

Article

From Hourly Irradiance Discrepancy to Day-Ahead Photovoltaic Energy: Evaluating Free Weather APIs for Urban Energy Management

Matej Cenky , Jozef Bendik  and Peter Janiga 

Department of Power and Applied Electrical Engineering, Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in Bratislava, Ilkovičova 3, 841 04 Bratislava, Slovakia; jozef.bendik@stuba.sk (J.B.); peter.janiga@stuba.sk (P.J.)

* Correspondence: matej.cenky@stuba.sk

Abstract

Freely accessible weather APIs are an attractive input for photovoltaic (PV) power prediction, yet how their reference-based error carries through—under temporal aggregation—into day-ahead PV energy is rarely quantified before model development. This study audits eight freely accessible weather-data services—seven forecast services and the Meteostat observational archive—and finds that, for the endpoints, subscription tiers, and collection period evaluated here, only two of the forecast services expose hourly global horizontal irradiance (GHI), while a third advertised irradiance field returns empty—itsself a material result for municipal integrators. The two irradiance-capable services are propagated through a common physical PV model of the 548 kWp east–west rooftop plant being built on a university campus in Bratislava, Slovakia, against independent references (CAMS irradiance; NASA POWER temperature and wind). Throughout, “error” denotes discrepancy against the reference, not measurement truth. Over a common 62-day spring–summer window, the day-ahead daily-energy mean absolute error (MAE, relative to mean reference daily energy) is 10–11% for both sources; Open-Meteo’s winter-inclusive own window raises its value to about 13%. On the matched window, Open-Meteo has the lower hourly GHI RMSE (121 vs. 128 $\text{W}\cdot\text{m}^{-2}$), yet this advantage does not carry through to day-ahead energy—the point-estimate ordering even changes—and neither matched-window difference is statistically resolved. Because daily aggregation rewards low bias over low scatter, this ordering change is already present in daily GHI energy—before the PV conversion—so hourly irradiance accuracy alone does not determine day-ahead energy. Propagated input-data auditing is therefore an advisable step before ML-based PV forecasting and day-ahead scheduling of urban distributed PV.

Keywords: photovoltaic power forecasting; urban PV systems; temporal error aggregation; open weather data; day-ahead scheduling



Academic Editor: Laura Cirrincione

Received: 15 July 2026

Revised: 31 July 2026

Accepted: 2 August 2026

Published: 3 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Photovoltaic generation is now a structural component of European power systems, and its variability places forecasting at the center of both grid operation and market participation. The accuracy of a PV power forecast is bounded first by the quality of its weather inputs and only then by the sophistication of the model that maps weather to power [1,2]. Machine-learning (ML) methods have come to dominate the PV-forecasting literature, but reviews of the field repeatedly find that input data and feature selection

govern accuracy at least as strongly as the choice of learning algorithm [1,3,4]. A model trained on biased or noisy weather inputs inherits that bias or noise unless the input error is explicitly characterized and corrected.

This problem is particularly relevant for smart-city energy systems [5]. Municipal buildings, university campuses, rooftop PV portfolios [6], local energy communities, electric-vehicle charging hubs, and urban microgrids increasingly require low-cost day-ahead forecasts for scheduling and flexibility management. Many such actors lack access to commercial numerical-weather-prediction products or dense on-site measurement networks. Freely accessible weather APIs are therefore attractive because they can provide forecast data for arbitrary coordinates with little deployment cost. However, treating such services as neutral inputs can be misleading: each provider has its own model chain, update cycle, spatial resolution, variable availability, and systematic error structure.

The economic cost of this omission is largest at the day-ahead horizon. Generators that bid PV output into a day-ahead market are penalized for the gap between scheduled and delivered energy, so error in the predicted next-day power profile can map directly to imbalance cost [7,8]. The economically meaningful way to evaluate a forecast therefore depends on how the forecast is used, not on a single statistical score computed in isolation [9]. A characterization of input error that stops at irradiance in $\text{W}\cdot\text{m}^{-2}$ does not directly answer the operational question of how much day-ahead PV energy an urban energy manager should schedule.

The solar-forecasting community has built a mature apparatus for verifying irradiance forecasts—standardized error metrics, skill scores referenced to climatology, persistence, or clear-sky models, and metrics sensitive to ramp timing [10–13]—and a tradition of comparing numerical weather prediction (NWP) models across regions [14,15]. Yet this apparatus is often applied to irradiance as a meteorological variable. It less frequently follows the error through the irradiance-to-power conversion to the kilowatt-hours a plant would deliver. Conversely, PV-modeling and ML-forecasting studies often assume a single weather input and do not examine how choosing among freely accessible providers changes the resulting day-ahead power profile. This leaves a practical gap for smart-city energy applications: before training or deploying an ML forecast model, the raw weather-input uncertainty should be quantified in the same unit and horizon used by the operator.

The study is guided by three research questions:

- RQ1. Which PV-relevant variables do freely accessible weather services expose with usable values, and in particular, how many deliver hourly irradiance?
- RQ2. How large is the day-ahead discrepancy of those services against independent references—both in the meteorological domain and in the operational domain of hourly AC power and daily energy?
- RQ3. Does a ranking established on hourly irradiance error survive propagation through a physical PV model and aggregation to day-ahead daily energy?

RQ3 carries the working hypothesis that it does. That assumption is implicit whenever weather sources are ranked by irradiance RMSE. The results reported below do not support it at this record length.

This paper addresses that gap by treating the weather forecast itself as the object of study. The contributions are:

1. A time-resolved input-data audit of eight freely accessible weather-data services—seven forecast services and the Meteostat observational archive. Error against independent references is quantified for the PV-relevant variables that the forecast services expose numerically—global horizontal irradiance (GHI), where available; 2 m air temperature; and wind speed—using the CAMS Radiation Service for GHI

and NASA POWER for temperature and wind. The audit additionally records which providers expose which cloud-cover-related fields, an availability finding rather than an error comparison.

2. A propagated PV-power discrepancy analysis in which the two services that expose hourly irradiance are passed through the same physical PV model, producing day-ahead hourly power and daily energy for the 548 kWp east–west rooftop system planned at the site.
3. A transparent, repeatable smart-city energy case study at the site of a rooftop PV plant being built in Bratislava, framed as the pre-commissioning stage of a campus PV digital twin: a proof of concept to be revalidated with a full-year record and measured plant output once the plant is commissioned.

The study does not provide a definitive ranking of all weather providers, nor does it measure final PV forecast error against an operating plant. It quantifies how raw, reference-based weather-forecast discrepancies propagate into PV-power discrepancies under a transparent and repeatable modeling chain—the input-characterization step that should precede ML training, source selection, and day-ahead scheduling in smart-city energy systems.

2. Literature Review

2.1. PV Power Forecasting and Weather-Input Dependence

PV power forecasting spans horizons from seconds to days and draws on ground measurements, sky imagery, satellite data, and NWP, with methods ranging from physical models to statistical and machine-learning approaches [1]. Recent reviews document a pronounced shift toward deep learning, with architectures such as LSTM and CNN variants applied to irradiance and power series [16,17]. A consistent finding across this literature is that performance is bounded by the weather input: feature selection and the quality of the NWP-derived predictors influence accuracy as much as the learning algorithm itself [3], and hybrid schemes that exploit physically meaningful transformations of the weather input can outperform purely data-driven approaches [4]. This motivates treating the weather input as a first-order determinant of forecast quality rather than as interchangeable external data.

2.2. Forecast Verification and Reference Datasets

The verification of deterministic solar forecasts has been consolidated into a community framework that prescribes standard accuracy measures—RMSE, MAE, and MBE—alongside skill scores that normalize performance against a reference forecast [10]. The choice of reference is itself a methodological decision: climatology, persistence, and their optimal convex combination provide rigorous baselines for skill computation [12], while clear-sky models supply a physically grounded alternative. Standardized metric suites extend beyond point error to ramp behavior, distributional fidelity, and economically oriented measures [11,18], and ramp/time-alignment metrics capture temporal-mismatch errors that RMSE and MAE miss [13]. Calibration and post-processing can materially change verification conclusions [19], so raw freely accessible forecasts should be characterized before they are used as ML inputs.

In the absence of on-site measurement, surface solar radiation is commonly referenced to reanalysis or satellite-derived products. ERA5 provides globally consistent hourly fields and is widely used in solar-energy studies [20], but validation studies show that reanalysis irradiance can have weather-regime-dependent bias [21,22]. Satellite-based products such as CM SAF SARA and the CAMS Radiation Service provide independent cross-checks and are linked to validated clear-sky and all-sky irradiance retrieval methods [23–27]. NASA POWER, based partly on MERRA-2 and related meteorological products, provides

a further independent source for temperature and wind variables [28,29]. These datasets are not measurement truth at the site; they are independent references that can identify relative forecast-reference discrepancies.

2.3. PV Modeling Chain and Propagated Error

Converting irradiance to PV power follows a well-established chain: decomposition of global irradiance into direct and diffuse components, projection onto the array plane, estimation of cell temperature, and conversion to DC and AC power. The open-source library pvlib implements this chain with community-validated models [30,31]. Transposition relies on anisotropic sky models such as Hay–Davies, Reindl, and Perez [32–34], whose accuracy on inclined surfaces has been empirically validated [35]; cell temperature can be represented by models such as Faiman’s [36]; and module or array behavior can be represented with models such as Sandia or PVWatts [37]. Each stage transforms input error rather than eliminating it, and comparisons of physical PV-forecasting models show that upstream weather-input and transformation errors can dominate the final power-forecast error budget [38]. Therefore, assessing forecast quality strictly in $\text{W}\cdot\text{m}^{-2}$ can be insufficient for PV applications.

2.4. Economic and Smart-City Relevance

Forecast error becomes operationally meaningful at the market and energy-management interface. In day-ahead markets, improved PV forecasting can reduce imbalance exposure [7], and recent work explicitly monetizes PV forecast errors produced by different input-data and model configurations [8]. The appropriate verification metric depends on the market and operational context rather than on RMSE alone [9]. For smart cities, the same logic applies beyond wholesale markets: municipal energy managers, campus operators, and local energy communities need forecasts that support storage scheduling, EV charging, demand response, and local flexibility decisions. Smart-city energy-management systems increasingly coordinate distributed renewable generation, storage, and flexible demand [5]; urban rooftop PV is a substantial and spatially distributed municipal resource [6]; and digital-twin platforms are emerging to plan and operate such urban energy systems [39]. The present study therefore evaluates raw forecast inputs in terms of propagated day-ahead PV energy and power profiles, not only meteorological discrepancy.

2.5. Research Gap

Taken together, the literature leaves a specific applied gap. Forecast verification is mature but usually stops at the meteorological variable; reference datasets are well studied but often used to evaluate one product at a time; the PV modeling chain is established but usually driven by a chosen weather input; and the economic literature shows that timing and energy-volume error matter. What is less often quantified is the propagated, time-resolved, day-ahead PV-power consequence of selecting among freely accessible weather APIs. This paper supplies that input-characterization step as a proof-of-concept framework for smart-city PV forecasting and scheduling.

3. Methodology

This section describes the study site, forecast data, independent references, error metrics, PV propagation model, and day-ahead protocol. The overall pipeline is summarized in Figure 1.

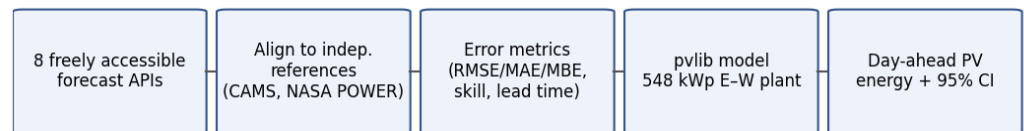


Figure 1. Analysis pipeline from freely accessible weather forecasts to day-ahead PV power.

