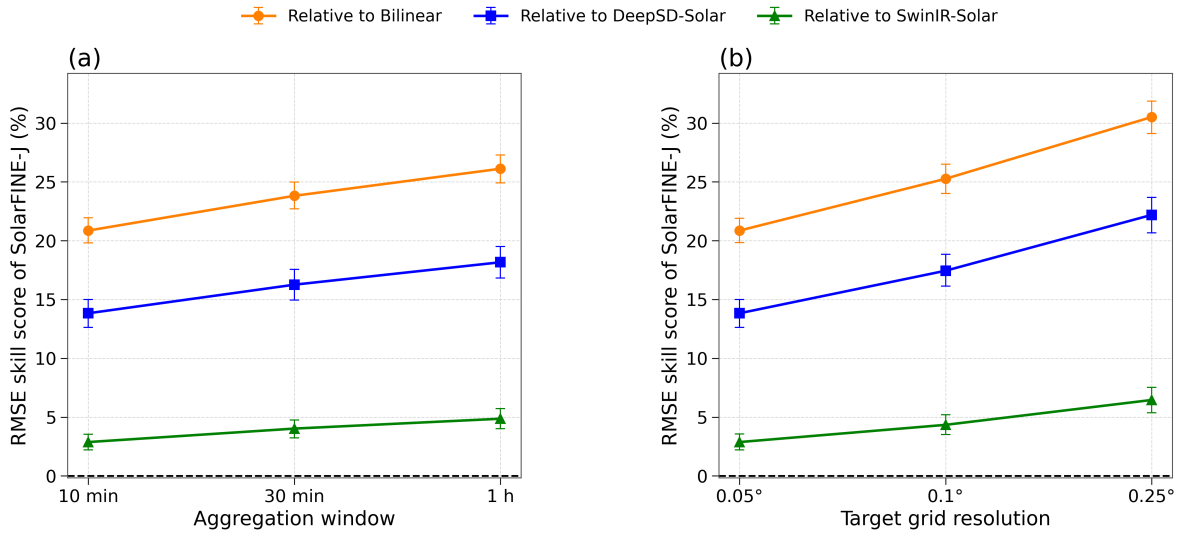


Figure 5. Scale dependence of SolarFINE-J downscaling performance. RMSE reductions from SolarFINE-J relative to bilinear interpolation, DeepSD-Solar, and SwinIR-Solar across (a) temporal and (b) spatial aggregation scales. Error bars denote 95% confidence intervals.

among the hourly means of the ensemble members was only 5.1×10^{-5} $\mathrm{W\,m^{-2}}$, indicating that the current ensemble contains almost no uncertainty in the hourly background. Increasing the number of members therefore cannot substantially correct the underdispersion of the full field. Conversely, the anomaly spread–skill ratio above 1 indicates that ensemble spread exceeds the anomaly RMSE of the ensemble mean, corresponding to overdispersion in this diagnostic.

Metrics were calculated after aggregating the six 10-min outputs within each hour. Hourly RMSE, spatial-SD error, and P99 bias are in $\mathrm{W\,m^{-2}}$. Increases in RMSE are expressed as percentages relative to the complete configuration. Spatial standard-deviation error and P99 bias describe differences from SWDR_EAP in hourly spatial variability and the high-irradiance tail, respectively.

The input ablations for the hourly spatial generator are summarized in Table 4. Relative to the complete configuration, which had an hourly RMSE of 113.23 $\mathrm{W\,m^{-2}}$, removing the hourly meteorological states increased RMSE by 9.0% to 123.42 $\mathrm{W\,m^{-2}}$. Using only radiation and solar conditions increased RMSE by 8.7% to 123.12 $\mathrm{W\,m^{-2}}$, while removing the DEM increased it by 5.9% to 119.89 $\mathrm{W\,m^{-2}}$. The largest RMSE increase therefore occurred when the hourly meteorological states were removed. The DEM also affected the spatial-variability error and the high-irradiance tail.

IG was evaluated for the hourly spatial generator's high-resolution residual $R_{q,h}$, which represents the spatial correction to the coarse irradiance background (Section 2.5.5). For this residual output, meteorological variables accounted for attribution shares of 0.783 in the dry season and 0.839 in the rainy season (Figure B2), substantially larger than the shares from coarse-resolution radiation, solar geometry, and the DEM. The attribution share from coarse-resolution radiation decreased from 0.117 in the dry season to 0.084 in the rainy season, while that from solar geometry decreased from 0.093 to 0.072. Several middle-

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

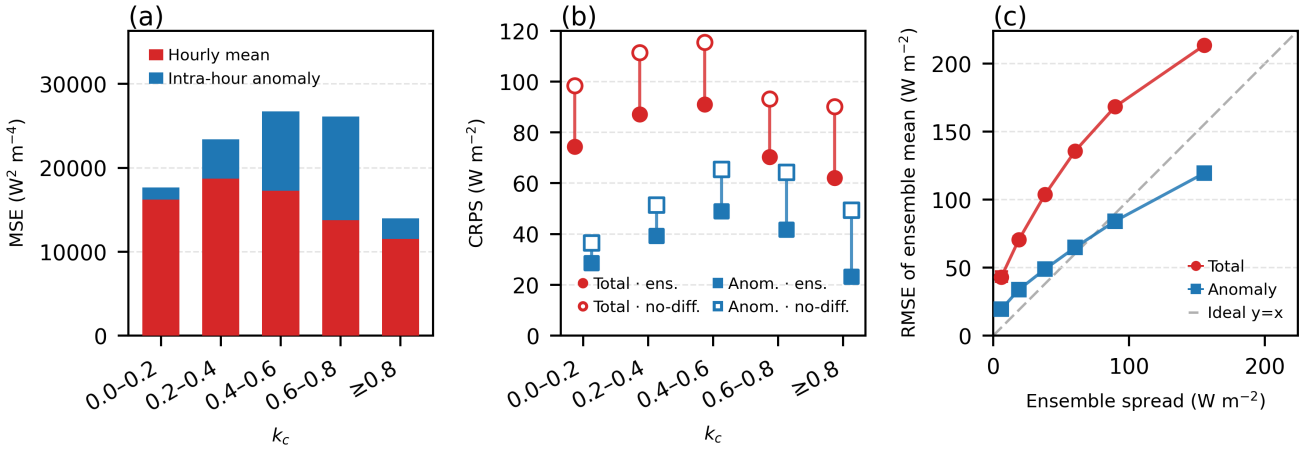


Figure 6. Error decomposition and ensemble calibration of SolarFINE-J. (a) Hourly-mean and intrahour-anomaly components of the total MSE. (b) CRPS of the diffusion ensemble and deterministic subhourly allocation. (c) Relationship between ensemble spread and the RMSE of the ensemble mean. In (b), filled and open symbols denote the diffusion ensemble and deterministic output, respectively, and connecting lines join the two CRPS values. The dashed line in (c) indicates ideal spread–skill agreement.

and upper-level meteorological variables received greater attribution during the rainy season: the shares increased from 0.045 to 0.105 for q500, from 0.043 to 0.076 for q850, and from 0.029 to 0.066 for t500.

The DEM had an IG attribution share below 0.01 in both seasons, while removing it increased hourly RMSE by 5.9%. These diagnostics quantify different aspects of the spatial generator. IG integrates gradients along the path from the reference input to the evaluated input and attributes the selected residual output to the input variables. The input ablation measures the change in hourly irradiance error when DEM information is removed. The results therefore show that the DEM contributes to hourly reconstruction performance despite its small share of the residual-output attribution.

Table 4. Input ablation experiments for SolarFINE-J.

Experiment	Configuration	Hourly RMSE	RMSE increase	Spatial-SD error	P99 bias
Complete	Radiation and solar + meteorology + DEM	113.23	—	26.68	-36.31
No DEM	DEM removed	119.89	+5.9	32.71	-56.43
No meteorology	Hourly meteorological states removed	123.42	+9.0	28.48	-43.58
Radiation and solar	Meteorology and DEM removed	123.12	+8.7	36.51	-52.69

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

3.4 End-to-End Station Evaluation Driven by AIFS Forecasts

End-to-end performance under forecast inputs was evaluated by driving the complete SolarFINE system with AIFS forecasts for 2024 and comparing the resulting hourly GHI with station observations (Figure 7). AIFS-SZA served as the primary physical baseline, while SolarFINE-noTune and SolarFINE-C were used to diagnose sensitivity to forecast-input adaptation and the treatment of accumulated radiation, respectively.

