Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

HelioTR-Atlas: coupling measured satellite direct irradiance with station-validated topographic thermodynamics for explainable photovoltaic yield assessment in complex terrain

Ertuğrul Gül

ertugrulgul@hakkari.edu.tr

Hakkari University

Article

Keywords: Photovoltaic yield modelling, Satellite direct irradiance, Topographic downscaling, Lapse rate, Reanalysis validation, Explainable machine learning

Posted Date: August 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10403998/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on November 13th, 2026. See the published version at https://doi.org/10.1038/s41598-026-75440-0.

Abstract

Terrain-resolved photovoltaic (PV) yield assessment needs three ingredients that are rarely combined: an irradiance source whose direct component is measured rather than assumed, thermodynamic fields corrected for elevation and validated against ground stations, and an attribution layer explaining why yield departs from irradiance. We present HelioTR-Atlas, a reproducible framework coupling SARAH-3 satellite irradiance, whose direct component is retrieved rather than inferred, with lapse-rate-corrected ERA5-Land temperature, dew point and wind, propagated through an anisotropic transposition, a physical cell-temperature model and an efficiency model calibrated to a 325 W module. The retrieved direct component avoids the Erbs diffuse-fraction assumption, which overestimated clear-sky direct irradiance by 22 W/m^2. Lapse-rate correction reduced ERA5-Land air-temperature RMSD against 62 stations from 3.25 to 2.52 degrees C, an improvement of 22.5%; dew point and wind were evaluated against a further 119 stations. Applied to Türkiye over 2000–2025, the framework gives a climatological mean yield of 1805 kWh/kWp across 32,479 land cells, an inter-annual coefficient of variation of 3.3%, and provincial means from 1334 kWh/kWp on the eastern Black Sea coast to 1994 kWh/kWp in the southeast. SHAP attribution separates two regimes: irradiance sets the spatial pattern of yield, whereas air temperature governs the performance ratio, most strongly at elevation extremes.

1. Introduction

Solar photovoltaics are central to Türkiye's energy transition and to its commitments on climate change, and the pace of deployment has made reliable, high-resolution maps of achievable PV yield a practical planning need rather than an academic one. Yield, however, is not set by irradiance alone. It also depends on air temperature, wind, and the terrain that modulates both, so a credible national atlas must couple satellite irradiance with surface thermodynamics at a resolution that resolves Türkiye's steep topography.

Existing resources fall short of this requirement in different ways. Recent Turkish studies have mapped solar suitability with GIS and multi-criteria analysis for site selection [1] or assessed the resource for downstream applications such as green hydrogen [2], rather than producing a validated PV-yield atlas. At the country and global scale, recent reanalysis-based assessments estimate PV potential or its future variability [3, 4], but they do not combine station-validated topographic thermodynamics with an independent satellite radiation source. Recent satellite-irradiance models based on deep learning target short-term forecasting rather than a climatological yield atlas [5], and recent explainable-ML studies apply attribution to site-specific power forecasting rather than to the spatial pattern of a national atlas [6]. Country-scale evaluations of solar-radiation datasets against ground measurements exist [7], yet they validate irradiance rather than PV yield. Reanalysis products have in fact been evaluated over Türkiye, and this study does not claim to validate them for the first time; dense station comparisons already exist for precipitation, with 256 stations reported by Amjad et al. [8], and for air temperature, with more than a thousand stations used by Karaman and Akyürek [9].

What has not been done is the combination that a PV yield atlas actually requires. A survey of a large reanalysis literature corpus found no country-scale validation at the scale of surface solar radiation or PV yield over Türkiye, no station-based evaluation of ERA5-Land dew point and wind, and no explainable attribution of the drivers of spatial yield differences; these three elements have never been brought together. We address this by developing HelioTR-Atlas, a physically based framework that couples independent SARAH-3 satellite irradiance with station-validated, lapse-rate–corrected ERA5-Land thermodynamics and a real module model, and adds an explainable attribution of the drivers of yield. Physically based PV-yield mapping of this kind has been demonstrated at the regional scale, most recently over Taiwan [10], which provides a useful quantitative benchmark for the thermodynamic correction. The present framework is built specifically for the Turkish setting and makes design choices suited to a continental, climatically diverse country: the measured SARAH-3 direct component is used in place of a diffuse split inferred from global irradiance with the Erbs model; the anisotropic Hay-Davies model replaces an isotropic transposition; the domain spans 26 years across the whole of Türkiye rather than a single island; and an explainability layer attributes the spatial pattern of yield to its physical drivers.

The contributions are therefore fourfold. A named, reproducible framework, HelioTR-Atlas, couples independent satellite irradiance with station-validated, lapse-rate–corrected reanalysis thermodynamics, including the dew-point and wind fields that had not previously been evaluated against Turkish stations. A country-scale, module-specific PV yield atlas for the Panasonic HIT 325 W module is produced for 2000–2025, with an inter-annual stability measure, a plus or minus one standard deviation classification, and provincial summaries. Finally, a SHAP analysis separates the irradiance control on raw yield from the thermal and topographic controls on the performance ratio, giving a physically interpretable account of where and why the resource departs from a pure-irradiance expectation.

To the best of our knowledge, HelioTR-Atlas is the first framework to close, within a single reproducible pipeline over Türkiye, three gaps that the literature has so far addressed only in isolation: the absence of a PV-yield–scale, station-validated resource assessment; the lack of any station-based evaluation of the ERA5-Land dew-point and wind fields on which such assessments depend; and the missing spatial attribution of the physical drivers of yield. The originality of the work lies not in a new learning algorithm but in coupling an independent satellite irradiance source with a validated, lapse-rate–corrected thermodynamic backbone and an explainable driver layer, so that the resulting national atlas is at once accurate, physically transparent, and directly usable for planning.

2. Data and study area

Türkiye offers a demanding testbed for topographic solar-resource modelling. The country spans roughly 783,000 km² across seven climate regions, from the humid Black Sea coast to the semi-arid southeast, and its terrain rises from sea level to above 5,000 m. These gradients produce sharp spatial contrasts in both irradiance and near-surface temperature, so a PV yield assessment that ignores elevation-driven thermodynamics will misrepresent large parts of the country (Fig. 1, Table 1).

Table 1
Input datasets used to build the HelioTR-Atlas PV-yield atlas.

Dataset	Variable(s)	Resolution	Period	Source
SARAH-3 (EUMETSAT CM SAF)	SIS (GHI), SID (direct)	0.05°, 30 min	2000–2025	[11]
ERA5-Land	2 m air temp., 2 m dew point, 10 m wind	0.1°, hourly	2000–2025	[12]
Copernicus GLO-30 DEM	Elevation	~ 30 m	static	[13]
GHCN-Daily / GSOD	Air temperature (62 stations)	daily	2000–2025	[14, 15]
GSOD (Türkiye)	Dew point, wind speed (119 stations)	daily	2000–2025	[15]

Surface solar irradiance was taken from the third edition of the Surface Solar Radiation Data Set – Heliosat (SARAH-3), produced by the EUMETSAT Satellite Application Facility on Climate Monitoring from Meteosat observations [11]. We used the surface incoming shortwave radiation (SIS, taken as global horizontal irradiance) and the surface incoming direct radiation (SID) products at their native 0.05° grid and 30-minute instantaneous sampling. Both variables were ordered for 2000–2025 over a box covering Türkiye (25.0–45.5° E, 35.0–42.5° N); the climate data record (1983–2020) and its interim extension (2021–2025) were concatenated into a continuous series. Retaining the measured direct component distinguishes this work from frameworks that infer the direct and diffuse split from global irradiance alone, and it allows the empirical separation model to be evaluated against the satellite observation rather than assumed.

Near-surface thermodynamics were drawn from ERA5-Land [12]. Hourly 2 m air temperature, 2 m dew-point temperature, and the 10 m wind components drove the PV model, and the corresponding daily archive was used for station validation. All ERA5-Land fields were kept at their native 0.1° resolution over 2000–2025. Because ERA5-Land is a land-only product, grid cells over water carry no data; the nearest valid land value was propagated to such cells before interpolation so that coastal Turkish cells received physically reasonable temperature and wind.