3.1. Study Site and Forecast Data Collection

The study site is the campus of the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology (STU FEI), in Bratislava, Slovakia, at approximately 48.15° N, 17.07° E, where a 548 kWp rooftop PV plant is under construction. Forecasts are collected for six closely spaced points, labeled a–e and p, that span the array footprint across the campus roofs. The points are separated by only a few tens of meters and are treated as a spatial cluster.

3.1.1. Forecast Sources

Hourly forecasts are retrieved from eight freely accessible or free-tier weather-data services: Open-Meteo [40] (<https://open-meteo.com> (accessed on 29 June 2026)), Meteosource (<https://www.meteosource.com> (accessed on 29 June 2026)), OpenWeatherMap (<https://openweathermap.org> (accessed on 29 June 2026)), Tomorrow.io (<https://www.tomorrow.io> (accessed on 29 June 2026)), Visual Crossing (<https://www.visualcrossing.com> (accessed on 29 June 2026)), WeatherAPI (<https://www.weatherapi.com> (accessed on 29 June 2026)), World Weather Online (<https://www.worldweatheronline.com> (accessed on 29 June 2026)), and Meteostat (<https://meteostat.net> (accessed on 29 June 2026)). Meteostat returns observed station history rather than a forecast; it is retained in the archive as a potential observational cross-check but is not analyzed in this manuscript, so the results below cover the seven forecast services. Responses are archived hourly as timestamped JSON and CSV files. Coverage is uneven across sources: Open-Meteo spans 19 November 2025 to 29 June 2026, including winter, whereas the other services were added later and span approximately 16 April 2026 to 29 June 2026. NASA POWER appears only as a small reference-side sample and is not used as a forecast source. After deduplication, the consolidated archive comprises approximately 184,600 files.

The term “freely accessible” covers both open-data services and commercial services with free access tiers. The analysis concerns the practical input data available to a user without paid proprietary forecast feeds, not the software license or internal model openness of each provider.

3.1.2. Variables and Irradiance Availability

The analysis targets variables that drive PV output and are commonly exposed by providers: GHI, 2 m air temperature, wind speed, and cloud-related variables. A material finding of the data audit is that, for the endpoints, subscription tiers, and collection period evaluated here (Table 1), hourly GHI is exposed by only two services: Open-Meteo, which provides a broader irradiance suite, and Visual Crossing, which provides solar radiation. Meteosource provides an irradiance field that is empty in the collected free-tier data, and OpenWeatherMap, Tomorrow.io, WeatherAPI, World Weather Online, and Meteostat do not expose hourly irradiance in the retrieved data. They provide cloud cover and, for some services, UV-related variables. Table 2 summarizes the audit.

Table 1. Configuration of the eight collected services: endpoint or product queried, native temporal resolution of the returned series, forecast horizon as requested by the collection scripts, and the relevant free-tier constraint. All services were polled once per hour over the collection period. Open-Meteo’s 16-day horizon corresponds to the 384 h ceiling of the self-consistency analysis. No service exposes a model-selection parameter on its free tier except Open-Meteo, which accepts an optional `models` argument; no such argument was sent, so its automatic selection applied (see text). Values were read from the retrieval scripts, not from provider documentation.

Service	Endpoint/Product	Native Res.	Horizon	Free-Tier Constraint
Open-Meteo	api.open-meteo.com/v1/forecast (accessed on 15 July 2026)	1 h	16 d	No key; fair use
Visual Crossing	weather.visualcrossing.com/VisualCrossingWebServices/rest/services/timeline (accessed on 15 July 2026)	1 h	7 d	1000 records/day
Meteosource	meteosource.com/api/v1/free/point (accessed on 15 July 2026)	1 h	24 h	24 h hourly cap
OpenWeatherMap	api.openweathermap.org/data/2.5/forecast (accessed on 15 July 2026)	3 h	5 d	60 calls/min
Tomorrow.io	api.tomorrow.io/v4/timelines (accessed on 15 July 2026)	1 h	5 d	500 calls/day, 25/h
WeatherAPI	api.weatherapi.com/v1/forecast.json (accessed on 15 July 2026)	1 h	3 d	3 d maximum
World Weather Online	api.worldweatheronline.com/premium/v1/weather.ashx (accessed on 15 July 2026)	1 h	14 d	500 requests/day
Meteostat	Python client, bulk data; station LZIB (WMO 11816)	1 h	72 h back	No key required

Table 1 documents the configuration under which each service was queried. Reproducing this study requires the endpoint, the model selection, and the free-tier constraints in force, because all three shape what a service delivers.

Table 2. Data-availability audit of the collected services: which PV-relevant fields each service exposed with usable values in the retrieved free-tier data (November 2025–June 2026). “Empty” denotes an advertised field that returned no data across the collection period.

Service	GHI	2 m Temp.	Wind	Cloud	Role
Open-Meteo	yes	yes	yes	yes	forecast
Visual Crossing	yes	yes	yes	yes	forecast
Meteosource	empty	yes	yes	yes	forecast
OpenWeatherMap	no	yes	yes	yes	forecast
Tomorrow.io	no	yes	yes	yes	forecast
WeatherAPI	no	yes	yes	yes	forecast
World Weather Online	no	yes	no	yes	forecast
Meteostat	no	yes	yes	yes	observations

Four points qualify the table. First, the PV-relevant fields were requested under provider-specific names and mapped to a common schema: GHI from `shortwave_radiation` (Open-Meteo), `solarradiation` (Visual Crossing), and `irradiance` (Meteosource, empty throughout); air temperature from `temperature_2m`, `temp`, `temperature`, `temp_c` or `tempC`; wind speed from `wind_speed_10m`, `windspeed`, `wind_speed`, `windSpeed`, `wind_kph` or

windSpeedMeterSec; and cloud cover from `cloud_cover`, `cloudcover`, `cloud_cover_total` or `cloudCover`. Wind was returned in $\text{km}\cdot\text{h}^{-1}$ by four services and in $\text{m}\cdot\text{s}^{-1}$ by the others, and was converted to $\text{m}\cdot\text{s}^{-1}$ on ingestion. Open-Meteo was queried for a broader irradiance suite than GHI alone—`direct_radiation`, `diffuse_radiation`, `direct_normal_irradiance`, and `global_tilted_irradiance`—although only GHI enters the analysis, so that every source passes through the identical decomposition.

Second, Open-Meteo does not operate its own numerical weather prediction model; it redistributes output from several centres and, in the absence of an explicit models parameter, selects among them automatically. The evaluated object is therefore the provider's default operational blend rather than any single NWP model. That is deliberate for this study, because it is what a municipal integrator obtains without special configuration, but it carries two implications: the underlying model mix may vary with region and horizon, and it may change over time. The results characterise the service as delivered.

Third, OpenWeatherMap's free-tier product is a 5-day forecast at a 3 h step rather than an hourly one. Its values were used as returned, without interpolation to an hourly grid.

Fourth, providers do not consistently publish a native grid resolution for their free-tier point endpoints, so that column is omitted rather than populated with values that could not be verified. The question it addresses is settled empirically instead: as reported in Section 3.1.3, the six collection points—separated by a few tens of metres—return numerically identical Open-Meteo daytime GHI, confirming that every provider's native grid is far coarser than the array footprint and that the cluster average is effectively lossless.

The comparison is therefore two-tiered:

- (a) a propagated irradiance-to-PV-power comparison for the direct-GHI sources, Open-Meteo and Visual Crossing;
- (b) a broader temperature, wind, and cloud-variable audit across the collected services.

Cloud-to-GHI reconstruction for irradiance-free services is not included in this manuscript and is left for future work.

3.1.3. Spatial Cluster Handling

Because the six collection points lie within the array footprint and the providers' native grids are much coarser than the inter-point spacing, the forecasts at these points are nearly identical. They are therefore averaged to a single cluster series per source and variable. In the collected data, the six points are numerically identical for Open-Meteo daytime GHI, with mean across-point standard deviation of $0.0 \text{ W}\cdot\text{m}^{-2}$; this exact check was performed for Open-Meteo, and comparable near-losslessness is expected for the other providers given their coarser native grids relative to the few-tens-of-meters point spacing. The cluster average is therefore effectively lossless while also providing robustness against intermittent data loss in the collection pipeline.

3.1.4. Data Quality and Gaps

The collection pipeline ran across multiple machines and contains outages. Missing intervals are recorded rather than imputed, and error statistics are computed over matched, jointly available timestamps only. After alignment to the references, the full day-ahead lead band contains approximately 58,000 matched hourly GHI forecast–reference pairs for Open-Meteo and approximately 17,000 for Visual Crossing, over their respective own windows. This band covers 18–30 h and all issue times, before the nearest-to-24 h down-selection defined in Section 3.5, which reduces the count to roughly one pair per valid hour. The propagated two-source PV comparison is evaluated on the common matched window of 62 days. Temperature and wind comparisons, after the same nearest-to-24 h selection,

contain of order 10^3 matched day-ahead pairs per service (roughly 500 to 5000), depending on provider coverage and window length.

3.2. Reference Datasets and Definition of Error

Because no single freely available archive provided complete, vendor-independent coverage of all reference variables over the study period, two independent references are combined. Surface GHI is taken from the CAMS Radiation Service [25–27]. Air temperature and wind speed are taken from the NASA POWER archive [28,29]. Both are independent of the evaluated forecast APIs and provide complete coverage for the relevant variables and periods used here.

Neither reference is treated as perfect site-level truth. CAMS is satellite-derived and NASA POWER is based on satellite and reanalysis products. Throughout the manuscript, therefore, “error” means discrepancy against the selected independent reference. The study measures how forecast-reference discrepancies propagate through a PV model, not the complete error of an operational PV plant against measured production. Once on-site PV measurements become available, the same framework can replace the reference-driven PV series with measured plant output.

A subtler concern is reference independence. Open-Meteo blends several NWP models (ICON, ECMWF-IFS, GFS), while the CAMS Radiation Service builds all-sky irradiance on the McClell clear-sky model driven by CAMS/IFS aerosols [25–27]; the shared ECMWF lineage could in principle correlate Open-Meteo’s GHI errors with the reference and flatter its scores. Two considerations bound this. First, the dominant source of day-ahead GHI error is cloud, and the CAMS cloud component is satellite-retrieved (Heliosat-4, from Meteosat), observationally anchored, and independent of both providers’ NWP cloud fields. Second, any residual shared-chain correlation would inflate Open-Meteo’s apparent meteorological advantage—which, as the results below show, still fails to transfer to day-ahead energy—so the confound is conservative for the central caution of this paper. Reference non-independence is nonetheless listed among the limitations.

The uncertainty of the references themselves bounds what this comparison can resolve. Validation against European BSRN stations indicates an hourly all-sky GHI RMSE of approximately $50\text{--}65\text{ W}\cdot\text{m}^{-2}$ for the CAMS Radiation Service, with larger relative errors at the cloudier northern and oceanic sites [41]. These statistics are pooled over all sky conditions and are not stratified by cloud regime, so no separate broken-cloud figure is quoted here. For NASA POWER, whose meteorological fields derive from MERRA-2, the project’s global hourly validation against 2599 surface stations reports an RMSE of $3.32\text{ }^{\circ}\text{C}$ for 2 m air temperature and $2.47\text{ m}\cdot\text{s}^{-1}$ for wind speed [42]. Wind in particular should be read as grid-scale regional flow rather than site-specific conditions, since buildings, surface roughness, and local topography are unresolved at the MERRA-2 grid scale—a caveat that applies directly to an urban rooftop.

