UC San Diego

UC San Diego Previously Published Works

Title

Localized Smart Persistence model for intra-hour solar irradiance forecasting across diverse climates

Permalink

https://escholarship.org/uc/item/68m5f21x

Journal

Solar Energy, 320

ISSN

0038-092X

Authors

Bahraminikoo, Parisa

Goh, Yong Kheng

Lim, Boon Han

et al.

Publication Date

2026-12-01

DOI

10.1016/j.solener.2026.115222

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Localized Smart Persistence Model for Intra-hour Solar Irradiance Forecasting across Diverse Climates

Parisa Bahraminikoo[a], Yong Kheng Goh[b], Boon Han Lim*[a], Sing Yee Chua[a], Kari Lappalainen[c], Jan Kleissl[d]

[a]*Department of Electrical and Electronic Engineering, Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, 43000 Kajang, Selangor, Malaysia*

[b]*Department of Mathematical and Actuarial Sciences, Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, 43000 Kajang, Selangor, Malaysia*

[c]*Electrical Engineering Unit, Tampere University, P. O. Box 692, Tampere FI-33101, Finland*

[d]*Center for Energy Research and Department of Mechanical and Aerospace Engineering, University of California at San Diego, La Jolla, California 92093, United States*

**Corresponding author: limbhan@utar.edu.my*

Abstract

Intra-hour solar irradiance forecasting is crucial for photovoltaic integration and grid operations. Smart Persistence (SP) is a widely used baseline, but its theoretical clear-sky scaling ratio may deviate from actual irradiance evolution under cloudy conditions and near sunrise/sunset. This study introduces Localized Smart Persistence (LSP), which replaces the clear-sky scaling ratio with a site-specific ratio derived from temporally averaged measured Global Horizontal Irradiance (GHI) profiles. The profiles are constructed by grouping historical daytime GHI measurements by time of day at the native temporal resolution of each dataset, and averaging them over a reference period. LSP was evaluated using temporal resolutions ranging from 30 seconds to 5 minutes and forecast horizons of 10 – 50 minutes across seven sites representing tropical, predominantly clear-sky, and high-latitude climatic conditions. Forecast performance was assessed using RMSE, MAE, R^2, and MBE, with nRMSE and the $NICE_\Sigma$ additionally evaluated at the common 5-minute resolution to facilitate cross-site comparisons. The results show that the effectiveness of LSP is strongly site-dependent. LSP generally reduced RMSE and MAE relative to SP at tropical and high-latitude sites, particularly at longer forecast horizons, while providing more representative irradiance evolution under low solar elevation conditions. In contrast, under predominantly persistence-dominated conditions, climatological scaling provided limited practical benefit. Approximately one year of historical irradiance data was generally sufficient to construct a representative climatological GHI profile for most sites. Overall, LSP should be regarded as a site-dependent extension of SP rather than a universally superior replacement, with potential extension to plane-of-array irradiance forecasting.

Keywords: Solar Irradiance Forecasting, Localized Smart Persistence, Tropical and Arid Climates, Temporal Resolution, Global Horizontal Irradiance, Persistence Models.

1. Introduction

Solar photovoltaics (PV) have been brought to the forefront in the global transition toward renewable energy. As the installed capacity of PV systems continues to increase, the variability of PV power generation caused by meteorological fluctuations has become a major challenge for power system operation and energy management. Sudden changes in solar irradiance, especially in areas with a high PV penetration, may affect grid stability and increase operational costs. Consequently, short-term solar irradiance

forecasting over timescales ranging from minutes to hours has become increasingly important [1–4]. Even modest improvements in forecast accuracy can produce meaningful operational benefits [5,6].

Persistence-based forecasting remains one of the most widely used baseline approaches for short-term solar irradiance prediction because of its simplicity, low computational cost, and competitive performance over short forecast horizons [6–8]. The simplest persistence formulation assumes that future irradiance equals current irradiance. This formulation is referred to here as conventional persistence.

Smart Persistence (SP) improves this assumption by persisting the clear-sky index (CSI) instead of irradiance directly [9,10]. By incorporating deterministic solar geometry through clear-sky normalization, SP generally performs better than simple persistence under stable atmospheric conditions and therefore continues to serve as a standard benchmark for evaluating more advanced forecasting approaches [6,10], including machine learning and physics-based models [5,7,11,12].

The performance of persistence-based forecasting methods depends strongly on climatic conditions, sky state, and temporal resolution. Under clear and steady conditions, SP remains a strong forecasting baseline, and more complex forecasting models often provide only limited improvement over short forecast horizons [1,13–15]. However, under intermittent cloud cover and rapidly changing atmospheric conditions, the performance differences between persistence-based forecasting approaches become much more pronounced [1,15].

Beyond SP, several persistence-based variants have also been proposed in the literature. These include climatology-based or hybrid climatology-persistence approaches that improve benchmark stability across different forecast horizons [16], stochastic persistence methods that smooth or statistically represent recent irradiance variability [1], and physics-informed or physically parameterized persistence approaches that incorporate cloud–radiation relationships or interactions among irradiance components [17]. These developments demonstrate that persistence forecasting has evolved in several directions to improve robustness, physical consistency, and forecast stability.

Nevertheless, important limitations still remain across persistence-based forecasting methods. Standard SP is sensitive to clear-sky mismatch, particularly in locations with variable cloud cover and persistently cloudy tropical climates, where theoretical clear-sky irradiance may substantially exceed the measured irradiance [21]. In high-latitude regions, prolonged periods of low solar elevation produce irradiance conditions similar to sunrise or sunset, during which the clear-sky irradiance may become very small, leading to increased sensitivity and reduced reliability of CSI scaling [2,11,17,18]. Additionally, forecast performance can be influenced by temporal resolution, which may smooth irradiance dynamics and affect the representation of short-term variability [19–21].

Pure climatological forecasting methods rely primarily on historical average irradiance profiles and may inadequately represent instantaneous sky conditions or rapidly changing atmospheric states during intra-hour forecasting. Hybrid climatology-persistence approaches partially address this issue by incorporating current irradiance information; however, climatological information is commonly introduced as an independent forecast component or weighted correction. Such formulations may smooth short-term irradiance variability, dilute transient atmospheric conditions, and treat climatological information as an external adjustment rather than as an intrinsic temporal evolution constraint within the persistence formulation itself.

Stochastic persistence methods can reduce noise or improve model robustness by averaging recent irradiance variability or representing it statistically in a compact form [1], but their performance depends strongly on the adopted stochastic formulation or smoothing strategy. Physically parameterized persistence approaches can further improve forecast skill by incorporating cloud-radiation relationships among irradiance components such as global horizontal irradiance (GHI), direct normal irradiance (DNI), and

diffuse horizontal irradiance (DHI) [17], or by applying cloud-property persistence assumptions. However, these approaches generally require additional irradiance components, physical parameterizations, or cloud-related assumptions [17,18].

More importantly, most existing persistence variants retain the conventional clear-sky-based scaling structure of SP, while climatological information is usually introduced as an external benchmark or auxiliary correction term. As a result, it remains unclear whether modifying the scaling reference inside the SP formulation itself can produce a forecasting baseline that remains simple while being more representative of local irradiance evolution. The aforementioned limitations highlight the need for a forecasting baseline that remains computationally simple while better preserving instantaneous irradiance states and capturing site-specific irradiance evolution [13,14,22].

To address this gap, this study introduces the Localized Smart Persistence (LSP) model as a site-specific modification of SP. Instead of using a theoretical clear-sky profile as the scaling reference, LSP constructs a locally averaged diurnal irradiance profile directly from historical high-resolution GHI measurements. The theoretical clear-sky scaling ratio in SP is then replaced by a climatological irradiance ratio derived from this local climatological GHI profile. In this sense, LSP can be interpreted as a climatology-informed extension of SP. However, unlike conventional climatology-persistence approaches, the climatological information is incorporated directly into the persistence scaling mechanism rather than being treated as a separate benchmark forecast or external correction.

This distinction is important because theoretical clear-sky normalization may not adequately reflect persistent local influences such as recurring cloud attenuation, aerosol loading, terrain masking, horizon obstruction, or systematic intraday asymmetries in irradiance behavior. These effects can become particularly relevant in tropical, coastal, mountainous, or high-latitude environments where actual irradiance evolution may deviate consistently from the idealized clear-sky assumptions [2].

In the present study, LSP is implemented using a single annual climatological irradiance profile in order to preserve methodological simplicity and maintain a transparent baseline formulation across diverse climates. However, the framework itself is not restricted to a single annual profile. Depending on climatic characteristics and application requirements, the climatological GHI profile may also be constructed using seasonal, monthly, or shorter adaptive time windows to better capture evolving irradiance patterns. This flexibility is particularly relevant for mid- and high-latitude environments where seasonal irradiance variations are more pronounced.

The proposed LSP model was evaluated across climatically diverse locations, including arid and mountainous regions in the United States, high-latitude conditions in Finland, and coastal environments, to assess its broader applicability [13,14]. This study makes two main contributions. First, it introduces a localized reformulation of SP in which the conventional clear-sky scaling ratio is replaced by a locally representative irradiance scaling ratio derived from measured local irradiance profiles. Second, it systematically evaluates the effectiveness and limitations of this reformulated baseline under diverse climatic conditions, thereby clarifying when locally representative persistence scaling may provide advantages over conventional SP.

In addition to GHI forecasting, the proposed framework can also be extended directly to global tilted irradiance (GTI) or plane-of-array (POA) irradiance forecasting when historical tilted irradiance measurements are available. In contrast, SP-based GTI forecasting generally requires additional clear-sky transposition procedures to convert horizontal irradiance into tilted irradiance, potentially introducing further modeling uncertainty associated with decomposition and sky transposition models. Thus, LSP may provide a simple and transparent framework for forecasting irradiance quantities more directly related to practical PV system operation.

To facilitate comparison among the persistence-based forecasting formulations discussed above, Table 1 summarizes their underlying reference structures, roles of climatological information, and principal limitations, including the proposed LSP framework.

Table 1 highlights that the novelty of LSP is not the use of climatological information itself, but its direct integration into the scaling mechanism of SP. LSP retains the current observed GHI while replacing the theoretical clear-sky scaling ratio with an empirical ratio derived from the site-specific historical diurnal irradiance profile. This localized reformulation preserves the simplicity of persistence-based forecasting while capturing recurring site-specific irradiance patterns.

Table 1. Comparison of persistence-based reference formulations for short-term solar irradiance forecasting.

Method family	Persisted quantity	Reference/scaling basis	Role of climatological information	Key characteristic and main limitation
Simple persistence	Current irradiance	No scaling reference	None	Ignores solar geometry and recurring intraday irradiance evolution
Smart persistence (SP)	Clear-sky index (CSI)	Theoretical clear-sky irradiance profile	None	Sensitive to clear-sky mismatch, particularly under variable cloud cover and low irradiance conditions
Climatology-based reference	Historical mean irradiance	Historical average irradiance profile	Primary forecasting reference	Does not preserve instantaneous sky conditions or transient irradiance variability
Hybrid climatology-persistence	Persistence and climatology combination	Weighted or auxiliary climatological correction	External climatological adjustment	May smooth short-term irradiance variability and requires a combination strategy
Stochastic persistence	Statistical representation of irradiance variability	Statistical or smoothing formulation	Limited/indirect	Performance depends on stochastic formulation and smoothing design
Physics-informed persistence	Cloud or irradiance-related physical predictors	Physically parameterized cloud-radiation relationships	Limited/indirect	Requires additional irradiance components, physical assumptions, or cloud-related parameterizations
Proposed LSP	Local irradiance transition ratio	Empirical local diurnal irradiance profile	Directly embedded within persistence scaling	Performance depends on the representativeness of the selected climatological profile

The paper is organized as follows: Section 2 presents the methodology, including data collection, model development, and the evaluation metrics, while Section 3 presents and discusses the results. Finally, Section 4 concludes the work.

