A CASE STUDY ON DEEP LEARNING FOR PHOTOVOLTAIC POWER FORECASTING COMBINING SATELLITE AND GROUND DATA

Luiz Henrique Buzzi [iD] **, Lucas Weihmann** [iD] **and Pablo Andretta Jaskowiak** [iD]

Joinville Technological Center (CTJ), Federal University of Santa Catarina (UFSC)

R. Dona Francisca, 8300 - Zona Industrial Norte, 89219-600 - Joinville, Santa Catarina, Brazil

luizbuzzi@gmail.com, lucas.weihmann@ufsc.br, pablo.andretta@ufsc.br

Abstract – The increasing demand for clean energy presents challenges in energy supply management, largely due to their intermittency. Photovoltaic power generation, in specific, is greatly affected by weather factors, which may render power grids susceptible to instability, quality and balance issues. In this context, photovoltaic power generation forecasting is crucial not only to enhance the management of diverse energy sources through generation planning, but also to ensure widespread adoption of photovoltaic energy. To address the predictability issue in generation, this study aims to investigate the combination of satellite data with meteorological data to predict the energy generation potential in photovoltaic panels within 30, 60, 120, and 180-minute horizons. For this purpose, images from the GOES-16 satellite are used in combination with data from a ground-based weather station, located at Florianópolis - Santa Catarina - Brazil. The data is fed to a convolutional neural network, where convolutions are employed to extract features from the satellite images, aiming to establish a relationship with solar irradiation. The output of the convolutional network serves as input for a multilayer perceptron network, which utilizes the data to predict the Global Horizontal Irradiance (GHI). Our results support that models incorporating satellite images provide forecasts approximately 41% better for the 30-minute horizon and 21% better for the 180-minute horizon, when compared to models without satellite images.

Keywords – Photovoltaic power forecasting, deep learning, convolutional neural network, satellite image, GOES-16.

1 INTRODUCTION

A substantial transition in the world's energy sources is needed to combat climate change. Its consequences and impacts, including escalating global temperatures, intensified extreme weather events, and the rapid melting of polar ice caps, have emphasized the urgency of moving from fossil fuels towards cleaner, renewable alternatives. Given the relevance of the subject, research and studies in the area lead to a clear differentiation between non-renewable and renewable energy sources, with extensive understanding of their respective advantages, limitations, and challenges, particularly for the latter.

A complete replacement of current energy sources by renewables is utopian, even in the long term. However, according to the Intergovernmental Panel on Climate Change (IPCC) [1], the integration of renewable energy into the grid is crucial for achieving climate change mitigation goals. Furthermore, research by Jacobson et al. [2] highlights the potential of renewable energy to not only decarbonize the energy sector, mitigate greenhouse gas emissions and reduce dependence on finite fossil fuel resources, but also to foster economic development and improve energy security. This trend underscores the increasing prominence of renewable energy sources, with solar energy emerging as a promising alternative, due to its substantially lower carbon emissions (in comparison to other energy sources) and inherently sustainability in the long run.

Photovoltaic (PV) cells are one of the main technologies used to generate electricity from solar energy. The power generation of photovoltaic cell panels is intrinsically related to solar irradiance [3]. According to Pedro et al. [4] both Direct Normal Irradiance (DHI) and Global Horizontal Irradiance (GHI) can be used to estimate energy generation in photovoltaic panels. Solar irradiance depends on various meteorological factors such as temperature, cloud cover, time of day and season, atmospheric aerosols, altitude and latitude. Therefore, the potential for electrical generation in PV panels at any given time depends on these factors [5, 6]. Factors, such as cloud cover, can vary significantly in the space of a few minutes and, in an unpredictable way, operators encounter challenges in managing electricity grids due to intermittencies and fluctuations in PV power generation [7].

In light of these circumstances, an accurate irradiance forecast allows for a more efficient integration of solar energy generated through PV panels into the energy grid, facilitating the process of managing hybrid energy systems. It is noteworthy that forecasting horizons may significantly vary depending on specific operational requirements, commonly categorized into intra-hour, intra-day, and day-ahead scenarios [6]. Intra-hour forecasts are relevant for dispatching, regulatory and load following purposes. However, in the context of power system operators that handle multiple load zones, the intra-day, especially considering the 1h to 6h ahead forecasting horizons, are of critical relevance/importance [8].

The successful application of a number of methods, including statistical, time series analysis, machine learning and deep learning, has been observed in the context of predicting Photovoltaic (PV) power generation [5]. Among the aforementioned approaches, those employing machine learning and deep learning have demonstrated the most promising results in recent studies [9]. In the field of deep learning, the notable achievements in image processing tasks provide promising avenues for integrating satellite imagery into photovoltaic power generation prediction models [10].

In this context, our study investigates the integration of machine learning/deep learning models with satellite imagery and ground measured data for the forecasting of photovoltaic power generation at intra-hour and intra-day horizons (30, 60, 120

and 180 minutes ahead). This investigation is conducted through a case study focusing on a photovoltaic plant situated in Florianópolis, Santa Catarina, Brazil. Satellite images utilized in the study are sourced from the ABI Imager of the GOES-16 satellite, featuring a 16-band state-of-the-art spectral radiometer [11]. In this work, we significantly expand our previous study [12], which considered solely satellite data for such forecasts. Indeed, to the best of our knowledge, our works are the first ones to employ data from the GOES-16 satellite for photovoltaic power forecasting within South Brazil. Finally, it is important to emphasize that no cloud identification or movement is explicitly extracted from satellite images, as in the work performed by Si et al. [13]. To that end, the models must extract useful features from images by themselves, implicitly.

The structure of the paper is outlined as follows: Section 2 provides a brief overview of related work on photovoltaic power generation forecasting, with a specific emphasis on satellite image-based approaches. Section 3 presents information on the photovoltaic plant considered in the study and the datasets of both ground variables and satellite images available. Section 4 presents the deep learning architectures employed and the experimental setup. The results obtained from the experiments carried out are presented in Section 5. Finally, Section 6 presents a summary of the principal findings of this work and offers insights into future research directions.

2 RELATED WORK

As illustrated in the preceding section, the prediction horizons are inherently related to the operational aspects of the energy management system, with particular consideration given to the integration of PV energy into the grid. Although literature offers certain categorizations of forecasting horizons, consensus regarding terminology remains elusive. According [6], the 30 and 60 minutes ahead can be defined as intra-hour forecast and the 120 and 180 minutes ahead forecasting horizons can by defined as intra-day forecasts. In order to facilitate comprehension and following the definition presented by Ahmed [5], the 30, 60, 120 and 180 minutes ahead forecasting horizons under scrutiny in this study will be referred as short-term forecasts.

The literature review by Kumar et al. [14] summarises the main data sources used in solar irradiance forecasting, considering the prediction horizon as a reference base. For prediction horizons of 10 minutes to one hour, they indicate that the use of ground measurements and cloud images are the most appropriate. For predictions between 1h and 5h, satellite images are the most recommended. And for prediction horizons above 5 hours, the highest accuracy rate is obtained using numerical weather prediction (NWP) models. Sobri et al. [7] also identified satellite imagery as a particularly suitable data resource for short-term photovoltaic (PV) power forecasting.

The number of articles published in the field of PV generation forecasting is significant. The brief review presented in this paper seeks to focus on works dealing with short-term horizon forecasting based on satellite images or hybrid techniques, using deep learning methods. For a more comprehensive understanding of the application of different data sources, prediction horizons and machine learning methods in PV power generation, we recommend checking out the works carried out by [5–7, 15].