The hourly GHI RMSE of SolarFINE increased with forecast lead time, from 136.7 $\mathrm{W\,m^{-2}}$ at D1 to 198.9 $\mathrm{W\,m^{-2}}$ at D15 (Figure 7a). Over the same period, its SS relative to AIFS-SZA increased from 0.084 to 0.174, with the 95% confidence intervals remaining above zero at all 15 lead days (Figure 7b). The increase in SS should be distinguished from the increase in absolute error. The D15 forecast remained less accurate in absolute terms than the D1 forecast; the higher SS indicates that its RMSE increased more slowly with lead time than that of AIFS-SZA.

SolarFINE-noTune behaved differently from the complete system. Its SS decreased from 0.047 at D1 to 0.009 at D2 and -0.011 at D3, remained below zero at most subsequent lead times, and reached -0.061 at D15. The complete SolarFINE system maintained positive SS throughout, with the difference between the two configurations becoming larger at medium and long lead times. These results indicate that the mapping learned from ERA5 did not transfer stably when applied directly to AIFS inputs and that forecast-input adaptation was an important component of the end-to-end system.

SolarFINE-C retains the relative distribution across the six hourly irradiance fields predicted by the adapted SolarFINE-T, but multiplies all six fields by a common factor so that their integrated total equals the original AIFS 6 h accumulation. Under this constraint, SS was below zero on most lead days, including a value of -0.015 at D15. Its lower performance relative to the complete system indicates that the station-level results are sensitive to how the 6 h accumulated-radiation total is treated.

The relative improvement varied with sky condition (Figure 7c). The most consistent improvement occurred at intermediate k_c, where SS remained positive on all 15 lead days and increased from 0.151 at D1 to 0.261 at D15. At low k_c, SS was also predominantly positive after D3 and reached 0.176 at D15. In contrast, SS was negative at most lead times from D2 to D10 under high-k_c conditions and approached zero only at longer lead times. Overall, the improvement was largest and most consistent at intermediate k_c, smaller at low k_c, and not consistent at high k_c.

SS also varied among stations (Figure 7d). At D15, station-level SS ranged from 0.138 to 0.221 and was positive at all five stations. The ordering among stations was not consistent across lead times; we therefore focus on the shared lead-time and sky-condition dependence without attributing the differences at individual stations.

Representative days provide a more direct comparison of the forecast time series under different sky conditions (Figure 8). Under clear skies, the observations and both forecasts exhibited smooth, bell-shaped diurnal cycles, with the main differences occurring in irradiance amplitude. Hourly variability increased under cloudy and overcast conditions. AIFS-SZA remained comparatively smooth, whereas SolarFINE followed some of the observed hourly fluctuations more closely. There were also representative days on which SolarFINE had larger errors. The diffusion members additionally show 10-min variability around the shared hourly background. Because the station observations are hourly, these subhourly trajectories illustrate the generated structure rather than independently validated 10-min variability.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

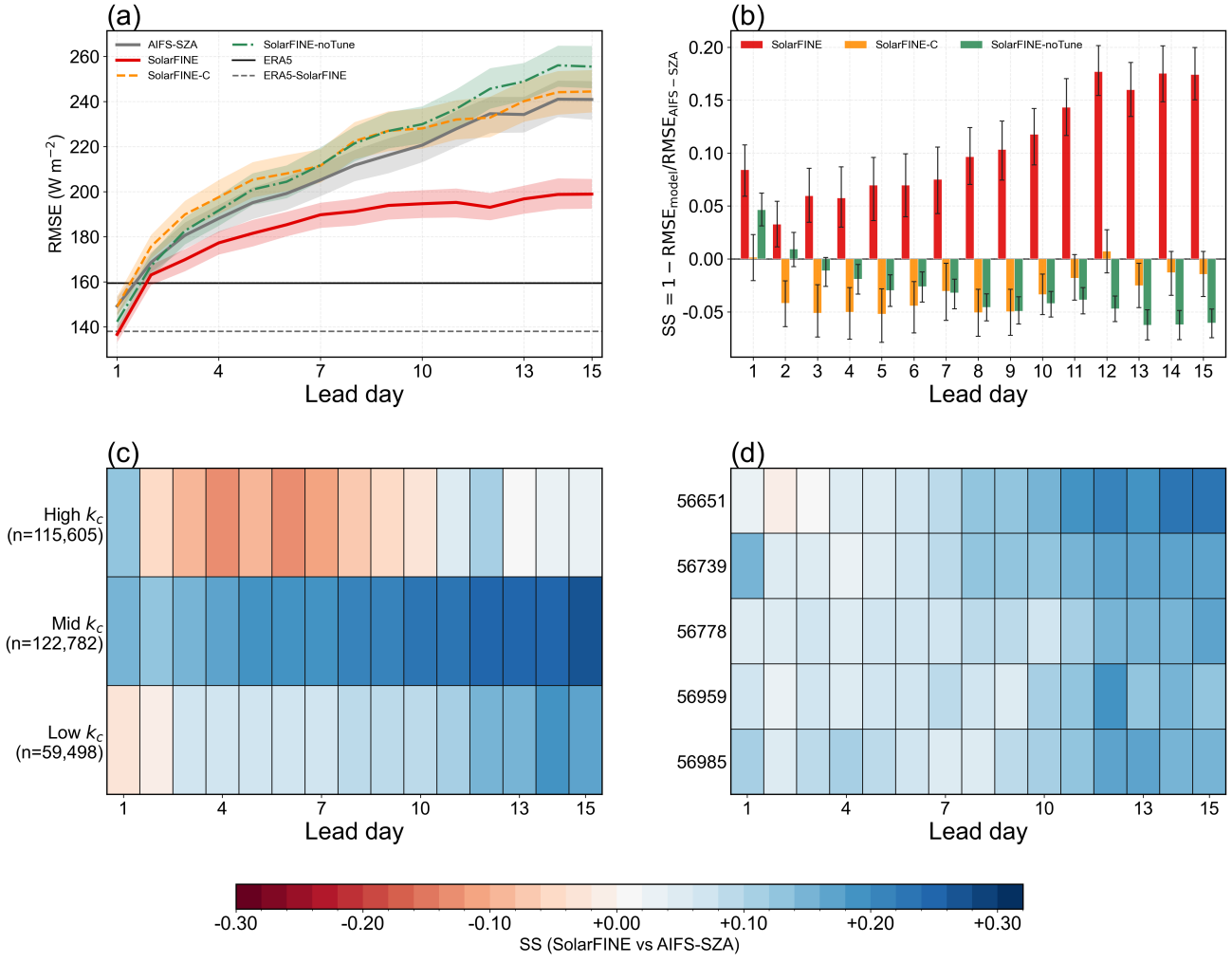


Figure 7. Station-level GHI forecast performance under AIFS inputs. (a) RMSE of SolarFINE and AIFS-SZA. (b) RMSE skill score of the SolarFINE configurations relative to AIFS-SZA. (c) Skill score stratified by clear-sky index; rows show high, intermediate, and low k_c, respectively. (d) Skill score at individual stations. Shading in (a) and error bars in (b) denote 95% confidence intervals; colors in (c) and (d) represent skill scores.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