Terrain elevation came from the Copernicus GLO-30 digital elevation model [13], from which mean cell elevation was computed on both the 0.05° SARAH-3 grid and the 0.1° ERA5-Land grid. Ground observations comprised 62 harmonized stations combining the Global Historical Climatology Network-Daily [14] and the Global Surface Summary of the Day [15] archives for air temperature, together with 119 Turkish GSOD stations that also report dew point and wind speed. These records were used only to evaluate the reanalysis thermodynamics; no ground radiation measurements were required, which sidesteps the sparse in-situ radiation coverage over Türkiye.

3. Methods

The pipeline couples the three input streams through a topographic downscaling of the thermodynamic fields, a measured direct-diffuse decomposition and transposition of the irradiance, a cell-temperature and efficiency chain, and a driver-attribution stage, as summarized in Fig. 2.

3.1. Topographic downscaling of the thermodynamic fields

ERA5-Land air temperature was corrected to local elevation with a constant environmental lapse rate (Eq. 1), and dew-point temperature was corrected analogously (Eq. 2):

$$T_{corr} = T_{2m} - \gamma_T (z_{DEM} - z_{ERA5}$$

1

$$T_{d,corr} = T_{d,2m} - \gamma_{Td} (z_{DEM} - z_{ERA5}$$

2

Here γ_T = 0.65°C per 100 m and γ_{Td} = 0.20°C per 100 m are the lapse rates for air and dew-point temperature, z_{DEM} is the GLO-30 cell elevation, and z_{ERA5} is the model orography from the ERA5-Land geopotential. These corrections follow the topographic-downscaling approach applied by Han et al. [10] and the ERA5 lapse-rate downscaling of Dutra et al. [16]. Wind speed was taken as the magnitude of the hourly 10 m components.

3.2. Radiation decomposition

Because SARAH-3 provides the measured direct horizontal irradiance (SID), the direct normal irradiance (DNI) and the diffuse horizontal irradiance (DHI) were obtained directly from the satellite fields (Eqs. 3 and 4):

$$DNI = \frac{SID}{cos_{Q_z}}$$

3

$$DNI = GHI - SID$$

4

where the solar zenith angle θ_z was limited away from the horizon. The Taiwan framework of Han et al. [10] instead inferred this split from global irradiance with the Erbs et al. [17] separation model; that correlation was retained here only as an internal check, so that the assumption could be evaluated directly against the satellite-measured direct component (Section 4.5).

3.3. Photovoltaic yield model

Hourly plane-of-array irradiance was computed with the Hay-Davies anisotropic transposition model [18] for a fixed array tilted at the local latitude and oriented due south; the sensitivity of the yield to this tilt choice is quantified in Section 4.7. Module temperature was estimated with the Faiman model [19] from plane-of-array irradiance, lapse-corrected air temperature, and wind speed. Relative module efficiency was represented with the ADR model [20], whose coefficients were calibrated to the Panasonic HIT 325 W module (VBHN325SA16), also used by Han et al. [10]. The calibration used the module single-diode current-voltage surface across irradiance and cell temperature, giving a standard-test-condition efficiency of 19.5% and a maximum-power temperature coefficient of − 0.310% per °C. Hourly module-DC specific yield in kWh kWp^{-1} was accumulated as the product of relative efficiency and plane-of-array irradiance normalized to standard conditions. All radiative and PV computations used pvlib [21]. System-level losses were not applied; a constant derate near 0.85 converts the reported module-DC yield to an as-installed estimate.

3.4. Climatological aggregation

The primary product is the long-term mean specific yield over 2000–2025, the standard resource metric used by operational atlases such as PVGIS [22] and more robust to inter-annual variability than a single typical meteorological year. The per-year yields were averaged to the climatological mean, their inter-annual coefficient of variation was retained as a stability indicator, and each cell was classified as poor, medium, or good using a plus or minus one standard deviation threshold about the national spatial mean. The atlas was clipped to Turkish territory using the dissolved provincial boundaries. A gridded typical meteorological year was not generated, because Han et al.[10] found it to differ from the long-term mean by only about 14 kWh m^{-2} for annual totals.

3.5. Station validation

The corrected fields were evaluated against the ground network following the group design of Han et al. [10], adapted to Türkiye. Group A comprises interior land cells, Group B comprises coastal or edge cells where the land-only reanalysis extrapolates across the shoreline, and Group C comprises stations above 1000 m. Mean bias deviation (MBD) and root-mean-square deviation (RMSD) were computed at daily resolution before and after the lapse-rate correction. Because the harmonized records are daily, the validation is reported at the daily scale, and dew point and wind were evaluated only at the GSOD stations that report them.

3.6. Driver attribution

To quantify why yield varies in space, a random-forest surrogate was trained on grid-cell features (global horizontal irradiance, lapse-corrected air temperature, elevation, wind speed, latitude) and explained with SHAP values [23]. Two targets were analysed, the specific yield Y itself and the performance ratio PR (Eq. 5), defined as the yield normalized by the plane-of-array irradiance H_{POA}:

$$PR = \frac{Y}{H_{POA}}$$

5

Attributing the performance ratio isolates the thermal and topographic controls that the irradiance-dominated raw yield obscures. Because the surrogate reproduced the physical yield and performance-ratio fields with coefficients of determination of 0.995 and 0.956 respectively (Section 4.4), its SHAP attributions are treated as a faithful proxy for the sensitivities of the physical model rather than as artefacts of the surrogate.

4. Results

4.1. Validation of the corrected thermodynamic fields

The lapse-rate correction brought ERA5-Land air temperature substantially closer to the station observations. Pooled over the 62 temperature stations, the MBD improved from − 2.01 to − 0.96°C and the RMSD fell from 3.25 to 2.52°C, a reduction of 22.5% (Table 2, Fig. 3). The improvement was largest for the coastal Group B (35.7%) and reached 20.3% for the high-elevation Group C, confirming that the correction is most effective where the elevation mismatch between the reanalysis grid and the surface is greatest; the per-station distributions in Fig. 4 show the same behaviour. The correction had a smaller effect on dew point: across the 119 GSOD stations the RMSD moved only from 2.19 to 2.13°C, because the raw ERA5-Land dew point already carried a small bias of − 0.42°C over Türkiye. For 10 m wind speed, evaluated without elevation correction, ERA5-Land showed an MBD of − 0.56 m s^{-1} and an RMSD of 1.49 m s^{-1}.

Table 2
Validation of the ERA5-Land thermodynamic fields against the Turkish station network, before and after the lapse-rate correction, by station group (2000–2025).

Variable	Group	n	MBD raw	RMSD raw	MBD corr.	RMSD corr.	ΔRMSD (%)
Air temperature (°C)	A (interior)	30	-1.89	3.02	-0.93	2.42	20.0
	B (coastal)	11	-2.33	3.02	-1.07	1.94	35.7
	C (> 1000 m)	21	-2.01	3.68	-0.96	2.93	20.3
	All	62	-2.01	3.25	-0.96	2.52	22.5
Dew point (°C)	A (interior)	63	-0.39	2.16	-0.16	2.10	2.5
	B (coastal)	31	-0.32	2.05	-0.05	1.94	5.5
	C (> 1000 m)	25	-0.62	2.46	-0.37	2.41	1.9
	All	119	-0.42	2.19	-0.17	2.13	3.1
Wind speed (m s^{-1})	All	119	-0.56	1.49	–	–	–

MBD, mean bias deviation; RMSD, root-mean-square deviation; corr., after lapse-rate correction. Wind speed is evaluated without elevation correction. Statistics are pooled over daily station–cell pairs for 2000–2025.

4.2. Photovoltaic yield atlas

The engine was verified by reproducing an independent single-cell calculation, which agreed with the grid atlas to within rounding. Over 2000–2025 the climatological module-DC specific yield averaged 1805 kWh kWp^{-1} across Turkish land, ranging from roughly 1200 to 2200 kWh kWp^{-1} (Fig. 5). A conventional system derate near 0.85 places the national mean close to 1500 kWh kWp^{-1} of installed capacity. The inter-annual coefficient of variation averaged 3.3%, indicating a stable resource, with the largest variability on the cloudier eastern Black Sea coast (Fig. 6). Under the classification, 4583 cells were rated good, 23,071 medium, and 4825 poor.