Two consequences follow, and they pull in opposite directions. First, part of the discrepancy reported here originates in the reference rather than the forecast, so the absolute values are upper bounds on the forecasts’ own error. Treating the two error sources as independent and subtracting a $60\text{ W}\cdot\text{m}^{-2}$ reference component in quadrature would reduce the matched-window GHI RMSE of 121 and $128\text{ W}\cdot\text{m}^{-2}$ to roughly 105 and $113\text{ W}\cdot\text{m}^{-2}$. Second, because the reference error is common to both providers, it largely cancels in the paired difference rather than adding to it. On the same first-order estimate, the gap between the two providers would be about 8 rather than $7\text{ W}\cdot\text{m}^{-2}$, so reference noise mildly *compresses* the apparent difference instead of manufacturing it.

The independence assumption behind that estimate is imperfect, for the reasons given above, so the corrected values are illustrative rather than definitive. The qualitative

implication is nonetheless robust and it sharpens the paper's central point: Open-Meteo's hourly irradiance advantage is unlikely to be an artefact of reference noise, and may be marginally larger than measured, which makes its failure to carry through to day-ahead energy the more notable. It is also a further argument for judging sources on propagated, aggregated quantities, which are less sensitive to per-hour reference noise than hourly point-error metrics.

All reference fields are aligned to forecast valid times in UTC. Forecast lead time is computed as the interval between issue time and valid time, allowing the discrepancy to be resolved by horizon.

3.3. Forecast-Error Metrics

For each source, variable, and lead-time bin, forecast values are compared with the time-aligned reference using standard deterministic-verification measures [10]: MBE, MAE, and RMSE. Let $\hat{x}_t$ denote the forecast and x_t the reference at matched timestamp t , with N the number of matched hourly pairs and $\Delta t = 1$ h the time step. The hourly discrepancy is

$$e_t = \hat{x}_t - x_t, \quad (1)$$

so that a negative e_t denotes a forecast below the reference. This sign convention is used throughout, including for every bias figure reported below. The three summary measures are

$$\text{MBE} = \frac{1}{N} \sum_{t=1}^N e_t, \quad \text{MAE} = \frac{1}{N} \sum_{t=1}^N |e_t|, \quad \text{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^N e_t^2}. \quad (2)$$

Daily quantities are simple sums over the matched hours of day d , with no gap-filling. Daily irradiation and daily PV energy are

$$H_d = \sum_{t \in d} \text{GHI}_t \Delta t, \quad E_d = \sum_{t \in d} P_{\text{ac},t} \Delta t, \quad (3)$$

where $P_{\text{ac},t}$ is the hourly AC power of Section 3.4. The daily-energy discrepancy is reported as a normalized MAE over the D evaluated days,

$$\text{nMAE} = \frac{\sum_{d=1}^D |\hat{E}_d - E_d|}{\sum_{d=1}^D E_d} \times 100\%, \quad (4)$$

that is, the sum of absolute per-day discrepancies divided by the sum of reference daily energies. It is not a mean of per-day ratios, and the two differ whenever daily energy varies across the window. The same definition applied to H_d gives the normalized daily-irradiation MAE used in the error ladder. The daily-energy bias is

$$\text{MBE}_E = \frac{1}{D} \sum_{d=1}^D (\hat{E}_d - E_d). \quad (5)$$

As pipeline diagnostics, the Pearson correlation coefficient and irradiance skill scores relative to two baselines—an independent clear-sky model implemented in pvlib using the Ineichen method [43] and 24 h persistence—are also computed, with skill defined as

$$\text{Skill} = 1 - \frac{\text{RMSE}_{\text{forecast}}}{\text{RMSE}_{\text{baseline}}}, \quad (6)$$

so that positive skill indicates lower RMSE than the baseline. The reported results lead with RMSE, MBE, and MAE, with clear-sky and 24 h-persistence skill reported below for irradiance; the Pearson correlation is retained only as an internal pipeline diagnostic and is not tabulated.

Two conventions govern which sample each denominator is taken over, and they are stated here once because the distinction changes the reported values. *Own-window* quantities use all matched timestamps available for that provider, so they span different periods for different providers and are season-confounded. *Matched-window* quantities use only the calendar days common to Open-Meteo and Visual Crossing, and are the only figures licensed for cross-provider comparison. Hourly metrics are computed over matched hours N ; daily metrics over evaluated days D . Both counts are given in the table captions.

No daytime mask is applied: hourly irradiance and power metrics are computed over all matched timestamps, so night hours (when both forecast and reference irradiance are near zero) are retained. This lowers the absolute $\text{W}\cdot\text{m}^{-2}$ and kW magnitudes relative to a daytime-only convention, but it is applied identically to every source and to the reference, so the cross-source comparisons are largely unaffected; a daylight-only sensitivity check (solar elevation above 5°) is reported with the results. To quantify sampling uncertainty under serially correlated daily errors, 95% confidence intervals are obtained with a day-block bootstrap using 2000 resamples. Whole days are resampled as blocks. Because the blocks are single days, multi-day synoptic persistence is not fully captured and the intervals are likely slightly anti-conservative; wider intervals would only reinforce the non-resolution reported below. For paired source-to-source comparisons, the same resampled day indices are used for both sources, preserving the paired structure. Confidence intervals are percentile intervals and are interpreted as uncertainty summaries for this finite sample rather than as a full multiple-comparison-adjusted hypothesis-testing framework.

Lead time is the primary axis for the reference-based results, binned as intraday 0–6 h, day-ahead around 24 h, and multi-day horizons up to 5–7 days. Two further axes are examined more selectively at the present record length. Diurnal structure is reported through the reference-free self-consistency analysis (Section 3.6) rather than as a full reference-based hour-of-day table. Seasonal structure is limited by the sub-annual record to a winter-inclusive versus spring–summer contrast for Open-Meteo. Fully stratified diurnal and seasonal reference-based tables are deferred to the full-year rerun.

Forecasts are evaluated as delivered, without any calibration or post-processing. Bias correction and provider blending are discussed as future operational extensions but are not applied to the main results.

3.4. PV Model and Error Propagation

Each irradiance-capable source is propagated through a common physical PV model implemented with pvlib [30,31], with the Sandia model cited as background for PV performance modeling [37]. The chain is given below in execution order. Solar position is computed first for each matched timestamp, giving the zenith angle θ_z .

Decomposition. Because Visual Crossing exposes only GHI, the global component is split with the Erbs correlation [44]. With the clearness index $k_t = \text{GHI}/(I_0 \cos \theta_z)$ and I_0 the extraterrestrial normal irradiance, the diffuse fraction is

$$\frac{\text{DHI}}{\text{GHI}} = \begin{cases} 1 - 0.09 k_t, & k_t \leq 0.22, \\ 0.9511 - 0.1604 k_t + 4.388 k_t^2 - 16.638 k_t^3 + 12.336 k_t^4, & 0.22 < k_t \leq 0.80, \\ 0.165, & k_t > 0.80, \end{cases} \quad (7)$$

$$\text{DNI} = \frac{\text{GHI} - \text{DHI}}{\cos \theta_z}. \quad (8)$$

Transposition. Plane-of-array irradiance is obtained with the Perez anisotropic sky model [33–35] as the sum of beam, sky-diffuse, and ground-reflected components,

$$G_{\text{POA}} = \text{DNI} \cos \theta_{\text{AOI}} + G_{d,\text{sky}} + \rho \text{GHI} \frac{1 - \cos \beta}{2}, \quad (9)$$

with θ_{AOI} the angle of incidence, $\beta = 12^\circ$ the surface tilt, and $\rho = 0.20$ the albedo. The sky-diffuse term $G_{d,\text{sky}}$ follows the Perez formulation, in which the isotropic, circumsolar, and horizon-brightening contributions are weighted by the empirical coefficients F_1 and F_2 ; the coefficient tables are those of Perez et al. [33] as implemented in pvlib. Reflection losses use the physical AOI model, and no spectral correction is applied.

Cell temperature. Module temperature follows Faiman [36],

$$T_{\text{cell}} = T_{\text{air}} + \frac{G_{\text{POA}}}{u_0 + u_1 v_w}, \quad (10)$$

with $u_0 = 25.0 \text{ W} \cdot \text{m}^{-2} \cdot \text{K}^{-1}$, $u_1 = 6.84 \text{ W} \cdot \text{m}^{-3} \cdot \text{s} \cdot \text{K}^{-1}$, and v_w the wind speed.

DC conversion and array summation. Each of the two faces $a \in \{\text{E}, \text{W}\}$ is modeled separately with the PVWatts DC relation [45],

$$P_{\text{dc},a} = \frac{G_{\text{POA},a}}{G_{\text{ref}}} P_{\text{dc0},a} [1 + \gamma(T_{\text{cell},a} - T_{\text{ref}})], \quad (11)$$

with $G_{\text{ref}} = 1000 \text{ W} \cdot \text{m}^{-2}$, $T_{\text{ref}} = 25^\circ \text{C}$, $\gamma = -0.29\%/^\circ \text{C}$ and $P_{\text{dc0},a} = 274.2 \text{ kW}$ per face. The two faces are summed and the PVWatts loss budget applied as a single multiplicative derate,

$$P_{\text{dc}} = (P_{\text{dc,E}} + P_{\text{dc,W}}) \prod_i (1 - L_i), \quad (12)$$

where the nine loss terms L_i of Table 3 combine to a derate of approximately 10.9%.

AC conversion. The two faces share one aggregate PVWatts inverter [45],

$$P_{\text{ac}} = \min(\eta(\zeta) P_{\text{dc}}, P_{\text{ac0}}), \quad (13)$$

with $P_{\text{ac0}} = 500 \text{ kW}$, nominal efficiency $\eta_{\text{nom}} = 0.984$, and $\eta(\zeta)$ the PVWatts part-load efficiency curve at loading $\zeta = P_{\text{dc}}/P_{\text{dc0}}$. Plant-level clipping is therefore represented explicitly by the $\min(\cdot)$ term. What is not resolved separately is *per-inverter* clipping across the four physical units, which is immaterial here because the 12° east–west geometry and the plant-level DC/AC ratio of 1.10 make inverter saturation rare. The exact pvlib version and the function called at each stage are listed in Appendix A.

The analysis uses the as-designed configuration of the plant, taken from its project documentation. The plant comprises 1192 Canadian Solar CS6.2-48TD-460 modules (460 Wp, n-type TOPCon, power temperature coefficient $-0.29\%/^\circ \text{C}$), totalling 548.3 kWp DC, distributed over thirteen flat-roof sections on an east–west ballasted racking system. Modules are mounted in east/west pairs at a 12° tilt (surface azimuths 90° and 270° , split evenly), a common choice on flat commercial roofs because it maximizes areal energy density and flattens the diurnal power profile. Module-level power optimizers aggregate the DC side into 42 strings feeding four Huawei string inverters (three rated 150 kW and one 50 kW), giving 500 kW nominal AC capacity and a DC/AC ratio of 1.10. The model represents the plant as two pvlib arrays—the east and west faces, 274.2 kWp each—with the Faiman cell-temperature model and a single aggregate PVWatts inverter at the 500 kW AC rating (capacity-weighted European efficiency 98.4%); the optimizer conversion loss is included

in the loss budget. Module counts, tilts, azimuths, inverter rating, thermal assumptions, and losses are held in an editable configuration file rather than hard-coded, so the analysis can be rerun for a revised design or for the commissioned plant. As a sanity check, the configured system yields a clear-sky midsummer AC peak of approximately 330 kW and approximately 2.9 MWh on a clear June day; the east–west split widens the production plateau while the low tilt keeps the peak well below the DC nameplate. The full parameter set is given in Table 3.

Table 3. As-designed PV-plant and model parameters, held in the editable `pv_config.yaml` and applied identically to every source and to the reference.