2. Methodology

2.1. Data Collection and Pre-processing

To evaluate the generalizability and performance of LSP, ground-based GHI measurement datasets were collected from seven geographically diverse sites with distinct climatic characteristics. These sites include three equatorial sites in Malaysia, which are characterized by limited seasonal variation, a high-latitude region in Finland with strong seasonal and diurnal contrasts, and three sites from the United States (New

Mexico, Colorado, Oregon) characterized by arid or mountainous climates and complex atmospheric dynamics. The diversity of sites ensures robust cross-climatic validation of the model.

The datasets were obtained from multiple sources. There are three datasets from the United States. The University of New Mexico dataset was accessed from the Dryad repository [23]. The Colorado (NREL Solar Radiation Research Laboratory) and Oregon (SRML) [24] datasets were collected from the NREL Measurement and Instrumentation Data Center (MIDC) [25], which provides high-quality and continuous meteorological and solar radiation records. The Finland dataset originates from the Solar PV Power Station Research Plant of Tampere University, as described by Torres Lobera et al. [26]. The datasets from Malaysia comprised measurements from Universiti Tunku Abdul Rahman (UTAR) and regional monitoring stations at Bukit Kayu Hitam (BKH) and Kampung Kota (EVN).

Temporal resolution varies depending on data availability and site characteristics. For UTAR, the University of New Mexico, and Tampere University, three temporal resolutions (30 seconds, 1 minute, and 5 minutes) were evaluated to investigate the influence of temporal resolution on forecasting performance. At NREL and SRML, both 1-minute and 5-minute datasets were available and analyzed, whereas only datasets with a 5-minute temporal resolution were available at BKH and EVN. For these datasets, lower-resolution series were generated through systematic subsampling rather than temporal averaging. For example, the 1-minute series at UTAR was constructed by selecting observations at 1-minute timestamps from the native 30-second measurements (e.g., 08:00:00, 08:01:00, 08:02:00, ...). This approach preserves instantaneous irradiance characteristics and avoids smoothing short-duration fluctuations that may be relevant to rapid irradiance variability and grid-related effects such as overvoltage, voltage sags, and inverter clipping behavior. However, subsampling may alter the representation of high-frequency variability compared with temporally averaged datasets, and this distinction should be considered when interpreting temporal-resolution effects.

Before constructing the forecasting model, the GHI records from each site were checked and cleaned. Obvious outliers, mainly caused by sensor malfunction or temporary shading from nearby objects, were removed. These points represented less than 2% of the data at most sites. Short gaps in the time series were filled by linear interpolation only when no more than a few consecutive minutes were missing. Longer contiguous missing segments were not interpolated and were excluded from profile construction. This interpolation was performed only to maintain continuity of the profile for time-based averaging purposes and was not used to replace long missing periods.

Following this, two datasets were prepared for each location: a reference dataset used to derive the site-specific climatological GHI profile, i.e., the time-of-occurrence averaging GHI profile for LSP calculation, and a test dataset for model validation. The time-of-occurrence averaging datasets spanned 180–720 days, capturing representative solar irradiance behavior while minimizing the influence of anomalous weather conditions. The test datasets comprised 60–365 days and were entirely disjoint from the reference datasets to avoid data leakage. The chronological separation between reference and test datasets was intentionally maintained to better reflect realistic forecasting conditions using historical irradiance observations. Shorter windows (<90 days) were found to introduce noise associated with anomalous weather events, whereas longer windows (>720 days) tended to dilute site-specific climatological patterns present in multi-year datasets [30]. Measurements were limited to daytime hours to align with solar availability. Timestamp consistency was additionally verified by comparing measured irradiance behavior against expected clear-sky diurnal patterns during clear-sky periods. The aggregation was implemented in Python using NumPy and Pandas, following the time-of-occurrence averaging procedure described above. Site specifications, including geographic coordinates, dataset durations, and temporal resolutions, are summarized in Table 2.

Table 2. Specifications of the seven study sites, including geographic coordinates, dataset durations, and temporal resolutions used in the evaluation of the LSP model.

Site	Location	Geographic coordinates		Dataset		
		Latitude (°)	Longitude (°)	No. of days used for time-of-occurrence averaging	No. of days used for testing	Temporal resolution
UTAR	Selangor, Malaysia	3.040	101.794	204 days Selected dates from 2023 and 2024	160 days 20 Jan 2025–17 Jul 2025	30 s, 1 min, 5 min
University of New Mexico	New Mexico, USA	35.082	-106.625	180 days 8 Dec 2017–14 Oct 2018	60 days 15 Oct 2018–21 Jan 2019	30 s, 1 min, 5 min
University of Tampere	Tampere, Finland	61.443	23.856	275 days 4 Feb 2019–3 Nov 2019	275 days 20 Feb 2021–21 Nov 2021	30 s, 1 min, 5 min
NREL	Golden, Colorado, USA	39.742	-105.178	720 days 2022–2023	365 days 2024	1 min, 5 min
SRML	Eugene, Oregon, USA	44.046	-123.074	720 days 2022–2023	365 days 2024	1 min, 5 min
BKH	Kedah, Malaysia	6.519	100.42	365 days 2018	180 days 1 Jan 2019–30 Jun 2019	5 min
EVN	Penang, Malaysia	5.434	100.33	265 days Selected dates in 2019	100 days Selected dates in 2019	5 min

The LSP reference profile was constructed separately for each site using time-of-occurrence averaging. For each time of day τ, all available daytime GHI measurements recorded at the same timestamp across the reference days were grouped and averaged to obtain a mean diurnal GHI profile:

$$GHI_{clim}(t) = \frac{1}{N_t}\sum_{d=1}^{N_t} GHI_d(t) \qquad (1)$$

where t denotes the time of day, $GHI_d(t)$ is the measured irradiance at timestamp t on day d, and N_t is the number of valid observations available at that timestamp across the reference period. This produced a local mean diurnal GHI curve at the native temporal resolution of each dataset. For a given site, both baselines used the same pre-processing steps, temporal resolution, reference period, and test period.

To distinguish the effect of the forecasting formulation from differences among the site datasets, validation was performed independently at each location. At each site, SP and LSP were evaluated using identical forecast–observation samples, preprocessing procedures, temporal resolutions, forecast horizons, and disjoint reference and test periods. Therefore, within-site differences between SP and LSP are not affected by inconsistent data handling.

2.2. Model Development

2.2.1. Smart Persistence (SP) Model

The starting point for building the LSP model is the SP model, which has established itself as a cornerstone in short-term solar irradiance forecasting due to its empirical basis in persistence principles, computational efficiency, and its reliance on minimalistic design logic [10,27]. The SP model relies on the observation that the CSI, a ratio comparing the observed solar irradiance to the idealized clear-sky conditions, does not change much over the short term. It forecasts GHI at time $t + h$ (h is the temporal forecast horizon) as:

$$\widehat{GHI}_{SP}(t+h) = GHI_{obs}(t) \times \frac{GHI_{clear}(t+h)}{GHI_{clear}(t)} = CSI(t) \times GHI_{clear}(t+h). \quad (2)$$

Here, $\widehat{GHI}_{SP}(t+h)$ denotes the SP forecast for the target time $t+h$; $GHI_{obs}(t)$ is the most recent observed GHI at the forecast issue time t; $GHI_{clear}(t)$ and $GHI_{clear}(t+h)$ are the modeled clear-sky GHI values at the issue and target times, respectively. The clear-sky values are derived from the REST2 radiative transfer model [27,28], which uses the solar zenith angle, aerosol optical depth, water vapor, and ozone content to simulate the irradiance under cloud-free conditions. Site-specific solar geometry parameters (latitude, longitude, elevation, and time) were obtained from site metadata, while atmospheric inputs such as aerosol optical depth, water vapor, and ozone were prescribed using standard climatological atmospheric profiles, following the default REST2 configuration as described in Gueymard [28] and Yang et al. [5]. This approach was adopted to ensure methodological consistency across all sites and to reflect common practice in baseline smart persistence implementations, rather than to optimize clear-sky accuracy at individual locations. While more advanced site-specific atmospheric datasets (e.g., reanalysis or AERONET-derived inputs) are available, their use would introduce additional data dependencies and complexity that are not aligned with the objective of establishing a simple and globally reproducible baseline model.

The *CSI* is defined as $CSI\ (t) = GHI_{obs}(t)/GHI_{clear}(t)$ and normalized for diurnal and seasonal solar geometry, enabling the SP model to project short-term GHI deviations due to solar geometry based on current observations [7]. However, the efficacy of the SP method deteriorates as the difference between the atmospheric conditions and the assumptions made for clear-sky conditions becomes systematic. Although SP is correct on average, the adjustment ratio for clear-sky conditions only considers the solar geometry for the corresponding time of day and implicitly assumes that cloud attenuation remains constant over the short forecast horizon. This hypothesis becomes less valid for areas whose cloud conditions change systematically throughout the day, such as areas near the tropics that often experience afternoon convection.

The main difference between the SP and LSP adjustment ratios lies in their conceptual treatment of irradiance evolution. While SP uses the ratio between clear-sky GHI values for adjusting the forecasting result, LSP uses an empirical mean diurnal GHI profile that naturally captures diurnal cloud evolution and solar-geometry effects. This means that the adjustment ratio of LSP can naturally describe the evolution pattern of GHI values even in cases where cloud cover does not change appreciably within the forecast window. This allows LSP to produce more reliable and physically consistent forecasts without relying on the assumption of constant cloud attenuation implicit in SP [5].

2.2.2. LSP Model

To mitigate these limitations, we propose the LSP model, which refines the SP framework by replacing the clear-sky–based adjustment ratio ($GHI_{clear}(t+h)/GHI_{clear}(t)$) with a site-specific, climatological GHI ratio. This modification shifts the model from a theoretical radiative baseline to a data-driven one, explicitly tailored to the local climatology. The LSP model is formulated as:

$$\widehat{GHI}_{LSP}(t+h) = GHI_{obs}(t) \times \frac{\overline{GHI}_{clim}(t+h)}{\overline{GHI}_{clim}(t)}. \quad (3)$$

Here, $\widehat{GHI}_{LSP}(t+h)$ denotes the LSP forecast for the target time t+h; $GHI_{obs}(t)$ is the most recently observed GHI at the forecast issue time t; and $GHI_{clim}(t)$ and $\overline{GHI}_{clim}(t+h)$ denote the values extracted from the precomputed local climatological mean GHI profile at the issue and target times, respectively. The ratio $GHI_{clim}(t+h)/GHI_{clim}(t)$, represents the dimensionless localized climatological scaling factor. Unlike SP's date-dependent clear-sky library, LSP generally employs a single annual climatological GHI profile computed over a representative period. However, season-specific profiles were used for Tampere because of its pronounced seasonal irradiance variation and the exclusion of low-irradiance winter periods, as discussed further in Section 3.4.

This climatological GHI profile embeds recurring local effects, such as mean diurnal changes in cloud attenuation and aerosol influences, into the scaling ratio $GHI_{clim}(t+h)/GHI_{clim}(t)$, thereby relaxing the clear-sky assumption, enhancing robustness under non-clear-sky-dominated conditions, and reducing sensitivity to uncertainties in clear-sky model parameterization.