4.3. Provincial summaries

Aggregating over the 81 provinces produced a ranking consistent with the known Turkish solar climatology (Fig. 7, Table 3). The highest provincial means occurred in the southeast and along the Mediterranean and Aegean coasts, led by Şanlıurfa (1994 kWh kWp^{-1}), Muğla (1988), Mersin (1980), Kilis (1979), and Aydın (1972). The lowest values formed a contiguous block on the eastern Black Sea coast, at Trabzon (1334), Rize (1344), Ordu (1416), Giresun (1504), and Düzce (1526).

Table 3
Ten highest- and ten lowest-yielding provinces by climatological mean module-DC specific yield (2000–2025). Full 81-province values are available from the corresponding author.

Rank	Province	Mean yield (kWh kWp⁻¹)	Rank	Province	Mean yield (kWh kWp⁻¹)
1	Şanlıurfa	1994	72	Bartın	1578
2	Muğla	1988	73	Kocaeli	1573
3	Mersin	1980	74	Zonguldak	1571
4	Kilis	1979	75	Sinop	1568
5	Aydın	1972	76	Samsun	1527
6	Gaziantep	1962	77	Düzce	1526
7	Mardin	1960	78	Giresun	1504
8	Karaman	1960	79	Ordu	1416
9	Antalya	1952	80	Rize	1344
10	Burdur	1951	81	Trabzon	1334

4.4. Driver attribution

The random-forest surrogate reproduced the spatial yield field almost exactly, with a hold-out coefficient of determination of 0.995. Explaining the raw yield, global horizontal irradiance dominated the SHAP importance (mean absolute SHAP of 119 kWh kWp⁻¹) far ahead of latitude, air temperature, elevation, and wind, confirming that irradiance sets the first-order spatial pattern. Attributing the performance ratio, which removes the irradiance signal, reversed the ranking: air temperature became the leading driver, followed by wind, while irradiance contributed almost nothing. The elevation-band breakdown is expressed as each driver's relative share of the total performance-ratio importance, because the absolute SHAP magnitudes are small. Air temperature accounted for about 54% of that importance in the low-elevation band, 46% at intermediate elevations, and 73% in the high band, so the thermal control was strongest at the extremes of the elevation distribution and largest in the cold highlands, while irradiance, elevation, and wind together made up the remainder (Fig. 8).

4.5. Measured versus inferred direct irradiance

Because comparable frameworks infer the direct and diffuse split from global irradiance, the Erbs correlation was evaluated against the measured SARAH-3 direct horizontal irradiance across the whole grid at 30-minute resolution for 2020 (54.5 million samples). The Erbs estimate reproduced the measured direct component with a mean bias of 12.7 W m⁻² and a root-mean-square deviation of 34.1 W m⁻². The agreement depended strongly on sky condition: under overcast skies (clearness index below 0.3) the correlation was essentially unbiased (MBD 0.5 W m⁻², RMSD 3.7 W m⁻²), and it remained close to unbiased at intermediate clearness (MBD 0.7 W m⁻², RMSD 31.3 W m⁻²). Under clear skies (clearness

index above 0.6), which contribute the bulk of the annual yield, the Erbs correlation overestimated the direct horizontal irradiance by 22.1 W m^{-2} (RMSD 40.1 W m^{-2}). Using the measured direct component therefore avoids a clear-sky bias in the transposition that an inferred split would introduce.

4.6. Cross-comparison with PVGIS

Türkiye has no operational ground-radiation network available for this study, so the atlas was benchmarked against PVGIS [22], an established product that combines PVGIS-SARAH2 irradiance and ERA5 meteorology with its own PV model. PVGIS was queried at the 81 provincial centroids for a 1 kWp south-facing array tilted at the local latitude with 14% system loss, and the atlas module-DC yield was converted to a system estimate with the same derate. The annual in-plane irradiation agreed closely, with a mean bias of 37.7 kWh m^{-2} (2.1%), a root-mean-square deviation of 73.8 kWh m^{-2}, and a spatial correlation of 0.95 (Fig. 9a), which indicates that the SARAH-3 radiation and the transposition reproduce the established product to within a few percent. The annual yield was 9.7% higher than PVGIS (MBD 134.9 kWh kWp^{-1}, RMSD 145.3 kWh kWp^{-1}) with a spatial correlation of 0.93 (Fig. 9b). The positive yield bias is consistent with the Panasonic HIT module, whose low temperature coefficient and low-light response raise the specific yield relative to the generic crystalline-silicon module assumed by PVGIS, while the high correlation confirms that the atlas reproduces the spatial pattern of the resource. Because both products ultimately draw on satellite irradiance, this is a cross-comparison rather than an independent validation; the absolute accuracy of the underlying SARAH-3 irradiance rests on its own extensive ground validation, which reports agreement of about 5 W m^{-2} for monthly surface irradiance [11].

4.7. Sensitivity to array tilt

The atlas fixes the array tilt at the local latitude, a common design rule. To quantify what this leaves unused, the full PV chain was run at six sites spanning the latitude, elevation and climate range of the country, sweeping the tilt from 0 to 50° over 2020 (Table 4). The optimum annual tilt fell between 31 and 35°, a few degrees below the latitude, consistent with the summer weighting of the resource and the diffuse contribution. Across the six sites the latitude tilt stayed within 0.3 to 1.0% of the site optimum (mean 0.5%), whereas a horizontal array lost 11 to 13%. Fixing the tilt at the local latitude therefore captures almost all of the achievable in-plane gain and does not materially change the spatial pattern of the atlas.

Table 4
Sensitivity of the annual module-DC specific yield (kWh kWp⁻¹) to array tilt at six representative sites (2020), for a south-facing array. φ is the local latitude. The last two columns give the shortfall of the latitude-tilt and of a horizontal array relative to the site optimum.

Site	Lat (°N)	Elev. (m)	Yield at tilt = φ	Optimum tilt (°)	Yield at optimum	φ vs opt. (%)	Horiz. vs opt. (%)
Mersin	36.8	15	2001	32	2008	-0.31	-11.4
Konya	37.9	1033	1988	31	1998	-0.48	-11.0
Şanlıurfa	37.2	540	2005	31	2013	-0.37	-10.7
Erzurum	39.9	1939	1880	35	1885	-0.29	-13.0
Rize	41.0	150	1349	35	1354	-0.34	-12.5
İstanbul	41.0	21	1685	31	1703	-1.03	-10.5

5. Discussion

The results confirm that elevation-aware thermodynamics are essential over Türkiye. The lapse-rate correction removed roughly half of the cold bias that ERA5-Land carries in mountainous terrain, and its 22.5% reduction in temperature error is close to the 25.5% reported by Han et al. [10] for Taiwan, which supports the transferability of the approach to a larger and more varied domain. The smaller gain for dew point is itself informative: ERA5-Land dew point is already well-calibrated over Türkiye, so the correction has little to remove, a finding that could only be established because the dew-point field was evaluated against stations here for the first time.

The relationship to the Taiwan study of Han et al. [10] is one of independent, parallel development rather than reproduction. Both works share the physically based logic of coupling satellite irradiance with lapse-rate–corrected ERA5-Land and a real module model, and the close agreement of the two lapse-rate results reported above is reassuring evidence that the underlying physics transfers across regions. The methodological choices diverge where the Turkish terrain demands it: the measured SARAH-3 direct component replaces the Erbs-inferred split, the anisotropic Hay-Davies transposition replaces the isotropic assumption, the record spans 26 years of a continental domain rather than a single island, and an explainable driver layer is added. The measured split is not a cosmetic change, because the Erbs correlation overestimates the direct horizontal irradiance by 22 W m⁻² under the clear skies that dominate the annual yield (Section 4.5), removing a systematic clear-sky bias from the transposition. Relative to general-purpose atlases such as PVGIS [22], the distinguishing features are the validated thermodynamic backbone, the independent satellite radiation source, and the driver attribution.

