Solar Irradiance Estimation from SEVIRI: A Feature Importance and Channel Selection Analysis

Solar Irradiance Estimation from SEVIRI: A Feature Importance and Channel Selection Analysis

1st Raimondo Gallo
Politecnico di Torino
Torino, Italia
raimondo.gallo@polito.it

2nd Marco Castangia
Politecnico di Torino
Torino, Italia
marco.castangia@polito.it

3rd Alberto Macii
Politecnico di Torino
Torino, Italia
alberto.macii@polito.it

4th Alessandro Aliberti
Politecnico di Torino
Torino, Italia
alessandro.aliberti@polito.it

5th Edoardo Patti
Politecnico di Torino
Torino, Italia
edoardo.patti@polito.it

***Abstract*—This study proposes a deep learning framework for estimating global horizontal irradiance (GHI) from SEVIRI satellite data, incorporating feature importance analysis and using Heliosat-4 as a reference benchmark. Building on prior work, we refine temporal alignment between satellite inputs and ground truth by aggregating 1-minute BSRN GHI measurements to 15-minute and hourly intervals. A multilayer perceptron (MLP), trained on 12 SEVIRI channels and McClear clear-sky GHI, is evaluated using leave-one-location-out cross-validation (LOLO-CV) across 16 diverse stations. Results consistently show improved accuracy over Heliosat-4 at both temporal scales. To enhance model interpretability and reduce computational complexity, we apply SHAP (SHapley Additive exPlanations) analysis to assess channel relevance. A reduced-input model that uses only the most informative channels maintains comparable performance, demonstrating the potential for efficient and scalable deployment. These findings support the feasibility of data-driven, interpretable, and resource-efficient GHI estimation systems for operational and forecasting applications.**

***Index Terms*—Solar Irradiance, GHI, satellite, SEVIRI, EUMETSAT, Deep Learning, SHAP**

I. Introduction

Accurate estimation of surface solar irradiance is critical for solar energy forecasting, climate research, and the operation of modern energy systems [1]. Among the principal components of solar irradiance, global horizontal irradiance (GHI) is particularly relevant due to its direct impact on photovoltaic (PV) performance. Although ground-based pyranometers offer high-precision GHI measurements, their deployment is limited by high costs and operational challenges, especially in remote or extreme environments. Conversely, satellite imagery provides wide spatial coverage and high temporal frequency, making it a valuable data source for developing data-driven models of solar irradiance estimation.

In previous work [2], we introduced a machine learning framework to estimate GHI using full-disk imagery from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) onboard the Meteosat Second Generation (MSG) satellite. The model combined multispectral SEVIRI data centered over target locations with clear-sky GHI values from the McClear model, and was evaluated via leave-one-location-out cross-validation (LOLO-CV) across 16 Baseline Surface Radiation Network (BSRN) stations. That study focused on instantaneous predictions at a 1-minute resolution.

In the present study, we extend this methodology through two main contributions. First, we enhance the temporal alignment between satellite observations and ground truth by aggregating the 1-minute GHI measurements into 15-minute and hourly averages, which are commonly used intervals in solar energy applications and allow for a comprehensive assessment of model performance. Second, we introduce an interpretability analysis using SHapley Additive exPlanations (SHAP) to quantify the contribution of each SEVIRI spectral channel to model performance. This enables us to investigate the potential for dimensionality reduction without compromising estimation accuracy. To validate our approach, we compare model outputs with those of Heliosat-4, a widely used physical-empirical model based on radiative transfer and satellite reflectance data. The results demonstrate that the proposed MLP-based method consistently outperforms Heliosat-4 in terms of accuracy at both 15-minute and hourly resolutions, even when trained with a reduced subset of SEVIRI channels selected via SHAP analysis.

The remainder of the paper is organized as follows: Section II reviews related work on solar radiation estimation techniques. Section III details the datasets, preprocessing steps, and the proposed MLP-based methodology, including the SHAP-based feature importance analysis. Section IV presents and discusses the experimental results, comparing performance across various data aggregation levels and channel configurations. Finally, Section V concludes the paper and outlines directions for future research.

II. Related works

Surface solar irradiance estimation from satellite imagery has traditionally relied on physical or semi-empirical models that incorporate radiative transfer principles. The Heliosat family of models, particularly Heliosat-2 and the more recent Heliosat-4 [3], are among the most widely adopted methods in Europe. These models estimate GHI by deriving cloud indices from visible satellite channels and applying clear-sky corrections based on atmospheric parameters. Also, such methods generally offer interpretable, physically consistent outputs, but they can be sensitive to cloud retrieval errors and often require multiple auxiliary datasets.