Consequently, the scaling mechanism in LSP differs conceptually from additive climatological correction approaches. The multiplicative formulation was retained because LSP aims to preserve the proportional irradiance transition behavior inherent to persistence-based forecasting. Instead of applying climatological information as an external additive correction, LSP scales the current irradiance according to the relative evolution of the local climatological irradiance profile. This allows the forecast to better preserve the instantaneous irradiance state and localized temporal evolution of sky conditions while adapting the persistence transition to site-specific irradiance behavior.

$$\widehat{GHI}_{LSP,add}(t+h) = GHI_{obs}(t) + [\overline{GHI}_{clim}(t+h) - \overline{GHI}_{clim}(t)]. \quad (4)$$

As a sensitivity alternative, an additive localized persistence formulation can be defined as shown in Equation (4), where the absolute climatological change between the forecast issue time and the target time is added to the current observed GHI. Unlike the multiplicative LSP formulation, this approach avoids division by potentially small climatological irradiance values. However, additive formulations implicitly assume that absolute climatological irradiance changes can be transferred directly to the current irradiance state, regardless of the prevailing irradiance magnitude or cloud attenuation level. This may become less representative during rapidly varying atmospheric conditions or under strong transient irradiance fluctuations. Therefore, the multiplicative formulation was adopted in the present study, while acknowledging that additive approaches may provide improved numerical stability under certain low-irradiance conditions.

A minimum tolerance of 1 W/m² was used as a numerical safeguard and daytime screening criterion for ratio-based persistence calculations, primarily to avoid division by zero or near-zero denominator values during very low irradiance periods near sunrise and sunset [11]. This threshold should not be interpreted as a physically defined irradiance boundary, but rather as a practical stability criterion commonly adopted in ratio-based irradiance formulations. For LSP, the denominator is a measured mean diurnal GHI profile rather than an instantaneous clear-sky irradiance. In the datasets used in this study, the climatological GHI profile for the retained daytime forecast periods generally remained well above the 1 W/m² tolerance, such that the threshold had negligible influence on normal daytime LSP forecasts and primarily affected the exclusion of very low-light boundary periods from evaluation.

Nevertheless, low irradiance conditions near sunrise, sunset, and high-latitude winter periods remain intrinsically challenging for multiplicative ratio-based forecasting approaches because small denominator values can increase numerical sensitivity and amplify relative forecast variability. Accordingly, the adopted

threshold serves mainly as a practical numerical stabilization measure rather than a complete physical solution to low-irradiance forecasting limitations.

The reference profile was subsequently integrated into a forecasting algorithm. The scaling ratio was dynamically applied to real-time $GHI_{obs}(t)$. Different forecast horizons of 10, 20, 30, 40, and 50 minutes were evaluated [3]. Sensitivity to temporal resolution was evaluated at 30-second, 1-minute, and 5-minute resolutions.

Figure 1 presents the mean diurnal GHI profiles for the UTAR site, along with the clear-sky GHI and the actual GHI values for one example day. It illustrates the core scaling mechanism used in both SP and LSP models. The figure contrasts the clear-sky profiles used in SP with the local mean diurnal GHI profile in LSP, applied to real-time GHI data from the UTAR site. The figure shows that the climatological GHI profile used in LSP more closely follows the measured irradiance behavior than the theoretical clear-sky GHI curve used in SP, particularly under persistently cloudy tropical conditions.

At 12:45 pm, for example, the clear-sky GHI increases by approximately 1% over the following 20 minutes, whereas the climatological GHI decreases by 4%, which better matches the observed GHI decrease of 7%. This illustrates how the LSP scaling ratio can better capture localized irradiance evolution associated with recurring afternoon cloud development. Over the full day, the mean diurnal GHI profile is shifted slightly toward earlier hours relative to the clear-sky curve, reflecting the tendency for increased afternoon cloudiness caused by tropical convective activity. Consequently, the SP model tends to overpredict irradiance during cloudier afternoon periods, whereas LSP more effectively captures the climatological evolution of cloud attenuation.

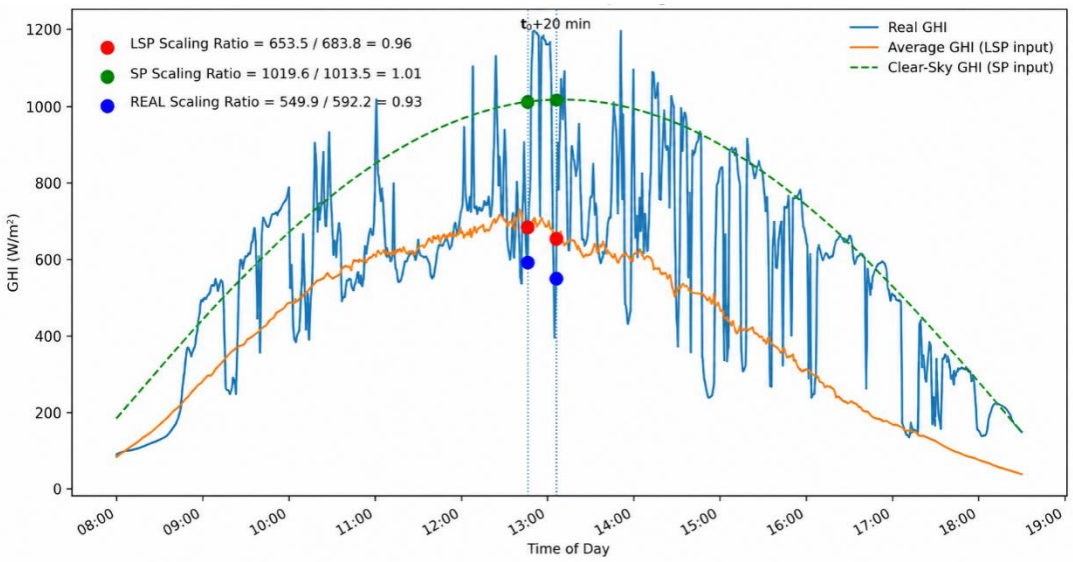


Figure 1. Relative representation of SP and LSP scaling ratios using real-time UTAR GHI data on 24 April 2025. SP scales by simulated clear-sky GHI values (green dashed line), while LSP scales forecasts by the mean local GHI profile (orange line).

In summary, the architecture of the LSP model combines the persistence principle with site-specific empirical irradiance data. The comparison between SP and LSP in this study is performed under identical site-specific conditions, using the same measured GHI input, forecast horizons, pre-processing pipeline, and test periods. In this sense, the reported differences are intended to reflect the effect of the modified scaling formulation rather than inconsistencies in data handling. At the same time, because LSP is explicitly

site-specific, the resulting performance should be interpreted as localized baseline behavior conditioned on local climatology rather than as evidence of universal cross-site generalization.

The present study focuses on deterministic point forecasting because the primary objective is to evaluate the effectiveness of the proposed LSP formulation as a simple baseline extension of Smart Persistence (SP). Accordingly, the analysis emphasizes direct comparison of deterministic forecast accuracy across different climatic conditions and temporal resolutions. Probabilistic forecasting and uncertainty quantification, including prediction intervals and confidence estimation, may provide additional operational insights and therefore represent potential directions for future work.

2.2.3. Simplified BLEND Persistence Benchmark

To provide an additional persistence-based benchmark, the Simplified Cyclostationary BLEND operator proposed by Voyant et al. [31] was implemented at UTAR. The model combines simple persistence with the phase-aligned irradiance observation from the previous daily cycle:

$$\widehat{GHI}_{BLEND}(t+h) = [1-\lambda(t,h)]GHI_{obs}(t-T+h) + \lambda(t,h)GHI_{obs}(t) \qquad (5)$$

where T was set to one day, and the phase-dependent weight was calculated as:

$$\lambda(t,h) = \frac{1}{2}[1+\rho(t,h)] \qquad (6)$$

where $\rho(t,h)$ is the Pearson correlation between the irradiance at the forecast issue time t and the irradiance at the target time t+h, estimated across the historical reference days for each time-of-day phase and forecast horizon.

For UTAR, the correlations were estimated from the 2023–2024 reference data at exact 5-minute timestamps without temporal averaging or interpolation. They were calculated separately for the 10-, 20-, 30-, 40-, and 50-minute forecast horizons using at least 30 valid historical pairs; when a correlation could not be estimated, $\rho(t,h)$ was set to 0, resulting in $\lambda(t,h) = 0.5$. LSP and Simplified BLEND were evaluated using identical common samples from the 2025 test data, and the performance metrics were calculated daily and then averaged across the valid test days.

2.3. Evaluation Metrics and Validation

To quantitatively assess the performance of the LSP and baseline SP models, four evaluation metrics were used: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Bias Error (MBE), and the Coefficient of Determination (R^2). These metrics are widely used in solar irradiance forecasting because they provide complementary information on forecast accuracy, bias, and the ability of the model to reproduce observed variability [5,9,29].

The RMSE measures the square root of the average squared differences between predicted and observed values, emphasizing larger errors as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(GHI_{obs,i} - GHI_{predicted,i}\right)^2} \qquad (7)$$

Here, n represents the number of observations in the test data set at the given temporal resolution. $GHI_{obs,i}$ denotes the measured GHI at time step i, and $GHI_{predicted,i}$ denotes the corresponding forecasted value. Lower RMSE values indicate higher prediction accuracy.

The MAE measures the average magnitude of the absolute forecast errors as:

$$MAE = \frac{1}{n}\sum \left| GHI_{obs,i} - GHI_{predicted,i} \right| \quad (8)$$

Unlike RMSE, MAE does not place as much weight on larger errors and therefore provides a more direct indication of the typical forecast deviation.

The MBE measures the average signed forecast error as:

$$MBE = \frac{1}{n}\sum \left(GHI_{predicted,i} - GHI_{obs,i} \right) \quad (9)$$

This metric indicates whether the forecast tends to overestimate or underestimate the observed irradiance systematically. Positive MBE values indicate overprediction on average, whereas negative values indicate underprediction.

The R^2 score evaluates the proportion of variance in the observed data explained by the model as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(GHI_{obs,i} - GHI_{predicted,i} \right)^2}{\sum_{i=1}^{n} \left(GHI_{obs,i} - GHI_{obs} \right)^2} \quad (10)$$

Here GHI_{obs} represents the average of the observed GHI data points. The statistic ranges from $-\infty$ to 1, with values closer to 1 suggesting better correspondence between prediction and observation. Negative R^2 values indicate that the forecast performs worse than using the mean observed irradiance as a predictor. Therefore, R^2 is used here as a supporting metric and is interpreted together with RMSE, MAE, and MBE rather than as a standalone indicator of forecast quality.

To support scale-independent and persistence-referenced comparisons, nRMSE and the empirical $NICE_{\Sigma}$ indicator [32] were additionally calculated. The nRMSE was normalized by the mean observed daytime GHI. $NICE_{\Sigma}$ was calculated relative to simple persistence, where values below 1 indicate performance better than simple persistence. To ensure a consistent comparison across all seven sites, these complementary metrics were evaluated at the common 5-minute temporal resolution for representative forecast horizons of 10, 30, and 50 minutes.