The choice to draw irradiance from SARAH-3 rather than from ERA5 itself is supported by a consistent body of validation evidence, and the magnitude of the reported ERA5 radiation error clarifies why the distinction matters for a terrain-resolved atlas. Benchmarking ERA5 global horizontal irradiance against 41 reference stations worldwide and 294 stations across Europe, Urraca et al. [24] measured a moderate

positive bias of 4.05 W m⁻² globally and 4.54 W m⁻² over Europe; although this halves the bias of the earlier ERA-Interim and MERRA-2 reanalyses by 50 to 75%, the 31 km ERA5 grid proved inadequate over coasts and mountains, where the absolute errors grow, and the reanalysis was found to overestimate under cloud while slightly underestimating under clear skies. At high latitudes Babar et al. [25] ranked the monthly ERA5 error, a mean absolute deviation of 6.8 W m⁻², above both satellite products they tested, SARAH at 5.0 W m⁻² and CLARA at 5.6 W m⁻², and traced it to a systematic misrepresentation of cloud that worsens as cloudiness rises. On the daily scale the same group measured an ERA5 root-mean-square deviation of 27.1 W m⁻² with a + 4.3 W m⁻² bias, against 17.9 W m⁻² for the CLARA satellite product [26].

Two features of this literature bear directly on Türkiye: the ERA5 radiation error is largest in exactly the coastal and mountainous settings that dominate the country, and the reanalysis mishandles the clear-sky regime that supplies most of the annual yield. Taking irradiance from the measured SARAH-3 fields while restricting ERA5-Land to the thermodynamic terms is therefore the appropriate division of labour. The agreement of the resulting in-plane irradiation with the independent PVGIS product (mean bias 2.1%, r = 0.95, Section 4.6) is consistent with the satellite-grade accuracy these studies attribute to SARAH-class products rather than to reanalysis radiation, and it reinforces the internal finding that an Erbs-inferred diffuse split would have injected a 22 W m⁻² clear-sky bias into the transposition (Section 4.5).

Where ERA5 and ERA5-Land have been used directly for solar-resource mapping, the emphasis has fallen on potential rather than validated, terrain-resolved yield, and the simplifications adopted there sharpen the contribution of the present work. Qiu et al. [3] exploited the ten-kilometre hourly ERA5-Land radiation and temperature fields to estimate China's PV potential, gaining finer spatial discrimination than coarser reanalyses such as MERRA. Jiang et al. [27] validated ERA5 hourly surface radiation against reference stations over 3.2 million samples, obtaining a correlation of 0.93 with a 6.72 W m⁻² bias and a 113.5 W m⁻² root-mean-square error, and used it to show that solar intermittency, measured by the coefficient of variation, rose across roughly 85% of global land between the 2000s and the 2010s.

Two contrasts are instructive. First, Jiang et al. [27] assumed a linear relationship between PV output and surface radiation and explicitly set aside air temperature, diffuse radiation and wind speed; HelioTR-Atlas instead carries the full thermodynamic chain, so that precisely those terms enter the yield and, through the attribution analysis, are shown to govern the performance ratio. Second, the national inter-annual coefficient of variation of 3.3% found here lies far below the rising intermittency reported for much of the globe, and the spatial-aggregation smoothing that Jiang et al. document, a coefficient of variation falling from 0.69 to 0.58 across a distributed European portfolio, implies that a geographically dispersed Turkish fleet would be steadier still, placing the country among the lower-risk mid-latitude regimes.

The use of ERA5-Land for the thermodynamic terms rests on independent support. Jani et al. [28] validated ERA5 wind against LIDAR measurements and obtained a correlation of 0.93, higher than for the MERRA-2, CFSR2 and other reanalyses they compared, and Ascencio-Vásquez et al. [29] reported near-

exact agreement of ERA5 with reference observations, with coefficients of determination of 0.99 for ambient temperature and 0.89 for global horizontal irradiance, which justified using it as a synthetic driver for PV modelling without ground measurements. The stronger agreement for temperature than for irradiance in that comparison itself echoes the division of labour adopted here, in which ERA5-Land supplies the thermodynamic terms while the irradiance is taken from the satellite. A recurring caveat, however, is that the reanalysis fields underpinning such assessments are seldom validated independently at the site of application: Soukissian et al. [30], for instance, noted a 20% relative deviation of ERA5 wind from measurements in the Mediterranean yet relied on the reanalysis without local correction. The station evaluation of temperature, dew point and wind reported here addresses that caveat directly for Türkiye.

The attribution analysis carries the main interpretive message. Raw yield is governed by irradiance, as expected, but the performance ratio, which strips out irradiance, is controlled primarily by air temperature and secondarily by wind, and the temperature control is strongest at the extremes of the elevation distribution. Hot lowlands lose performance to elevated cell temperatures, whereas cold highlands gain it, so the topographic correction feeds directly into a spatially varying thermal signal that a pure-irradiance atlas cannot express, indicating where thermal and elevation effects most influence achievable output. This is the mechanism that the linear PV-radiation assumption of the ERA5-based assessments discussed above cannot resolve.

Several boundaries of the analysis should be read alongside these findings. The reported yields are module-DC quantities from the ADR and Faiman models, so an as-installed estimate requires the additional system derate noted above. The good, medium, and poor labels are a relative classification defined by plus or minus one standard deviation about the national mean rather than an absolute economic threshold, and are meant for comparison within the country rather than as bankability categories. The fixed latitude tilt is a design convention rather than a constraint: a full sweep at representative sites placed it within one percent of the site optimum (Section 4.7), so the atlas is insensitive to this choice. The station comparison is performed at daily resolution because the harmonized records are daily, so the sub-daily structure of the thermodynamic error is not validated directly; the daily comparison nonetheless constrains the dominant thermal term of the PV model, and the satellite irradiance that drives the diurnal cycle carries its own validated sub-daily behaviour in the SARAH-3 product [11], which together bound the effect on the annual yield. Reporting a long-term climatological mean together with an explicit inter-annual coefficient of variation, rather than a single typical meteorological year, is a deliberate choice: a typical year collapses the distribution to one representative sequence and cannot convey the year-to-year variability that climate variability imparts to PV output [4], whereas the mean-and-variability summary preserves both the central estimate and its stability, which the atlas shows to be high across most of the country. The radiation record joins the validated climate data record with its interim extension, and the lapse rates are held constant rather than allowed to vary by season, both of which are reasonable for a national climatology but leave room for refinement. The ADR coefficients were calibrated from the module single-diode surface rather than from outdoor measurements. Future work can incorporate ground radiation once national measurements

become available, allow the lapse rates to vary seasonally, and release a gridded typical meteorological year as a simulation-ready complement to the climatological maps.

6. Conclusions

HelioTR-Atlas brings together, for the first time over Türkiye, independent satellite irradiance, station-validated and lapse-rate–corrected reanalysis thermodynamics, and an explainable attribution of yield drivers within a single reproducible framework. The resulting 2000–2025 atlas resolves the country's topographic contrasts, ranks provinces in agreement with the known solar climatology, and shows that while irradiance sets the resource, temperature governs its departure from a pure-irradiance expectation most strongly at the extremes of elevation. The framework is portable to other complex-terrain regions and provides a validated basis for national and provincial solar planning.

Declarations

CRediT authorship contribution statement

Ertuğrul Gül: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization.

Competing interest

The author declares no competing interests.

Funding

This study was supported by the Scientific Research Projects Coordination Office of Hakkari University, Türkiye (Project No: FM19BAP).

Data availability

All primary datasets are openly available from their providers. SARAH-3 surface incoming shortwave and direct irradiance are distributed by the EUMETSAT CM SAF [31]; ERA5-Land hourly fields are available from the Copernicus Climate Data Store [12]; the Copernicus GLO-30 digital elevation model is available from the Copernicus Data Space Ecosystem; and the ground observations are the NOAA Global Surface Summary of the Day (GSOD) together with the Turkish State Meteorological Service station records used under their terms of access. The derived atlas and the tabulated results supporting the findings are archived on Zenodo under a CC BY 4.0 licence [32], comprising the climatological yield grids, the provincial statistics, the tilt-sensitivity sweep, the PVGIS cross-comparison and the station-validation summaries. The Turkish State Meteorological Service station records themselves are not redistributed, in line with their terms of access.

Code availability

The code implementing the HelioTR-Atlas pipeline and reproducing all results and figures is contained in a single notebook, available from the author upon reasonable request.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work the author used Anthropic Claude to improve the language and readability of the manuscript and to support the development and debugging of the Python data-processing and modeling pipeline. After using this tool, the author reviewed and edited the content as needed, verified all code, outputs, and analyses, and takes full responsibility for the content of the publication.