A method for integrating meteorological factors and satellite imagery for estimating Direct Normal Irradiance (DNI), Direct Horizontal Irradiance (DHI), and GHI is presented by Eissa et al. [16]. This integration is achieved by employing distict network according to weather conditions. The authors choose to divide the problem into two parts: the estimation of direct normal incidence (DNI) and direct horizontal incidence (DHI). Both networks use the $10.8\mu m$ and $12\mu m$ satellite spectral channels to create a cloud mask that divides images into cloudless and cloudy states. The program then uses the results to dynamically choose the right network and adds other sources of information, such as zenith angle and supplementary channels. Consequently, four distinct models are trained: two for DNI estimation (one for cloudy and one for cloudless conditions) and two for DHI estimation (similarly, one for cloudy and one for cloudless conditions). The fusion of DNI and DHI results via analytical means estimates the Global Horizontal Irradiance (GHI), yielding a relative Root Mean Square Error (rRMSE) of approximately 12% for GHI predictions.

Unlike other approaches, the work of Wang et al. [17] integrates satellite images with PV power generation data from a neighbouring plant to forecast the PV power generation of the target plant. Support Vector Machines (SVM) and Gradient Boosting Decision Trees (GBDT) are then used to verify the effectiveness of the proposed approach. Using linear extrapolation, cloud movement over the plants is predicted using satellite images. Data from two real plants are employed as a case study, with forecast horizon ranging from 15 minutes to 4 hours. The authors found that the proposed method outperformed existing benchmarks in the literature, based on Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) metrics.

Cheng et al. [18] highlight a key distinction between their study and the majority of literature in the field: they employ satellite images directly to predict PV energy generation, whereas most other studies derive their irradiance or cloud cover predictions from pre-processed satellite images. A method based on autoencoders is used to integrate satellite images and dynamically define a region of interest for prediction horizons of up to 3 hours. The comparative results presented by the authors indicate that the proposed method is superior to smart persistence and optical flow extrapolation approaches and can be used to improve the performance of deep learning models. However, the authors comment that the large size of the satellite data processed imposes computational limitations on the method in the training phase, which represents a challenge in the application of this method.

In order to estimate the Global Horizontal Irradiance (GHI) based on satellite data that captures cloud behavior in the visible spectrum, Si et al. [13] suggested using a Convolutional Neural Network (CNN). Pre-processing is applied to images taken at three different times, t, $t-1$, and $t-2$, where t is the current time, in order to normalize the data and adjust for differences in zenith angles. After that, the pre-processed images serve as input into a CNN for feature extraction. This new features are then fed into a Multilayer Perceptron (MLP) and merged with weather station data for GHI prediction. The authors stress the need of image pre-processing and demonstrate how well their method works to reliably identify cloud movements based on wind

speed forecasts.

Marquez et al. [19] introduced a method for the prediction of PV power generation using satellite images in the visible and infrared spectrum. Their approach uses visible spectrum imagery to produce a cloud cover index, and then applies normalization and zenith angle correction to account for ground albedo. Using the Particle Image Velocimetry (PIV) technique, infrared pictures are used to obtain a vector field that describes the direction and velocity of clouds. After integration, cloud velocity and cover data are fed into an artificial neural network (ANN) to estimate GHI at different forecasting horizons.

The study conducted by Pelisson et al. [20], while exclusively relying on ground data from a solarimetric station and PV plant, is pertinent to our investigation due to its reliance on the same ground meteorological data employed in our work. The authors explore various machine learning techniques and feature space configurations for short-term PV power forecasting, observing optimal performance with Multilayer Perceptron (MLP) and Support Vector Regression (SVR). Notably, the optimal feature space encompasses a combination of Global Horizontal Irradiance (GHI), Direct Normal Irradiance (DNI), Direct Horizontal Irradiance (DHI), Ultraviolet (UV) radiation, Relative Humidity (RH), Ambient Temperature (AT), and Wind Speed (WS). Despite the robust correlation observed between GHI and PV power output, the authors highlight that the inclusion of additional variables in the model resulted in a remarkable reduction of nearly 30% in Root Mean Square Error (RMSE). This observation underscores the potential for further enhancement in forecasting accuracy through the incorporation of alternative techniques, including the integration of satellite imagery.

Even though related to the works just reviewed, our work differs in some aspects. First, we consider both satellite and ground data, whereas most works focus mostly on one data source alone in order to build models. Regarding satellite data, we employ data from GOES-16 by itself, without zenith angle correction, cloud coverage estimation and/or cloud movement forecasting, as in some of the works reviewed. That is, the models we build and evaluate in this work are trained to implicitly extract information from satellite images, whereas in other works, this is done explicitly. Although Pelisson et al. [20] considered the same location, only ground data were considered.

3 CASE STUDY

Our case study can be briefly summarized as the forecasting of Global Horizintal Irradiance (GHI) for a photovoltaic power generation facility located at Florianópolis — Santa Catarina — Brazil. The location is shown in Figure 1. Forecasts consider a two year time frame, for the years 2018 and 2019. Both ground and satellite data are considered and evaluated for building models. In this section details of the case study are presented. First, we describe the dataset employed in the work, which comprises two sources: ground and satellite data. Ground data, which is described in Section 3.1, was obtained from a solarimetric station located at Laboratório Fotovoltaica — Florianópolis — Santa Catarina — Brazil. This site also houses the photovoltaic plant, which is the focus of our forecasting efforts. Section 3.2, provides details regarding the satellite source, which was obtained from the GOES-16 satellite. This section also presents a spatial analysis to define the area of interest for the collected imagery. Section 3.3 discusses forecasting horizons, data fusion, and granularity.

3.1 Ground Data

The ground station database was obtained from Laboratório Fotovoltaica[1] (UFSC Solar Energy Research Laboratory), comprising weather related data and power generation data of the photovoltaic plant. The data was obtained with a frequency of 1min with a Campbell CR6 datalogger [20]. Besides Global Horizontal Irradiance (GHI), which is our forecasting target, other variables are collected from the solarimetric station, as presented by Table 1. For more information regarding ground data, the reader may refer to Pelisson et al. [20]. At this point, it is important to discuss our selection of the forecasting variable (target variable), namely, GHI. Even though we could adopt the actual power output of the photovoltaic plant as our forecasting variable, such values would be completely tied to the actual configuration of the plant. The addition or removal of photovoltaic panels, for instance, would require new forecasting models. In order to obtain plant agnostic forecasts, we selected GHI as our forecasting variable. Among the three available irradiance components (GHI, DNI, and DHI), GHI is the one with the highest value of pearson correlation w.r.t. photovoltaic power generation, i.e., 0.97. This gives an advantage to the model, which can be applied to different generating units in the same park through a relationship between the GHI and the generation potential.

Before actually using the database, pre-processing work had to be done. First, the 2018 and 2019 data were unified, removing eventual duplicated readings. For convenience, the timestamps from the data were converted to GMT-0, which is the same time zone used for the satellite. After the removal of some inconsistent data points (negative values and values above the maximum observable threshold), the database was left with 871,709 time points comprising the years 2018 and 2019. After an analysis of the average distribution of GHI over the day, we delimited our forecast between 9:00am and 5:30pm. Considering such a forecasting interval and the fact that all the models considered for analysis in this work will output estimates for GHI, the final ground dataset was reduced to a total of 218,549 readings (from an initial value of 871,709) with 1min granularity.