The nRMSE for forecasting model X was calculated as:

$$nRMSE_X(\%) = \frac{RMSE_X}{\overline{GHI}_{obs}} \times 100 \quad (11)$$

where $\overline{GHI}_{obs}$ is the mean observed daytime GHI. Simple persistence was defined as:

$$\widehat{GHI}_p(t+h) = GHI_{obs}(t) \quad (12)$$

The empirical $NICE^k$ value for model X was calculated relative to simple persistence as:

$$NICE_X^k = \frac{\left[\frac{1}{N}\sum_{i=1}^{N}|GHI_{obs,i}-GHI_{X,i}|^k\right]^{1/k}}{\left[\frac{1}{N}\sum_{i=1}^{N}|GHI_{obs,i}-GHI_{P,i}|^k\right]^{1/k}}, k = 1,2,3 \quad (13)$$

The equally weighted combined indicator was calculated as:

$$NICE_{\Sigma,X} = \frac{NICE_X^1+NICE_X^2+NICE_X^3}{3} \quad (14)$$

Here, X denotes the evaluated forecasting model, either SP or LSP; N is the number of valid forecast-observation pairs; i denotes the observation index; h is the forecast horizon; and k is the error-order parameter, with k=1,2 and 3. $GHI_{obs,i}$, $GHI_{X,i}$, and $GHI_{P,i}$ denote the observed GHI, the forecast produced by model X, and the simple-persistence forecast, respectively.

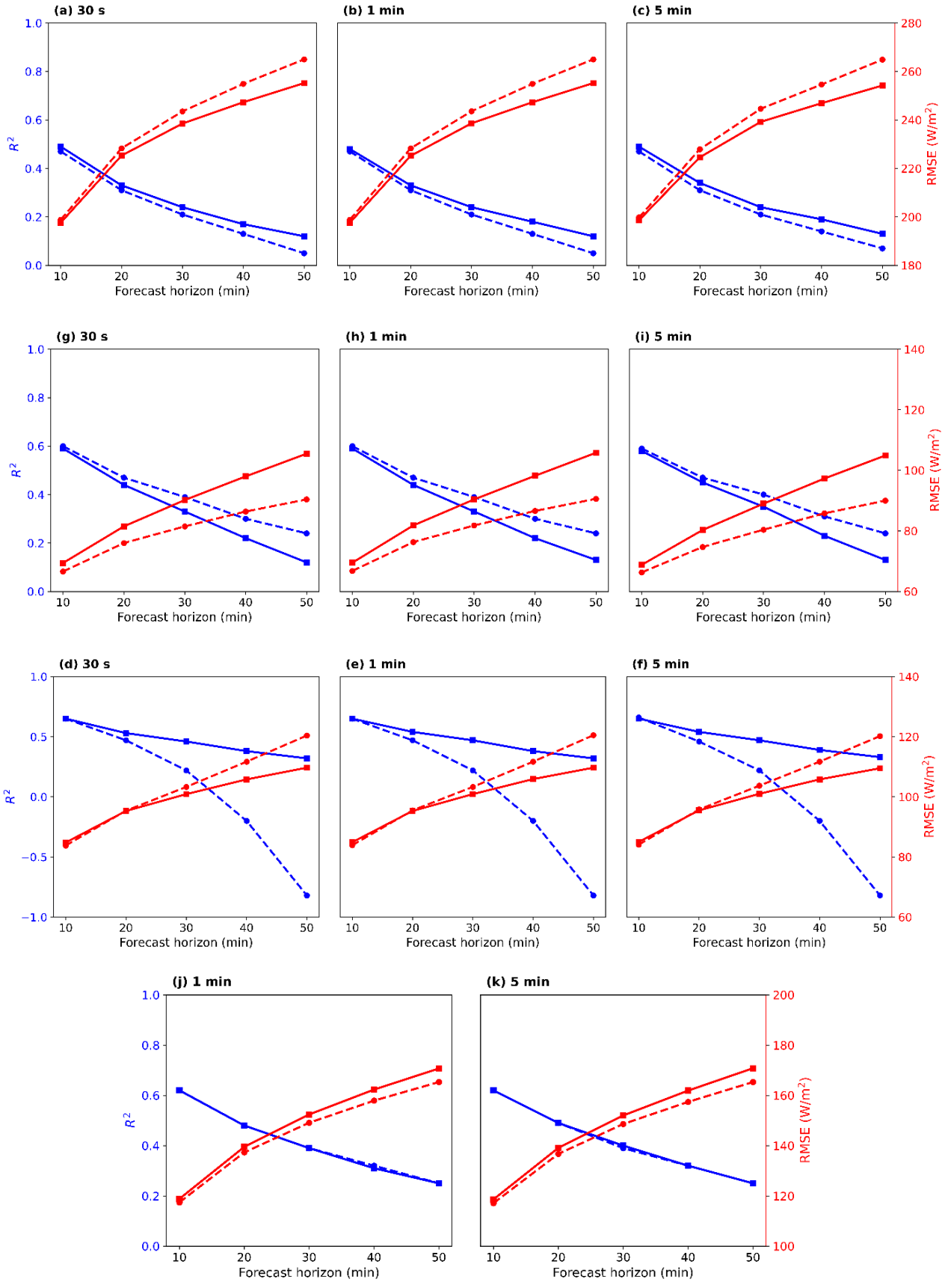
Consistent with the existing site-level evaluation procedure, nRMSE and $NICE_\Sigma$ were calculated separately for each valid test day and subsequently averaged across all valid days for each site and forecast horizon. To ensure a consistent cross-site comparison, nRMSE and $NICE_\Sigma$ were evaluated at the common 5-minute temporal resolution for representative forecast horizons of 10, 30, and 50 minutes. Lower $NICE_\Sigma$ values indicate better performance: values below 1 indicate improvement over simple persistence, whereas values above 1 indicate poorer performance.

3. Results and Discussion

3.1. Forecast Performance of SP and LSP across Sites

The LSP model was tested using stations that represent different climatic and geographic conditions, including tropical, arid, temperate, predominantly clear-sky, and high-latitude environments. The objective was not only to quantify the average improvement relative to SP, but also to identify the climatic conditions under which LSP provides meaningful benefits and those where it offers limited practical advantage. The main comparisons are presented in Figures 2 and 3 for temporal resolutions ranging from 30 seconds to 5 minutes and forecast horizons from 10 to 50 minutes, while the full numerical results are provided in Table A1. In Figure 3, the SP/LSP RMSE ratio is used as a relative performance indicator, where values greater than 1 indicate lower RMSE for LSP. Additional MAE and MBE results are reported in Table A2 in the appendix. In general, MAE supports the main RMSE-based trends, while MBE provides complementary information regarding the direction and magnitude of forecast bias.

Complementary nRMSE and $NICE_\Sigma$ results at the common 5-minute temporal resolution are presented in Table A3. Additional analyses of the LSP formulation, the Simplified BLEND benchmark, and the climatological averaging strategy are presented separately in Sections 3.2–3.4.


(a) 30 s
(b) 1 min
(c) 5 min
(g) 30 s
(h) 1 min
(i) 5 min
(d) 30 s
(e) 1 min
(f) 5 min
(j) 1 min
(k) 5 min
R^2
RMSE (W/m^2)
Forecast horizon (min)

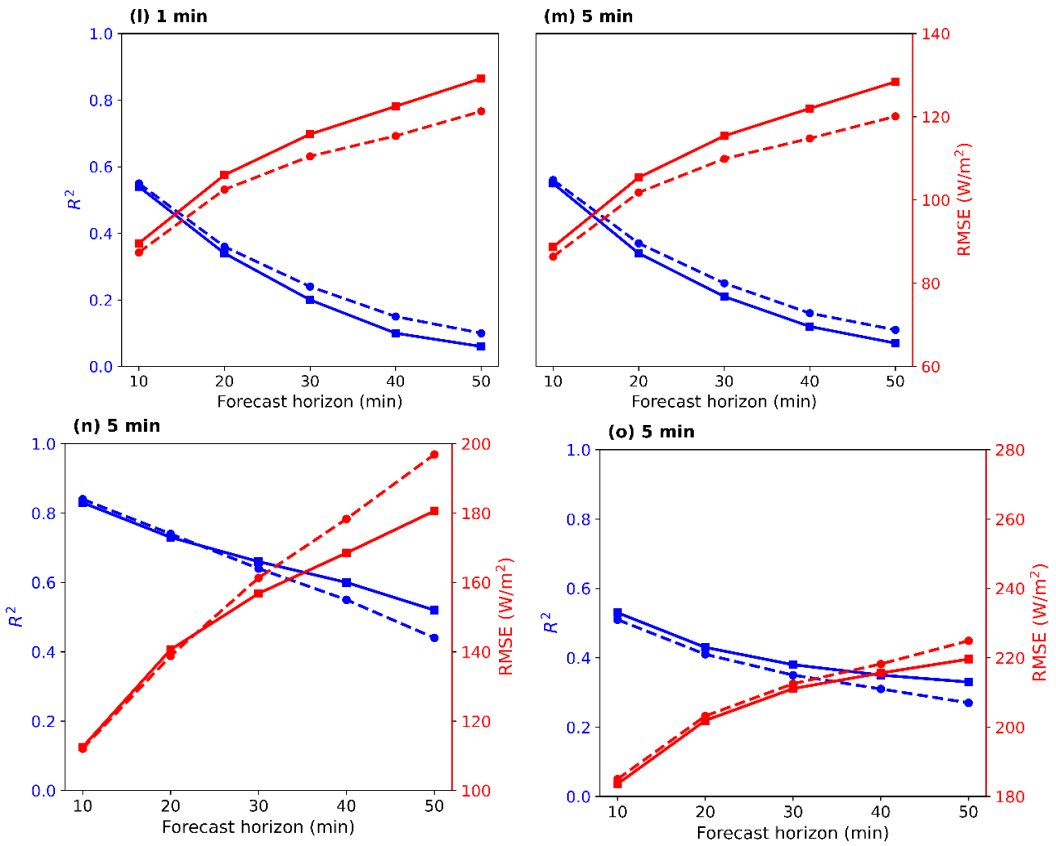


Figure 2. Performance comparison of SP and LSP across forecast horizons using R^2 (left axis) and RMSE (right axis). Panels (a–c), UTAR; panels (d–f), Tampere; panels (g–i), New Mexico; panels (j–k), NREL; panels (l–m), SRML; panel (n), EVN; and panel (o), BKH. The temporal resolution is identified above each panel. Dashed lines with circular markers denote SP, whereas solid lines with square markers denote LSP. Blue and red lines represent R^2 and RMSE, respectively.

