Deep Learning for Satellite-Based Solar Irradiance Forecasting

Andrés Herrera[1], Franco Marchesoni-Acland[2] and Rodrigo Alonso-Suárez[1,3]

[1] Laboratorio de Energía Solar, Facultad de Ingeniería, UDELAR, Montevideo, Uruguay

[2] Université Paris-Saclay, ENS Paris-Saclay, CNRS, Centre Borelli, Gif-sur-Yvette, France

[3] Laboratorio de Energía Solar, CENUR Litoral Norte, UDELAR, Salto, Uruguay

Abstract

This work presents the solar irradiance evaluation of a novel intraday solar forecasting method that applies deep learning to geostationary satellite imagery. The U-Net architecture is used to predict the future cloud position up to 5 hours ahead, which is then converted into a solar irradiance forecast using a semi-empirical satellite-to-irradiance model. The method is implemented using satellite earth reflectance images from the GOES-16 visible channel. It is trained to maximize reflectance prediction accuracy (image prediction) and is evaluated for both satellite reflectance and solar irradiance forecast (Global Horizontal Irradiance, GHI). The solar irradiance prediction is evaluated using high-quality ground measurement for a site located in the Pampa Húmeda region of southeastern South America. The research demonstrates significant improvements over traditional methods such as those based on the cloud motion vector estimation and extends previous research reported in Marchesoni-Acland et al. (2023).

Keywords: Solar forecasting, deterministic forecast, satellite imagery, GOES16, GHI.

1. Introduction

Integrating solar energy into electricity grids presents significant challenges due to its inherent variability, largely driven by cloud dynamics that affect solar irradiance over photovoltaic power stations. This variability necessitates accurate forecasting of solar resource availability for effective grid management and energy trading. This research addresses satellite intraday solar irradiance forecasting up to 5 hours ahead, usually referred to as nowcasting or very short-term forecasting.

Solar irradiance prediction is mainly driven by cloud formation, deformation, displacement, and dissipation, which are difficult to predict. Numerical Weather Prediction (NWP) models have traditionally provided cloudiness and solar irradiance forecasts through physical modelling. These models remain the best option for day-ahead forecasts, typically enhanced by post-processing techniques (Yang et al., 2022). However, for intraday solar irradiance forecasts, satellite-based methods have demonstrated superior performance compared to regular NWP runs for horizons up to 4-5 hours ahead (Kühnert et al., 2013; Perez et al., 2010).

The main technique for satellite-based solar irradiance prediction is cloud motion vectors (CMV), which estimate a two-dimensional cloud motion field from the image sequence and use it to extrapolate the future cloud position. In Aicardi et al. (2022) several CMV methods were compared, including the block-matching technique of Lorenz et al. (2004) and various optical flow methods (Horn & Schunck, 1981; Lucas & Kanade, 1981; Farnebäck, 2003; Zach et al., 2007). This work found that the TVL1 algorithm is the best CMV strategy for image and solar irradiance prediction. However, CMV strategies have several limitations, the main ones being motion extrapolation based on static motion fields and the inability to accurately predict cloud formation and/or extinction. In fact, pure CMV strategies can only reallocate pixels of satellite images but cannot create new values.

Deep learning methods have recently been introduced in several areas related to meteorology (Ren et al., 2021; Ravuri et al., 2021; Schneider et al., 2022). The forecasting problem is inherently self-supervised, as the targets

exist naturally in the historical record without the need for manual annotation. Geostationary meteorological satellites provide large amounts of data (typically tens of terabytes) at regular 10-15 minute intervals, making intraday satellite cloudiness forecasts particularly suitable for deep learning applications. For this reason, recent research has applied deep learning to cloudiness and solar forecasts. Marchesoni et al. (2023) and Berthomier et al. (2020) used U-Net architectures for cloud prediction in different ways, while Nielsen et al. (2021) proposed a custom architecture called IrradianceNet for solar irradiance prediction. Marchesoni et al. compared these approaches on the same dataset and showed that the U-Net is the best performing deep learning method for the albedo prediction problem (cloudiness as seen from the satellite visible channel). The present work extends that of Marchesoni et al. (2023) to solar irradiance prediction, a research gap that was not filled in the previous work.

2. Data

The solar irradiance ground truth in this work corresponds to the Global Horizontal Irradiance (GHI) measurement at the LES experimental facility, located in northwestern Uruguay. This site operates a Kipp & Zonen SOLYS2 ground station equipped with class A spectrally flat pyranometers (ISO 9060:2018 standard). Measurements are recorded at 1-minute intervals as averages of 10-second samples. The station undergoes daily maintenance, including dome cleaning and horizontal plane checks, with instruments calibrated every two years against a Kipp & Zonen CMP22 reference pyranometer with World Radiometric Reference traceability. The facility maintains high quality standards in accordance with Baseline Surface Radiation Network (BSRN) requirements. The hourly ground data for this work were calculated by averaging the corresponding 1-minute samples for each forecast horizon based on the satellite time stamp. Data quality control was applied to the 1-minute data in accordance with the BSRN guidelines (McArthur, 2005).