979-8-3315-9515-9/25/$31.00

The growing availability of high-frequency satellite imagery, coupled with advances in computational capabilities, has accelerated the adoption of machine learning techniques for solar irradiance estimation. Early works employed random forests [4] and support vector machines [5], often trained on features extracted from geostationary satellite images. More recently, deep learning approaches, including convolutional neural networks (CNNs) [6], [7], recurrent models [8], and hybrid architectures, have been explored to capture spatial and temporal patterns directly from raw satellite inputs. These models have shown strong performance in short-term forecasting tasks and in conditions where traditional physical models often underperform, particularly in the presence of dynamic cloud patterns and atmospheric variability.

In our previous work [2], we developed a multilayer perceptron (MLP) model for the estimation of GHI, trained on all 12 spectral channels from the SEVIRI instrument and clear-sky GHI values derived from the McClear model. The approach demonstrated high accuracy on 16 BSRN stations, evaluated using a LOLO-CV framework. However, that study was limited to 1-minute resolution data and did not investigate the interpretability of the model or the potential benefits of reducing input dimensionality.

In this study, we refine and extend the previous methodology through two key enhancements: i) we align the temporal resolution of ground-based GHI measurements with SEVIRI's 15-minute acquisition cadence, enabling both 15-minute and hourly aggregation analyses; and ii) we incorporate SHAP [9] to quantify the relative importance of each SEVIRI input channel. The proposed model is evaluated via LOLO-CV and benchmarked against Heliosat4 estimates. The obtained results show an average RMSE reduction of 12% at the 15-minute level compared to Heliosat-4. Furthermore, SHAP-based feature selection reveals that the model preserves predictive performance even when retrained on only the four or five most informative SEVIRI channels, highlighting both the robustness and efficiency of the proposed approach.

III. Materials and methods

A. Dataset and Preprocessing

This study leverages multiple complementary datasets to train, validate, and benchmark the proposed GHI estimation model. The primary input consists of full-disk satellite imagery acquired by the SEVIRI onboard the MSG satellite positioned at 0° longitude. All 12 SEVIRI spectral channels, including the High Resolution Visible (HRV) channel, were extracted at 15-minute intervals, corresponding to the dissemination times of the full-disk image scans from the MSG satellite. Table I provides the characterization of the 12 SEVIRI channels [10] along with their major scope. Using specific formulas [11], the pixel-calibrated radiances have been converted to reflectances and brightness temperature for the warm and cold channels, respectively. For each timestamp, image patches were cropped such that the central pixel aligns with one of the 16 selected ground stations. Specifically, 6×6 pixel patches were extracted for the 11 standard-resolution channels, while a 16×16 pixel patch was used for the HRV channel to preserve the same geographical coverage. Additionally, SEVIRI scenes affected by known artifacts (such as missing or negative pixel values) or retrieval issues were excluded to ensure a clean and reliable dataset.

The selected 16 stations belong to the BSRN, whose characterization is reported in Table II. The BSRN provides high-quality, 1-minute resolution measurements of GHI. All BSRN GHI measurements were quality-controlled using the BSRN Toolbox, applying the Solpos algorithm without refraction. Only records labeled with quality codes different from 8, 32, 40, and 42 were retained for training and evaluation. To ensure consistency with the SEVIRI data cadence, the BSRN measurements were averaged into non-overlapping 15-minute intervals only if at least 13 measurements out of the 15 available values were valid. While, hourly averages were computed if at least 52 measurements out of the 60 possible values were valid.

The McClear model, configured to match the 15-minute temporal resolution, was employed to retrieve clear-sky irradiance values via the PVLib Python library. These clear-sky values were used as a model input along with the extracted patches from the SEVIRI images. For benchmark comparison, GHI estimates from the Heliosat-4 model were also obtained using PVLib, ensuring the same temporal resolution and site-specific alignment. Also, we retrieved Heliosat-4 hourly GHI estimates to further enrich the benchmarking of our methodology.

The final dataset spans two years, from 2016 to 2017, covering all seasons and diverse atmospheric conditions. Plus, the spatial diversity of the BSRN stations, ranging from desert to maritime and alpine climates, guarantees a robust basis for evaluating the generalization capabilities of the model through the LOLO-CV strategy.

B. Model Architecture and Training

The proposed estimation model is based on an MLP architecture designed to predict GHI from satellite observations and clear-sky irradiance data. Each training sample corresponds to a specific time instant and station location, and it includes the processed SEVIRI satellite channels and the McClear clear-sky GHI value.