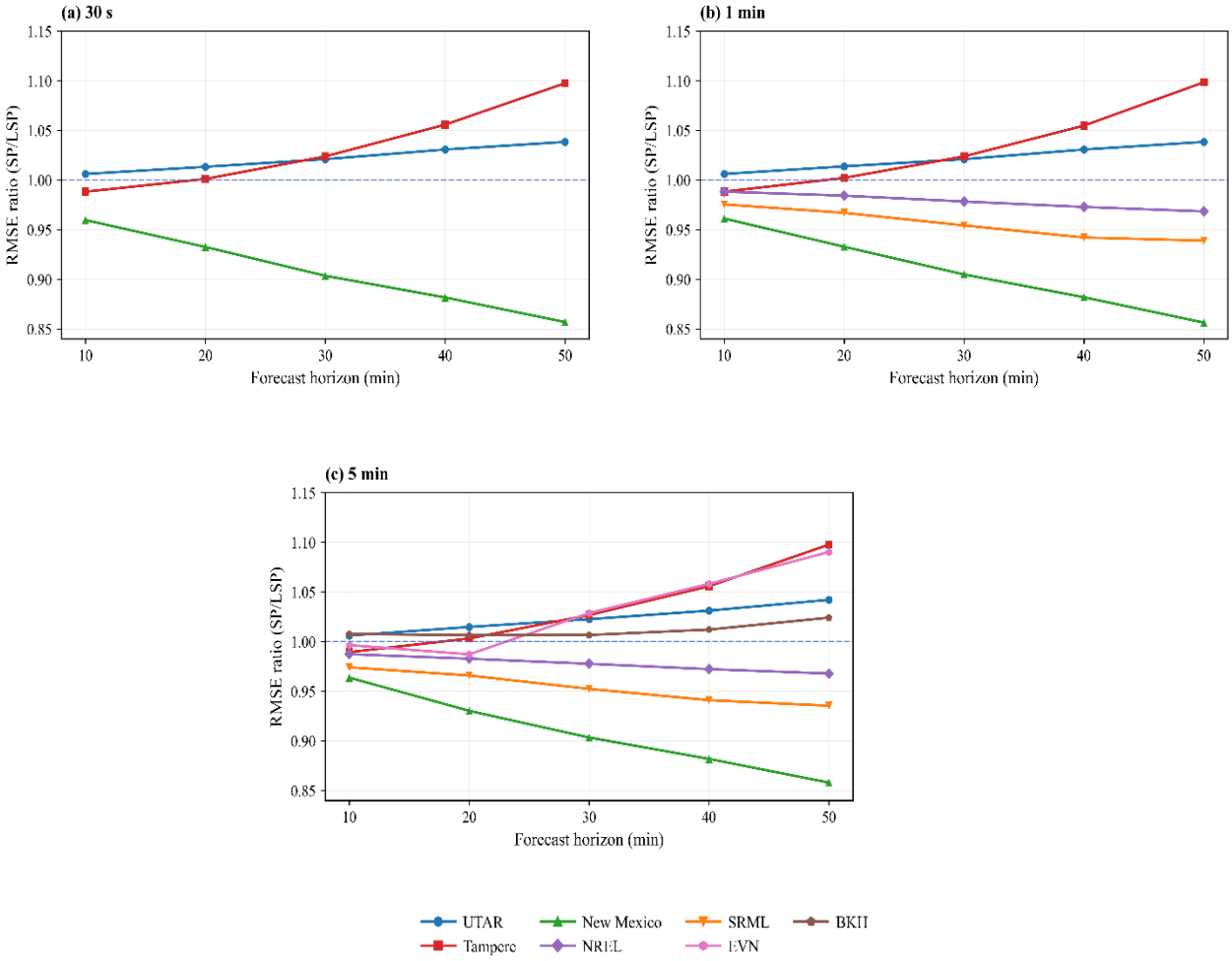


Figure 3. RMSE ratio (SP/LSP) across forecast horizons at temporal resolutions of (a) 30 s, (b) 1 min, and (c) 5 min. Values greater than 1 indicate lower RMSE and better performance for LSP, whereas values below 1 indicate lower RMSE and better performance for SP. The horizontal dashed line represents equal RMSE for the two models. Colors and markers identify the study sites, as indicated in the shared legend.

At UTAR, LSP performs consistently better than SP, although the improvement remains modest. In terms of RMSE, the reduction is approximately 0.6–4% and becomes more noticeable as the forecast horizon increases, particularly beyond 30 minutes. A similar trend is observed in MAE, where the difference between the two models is minimal at shorter horizons but gradually favors LSP at longer lead times. The R^2 values are also slightly higher for LSP, indicating a small but consistent improvement in producing observed irradiance variability. At the same time, the MBE results show that SP tends to overpredict irradiance at this site, whereas LSP shifts the bias toward slight underprediction. This suggests that the use of a local climatological GHI profile not only moderately reduces forecast error, but also alters the direction of the systematic forecast bias. The complementary nRMSE and $NICE_{\Sigma}$ results at the common 5-minute temporal resolution showed a similar overall tendency, with LSP generally favored over SP, particularly at the longer forecast horizons. Overall, the UTAR results indicate that the advantage of the climatological scaling used in LSP becomes more pronounced as the forecast horizon increases and the assumption of direct persistence becomes less adequate.

A similar behavior can be observed at the two Malaysian monitoring stations, although the magnitude of improvement differs between them. At EVN, the advantage of LSP is limited at the shorter forecast horizons but becomes more evident as the lead time increases. RMSE decreases by approximately 2.8–8.3% between the 30- and 50-minute horizons, and MAE also shows a corresponding improvement at longer lead times. The R^2 values also increase slightly, suggesting that the local climatological GHI profile becomes more informative as short-term irradiance evolution becomes more difficult for standard persistence to represent. The MBE values also indicate a modest reduction in the positive bias of SP at longer horizons. Similar to UTAR, the nRMSE and $NICE_\Sigma$ results also showed an increasing advantage for LSP at longer forecast horizons.

At BKH, LSP produced marginally lower RMSE, MAE, and nRMSE values and slightly higher R^2 values across the evaluated horizons, whereas $NICE_\Sigma$ consistently favored SP. The absolute MBE was slightly lower for SP at the 10-minute horizon but lower for LSP from 20 to 50 minutes, with the bias shifting from slight overprediction under SP to slight underprediction under LSP. Therefore, BKH indicates a small and metric-dependent advantage rather than a uniform improvement.

Tampere University results show a broadly similar overall trend to the tropical sites, in that the advantage of LSP becomes more visible at longer forecast horizons. However, in Tampere the main benefit is more clearly reflected in the preservation of positive R^2 values and more balanced long-horizon behavior under high-latitude conditions. The RMSE difference between SP and LSP remains small at short lead times, but the relative benefit of LSP becomes increasingly visible as the forecast horizon increases. A similar tendency is observed in MAE, which is slightly higher for LSP at the shortest horizon but becomes lower than SP beyond 30 minutes, with the clearest improvement at 40- and 50-minute horizons.

The contrast was more pronounced in R^2. At the 5-minute temporal resolution, SP produced negative R^2 values at the 40- and 50-minute horizons, whereas LSP maintained positive values. This result indicates that the local climatological GHI profile provided a more representative description of long-horizon irradiance evolution under conditions characterized by strong seasonal variation and low solar elevation. The complementary nRMSE and $NICE_\Sigma$ results reflected the same horizon-dependent interpretation: SP retained a marginal advantage at the 10-minute horizon, whereas LSP performed substantially better at the 30- and 50-minute horizons. The MBE results also indicated more balanced long-horizon forecast bias for LSP.

A known limitation of classical SP stems from its reliance on theoretical clear-sky irradiance in computing the clear-sky index, particularly near sunrise and sunset. Under low solar elevation angles, mismatches between clear-sky estimates and measured GHI in clear conditions can lead to highly sensitive CSI values and spurious forecast spikes. As illustrated in Figure 4 for representative days at Tampere University and EVN, LSP more closely follows the observed irradiance evolution, whereas SP exhibits pronounced early-morning and late-evening forecast spikes. This behavior arises from the use of empirical local climatological GHI profiles constrained by historical observations. In the present study, however, R^2 was evaluated over much longer periods, so this local low-irradiance effect is not expected to dominate the overall R^2 values. Therefore, the stronger performance of LSP in such periods is more appropriately interpreted through its more physically representative behavior and lower absolute forecast error, rather than through R^2 alone.

However, it should be noted that LSP does not entirely eliminate the general limitations of ratio-based persistence models under extremely low irradiance conditions. During sunrise and sunset, particularly during the winter months at high latitudes, the solar signal may become insufficient to maintain physically meaningful relative scaling ratios. Therefore, periods of very low irradiance should either be excluded, treated separately, or interpreted with caution in operational evaluations.

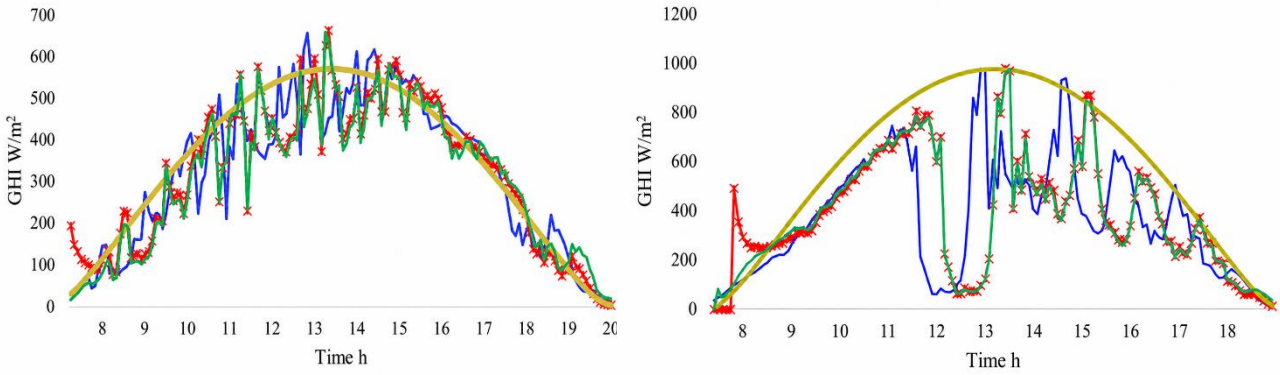


Figure 4. Examples from Tampere (4 April 2021, left) and EVN (8 September 2019, right) at a 30-minute forecast horizon, illustrating the near-sunrise bias of SP caused by clear-sky normalization under low-solar-elevation conditions. The blue solid line represents the observed GHI, the red line with cross markers represents the SP forecast, the green solid line represents the LSP forecast, and the thick yellow line represents the modeled clear-sky GHI.

The NREL results represent a comparatively more neutral case in which SP already performs strongly and the benefit of LSP remains limited. Across the tested temporal resolutions and forecast horizons, the RMSE, MAE and nRMSE values of the two methods remain close, with SP generally retaining a slight advantage. Meanwhile, the $NICE_{\Sigma}$ values also favored SP. Although some seasonal LSP configurations produce marginally higher R^2 values, the differences remain small and do not indicate a clear practical improvement. The MBE results show that LSP shifts the forecast bias from a small positive value in SP toward slight underprediction, but this bias rebalancing does not translate into lower absolute forecast error. Overall, the NREL results suggest that under predominantly clear-sky conditions with limited short-term variability, SP already provides a highly effective baseline and the introduction of a local climatological scaling offers little additional forecasting benefit.

New Mexico represents the clearest negative case for LSP. Across all tested forecast horizons, LSP produces higher RMSE, MAE, nRMSE, and $NICE_{\Sigma}$ values than SP, while the corresponding R^2 values are also consistently lower. The MBE results further indicate that LSP introduces a substantially larger positive forecast bias, suggesting a stronger tendency toward overprediction under these relatively steady arid conditions. In this case, the climatological scaling appears to deviate from the local short-term irradiance evolution rather than improve it.

SRML also did not provide a consistent improvement case for LSP. SP performed better at the 10- and 30-minute horizons, whereas the 50-minute results were metric-dependent: SP achieved a lower nRMSE, while LSP obtained a slightly lower $NICE_{\Sigma}$. Across the wider set of absolute-error metrics, LSP generally remained less accurate and shifted the forecast bias from slight underprediction under SP toward positive bias. These results indicate that when irradiance persistence already dominates short-term evolution, localized climatological scaling may provide limited additional information and can reduce forecast accuracy.