The satellite image dataset consists of imagery from GOES16's Advanced Baseline Imager (ABI) captured on Channel 02, covering the period from April 2019 to March 2025. These images provide reflectance factor in the spectral range of 0.59–0.69 μm and are distributed as level-2 Cloud and Moisture Imagery (CMI), also known as the 'red' channel. The images used in this work are a crop of the full disk images centered at 31.39°S, 57.95°W, the location of the LES facility. The images have a shape of 512 × 512 pixels and cover the spatial region from 28.40°S to 34.58°S in latitude and from 61.39°W to 54.04°W in longitude, resulting in a pixel resolution of approximately 1 km. The images are separated by a ten-minute interval. Only images in which all pixels have at least a cosine of the solar zenith angle greater than 0.15 are included in the dataset. The first 70% of the recorded images were used for training purposes, with the remaining 30% were reserved for evaluating performance on unseen images. The dataset is divided by day. Intraday division is inadequate and should not be performed, as this would maintain correlations between the training and evaluation sets. This would result in an inaccurate and unrepresentative evaluation of performance.

3. Methods

This work compares two distinct approaches to predict future satellite imagery and site-specific GHI: a deep learning technique using a U-Net architecture and traditional Cloud Motion Vector (CMV) methods based on optical flow algorithms. Both selected optical flow methods provide dense motion estimation without the need for interpolation techniques, and all methods are capable of pixel-wise image prediction. The methods are evaluated against persistence baselines at image and GHI levels.

Figure 1 provides an outline of the process. The first step is to convert the images from reflectance factor to Earth reflectance, also known as Earth albedo. Then, the image forecast for up to five hours ahead is made using the most recently available Earth albedo image. Here, different methods for image forecasting are tested, e.g. the U-Net and CMV techniques. The final step is to convert the predicted images into a GHI forecast. For this, a simple Cloud Index Model (CIM) is used in which the clear sky estimates from the ESRA model (Rigollier et al., 2004) are attenuated by a locally adjusted linear function of the cloud index (Cano et al., 1986), calculated from the predicted Earth albedo images. This approach, known as CIM-ESRA, is a standard

semi-empirical method for converting satellite images to GHI. Full details can be found in Laguarda et al., (2020). In particular, spatial averaging is used to improve the representation of hourly GHI, as described in Laguarda et al., but adapted to GOES16.

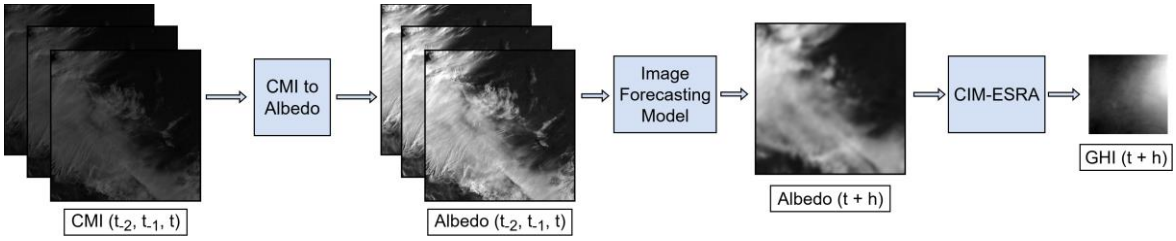

Figure 1: Diagram of the process by which the raw GOES-16 Channel 02 images are used to forecast solar irradiance.

3.1. Deep Learning Approach: U-Net

The deep learning method uses a U-Net architecture (Ronneberger et al., 2015) adapted for spatio-temporal image prediction of GOES-16 satellite imagery. The model is able to directly predict future cloud patterns and reflectance values.

The U-Net architecture employs an encoder-decoder structure with skip connections, enabling it to learn multi-scale features and patterns in image sequences. The model uses the three most recent images as input (at times *t-2*, *t-1*, and *t*, with a time step of 10 minutes) and was trained using Mean Absolute Error (MAE) as the loss function. Separate models were trained to predict images at each forecast time horizon, ranging from one to five hours ahead.

The Deep Learning approach aims to capture complex cloud dynamics, including movement, deformation, and critically, cloud formation and extinction—capabilities that are notably absent in traditional CMV algorithms. This U-Net method for geostationary satellite image prediction was proposed and tested in Marchesoni-Acland et al. (2023), where full details are provided. This work extends the performance assessment to GHI, since in previous work the method was only evaluated at image level.

3.2. Cloud Motion Vector Methods

Two optical flow algorithms were implemented as benchmark methods: the TVL1 and Farnebäck methods. Both algorithms estimate a static cloud motion field from two consecutive satellite images and extrapolate this motion to generate future images. In the following two subsections both methods are described.