Parameter	Value
Modules	1192 × Canadian Solar CS6.2-48TD-460 (460 Wp, n-type TOPCon)
DC capacity	548.3 kWp
Array geometry	East and West faces, 12° tilt, azimuths 90°/270°, 274.2 kWp each
Power temperature coefficient	−0.29%/°C
Decomposition/transposition	Erbs/Perez
Cell-temperature model	Faiman ($u_0 = 25.0$, $u_1 = 6.84$)
AOI/spectral model	physical/no loss
Inverter	single aggregate PVWatts, 500 kW AC, $\eta_{\text{nom}} = 0.984$
DC/AC ratio	1.10
Albedo	0.20
PVWatts system losses (%)	soiling 2, shading 1, snow 0.5, mismatch 1, wiring 2, connections 1.5, LID 0.5, nameplate 1, availability 2; combined derate $\approx 10.9\%$

The same PV model and configuration are applied to every forecast source and to the reference. For a strictly like-for-like comparison, the same Erbs decomposition is applied to Open-Meteo GHI—and to the CAMS reference GHI—rather than using any native direct and diffuse components. Using a single decomposition for all series isolates the effect of the GHI forecast itself rather than of each provider’s internal component split. Because Open-Meteo’s native components are plausibly better than a generic decomposition, this choice, if anything, handicaps the source with the lower matched-window GHI RMSE point estimate and is therefore conservative. Each source’s PV run is driven by that source’s own forecast temperature and wind speed, with the NASA POWER reference values substituted only where a provider does not expose those fields; the reference-driven PV run uses NASA POWER temperature and wind together with CAMS GHI. Because the reference GHI passes through the identical decomposition, transposition, thermal, and inverter models, the resulting power discrepancy is attributable to the weather input and its transformation through the shared PV chain, not to differing decomposition treatments.

The array geometry and DC/AC ratio are fixed by the plant design; no sensitivity sweep over the decomposition model, transposition model, or loss assumptions is included in the present manuscript. Such sensitivity analysis is planned as the next extension and is listed as a limitation.

3.5. Day-Ahead Analysis Protocol

The headline analysis targets the day-ahead power profile. For each valid hour, the forecast nearest to 24 h lead is selected from the 18–30 h lead-time band. This constitutes a 24 h lead-time-controlled evaluation that approximates a day-ahead forecast rather than a fixed gate-closure schedule; real market gate-closure rules differ. Because the nearest-to-24 h rule can draw consecutive valid hours from different issue times, the assembled day-ahead

profile mixes issue times across the day. Each selected forecast series is propagated into an hourly PV power profile and compared with the reference-driven profile under the same system configuration. For each source's own-window daily-energy series, days with fewer than 20 matched hours are excluded; the matched two-source comparison is evaluated over the days common to both sources. Daily energy and daily irradiation are simple sums of the available matched hourly values, with no gap-filling applied.

A second protocol is evaluated as a robustness check. Under this *fixed gate-closure* protocol, the entire next day is taken from the single forecast run issued at or immediately before day-ahead gate closure. The gate is set at 12:00 CET on D−1, the gate closure of the single day-ahead market coupling, which corresponds to 10:00 UTC over the spring–summer matched window. A day is retained only if a run is available within 6 h of its gate; this holds for 59 of the 62 matched days, with a median gate age of 1.0 h for Open-Meteo and 0.6 h for Visual Crossing.

The two protocols trade off against each other and neither dominates. The nearest-to-24 h rule holds the effective lead close to 24 h, but mixes issue times within a day, so the assembled profile is not one an operator ever held. The fixed-gate protocol is operationally exact, but the effective lead then varies across the target day—typically from 15 to 38 h—which confounds lead time with hour of day. Both are therefore reported, and Section 4.2 gives the results side by side.

Two error views are reported: daily-energy discrepancy, relevant to volume settlement, and hourly profile-shape discrepancy, relevant to schedule accuracy and ramp-related operation [7,8]. Because only Open-Meteo and Visual Crossing expose hourly irradiance, the propagated power comparison is limited to those two services and is evaluated on their common set of days.

3.6. Complementary Self-Consistency Analysis

As a reference-free robustness check, forecast self-consistency is evaluated for the Open-Meteo archive, the only source spanning the full record with all relevant parameter groups. For each valid time, the shortest-lead forecast serves as a nearest-cast reference, and longer-lead forecasts are compared with it. The resulting statistics measure forecast drift, not absolute accuracy. They are reported as a consistency metric complementary to the independent-reference analysis and are computed for GHI, DNI, GTI, 2 m temperature, 10 m wind, and total cloud cover over lead times up to 384 h.

4. Results

Results are reported against the independent references: CAMS for GHI and NASA POWER for temperature and wind. The Open-Meteo record spans late November 2025 to late June 2026, while the remaining forecast providers mainly span mid-April to late June 2026. Unless otherwise stated, results refer to the day-ahead lead band, defined as 18–30 h lead with selection nearest to 24 h.

4.1. Meteorological Forecast-Reference Discrepancy

For GHI, available only from Open-Meteo and Visual Crossing in the collected data, Open-Meteo shows the lower discrepancy against CAMS. Over each provider's own window, Open-Meteo has a day-ahead GHI RMSE of $84.0 \text{ W}\cdot\text{m}^{-2}$, while Visual Crossing has $128.7 \text{ W}\cdot\text{m}^{-2}$. However, these own-window values are not directly comparable and are season-confounded: Open-Meteo's window includes low-irradiance winter months, which mechanically deflate its absolute $\text{W}\cdot\text{m}^{-2}$ RMSE relative to Visual Crossing's spring–summer-only window. Relative to an Ineichen clear-sky baseline, the own-window day-ahead GHI skill is 0.27 for Open-Meteo and 0.11 for Visual Crossing; relative to 24 h

persistence it is 0.27 and 0.26. As own-window quantities, these carry the same season confound. The clear-sky value is diagnostic only, because a clear-sky curve is an upper envelope rather than an all-weather forecast. On the common matched window of 62 days, Open-Meteo has a day-ahead GHI RMSE of $121 \text{ W}\cdot\text{m}^{-2}$ with a 95% CI of [111, 132], compared with $128 \text{ W}\cdot\text{m}^{-2}$ for Visual Crossing with a 95% CI of [117, 140]. The paired difference is $-6.9 \text{ W}\cdot\text{m}^{-2}$ with a 95% CI of $[-16.4, +2.4]$, which includes zero: the matched-window advantage is consistent in sign but is not statistically resolved at the current record length. Restricting to daylight hours (solar elevation above 5°) raises the matched-window GHI RMSE to approximately 157 (Open-Meteo) and $166 \text{ W}\cdot\text{m}^{-2}$ (Visual Crossing) but preserves both the ordering and the small gap, confirming that the retained night hours do not drive the comparison.

Discrepancy nevertheless increases with lead time (Figure 2). Lead-time behavior is shown in two panels, because the own-window and matched-window views answer different questions. Panel (a) gives each provider's error growth over its own observation window. Those curves are season-confounded and describe individual degradation patterns, not a ranking. Panel (b) restricts both providers to the 62 common days and is the only view that supports a direct comparison. Computed on each provider's own window, Open-Meteo's GHI RMSE rises from approximately $80 \text{ W}\cdot\text{m}^{-2}$ at 0–6 h to approximately $123 \text{ W}\cdot\text{m}^{-2}$ at 5–7 days, while Visual Crossing rises from approximately 121 to $210 \text{ W}\cdot\text{m}^{-2}$. The point-estimate ordering is the same at every horizon. Being own-window, it carries the same season confound.

The size of that confound is worth quantifying. In the 0–6 h bin Open-Meteo's GHI RMSE is $79.6 \text{ W}\cdot\text{m}^{-2}$ over its own winter-inclusive window but $113.8 \text{ W}\cdot\text{m}^{-2}$ over the matched spring–summer window—a difference produced entirely by the evaluation period, not by the forecasts. On the matched window Open-Meteo has the lower RMSE in every lead bin, and the gap widens with horizon: $7.3 \text{ W}\cdot\text{m}^{-2}$ at 0–6 h, $6.9 \text{ W}\cdot\text{m}^{-2}$ at 24–48 h, and $33.4 \text{ W}\cdot\text{m}^{-2}$ at 120–168 h (Table 4). These are point estimates without confidence intervals, and they are reported as a consistent sign rather than as a resolved difference. This apparent consistency in the meteorological domain contrasts with the propagated energy results below.

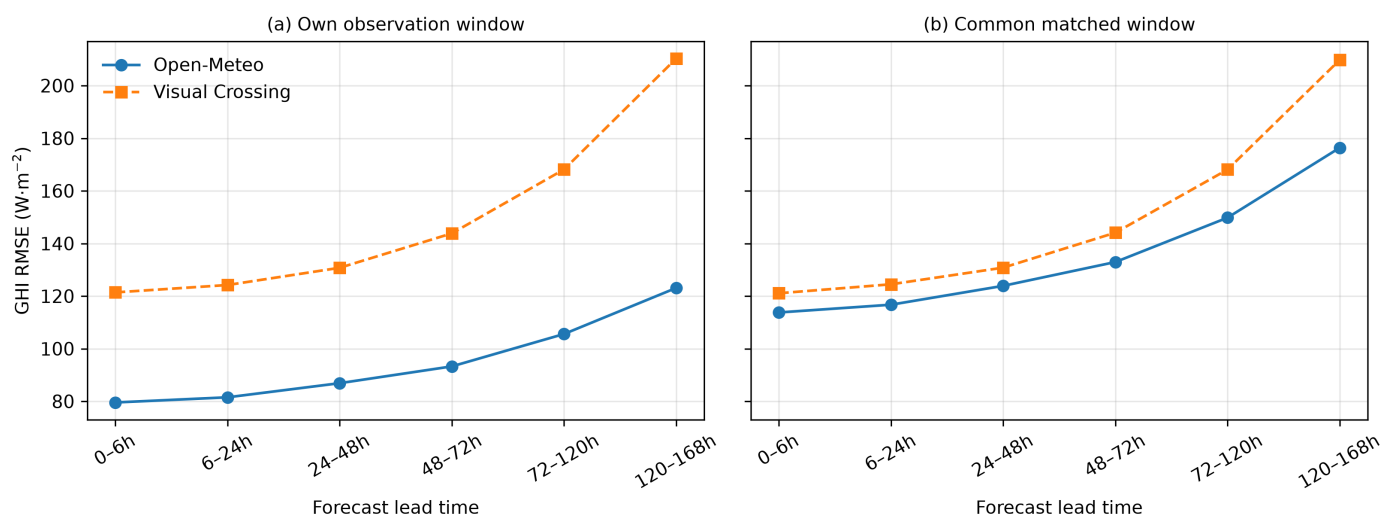


Figure 2. GHI RMSE versus lead time against the CAMS reference. (a) Each provider over its own observation window; these curves are season-confounded and describe per-provider degradation patterns rather than a comparison. (b) Both providers over the 62 common matched days, the view that supports a direct comparison. Matched-pair counts per bin are given in Table 4.

Table 4. Lead-time GHI RMSE ($\text{W}\cdot\text{m}^{-2}$) against the CAMS reference, with the number of matched hourly pairs N per bin. Own-window values use each provider’s full record and are season-confounded; matched-window values use the 62 common days and are directly comparable. N is the smaller of the two providers’ counts.