At the common 5-minute temporal resolution, the nRMSE results were consistent with the principal RMSE trends because both models were normalized using the same mean observed daytime GHI at each site. $NICE_{\Sigma}$ broadly reflected the same site- and horizon-dependent patterns while providing complementary information relative to simple persistence. LSP was favored at UTAR and EVN and at the longer horizons in Tampere, whereas SP remained more competitive at NREL and New Mexico. BKH remained metric-dependent, and SRML showed mixed behavior at the longest horizon. Thus, the normalized and persistence-

referenced metrics do not contradict the principal RMSE and MAE findings but help clarify the magnitude and practical relevance of the observed differences.

Overall, the cross-site results demonstrate that the effectiveness of LSP depends strongly on local climatic characteristics and irradiance dynamics. LSP performs most effectively when recurring diurnal cloud evolution or other local effects cause systematic deviations from theoretical clear-sky scaling. Under such conditions, the empirical climatological GHI profile provides a more representative reference for intra-hour irradiance evolution. In contrast, at predominantly clear-sky or strongly persistence-dominated sites, climatological scaling provides limited additional information and may reduce forecast accuracy. Therefore, LSP should be interpreted as a site-dependent extension of SP rather than a universally superior replacement.

3.2. Sensitivity Analysis of Multiplicative and Additive LSP Formulations

To evaluate whether the multiplicative structure of LSP should be retained, a sensitivity analysis was performed by comparing the multiplicative and additive formulations at UTAR. UTAR was selected as the principal site because it provided the complete reference and test datasets required for the analysis. The common 5-minute temporal resolution was adopted, and all five forecast horizons from 10 to 50 minutes were evaluated.

Table 3. Sensitivity comparison between the multiplicative and additive LSP formulations at UTAR using data at a 5-minute temporal resolution over the 10–50-minute forecast horizons.

Temporal forecast horizon (min)	Model	RMSE (W/m^2)	MAE (W/m^2)	nRMSE (%)	$NICE_\Sigma$
10	Additive LSP	200.6	119	45.07	0.994
	Multiplicative LSP	**198.6**	**117.1**	**44.59**	**0.979**
20	Additive LSP	229.5	148.1	51.55	0.979
	Multiplicative LSP	**224.6**	**143.4**	**50.42**	**0.953**
30	Additive LSP	247.2	166.8	55.19	0.963
	Multiplicative LSP	**239.2**	**159.4**	**53.36**	**0.925**
40	Additive LSP	257.7	179.4	57.36	0.944
	Multiplicative LSP	**246.9**	**169.5**	**54.87**	**0.897**
50	Additive LSP	268.4	191.3	59.36	0.924
	Multiplicative LSP	**254.2**	**178.9**	**56.26**	**0.870**

Lower RMSE, MAE, nRMSE, and $NICE_\Sigma$ values indicate better performance. Bold values indicate the better-performing formulation for each metric and forecast horizon.

The multiplicative formulation achieved lower RMSE, MAE, nRMSE, and $NICE_\Sigma$ values at every evaluated forecast horizon. The performance difference generally increased with forecast horizon, suggesting that proportional climatological scaling becomes increasingly beneficial for intermediate and longer intra-hour forecasts in the UTAR dataset.

The additive formulation avoids division by potentially small climatological irradiance values and may therefore provide numerical advantages under specific low-irradiance conditions. However, it assumes that the absolute climatological change can be transferred directly to the current irradiance state regardless of the prevailing irradiance magnitude or cloud attenuation. The results indicate that this assumption was less

representative than proportional scaling for the UTAR data. Therefore, the multiplicative formulation was retained as the proposed LSP formulation.

3.3. Benchmark Comparison with Simplified BLEND

To evaluate the added value of LSP relative to another persistence-based benchmark, a separate comparison with Simplified BLEND was conducted for UTAR dataset. The comparison used the common 5-minute temporal resolution and all forecast horizons from 10 to 50 minutes. UTAR was selected because it provided the complete historical reference data and phase-aligned previous-day observations required to implement both methods.

Because Simplified BLEND requires an irradiance observation from the corresponding phase of the previous day, the analysis was restricted to identical forecast–observation samples available to both LSP and Simplified BLEND. The performance metrics were calculated separately for each valid test day and subsequently averaged across the evaluation period.

Table 4. Comparison of LSP and Simplified BLEND at UTAR using data at a 5-minute temporal resolution over the 10–50-minute forecast horizons.

Temporal forecast horizon (min)	Model	RMSE (W/m^2)	MAE (W/m^2)	nRMSE (%)	$NICE_\Sigma$
10	LSP	198.7	**117.2**	44.61	**0.98**
	Simplified BLEND	**197.2**	124.7	**44.12**	1.01
20	LSP	**224.6**	**143.5**	**49.53**	**0.95**
	Simplified BLEND	224.7	152	49.55	0.97
30	LSP	**239.3**	**159.4**	**53.40**	**0.92**
	Simplified BLEND	240.1	170.2	53.82	0.94
40	LSP	**247**	**169.6**	**55.46**	**0.89**
	Simplified BLEND	249.4	181.8	56.13	0.92
50	LSP	**254.2**	**178.9**	**56.28**	**0.87**
	Simplified BLEND	258.3	192.7	57.44	0.90

Lower RMSE, MAE, nRMSE, and $NICE_\Sigma$ values indicate better performance. Bold values indicate the better-performing model for each metric and forecast horizon. The reported values are arithmetic means of daily metrics calculated using identical common valid samples for LSP and Simplified BLEND. Because Simplified BLEND requires the phase-aligned previous-day observation, the common-sample restriction causes small differences between the LSP values in this table and those reported in Tables 3 and A1.

At the 10-minute horizon, Simplified BLEND achieved marginally lower RMSE and nRMSE values, whereas LSP produced lower MAE and $NICE_\Sigma$. From the 20- to 50-minute horizons, LSP consistently achieved lower values for all four reported metrics. The difference became progressively more visible at the longer horizons, particularly for MAE.

These results indicate that the phase-aligned previous-day observation used by Simplified BLEND remains useful at very short horizons, but the localized climatological scaling of LSP provides a more stable reference as the forecast horizon increases. The comparison therefore suggests that LSP can provide a competitive persistence-based benchmark for intermediate and longer intra-hour forecasting at UTAR, while also showing that no single persistence formulation is necessarily superior at every horizon and according to every metric.

3.4. Sensitivity to the Climatological Averaging Strategy

The sensitivity of LSP to the climatological averaging strategy was evaluated through two comparisons. First, a single annual climatological GHI profile was compared with season-specific profiles. Second, one-year and two-year historical averaging periods were compared. The detailed results are reported in Tables A1 and A2. This analysis focused primarily on NREL, SRML, and Tampere, where seasonal irradiance variation was expected to have a stronger influence on the representativeness of the climatological profile.

For NREL and SRML, season-specific profiles did not provide substantial improvements relative to a single annual profile. At NREL, for example, the RMSE at the 20-minute forecast horizon decreased only from 139.5 W/m² to 138.9 W/m² when seasonal profiles were applied, while R^2 increased marginally from 0.48 to 0.49. At the 50-minute horizon, RMSE decreased from 170.7 W/m² to 167.4 W/m², and R^2 increased from 0.25 to 0.26. The corresponding MAE and MBE values in Table A2 supported the same conclusion. These relatively small differences indicate that a single annual climatological profile is generally sufficient for representing irradiance behavior at NREL and SRML.

The situation was different at Tampere because of its high-latitude location and strong seasonal variation in daylight duration and solar elevation. Under these conditions, a single annual profile was less capable of consistently representing irradiance evolution throughout the year. Therefore, season-specific climatological profiles were adopted at Tampere, while winter periods were excluded because of their extremely low irradiance levels and limited practical value for short-term solar forecasting.

For the retained seasons, seasonal averaging produced more representative long-horizon behavior. At the 5-minute temporal resolution, SP produced R^2 values of -0.20 and -0.82 at the 40- and 50-minute horizons, respectively, whereas LSP maintained positive R^2 values of 0.39 and 0.33. The corresponding MAE and MBE values also indicated more favorable long-horizon error characteristics for LSP under seasonal averaging.

The effect of the historical averaging period was examined by comparing one-year and two-year climatological profiles. Extending the historical period to two years did not produce a consistent improvement in RMSE, MAE, or R^2 across the evaluated sites. In most cases, approximately one year of historical irradiance measurements was sufficient to construct a representative climatological profile. This finding represents a practical advantage because LSP can be implemented at sites where only a limited period of historical irradiance measurements is available. Overall, seasonal averaging appears most relevant in regions with pronounced seasonal irradiance variation, whereas a single annual profile based on approximately one year of data remains adequate for most other climatic conditions.

3.5. Practical Implications and Baseline Selection

The forecast errors observed across the seven sites should be interpreted together with the intrinsic variability and forecastability of the local irradiance series. Under highly stable and highly predictable conditions, particularly at very short forecast horizons, simple persistence remains an appropriate baseline because the most recent observation often contains most of the available predictive information.

SP remains a strong baseline when deterministic solar-geometry evolution is adequately represented by the theoretical clear-sky profile and atmospheric conditions remain relatively stable. Its low computational requirements and absence of a historical training requirement make it particularly useful for new sites or locations with reliable clear-sky model inputs.

LSP is most relevant when measured irradiance exhibits recurring site-specific diurnal patterns that systematically differs from theoretical clear-sky scaling. These differences may arise from recurring cloud

attenuation, afternoon convection, aerosols, horizon obstruction, terrain effects, or other localized factors. Under such conditions, LSP provides a locally representative scaling reference while retaining the current observed irradiance state. Its benefits become more apparent at intermediate and longer intra-hour forecast horizons, where direct persistence assumptions become less reliable.

Nevertheless, all persistence-based approaches remain inherently limited under rapidly evolving and weakly predictable cloud conditions. When irradiance changes are dominated by transient cloud motion rather than recurring local patterns, forecasting systems incorporating sky imagery, cloud-motion vectors, satellite observations, numerical weather prediction, or machine-learning models may be more appropriate.

Therefore, the selection of a forecasting baseline should depend on the local irradiance regime, forecast horizon, availability of historical measurements, acceptable computational complexity, and operational objective. Simple persistence, SP, and LSP should be viewed as complementary benchmarks rather than as universally ordered alternatives.

4. Conclusion

This study set out to address two key limitations of conventional SP forecasting. First, SP may produce unreliable behavior near sunrise and sunset because clear-sky normalization becomes highly sensitive under low-solar-elevation conditions. Second, the clear-sky scaling ratios used in SP may differ systematically from the actual local irradiance evolution at a given site. To address these limitations, the proposed Localized Smart Persistence (LSP) replaces theoretical clear-sky scaling with empirical site-specific climatological GHI profiles derived from historical measurements.

The proposed LSP model was evaluated using temporal resolutions ranging from 30 seconds to 5 minutes and forecast horizons between 10 and 50 minutes across seven sites representing tropical, arid, temperate, predominantly clear-sky, and high-latitude climatic conditions. Forecast performance was assessed using RMSE, MAE, R^2, and MBE, while mean daily nRMSE and $NICE_\Sigma$ were additionally reported at the common 5-minute temporal resolution for representative 10-, 30-, and 50-minute forecast horizons.

The results show that the effect of LSP is strongly site-dependent. Under conditions characterized by recurring diurnal irradiance evolution that differs systematically from theoretical clear-sky behavior, particularly at tropical sites influenced by persistent afternoon convection, LSP generally reduced RMSE and MAE relative to SP, especially at longer forecast horizons. For example, at the UTAR site using 1-minute data, LSP reduced RMSE from 243.6 W/m² to 238.6 W/m² at a 30-minute forecast horizon.