The MLP network, shown in Fig. 1, is composed of three parallel branches, each processing a distinct type of input. Referring to Fig. 1, *Branch 1* receives the 11 standard SEVIRI channels. The input $6 \times 6 \times 11$ pixel image centered over the station location is flattened and concatenated across channels, resulting in a single feature vector. This vector is passed through a fully connected layer with 16 neurons and Linear activation. *Branch 2* processes the HRV channel, in which the 16×16 pixel window is flattened. This higher-resolution input is similarly passed through a fully connected layer with 16 neurons and Linear activation. *Branch 3* takes as input the McClear clear-sky GHI estimate corresponding to the same location and time. This scalar input is passed through its own fully connected layer with 16 neurons and Linear activation. The outputs of the three branches are concatenated and fed into a final fully connected layer with 1024 neurons, followed by a dropout layer with a rate of 0.5 to reduce overfitting. The network outputs a single value, an estimate of the all-sky GHI, through a final output layer with a single neuron and linear activation.

The model is trained using the Adam optimizer to minimize the mean squared error (MSE) loss. Moreover, through a grid search, we determined the optimal hyperparameters, including batch size, learning rate, and number of training

TABLE I
AN OVERVIEW OF THE 12 CHANNELS INCLUDING CENTRAL, MINIMUM AND MAXIMUM WAVELENGTH OF THE SEVIRI CHANNELS, THEIR IDENTIFIERS AND THEIR MAIN AIM.

Channel identifiers		**Characteristics of spectral band (μm)**			**Additional information**
No.	Name	λ_{cen}	λ_{min}	λ_{max}	Aim
1	VIS0.6	0.635	0.56	0.71	Cloud detection and tracking, land surface and vegetation monitoring
2	VIS0.8	0.81	0.74	0.88	Cloud detection and tracking, land surface and vegetation monitoring
3	NIR1.6	1.64	1.50	1.78	Discrimination between snow and cloud, ice and water clouds
4	IR3.9	3.90	3.48	4.36	Low cloud and fog detection
5	WV6.2	6.25	5.35	7.15	Water vapor, winds and semitransparent clouds
6	WV7.3	7.35	6.85	7.85	Water vapor, winds and semitransparent clouds
7	IR8.7	8.70	8.30	9.10	Thin cirrus clouds and discrimination between ice and water clouds
8	IR9.7	9.66	9.38	9.94	Ozone patterns tracking
9	IR10.8	10.80	9.80	11.80	Cirrus cloud and volcanic ash clouds detection
10	IR12.0	12.00	11.00	13.00	Cirrus cloud and volcanic ash clouds detection
11	IR13.4	13.40	12.40	14.40	Cirrus discrimination, cloud top pressure evaluation
12	HRV	Broadband (about 0.4 – 1.1)			Monitoring of convection in high resolution

TABLE II
THE GENERAL INFORMATION OF THE SELECTED BSRN STATIONS ORDERED BY DECREASING LATITUDE.

Location	**Abbr.**	**Latitude (°)**	**Longitude (°)**	**Altitude (m)**
Toravere (EST)	TOR	58.2540	26.4620	70.0
Lindenberg (DEU)	LIN	52.2100	14.1220	125.0
Cabauw (NLD)	CAB	51.9711	4.9267	0
Camborne (GBR)	CAM	50.2167	-5.3167	88.0
Palaiseau (FRA)	PAL	48.7130	2.2080	156.0
Sonnblick (AUT)	SON	47.0540	12.9577	3108.9
Payerne (CHE)	PAY	46.8123	6.9422	491.0
Carpentras (FRA)	CAR	44.0830	5.0590	100.0
Cener (ESP)	CNR	42.8160	-1.6010	471.0
Izaña (ESP)	IZA	28.3093	-16.4993	2372.9
Tamanrasset (DZA)	TAM	22.7903	5.5292	1385.0
Petrolina (BRA)	PTR	-9.0690	-40.3200	387.0
Gobabeb (NAM)	GOB	-23.5614	15.0420	407.0
Florianopolis (BRA)	FLO	27.6047	-48.5227	11.0
São Martinho da Serra (BRA)	SMS	-29.4428	-53.8231	489.0
De Aar (ZAF)	DAA	-30.6667	23.9930	1287.0

epochs [2]. The MLP model, implemented using Tensorflow, has approximately 62K parameters.

We adopted a LOLO-CV strategy to assess the generalization performance of the methodology. At each iteration, the model was trained on data from 15 of the 16 BSRN stations and evaluated on the remaining one. This configuration offers insight into spatial robustness and generalization capability while simulating real-world deployment.