3.2.1. TVL1 Method

The TVL1 algorithm (Zach et al., 2007; Sánchez et al., 2013) combines an L_1 data penalty term on the optical flow equation with a Total Variation (TV) regularization. Its mathematical expression is formulated as the following optimization problem:

$$\arg\min_{u,v}\left\{\int_I |\nabla u| + |\nabla v| + \lambda \cdot \left|I_x \cdot u + I_y \cdot v + I_t\right|\right\}, \qquad \text{(eq. 1)}$$

where u and v represent the horizontal components of the motion field in the x and y directions, respectively, and I_x, I_y, and I_t are the spatial and temporal derivatives of the image. λ is a trade-off parameter between the regularization and data terms. The TV regularization promotes piecewise-smooth solutions while preserving discontinuities in the motion field, making it particularly suitable for capturing interfaces between clouds at different altitudes moving in different directions.

The implementation parameters were set according to the optimal values found in Aicardi et al., (2022). These are $\lambda = 0.055$ and $M = 6$ (down-scaling levels). This is done to maintain the implementation of the algorithm that had the best performance in that work.

3.2.2. Farnebäck Method

The Farnebäck (2003) method is based on a second-order polynomial expansion of the neighborhood of each pixel, designed to handle noisy sequences with high-frequency and abrupt variations in the motion field. The implementation of this algorithm for this work uses the OpenCV calcOpticalFlowFarneback function with locally optimized parameters from Marchesoni-Acland et al. (2023). The parameters that were used here are presented in Table 1.

Table 1: Optimized parameters for the Farnebäck optical flow method.

Parameter	pyr_scale	levels	winsize	iterations	poly_n	poly_sigma
Value	0.3987	4	22	3	5	0.8480

3.2.3. Image Extrapolation

For both CMV methods, predicted images are generated using a backward search projection algorithm, where each pixel in the predicted image is constructed through bi-linear interpolation of the previous image using the inverted and scaled motion field. The prediction process is iterative: for each forecast horizon, the CMV is applied to the previously predicted image, starting from the current time *t* image.

3.3. Baseline Methods

In order to decide whether methods' performance is acceptable, they need to be compared to simple, naive baselines. Methods that do not outperform simple baselines should be disregarded. Also, forecasting skill scores need a reference for computation and this is provided by baseline methods. In this work, baselines are implemented at image and GHI levels, both of which are based on the persistence principle.

3.3.1. Image Persistence

At the image level, a simple persistence baseline is implemented by maintaining constant the Earth reflectance image at time *t* across all forecast horizons. This is the same image-level baseline that was used in Aicardi et al. (2022) and Marchesoni-Acland et al. (2023). This naive approach assumes no change in cloud patterns and provides a fundamental performance reference against which all forecasting methods can be evaluated. This is a reasonable persistence implementation at image level, given that the Earth's albedo is a geometrically normalized, dimensionless quantity with little to no geometrical influence (e.g. the absence of spatial structures associated with the relative position between the Sun and the Earth). For the same reason, the image prediction methods implemented in this study (U-Net and CMV) use the Earth's albedo as both the input and the output. This allows the methods to avoid dealing with geometrical variations and focus solely on the optical properties and structures of the clouds and the background (Earth surface) that are present in the image sequence.

3.3.2. Clear-Sky Index Persistence

The GHI baseline is implemented through the persistence of the clear-sky index (k_c), which is calculated from solar irradiance ground measurements. k_c is also a geometrically normalized, dimensionless quantity, which is defined as the ratio of the GHI to that of a clear-sky model. The k_c persistence is implemented on an hourly timescale, in line with the time step of the forecast horizon. The clear-sky model used is the McClear model (Lefèvre et al., 2013). The "persisted" GHI for all hourly forecast horizons is computed as $\mathrm{GHI}(t+h) = k_c(t) \times \mathrm{GHI}_{csk}(t+h)$, where h is the forecast horizon and $\mathrm{GHI}_{csk}(t+h)$ are the clear-sky model estimates that can be computed in advance for any *(t+h)*.

3.4. Evaluation Framework

The forecasting methods are evaluated at two levels: image-level and GHI-level performance. Image-level evaluation compares predicted albedo images pixel-by-pixel with the ground truth image for each forecast horizon. All image pixels are considered for persistence and U-Net models prediction. For CMV models, some border pixels are excluded because the method is unable to provide predictions in situations where out-of-area pixels are needed to generate the borders of the predicted image. The metrics employed for performance evaluation at both levels are the Mean Bias Error (MBE), the Mean Absolute Error (MAE), the Root Mean

Square Error (RMSE), and the Forecasting Skill (FS). The first three metrics can be expressed as a percentage of the average value for each magnitude. For image-level evaluation, this is the mean Earth albedo, and for irradiance-level evaluation, this is the mean measured GHI. The relative versions of these metrics are rMBE, rMAE, and rRMSE, respectively.

4. Results

This section presents a comparative performance analysis of the U-Net deep learning model against traditional Cloud Motion Vector (CMV) methods and persistence benchmarks. Following the two-level evaluation framework suggested by Aicardi et al. (2022), the assessment is first conducted at the image level to evaluate the accuracy of cloud prediction and subsequently at the irradiance level to determine the practical performance for Global Horizontal Irradiance (GHI) forecasting. The next two subsections each address the main discussion of these two evaluations. Full quantitative results can be found in the Appendix (Tables 2 to 7).