References

1. Coruhlu, Y. E., Solgun, N., Baser, V. & Terzi, F. Revealing the solar energy potential by integration of GIS and AHP in order to compare decisions of the land use on the environmental plans. *Land. Use Policy*. **113**, 105899 (2022).
2. Karayel, G. K., Javani, N. & Dincer, I. Green hydrogen production potential for Turkey with solar energy. *Int. J. Hydrog. Energy*. **47**, 19354–19364 (2022).
3. Qiu, T. et al. Potential assessment of photovoltaic power generation in China. *Renew. Sustain. Energy Rev.* **154**, 111900–111900 (2021).
4. Chen, X. et al. Climate change impacts on global photovoltaic variability. *Appl. Energy*. **374**, 124087 (2024).
5. Nielsen, A. H., Iosifidis, A., Karstoft, H. & IrradianceNet Spatiotemporal deep learning model for satellite-derived solar irradiance short-term forecasting. *Sol. Energy*. **228**, 659–669 (2021).
6. Mitrentsis, G. & Lens, H. An interpretable probabilistic model for short-term solar power forecasting using natural gradient boosting. *Appl. Energy*. **309**, 118473 (2022).
7. Cao, Q., Liu, Y., Sun, X. & Yang, L. Country-level evaluation of solar radiation data sets using ground measurements in China. *Energy* **241**, 122938 (2022).
8. Amjad, M., Yılmaz, M. T., Yücel, İ. & Yılmaz, K. K. Performance evaluation of satellite- and model-based precipitation products over varying climate and complex topography. *J. Hydrol.* **584**, 124707–124707 (2020).
9. Hasan Karaman, Ç. & Akyürek, Z. Evaluation of near-surface air temperature reanalysis datasets and downscaling with machine learning based Random Forest method for complex terrain of Turkey. *Adv. Space Res.* **71**, 5256–5281 (2023).
10. Han, J. Y., Kan, Y. H. & Vohnicky, P. A framework for PV yield-based solar energy mapping in mid-low latitude regions with complex topography. *Renew. Energy*. **256**, 124512 (2026).
11. Pfeifroth, U. et al. SARAH-3 – satellite-based climate data records of surface solar radiation. *Earth Syst. Sci. Data*. **16**, 5243–5265 (2024).

12. Muñoz-Sabater, J. et al. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth Syst. Sci. Data*. **13**, 4349–4383 (2021).
13. European Space Agency & Airbus. Copernicus DEM. (2022). https://doi.org/10.5270/ESA-c5d3d65
14. Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E. & Houston, T. G. An Overview of the Global Historical Climatology Network-Daily Database. *J. Atmos. Ocean. Technol.* **29**, 897–910 (2012).
15. NOAA National Centers for Environmental Information. Global Surface Summary of the Day - GSOD. (2018). https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00516
16. Dutra, E. et al. Environmental Lapse Rate for High-Resolution Land Surface Downscaling: An Application to ERA5. *Earth Space Science* **7**, (2020).
17. 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*. **28**, 293–302 (1982).
18. Hay, J. E. Calculating solar radiation for inclined surfaces: Practical approaches. *Renew. Energy*. **3**, 373–380 (1993).
19. Faiman, D. Assessing the outdoor operating temperature of photovoltaic modules. *Progress Photovoltaics*. **16**, 307–315 (2008).
20. Driesse, A., Theristis, M. & Stein, J. S. A New Photovoltaic Module Efficiency Model for Energy Prediction and Rating. *IEEE J. Photovoltaics*. **11**, 527–534 (2021).
21. Holmgren, F., Hansen, W. W., Mikofski, A. & C. & M. pvlib python: a python package for modeling solar energy systems. *JOSS* **3**, 884 (2018).
22. Huld, T., Müller, R. & Gambardella, A. A new solar radiation database for estimating PV performance in Europe and Africa. *Sol. Energy*. **86**, 1803–1815 (2012).
23. Lundberg, S. & Lee, S. I. A Unified Approach to Interpreting Model Predictions. 30, (2017).
24. Urraca, R. et al. Evaluation of global horizontal irradiance estimates from ERA5 and COSMO-REA6 reanalyses using ground and satellite-based data. *Sol. Energy*. **164**, 339–354 (2018).
25. Babar, B., Graversen, R. G. & Boström, T. Solar radiation estimation at high latitudes: Assessment of the CMSAF databases, ASR and ERA5. *Sol. Energy*. **182**, 397–411 (2019).
26. Babar, B., Luppino, L. T., Boström, T. & Anfinsen, S. N. Random forest regression for improved mapping of solar irradiance at high latitudes. *Sol. Energy*. **198**, 81–92 (2020).
27. Jiang, H., Lu, N., Yao, L., Qin, J. & Liu, T. Impact of climate changes on the stability of solar energy: Evidence from observations and reanalysis. *Renew. Energy*. **208**, 726–736 (2023).
28. Jani, H. K., Kachhwaha, S. S., Nagababu, G. & Das, A. Temporal and spatial simultaneity assessment of wind-solar energy resources in India by statistical analysis and machine learning clustering approach. *Energy* **248**, 123586–123586 (2022).
29. Ascencio-Vásquez, J., Kaaya, I., Brecl, K., Weiß, K. & Topič, M. Global Climate Data Processing and Mapping of Degradation Mechanisms and Degradation Rates of PV Modules. *Energies* **12**, 4749–4749 (2019).

30. Soukissian, T., Karathanasi, F. & Zaragkas, D. K. Exploiting offshore wind and solar resources in the Mediterranean using ERA5 reanalysis data. *Energy. Conv. Manag.* **237**, 114092–114092 (2021).
31. Pfeifroth, U. et al. Surface Radiation Data Set - Heliosat (SARAH) - Edition 3. 12.3 TiB Satellite Application Facility on Climate Monitoring (CM SAF) https://doi.org/10.5676/EUM_SAF_CM/SARAH/V003 (2023).
32. Gül, E. HelioTR-Atlas: station-validated, explainable photovoltaic yield atlas of Türkiye from SARAH-3 satellite irradiance and lapse-rate-corrected ERA5-Land (2000–2025). Zenodo (2026). https://doi.org/10.5281/ZENODO.21385828

Figures

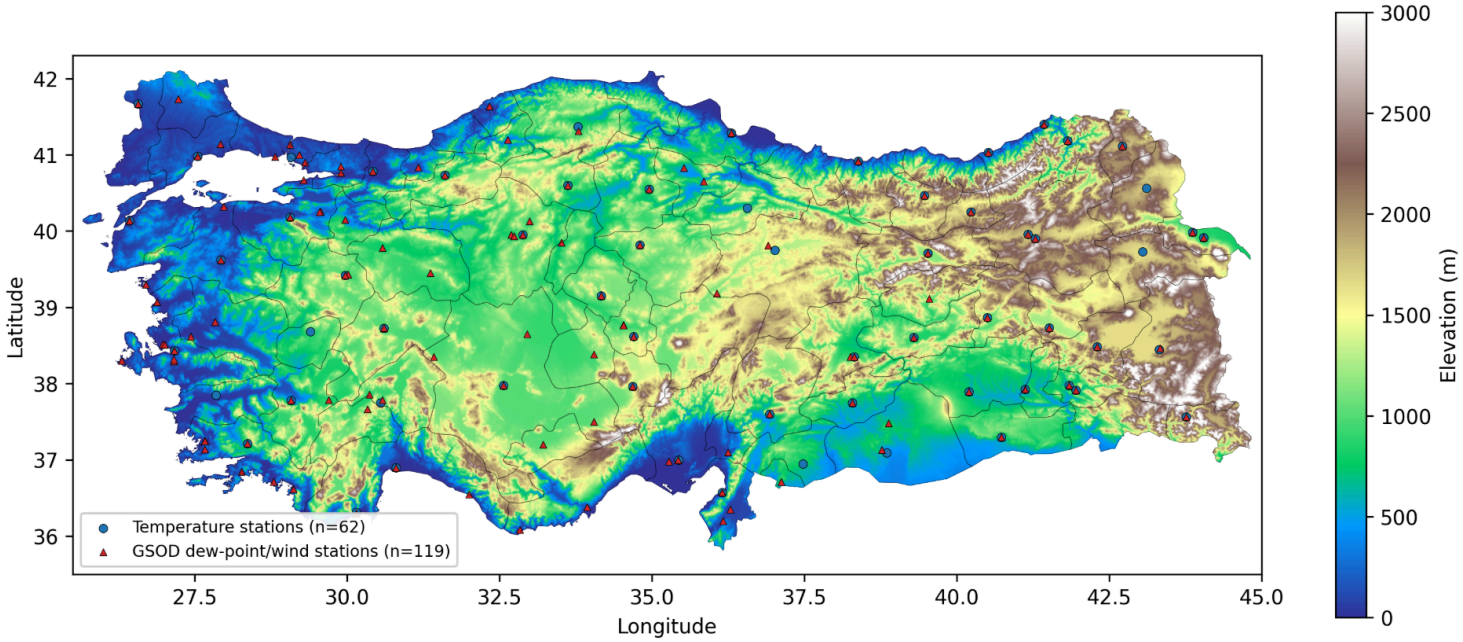


Figure 1

Study area. Topography of Türkiye from the Copernicus GLO-30 digital elevation model, with the validation network: 62 stations used for the air-temperature comparison and 119 NOAA GSOD stations used for the dew-point and wind comparison.