C. SHAP Analysis

SHAP analysis is a unified framework for interpreting machine learning predictions based on cooperative game theory [9]. In this work, we used SHAP to better understand the relative impact of each SEVIRI channel on the predictions of our model. SHAP analysis evaluates the contribution of each input feature to the final model output, offering a quantitative measure of each feature's importance. In more detail, given the trained model f and an input feature vector x, the SHAP value ϕ_i for feature i is defined as:

$$\phi_i(f,x) = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} \left[f_{S\cup\{i\}}(S \cup \{i\}) - f_S(x_S) \right] \tag{1}$$

where F is the full set of input features, while S represents all possible subsets of features excluding feature i. The $f_S(x_S)$ represents the model's output with the *missing* feature i; while $f_{S\cup\{i\}}(S \cup \{i\})$ is the model's prediction including also feature i. These marginal contributions are then weighted across all subsets S. However, evaluating the sum over all subsets S is computationally intractable, especially in our case study, having a total of 653 input features ($6 \times 6 \times 11$ for the 11 channels, 16×16 for the HRV channel, and 1 for the GHI in clear-sky). Therefore, we utilize the DeepExplainer implementation provided in the SHAP Python library, which efficiently approximates SHAP values for deep neural networks. This method builds on DeepLIFT and uses a background dataset to simulate the effect of missing features, allowing the estimation of SHAP values without evaluating all $2^{|N|}$ subsets.

In our experiments, the SHAP analysis was conducted

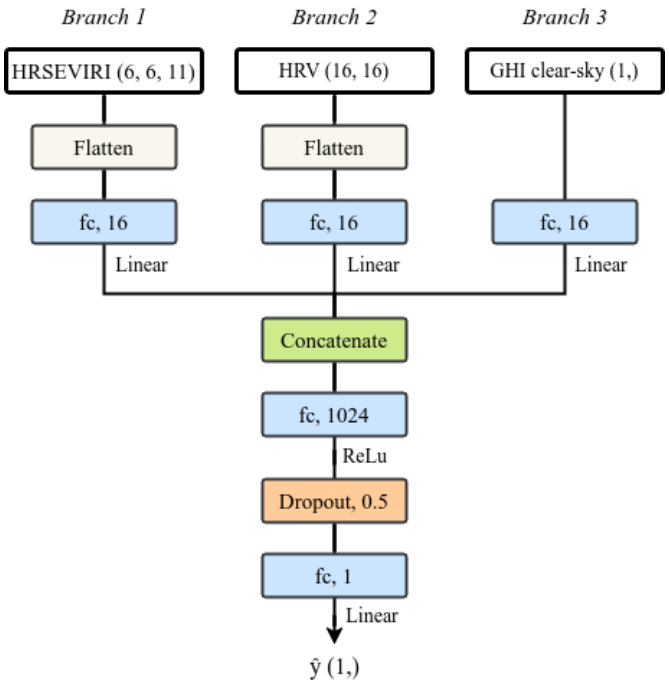

Fig. 1. The structure of the MLP model, where fully connected layers and their output features are shown in light blu boxes, along with their activation functions.

independently for each round of the LOLO-CV procedure. For each trained model in the LOLO-CV procedure, we defined the background dataset by selecting a representative subset of samples drawn from the 15 stations used for training. Specifically, for each training station, we selected data from five days, each characterized by different variability conditions (when available) [2], namely *clear*, *mild*, *moderate*, *high*, and *overcast*. In this way, we preserved the diversity and statistical properties of the training distribution.

SHAP values were then computed over a subset of samples, still including five days with different variability conditions, from the held-out test station, enabling an analysis of the contribution of each SEVIRI channel to the model's predictions. Finally, the resulting SHAP values were averaged to determine the global importance of each SEVIRI channel for all stations. This setup ensured that the feature importance analysis stayed consistent with the cross-validation process. The overall top-ranked channels were then used to train a reduced-input version of the model to empirically determine the importance of the most significant SEVIRI channels.

IV. Results

A. Model evaluation

Through the LOLO-CV strategy, described in Section III-B, we trained the MLP model on 15-minute averaged data. The resulting 15-minute predictions are then either kept at 15-minute resolution or averaged into hourly estimates. These two sets of outputs are subsequently compared with Heliosat-4's all-sky GHI estimates at the corresponding temporal resolutions, 15-minute and 1-hour, respectively.

The evaluation metrics used were the root mean square error (RMSE) and mean bias error (MBE), both expressed in watts per square meter (W/m^2). These metrics quantify, respectively, the average magnitude of the prediction error and the sensitivity to larger deviations, and the average difference between predicted and observed values. Also, we report the coefficient of determination R^2 as a normalized indicator of how well the model captures the variance of the observed GHI. In the following, the mathematical formulation of the evaluation metrics is reported, where y_{pred} is the predicted value, y_{test} is the observed value, $\overline{y}_{test}$ is the mean value of the observed values, and N is the total number of observations.

$$RMSE = \sqrt{\frac{\sum_{n=1}^{N}(y_{pred,n} - y_{test,n})^2}{N}} \tag{2}$$

$$MBE = \frac{\sum_{n=1}^{N} y_{pred,n} - y_{test,n}}{N} \tag{3}$$

$$R^2 = 1 - \frac{\sum_{n=1}^{N}(y_{test,n} - y_{pred,n})^2}{\sum_{n=1}^{N}(y_{test,n} - \overline{y}_{test})^2} \tag{4}$$

Table III reports the performance comparison between the proposed MLP-based model and the Heliosat-4 method for each one of the 16 BSRN stations, based on 15-minute averaged GHI data. The MLP consistently outperforms Heliosat-4 in terms of RMSE and R^2 across almost all sites. For example, the RMSE in Toravere (TOR) decreases from Heliosat-4's 61.189 W/m² to the MLP's 48.659 W/m², and R^2 increases from 0.897 to 0.935. Comparable improvements can be observed in the other stations with the exception of Sonnblick (SON), where Heliosat-4 achieves a lower RMSE. The site's topography and weather may be the cause of this variation, which might pose challenges for data-driven methods.

