Deep Learning and Satellite Images for Photovoltaic Power Forecasting: A Case Study

Luiz Henrique Buzzi, Lucas Weihmann and Pablo Andretta Jaskowiak
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 growing demand for renewable energy resources presents a supply management challenge, as photovoltaic (PV) energy exhibits intermittent generation due to meteorological factors. The unpredictability of these variations leaves power grids vulnerable to instability, quality, and balance issues. In this context, accurate forecasting of PV power generation can improve management through generation planning, allowing for the balancing of different energy sources, which is crucial for achieving widespread PV energy adoption. The rapid development and significant advancements in deep learning present new possibilities for the use of satellite imagery in PV power forecasting. In this work we build and evaluate several deep learning models in the context of PV power forecasting, aiming at 30 and 60 minutes horizons. Our models are built for the prediction of the Global Horizontal Irradiance (GHI) component which, due to its strong correlation with PV power generation, can be employed not only to derive the actual PV plant output, but also as a measure generation potential, regardless of the actual PV plant. The models take as input images from the GOES-16 satellite and ground-based meteorological measurements, which are considered as desired outcomes. Several model configurations demonstrated the viability of GHI forecasting based on satellite imagery, with the best models achieving relative root mean squared errors (rRMSE) of 15.6% and 17.2% for 30-minute and 60-minute forecast horizons, respectively.**

***Index Terms*—photovoltaic power forecasting, deep learning, convolutional neural network, satellite image, GOES-16**

I. Introduction

Electricity is ubiquitous in our daily lives. In the Net Zero Emissions (NZE) by 2050 Scenario, electricity accounts for approximately 50% of all energy consumed worldwide [1]. Due to environmental concerns, a great deal of attention has been directed towards electricity generation. More than ever before, there is now a clearly established distinction between non-renewable and renewable energy resources, with a vast accumulated knowledge on their corresponding advantages, limitations, and challenges, specifically for the latter.

Currently, one of the most promising renewable energy resources available is, perhaps, solar energy, which can be converted into electricity with the use of photovoltaic panels. Despite its clean status, photovoltaic generation is strongly dependent of solar irradiation. The irradiation that reaches photovoltaic panels is, in turn, affected by several meteorological factors that, combined, determine the generation potential of a given photovoltaic power plant at a given time [2], [3]. Indeed, grid managers may face challenges due to the inherent nature of photovoltaic power generation, which involves inevitable and sometimes sudden alterations in power output [4].

In this scenario, precise forecasting of photovoltaic power generation has emerged as a critical factor, whether to obtain a better integration with other energy resources or to provide efficient grid management as a whole. It is important to note that the forecasting horizons may vary significantly depending on the specific requirements being considered, with three typical scenarios defined in the literature: intra-hour, intra-day, and day-ahead [3]. Regardless of the forecasting scenario, a series of techniques from different areas have been successfully employed to such a task, including statistics, time-series analysis, machine learning, and deep learning [2].

Machine learning and deep learning based forecasting have emerged as critical components in the realm of photovoltaic power generation [5]. By leveraging methods from these fields, photovoltaic power generation forecasts can be significantly improved, allowing for better integration of solar energy into the grid. Regarding deep learning, specifically, the remarkable results attained in several image processing tasks by these models offer novel prospects for incorporating satellite images into the prediction of photovoltaic (PV) power generation [6].

Taking this scenario into account, this paper investigates the combination of machine learning/deep learning models and satellite images for the forecasting of PV power generation at daily horizons (30 and 60 minutes ahead). Such investigation is performed in a case study, considering a PV plant located in the city of Florianópolis - Santa Catarina - Brazil. Satellite images are obtained from the ABI Imager of GOES-16 satellite, a 16-band latest generation spectral radiometer [7].

The remainder of the paper is organized as follows. In Section II we briefly review related work on the forecasting of photovoltaic power generation, with focus on satellite image based approaches. In Section III we introduce and discuss relevant information regarding the databases employed in our case study. In Section IV the deep learning approach developed for the task and its hyperparameters are presented. In Section V results are discussed. Finally, in Section VI, conclusions and future work directions are taken into account.

II. Related Work

As previously discussed, depending on the operational conditions under consideration, a particular forecasting horizon might be of greater interest. Even though there as some

categorizations of forecasting horizons in the literature, there is no clear consensus regarding their actual naming. Bearing that in mind, the two distinct forecasting horizons we consider in this work (30 and 60 minutes ahead) may be categorized as short-term [2] or intra-hour [3], just to mention a few.

According to Sobri et al. [4], satellite imagery is one of the most appropriate data resources for short-term PV power forecasting. Considering this scenario, this brief review presents works that employ satellite images, whether combined or not with other data sources (e.g., meteorological stations) for the short-term forecasting of PV power generation. Special attention is given to works employing machine learning and deep learning techniques. It is worth noting that due to space constraints, our review is no exhaustive. For extensive surveys on the subject, we refer the reader to [2]–[4], [8].