At the high-latitude Tampere site, SP produced negative R^2 values at the 40- and 50-minute forecast horizons, whereas LSP maintained positive R^2 values under the same low solar elevation conditions. In contrast, under predominantly steady clear-sky or highly persistence-dominated conditions, climatological scaling provided little or no practical forecasting advantage.

The sensitivity analysis showed that the multiplicative LSP formulation outperformed the additive formulation across all evaluated forecast horizons at UTAR. In the comparison with Simplified BLEND, BLEND achieved marginally lower RMSE and nRMSE at the 10-minute horizon, whereas LSP provided better overall performance from 20 to 50 minutes.

Because the LSP reference profile is derived directly from historical measurements, the method can implicitly capture site-specific effects such as horizon obstruction, terrain masking, nearby structures, recurring shading patterns, aerosols, and local cloud attenuation that are not represented in theoretical clear-sky normalization. In the present study, approximately one year of historical irradiance data was generally sufficient to construct a representative climatological GHI profile, while sites with weaker seasonal variability may require shorter averaging periods. This makes LSP relatively simple to implement at

locations where local measurements are available and where clear-sky normalization is less representative of actual irradiance behavior.

The empirical structure of LSP also allows the framework to be extended beyond horizontal irradiance forecasting to variables such as plane-of-array irradiance, PV power production, or AC power output using corresponding historical measurements. Such applications may be particularly useful in practical PV systems where conventional clear-sky-based SP requires additional sky transposition modelling, which may introduce additional uncertainty. However, these extensions still depend on the availability of representative historical datasets and should account for operational factors, such as system degradation, curtailment, and configuration changes.

Overall, LSP should be interpreted as a site-dependent persistence-based baseline extension rather than a universally superior forecasting model. Its principal value lies in providing a locally representative forecasting reference under climatic conditions where clear-sky normalization is less physically representative of actual irradiance evolution. Future benchmarking studies should thus consider the suitability of SP, LSP, or alternative persistence formulations according to the prevailing climatic conditions, irradiance characteristics, and forecasting timescales of the target application.

Acknowledgment

The research project was supported by the Ministry of Higher Education Malaysia (MoHE) through the Fundamental Research Grant Scheme (FRGS/1/2022/TK08/UTAR/02/8) and by the Ministry of Science, Technology and Innovation (MOSTI) through the Malaysia–China Intergovernmental Science, Technology and Innovation Cooperative Programme 2022, with additional funding provided by the National Key R&D Program of China (No. 2022YFE0196100). The authors gratefully acknowledge this financial support.

Appendix

Table A1. RMSE (W/m²) and R^2 values for SP and LSP across the sites and temporal forecast horizons.

Site	Temporal resolution	Model	Temporal forecast horizon (minutes ahead)									
			10		20		30		40		50	
			RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
UTAR	30 s	SP	198.8	0.47	228.3	0.31	243.5	0.21	254.9	0.13	265	0.05
		LSP	197.6	0.49	225.3	0.33	238.5	0.24	247.3	0.17	255.2	0.12
	1 min	SP	198.8	0.47	228.3	0.31	243.6	0.21	254.9	0.13	265	0.05
		LSP	197.6	0.48	225.2	0.33	238.6	0.24	247.3	0.18	255.2	0.12
	5 min	SP	199.8	0.47	227.9	0.31	244.6	0.21	254.6	0.14	264.9	0.07
		LSP	198.6	0.49	224.6	0.34	239.2	0.24	246.9	0.19	254.2	0.13
EVN	5 min	SP	112.1	0.84	138.8	0.74	161.3	0.64	178.3	0.55	196.9	0.44
		LSP	112.5	0.83	140.6	0.73	156.8	0.66	168.5	0.6	180.6	0.52
BKH	5 min	SP	185	0.51	203.2	0.41	212.5	0.35	218.2	0.31	224.9	0.27
		LSP	183.6	0.53	201.9	0.43	211.1	0.38	215.6	0.35	219.6	0.33
Tampere	30 s	SP	83.8	0.65	95.4	0.47	103.3	0.22	111.7	-0.2	120.4	-0. 82
		LSP	84.8	0.65	95.3	0.53	100.9	0.46	105.8	0.38	109.7	0.32

	1 min	SP	83.9	0.65	95.5	0.47	103.3	0.22	111.7	-0.2	120.5	-0.82
		LSP	84.9	0.65	95.3	0.54	100.9	0.47	105.9	0.38	109.7	0.32
	5 min	SP	84.1	0.66	95.8	0.46	103.7	0.22	111.7	-0.2	120.2	-0.82
		LSP	85	0.65	95.5	0.54	101	0.47	105.8	0.39	109.5	0.33
NREL	1 min	SP	117.4	0.62	137.3	0.48	149.1	0.39	157.9	0.32	165.3	0.25
		LSP -1 year mean profile	118.8	0.62	139.5	0.48	152.4	0.39	162.3	0.31	170.7	0.25
		LSP- 2 year mean profile	118.5	0.62	139.2	0.49	152.1	0.4	162	0.32	170.5	0.25
		LSP-separate seasons	118.9	0.63	138.9	0.49	151	0.4	159.9	0.33	167.4	0.26
	5 min	SP	117.1	0.62	136.7	0.49	148.6	0.39	157.4	0.32	165.3	0.25
		LSP-1 year mean profile	118.6	0.62	139.1	0.49	152	0.4	161.9	0.32	170.8	0.25
		LSP-2 year mean profile	118.2	0.63	138.7	0.5	151.7	0.4	161.6	0.32	170.6	0.25
		LSP-separate seasons	118.6	0.63	138.4	0.5	150.6	0.4	159.5	0.33	167.5	0.26
SRML	1 min	SP	87.4	0.55	102.5	0.36	110.5	0.24	115.4	0.15	121.3	0.1
		LSP-1 year mean profile	89.6	0.54	106	0.34	115.8	0.2	122.5	0.1	129.2	0.06
		LSP-2 year mean profile	89.3	0.54	105.8	0.34	115.7	0.2	122.5	0.1	129	0.06
		LSP-separate seasons	90.3	0.54	106.1	0.34	115	0.22	120.9	0.12	127.6	0.08
	5 min	SP	86.4	0.56	101.8	0.37	109.9	0.25	114.8	0.16	120.1	0.11
		LSP-1 year mean profile	88.7	0.55	105.4	0.34	115.4	0.21	122	0.12	128.4	0.07
		LSP-2 year mean profile	88.3	0.55	105.2	0.34	115.2	0.22	122	0.12	128.4	0.06
		LSP-separate seasons	89.5	0.55	105.4	0.35	114.6	0.23	120.3	0.14	127.1	0.1
New Mexico	30 s	SP	66.6	0.6	76	0.47	81.5	0.39	86.4	0.3	90.4	0.24
		LSP	69.4	0.59	81.5	0.44	90.2	0.33	98	0.22	105.5	0.12
	1 min	SP	66.8	0.6	76.3	0.47	81.8	0.39	86.6	0.3	90.6	0.24
		LSP	69.5	0.59	81.8	0.44	90.4	0.33	98.2	0.22	105.8	0.13

	5 min	SP	66.3	0.59	74.7	0.47	80.4	0.4	85.8	0.31	90	0.24
		LSP	68.8	0.58	80.3	0.45	89	0.35	97.3	0.23	104.9	0.13

To facilitate interpretation of Table A1, it should be noted that the climatological averaging periods differ across sites (ranging from seasonal to two-year periods). For clarity, readers are advised to cross-reference these tabulated values with Figure 2, which provides a visual comparison of forecasting performance across different horizons and temporal resolutions.

Table A2. MAE (W/m²) and MBE (W/m²) values for SP and LSP across the sites and temporal forecast horizons.

Site	**Temporal resolution**	**Model**	**Temporal forecast horizon (minutes ahead)**									
			10		**20**		**30**		**40**		**50**	
			MBE	**MAE**	**MBE**	**MAE**	**MBE**	**MAE**	**MBE**	**MAE**	**MBE**	**MAE**
UTAR	30 s	SP	4.1	116.4	8.1	144.7	12.1	160.2	18.9	172.7	23.9	183.5
		LSP	-2.6	116.6	-6.5	144.2	-9.8	159	-13.1	169.9	-16.2	179.7
	1 min	SP	4	116.3	8.1	144.6	12.1	160.3	18.9	172.7	23.9	183.5
		LSP	-2.5	116.6	-5.2	144.1	-8	159	-13.1	169.9	-16.3	179.6
	5 min	SP	4.8	116.9	9.7	144.2	14.7	161.2	19.8	172.3	25	182.7
		LSP	-3.2	117.1	-6.5	143.4	-9.5	159.4	-12.5	169.5	-15.4	178.9
EVN	5 min	SP	3.2	61.8	6.9	82.4	11.6	100.8	16.3	113.2	20.9	125.5
		LSP	3.2	61.3	7.2	83.2	9.9	97.5	13.3	108.1	16.7	118.5
BKH	5 min	SP	1	103.1	3	122.3	3.9	136	5.5	143.2	7.4	150.4
		LSP	-1.4	102.3	-2.8	121.7	-3.5	134.4	-4.9	142.1	-5.6	146.6
Tampere	30 s	SP	0.3	45	1.3	55.2	2.7	62	4.4	68.1	6.4	73.7
		LSP	0.5	46.3	0.8	55.8	1.2	61.6	1.6	66.4	2	70.5
	1 min	SP	0.3	45	1.3	55.3	2.7	62	4.4	68.2	6.4	73.7
		LSP	0.5	46.3	0.8	55.8	1.2	61.6	1.6	66.5	2	70.6
	5 min	SP	0.3	45.3	1.3	55.6	2.7	62.4	4.4	68.3	6.4	73.8
		LSP	0.5	46.3	0.8	55.8	1.2	61.6	1.6	66.4	2	70.6
NREL	1 min	SP	1.9	66.2	4	83.4	6.2	94.5	8.4	102.9	10.7	110.1
		LSP-1 year mean profile	-0.4	69.5	-0.9	88.5	-1.3	101.7	-1.8	112.3	-2.6	121.5
		LSP-2 year mean profile	-0.3	68.9	-0.8	88.1	-1.1	101.5	-1.6	112.2	-2.1	121.4
		LSP-separate seasons	-0.2	69.4	-0.4	87.3	-0.6	99.4	-1	108.7	-1.4	116.7
	5 min	SP	1.9	66.1	4	83.1	6.2	94.3	8.4	102.6	10.7	109.9
		LSP-1 year mean profile	-0.4	69.3	-0.8	88.2	-1.3	101.5	-1.8	112.1	-2.5	121.3
		LSP-2 year	-0.4	68.8	-0.7	87.9	-1.1	101.3	-1.6	112	-2.1	121.3

		mean profile										
		LSP-separate seasons	-0.2	69.2	-0.4	87	-0.6	99.2	-1	108.5	-1.4	116.5
SRML	1 min	SP	-0.5	55.3	-1.2	68.4	-1.7	76.3	-2.3	81.1	-2.9	86.3
		LSP-1 year mean profile	0.6	58.1	1.2	73	1.9	82.6	2.7	89.5	3.4	92.7
		LSP-2 year mean profile	0.6	57.7	1.1	72.8	1.8	82.5	2.6	89.5	3.4	97.6
		LSP-separate seasons	0.9	58.7	1.6	72.7	2.5	81.4	3.6	87.2	4.1	95.2
	5 min	SP	-0.5	54.7	-1.2	68.1	-1.8	76	-2.3	89.6	-2.9	93
		LSP-1 year mean profile	0.6	57.6	1.2	72.5	1.9	82.4	2.7	89	3.4	97.6
		LSP-2 year mean profile	0.6	57.1	1.8	72.3	1.8	82.2	2.6	89	3.3	97.6
		LSP-separate seasons	0.9	58.1	1.6	72.1	2.5	81	3.6	86.7	4.3	95.2
New Mexico	30 s	SP	0.8	39.6	1.6	48.3	2.3	53.9	2.9	58.7	3.9	62.6
		LSP	4.2	42.9	8.2	54.3	12.1	62.7	15.7	70.3	18.6	77.9
	1 min	SP	0.8	39.6	1.6	48.5	2.3	53.9	2.9	58.7	3.9	62.6
		LSP	4.2	42.9	8.2	54.5	12.1	62.8	15.7	70.3	18.5	77.8
	5 min	SP	0.7	39.7	1.5	47.8	2.1	53.4	2.8	58.3	3.6	62.2
		LSP	4.1	42.8	8.1	53.6	11.8	62.1	15.4	69.8	18.1	77.2

Table A3. Mean daily nRMSE and $NICE_{\Sigma}$ values for SP and LSP at the common 5-minute temporal resolution for the 10-, 30-, and 50-minute forecast horizons. Lower values indicate better performance. A $NICE_{\Sigma}$ value below 1 indicates better performance than simple persistence. Bold values indicate the lower value for each site and forecast horizon.