In terms of bias error, the MLP model frequently yields GHI estimates with little systematic error, highlighted by MBE values frequently close to zero. In contrast, Heliosat-4 has a tendency to exhibit slightly greater biases at several stations, such as Payerne (5.93 W/m²) and Petrolina (9.96 W/m²). Interestingly, observing Table III, the MLP model reveals a general tendency to underestimate the predictions. Most stations present R^2 values above 0.95, with peaks at Gobabeb (0.993) and Carpentras (0.983), reflecting the model's strong ability to follow temporal fluctuations in GHI. Fig. 2 shows, for three sample stations, the MLP and Heliosat-4 estimates against the target measured GHI, and it underlines the higher accuracy just discussed, with the MLP estimates closely following the temporal progression of the target, and correctly detecting sudden and steep variations.

In accordance with the LOLO-CV setup, Table IV displays the estimation results for hourly averaged GHI in addition to the 15-minute averaged analysis. Consistent trends can be seen in the comparison between MLP and Heliosat-4, with the suggested model outperforming the benchmark across almost all stations. These results prove that averaging predictions from finer temporal resolutions enhances the accuracy of hourly estimates, preserving the generalization capabilities exhibited in the higher-resolution setting.

These results demonstrate the proposed MLP model's accuracy and generalizability for GHI estimation using satellite imagery. Even when applied to previously unseen locations, the model demonstrates its effectiveness, proving its potential for scalable and high-resolution solar irradiance estimation.

B. Feature Importance

To gain further insights into the decision-making process of our model and to reduce model complexity, we employed SHAP analysis to evaluate the contribution of each SEVIRI channel to the GHI estimation task. Following the procedure discussed in Section III-C, for each round of the LOLO-CV strategy, we compute the SHAP values using a background

TABLE III
Accuracy of MLP and Heliosat-4 estimates across BSRN stations computed on 15-minute averages.

Stations	MLP			Heliosat-4		
	RMSE	MBE	R^2	RMSE	MBE	R^2
TOR	**48.659**	**-0.088**	**0.935**	61.189	-4.005	0.897
LIN	**45.481**	**-0.099**	**0.951**	50.423	0.296	0.940
CAB	**42.692**	-1.376	**0.957**	51.728	**-0.226**	0.936
CAM	**46.094**	**1.740**	**0.952**	53.937	2.780	0.935
PAL	**45.511**	2.472	**0.957**	54.495	**0.967**	0.938
SON	116.378	-18.309	0.766	**107.656**	**-14.721**	**0.800**
PAY	**44.157**	**-0.338**	**0.966**	55.628	5.933	0.947
CAR	**34.979**	**-0.305**	**0.983**	42.659	6.941	0.975
CNR	**46.644**	**-1.314**	**0.969**	53.686	3.589	0.959
IZA	**66.267**	**-2.939**	**0.968**	67.913	-8.292	0.967
TAM	**42.591**	**0.670**	**0.984**	60.689	-4.544	0.968
PTR	**65.070**	**2.403**	**0.960**	74.897	9.964	0.947
GOB	**31.615**	-3.273	**0.993**	38.272	**-2.903**	0.989
FLO	**75.991**	**2.126**	**0.927**	81.365	4.395	0.917
SMS	**57.042**	**0.893**	**0.962**	67.828	-2.273	0.947
DAA	**46.301**	**-0.109**	**0.983**	52.925	1.764	0.977