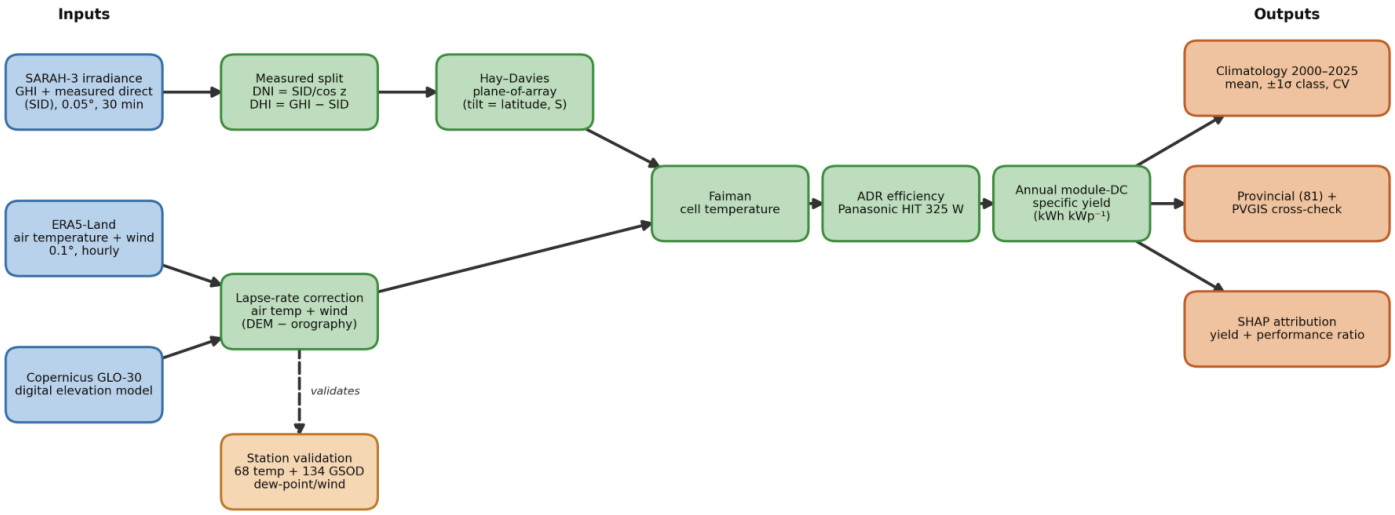


Figure 2

Methodology. Processing chain from SARAH-3 satellite irradiance, ERA5-Land thermodynamic fields and the GLO-30 elevation model through the lapse-rate correction, the Hay-Davies transposition, the Faiman cell-temperature and ADR efficiency models, to the specific-yield atlas and the SHAP driver attribution.

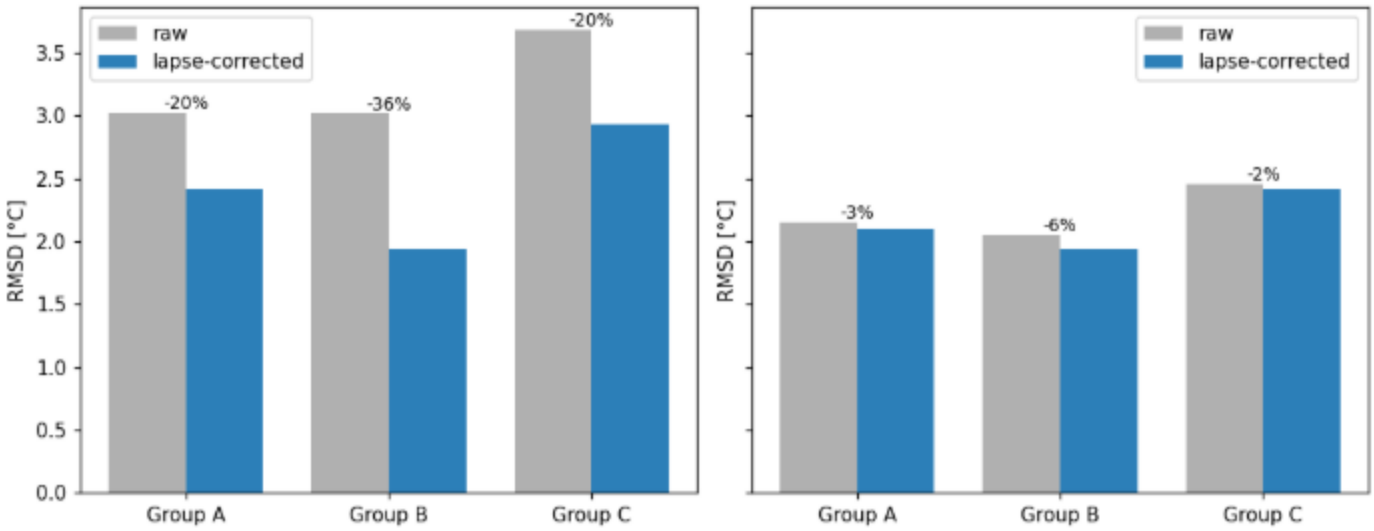


Figure 3

Effect of the lapse-rate correction on the ERA5-Land error against stations, shown as the mean RMSD before and after correction by station group.