Si et al. [9] proposed the adoption of a Convolutional Neural Network (CNN) to estimate the Global Horizontal Irradiance (GHI) based on cloud behaviour obtained from satellite photos in the visible spectrum. Images are obtained at three distinct time instants, t, $t-1$, and $t-2$, with t indicating the current time. Pre-processing is then performed in order to normalize the data and correct the zenith angle variation. Finally, the pre-processed images are provided as input to a CNN that will perform feature extraction to relate them to the actual problem. The output of the CNN is combined with data from weather stations and fed as input to a Multilayer Perceptron (MLP), which is responsible for actually predicting the GHI. The authors highlight the importance of pre-processing the images and the effectiveness of the approach in correctly identifying cloud motion based on wind speed predictions.

Marquez et al. [10], presents a method for the prediction of PV power generation through satellite images of the visible and infrared spectrum. Their work employs satellite images of the visible spectrum to generate a cloud cover index. This is achieved by applying a zenith angle correction and a procedure to normalize the images considering ground albedo. The infrared images are used to generate a vector field that determines the direction and velocity vector of the clouds using the Particle Image Velocimetry (PIV) algorithm. The cloud velocity and cloud cover data are combined as input to an Artificial Neural Network (ANN) for the prediction of GHI at 30, 60, 90, and 120 minute horizons with 30min granularity.

Eissa et al. [11] introduces an approach using distinct networks for different weather conditions aiming to combine meteorological parameters and satellite images to determine the DNI (Direct Normal Irradiance), DHI (Direct Horizontal Irradiance), and GHI (Global Horizontal Irradiance). The authors divide the problem into DNI and DHI estimation, processing each of the variables separately. The two networks apply the same strategy, using 10.8µm and 12µm satellite spectral channels to generate a cloud mask that divides the images into cloudless and cloudy. Depending on the results, the algorithm selects which network to use and incorporates other sources of information prior to its application, such as additional channels and zenith angle, to mention a few. Four models are actually trained: two for DNI (cloud and cloudless) and two for DHI (cloud and cloudless). To obtain GHI estimates, results from DNI and DHI are combined analytically, leading to a rRMSE close to 12% for GHI estimates.

The work of Pelisson et al. [12] considered solely ground data from a solarimetric station and PV Plant located at Fotovoltaica - UFSC. Even though this particular study does not rely on satellite imagery, it is reported here since the very same ground data was employed in our work. The authors compared several machine learning techniques and explored different feature space configurations for next day power forecasting, based on ground meteorological data. Based on their results, the authors suggest that the best results were obtained with Multilayer Perceptron (MLP) and Support Vector Regression (SVR). Regarding feature space, the best performance was obtained with the combination of Global Horizon Irradiance (GHI), Direct Normal Irradiance (DNI), Direct Horizontal Irradiance (DHI), Utra Violet (UV), Relative Humidity (RH), Ambient Temperature (AT) and Wind Speed (WS). Despite the strong correlation between GHI and PV power output, the authors noted that by adding other variables to the model a reduction of almost 30% in RMSE was achieved. This led the authors to conclude that other forecasting techniques, taking into account the use of satellite images, for instance, can be used to achieve even better results.

III. Databases

This section presents briefly the data sources considered in this study. First, we take into account ground data, which was obtained from Fotovoltaica UFSC (UFSC Solar Energy Research Laboratory)[1]. Besides a PV power generation facility, the laboratory also has a solarimetric station, capable of collecting data from different parameters, such as: Global Horizon Irradiance (GHI), Direct Normal Irradiance (DNI), Direct Horizontal Irradiance (DHI), Utra Violet (UV), Relative Humidity (RH), Ambient Temperature (AT) and Wind Speed (WS). We then consider and discuss the main data source for our work, that is, data from the GOES-16 satellite. All data considered in this work comprises the years of 2018 and 2019.

A. Ground Station Data

The ground station database was obtained from Fotovoltaica UFSC (UFSC Solar Energy Research Laboratory), comprising weather related data and power generation data by the photovoltaic plant. The data was obtained with a frequency of 1min with a Campbell CR6 datalogger [12]. Despite the fact that the solarimetric station collects several variables, for our work, the only variable taken into account is the Global Horizontal Irradiance (GHI), therefore, we refrain from discussing the remaining ones. For more information, the reader may refer to Pelisson et al. [12]. There are two reasons for considering solely GHI and discarding all the remaining variables from the solarimetric station: (i) we want to evaluate the use of satellite images for power forecasting and (ii) from the three available irradiance components (GHI, DNI, and DHI), GHI

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

is the one with the highest value of pearson correlation w.r.t. PV power generation, as shown by Table I (correlation values presented in the table are w.r.t the clean database, see below for details on it). Indeed, GHI has a very strong correlation with the generated power, as expected. Such an observation holds true with data surveyed in the literature and allows us to build a prediction model that has the GHI as output. 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.

TABLE I
CORRELATION BETWEEN IRRADIANCE AND POWER GENERATION.