4.1. Image-level Prediction Performance

The initial evaluation focuses on the core task of predicting future satellite reflectance images. Figure 2 displays all metrics as a function of the forecast horizon for the U-Net model compared to the TVL1 and Farnebäck optical flow algorithms, as well as the image persistence procedure. The MBE, MAE and RMSE are shown in their absolute form (i.e. not expressed as a percentage of the mean value). The relative rMBE, rMAE and rRMSE plots are shown in the Appendix in Figure 4. The results clearly demonstrate the superiority of the U-Net architecture in predicting cloud patterns. Across all forecast horizons, the U-Net model achieves the lowest MAE and RMSE, indicating a significant improvement in accuracy over traditional methods. The TVL1 algorithm, identified by Aicardi et al. as the best-performing CMV method, serves as the strongest optical flow alternative, yet the U-Net consistently outperforms it. This highlights the deep learning model's enhanced ability to capture complex, non-linear dynamics of cloud evolution, including dissipation, as shown before in Marchesoni-Acland et al. (2023), which CMV methods cannot model.

The MBE plot, however, reveals a different behavior. The TVL1 and Farnebäck models exhibit smaller bias than the U-Net model for all time horizons. Furthermore, the U-Net model exhibits a consistent negative bias that increases with the forecast horizon. In contrast, the optical flow methods demonstrate mixed results: positive for shorter time horizons and negative for larger ones. This suggests that the U-Net model tends to slightly underpredict satellite reflectance more than the others, particularly for longer lead times. This characteristic warrants further investigation in future work.

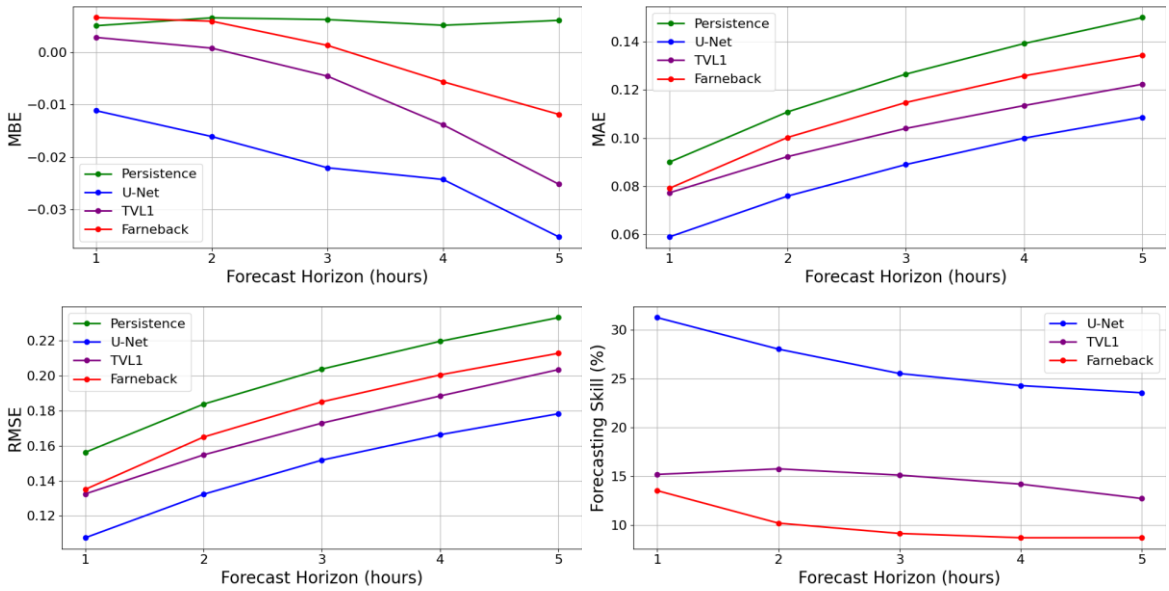

Figure 2: Image-level performance evaluation as a function of the forecasting horizon. MBE, MAE and RMSE are dimensionless as the image prediction is for Earth albedo. The FS is expressed as a percentage.

4.2. GHI Prediction Performance

Building upon the image-level analysis, the predicted albedo images were converted to GHI forecasts using the CIM-ESRA model. This step evaluates the practical utility of each method for solar energy applications. Figure 3 presents all the performance metrics for the GHI forecast at the LES experimental facility for horizons from 1 to 5 hours. The MBE, MAE, RMSE, and FS for the irradiance-level evaluation is shown in Figure 3. The relative versions of the first three metrics are shown in the Appendix in Figure 5.

The GHI forecasting results confirm the findings from the image-level analysis. The U-Net model delivers the most accurate GHI predictions, achieving the lowest RMSE at every forecast horizon. This result validates the hypothesis that higher accuracy in image prediction directly translates to improved solar irradiance forecasting, as found in Aicardi et al. (2022). The underestimation bias observed in the image-level evaluation of the U-Net is converted to an overestimation bias in the irradiance-level evaluation, which is reasonable and expected. The FS decreased from a range of 24-31% to a stable level of 22-23% across all forecasting time horizons. Note that the persistence baseline differs between the two evaluations; therefore, the FS behavior may change from one to the other.