3.2 Satellite Data

The satellite images employed in this work are obtained from one of the satellites of the GOES series. Given the location of the photovoltaic power station considered in this work, the GOES-16 satellite is of our interest. The satellite data were obtained
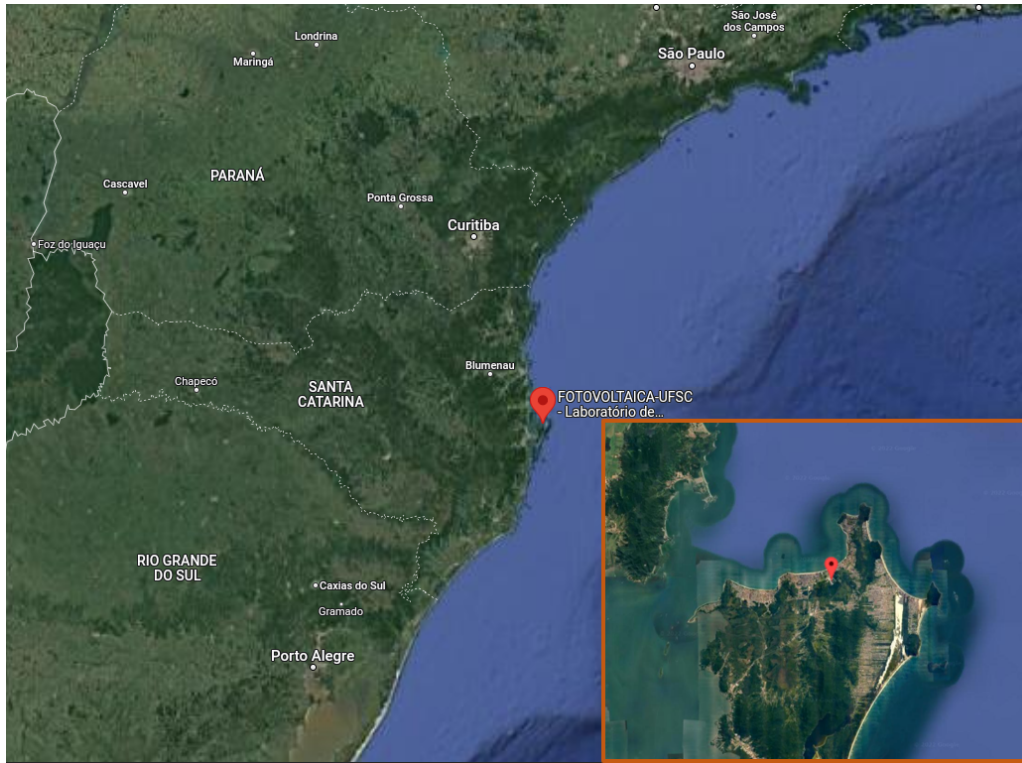
---

[1]https://fotovoltaica.ufsc.br/

Figure 1: Location of the photovoltaic power plant w.r.t. south Brazil. In the zoomed area (bottom right) a crop of $10°$ latitude and $10°$ longitude centered at the site is presented. The satellite images employed in this were cropped to this very same area.

Table 1: Data from the UFSC laboratory and photovoltaic plant.

Parameter	Acronym	Sensor	Unit
Direct Normal Irradiance	DNI	SHP1	W/m^2
Diffuse Horizontal Irradiance	DHI	SMP11	W/m^2
Ultraviolet Index	UV	CUV 5	W/m^2
Ambient temperature	AT	HMP155	^{0}C
Wind speed	WS	WindSonic1	m/s
Relative humidity	RH	HMP155	%
Global Horizontal Irradiance	GHI	SMP22	W/m^2

in the same interval as the ground station data, for the years 2018 and 2019. Initially, the images were obtained in full disk format, covering the entire capture area of GOES-16. For such a reason, the satellite data also went through several pre-processing stages.

The GOES-16 satellite has 16 spectral bands, with 6 in the visible range and 12 in the non-visible range. As our goal is to relate satellite images to GHI (and consequently photovoltaic power generation), the ABI imager's Channel 2 is a reasonable option. Besides being recommended for this application, this is the spectral channel with the highest spatial resolution, 0.5 km/pixel, allowing for a more detailed evaluation of our point of interest. Images from this channel provide reflectance data, which can be used to estimate the amount of irradiance at the ground level (a region covered with clouds, for example, usually exhibits a higher reflectance than the ground). A point that deserves particular attention concerns changes in the satellite's operational mode after its launch. Between the years 2018 and 2019, there was a shift in the operational mode of the GOES-16 satellite from Mode M3 to Mode M6, altering the data capture frequency. In Mode M3, the satellite executed a routine that captured an image every 15 minutes, whereas in Mode M6, images are captured every 10 minutes. Thus, for the sake of data compatibility, the minimum capture frequency we can work with for both years is 30 minutes, which we adopted.

Each raw image from the ABI imager's Channel 2 is approximately 500MB in size. To reduce the data volume and focus the analysis, the interval from 8:00 AM to 6:00 PM GMT-3 was defined, as discussed in the previous section. To facilitate image processing, the database was obtained within a window of $10°$ latitude and $10°$ longitude, centered on the point of interest at coordinates -27.4310° latitude and -48.4414° longitude. After this crop, the images cover a region centered at the location of interest, as show in the right bottom of Figure 1. It is worth noticing that even after the cropping, each image has a resolution of 2124×875 pixels, meaning that, if provided without any pre-processing to the models, these would have to deal with roughly 1.8 million input features. Based on the work of Chow et al. 21, we assumed that the types of clouds with the greatest influence to

the observed GHI woudld have a speed of no more than 40km/h. Considering that the satellite's spatial resolution is 0.5km/pixel and the granularity of the data we work with is 30 minutes, we estimated that images with 40×40 pixels (an area of 20×20 km) would provide enough information to the models and, at the same time, reduce the computational time needed to training.

Therefore, we were able to reduce the size of the satellite images, to the zoomed area shown in Figure 2. It is worth noticing that Figure 2(a) was obtained from Google Maps. Actual satellite images from GOES-16 are depicted in Figure 2(b) and Figure 2(c). These are the actual satellite images employed in our work. Having made such considerations, the final pre-processed dataset of satellite images consisted of a total of 14,329 images, with 7,136 images for the year 2018 and 7,193 images for the year 2019.

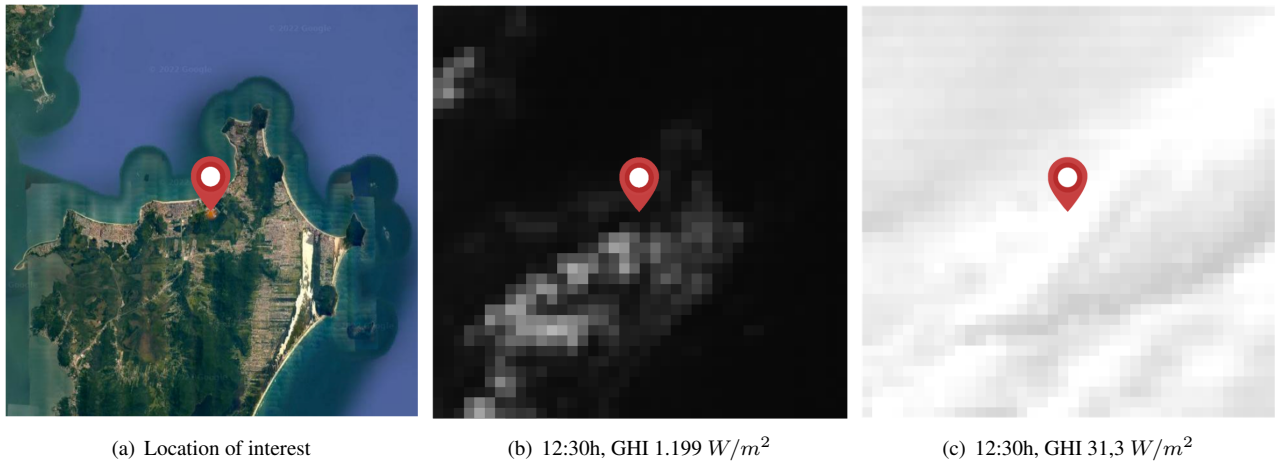(a) Location of interest   (b) 12:30h, GHI 1.199 W/m^2 (c) 12:30h, GHI 31,3 W/m^2

Figure 2: Image (a) is a representation of the location of interest, obtained from Google Maps 22. Images (b) and (c) were obtained from the GOES-16 satellite, channel 2, for two different days, at same time (12:30h). These are actual images we employ in our work. Note that lower/higher reflectances (darker/brighter images) translate in higher/lower GHI values.