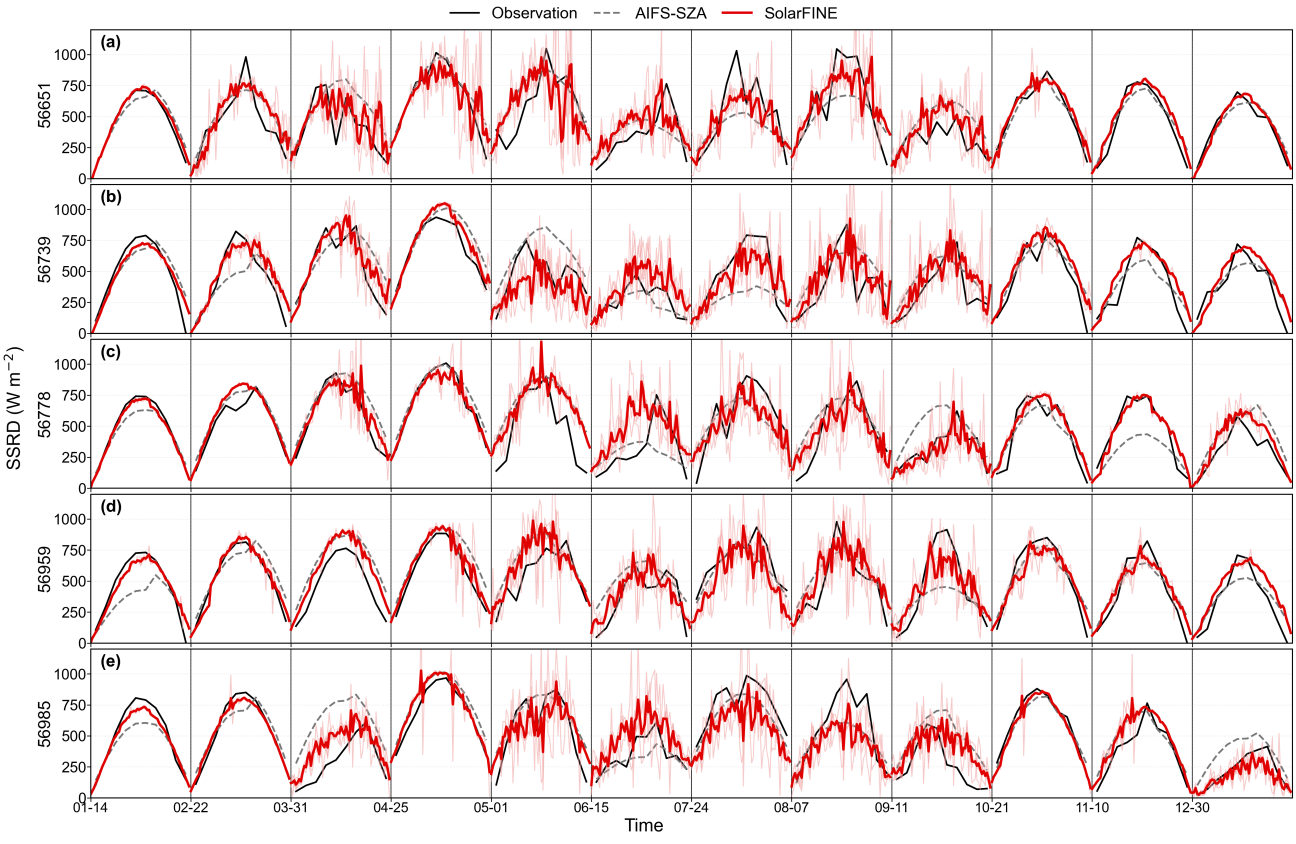


Figure 8. Representative station-level GHI forecasts under different sky conditions. Black lines show hourly observations, dashed gray lines show AIFS-SZA, light-colored lines show the 10-min SolarFINE ensemble members, and dark-colored lines show the ensemble mean. The panels show representative examples from different stations and months to illustrate the hourly forecast variations.

Figure 9 compares the joint distributions of observed and forecast GHI across lead times. At D1, station-level correlations for SolarFINE ranged from 0.842 to 0.869, with RMSE ranging from 130.0 to 150.0 $\mathrm{W\,m^{-2}}$. The corresponding ranges for AIFS-SZA were 0.808–0.850 and 138.4–175.5 $\mathrm{W\,m^{-2}}$, respectively. The scatter distributions broadened for both methods as lead time increased. By D15, correlations for SolarFINE had decreased to 0.682–0.769 and RMSE had increased to 191.6–211.4 $\mathrm{W\,m^{-2}}$, compared with 0.585–0.725 and 222.2–262.6 $\mathrm{W\,m^{-2}}$ for AIFS-SZA.

The station-specific RMSE difference between AIFS-SZA and SolarFINE widened from approximately 4–25 $\mathrm{W\,m^{-2}}$ at D1 to 31–58 $\mathrm{W\,m^{-2}}$ at D15, consistent with the increase in SS shown in Figure 7. Forecast-to-observed variance ratios were generally close to 1 for both methods. At D1, the ratios were 0.85–0.89 for SolarFINE and 0.89–0.95 for AIFS-SZA. At longer lead times, the SolarFINE ratios decreased to approximately 0.83–0.89, whereas those for AIFS-SZA remained within approximately 0.90–1.01.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

Lower RMSE did not imply improvement in every statistic. At some stations and lead times, the correlation from SolarFINE was comparable to, or slightly lower than, that from AIFS-SZA, and the AIFS-SZA variance ratio was sometimes closer to 1. These metrics describe different aspects of forecast performance. In the present station evaluation, the most consistent difference from SolarFINE was a reduction in error magnitude at individual valid times; improvements in correlation and variance amplitude did not occur under every condition.

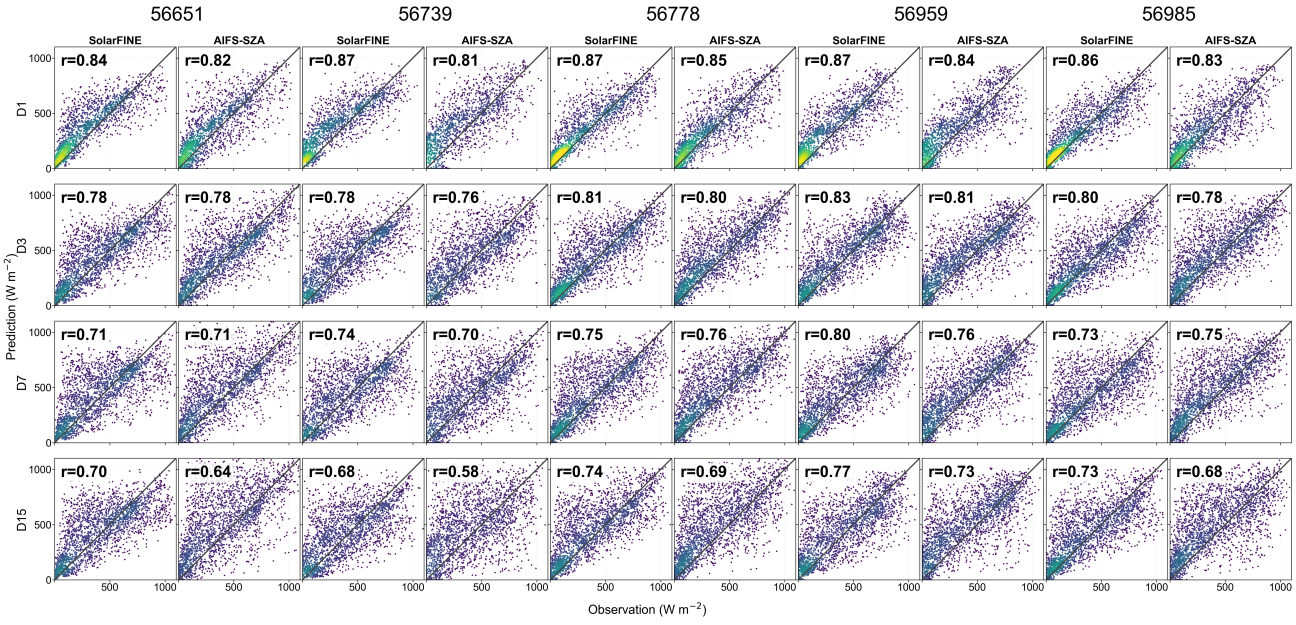


Figure 9. Joint distributions of observed and forecast station-level GHI across lead times. Rows indicate forecast lead times and columns indicate stations, with SolarFINE and AIFS-SZA shown separately. The horizontal axis gives observed GHI, the vertical axis gives forecast GHI, colors indicate sample density, and gray lines show the 1:1 relationship. Pearson correlation coefficients are reported within each panel.

4 Discussion and Conclusions

4.1 Hourly Information Recoverable From Accumulated Radiation

A 6 h radiation accumulation preserves the total energy within a time window but compresses information about its hourly allocation. Solar geometry constrains the broad shape of the irradiance diurnal cycle, but it cannot describe hourly modulation associated with changes in clouds, water vapor, and large-scale circulation. The lower hourly reconstruction error of SolarFINE-T relative to SZA-based allocation indicates that contemporaneous meteorological states provide useful information for recovering the temporal evolution within the window. The difference was most evident near the beginning and end of each 6 h window, where errors from SZA-based allocation and ConvLSTM increased while those from SolarFINE-T var-

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

ied less with hourly position. This result shows the value of combining target-hour solar geometry and atmospheric states to estimate the evolution within a 6 h window under an integral-consistency constraint.

The sequential ablations further show that the physical designs affected different aspects of reconstruction performance. Adding solar zenith angle reduced RMSE, whereas the transmittance parameterization primarily improved temporal correlation without producing a further reduction in RMSE. Within this sequence of configurations, the complete model yielded both the lowest RMSE and the highest correlation, indicating that the combined design improved both error magnitude and the consistency of the reconstructed hourly evolution.