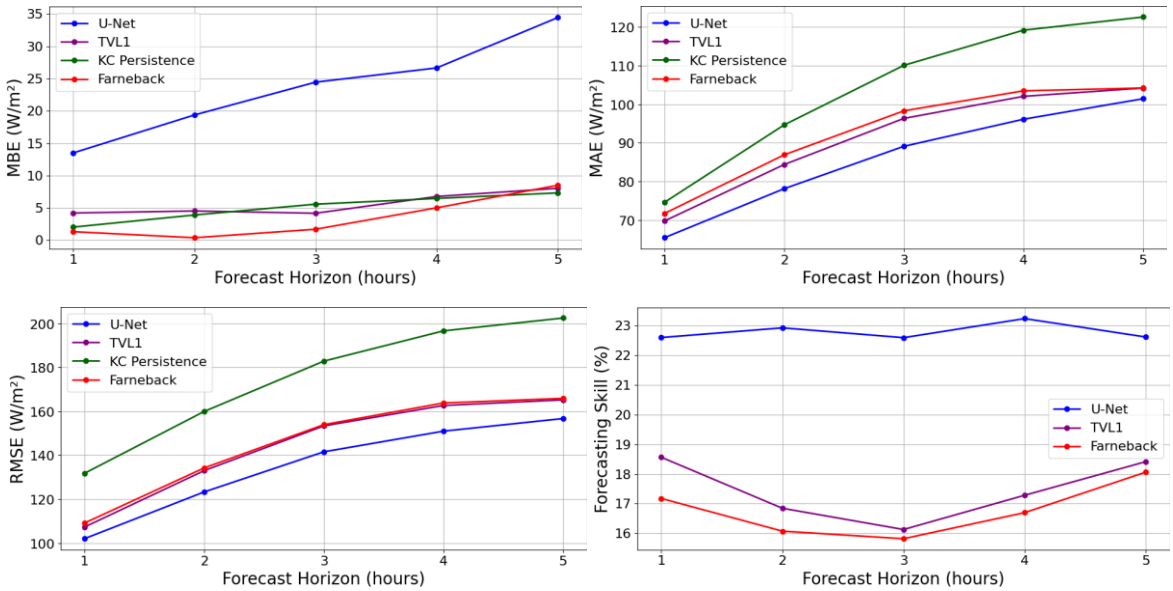

Figure 3: GHI-level performance evaluation as a function of the forecasting horizon. MBE, MAE and RMSE are expressed in W/m². The FS is expressed as a percentage.

All CMV-based and deep learning-based methods successfully outperform the standard GHI persistence benchmark based on k_c, demonstrating their value in providing skillful forecasts. More importantly, the U-Net's performance margin over the TV-L1 and Farnebäck models is consistent across the time horizons. This underscores the practical advantage of a deep learning approach. The U-Net can better anticipate significant changes in sky conditions that directly impact surface irradiance. The static motion fields used by CMV methods become progressively less accurate over longer horizons, whereas the U-Net's spatiotemporal learning capability appears more robust over the entire 5-hour nowcasting window.

5. Conclusions

This work successfully demonstrated the effectiveness of a U-Net deep learning architecture for intraday solar irradiance prediction for hourly forecast horizons up to 5 hours ahead. The evaluation, conducted at both the satellite image and ground irradiance levels, consistently showed the superiority of the deep learning approach over traditional CMV methods, including the high-performing TVL1 cloud motion vector algorithm and persistence benchmarks.

The key finding is that the U-Net model's enhanced accuracy in predicting future cloud reflectance, e.g. lowest MAE and RMSE at all forecast horizons, directly translates into more precise GHI forecasts. This performance advantage stems from the model's ability to capture the complex, non-linear spatiotemporal dynamics of cloud evolution, including dissipation phenomena that static CMV methods inherently struggle to predict. Accurate

cloud dissipation forecasting has been observed for the U-Net implementation in Marchesoni-Acland et al. (2023). However, we have not observed yet accurate cloud formation forecasting.

The results validate the hypothesis that deep learning provides a more robust and accurate solution for solar irradiance satellite-based nowcasting. This methodology holds significant potential for improving the operational management of solar power and its integration into electricity grids, offering more reliable inputs for grid stability and energy trading.

6. Acknowledgments

The authors acknowledge financial from the ANII FSE 1 2022_1_175459 project, Uruguay. R. Alonso-Suárez acknowledge financial support from Uruguay's CSIC-UDELAR research program.

7. References

Aicardi, D., Musé, P., & Alonso-Suárez, R., 2022. A comparison of satellite cloud motion vectors techniques to forecast intra-day hourly solar global horizontal irradiation. Solar Energy, 233, 46–60.

Berthomier, L., Pradel, B., & Perez, L., 2020. Cloud cover nowcasting with deep learning. In Proceedings of the 10th International Conference on Image Processing Theory, Tools and Applications (IPTA) (pp. 1–6).