3.3 Data Fusion and Forecasting Horizon

The two data sources previously described were obtained with different granularities (resolutions). For ground data, measurements were taken at every minute. For satellite data, measurements were obtained every 15 or 10 minutes, depending on the mode of operation for the GOES-16 satellite, as discussed in the previous section. In order to perform data fusion and obtain a consistent dataset comprising both data sources, both data sources went through additional pre-processing steps.

As discussed earlier, between the years 2018 and 2019, the GOES-16 satellite changed from Mode M3 to Mode M6, altering the data capture frequency from 15 minutes to 10 minutes. For consistency, we defined a granularity/frequency of 30 minutes for all of our data. Therefore, intermediate readings from this data source were discarded, namely: 15 minutes and 45 minutes for Mode M3, and 10 minutes, 20 minutes, 40 minutes, and 50 minutes for Mode M6. For the ground data, we performed a downsampling, by averaging values of readings obtained every minute to a the same 30 minute resolution as the satellite data.

After these steps, the final dataset comprising the years 2018 and 2019 has a total of 14,329 objects, with a 30 minute resolution. Such a resolution also provides the minimum horizon and resolution of GHI forecasts. Given our data resolution and the fact that, as discussed by Sobri et al. [7], satellite data is appropriate for intra-hour and intra-day forecasts, we build and evaluate models for forecasting GHI considering 30, 60, 120, and 180 minutes ahead.

4 Methods and Experimental Setup

This section describes in detail the deep learning approaches employed in our evaluation and the entire experimental setup. In section 4.1 we discuss the deep learning architectures, the configuration of the convolutional layers, and the approach we adopted in order to integrate ground and satellite data as input to the models. It is worth noticing that the last layer of all approaches consists of a Multilayer Perceptron (MLP), which aims to relate all input data to the output, thus producing a GHI forecast. ReLU was defined as the activation function of the MLP, while the dropout rate, the number of layers, and neurons were limited based on a preliminary study

In Section 4.2, the whole experimental setup is discussed, including training, validation, and testing. This section also discusses the optimizer, validation functions, decay, and the framework used for network programming.

4.1 Forecasting Architectures

We evaluate three distinct forecasting architectures in our work. These architectures difference themselves with respect to the input data that is provided to each one of them. Figure 3 provides an overview of the three forecasting architectures, which

are obtained on the basis of two blocks of input data, namely, satellite data (Block A) and ground data (Block B). The three forecasting architectures are obtained with the following configurations : (i) satellite data (Block A); (ii) ground data (Block B); and (iii) ground and satellite data combined (Block A + Block B). Further details on each one are provided in the following.
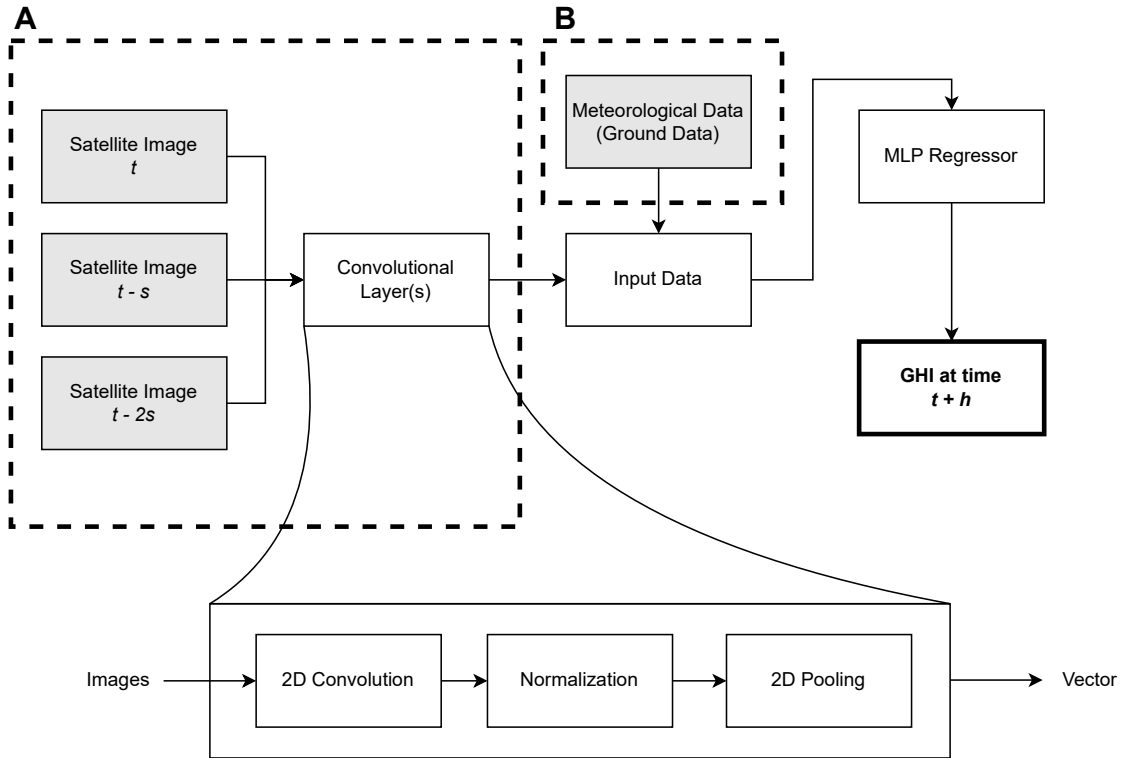

Figure 3: Building blocks that comprise the different ANN architectures employed in this work.

4.1.1 Forecasting Based on Satellite Images

The first architecture comprises solely input data from Block A, that is, satellite data. In order to define the configuration and number of satellite images provided as input, we follow the work of Si et. al [13], in which the authors showed that, for a given location, it is possible to estimate ground GHI values with a Convolutional Neural Network (CNN) and three input satellite images from consecutive time instants. More specifically, given a current time t, our aim is to provide a forecast for a future time instant, given by $t + h$, with h defining the forecasting horizon. For that, besides the satellite image for time instant t, we also consider two additional past satellite images, obtained at times $t - s$, $t - 2s$, where s defines the data granularity.

Given that we take into account three images in order to estimate GHI, each one with 40×40 pixels, there are 4,800 inputs (1,600 from each image). Providing this number of inputs directly to the MLP would be computational expensive. Thus, we employ a convolutional layer, in order to extract the main features from the three images and reduce their combined dimensionality. In such a way, we expected to improve the performance of the MLP network. As depicted in the diagram in Figure 3 (bottom), each convolutional layer consists of: (i) a 2D convolution, (ii) a normalization layer, and (iii) a 2D pooling layer. In the first layer of the convolutional set, the ReLU activation was applied. After normalization, a 2D average pooling was employed, with a step of 2x2. This topology was adapted from the work of Si et al. [13].

For this particular architecture, the set of parameters presented below was considered. Again, the selection of parameters considered in our experiments within the convolutional layer were based on the work of Si et al. [13]. Preliminary experiments also indicated that these provided a good compromise between forecasting quality and training time. It is important to keep in mind that input images have only 40×40 pixels, limiting the search space due to its small size. In this case, the output of the convolution was just flattened before being provided as input to the MLP, since a dense layer on the output was not necessary.

Convolutional Layer(s)			Multilayer Perceptron Regressor		
# Layers	Filter Size	Kernel Size	Layers	Dropout	Neurons
1, 2	16, 32	3,5	1, 2	0.15	64, 128, 256

With regard to this forecasting topology, a total of 48 forecasting models were considered for evaluation, employing a combination of the presented parameters. All models based solely on satellite images will be identified by `SatImg` in results, followed by their respective parameters, when applicable or necessary.