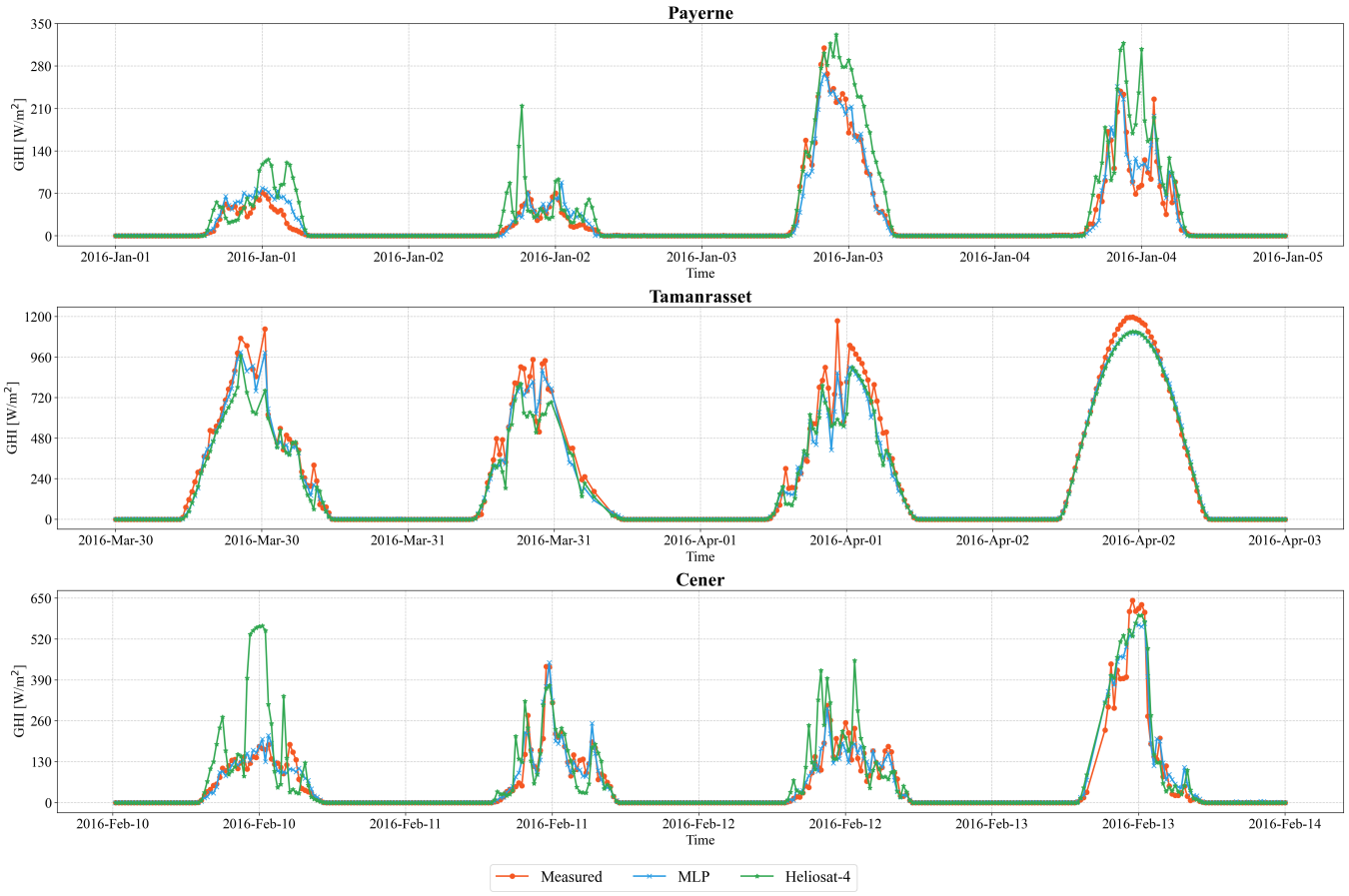

Fig. 2. Comparison of 15-minute averaged GHI prediction curves for the proposed MLP model and the Heliosat-4 method against ground measurements at three representative stations. Each subplot illustrates the temporal alignment and accuracy of both models in capturing irradiance dynamics across different climatic and geographical conditions.

TABLE IV
ACCURACY OF MLP AND HELIOSAT-4 ESTIMATES ACROSS BSRN STATIONS COMPUTED ON HOURLY AVERAGES.

Stations	MLP			Heliosat-4		
	RMSE	MBE	R^2	RMSE	MBE	R^2
TOR	**34.192**	**-0.136**	**0.966**	46.276	-4.010	0.938
LIN	**29.029**	**-0.125**	**0.979**	34.913	0.276	0.970
CAB	**26.438**	-1.383	**0.983**	35.257	**-0.245**	0.969
CAM	**33.649**	**1.764**	**0.974**	40.371	2.784	0.962
PAL	**29.859**	2.518	**0.980**	38.086	**1.002**	0.968
SON	108.099	-17.846	0.789	**95.226**	**-13.997**	**0.836**
PAY	**30.500**	**-0.314**	**0.983**	41.023	5.938	0.970
CAR	**24.960**	**-0.336**	**0.992**	32.018	7.112	0.986
CEN	**33.949**	**-1.318**	**0.983**	39.244	3.596	0.977
IZA	**59.920**	**-2.639**	**0.974**	61.222	-7.666	0.972
TAM	**29.813**	**1.610**	**0.992**	45.242	-3.037	0.981
PTR	**43.335**	**2.054**	**0.982**	52.729	9.483	0.973
GOB	**25.334**	-3.273	**0.995**	30.189	**-2.897**	0.993
FLO	**62.378**	**2.387**	**0.949**	66.087	4.654	0.943
SMS	**39.792**	**0.995**	**0.981**	51.877	-1.994	0.968
DAA	**29.039**	**0.015**	**0.993**	35.693	1.788	0.990

dataset drawn from the training stations and a subset of samples from the corresponding held-out test station.