Representing the hourly state as an explicit intermediate layer allows the framework to make fuller use of the hourly meteorological fields available from ERA5. ERA5 provides not only hourly irradiance but also cloud, water-vapor, temperature, and circulation variables that characterize the weather conditions influencing the spatial distribution of surface irradiance. SolarFINE-T therefore reconstructs both hourly irradiance and the corresponding meteorological states and passes them jointly to SolarFINE-J. Using high-resolution radiation from SWDR_EAP as supervision, SolarFINE-J then learns how local irradiance structure varies with the hourly weather state. The intermediate representation thus allows the hourly meteorological information from ERA5 to constrain both the temporal reconstruction of accumulated radiation and the subsequent spatial downscaling.

Generating 10-min high-resolution irradiance directly from a 6 h accumulation would require a model to address information compressed by temporal accumulation, local structure lost through spatial coarsening, and unresolved variability within each hour at the same time. SolarFINE separates these transformations so that each stage can use information and constraints matched to its target scale, while allowing temporal reconstruction error to be evaluated before spatial generation. For other accumulated meteorological variables, these results suggest that information losses from temporal accumulation and spatial coarsening should be considered separately: contemporaneous atmospheric states may help recover the temporal allocation, whereas local spatial structure may still need to be learned from high-resolution reference data.

4.2 Fine-Scale Structure, Pointwise Error, and Ensemble Uncertainty

Pointwise error and spatial structure describe different properties of a high-resolution irradiance field. At the native $0.05°$, 10-min resolution, SolarFINE-J reduced RMSE by 20.9% relative to bilinear interpolation and by 2.9% relative to SwinIR-Solar, while its spatial power spectrum, spatial standard deviation, and high-irradiance tail were closer to SWDR_EAP.

At the native resolution, RMSE compares the prediction and reference at every grid point and valid time. Temporal or spatial aggregation compares them at a coarser evaluation scale. After temporal aggregation to 1 h, the RMSE reductions relative to bilinear interpolation and SwinIR-Solar increased to 26.1% and 4.9%, respectively; after spatial aggregation to $0.25°$, they increased to 30.5% and 6.5%. Pointwise error, probabilistic scores, and spatial spectra together describe the generated fields across scales (Harris et al., 2022).

All SolarFINE-J ensemble members share the same hourly spatial background. Differences among the members arise only from how the six 10-min fields vary within that hour. The current diffusion ensemble therefore represents uncertainty in

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

intrahour variability conditional on a given hourly background, but it does not represent uncertainty in the hourly background itself.

For the complete irradiance field, RMSE includes errors in both the hourly background and the intrahour variability, whereas the ensemble spread arises mainly from the latter. Because errors in the shared hourly background have no corresponding ensemble spread, the full field is underdispersive. After the hourly mean is removed, the evaluation isolates the intrahour anomalies. Their spread–skill ratio is slightly above 1, indicating that anomaly ensemble spread exceeds the ensemble-mean error. Spread across members describes ensemble uncertainty, whereas temporal variance or irradiance increments describe fluctuations within a member. Although the diffusion ensemble yielded lower CRPS than the deterministic subhourly output, the hourly background and intrahour variability remain distinct calibration targets. Increasing the number of diffusion members can sample the conditional distribution around that background more completely, but it cannot represent errors shared by all members.

In the input ablations, removing the hourly meteorological states produced the largest increase in hourly RMSE. IG assigned the largest attribution share to meteorological variables for the hourly spatial generator's high-resolution residual. Together, these diagnostics link meteorological inputs to both the learned spatial correction and hourly reconstruction performance. The DEM had a low IG share, while removing it increased hourly RMSE by 5.9%. IG attributes the selected residual output along the path from the reference input to the evaluated input; ablation measures the change in hourly irradiance error when an information source is removed. The two diagnostics therefore describe complementary aspects of the spatial generator.

4.3 From Controlled Downscaling to End-to-End Performance Under AIFS Forecast Inputs

The controlled ERA5 experiments and the AIFS-driven evaluation address different questions. The former tests which temporal and spatial information can be recovered when the inputs and reference data are comparatively well matched. The latter tests whether this capability produces forecast improvements against independent observations when upstream forecast errors and input-distribution differences are present. Strong reconstruction and downscaling performance under controlled conditions therefore does not by itself establish improvement under forecast inputs.

In the end-to-end station evaluation, SolarFINE produced its largest and most consistent improvement at intermediate k_c. At low k_c, SS was generally positive from D3 onward. At high k_c, SS was negative on most lead days from D2 to D10 and approached zero at longer leads, indicating no consistent improvement over AIFS-SZA. One possible explanation is that, at high k_c, where irradiance is close to its clear-sky estimate, solar geometry already describes much of the diurnal variation, leaving less scope for additional improvement. At low k_c, strong irradiance attenuation and errors in its forecast may limit downstream performance. At intermediate k_c, variations beyond solar geometry and useful conditions retained in the upstream meteorological fields may together offer greater scope for improvement.

Although ERA5 and AIFS use the same variables and spatial grid, their cloud, water-vapor, and radiation fields may differ in their joint distributions and forecast-error characteristics. SolarFINE-noTune lost relative skill at medium and long lead times, whereas the adapted system retained positive SS across all lead days. This comparison shows the value of adapting the ERA5-trained temporal stage to AIFS inputs.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

SolarFINE-C applies a common scaling factor to the six hourly irradiance fields so that their integrated total matches the original AIFS 6 h accumulation. Its skill score was below zero at most lead times. Relative to SolarFINE, this comparison shows that the treatment of the accumulated total affects the final hourly station forecasts. The adapted temporal stage jointly adjusts the radiation total and its allocation among the six hours, and the spatial stage subsequently maps these fields to the high-resolution background. The station-level gains reflect this combined processing of the forecast inputs.

4.4 Limitations and Future Work

The current end-to-end evaluation uses observations from five stations in Yunnan during 2024. The resulting forecast improvements are therefore established for one region and one independent forecast year. Differences in cloud regimes, solar geometry, terrain, and surface conditions may alter both the information available to the model and the propagation of forecast errors. Evaluation across additional regions and consecutive forecast years is needed to determine the stability of the end-to-end performance across climatic and interannual conditions.

The current high-resolution evaluation against SWDR_EAP characterizes pointwise errors, spatial spectra, irradiance distributions, and conditional probabilistic scores at 0.05° and 10-min resolution. Independent ground observations evaluate the hourly forecasts. Temporal diagnostics of irradiance increments, autocorrelation, and transitions between successive hours would extend the evaluation of the generated sequences. Minute-scale ground observations would provide an independent test of the 10-min trajectories in Figure 8, cloud-shadow passages, and rapid irradiance ramps.

A probabilistic hourly stage could add uncertainty in the shared background while retaining hourly-mean consistency within each member.

SolarFINE-noTune and SolarFINE-C diagnose the effects of forecast-input adaptation and the treatment of 6 h accumulated radiation within the current forecast chain. The observed station-level gains combine accumulated-radiation adjustment, hourly reconstruction, and spatial refinement. These comparisons establish the performance of the combined system relative to AIFS-SZA. Separate contributions from radiation-bias correction, hourly allocation, and spatial refinement remain a question for component-controlled comparisons.

This study evaluates surface solar irradiance. Photovoltaic operation also depends on module conversion, temperature response, and power-system characteristics, so improvements in irradiance metrics may not transfer proportionally to photovoltaic power, ramp events, or scheduling errors. Coupling SolarFINE output to an independent photovoltaic power model and operational verification framework would provide a direct evaluation for downstream solar-energy forecasting.

4.5 Conclusions

SolarFINE converts coarse-grid 6 h accumulated solar radiation from MLWP output into hourly states and 0.05°, 10-min irradiance fields. The framework treats temporal reconstruction, spatial downscaling, and subhourly generation as three connected scale transformations and applies information and constraints matched to the target scale of each stage.

The controlled ERA5 experiments show that combining accumulated radiation with concurrent meteorological states recovers hourly variability more accurately than allocation based on solar geometry alone. SolarFINE-T reduced daytime hourly

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

RMSE by 29.9% relative to SZA-based allocation, indicating that hourly information on clouds, water vapor, temperature, and circulation provides useful constraints on the temporal allocation of accumulated radiation.

In the high-resolution experiments, SolarFINE-J reduced RMSE at the native $0.05°$, 10-min resolution by 20.9% relative to bilinear interpolation and by 2.9% relative to SwinIR-Solar, while more closely reproducing the spatial power spectrum, spatial variability, and high-irradiance tail. The diffusion ensemble further improved conditional CRPS and describes intrahour uncertainty given the hourly background.