Cano, D., Monget, J., Albuisson, M., Guillard, H., Regas, N., & Wald, L., 1986. A method for the determination of the global solar radiation from meteorological satellite data. Solar Energy, 37, 31–39.

Farnebäck, G., 2003. Two-frame motion estimation based on polynomial expansion. In J. Bigun, & T. Gustavsson (Eds.), Image Analysis (pp. 363–370). Berlin, Heidelberg: Springer Berlin Heidelberg.

Horn, B. K. P., & Schunck, B. G., 1981. Determining optical flow. Artif. Intell., 17, 185–203.

Kühnert, J., Lorenz, E., & Heinemann, D., 2013. Chapter 11 - Satellite-based irradiance and power forecasting for the german energy market. In J. Kleissl (Ed.), Solar Energy Forecasting and Resource Assessment (pp. 267–297). Boston: Academic Press.

Laguarda, A., Giacosa, G., Alonso-Suárez, R., & Abal, G., 2020. Performance of the site-adapted CAMS database and locally adjusted cloud index models for estimating global solar horizontal irradiation over the Pampa Húmeda. Solar Energy, 199, 295–307.

Lefèvre et al., 2013. Mcclear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions. Atmospheric Measurement Techniques, 6, 2403–2418.

Lorenz, E., Hammer, A., & Heinemann, D., 2004. Short term forecasting of solar radiation based on satellite data. In EUROSUN2004 (ISES Europe Solar Congress) (pp. 841–848). Freiburg, Germany.

Lucas, B. D., & Kanade, T., 1981. An iterative image registration technique with an application to stereo vision. In Proceedings of the 7th International Joint Conference on Artificial Intelligence - Volume 2 IJCAI'81 (pp. 674–679). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Marchesoni-Acland, F., Herrera, A., Mozo, F., Camiruaga, I., Castro, A., & Alonso-Suárez, R., 2023. Deep learning methods for intra-day cloudiness prediction using geostationary satellite images in a solar forecasting framework. Solar Energy, 262, 111820.

McArthur, L., 2005. Baseline Surface Radiation Network Operations Manual. Technical Report WCRP-121/WMO TD-No. 1274 World Climate Research Programme – WMO.

Nielsen, A. H., Iosifidis, A., & Karstoft, H., 2021. IrradianceNet: Spatiotemporal deep learning model for satellite-derived solar irradiance short-term forecasting. Solar Energy, 228, 659–669.

Perez, R., Kivalov, S., Schlemmer, J., Hemker, K., Renné, D., Hoff, T.E., 2010. Validation of short and medium term operational solar radiation forecasts in the US, Solar Energy, 84(12), 2161-2172.

Ravuri, S., Lenc, K., Willson, M., Kangin, D., Lam, R., Mirowski, P., Fitzsimons, M., Athanassiadou, M., Kashem, S., Madge, S. et al., 2021. Skilful precipitation nowcasting using deep generative models of radar. Nature, 597, 672–677.

Ren, X., Li, X., Ren, K., Song, J., Xu, Z., Deng, K., & Wang, X., 2021. Deep learning-based weather prediction: A survey. Big Data Research, 23, 100178.

Rigollier, C., Lefèvre, M., & Wald, L., 2004. The method Heliosat-2 for deriving shortwave solar radiation from satellite images. Solar Energy, 77, 159–169.

Ronneberger, O., Fischer, P., & Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 (pp. 234–241). Cham: Springer International Publishing.

Schneider, R., Bonavita, M., Geer, A., Arcucci, R., Dueben, P., Vitolo, C., Le Saux, B., Demir, B., & Mathieu, P.-P., 2022. ESA-ECMWF report on recent progress and research directions in machine learning for Earth system observation and prediction. npj Climate and Atmospheric Science, 5, 1–5.

Sánchez, J., Meinhardt-Llopis, E., & Facciolo, G, 2013. TV-L1 optical flow estimation. Image Processing Online (IPOL), .

Yang et al., 2022. A review of solar forecasting, its dependence on atmospheric sciences and implications for grid integration: Towards carbon neutrality. Renewable and Sustainable Energy Reviews, 161, 112348.

Zach, C., Pock, T., & Bischof, H., 2007. A duality based approach for realtime TV-L1 optical flow. In Proceedings of the 29th DAGM Conference on Pattern Recognition (pp. 214–223).

8. Appendix – Performance Metrics

8.1. Image-Level Complementary Evaluation

Table 2: Performance metrics of the U-Net algorithm at image level. The average albedo for normalized metrics is 0.29.

Forecast horizon	MBE (no unit)	rMBE (%)	MAE (no unit)	rMAE (%)	RMSE (no unit)	rRMSE (%)	FS (%)
1 hour	-0.011	-3.8	0.059	20.2	0.108	36.8	31.2
2 hours	-0.016	-5.5	0.076	26.0	0.132	45.3	28.0
3 hours	-0.022	-7.5	0.089	30.5	0.152	52.0	25.5
4 hours	-0.024	-8.3	0.010	34.2	0.166	57.0	24.3
5 hours	-0.0352	-12.1	0.109	36.1	0.178	61.1	23.6

Table 2: Performance metrics of the TVL1 algorithm at image level. The average albedo for relative metrics is 0.29.