Lead Bin	Own Window		Matched Window		N
	OM	VC	OM	VC	
0–6 h	79.6	121.4	113.8	121.1	8268
6–24 h	81.5	124.2	116.7	124.5	25,128
24–48 h	86.9	130.8	123.9	130.8	33,113
48–72 h	93.3	143.8	132.9	144.1	31,806
72–120 h	105.6	168.2	149.8	168.1	59,396
120–168 h	123.1	210.3	176.4	209.8	40,515

For air temperature and wind speed, available from most evaluated services, the spread across providers is smaller than for irradiance (Table 5). Day-ahead temperature RMSE against NASA POWER is between about 1.8 and 2.6 °C; most providers carry a cold bias (down to about −2.1 °C), while Open-Meteo carries a warm bias (+0.7 °C). Wind-speed RMSE is between about 1.2 and 1.6 $\text{m}\cdot\text{s}^{-1}$ for providers with usable wind fields, with a general tendency to under-forecast wind. Two cautions apply. First, each provider is characterized over its own collection window, so the cross-provider ordering is indicative only. Second, NASA POWER is a coarse ($\sim 0.5^\circ$) MERRA-2-based reference that cannot arbitrate sub-degree temperature differences at an urban site; the cold-bias-versus-warm-bias pattern may partly reflect the reference’s spatial resolution and its blindness to the urban heat island rather than true forecast bias. These variables feed PV cell temperature, but the propagated energy comparison is dominated by irradiance availability and discrepancy.

Table 5. Day-ahead temperature and wind discrepancy against NASA POWER, computed on the same protocol as Table 6 (18–30 h band, forecast nearest 24 h), with the number of matched forecast-reference pairs (N) per provider. Windows differ across providers, so each row characterizes a provider individually and the cross-provider ordering is indicative only. World Weather Online exposes no usable wind field. NASA POWER is a $\sim 0.5^\circ$ MERRA-2-based reference and cannot resolve sub-degree differences at an urban site.

Source	T RMSE (°C)	T MBE (°C)	Wind RMSE ($\text{m}\cdot\text{s}^{-1}$)	Wind MBE ($\text{m}\cdot\text{s}^{-1}$)	N (Pairs)
Open-Meteo	2.19	+0.72	1.62	−0.62	4955
Visual Crossing	1.88	−0.73	1.39	−0.84	1414
Meteosource	2.09	−1.51	1.33	−0.82	1386
OpenWeatherMap	2.21	−1.21	1.21	−0.30	471
Tomorrow.io	1.79	−0.47	1.53	−1.00	1411
WeatherAPI	2.54	−2.07	1.28	−0.42	1415
World Weather Online	2.56	−2.09	–	–	1396

Table 6. Day-ahead GHI discrepancy against CAMS. Own-window values use each provider’s available period; matched-window values use the common 62-day period.

Source	Own-Window RMSE ($\text{W}\cdot\text{m}^{-2}$)	Own-Window MBE ($\text{W}\cdot\text{m}^{-2}$)	Matched RMSE ($\text{W}\cdot\text{m}^{-2}$)	Matched 95% CI ($\text{W}\cdot\text{m}^{-2}$)	Matched Days
Open-Meteo	84.0	−3.3	121	[111, 132]	62
Visual Crossing	128.7	−11.1	128	[117, 140]	62

4.2. Propagated PV-Power Discrepancy

Driving the common 548 kWp (DC)/500 kW (AC) pvlib system with each GHI source's day-ahead forecast and comparing the resulting power with the reference-driven power over the matched window (the 62 days common to both sources) yields similar daily-energy discrepancies for both providers (Figure 3). Daily-energy MAE is reported as a normalized MAE: the sum of absolute per-day energy discrepancies divided by the sum of reference daily energies over the evaluated days, equivalent to the mean absolute daily-energy error expressed as a percentage of mean reference daily energy. On this definition, the day-ahead daily-energy MAE is 10.9% for Open-Meteo, with a 95% CI of [7.7, 16.2], and 10.1% for Visual Crossing, with a 95% CI of [6.7, 15.4]. The mean reference-driven daily energy over the matched window is approximately 2450 kWh day⁻¹; the winter-inclusive 13% value quoted for Open-Meteo in the Discussion uses its own window's reference denominator, not this matched-window value.

The paired difference in daily-energy MAE, Open-Meteo minus Visual Crossing, is +0.8 percentage points with a 95% CI of [−1.4, +2.8]. It is essentially unchanged, and still spans zero, when the day-block bootstrap uses 3-, 5-, or 7-day blocks instead of single days. Because the interval includes zero, the two sources are not statistically distinguishable in propagated day-ahead daily-energy MAE over this sample. The point estimate nonetheless favors Visual Crossing, the provider with the larger hourly GHI discrepancy.

The sign and size of the energy bias also differ, reported here as a point-estimate diagnosis without attached confidence intervals. Open-Meteo systematically under-forecasts daily energy, with MBE of approximately −76 kWh day⁻¹, whereas Visual Crossing is far less biased, at approximately −14 kWh day⁻¹. On the matched window, the GHI bias is coherent with this energy bias: Open-Meteo's matched-window GHI MBE (−15.7 W·m⁻²) is more negative than Visual Crossing's (−8.6 W·m⁻²), in the same direction as its larger daily-energy under-forecast. The less-biased appearance in Table 6 (−3.3 versus −11.1 W·m⁻²) is an own-window artifact, because Open-Meteo's record includes long low-irradiance winter nights that dilute its daytime bias. Meteorological and energy bias therefore agree once reported over the same window.

On the matched 62-day window, the hourly power-profile RMSE is 50 kW for Open-Meteo (95% CI [42, 61]) and 54 kW for Visual Crossing ([45, 65]) on the 500 kW AC plant, and the two overlap. Open-Meteo's smaller own-window value (approximately 34 kW) reflects the low-power winter days in its longer record. Matching the window removes that apparent advantage and brings the two providers close, consistent with the daily-energy result.

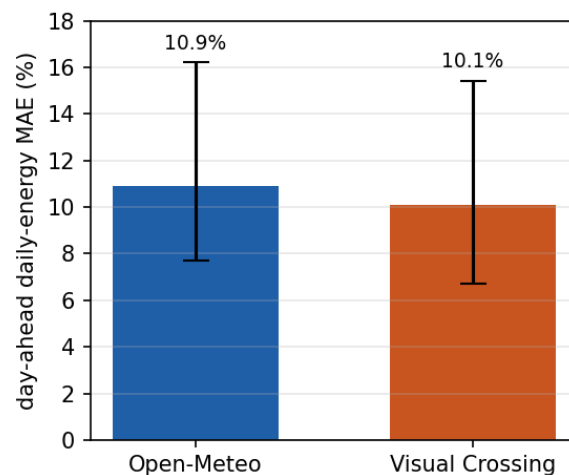


Figure 3. Day-ahead PV daily-energy MAE on the matched 62-day window. Error bars indicate 95% day-block bootstrap confidence intervals.

Because daily energy is summed over the available matched hours, the completeness of each day matters. Over the 62 matched days, 56 days have a full 24 matched hours in both providers, one day has 22, and five days fall below the 20 h threshold applied to the own-window series. The corresponding own-window figures are 202 complete days of 205 evaluated for Open-Meteo and 56 of 57 for Visual Crossing. Position matters more than count, since a missing midday hour removes far more energy than a missing night hour. Measured as a share of the reference daily energy, the hours missing from an incomplete day carry a median of 0.0% for Open-Meteo and 4.2% for Visual Crossing, so the gaps fall predominantly in low-power hours.

The daily-energy comparison is insensitive to these choices. Table 7 repeats it over three nested subsets, from all common days to days with a full 24 matched hours in both providers. The nMAE drifts by 0.4 percentage points at most, the paired difference stays between +0.69 and +0.80 percentage points, the ordering is unchanged, and every paired interval includes zero.

Table 7. Day-ahead daily-energy discrepancy over three nested completeness criteria. Paired differences (Open-Meteo – Visual Crossing) carry 95% day-block bootstrap CIs. The first row is the headline window used elsewhere in the paper.

Subset	D	OM nMAE (%)	VC nMAE (%)	Paired Diff. (pp)
All common days	62	10.9 [7.7, 16.2]	10.1 [6.7, 15.4]	+0.80 [−1.4, +2.8]
≥20 h in both	57	11.1 [7.7, 16.7]	10.4 [6.7, 16.5]	+0.69 [−1.6, +2.8]
24/24 h in both	56	11.3 [7.8, 16.8]	10.5 [6.8, 16.4]	+0.77 [−1.6, +3.0]

The same comparison under the fixed gate-closure protocol of Section 3.5 is given in Table 8. Both protocols are evaluated on the 59 days for which a run is available near the gate. Moving to a single issue time raises both providers' daily-energy nMAE, to 11.5% for Open-Meteo and 10.9% for Visual Crossing, which is expected because the effective lead is longer than 24 h for most of the target day. The paired difference narrows from +0.78 to +0.62 percentage points and continues to include zero.

The substantive point is that the pattern is reproduced. Under the operationally exact protocol, Open-Meteo still has the lower hourly GHI RMSE, 119 against 129 $\text{W}\cdot\text{m}^{-2}$, and still the higher daily-energy nMAE. The decoupling between the meteorological and the propagated-energy domain is therefore not an artefact of the lead-controlled selection rule, and neither protocol resolves a difference between the two providers.

Table 8. Day-ahead results under both selection protocols, on the same 59 matched days. The nearest-to-24 h protocol holds lead time near 24 h but mixes issue times; the fixed-gate protocol takes the whole target day from the run issued at or before 12:00 CET on D−1, with the effective lead spanning approximately 15–38 h. Paired 95% day-block bootstrap CIs; both include zero.

Protocol	OM nMAE (%)	VC nMAE (%)	Paired Diff. (pp)	OM GHI RMSE	VC GHI RMSE
Nearest to 24 h	11.1 [7.9, 16.5]	10.3 [6.8, 15.9]	+0.78 [−1.4, +2.8]	122	129
Fixed gate closure	11.5 [8.2, 17.0]	10.9 [7.2, 16.9]	+0.62 [−1.6, +2.7]	119	129

4.3. Meteorological Ranking Versus Propagated Energy Ranking

The central result is a decoupling between the meteorological and propagated-energy domains. Open-Meteo’s hourly GHI advantage is large and stable with lead time on its own record and consistent in sign on the matched window, yet it does not carry through to day-ahead daily PV energy: the point-estimate ordering changes in favor of Visual Crossing, and the paired difference includes zero. The mechanism is daily aggregation. Daily energy depends on the net daily irradiance-to-power contribution, so positive and negative hourly deviations partly cancel; Open-Meteo’s systematic under-forecast survives this cancellation and accumulates into a daily-energy bias of about -67 kWh day^{-1} , while Visual Crossing’s larger but more symmetric hourly scatter cancels almost completely. Hourly irradiance RMSE alone is therefore not sufficient to predict day-ahead energy performance; because aggregation rewards low bias over low scatter, the bias and its diurnal structure matter at least as much as point-wise precision.

To locate where the ranking changes, Table 9 follows the two sources through the chain on the matched 62-day window, from hourly GHI to day-ahead AC energy. The ordering change is already present at *daily* GHI energy—Open-Meteo’s normalized daily-GHI MAE (9.4%) exceeds Visual Crossing’s (8.3%)—before any plane-of-array transposition or PV conversion. Propagation through the PV chain slightly *narrows* the point-estimate gap (from +1.1 to +0.8 percentage points) rather than creating it. As in Section 4.2, no stage-wise difference is statistically resolved: every paired 95% CI includes zero (Table 9). This is a change in point-estimate ordering, not a demonstrated reversal. The change is therefore primarily a consequence of moving from an hourly point-error metric (RMSE) to a daily energy-volume metric (normalized MAE), in which low bias is rewarded over low scatter; the irradiance-to-power model is a secondary modifier at this plant’s low tilt and DC/AC ratio.