In the AIFS-driven station evaluation, the hourly spatial backgrounds produced by SolarFINE maintained positive skill relative to AIFS-SZA from D1 to D15. Forecast-input adaptation improved station performance relative to SolarFINE-noTune, whereas rescaling the hourly fields to match the original AIFS 6 h accumulation reduced skill. Improvements were largest and most consistent at intermediate k_c. Together with the controlled reconstruction and high-resolution evaluations, these results establish the framework's performance across hourly reconstruction, conditional intrahour generation, and hourly forecasting under AIFS inputs.

Code and data availability. ERA5 reanalysis data are publicly available from the Copernicus Climate Data Store. The AIFS Single v1.0 model checkpoint used to generate the retrospective forecasts is publicly available from ECMWF through Hugging Face (ECMWF, 2025), DOI: https://doi.org/10.57967/hf/4629. The high-resolution radiation data are from the East Asia–Pacific surface downward longwave and shortwave radiation dataset distributed by the National Tibetan Plateau Data Center (Husi et al., 2021), DOI: https://doi.org/10.11888/Meteoro.tpdc.271729. SRTM elevation data are publicly available from the NOAA National Centers for Environmental Information (NCEI). The SolarFINE v0.1.0 source code and trained checkpoints are publicly available from Zenodo under the MIT License (Dai et al., 2026), all-versions DOI: https://doi.org/10.5281/zenodo.23009436. Hourly surface solar radiation observations for 2024 were provided by the Yunnan Provincial Meteorological Bureau under restricted access. They are not publicly available; requests for access should be directed to the Bureau and require its approval. The authors cannot redistribute these observations.

Appendix A: Model Architecture and Implementation Details

A1 SolarFINE-T Temporal Reconstruction Network

SolarFINE-T reconstructs hourly evolution within each 6 h window under the accumulated-radiation constraint and supplies concurrent meteorological states to the downstream spatial model. It operates on a 36×36 coarse grid. The inputs are divided into 2×2 patches and embedded as 128-dimensional tokens, which are processed by four ModAFNO blocks containing four spectral subblocks each. The target-hour position and forecast lead time form the FiLM condition. The temporal modulation has 64 dimensions, and the condition vector is encoded using a 64-dimensional sinusoidal embedding followed by a two-layer multilayer perceptron. A shared decoder produces coarse-grid irradiance and concurrent meteorological states for each of the six target hours. The meteorological input and output variables are identical, standardized using the training-set statistics, and optimized with an MSE loss. All reconstructed meteorological channels are passed to SolarFINE-J. The irradiance branch

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

uses the transmittance parameterization described in Section 2.3.1, and the 6 h integral consistency is imposed through the corresponding loss term.

The retrospective AIFS forecasts were generated with the AIFS Single v1.0 checkpoint released by ECMWF through Hugging Face and the anemoi-inference software on two NVIDIA RTX 4090 GPUs. Each retrospective forecast was initialized from the corresponding ERA5 atmospheric state.

A2 SolarFINE-J Hourly Spatial Stage

The hourly spatial stage recovers local spatial differences while retaining the coarse-grid radiation background. It takes the hourly coarse-grid irradiance and meteorological states reconstructed by SolarFINE-T as dynamic inputs and combines them with solar geometry, clear-sky irradiance, and DEM to produce an hourly high-resolution spatial background. The spatial network is a ResUNet with encoder widths of 64, 128, and 256 channels and an eight-head global-attention block at the 9×9 bottleneck. Decoder features are interpolated to the 176×176 target grid, fused with three auxiliary channels, and processed by four 64-channel residual blocks. The network predicts a spatial correction to the bilinearly interpolated coarse-grid irradiance field; their sum gives the hourly spatial background in Equation (4).

A3 SolarFINE-J Conditional Subhourly Generation Stage

The subhourly stage represents intrahour variability that remains unresolved after conditioning on the hourly weather and radiation background. It generates six consecutive 10-min irradiance fields within each hour. Additional conditions include the hourly meteorological states reconstructed by SolarFINE-T, the clear-sky index calculated from the hourly spatial background and theoretical clear-sky GHI, solar geometry and clear-sky irradiance at the six target times, and DEM. The denoiser is a spatiotemporal U-Net with a base width of 64, three channel levels of 64, 128, and 256, and a dropout rate of 0.1. Two-dimensional residual blocks process the spatial features, and temporal convolutions represent relationships among the six frames. Diffusion training uses a 1000-step cosine noise schedule and a v-prediction objective. Inference uses the DPM-Solver++(2M) SDE sampler with 20 sampling steps. Samples below the clear-sky irradiance threshold skip stochastic sampling and retain the nonnegative hourly background in all six frames; the output is zero where that background is zero. The generated frames are transformed into nonnegative normalized weights so that their hourly arithmetic mean equals the hourly spatial background exactly, as defined in Equation (6).

Appendix B: Training and Diagnostic Details

B1 Data Splits and Standardization

SolarFINE-T and SolarFINE-J use the same temporal split for training, validation, and controlled testing. Data from 2016–2018 and days 1–15 of each month in 2019 are used for training, the remaining days in 2019 are used for validation, and 2020 is reserved for independent testing. The irradiance output of SolarFINE-T is standardized using a training-set hourly mean

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

of 191.67 $\mathrm{W\,m^{-2}}$ and a standard deviation of 270.0 $\mathrm{W\,m^{-2}}$. SolarFINE-J takes the hourly states generated by the frozen SolarFINE-T as dynamic inputs. Its hourly spatial and subhourly diffusion stages are trained sequentially. The paired 2023 AIFS–ERA5 samples are used only for forecast-input adaptation of SolarFINE-T; SolarFINE-J remains frozen during this stage.

B2 Optimization

All three training stages use AdamW with cosine-annealing learning-rate schedules. SolarFINE-T, the hourly spatial stage of SolarFINE-J, and the subhourly diffusion stage are trained separately. The spatial-stage parameters are frozen before diffusion training. The learning rate, batch size, and number of epochs are listed in Table B1. Mixed-precision and run-specific training details are described below.

Table B1. Training configurations for the three SolarFINE stages.

Stage	Optimizer/schedule	Initial LR	Weight decay	Batch	Epochs
SolarFINE-T	AdamW/cosine	5×10^{-4}	10^{-5}	64	100
Hourly spatial stage	AdamW/cosine	10^{-4}	10^{-4}	16	100
Subhourly diffusion	AdamW/cosine	10^{-4}	10^{-5}	16	30

The hourly spatial stage reduces its learning rate to 10^{-6}, uses automatic mixed precision, and stopped after approximately 14 epochs in the reported run. The diffusion stage uses an exponential moving average coefficient of 0.9999 and does not use automatic mixed precision.

B3 Cross-Scale Consistency

The SolarFINE-T loss combines hourly meteorological-state error, hourly irradiance error, and a 6 h integral-consistency term with $\lambda_E = 0.15$. The six hourly irradiance fields are multiplied by 3600 s and summed before comparison with the accumulated radiation corresponding to the six supervisory targets. The integral term is normalized by the standard deviation of accumulated radiation in the training samples. It is a soft training constraint, so the predicted hourly fields are not forced to match the input accumulation exactly for every window. In SolarFINE-J, hourly-mean consistency is imposed by normalizing the six frame weights: every 10-min field is nonnegative and the arithmetic mean of the six frames equals the hourly spatial background exactly.

B4 Integrated Gradients

Integrated Gradients is applied only to the hourly spatial stage of SolarFINE-J and does not attribute the diffusion sampling path. The attribution target is the spatially aggregated high-resolution correction $R_{q,h}$ over a fixed region. Absolute attributions are normalized within each sample and then grouped into meteorological state, coarse-grid radiation, solar conditions, and DEM. Dry- and rainy-season samples use the same input standardization and are matched by solar-zenith-angle range.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

Table B2. Hourly reconstruction errors for the 19 accompanying meteorological variables. RMSE and Pearson r are evaluated against hourly ERA5 fields from 2020.