Forecast horizon	MBE (no unit)	rMBE (%)	MAE (no unit)	rMAE (%)	RMSE (no unit)	rRMSE (%)	FS (%)
1 hour	+0.003	+1.0	0.077	26.5	0.133	45.4	15.2
2 hours	+0.001	+0.3	0.092	31.6	0.155	53.0	15.8
3 hours	-0.005	-1.6	0.104	35.6	0.173	59.2	15.1
4 hours	-0.014	-4.7	0.114	38.9	0.189	64.5	14.2
5 hours	-0.025	-4.6	0.122	41.9	0.204	69.7	12.7

Table 2: Performance metrics of the Farnebäck algorithm at image level. The average albedo for relative metrics is 0.29.

Forecast horizon	MBE (no unit)	rMBE (%)	MAE (no unit)	rMAE (%)	RMSE (no unit)	rRMSE (%)	FS (%)
1 hour	+0.007	+2.3	0.079	27.1	0.135	46.3	13.6
2 hours	+0.006	+2.0	0.100	34.3	0.165	56.5	10.2
3 hours	+0.001	+0.5	0.115	39.3	0.185	63.4	9.2
4 hours	-0.006	-1.9	0.126	43.1	0.201	68.7	8.7
5 hours	-0.012	-4.1	0.134	46.0	0.213	72.9	8.7

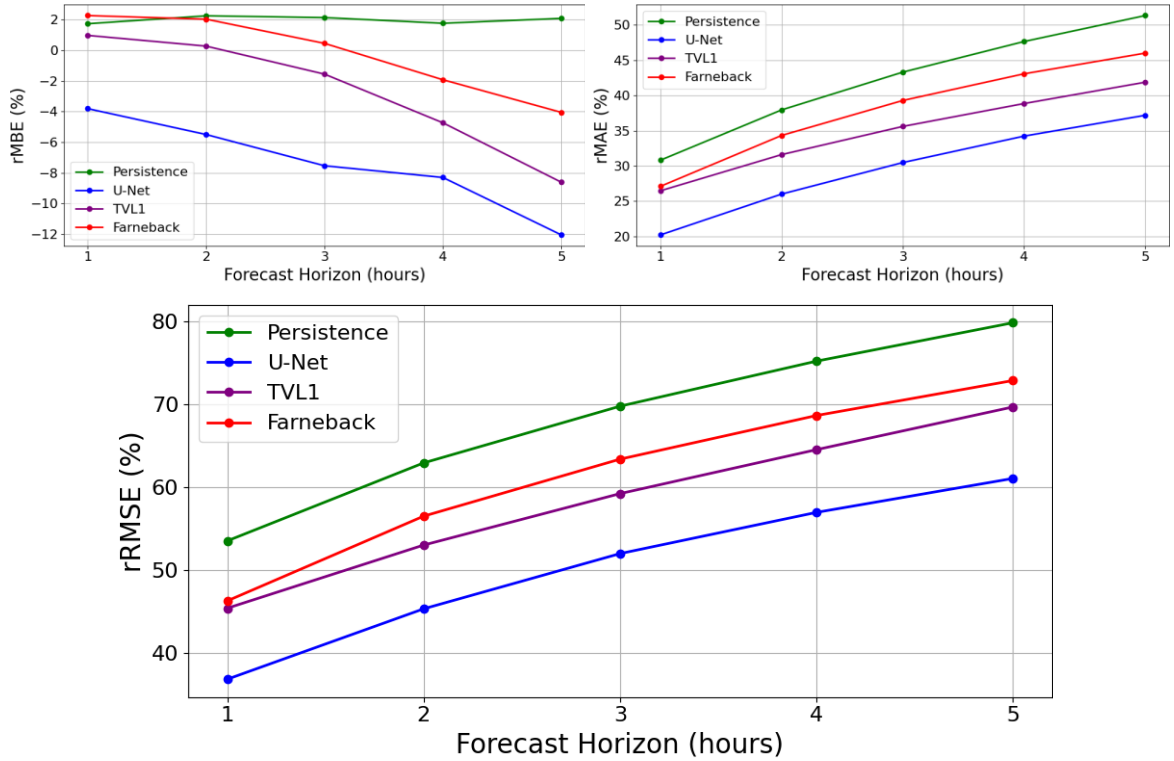

Figure 4: rMBE, rMAE and rRMSE metrics for the image-level evaluation. The normalizing average albedo value is 0.29.

8.2. GHI-Level Complementary Evaluation

Table 2: Performance metrics of the U-Net algorithm at GHI level. The average value for relative metrics is 432 W/m².