4.1.2 Forecasting Based on Ground Data

The second architecture consists of data from Block B (see Figure 3), that is, ground data from the solarimetric station, and temporal/calendar data. The features measured at ground were already described in Section 3.1, Table 1. Regarding temporal/calendar data, we consider the day of the year and the minute of the day. Note that these two temporal/calendar features help to incorporate variations in solar irradiance observed due to the season of the year and the hour of the day. It is clear that some parameters/features, such as UV incidence, have greater impact on photovoltaic power output than others. In order to verify if forecasts based on UV can be enhanced or even surpassed by different feature sets, we evaluate several combinations of the available features. In brief, we obtain four feature subsets, each representing a combination of ground station data, see Table 2.

Table 2: Table with details for the four different subsets of features considered for ground data.

Feature Subset Name	Features
UV + CAL	UV Irradiation and Calendar Data
UV + AT + RH + WS + CAL	UV Irradiation, Air Temperature, Relative Humidity, Wind Speed, and Calendar ata
AT + RH + WS + CAL	Air Temperature, Relative Humidity, Wind Speed, and Calendar Data
AT + RH + CAL	Air Temperature, Relative Humidity, and Calendar Data

It is worth noticing that in this case, parameters related to convolutional layers are not present given that the forecasting topology does not require satellite images. Therefore, convolution is not present for ground data.

For each one of these feature subsets, we considered the very same parameters employed for the MLP network as described in the case of satellite data. Consequently, a total of six distinct model configuration, varying in terms of the number of layers and the number of neurons per layers, were evaluated for each of the four feature subsets. This resulted in a total of 24 forecasting models.

Models based solely on ground data will be identified by the feature subset name defined in Table 2, followed by the respective parameters, when necessary for further details.

4.1.3 Forecasting Based on Ground and Satellite Data

In the latter case, a strategy for combining the data was devised. Following the same data combination proposal as Marquez et al. 19, the resulting vector from satellite images, as obtained from the final convolution and the meteorological station data are concatenated to form a single data vector, which is then provided as input to the MLP. A first concern in this case, was whether there would be any bias in favor of satellite data in the forecasting procedure, given the larger number of features, in comparison to that of meteorological data. To address this concern, besides feeding the output from the convolution "as is" to the MLP, we also considered/evaluated a topology with a dense layer after the convolution, in which a reduced set of 16 features was extracted.

Note that this forecasting architecture arises from the combination of Blocks A and B from Figure 3, for which further details were provided in the previous sections. Indeed, it is worth noticing that for the case of ground data, different feature subset configurations are possible. In this case, we adopted the same four sets as already described by Table 2, with the addition of features obtained from the satellite images. This allows for a fair comparison between the two cases. For this particular architecture, in which satellite and ground data were combined, the set of parameters presented below was considered.

Convolutional Layer(s)			Dense	Multilayer Perceptron		
# Layers	Filter Size	Kernel Size	None, 16	Layers	Dropout	Neurons
1, 2	16, 32	3,5		1, 2	0.15	64, 128, 256

Regarding this forecasting topology, a total of 96 different model configurations were considered for each of the four distinct feature subsets under evaluation, that is, a total of 384 forecasting models. Models that combine satellite images and ground data will be referred to as `SatImg + Ground Features` in the results, e.g., `SatImg + UV + Cal`.

4.2 Experimental Setup

As previously discussed, a granularity of 30 minutes and forecasts considering 30, 60, 120, and 180 minutes ahead (henceforth referred to as h — horizon) were considered. A forecasting window was defined between 9:00 AM and 5:30 PM. In brief, the models we build and evaluate work as regressors, for which, given a set of descriptive features (ground and/or satellite data) a forecast is produced (estimated GHI). Therefore, given a current time instant t, and a desired forecast $t + h$, for forecasting models based on: (i) satellite data alone, images from instants t, $t - 30$, and $t - 60$ are provided as input; (ii) ground data alone, meteorological features from time instant t are provided as input; and (iii) combined satellite and ground data, images from time instants t, $t - 30$, and $t - 60$, within meteorological features from time instant t are provided as input. Recall that, in the case of meteorological features alone, four distinct subsets are evaluated — as we already discussed within Table 2.

The dataset was split into 3 sets for evaluation of the models, simulating a real application scenario, as we discuss in the following. The first one is the training set, which is composed of 70% of the 2018 data. It is worth noticing that this subset was balanced homogeneously across months, which means that it includes 70% of data from January, 70% from February, and so forth. Note that, in this way, we guarantee that the training set contains data from the different seasons of the year and also a considerable weather variability. The second set, comprised the remaining 30% of the 2018 data were reserved for validation and hyperparameter selection, that is, the determination of the best parameters for the forecasting models when considering distinct feature inputs. Finally, the third set, accounts for the test data, which was not seen during training or hyperparameter estimation and, therefore, can be employed to estimate the quality of the models as one would expect to obtain in a real application scenario. This set is composed of all data points from the year 2019. All data is available under request at Zenodo [2].

All codes for the evaluation were developed in Python, using the TensorFlow framework [23] with the Keras API [24]. Training of the MLP models was performed using Adam optimization with exponential decay. The learning rate parameter was set to 0.001, and a decay rate of $1e^{-7}$. The loss function used was Mean Squared Error (MSE). All neural networks were trained up to a limit of 10000 epochs, but with early stopping implemented, if persisting for 500 epochs. The batch size was set to 512.

To identify the best set of parameters, for each forecasting horizon a total of 456 different model configurations were evaluated, i.e., 48 models based only on satellite data; 24 models based only on ground data; and 384 models for satellite and ground data. All evaluations considered the Root Mean Square Error (RMSE) as the evaluation metric, which is defined in Equation 1.

$$\text{RMSE}(y, \hat{y}) = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{1}$$

In the equation above, $\hat{y}$ accounts for the forecasts, as provided by a given model, y account for the desired output (ground truth), and n accounts for the total number of forecasts. A perfect model would yield a RMSE equal to zero — it is a minimization metric.

5 Results and Discussion

In order to gain some insight about the problem, we analyzed how satellite images correlated with GHI values measured at ground. For that, we considered all data points from the dataset within the years of 2018 and 2019. Given that each pixel in the satellite image accounts for 0.5km in height and width, for the solarimetric station, the location of interest, we considered the average of four pixels centered at it (an area of 1 km^2). At this point, it is important to recall that the satellite images employed in this work provide information related with reflectance, therefore, one would expect that the lower/higher the reflectance the higher/lower the GHI, which translates in a negative correlation between both variables. Indeed, that is the case observed with such an analysis, with a Pearson correlation of -0.34, for the area centered at the soalrimetric station. Even though the negative correlation follows the expected pattern, it is clear that reflectance data alone has a low linear correlation with GHI. This indicates that: (i) taking into account solely the reflectance of the pixels within the location of interest might not be enough and; (ii) different data sources might be necessary to enhance forecasts based on satellite data. We also correlated each one of the pixels of the cropped area (which will be later be input of the models) with the GHI (note that for each time instant we only have a single GHI value, as measured at the solarimetric station), resulting in Figure 4. From the figure, one can observe that, in absolute terms, correlation values obtained for the solarimetric station are amongst the highest ones.
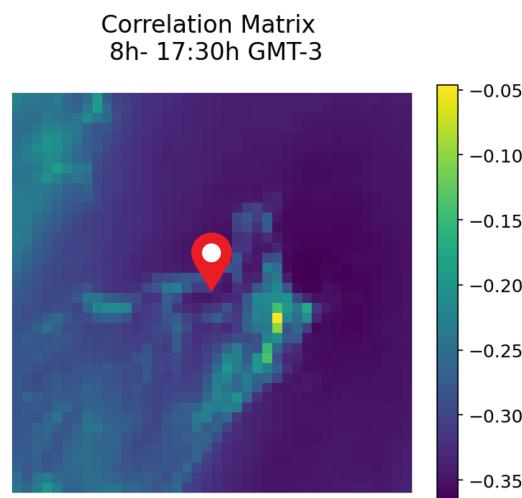

Figure 4: Image depicts in each pixel the average correlation between each pixel of the satellite image (these account for reflectance) with respect to the GHI obtained by the ground solarimetric station, considering all data points from 2018 and 2019.