Table 9. Error ladder over the matched window—the same 62 common days and matched timestamps at every stage—giving Open-Meteo and Visual Crossing discrepancy against the reference from hourly GHI to day-ahead AC energy. Daily quantities are simple sums of the matched hourly values; normalized MAE (nMAE) is the sum of absolute per-day discrepancies divided by the sum of reference daily values. Paired differences (Open-Meteo–Visual Crossing) carry 95% day-block bootstrap CIs; every interval includes zero, so the point-estimate ordering change at daily GHI is not statistically resolved. The change appears before PV conversion.

Evaluation Stage	Open-Meteo	Visual Crossing	Difference	95% CI (Diff.)
Hourly GHI RMSE ($\text{W}\cdot\text{m}^{-2}$)	121	128	−7	[−16, +2]
Daily GHI energy nMAE (%)	9.4	8.3	+1.1	[−1.0, +3.2]
Daily plane-of-array energy nMAE (%)	11.4	10.3	+1.0	[−1.2, +3.0]
Daily DC-energy nMAE (%)	10.9	10.0	+0.8	[−1.2, +2.9]
Daily AC-energy nMAE (%)	10.9	10.1	+0.8	[−1.4, +2.9]

4.4. Complementary Self-Consistency Analysis

The reference-free nearest-cast analysis for Open-Meteo supports the lead-time pattern by an independent route. Forecast drift grows with lead time for all evaluated parameter groups (Figure 4). The day-ahead GHI consistency MAE is approximately $33 \text{ W}\cdot\text{m}^{-2}$, of the

same order as Open-Meteo's day-ahead GHI MAE against the CAMS reference over its own window, approximately $44 \text{ W} \cdot \text{m}^{-2}$. The diurnal pattern in Figure 5 shows that irradiance drift concentrates in daylight hours and peaks near solar noon, when the PV plant is generating. DNI is less self-consistent than GHI, while GTI behaves more similarly to GHI. These results quantify forecast drift rather than absolute accuracy, so they are interpreted as a robustness check rather than as a replacement for reference-based verification.

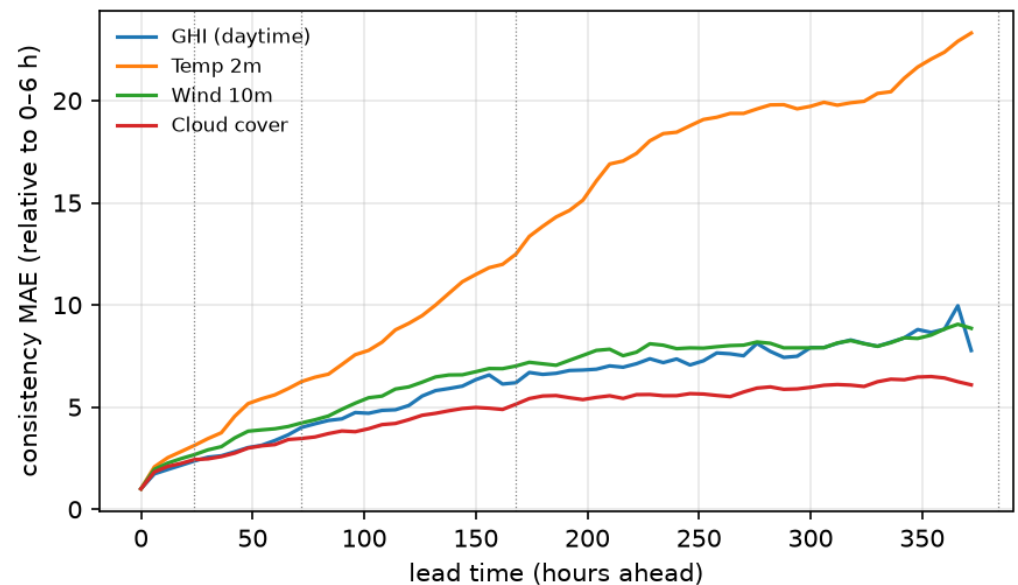


Figure 4. Forecast self-consistency error growth with lead time. Each curve is the consistency MAE at a given lead divided by its value in the 0–6 h bin (a dimensionless growth ratio starting at 1.0); solar variables use daytime hours only.

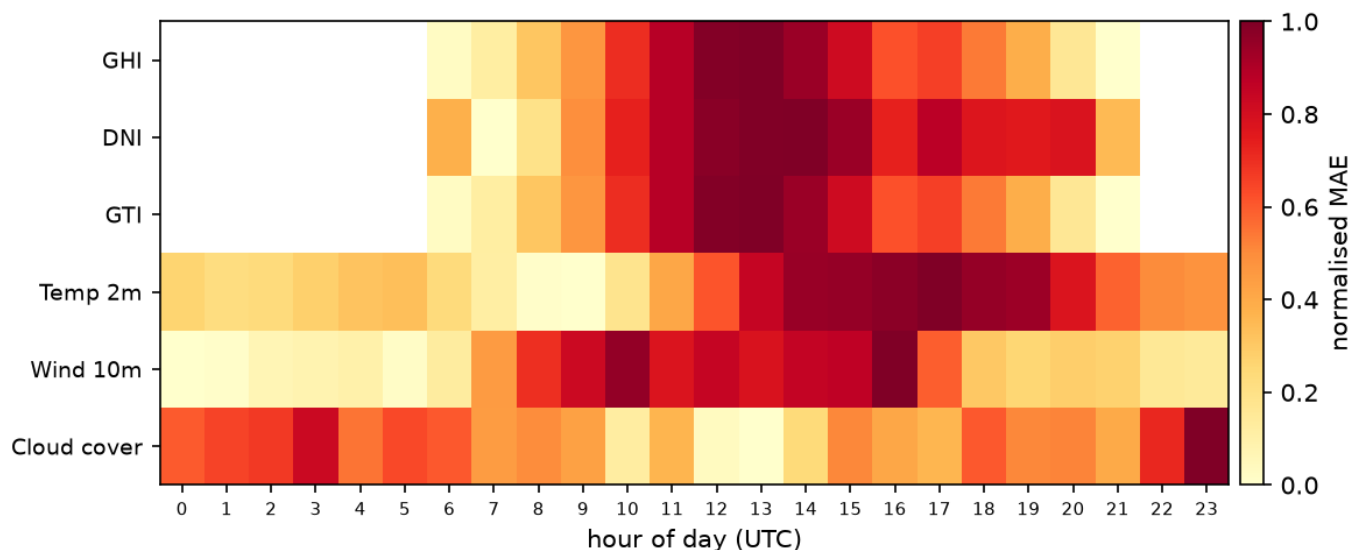


Figure 5. Diurnal forecast self-consistency based on the Open-Meteo nearest-cast comparison. Cell values are the mean absolute nearest-cast difference (MAE) by hour of day, min–max normalized within each row (each parameter rescaled to [0, 1] across the 24 h, so 0 marks that parameter's quietest hour and 1 its peak); solar variables use daytime hours only. Irradiance drift peaks during daylight and near solar noon.

5. Discussion

5.1. Do Better Irradiance Forecasts Yield Better Day-Ahead PV Energy Forecasts?

In this dataset, no. Intercomparisons of NWP models and commercial providers conventionally rank sources by an irradiance error metric [14,15]. On that basis Open-Meteo is the conventionally ranked winner here: its own-record day-ahead GHI RMSE is 84 against 129 $\text{W}\cdot\text{m}^{-2}$, and the point-estimate ordering holds at every lead time. That ranking rests on own-window metrics, however, and is itself season-confounded. Once the same forecasts are propagated through an identical PV model and integrated to day-ahead energy—the quantity an operator schedules and settles—the ordering is no longer reproduced: the daily-energy MAEs are not statistically distinguishable over this sample, and the point estimate favors the provider with the larger hourly irradiance discrepancy. This ordering change arises because daily aggregation rewards low bias rather than low scatter, so provider selection for PV applications should rest on propagated energy and power-profile metrics, not on meteorological metrics alone. An error ladder over the matched window (Table 9) localizes the change: the ordering change is already present in daily GHI energy, before plane-of-array transposition or PV conversion, and the PV chain narrows the point-estimate gap slightly rather than creating it. The dominant cause is therefore the shift from an hourly point-error metric to a daily energy-volume metric, with the physical PV model a secondary modifier.

Two cautions bound this interpretation. First, as set out in Section 4.2, the common window contains 62 spring and early-summer days and no difference on it is statistically resolved. The result documents a rank inversion in point estimates, not evidence that Visual Crossing is the better PV input. Second, the finding is consistent with reports that upstream weather-input and transformation error dominate the PV power-forecast budget [38], but it adds a sharper point: the metric domain in which a source is judged can change the conclusion, so ranking freely accessible forecasts in $\text{W}\cdot\text{m}^{-2}$ alone can be misleading for a PV application.

5.2. Implications for Smart-City Energy Management

The results are directly relevant to smart-city energy applications in which urban actors rely on low-cost data infrastructure. A municipal energy manager, campus operator, or local energy community may select a freely accessible weather API because it is easy to integrate into an energy-management platform. This paper suggests that such a choice should be audited, in a fixed order, before the data are treated as a reliable ML input:

- (i) **Availability.** Establish which PV-relevant fields the provider actually exposes; here, only two of eight services delivered usable hourly irradiance, and one advertised irradiance field was empty. For many municipal integrators this step alone is decisive.
- (ii) **Matched-window magnitude.** Compare providers only over a common period and in the operational unit; a raw day-ahead daily-energy discrepancy of roughly 10–11% over the common spring–summer window is a non-negligible scheduling uncertainty for PV portfolios, storage operation, EV-charging coordination, and day-ahead market participation.
- (iii) **Bias-versus-scatter structure.** Characterize, in the propagated energy domain, the systematic (bias) component relative to the day-to-day scatter. In this sample, both providers are scatter-dominated at the daily scale, but Open-Meteo carries the larger systematic component ($|\text{MBE}|/\text{SD} \approx 0.2$) while Visual Crossing is close to unbiased (≈ 0.03).
- (iv) **Correct or blend.** Use that structure to decide whether to bias-correct the more systematically biased provider or to blend complementary ones. A more-biased and a near-unbiased source are in principle a combination that could favor a blend, though

on the present 62-day record, this remains a point-estimate observation to be confirmed on a longer series.

The same audit outputs—systematic bias to correct before ML training, complementary error structures, and the time-of-day and lead-time features worth passing downstream—are the natural inputs to a smart-city digital twin [39]. Framed this way, the present study is the pre-commissioning stage of such a twin for the campus plant, to be revalidated against measured output once the plant is operating.

5.3. Magnitude and Interpretation of Day-Ahead Discrepancy

The approximately 10–11% day-ahead daily-energy MAE over the matched window should be interpreted as raw, uncorrected, reference-driven PV-energy discrepancy. It is not an unavoidable lower bound on future ML performance. A downstream model with sufficient historical target data may learn bias corrections, temporal patterns, or provider-specific error structures. However, the result quantifies the level of input-driven uncertainty that such a model must confront. Including winter months available for Open-Meteo increases its own-window day-ahead daily-energy MAE to approximately 13%, indicating that the spring–summer value is optimistic relative to full-year operation.

These magnitudes are broadly consistent with the published literature. A day-ahead GHI RMSE in the 80–130 $\text{W}\cdot\text{m}^{-2}$ range for a Central European site is comparable to the spread reported in multi-model NWP irradiance intercomparisons over Europe [14,15], and a day-ahead PV-energy MAE near 10% is of the order reported in physical-model PV-forecast error budgets [8,38]. The audit therefore does not flag freely accessible data as anomalously poor; it locates these APIs within the same error band as more elaborate inputs, which is precisely why the propagated, operational-unit comparison—rather than the headline $\text{W}\cdot\text{m}^{-2}$ figure—is the informative one.