Forecast horizon	MBE (W/m²)	rMBE (%)	MAE (W/m²)	rMAE (%)	RMSE (W/m²)	rRMSE (%)	FS (%)
1 hour	+13.5	+3.1	65.5	15.2	102.0	23.6	22.6
2 hours	+19.4	+4.5	78.2	18.1	123.3	28.5	22.9
3 hours	+24.5	+5.7	89.1	20.6	141.6	32.8	22.6
4 hours	+26.7	+6.2	96.1	22.3	151.0	35.0	23.2
5 hours	+34.4	+8.0	101.4	23.5	156.7	26.3	22.6

Table 2: Performance metrics of the TVL1 algorithm at GHI level. The average value for relative metrics is 432 W/m².

Forecast horizon	MBE (W/m²)	rMBE (%)	MAE (W/m²)	rMAE (%)	RMSE (W/m²)	rRMSE (%)	FS (%)
1 hour	+4.2	+1.0	69.8	16.2	107.3	24.9	18.6
2 hours	+4.5	+1.1	84.4	19.5	133.0	30.8	16.9
3 hours	+4.2	+1.0	96.3	22.3	153.4	35.5	16.1
4 hours	+6.8	+1.6	102.1	23.6	162.6	37.7	17.3
5 hours	+8.0	+1.9	104.2	24.1	165.2	38.2	18.4

Table 2: Performance metrics of the Farnebäck algorithm at image level. The average value for relative metrics is 432 W/m².

Forecast horizon	MBE (W/m²)	rMBE (%)	MAE (W/m²)	rMAE (%)	RMSE W/m²)	rRMSE (%)	FS (%)
1 hour	+1.3	+0.3	71.4	16.6	109.2	25.3	17.2
2 hours	+0.4	+0.1	86.9	20.1	134.2	31.1	16.1
3 hours	+1.7	+0.4	98.3	22.8	153.9	35.6	15.8
4 hours	+5.0	+1.2	103.5	24.0	163.8	37.9	16.7
5 hours	+8.5	+2.0	104.2	24.1	165.9	38.4	18.1

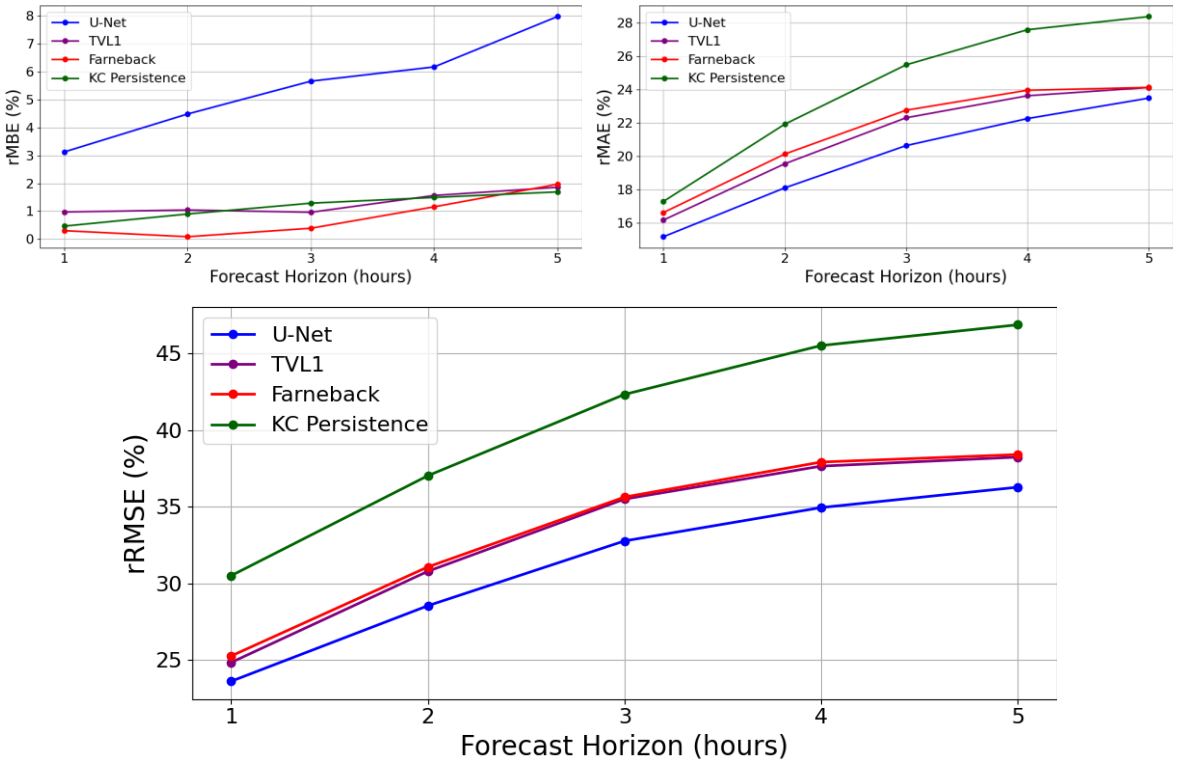

Figure 5: rMBE, rMAE and rRMSE metrics for the GHI-level evaluation. The normalizing average GHI value is 432 W/m².