The SHAP analysis reveals that not all channels equally contribute to the model's output. As reported in Table V, channels in the visible range (i.e, 1, 2, and 3) consistently emerged as the most influential across stations, followed by channel 8 and, to a slightly lesser extent, channels 6, 7, and 10. Channel 12 also showed a stable, albeit moderate, relevance. In contrast, channels 5 and 9 contributed minimally in most cases.

To empirically verify the results of SHAP analysis, we tested two reduced model configurations using only the top-ranked channels, i.e., channels 1, 2, 3, 8, and 12. We included channel 12 in the analysis, since it exhibits a stable and decent relevance across all stations. In the first configuration, we solely use channels 1, 2, 3, and 8, out of the available SEVIRI channels; in the second configuration, we also included channel 12. As shown in Figure 3, the predictive performance of the original model employing all channels is preserved with little degradation when using smaller sets of inputs. Indeed, compared with original model, the first reduced model configuration registers an overall average RMSE worsening of 1.359 W/m^2; while, for the second reduced model configuration the overall average RMSE deteriorates by just 0.575 W/m^2.

These reults suggest that a more compact input representation, informed by SHAP analysis, can be adopted without compromising the accuracy. Moreover, this feature selection strategy offers practical advantages, such as reducing input dimensionality, which can simplify deployment in operational settings or speed up inference when bandwidth or storage is limited.

TABLE V

THE RELATIVE SHAP VALUES FOR EACH OF THE 12 SEVIRI CHANNELS, ACROSS ALL SELECTED BSRN STATIONS. HIGHLIGHTED IN GREEN ARE THE AVERAGE MORE INFORMATIVE CHANNELS, SELECTED FOR THE RETRAINING; IN RED ARE THE RELATIVELY LESS INFORMATIVE CHANNELS, EXCLUDED FROM THE RETRAINING; IN YELLOW CHANNEL 12 (HRV) WHICH MAINTAINS A STABLE IMPORTANCE ACROSS ALL STATIONS.

Channel	**SHAP values**																
	TOR	LIN	CAB	CAM	PAL	SON	PAY	CAR	CNR	IZA	TAM	PTR	GOB	FLO	SMS	DAA	**Mean**
1	0.231	0.273	0.239	0.209	0.216	0.405	0.273	0.231	0.271	0.365	0.176	0.190	0.166	0.224	0.287	0.213	0.248
2	0.221	0.230	0.241	0.185	0.220	0.071	0.227	0.219	0.267	0.095	0.172	0.166	0.152	0.214	0.249	0.182	0.194
3	0.124	0.155	0.130	0.120	0.148	0.136	0.196	0.136	0.112	0.091	0.158	0.156	0.165	0.159	0.153	0.137	0.142
4	0.050	0.044	0.051	0.058	0.058	0.034	0.048	0.055	0.036	0.031	0.048	0.057	0.072	0.051	0.043	0.056	0.049
5	0.024	0.017	0.018	0.026	0.016	0.013	0.016	0.016	0.017	0.013	0.023	0.02	0.032	0.041	0.021	0.026	0.021
6	0.061	0.035	0.040	0.068	0.049	0.056	0.025	0.046	0.046	0.068	0.064	0.059	0.060	0.060	0.039	0.060	0.052
7	0.052	0.046	0.055	0.063	0.064	0.036	0.043	0.063	0.050	0.036	0.056	0.067	0.063	0.060	0.044	0.070	0.054
8	0.095	0.073	0.088	0.112	0.077	0.052	0.064	0.088	0.062	0.065	0.114	0.133	0.143	0.065	0.052	0.113	0.087
9	0.030	0.018	0.025	0.026	0.026	0.017	0.015	0.028	0.025	0.026	0.034	0.023	0.041	0.018	0.017	0.023	0.025
10	0.047	0.039	0.045	0.055	0.057	0.085	0.033	0.050	0.043	0.092	0.069	0.048	0.047	0.048	0.035	0.052	0.053
11	0.031	0.029	0.030	0.045	0.036	0.036	0.025	0.032	0.030	0.035	0.055	0.051	0.036	0.030	0.027	0.043	0.036
12	0.034	0.042	0.036	0.034	0.035	0.059	0.035	0.033	0.042	0.085	0.030	0.031	0.023	0.030	0.033	0.026	0.038