Variable	Unit	RMSE			Pearson r		
		Linear	ModAFNO	SolarFINE-T	Linear	ModAFNO	SolarFINE-T
2-m temperature	K	1.351	0.987	1.038	0.988	0.994	0.993
Mean sea-level pressure	Pa	91.506	172.085	212.568	0.990	0.965	0.946
Total column water vapour	kg m^{-2}	0.656	0.646	0.705	0.999	0.999	0.999
Total cloud cover	1	0.126	0.099	0.100	0.944	0.966	0.965
High cloud cover	1	0.119	0.100	0.103	0.958	0.971	0.969
Medium cloud cover	1	0.126	0.105	0.107	0.923	0.947	0.945
Low cloud cover	1	0.148	0.119	0.121	0.927	0.953	0.952
850-hPa temperature	K	0.576	0.547	0.611	0.995	0.995	0.994
850-hPa specific humidity	g kg^{-1}	0.336	0.291	0.313	0.997	0.998	0.997
850-hPa u wind	m s^{-1}	0.504	0.396	0.413	0.977	0.986	0.984
850-hPa v wind	m s^{-1}	0.489	0.399	0.417	0.981	0.987	0.986
500-hPa temperature	K	0.291	0.309	0.338	0.998	0.998	0.997
500-hPa specific humidity	g kg^{-1}	0.223	0.193	0.207	0.995	0.996	0.995
500-hPa u wind	m s^{-1}	0.896	0.805	0.853	0.996	0.997	0.997
500-hPa v wind	m s^{-1}	0.944	0.717	0.757	0.978	0.987	0.986
200-hPa temperature	K	0.265	0.332	0.376	0.995	0.992	0.990
200-hPa specific humidity	g kg^{-1}	0.011	0.008	0.009	0.980	0.988	0.987
200-hPa u wind	m s^{-1}	1.118	1.163	1.191	0.999	0.999	0.999
200-hPa v wind	m s^{-1}	1.215	1.119	1.195	0.995	0.995	0.995

Figure B2 reports the seasonal group- and channel-level attributions. These attribution shares describe sensitivity of the model output to its inputs and are interpreted separately from the input-removal experiments in Table 4.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

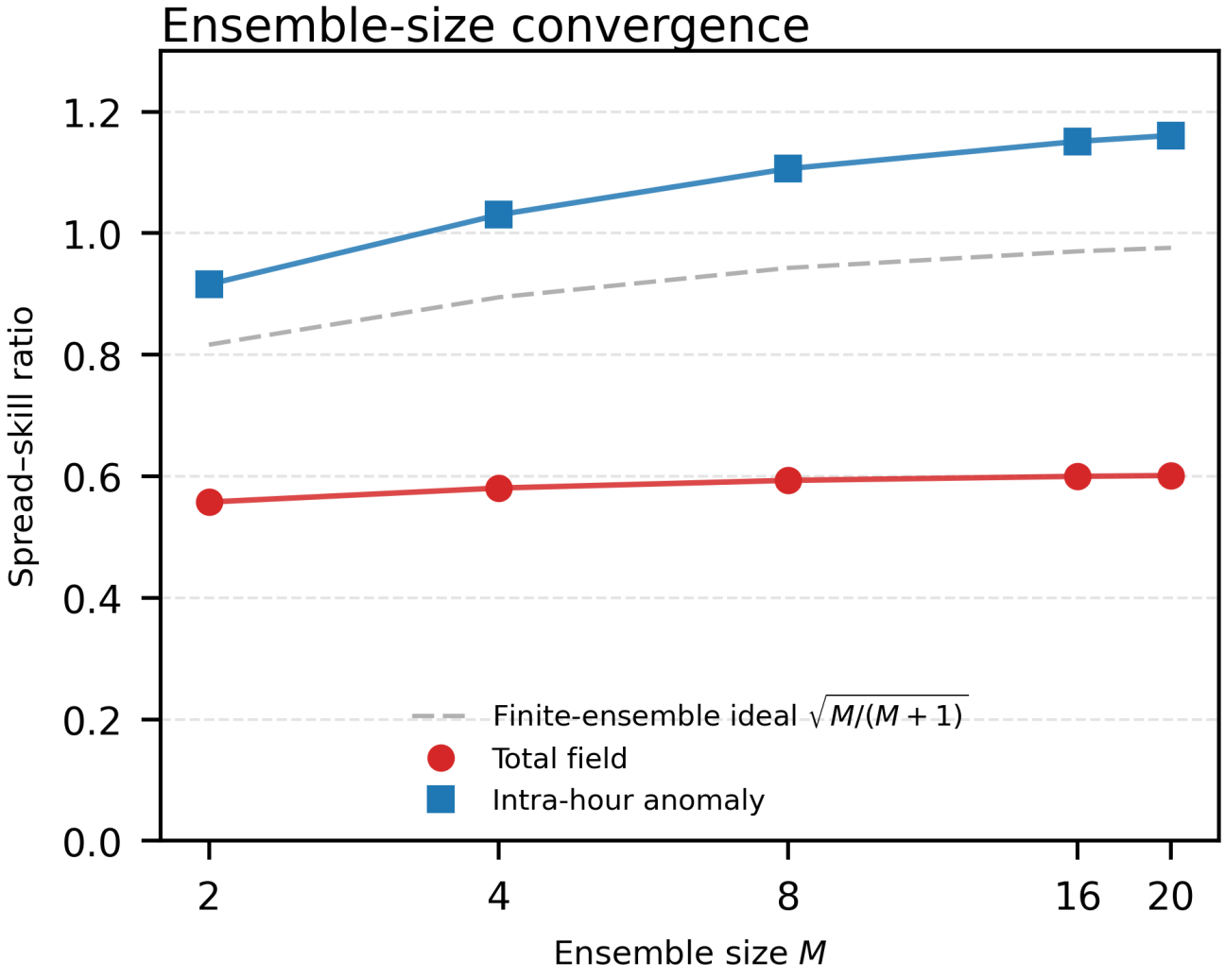


Figure B1. Ensemble size and spread–skill ratio. Red denotes the full 10-min field, and blue denotes intrahour anomalies after removal of the hourly mean. The gray dashed line shows the finite-member reference relation $\sqrt{M/(M+1)}$.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

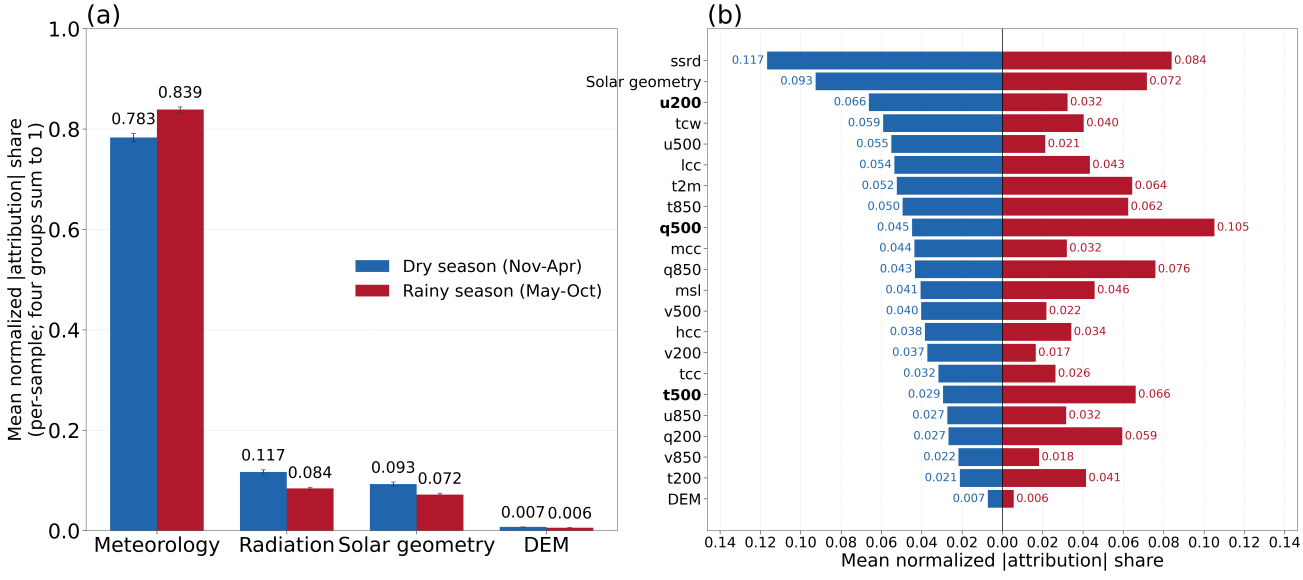


Figure B2. Seasonal input attribution for the hourly spatial stage. (a) Normalized absolute attribution shares for meteorological states, coarse-grid radiation, solar conditions, and DEM. (b) Attribution shares for individual input channels. Blue denotes the dry season (November–April), and red denotes the rainy season (May–October).

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

Author contributions. Zekai Dai conceptualized the study, designed the experiments and numerical case studies, conducted the experiments, and wrote the original draft. Man Wang provided the observational data. Yongwei Wang and Yuanjian Yang acquired funding, coordinated the project, and reviewed and edited the manuscript. Xindin Li, Xiaoying Jiang, Yajun Wang, and Jiangwei Ji contributed to the investigation and supporting research tasks.

Competing interests. The authors declare there are no conflicts of interest for this manuscript.