In concrete terms, for the 548 kWp campus plant, a 10–11% day-ahead daily-energy discrepancy is on the order of 250–270 kWh day^{-1} of scheduling uncertainty against a mean of about 2450 kWh day^{-1} , and the matched-window hourly power-profile RMSE of about 50 kW is roughly one-tenth of the 500 kW inverter rating. These magnitudes are non-negligible for sizing storage reserve, self-consumption targeting, and EV-charging coordination, although converting them into a specific reserve or market bid requires a defined storage model and pricing scheme and is left to operational follow-up work.

Market relevance is discussed in energy terms rather than currency because monetizing the discrepancy would require a specific imbalance-pricing scheme, bidding strategy, and regulatory framework. The present paper therefore frames market relevance as operational exposure rather than as a financial settlement calculation.

5.4. Practical Consequences for ML-Based Forecasting

The practical recommendation is that weather-input characterization should precede ML model training. If the raw input has a known lead-time-dependent bias or higher uncertainty during certain hours, the ML model should either correct for it, weight it accordingly, or use it as an uncertainty feature. Treating the weather API as ground truth hides this information. Likewise, if two freely accessible weather services show different bias structures, a simple bias correction or source blend may improve the input before complex ML architectures are considered.

5.5. Limitations

Several limitations bound the interpretation of the results.

First, the temporal record is short and incomplete. Open-Meteo spans late November 2025 to late June 2026, while the other services mainly cover mid-April to late June 2026.

The matched propagated comparison contains 62 common days. The results therefore demonstrate the framework and provide indicative magnitudes, not complete climatological conclusions.

The direction of the seasonal effect can be bounded using the record itself. The matched window falls in spring and early summer, a high-irradiance regime dominated by convective cloud. Autumn and winter in Bratislava bring low solar elevation and persistent stratiform cloud, which reduce the day-to-day variance of GHI and therefore deflate absolute $\text{W}\cdot\text{m}^{-2}$ error. The size of that effect is visible in our own data: the same Open-Meteo forecasts give a day-ahead GHI RMSE of $84.0 \text{ W}\cdot\text{m}^{-2}$ over the winter-inclusive own window against $121 \text{ W}\cdot\text{m}^{-2}$ over the matched spring–summer window. The absolute magnitudes reported here are thus close to a seasonal upper envelope. The relative daily-energy error could move in either direction, because fog and stratus constitute a different error regime from convective cloud and the aggregation cancellation depends on the bias structure rather than the scatter alone. The aggregation mechanism itself—daily summation rewarding low bias over low scatter—is not season-specific. The full-year rerun will test the point estimates.

Second, the references are not site measurements, and they are not perfectly independent of the forecasts. CAMS and NASA POWER are satellite/reanalysis-based references, so the reported errors are forecast-reference discrepancies. Their own hourly error is substantial—approximately $50\text{--}65 \text{ W}\cdot\text{m}^{-2}$ for CAMS GHI and $2.47 \text{ m}\cdot\text{s}^{-1}$ for NASA POWER wind (Section 3.2)—which places an upper bound on the absolute magnitudes reported here, although it largely cancels in the paired provider comparison. Moreover, Open-Meteo and the CAMS radiation chain share an ECMWF lineage (discussed above), a residual coupling that would, if anything, flatter Open-Meteo’s meteorological scores and is therefore conservative for the paper’s central caution. Once the plant is commissioned, the framework can be rerun against measured plant output.

Third, the propagated PV comparison includes only two irradiance-capable providers. The other collected services are relevant for a data-availability audit and for temperature, wind, and cloud comparisons, but they cannot be directly propagated to PV power without deriving irradiance from cloud-cover variables. This extension is left for future work.

Fourth, the PV system is the as-designed configuration from the project documentation, not yet a measured plant. The model aggregates the four inverters into a single PVWatts inverter and represents the module-level optimizers as a fixed loss component, and the absolute power and energy values remain simulation-based until commissioning. Because every source is propagated through the identical configuration, the relative spread among sources is robust to these simplifications, and the configuration file can be updated against as-built and measured data without code changes.

Fifth, PV-model sensitivities are not explored in this manuscript. The transposition model, decomposition model, loss assumptions, and thermal parameters could affect the propagated discrepancy; the array geometry and DC/AC ratio are fixed by the plant design but would vary across other plants. A sensitivity analysis is planned before final operational use of the framework.

Sixth, the analysis covers a single urban site in Central Europe. The framework generalizes to other cities and climates, but the numerical discrepancy magnitudes should not be transferred without rerunning the analysis locally. The same caution applies to the plant configuration and not only to the climate. The 12° east–west geometry and the DC/AC ratio of 1.10 make inverter clipping rare and limit the influence of the transposition step. A steeper south-facing array with a higher DC/AC ratio would weight those stages differently, so the propagated magnitudes reported here should not be transferred to a differently configured plant without rerunning the chain.

6. Conclusions

This paper presented a transparent, repeatable framework for auditing freely accessible weather-forecast inputs and propagating their reference-based discrepancies into day-ahead PV power. Its first result is a data-availability finding: of the eight collected weather services, and for the endpoints, subscription tiers, and collection period evaluated here, only two exposed usable hourly irradiance and one advertised irradiance field was empty—for a municipal integrator, often the most actionable outcome. Applied to the 548 kWp urban rooftop PV plant being built in Bratislava, the framework showed that for those two irradiance-capable sources, the raw day-ahead daily PV-energy discrepancy over the common 62-day spring–summer window was approximately 10–11%. Open-Meteo had the lower hourly GHI RMSE on its own, season-confounded record, but after propagation through the PV model, the daily-energy point estimate favored the other provider and the paired difference included zero.

Returning to the research questions: on RQ1, only two of the seven forecast services delivered usable hourly irradiance under the conditions evaluated. On RQ2, the day-ahead discrepancy is 10–11% of daily energy over the matched window, against 121 and 128 $\text{W}\cdot\text{m}^{-2}$ of hourly GHI RMSE in the meteorological domain. On RQ3, the irradiance-based ranking did not survive propagation and aggregation: the point-estimate ordering changed, and no difference was statistically resolved. The working hypothesis behind RQ3—that an hourly irradiance ranking transfers to day-ahead energy—is therefore not supported by this record, and the same conclusion holds under a fixed gate-closure protocol.

The main conclusion is not that one provider is universally superior. The study shows that meteorological forecast rankings need not survive aggregation into the operational quantity—here the apparent hourly irradiance advantage did not carry through and the point-estimate ordering even changed. An error ladder localizes this change to daily GHI aggregation, before the PV model, so it reflects the metric domain (daily energy volume versus hourly point error) rather than the irradiance-to-power conversion, which at this plant's low tilt and DC/AC ratio, only modestly modifies the gap. For smart-city energy management, this means that weather APIs should be evaluated in the operational metric domain that matters: day-ahead power profiles and daily energy, not only irradiance in $\text{W}\cdot\text{m}^{-2}$.

Future work will extend the record to a full year, replace reference-driven PV power with measured plant output, add sensitivity analysis over PV-system configuration and irradiance decomposition, derive irradiance from cloud-cover data to include the remaining providers in the power comparison, and test bias correction and provider blending as low-cost improvements before ML model training.

Author Contributions: Conceptualization, M.C.; methodology, M.C.; software, M.C.; validation, M.C.; formal analysis, M.C.; investigation, M.C.; data curation, M.C.; writing—original draft preparation, M.C.; writing—review and editing, J.B. and P.J.; supervision, J.B. and P.J. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Faculty of Electrical Engineering and Information Technology of the Slovak University of Technology in Bratislava, Slovakia under the contract No. FEI-VZ-PIT-2025-04.

Data Availability Statement: The analysis code, the plant configuration file (`pv_config.yaml`), the field dictionary and reproduction guide (`REPRODUCE.md`), the environment specification, and the processed matched datasets (matched hourly forecast–reference series and per-day energy series) are openly available and archived at <https://doi.org/10.5281/zenodo.21379666>. The raw forecast archive is provided as retrieval scripts and a file manifest with checksums where provider licensing permits redistribution. Every reported table and figure is regenerated from the archived dataset by the scripts documented in `REPRODUCE.md`.

Acknowledgments: The authors acknowledge the use of freely accessible weather and reference datasets described in the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

API	Application programming interface
CAMS	Copernicus Atmosphere Monitoring Service
CI	Confidence interval
DNI	Direct normal irradiance
GHI	Global horizontal irradiance
GTI	Global tilted irradiance
MAE	Mean absolute error
MBE	Mean bias error
ML	Machine learning
NWP	Numerical weather prediction
PV	Photovoltaic
RMSE	Root mean square error

Appendix A. Implementation Reference

The modeling chain of Section 3.4 is implemented with pvlib python 0.15.2 under Python 3.10, with pandas 2.3.3, numpy 2.2.6, and pyarrow 25.0.0. All bootstrap resampling uses seed 0. Table A1 lists the function called at each stage. The complete parameter set is held in `pv_config.yaml` and applied identically to every source and to the reference; both the configuration and the analysis scripts are archived at the DOI given in the Data Availability statement.

Table A1. pvlib functions used at each stage of the modeling chain (pvlib 0.15.2).

Stage	pvlib Call
Solar position	<code>solarposition.get_solarposition</code>
Decomposition (Erbs)	<code>irradiance.erbs</code>
Transposition (Perez)	<code>irradiance.get_total_irradiance (model = 'perez')</code>
AOI loss	<code>iam.physical</code>
Cell temperature	<code>temperature.faiman ($u_0 = 25.0$, $u_1 = 6.84$)</code>
DC power	<code>pvsystem.pvwatts_dc ($\gamma = -0.0029 \text{ K}^{-1}$)</code>
System losses	<code>pvsystem.PVSystem(losses_parameters = ...)</code>
AC power	<code>inverter.pvwatts ($P_{ac0} = 500 \text{ kW}$, $\eta_{nom} = 0.984$)</code>
Clear-sky baseline	<code>clearsky.ineichen</code>

References

1. Antonanzas, J.; Osorio, N.; Escobar, R.; Urraca, R.; Martinez-de Pison, F.J.; Antonanzas-Torres, F. Review of photovoltaic power forecasting. *Sol. Energy* **2016**, *136*, 78–111. <https://doi.org/10.1016/j.solener.2016.06.069>.
2. Yang, D.; Wang, W.; Gueymard, C.A.; Hong, T.; Kleissl, J.; Huang, J.; Perez, M.J.; Perez, R.; Bright, J.M.; Xia, X.; et al. A review of solar forecasting, its dependence on atmospheric sciences and implications for grid integration: Towards carbon neutrality. *Renew. Sustain. Energy Rev.* **2022**, *161*, 112348. <https://doi.org/10.1016/j.rser.2022.112348>.
3. Markovics, D.; Mayer, M.J. Comparison of machine learning methods for photovoltaic power forecasting based on numerical weather prediction. *Renew. Sustain. Energy Rev.* **2022**, *161*, 112364. <https://doi.org/10.1016/j.rser.2022.112364>.
4. Mayer, M.J. Benefits of physical and machine learning hybridization for photovoltaic power forecasting. *Renew. Sustain. Energy Rev.* **2022**, *168*, 112772. <https://doi.org/10.1016/j.rser.2022.112772>.
5. Sharma, S.; Dua, A.; Singh, M.; Kumar, N.; Prakash, S. Fuzzy rough set based energy management system for self-sustainable smart city. *Renew. Sustain. Energy Rev.* **2018**, *82*, 3633–3644. <https://doi.org/10.1016/j.rser.2017.10.099>.