It is important to recall that before being provided as input to the forecasting models, satellite images are subjected only to cropping, in order to reduce their size. This is worth mentioning, given that, according to several studies (see, for instance, Si et al. 13) the solar zenith angle has a considerable influence in the estimation of GHI, when satellite images are employed. In brief, the solar zenith angle is the angle observed between the incidence of the sun's rays and a vertical line at that point. Given that we do not perform any zenith angle correction, we investigate its impacts in the observed correlations between irradiance (satellite images) and GHI (measured at ground). Such an investigation is performed only for the four pixels from the satellite image centered at the solarimetric station, where GHI is measured. This result is shown in Figure 5, for which each point accounts for the average correlation between irradiance and GHI considering data from 2018 and 2019. It is evident that the sun's position relative to the plant strongly influences generation data, and the relationship is stronger around noon, when the solar zenith angle is minimized, as the sun is positioned directly above the plant. Note that the y axis in the plot is reversed around zero, given that we place emphasis in negative correlations, which are the expected in this particular application scenario.
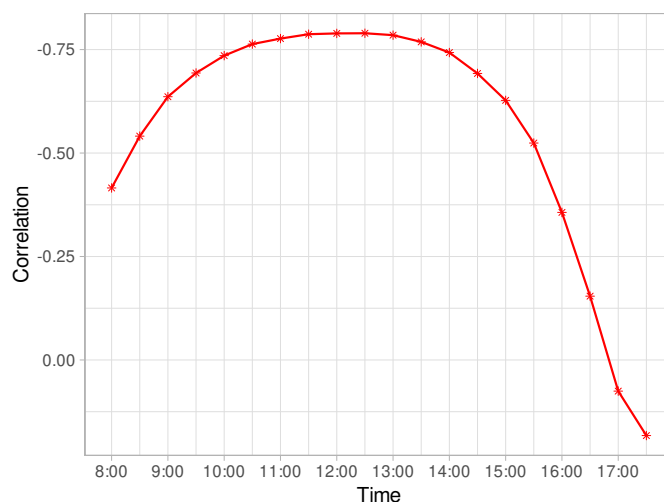

Figure 5: Average correlation considering the years of 2018 and 2019 for reflectance considering the four pixels centered at the solarimetric station (satellite images) and the GHI values measured at the site. Points are plotted with a 30 minute resolution.

From these analyses we can conclude that there is correlation between irradiance, as estimated by satellite images, and photovoltaic power generation. This is indeed verifiable on the basis of a strong correlation between GHI and photovoltaic power output. Even though only weak to moderate negative correlations were observed between reflectance estimated from satellite images and GHI, these account only for linear relationships, and can hide more complex relationships. Moreover, given the great impact of the solar zenith angle in such a correlation, satellite images will be fused with data from different sources, in order to account for such factors and obtain better forecasts.

For the evaluation of the forecasting models we consider the observed RMSE regarding the test set, for the four forecasting horizons, that is, 30, 60, 120, and 180 minutes. As we discussed previously when presenting each of the data sources, in particular for the case of satellite data, we did not perform any correction and/or adjustment in order to account for the zenithal angle. This, in turn, might influence the quality of the forecasts made by the models. The addition of calendar data, that is, the day of the year and minute of the day, was conceived as an attempt to circumvent such a factor. In order to assess whether such an addition was effective or not, we depict in Figure 6 boxplots for RMSE regarding *all* models under evaluation that considered only satellite images and satellite images in addition with calendar data, respectively. We can observe that for all configurations in all time horizons, the inclusion of the day of the year and minutes of the day significantly improved the result (lower RMSE values), with percentual improvements of 29.7%, 36.4%, 36.6%, and 36.6% for the horizons of 30, 60, 120, and 180 minutes respectively. Such improvements are computed considering the models after hyperparameter tuning, that is, the best parameter combination in each scenario. It is clear that by adding calendar data better forecasts are obtained, in general. In view of these results, we proceed with a detailed analysis considering only forecasting models that take this particular information into account.

We will split subsequent analysis in two parts. The first one contrasts the results obtained with satellite image data with those of satellite and ground data, without considering the UV feature (Ultraviolet Index). The second part compares the same data sources, in this case, however, taking the UV feature, as obtained from ground data into consideration.

The first part of the analysis is summarized by the boxplots depicted in Figure 7. Note that for all forecasting horizons, the forecasting models that take as input satellite images (namely `SatImg + CAL`, `SatImg + AT + RH + WS + CAL`, and `SatImg + AT + RH + CAL`) present, overall, similar results in terms of RMSE distribution. In fact, there is only a slight decrease in the median RMSE (this is also true for the overall boxplot distribution) when meteorological features are added to the satellite images. This observation corroborates with our assumption that calendar data is an important aspect to the models, aiding their mapping when forecasting GHI values. Furthermore, these results indicate that the addition of ground station data did not provided significant improvements. Exception can be observed for the forecasting horizons of 60 and 120 minutes, in which models that take only satellite images and calendar data provided slightly superior results w.r.t. the others. From these
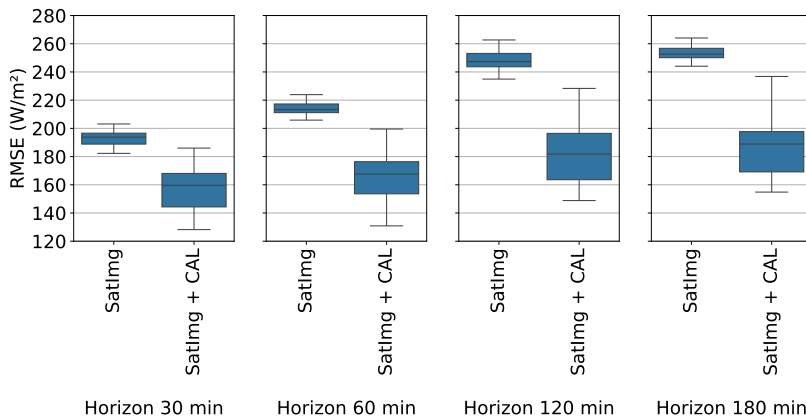
Figure 6: Distribution of RMSE results for all models trained in the presented configurations, in 2019.

plots, one can also observe that, as expected, the RMSE of models tend increase as the forecasting horizon is expanded.
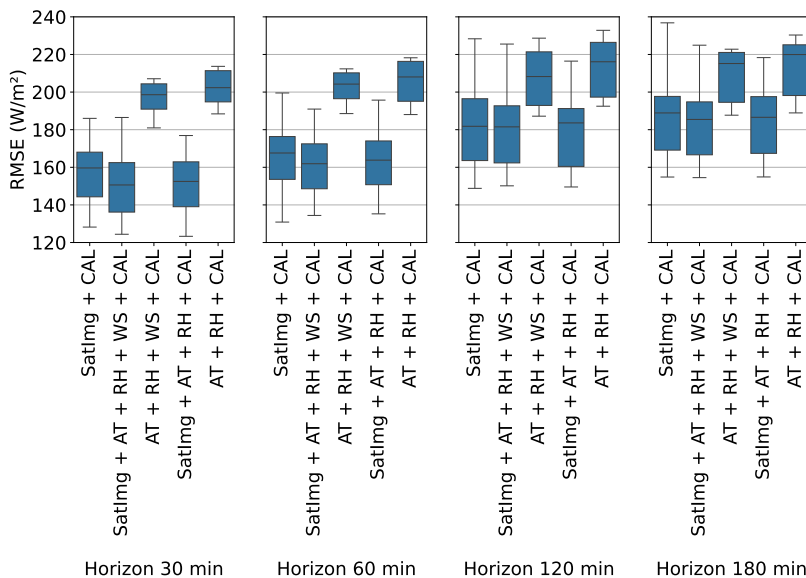

Figure 7: Distribution of RMSE results for all models trained in the presented configurations, in 2019.