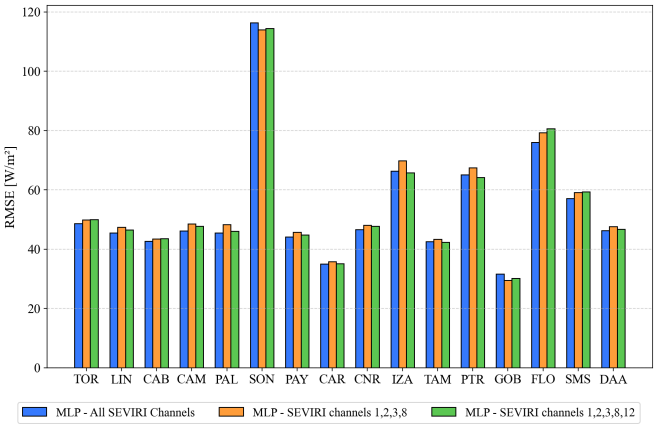

Fig. 3. Comparison of RMSE values for GHI estimates computed on 15-minute averages across stations using three SEVIRI channel configurations: all channels, channels 1, 2, 3, 8, and channels 1, 2, 3, 8, 12.

V. CONCLUSIONS

This work presents a deep-learning approach for high-resolution GHI estimation using SEVIRI satellite data. Employing an MLP model evaluated through a LOLO-CV strategy, we consistently achieve strong predictive performance across a diverse set of 16 BSRN stations. Significant gains in accuracy and bias metrics are shown in the comparisons with the Heliosat-4 method for both 15-minute and hourly resolutions. Moreover, we leverage SHAP values to guide feature selection and interpret the behavior of the model. This analysis reveals that model performance is strongly influenced by a small subset of SEVIRI channels. Indeed, retraining the model with only the top-ranked channels by SHAP confirms that we achieve the same performance with a reduced set of channels. The high accuracy, interpretability, and input reduction jointly suggest that the proposed methodology is not only effective but also practical for scalable solar irradiance estimation. In future work, we plan to explore alternative clear-sky models, tailored to specific climate zones, that reduce the dependence on indirect inputs inherent to McClear and further improve the accuracy of the estimation across diverse atmospheric and climatic conditions.

REFERENCES

[1] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F.J. Martinez-de-Pison, F. Antonanzas-Torres, Review of photovoltaic power forecasting, Solar Energy, 2016, vol. 136, pp.78–111.

[2] R. Gallo, M. Castangia, A. Macii, E. Patti, A. Aliberti, Neural networks for estimating surface solar irradiation from satellite images, Engineering Applications of Artificial Intelligence, 2025, vol. 144, pp.110101., ISSN 0952-1976, https://doi.org/10.1016/j.engappai.2025.110101.

[3] Z. Qu, A. Oumbe, P. Blanc, B. Espinar, et al. Fast radiative transfer parameterisation for assessing the surface solar irradiance: The Heliosat-4 method. Meteorologische Zeitschrift, 2017, vol. 26(1), pp.33–57.

[4] B. Babar, L. T. Luppino, T. Boström, S. N. Anfinsen, Random forest regression for improved mapping of solar irradiance at high latitudes, Solar Energy, 2020, vol. 198, pp.81–92.

[5] C. Voyant, G. Notton, S. Kalogirou, M. Nivet, C. Paoli, F. Motte, A. Fouilloy, Machine learning methods for solar radiation forecasting: A review. Renewable Energy, 2017, vol. 105, pp.569-–582.

[6] S. Chen, C. Li, Y. Xie, M. Li, Global and direct solar irradiance estimation using deep learning and selected spectral satellite images, Applied Energy, 2023, vol. 352, pp.121979.

[7] E. O. Yuzer, A. Bozkurt, Deep learning model for regional solar radiation estimation using satellite images, 2023, vol. 14(8), pp.102057.

[8] F. A. Suwadana, P. M. P. Garniwa, D. A. Putera, D. Puspita, A. Gufron, I. A. Aditya, H. Lee, & I. Garniwa, Solar Irradiance Estimation in Tropical Regions Using Recurrent Neural Networks and WRF Models. Energies, 2025, vol. 18(4), 925.

[9] S. Lundberg, S. Lee, A unified approach to interpreting model predictions, NeurIPS, 2017.

[10] Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., & Ratier, A. (2002). AN INTRODUCTION TO METEOSAT SECOND GENERATION (MSG). Bulletin of the American Meteorological Society, vol. 83(7), pp.977-992.

[11] EUMETSAT, MSG High Rate SEVIRI level 1.5 data guide, https://user.eumetsat.int/resources/user-guides/msg-high-rate-seviri-level-1-5-data-guide#ID-Calibrated-radiances, Accessed: 2025-05-19.