6. Yuan, J.; Farnham, C.; Emura, K.; Lu, S. A method to estimate the potential of rooftop photovoltaic power generation for a region. *Urban Clim.* **2016**, *17*, 1–19. <https://doi.org/10.1016/j.uclim.2016.03.001>.
7. Antonanzas, J.; Pozo-Vázquez, D.; Fernandez-Jimenez, L.A.; Martinez-de Pison, F.J. The value of day-ahead forecasting for photovoltaics in the Spanish electricity market. *Sol. Energy* **2017**, *158*, 140–146. <https://doi.org/10.1016/j.solener.2017.09.043>.
8. Markovics, D.; Mayer, M.J. Cost of photovoltaic power forecast errors based on machine learning, numerical weather prediction and real-time measurements. In *Proceedings of the 2025 IEEE 53rd Photovoltaic Specialists Conference (PVSC)*; IEEE: New York, NJ, USA, 2025; pp. 0703–0708. <https://doi.org/10.1109/PVSC59419.2025.11133201>.
9. Antonanzas, J.; Perpiñán-Lamigueiro, O.; Urraca, R.; Antonanzas-Torres, F. Influence of electricity market structures on deterministic solar forecasting verification. *Sol. Energy* **2020**, *210*, 44–46. <https://doi.org/10.1016/j.solener.2020.04.017>.
10. Yang, D.; Alessandrini, S.; Antonanzas, J.; Antonanzas-Torres, F.; Badescu, V.; Beyer, H.G.; Blaga, R.; Boland, J.; Bright, J.M.; Coimbra, C.F.M.; et al. Verification of deterministic solar forecasts. *Sol. Energy* **2020**, *210*, 20–37. <https://doi.org/10.1016/j.solener.2020.04.019>.
11. Zhang, J.; Florita, A.; Hodge, B.M.; Lu, S.; Hamann, H.F.; Banunarayanan, V.; Brockway, A.M. A suite of metrics for assessing the performance of solar power forecasting. *Sol. Energy* **2015**, *111*, 157–175. <https://doi.org/10.1016/j.solener.2014.10.016>.
12. Yang, D. Making reference solar forecasts with climatology, persistence, and their optimal convex combination. *Sol. Energy* **2019**, *193*, 981–985. <https://doi.org/10.1016/j.solener.2019.10.006>.
13. Vallance, L.; Charbonnier, B.; Paul, N.; Dubost, S.; Blanc, P. Towards a standardized procedure to assess solar forecast accuracy: A new ramp and time alignment metric. *Sol. Energy* **2017**, *150*, 408–422. <https://doi.org/10.1016/j.solener.2017.04.064>.
14. Perez, R.; Lorenz, E.; Pelland, S.; Beauharnois, M.; Van Knowe, G.; Hemker, K., Jr.; Heinemann, D.; Remund, J.; Müller, S.C.; Traummüller, W.; et al. Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe. *Sol. Energy* **2013**, *94*, 305–326. <https://doi.org/10.1016/j.solener.2013.05.005>.
15. Lorenz, E.; Kühnert, J.; Heinemann, D.; Nielsen, K.P.; Remund, J.; Müller, S.C. Comparison of global horizontal irradiance forecasts based on NWP models with different spatio-temporal resolutions. *Prog. Photovoltaics: Res. Appl.* **2016**, *24*, 1626–1640. <https://doi.org/10.1002/pip.2799>.
16. Kumari, P.; Toshniwal, D. Deep learning models for solar irradiance forecasting: A comprehensive review. *J. Clean. Prod.* **2021**, *318*, 128566. <https://doi.org/10.1016/j.jclepro.2021.128566>.
17. Iheanetu, K.J. Solar photovoltaic power forecasting: A review. *Sustainability* **2022**, *14*, 17005. <https://doi.org/10.3390/su142417005>.
18. Marquez, R.; Coimbra, C.F.M. Proposed metric for evaluation of solar forecasting models. *J. Sol. Energy Eng.* **2013**, *135*, 011016. <https://doi.org/10.1115/1.4007496>.
19. Mayer, M.J.; Yang, D. Calibration of deterministic NWP forecasts and its impact on verification. *Int. J. Forecast.* **2023**, *39*, 981–991. <https://doi.org/10.1016/j.ijforecast.2022.03.008>.
20. Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horányi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **2020**, *146*, 1999–2049. <https://doi.org/10.1002/qj.3803>.
21. Urraca, R.; Huld, T.; Gracia-Amillo, A.; Martinez-de Pison, F.J.; Kaspar, F.; Sanz-Garcia, A. Evaluation of global horizontal irradiance estimates from ERA5 and COSMO-REA6 reanalyses using ground and satellite-based data. *Sol. Energy* **2018**, *164*, 339–354. <https://doi.org/10.1016/j.solener.2018.02.059>.
22. Babar, B.; Graversen, R.; Boström, T. Solar radiation estimation at high latitudes: Assessment of the CMSAF databases, ASR and ERA5. *Sol. Energy* **2019**, *182*, 397–411. <https://doi.org/10.1016/j.solener.2019.02.058>.
23. Pfeifroth, U.; Sanchez-Lorenzo, A.; Manara, V.; Trentmann, J.; Hollmann, R. Trends and variability of surface solar radiation in Europe based on surface- and satellite-based data records. *J. Geophys. Res. Atmos.* **2018**, *123*, 1735–1754. <https://doi.org/10.1002/2017JD027418>.
24. Driemel, A.; Augustine, J.; Behrens, K.; Colle, S.; Cox, C.; Cuevas-Agulló, E.; Denn, F.M.; Duprat, T.; Fukuda, M.; Grobe, H.; et al. Baseline Surface Radiation Network (BSRN): Structure and data description (1992–2017). *Earth Syst. Sci. Data* **2018**, *10*, 1491–1501. <https://doi.org/10.5194/essd-10-1491-2018>.
25. Lefèvre, M.; Oumbe, A.; Blanc, P.; Espinar, B.; Gschwind, B.; Qu, Z.; Wald, L.; Schroedter-Homscheidt, M.; Hoyer-Klick, C.; Arola, A.; et al. McClear: A new model estimating downwelling solar radiation at ground level in clear-sky conditions. *Atmos. Meas. Tech.* **2013**, *6*, 2403–2418. <https://doi.org/10.5194/amt-6-2403-2013>.
26. Qu, Z.; Oumbe, A.; Blanc, P.; Espinar, B.; Gesell, G.; Gschwind, B.; Klüser, L.; Lefèvre, M.; Saboret, L.; Schroedter-Homscheidt, M.; et al. Fast radiative transfer parameterisation for assessing the surface solar irradiance: The Heliosat-4 method. *Meteorol. Z.* **2017**, *26*, 33–57. <https://doi.org/10.1127/metz/2016/0781>.
27. Gschwind, B.; Wald, L.; Blanc, P.; Lefèvre, M.; Schroedter-Homscheidt, M.; Arola, A. Improving the McClear model estimating the downwelling solar radiation at ground level in cloud-free conditions: McClear-v3. *Meteorol. Z.* **2019**, *28*, 147–163. <https://doi.org/10.1127/metz/2019/0946>.

28. Gelaro, R.; McCarty, W.; Suárez, M.J.; Todling, R.; Molod, A.; Takacs, L.; Randles, C.A.; Darmenov, A.; Bosilovich, M.G.; Reichle, R.; et al. The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Clim.* **2017**, *30*, 5419–5454. <https://doi.org/10.1175/JCLI-D-16-0758.1>.
29. Sparks, A.H. nasapower: A NASA POWER global meteorology, surface solar energy and climatology data client for R. *J. Open Source Softw.* **2018**, *3*, 1035. <https://doi.org/10.21105/joss.01035>.
30. Holmgren, W.F.; Hansen, C.W.; Mikofski, M.A. pvlib python: A python package for modeling solar energy systems. *J. Open Source Softw.* **2018**, *3*, 884. <https://doi.org/10.21105/joss.00884>.
31. Anderson, K.S.; Hansen, C.W.; Holmgren, W.F.; Jensen, A.R.; Mikofski, M.A.; Driesse, A. pvlib python: 2023 project update. *J. Open Source Softw.* **2023**, *8*, 5994. <https://doi.org/10.21105/joss.05994>.
32. Reindl, D.T.; Beckman, W.A.; Duffie, J.A. Evaluation of hourly tilted surface radiation models. *Sol. Energy* **1990**, *45*, 9–17. [https://doi.org/10.1016/0038-092X\(90\)90061-G](https://doi.org/10.1016/0038-092X(90)90061-G).
33. Perez, R.; Ineichen, P.; Seals, R.; Michalsky, J.; Stewart, R. Modeling daylight availability and irradiance components from direct and global irradiance. *Sol. Energy* **1990**, *44*, 271–289. [https://doi.org/10.1016/0038-092X\(90\)90055-H](https://doi.org/10.1016/0038-092X(90)90055-H).
34. Hay, J.E.; Davies, J.A. Calculation of the solar radiation incident on an inclined surface. In *Proceedings of the 1st Canadian Solar Radiation Data Workshop*; Ministry of Supply and Services: Ottawa, ON, Canada, 1980; p. 59.
35. Loutzenhiser, P.G.; Manz, H.; Felsmann, C.; Strachan, P.A.; Frank, T.; Maxwell, G.M. Empirical validation of models to compute solar irradiance on inclined surfaces for building energy simulation. *Sol. Energy* **2007**, *81*, 254–267. <https://doi.org/10.1016/j.solener.2006.03.009>.
36. Faiman, D. Assessing the outdoor operating temperature of photovoltaic modules. *Prog. Photovoltaics Res. Appl.* **2008**, *16*, 307–315. <https://doi.org/10.1002/pip.813>.
37. King, D.L.; Boyson, W.E.; Kratochvil, J.A. *Photovoltaic Array Performance Model*; Technical Report SAND2004-3535; Sandia National Laboratories: Albuquerque, NM, USA, 2004. <https://doi.org/10.2172/919131>.
38. Mayer, M.J.; Gróf, G. Extensive comparison of physical models for photovoltaic power forecasting. *Appl. Energy* **2021**, *283*, 116239. <https://doi.org/10.1016/j.apenergy.2020.116239>.
39. Xu, W.; Liu, S. Novel economic models for advancing urban energy management and transition: Simulation of urban energy system in digital twin. *Sustain. Cities Soc.* **2024**, *101*, 105154. <https://doi.org/10.1016/j.scs.2023.105154>.
40. Zippenfenig, P. *Open-Meteo.com Weather API*, Version 1.4.0.; Zenodo: Geneva, Switzerland, 2023. <https://doi.org/10.5281/zenodo.7970649>.
41. SoDa Service. Validation of CAMS Radiation Service Against BSRN Stations. CAMS Radiation Service Version 4.0; Validation for 2004–2020; The Current Service Uses a Newer Implementation of the Same Heliosat-4/APOLLO Framework. 2021. Available online: <https://www.soda-pro.com/en/help/validation/cams-radiation-bsrn> (accessed on 27 July 2026).
42. NASA POWER Project. Meteorological Data Assessment Process. Hourly MERRA-2 Assessment Against 2,599 NOAA/NCEI Integrated Surface Database Stations for 2001–2020. Available online: <https://power.larc.nasa.gov/docs/methodology/meteorology/assessment/> (accessed on 27 July 2026).
43. Ineichen, P.; Perez, R. A new airmass independent formulation for the Linke turbidity coefficient. *Sol. Energy* **2002**, *73*, 151–157. [https://doi.org/10.1016/S0038-092X\(02\)00045-2](https://doi.org/10.1016/S0038-092X(02)00045-2).
44. Erbs, D.G.; Klein, S.A.; Duffie, J.A. Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation. *Sol. Energy* **1982**, *28*, 293–302. [https://doi.org/10.1016/0038-092X\(82\)90302-4](https://doi.org/10.1016/0038-092X(82)90302-4).
45. Dobos, A.P. *PVWatts Version 5 Manual*; Technical Report NREL/TP-6A20-62641; National Renewable Energy Laboratory: Golden, CO, USA, 2014.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.