As the UV feature (Ultraviolet Radiation) has a strong correlation with GHI, we choose to analyze its influence on the models separately. In the graph in Figure 8, we depict results for the forecasting models with and without considering the UV feature. In this figure, one can note a significant increase in the model's forecasting quality across all horizons with the addition of the UV component. However, this influence is much stronger at the 30-minute horizon, which is an expected result. Indeed, the gain in forecasting quality at the 30 minute horizon is of 21.16% using only the UV component and 26.4% when combining UV with satellite data, when compared to the counterparts without UV. However, for longer horizons, the UV component begins to reduce its influence on the model. Specifically, considering the 180 minute horizon, it shows only a slightly superior forecasting quality when compared to those of the satellite images alone, with a 2.9% improvement.

Finally, we take into further consideration only models that take into account UV as input parameter, alongside satellite images and ground data. Results are shown by month in Figure 9. When we evaluate the series throughout the months of the test set (2019 year), as shown, one can notice a deterioration in performance of the forecasts for the months in the interval from August to December for the 30-minute and 60-minute horizons, and from October to December for the 120-minute to 180-minute horizons. Apart from that, it becomes evident that, once the UV parameter is added to any model, the use of additional features has only marginal effect on the forecasts. It is important to note, however, that UV measurements may not be readily available and, furthermore, the use of satellite images opens new possibilities, allowing for the estimation of GHI in areas without the possibility of ground measurements, for instance. Moreover, as we already discussed, the significance of UV values for longer forecasting horizons tends to decay.

We close our discussion of the results by depicting the best results for each model configuration, in Table 3. Once again, note that the addition of the UV feature greatly reduces RMSE values. Furthermore, even considering the best models, the addition of
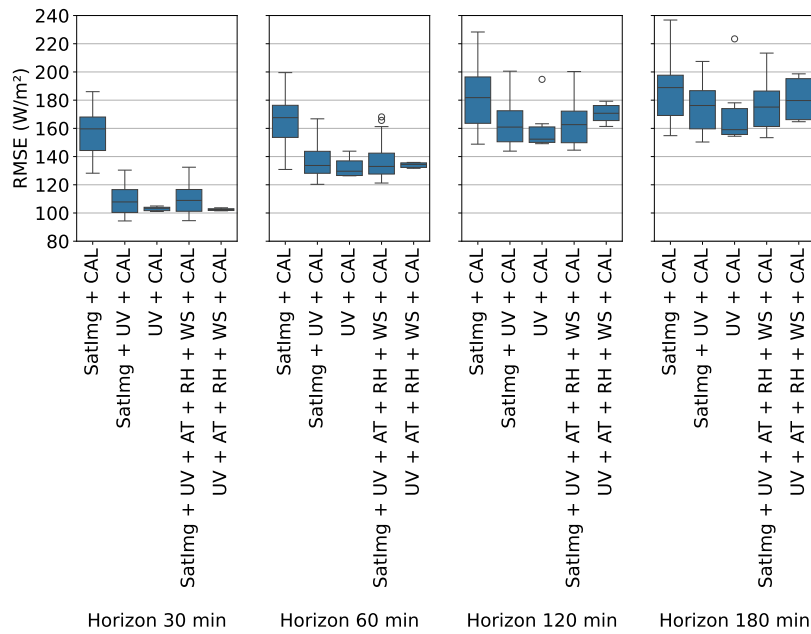
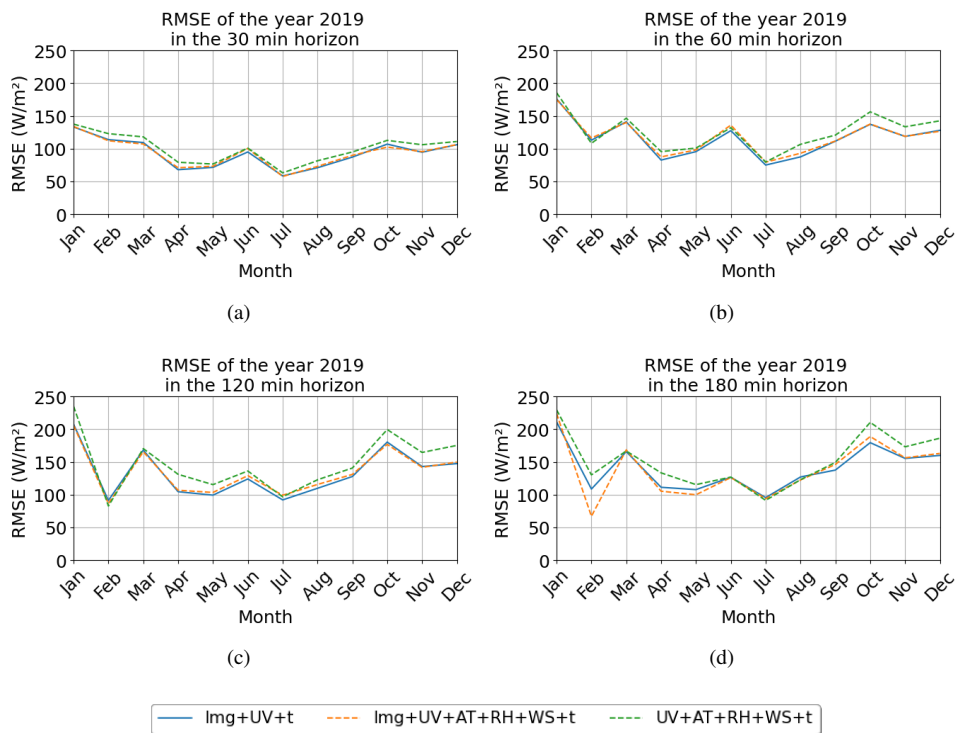Figure 8: Distribution of RMSE results for all models trained in the presented configurations, in 2019.

Figure 9: Monthly RMSE for the year 2019.

calendar data provides a significant improve when considering the baseline of only satellite images alone. Although it is difficult to directly compare our results to those of similar works (given that solar irradiance is highly dependent on the actual location of the plant), we observe that the RMSE values obtained in our work behave similar to those reported by [19] and [13] across different forecasting horizons, that is, the higher to horizon, the higher the RMSE. In perspective with these two works, the RMSE values obtained in our work are slightly higher. Once again, however, we emphasize that locations and model architectures differ, making it difficult to establish a direct and unbiased comparison with these related, yet different, works.

Table 3: Summary of the best RMSE results for each feature subset considering the different forecasting horizons.