Site	Temporal forecast horizon (min)	nRMSE (%)		$NICE_{\Sigma}$	
		SP	LSP	SP	LSP
UTAR	10	44.84	**44.59**	0.981	**0.979**
	30	54.52	**53.36**	0.939	**0.925**
	50	58.42	**56.26**	0.896	**0.870**
EVN	10	29.29	**29.15**	0.943	**0.939**
	30	41.74	**40.47**	0.892	**0.857**
	50	49.90	**45.79**	0.879	**0.790**
BKH	10	40.43	**40.39**	**0.949**	0.956
	30	45.96	**45.91**	**0.876**	0.884
	50	47.62	**47.36**	**0.812**	0.818

Tampere	10	**39.09**	39.11	**0.978**	0.996
	30	51.99	**47.51**	1.039	**0.916**
	50	65.10	**51.18**	1.176	**0.865**
NREL	10	**24.71**	24.99	**0.823**	0.909
	30	**31.95**	32.67	**0.741**	0.823
	50	**35.77**	36.93	**0.692**	0.776
New Mexico	10	**22.01**	22.61	**0.784**	0.897
	30	**26.61**	28.58	**0.676**	0.818
	50	**29.62**	32.86	**0.624**	0.777
SRML	10	**28.50**	29.01	**0.825**	0.934
	30	**37.66**	39.10	**0.756**	0.851
	50	**42.38**	43.68	0.843	**0.821**

References

[1] Voyant C, Notton G. Solar irradiation nowcasting by stochastic persistence: A new parsimonious, simple and efficient forecasting tool. *Renewable and Sustainable Energy Reviews* 2018;92:343–52. https://doi.org/10.1016/j.rser.2018.04.116.

[2] Sengupta M, Habte A, Wilbert S, Gueymard C, Remund J. *Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications: Third Edition*. Golden, CO: National Renewable Energy Laboratory; 2021. NREL/TP-5D00-77635.

[3] Ela E, Milligan M, Kirby B. *Operating Reserves and Variable Generation: A comprehensive review of current strategies, studies, and fundamental research on the impact that increased penetration of variable renewable generation has on power system operating reserves*. Golden, CO: National Renewable Energy Laboratory; 2011. NREL/TP-5500-51978.

[4] Nouri B, Lezaca J, Fabel Y, Hammer A, Blum N, Wilbert S. Enhancing intra-hour solar irradiance forecasting for solar applications: A blended model of satellite, sky imager, and persistence. *Solar RRL*. 2025;9:e202500486. https://doi.org/10.1002/solr.202500486.

[5] Yang D, Kleissl J, Gueymard CA, Pedro HTC, Coimbra CFM. History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining. *Solar Energy* 2018;168:60–101. https://doi.org/10.1016/j.solener.2017.11.023.

[6] Antonanzas J, Osorio N, Escobar R, Urraca R, Martinez-de-Pison FJ, Antonanzas-Torres F. Review of photovoltaic power forecasting. *Solar Energy* 2016;136:78–111. https://doi.org/10.1016/j.solener.2016.06.069.

[7] Lauret P, Voyant C, Soubdhan T, David M, Poggi P. A benchmarking of machine learning techniques for solar radiation forecasting in an insular context. *Solar Energy* 2015;112:446–57. https://doi.org/10.1016/j.solener.2014.12.014.

[8] Voyant C, Notton G, Kalogirou S, Nivet ML, Paoli C, Motte F, et al. Machine learning methods for solar radiation forecasting: A review. *Renewable Energy* 2017;105:569–82. https://doi.org/10.1016/j.renene.2016.12.095.

[9] Urraca R, Antonanzas J, Alia-Martinez M, Martinez-De-Pison FJ, Antonanzas-Torres F. Smart baseline models for solar irradiation forecasting. *Energy Convers Manag* 2016;108:539–48. https://doi.org/10.1016/j.enconman.2015.11.033.

[10] Kleissl J. *Solar Energy Forecasting and Resource Assessment*. Academic Press; 2013. https://doi.org/10.1016/C2011-0-07022-9.

[11] Cui Y, Wang P, Meirink JF, Ntantis N, Wijnands JS. Solar radiation nowcasting based on geostationary satellite images and deep learning models. *Solar Energy*. 2024;282:112866. https://doi.org/10.1016/j.solener.2024.112866.

[12] Ruan G, Chen X, Lim EG, Fang L, Su Q, Jiang L, et al. On the use of sky images for intra-hour solar forecasting benchmarking: Comparison of indirect and direct approaches. *Solar Energy* 2024;276:112649. https://doi.org/10.1016/j.solener.2024.112649.

[13] Martín-Pomares L, Martínez D, Polo J, Perez-Astudillo D, Bachour D, Sanfilippo A. Analysis of the long-term solar potential for electricity generation in Qatar. *Renewable and Sustainable Energy Reviews* 2017;73:1231–1246. https://doi.org/10.1016/j.rser.2017.01.125.

[14] Wang Y, Wu L. On practical challenges of decomposition-based hybrid forecasting algorithms for wind speed and solar irradiation. *Energy*. 2016;112:208–220. https://doi.org/10.1016/j.energy.2016.06.075.

[15] Rahimi N, Park S, Choi W, Oh B, Kim S, Cho Y-H, et al. A Comprehensive Review on Ensemble Solar Power Forecasting Algorithms. *Journal of Electrical Engineering and Technology* 2023;18:719–33. https://doi.org/10.1007/s42835-023-01378-2.

[16] Yang D. Making reference solar forecasts with climatology, persistence, and their optimal convex combination. *Solar Energy* 2019;193:981–985. https://doi.org/10.1016/j.solener.2019.10.006.

[17] Liu W, Liu Y, Zhou X, Xie Y, Han Y, Yoo S, et al. Use of physics to improve solar forecast: Physics-informed persistence models for simultaneously forecasting GHI, DNI, and DHI. *Solar Energy* 2021;215:252–65. https://doi.org/10.1016/j.solener.2020.12.045.

[18] Yelchuri S, Rangaraj AG, Xie Y, Habte A, Joshi MC, Boopathi K, Sengupta M, Balaraman K. *A short-term solar forecasting platform using a physics-based smart persistence model and data imputation method*. Golden, CO: National Renewable Energy Laboratory; 2021. NREL/TP-5D00-81421.

[19] Kiss VM, Hetesi Z, Kiss T. The effect of time resolution on energy system simulation in case of intermittent energies. *Renewable and Sustainable Energy Reviews* 2024;191:114099. https://doi.org/10.1016/j.rser.2023.114099.

[20] Omoyele O, Hoffmann M, Koivisto M, Larrañeta M, Weinand JM, Linßen J, et al. Increasing the resolution of solar and wind time series for energy system modeling: A review. *Renewable and Sustainable Energy Reviews*. 2024;189:113792. https://doi.org/10.1016/j.rser.2023.113792.

[21] Eroshenko SA, Khalyasmaa AI, Snegirev DA, Dubailova VV, Romanov AM, Butusov DN. The impact of data filtration on the accuracy of multiple time-domain forecasting for photovoltaic power plants generation. *Applied Sciences*. 2020;10(22):8265. https://doi.org/10.3390/app10228265.

[22] Xie Y, Sengupta M. A Fast All-sky Radiation Model for Solar applications with Narrowband Irradiances on Tilted surfaces (FARMS-NIT): Part I. The clear-sky model. *Solar Energy* 2018;174:691–702. https://doi.org/10.1016/j.solener.2018.09.056.

[23] Terrén-Serrano G, Martínez-Ramón M. Multi-layer wind velocity field visualization in infrared images of clouds for solar irradiance forecasting. *Applied Energy* 2021;288:116656. https://doi.org/10.1016/j.apenergy.2021.116656.

[24] Vignola F, Andreas A. *University of Oregon: Solar Radiation Monitoring Lab (Data)*. 2013. https://doi.org/10.7799/1183467.

[25] Stoffel T, Andreas A. *NREL Solar Radiation Research Laboratory (SRRL): Baseline Measurement System (BMS), Golden, Colorado*. NLR Data Catalog. Golden, CO: National Laboratory of the Rockies; 2015. https://doi.org/10.7799/1052221.

[26] Torres Lobera D, Mäki A, Huusari J, Lappalainen K, Suntio T, Valkealahti S. Operation of TUT solar PV power station research plant under partial shading caused by snow and buildings. *International Journal of Photoenergy*. 2013;2013:837310. https://doi.org/10.1155/2013/837310.

[27] Ineichen P. A broadband simplified version of the Solis clear sky model. *Solar Energy* 2008;82:758–62. https://doi.org/10.1016/j.solener.2008.02.009.

[28] Gueymard CA. REST2: High-performance solar radiation model for cloudless-sky irradiance, illuminance, and photosynthetically active radiation - Validation with a benchmark dataset. *Solar Energy*. 2008;82(3):272–285. https://doi.org/10.1016/j.solener.2007.04.008.

[29] Zhang J, Florita A, Hodge BM, Lu S, Hamann HF, Banunarayanan V, et al. A suite of metrics for assessing the performance of solar power forecasting. *Solar Energy* 2015;111:157–75. https://doi.org/10.1016/j.solener.2014.10.016.

[30] Yang D, Bright JM. Worldwide validation of 8 satellite-derived and reanalysis solar radiation products: A preliminary evaluation and overall metrics for hourly data over 27 years. *Solar Energy*. 2020;210:3–19. https://doi.org/10.1016/j.solener.2020.04.016.

[31] Voyant C, Banes C, Garcia-Gutierrez L, Notton G, Despotovic M, Yaseen ZM. Symmetry-constrained forecasting of periodically correlated energy processes. Applied Mathematical Modelling 2026;157:116988. doi:10.1016/j.apm.2026.116988

[32] Voyant C, Despotovic M, Garcia-Gutierrez L, Amaro e Silva R, Lauret P, Soubdhan T, Bailek N. NICE^{k} metrics: Unified and multidimensional framework for evaluating deterministic solar forecasting accuracy. Sustainable Energy Technologies and Assessments 2025;83:104588. https://doi.org/10.1016/j.seta.2025.104588.