Acknowledgements. This study was supported by the China Meteorological Administration Innovation Development Project (CXFZ2026 J139) and the Open Research Fund of Hubei Key Laboratory of Intelligent Yangtze and Hydroelectric Science, China Yangtze Power Co., Ltd. (Project No. 2425020015). We thank the Yunnan Provincial Meteorological Bureau for providing the ground-based solar radiation observations. We acknowledge the High Performance Computing Center of Nanjing University of Information Science and Technology for their support of this work.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

References

Addison, H., Kendon, E. J., Ravuri, S., Aitchison, L., & Watson, P. A. G. (2026). Machine Learning Emulation of Precipitation From km-Scale UK Regional Climate Simulations Using a Diffusion Model. Journal of Advances in Modeling Earth Systems, 18, e2025MS005140. https://doi.org/10.1029/2025MS005140.

Aich, M., Hess, P., Pan, B., Bathiany, S., Huang, Y., & Boers, N. (2026). Conditional diffusion models for downscaling and bias correction of Earth system model precipitation. Geoscientific Model Development, 19, 1791–1808. https://doi.org/10.5194/gmd-19-1791-2026.

Annau, N. J., Cannon, A. J., & Monahan, A. H. (2023). Algorithmic hallucinations of near-surface winds: Statistical downscaling with generative adversarial networks to convection-permitting scales. Artificial Intelligence for the Earth Systems, 2(4). https://doi.org/10.1175/AIES-D-23-0015.1.

Appenzeller, T. (2025). Good morning, sunshine. Science, 390(6779), 1208–1209. https://doi.org/10.1126/science.aee8000.

Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. Nature, 525, 47-55. https://doi.org/10.1038/nature14956.

Baño-Medina, J., Manzanas, R., & Gutiérrez, J. M. (2020). Configuration and intercomparison of deep learning neural models for statistical downscaling. Geoscientific Model Development, 13, 2109-2124. https://doi.org/10.5194/gmd-13-2109-2020.

Bi, K., Xie, L., Zhang, H., et al. (2023). Accurate medium-range global weather forecasting with 3D neural networks. Nature, 619, 533-538. https://doi.org/10.1038/s41586-023-06185-3.

Bischoff, T., & Deck, K. (2024). Unpaired downscaling of fluid flows with diffusion bridges. Artificial Intelligence for the Earth Systems. https://doi.org/10.1175/AIES-D-23-0039.1.

Bodnar, C., Bruinsma, W. P., Lucic, A., et al. (2025). A Foundation Model for the Earth System. Nature. https://doi.org/10.1038/s41586-025-09005-y.

Chen, L., Zhong, X., Zhang, F., et al. (2023). FuXi: a cascade machine learning forecasting system for 15-day global weather forecast. npj Climate and Atmospheric Science, 6, 190. https://doi.org/10.1038/s41612-023-00512-1.

Chen, S., Li, L., Wei, Z., et al. (2024). Exploring topography downscaling methods for hyper-resolution land surface modeling. Journal of Geophysical Research: Atmospheres. https://doi.org/10.1029/2024JD041338.

Cheng, F., Zhang, S., Tian, Q., & Duan, W. (2024). Spatial downscaling of downward surface shortwave radiation based on image super-resolution. IEEE Transactions on Geoscience and Remote Sensing. https://doi.org/10.1109/TGRS.2024.3415051.

Chu, Y., Li, M., Coimbra, C. F. M., et al. (2021). Intra-hour irradiance forecasting techniques for solar power integration: A review. iScience, 24, 103136. https://doi.org/10.1016/j.isci.2021.103136.

Dai, Z. (2026). SolarFINE-MLWP (v0.1.0) [Software]. Zenodo. https://doi.org/10.5281/zenodo.23009436.

de Roda Husman, S., Hu, Z., van Tiggelen, M., et al. (2024). Physically-informed super-resolution downscaling of Antarctic surface melt. Journal of Advances in Modeling Earth Systems, 16, e2023MS004212. https://doi.org/10.1029/2023MS004212.

ECMWF. (2025). AIFS Single v1.0 [Model]. Hugging Face. https://doi.org/10.57967/hf/4629.

Fallah, B., Rostami, M., Russo, E., et al. (2025). Climate model downscaling in central Asia: a dynamical and a neural network approach. Geoscientific Model Development, 18, 161-180. https://doi.org/10.5194/gmd-18-161-2025.

Fuentes-Franco, R., Krus, K., Ivanov, M., et al. (2025). Pan-European high-resolution downscaling using deep learning. Journal of Geophysical Research: Machine Learning and Computation. https://doi.org/10.1029/2025JH000630.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.

Getter, D., Bessac, J., Rudi, J., & Feng, Y. (2024). Statistical treatment of convolutional neural network superresolution of inland surface wind for subgrid-scale variability quantification. Artificial Intelligence for the Earth Systems. https://doi.org/10.1175/AIES-D-23-0009.1.

Giroux, J.-S., Breton, S.-P., & Carreau, J. (2025). Interpolation-free deep learning for meteorological downscaling on unaligned grids across multiple domains with application to wind power. Artificial Intelligence for the Earth Systems, 4. https://doi.org/10.1175/AIES-D-24-0097.1.

Glawion, L., Polz, J., Kunstmann, H., Fersch, B., & Chwala, C. (2023). spateGAN: Spatio-temporal downscaling of rainfall fields using a cGAN approach. Earth and Space Science, 10, e2023EA002906. https://doi.org/10.1029/2023EA002906.

Glawion, L., Polz, J., Kunstmann, H., Fersch, B., & Chwala, C. (2025). Global spatio-temporal ERA5 precipitation downscaling to km and sub-hourly scale using generative AI. npj Climate and Atmospheric Science. https://doi.org/10.1038/s41612-025-01103-y.

González-Abad, J., & Gutiérrez, J. M. (2025). Are Deep Learning Methods Suitable for Downscaling Global Climate Projections? An Intercomparison for Temperature and Precipitation over Spain. Artificial Intelligence for the Earth Systems, 4. https://doi.org/10.1175/AIES-D-24-0121.1.

González-Abad, J., Baño-Medina, J., & Gutiérrez, J. M. (2023). Using explainability to inform statistical downscaling based on deep learning beyond standard validation approaches. Journal of Advances in Modeling Earth Systems, 15, e2023MS003641. https://doi.org/10.1029/2023MS003641.

Guo, A., Liu, Q., Wang, J., et al. (2025). A lightweight terrain-constraint model for wind spatial downscaling. Journal of Geophysical Research: Machine Learning and Computation. https://doi.org/10.1029/2024JH000147.

Harder, P., Hernandez-Garcia, A., Ramesh, V., et al. (2023). Hard-Constrained Deep Learning for Climate Downscaling. Journal of Machine Learning Research, 24(365), 1–40. https://jmlr.org/papers/v24/23-0158.html.

Harris, L., McRae, A. T. T., Chantry, M., Dueben, P. D., & Palmer, T. N. (2022). A generative deep learning approach to stochastic downscaling of precipitation forecasts. Journal of Advances in Modeling Earth Systems, 14, e2022MS003120. https://doi.org/10.1029/2022MS003120.

Hatanaka, Y., Glaser, Y., Galgon, G., Torri, G., & Sadowski, P. (2023). Diffusion models for high-resolution solar forecasts [Preprint]. arXiv:2302.00170. https://doi.org/10.48550/arXiv.2302.00170.

Haurwitz, B. (1945). Insolation in relation to cloudiness and cloud density. Journal of Meteorology, 2, 154-166. https://doi.org/10.1175/1520-0469(1945)002<0154:IIRTCA>2.0.CO;2.

Hersbach, H., Bell, B., Berrisford, P., et al. (2020). The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society, 146, 1999-2049. https://doi.org/10.1002/qj.3803.

Hu, W., Scholz, Y., Yeligeti, M., von Bremen, L., & Deng, Y. (2023). Downscaling ERA5 wind speed data: a machine learning approach considering topographic influences. Environmental Research Letters. https://doi.org/10.1088/1748-9326/aceb0a.

Husi, L., Wang, T., & Du, Y. (2021). East Asia-Pacific longwave/shortwave downward radiation at the surface data set (2016-2020) [Dataset]. National Tibetan Plateau / Third Pole Environment Data Center. https://doi.org/10.11888/Meteoro.tpdc.271729.

Jiang, P., Yang, Z., Wang, J., et al. (2023). Efficient super-resolution of near-surface climate modeling using the Fourier Neural Operator. Journal of Advances in Modeling Earth Systems, 15, e2023MS003800. https://doi.org/10.1029/2023MS003800.