Model Configuration	RMSE 30min W/m^2	RMSE 60min W/m^2	RMSE 120min W/m^2	RMSE 180min W/m^2
SatImg	182.27	205.85	234.94	244.06
SatImg + CAL	128.21	130.87	148.80	154.82
UV + CAL	101.08	126.38	149.14	154.43
UV + AT + RH + WS + CAL	101.54	131.67	161.38	164.72
AT + RH + WS + CAL	180.95	188.56	187.18	187.71
AT + RH + CAL	188.41	188.04	192.48	188.92
SatImg + UV + CAL	**94.36**	**120.33**	**143.86**	**150.33**
SatImg + UV + AT + RH + WS + CAL	94.55	121.24	144.56	153.40
SatImg + AT + RH + WS + CAL	124.39	134.40	150.15	154.50
SatImg + AT + RH + CAL	123.30	135.24	149.54	154.87

6 CONCLUSIONS

This study evaluated different deep learning approaches for the forecast of GHI values at 30, 60, 120, and 180-minute horizons, integrating satellite images with meteorological data (obtained from a ground station) to investigate their influence on the results. For this purpose, research scenarios were presented where the influence of the variables with the greatest weight, namely, day of the year, minutes of the day, satellite images, and UV radiation, were analyzed for each horizon.

We observed that adding temporal data such as time of day and day of the year resulted in a significant improvement in the results for all prediction horizons. The proposed models benefited from these components to adjust network parameters to make it more sensitive to GHI variations caused by the sun's movement relative to the plant. This gain intensified with the increase in the prediction horizon, demonstrating that for longer horizons, temporal data has greater influence. These results show that temporal data tends to positively influence prediction.

We identified that models using only climatic data from the meteorological station, including air temperature, relative humidity, wind speed, and temporal data, provided inferior forecasts compared to models using only satellite data and temporal data, with improvements of 41.14% at the 30-minute horizon and 21.24% at the 180-minute horizon when compared to models without satellite images. Thus, we observed that satellite images significantly improve these models at all proposed horizons. However, with an increase in the forecasting horizon, the influence of satellite images on the results begins to decrease, with inferior performance as the horizon increases. When combining meteorological data with satellite data, we noticed only a slight performance gain compared to models using only satellite images and temporal data.

We observed that the UV component has a significant influence on the network's performance in predictions with shorter horizons. When comparing prediction models using only satellite images with those using only UV radiation, the latter had a performance 21.16% higher compared to only satellite images and temporal data at the 30-minute horizon. This result is attenuated with the increase in the horizon, where the gains from adding the UV component end up being smaller.

Finally, the results obtained here confirm that satellite images positively influence neural network models for predicting energy generation in photovoltaic panels. The addition of satellite images resulted in improved models, especially at longer horizons. It is worth noting that although the UV component or GHI itself is ideal, it is not always possible to install a solarimetric station to analyze the energy potential of a specific area. In this condition, satellites are accessible and provide a cheap alternative for mapping and have the capability to function as well as the station for GHI prediction.

As future work, we plan to investigate if the estimates obtained in a particular location can be transposed to different ones, of course, with the addition of further features that would minimally describe locations. In this sense, this would allow for the generation of location agnostic models, which, in turn, would be able to provide the photovoltaic potential of a particular region of interest, without the need of any ground data.

References

[1] V. Masson-Delmotte, P. Zhai, H. O. Pörtner, D. Roberts, J. Skea, P. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J. B. R. Matthews, Y. Chen, X. Zhou, M. I. Gomis, E. Lonnoy, T. Maycock, M. Tignor and T. Waterfield. "IPCC, 2018: Global warming of 1.5°C.", 2018.

[2] M. Z. Jacobson, A.-K. von Krauland, Z. F. Burton, S. J. Coughlin, C. Jaeggli, D. Nelli, A. J. H. Nelson, Y. Shu, M. Smith, C. Tan, C. D. Wood and K. D. Wood. "Transitioning All Energy in 74 Metropolitan Areas, Including 30 Megacities, to 100% Clean and Renewable Wind, Water, and Sunlight (WWS)". vol. 13, no. 18.

[3] T.-P. Chu, J.-H. Guo, Y.-G. Leu and L.-F. Chou. "Estimation of solar irradiance and solar power based on all-sky images". vol. 249, pp. 495–506.

[4] H. T. Pedro, C. F. Coimbra, M. David and P. Lauret. "Assessment of machine learning techniques for deterministic and probabilistic intra-hour solar forecasts". vol. 123, pp. 191–203.

[5] R. Ahmed, V. Sreeram, Y. Mishra and M. D. Arif. "A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization". *Renewable and Sustainable Energy Reviews*, vol. 124, pp. 109792, 2020.

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

[7] S. Sobri, S. Koohi-Kamali and N. A. Rahim. "Solar photovoltaic generation forecasting methods: A review". *Energy Conversion and Management*, vol. 156, pp. 459–497, 2018.

[8] H. T. Pedro and C. F. Coimbra. "Assessment of forecasting techniques for solar power production with no exogenous inputs". vol. 86, no. 7, pp. 2017–2028.

[9] A. Alcaniz, D. Grzebyk, H. Ziar and O. Isabella. "Trends and gaps in photovoltaic power forecasting with machine learning". *Energy Reports*, vol. 9, pp. 447–471, 2023.

[10] J. Qin, H. Jiang, N. Lu, L. Yao and C. Zhou. "Enhancing solar PV output forecast by integrating ground and satellite observations with deep learning". *Renewable and Sustainable Energy Reviews*, vol. 167, pp. 112680, 2022.

[11] N. N. STAR. "GOES ABI (Advanced Baseline Imager) Realtime Imagery", 2017.

[12] L. W. Luiz Henrique Buzzi and P. A. Jaskowiak. "Deep Learning and Satellite Images for Photovoltaic Power Forecasting: A Case Study". In *Anais do XVI Congresso Brasileiro de Inteligência Computacional (CBIC'2023)*, edited by E. Simas, D. D. Ferreira and L. R. Oliveira, pp. 1–8, Salvador, BA, 2023. SBIC.

[13] Z. Si, Y. Yu, M. Yang and P. Li. "Hybrid Solar Forecasting Method Using Satellite Visible Images and Modified Convolutional Neural Networks". *IEEE Transactions on Industry Applications*, vol. 57, pp. 5–16, 2021.

[14] D. S. Kumar, G. M. Yagli, M. Kashyap and D. Srinivasan. "Solar irradiance resource and forecasting: a comprehensive review". *IET Renewable Power Generation*, vol. 14, no. 10, pp. 1641–1656, 2020.

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

[16] Y. Eissa, P. R. Marpu, I. Gherboudj, H. Ghedira, T. B. Ouarda and M. Chiesa. "Artificial neural network based model for retrieval of the direct normal, diffuse horizontal and global horizontal irradiances using SEVIRI images". *Solar Energy*, vol. 89, pp. 1–16, 2013.

[17] F. Wang, X. Lu, S. Mei, Y. Su, Z. Zhen, Z. Zou, X. Zhang, R. Yin, N. Duić, M. Shafie-khah and J. P. S. Catalão. "A satellite image data based ultra-short-term solar PV power forecasting method considering cloud information from neighboring plant". vol. 238, pp. 121946.

[18] L. Cheng, H. Zang, Z. Wei, T. Ding, R. Xu and G. Sun. "Short-term Solar Power Prediction Learning Directly from Satellite Images With Regions of Interest". vol. 13, no. 1, pp. 629–639. Conference Name: IEEE Transactions on Sustainable Energy.

[19] R. Marquez, H. T. Pedro and C. F. Coimbra. "Hybrid solar forecasting method uses satellite imaging and ground telemetry as inputs to ANNs". *Solar Energy*, vol. 92, pp. 176–188, 2013.

[20] A. A. Pelisson, T. F. Covões, A. Spengler and P. A. Jaskowiak. "Comparative Study of Photovoltaic Power Forecasting Methods". pp. 555–566, 2020.

[21] C. W. Chow, B. Urquhart, M. Lave, A. Dominguez, J. Kleissl, J. Shields and B. Washom. "Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed". *Solar Energy*, vol. 85, pp. 2881–2893, 11 2011.

[22] Google. "Google Maps", 2022.

[23] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu and X. Zheng. "TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems", 2015. Software available from tensorflow.org.

[24] F. Chollet *et al.*. "Keras", 2015.