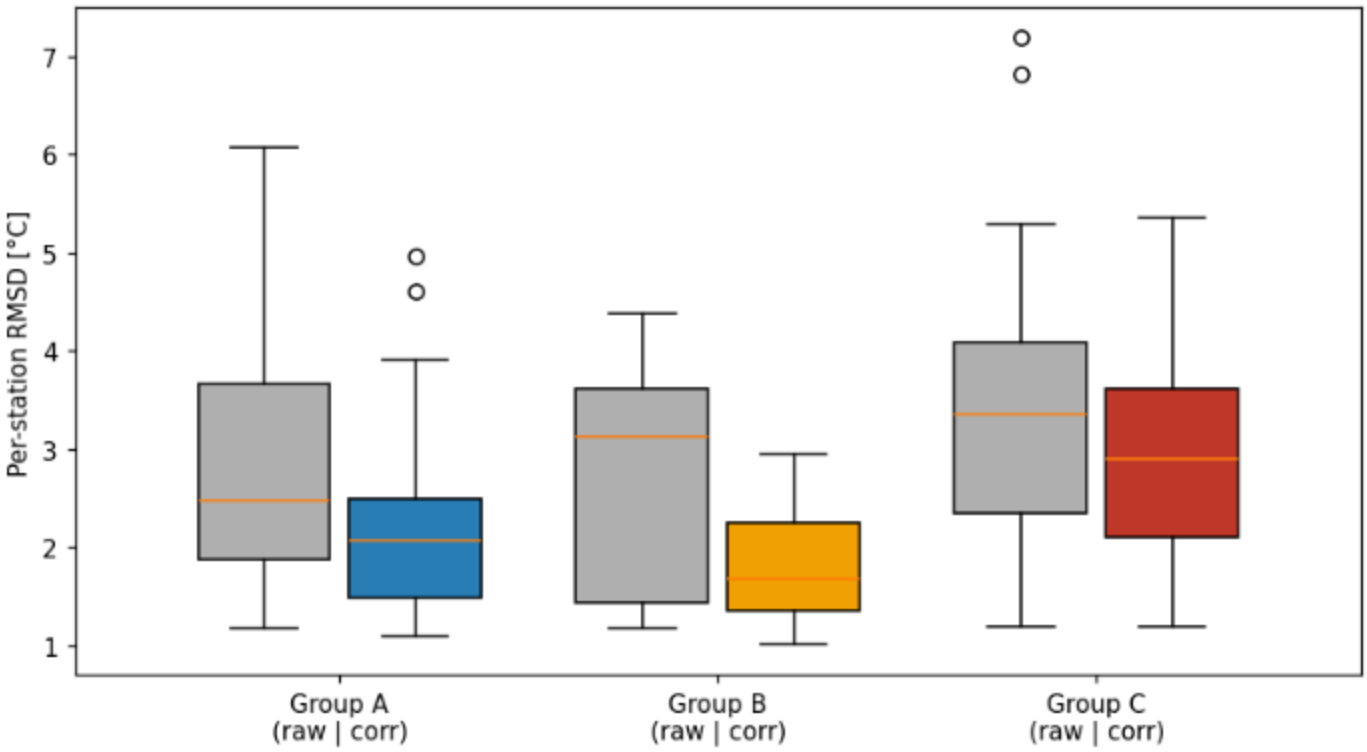


Figure 4

Per-station RMSD distributions before and after the lapse-rate correction for the interior, coastal, and high-elevation station groups.

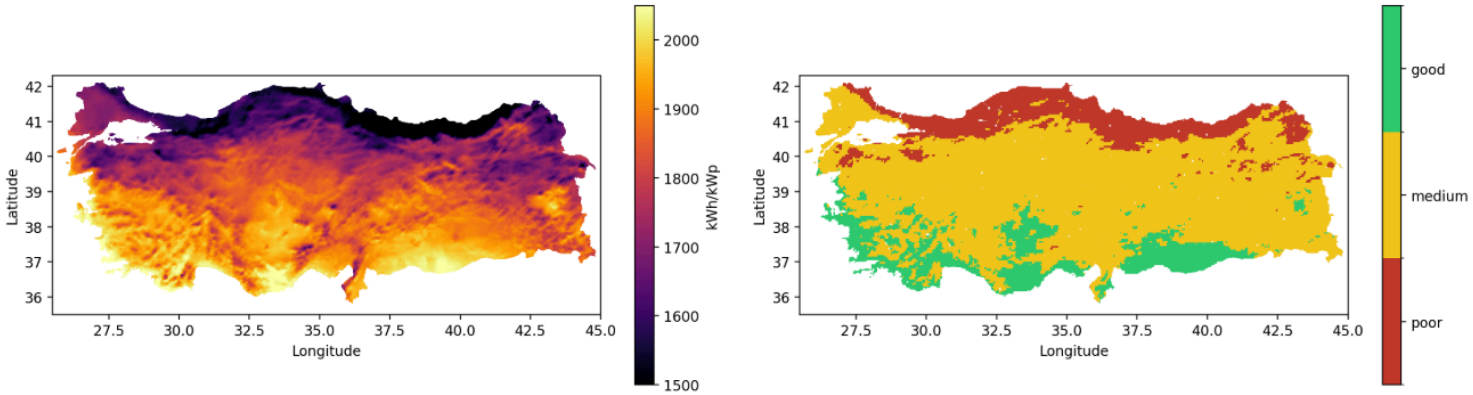


Figure 5

Photovoltaic yield atlas for 2000–2025. (a) Climatological module-DC specific yield (kWh kWp⁻¹). (b) Relative good, medium, and poor classification defined by plus or minus one standard deviation about the national mean.

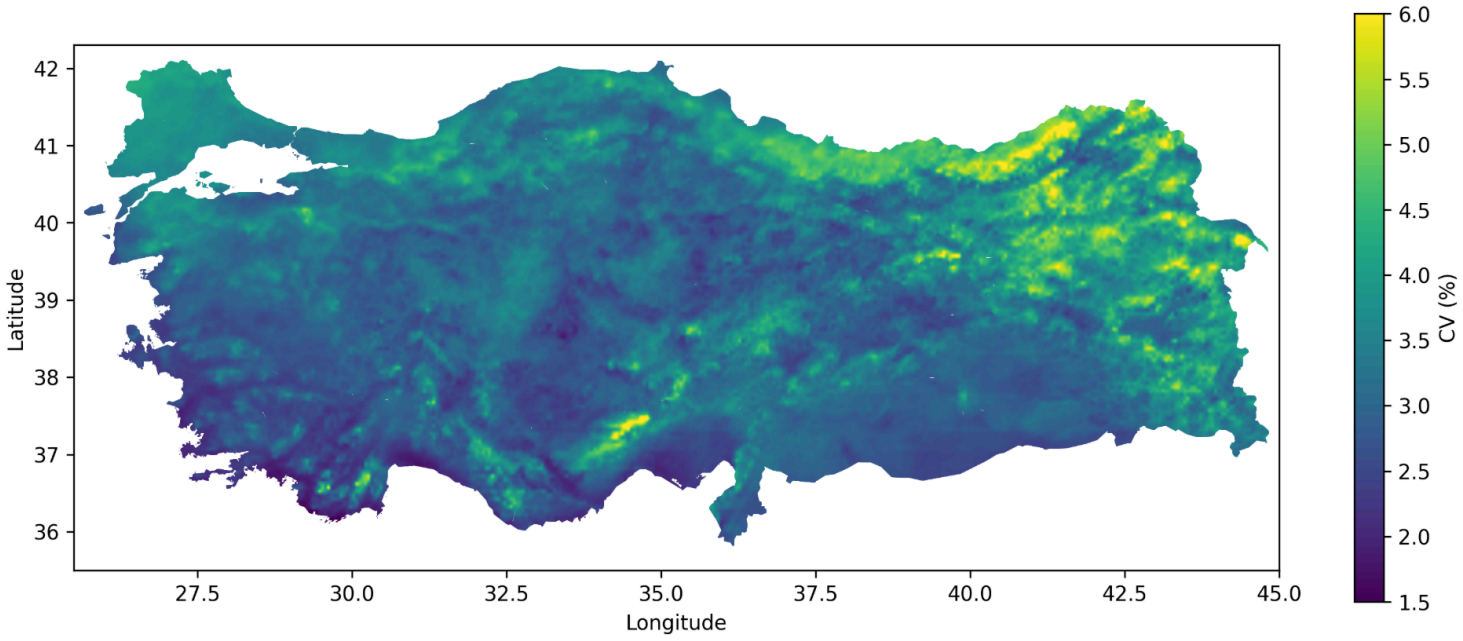


Figure 6

Inter-annual coefficient of variation of the annual specific yield over 2000–2025.

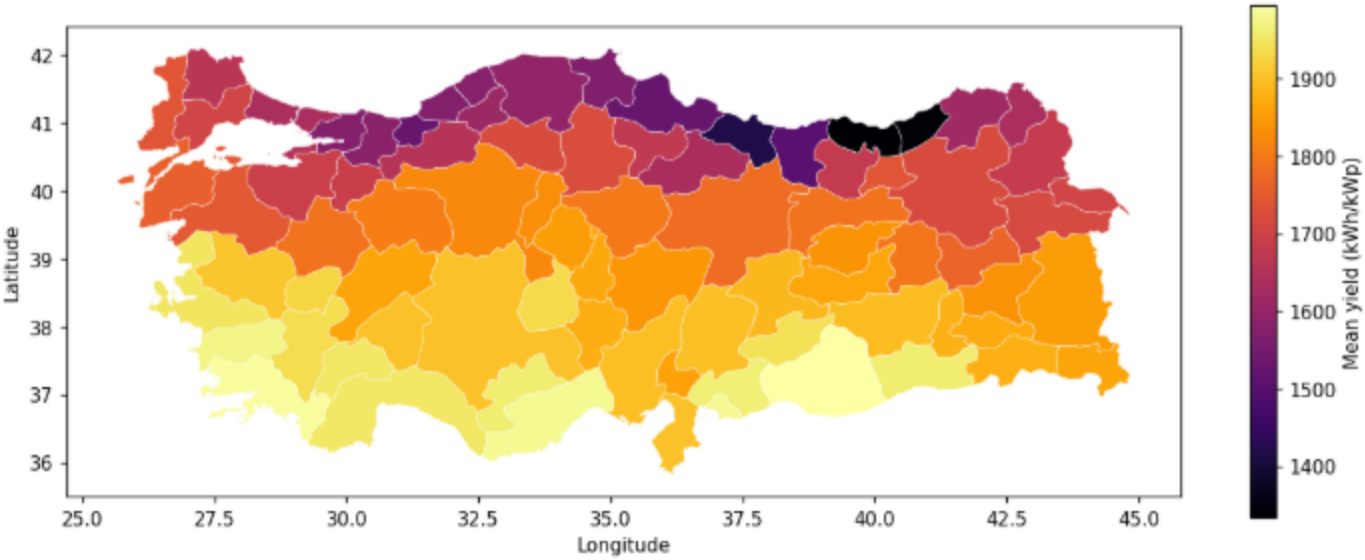


Figure 7

Provincial mean specific yield aggregated over the 81 provinces.