Jimenez, P. A., Hacker, J. P., Dudhia, J., et al. (2016). WRF-Solar: Description and clear-sky assessment of an augmented NWP model for solar power prediction. Bulletin of the American Meteorological Society, 97(7), 1249-1264. https://doi.org/10.1175/BAMS-D-14-00279.1.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.



Lafferty, D. C., & Sriver, R. L. (2023). Downscaling and bias-correction contribute considerable uncertainty to local climate projections in CMIP6. npj Climate and Atmospheric Science, 6, 158. https://doi.org/10.1038/s41612-023-00486-0.

Lam, R., Sanchez-Gonzalez, A., Willson, M., et al. (2023). Learning skillful medium-range global weather forecasting. Science, 382, 1416-1421. https://doi.org/10.1126/science.adi2336.

Lang, S., Alexe, M., Chantry, M., et al. (2024). AIFS — ECMWF's data-driven forecasting system [Preprint]. arXiv:2406.01465. https://doi.org/10.48550/arXiv.2406.01465.

Leinonen, J., Nerini, D., & Berne, A. (2021). Stochastic super-resolution for downscaling time-evolving atmospheric fields with a generative adversarial network. IEEE Transactions on Geoscience and Remote Sensing. https://doi.org/10.1109/TGRS.2020.3032790.

Letu, H., Nakajima, T. Y., Wang, T., et al. (2022). A New Benchmark for Surface Radiation Products over the East Asia–Pacific Region Retrieved from the Himawari-8/AHI Next-Generation Geostationary Satellite. Bulletin of the American Meteorological Society, 103, E873–E888. https://doi.org/10.1175/BAMS-D-20-0148.1.

Liang, J., Cao, J., Sun, G., Zhang, K., Van Gool, L., & Timofte, R. (2021). SwinIR: Image Restoration Using Swin Transformer. Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 1833–1844. https://doi.org/10.1109/ICCVW54120.2021.00210.

Ling, F., Lu, Z., Luo, J.-J., et al. (2024). Diffusion model-based probabilistic downscaling for 180-year East Asian climate reconstruction. npj Climate and Atmospheric Science, 7, 131. https://doi.org/10.1038/s41612-024-00679-1.

Liu, Z., Chen, H., Bai, L., et al. (2024). MambaDS: Near-Surface Meteorological Field Downscaling With Topography Constrained Selective State-Space Modeling. IEEE Transactions on Geoscience and Remote Sensing, 62, 4112615. https://doi.org/10.1109/TGRS.2024.3496895.

Liu, Z., Pan, B., Chao, J., et al. (2026). Km-scale solar irradiance ensemble forecasting with generative deep learning. npj Artificial Intelligence. https://doi.org/10.1038/s44387-026-00133-y.

Mardani, M., Brenowitz, N., Cohen, Y., et al. (2025). Residual corrective diffusion modeling for km-scale atmospheric downscaling. Communications Earth & Environment, 6, 124. https://doi.org/10.1038/s43247-025-02042-5.

Mayer, M. J., Yang, D., & Szintai, B. (2023). Comparing global and regional downscaled NWP models for irradiance and photovoltaic power forecasting: ECMWF versus AROME. Applied Energy, 352, 121958. https://doi.org/10.1016/j.apenergy.2023.121958.

Moldovan, G., Pinnington, E., Prieto Nemesio, A., et al. (2026). AIFS Single 1.1.0: an update to ECMWF's machine-learned weather forecast model AIFS. Geoscientific Model Development, 19, 4703-4724. https://doi.org/10.5194/gmd-19-4703-2026.

Nijsse, F. J. M. M., Mercure, J.-F., Ameli, N., et al. (2023). The momentum of the solar energy transition. Nature Communications, 14, 6542. https://doi.org/10.1038/s41467-023-41971-7.

Niu, Z., Huang, W., Huang, S., et al. (2026). Intelligent Shanghai Typhoon Model (ISTM): A diffusion-based downscaling emulator from AIFS to kilometer-scale regional reanalysis for typhoon forecasts. Journal of Geophysical Research: Machine Learning and Computation, 3, e2025JH001072. https://doi.org/10.1029/2025JH001072.

Pathak, J., Subramanian, S., Harrington, P., et al. (2022). FourCastNet: A global data-driven high-resolution weather model using Adaptive Fourier Neural Operators [Preprint]. arXiv:2202.11214. https://doi.org/10.48550/arXiv.2202.11214.

Pinnington, E., Prieto Nemesio, A., & Moldovan, G. (2025). Operational release of AIFS Single 1.0. ECMWF Newsletter, 183. https://www.ecmwf.int/en/newsletter/183/news/operational-release-aifs-single-10.

Price, I., Sanchez-Gonzalez, A., Alet, F., et al. (2025). Probabilistic weather forecasting with machine learning. Nature, 637, 84–90. https://doi.org/10.1038/s41586-024-08252-9.

https://doi.org/10.5194/egusphere-2026-6594
Preprint. Discussion started: 16 November 2026
© Author(s) 2026. CC BY 4.0 License.



Rampal, N., Gibson, P. B., Sherwood, S., Abramowitz, G., & Hobeichi, S. (2025). A reliable generative adversarial network approach for climate downscaling and weather generation. Journal of Advances in Modeling Earth Systems, 17, e2024MS004668. https://doi.org/10.1029/2024MS004668.

Rampal, N., Hobeichi, S., Gibson, P. B., et al. (2024). Enhancing regional climate downscaling through advances in machine learning. Artificial Intelligence for the Earth Systems, 3, e230066. https://doi.org/10.1175/AIES-D-23-0066.1.

Schillinger, M., Samarin, M., Shen, X., Knutti, R., & Meinshausen, N. (2026). EnScale: Temporally-Consistent Multivariate Generative Downscaling via Proper Scoring Rules. Journal of Advances in Modeling Earth Systems, 18, e2025MS005593. https://doi.org/10.1029/2025MS005593.

Smith, O., Cattell, O., Farcot, E., O'Dea, R. D., & Hopcraft, K. I. (2022). The effect of renewable energy incorporation on power grid stability and resilience. Science Advances, 8, eabj6734. https://doi.org/10.1126/sciadv.abj6734.

Springenberg, M., Otero, N., Xue, Y., & Ma, J. (2026). DiffScale: Continuous Downscaling and Bias Correction of Subseasonal Wind Speed Forecasts Using Diffusion Models. Journal of Advances in Modeling Earth Systems, 18, e2025MS005282. https://doi.org/10.1029/2025MS005282.

Stengel, K., Glaws, A., Hettinger, D., & King, R. N. (2020). Adversarial super-resolution of climatological wind and solar data. Proceedings of the National Academy of Sciences, 117, 16805-16815. https://doi.org/10.1073/pnas.1918964117.

Tomasi, E., Franch, G., & Cristoforetti, M. (2025). Can AI be enabled to perform dynamical downscaling? A latent diffusion model to mimic kilometer-scale COSMO5.0_CLM9 simulations. Geoscientific Model Development, 18, 2051–2078. https://doi.org/10.5194/gmd-18-2051-2025.

Vandal, T., Kodra, E., Ganguly, S., et al. (2017). DeepSD: Generating high resolution climate change projections through single image super-resolution. Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1663–1672. https://doi.org/10.1145/3097983.3098004.

Weyn, J. A., Ni, Z., Misra, A., Fein, W., Dong, H., Bruinsma, W. P., Turner, R. E., Corey, M., Thambiratnam, K., White, K., Takeda, K., & Sun, H. (2026). Aurora 1.5: Fine-Tuning a Foundation Model for Medium-Range Ensemble Weather Prediction [Preprint]. Microsoft Research, 9 July 2026. https://www.microsoft.com/en-us/research/wp-content/uploads/2026/07/Aurora_1_5_Paper.pdf.

Wolvin, S., Strong, C., Brewer, S., Warner, A., & Steenburgh, W. J. (2026). Evaluation of a Convolutional Neural Network to Predict Wintertime Orographic Precipitation Gradients of the CONUS Northern Rockies. Artificial Intelligence for the Earth Systems, 5. https://doi.org/10.1175/AIES-D-25-0066.1.

Yi, C., Yu, M., Qian, W., Wen, Y., & Yang, H. (2026). Efficient kilometer-scale precipitation downscaling with conditional wavelet diffusion. Journal of Geophysical Research: Machine Learning and Computation, 3, e2025JH000941. https://doi.org/10.1029/2025JH000941.

Zhong, X., Chen, L., Fan, X., Qian, W., Liu, J., & Li, H. (2024). FuXi-2.0: Advancing machine learning weather forecasting model for practical applications [Preprint]. arXiv:2409.07188. https://doi.org/10.48550/arXiv.2409.07188.