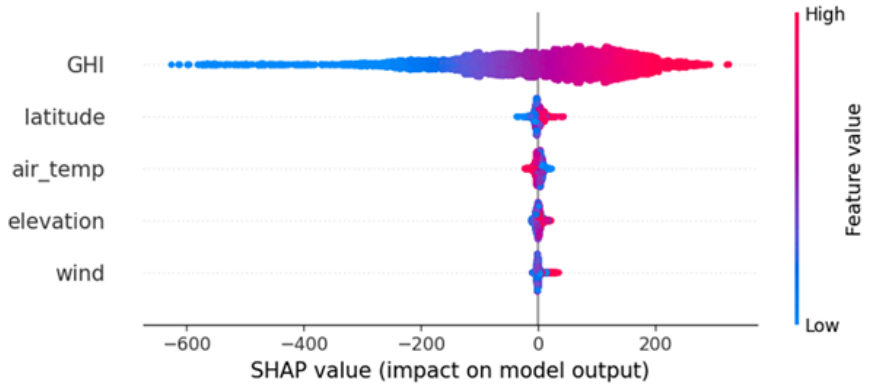


Fig. 8a. SHAP beeswarm for the raw specific yield.

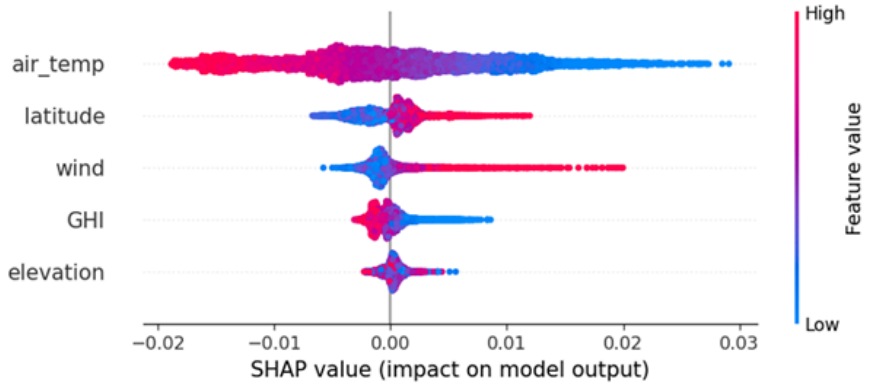


Fig. 8b. SHAP beeswarm for the performance ratio.

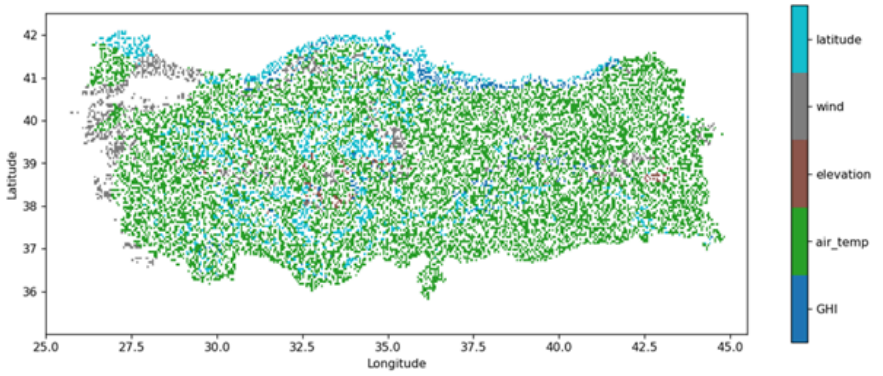


Fig. 8c. Map of the leading performance-ratio driver per grid cell. Panels (a)–(c) together form Fig. 8.

Figure 8

See image above for figure legend

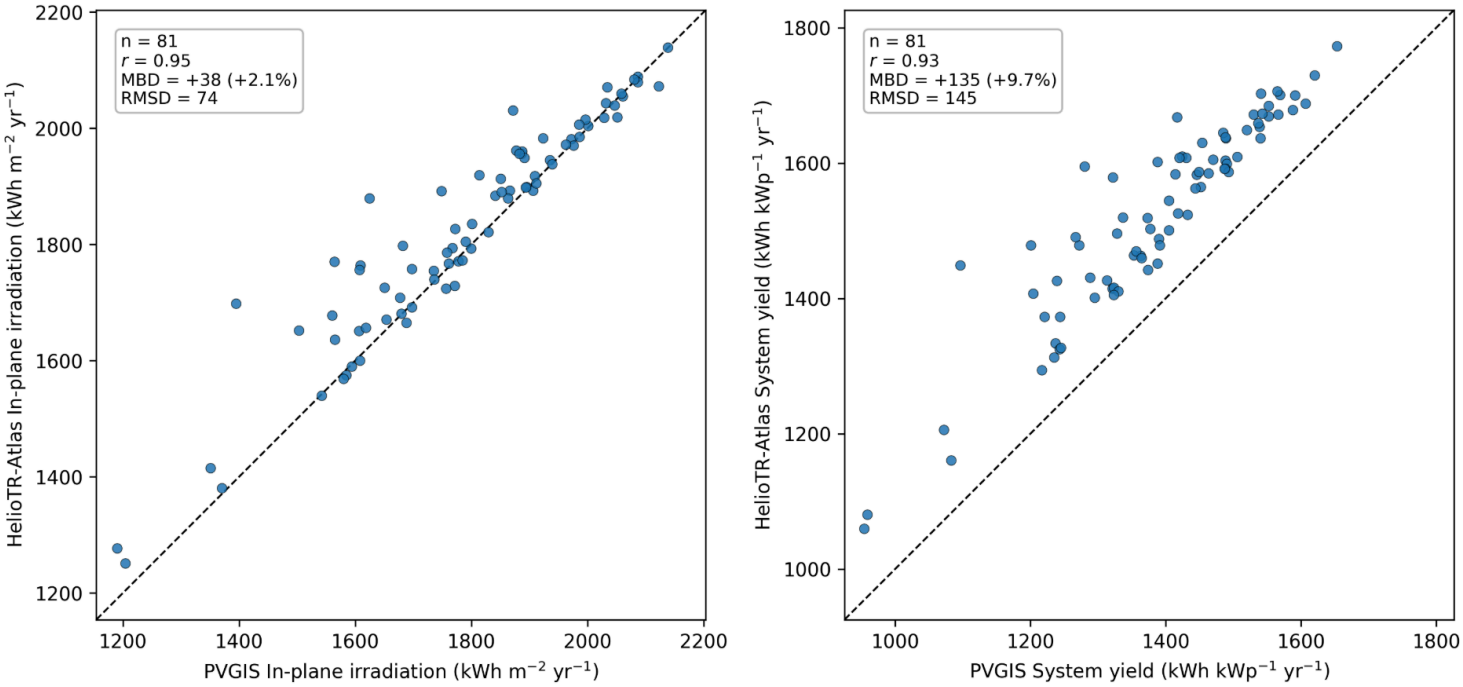


Figure 9

Cross-comparison with PVGIS at the 81 provincial centroids. (a) In-plane irradiation; (b) system yield, each plotted against the 1:1 line with the mean bias, root-mean-square deviation, and spatial correlation inset.