Irradiation Component	Correlation w.r.t. power generation
GHI	0.97
DNI	0.77
DHI	0.36

Before actually using the database, pre-processing work had to be done. 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.

An analysis of the average distribution of GHI over the day is shown in the graph in Figure 1. In order to delimit our forecast ranges, we took data represented in this image into account. It is possible to observe that the solarimetric station (and consequently the PV Power station, given that they are co-located) starts to receive solar radiation around 6am, remaining active until 5pm. Note that this Figure depicts average behaviour, with significant changes occurring due to seasons of the year. Considering the most significant measurements of GHI from the database, we initially defined a time frame of analysis between 9:00am and 5:30pm. This first interval will be further narrowed after exploratory data analysis.

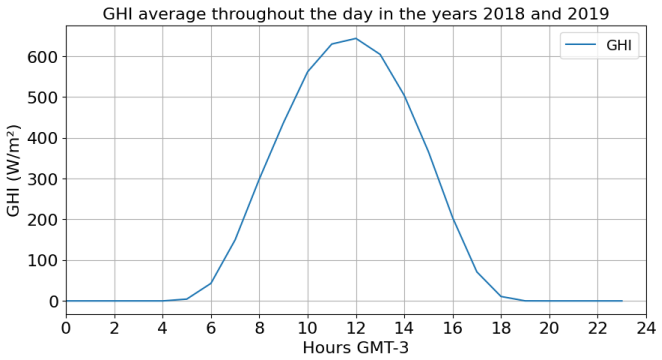


Fig. 1. Mean GHI distribution over 24 hours in the years 2018 and 2019.

Finally, considering the given prediction interval and the fact that our models will output estimates for GHI, the final database was reduced to a total of 218,549 readings (from an initial value of 871,709) with 1min granularity.

B. Satellite Data

This work takes advantage of the improvements implemented in the GOES (Geostationary Operational Environmental Satellite) satellite series[2] to build and evaluate deep learning models for GHI forecasting. It is important to make it clear that GHI data discussed in the previous section are employed only as targets (desired values) to train and evaluate our models, for which inputs are based solely on satellite data. Once trained, the models provide forecasts, once more, relying only on satellite data, described in this section.

The PV power station considered for the development of this work is located in Florianópolis - Santa Catarina - Brazil. For this particular location, the GOES-16 satellite is able to capture the desired images. The satellite data consider in this work is within the same interval as the ground station data (2018 and 2019). The processing of the data began with *full disk* images, which cover the entire GOES-16 capture area. It is important to note that for the time period under consideration in this study, GOES-16 provides captures (images) every 30min.

Images from the GOES-16 satellite are provided by the Advanced Baseline Imager (ABI). We considered images from channel 2 of the ABI, which provide greater spatial resolution and provides in its images data relevant to our application [7]. Each raw ABI image from channel 2 has approximately 500MB. To reduce the volume of downloaded data, only data from 9:00am to 5:30pm GMT-3 was considered, as previously discussed. For each image, a region of interest of $\pm 5^{\circ}$ latitude and $\pm 5^{\circ}$ longitude was centered over the point of interest (location of the power station), which has coordinates -27.4310° latitude and -48.4414° longitude.

Even with the cropping, the resulting images are still considerably large (2124x875 pixels), resulting in roughly 1.8 million pixel values. In the case of no further processing, this would be the required input size of the deep learning model for our study, demanding massive computational processing.

In order to reduce the size of the images and consequently the computational processing necessary to train and evaluate the models, we took into account cloud speeds. Based on previous work from Chow et al. [13], in which clouds that most affect forecasting were found to travel below 28km/h, we conjectured that an upper limit of 40 km/h would be enough to capture significant cloud movements. Considering the values of maximum cloud speed (40km/h), pixel granularity (0.5km in side per pixel), and the frequency of capture (30min), we estimated that images comprising 40×40 pixels ($20km \times 20km$) would be adequate for our forecasting horizons (30/60 min). The area corresponding to the final crop is shown in Fig. 2. The final images provided to the models have a total of 1600 points. Note that this picture is shown for illustrative purposes only. Satellite imagery employed by the models in this work account for reflectance information, as we detail in the following.

[2]https://www.goes-r.gov/

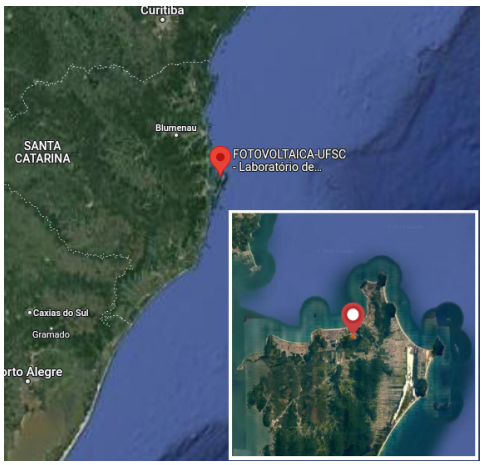
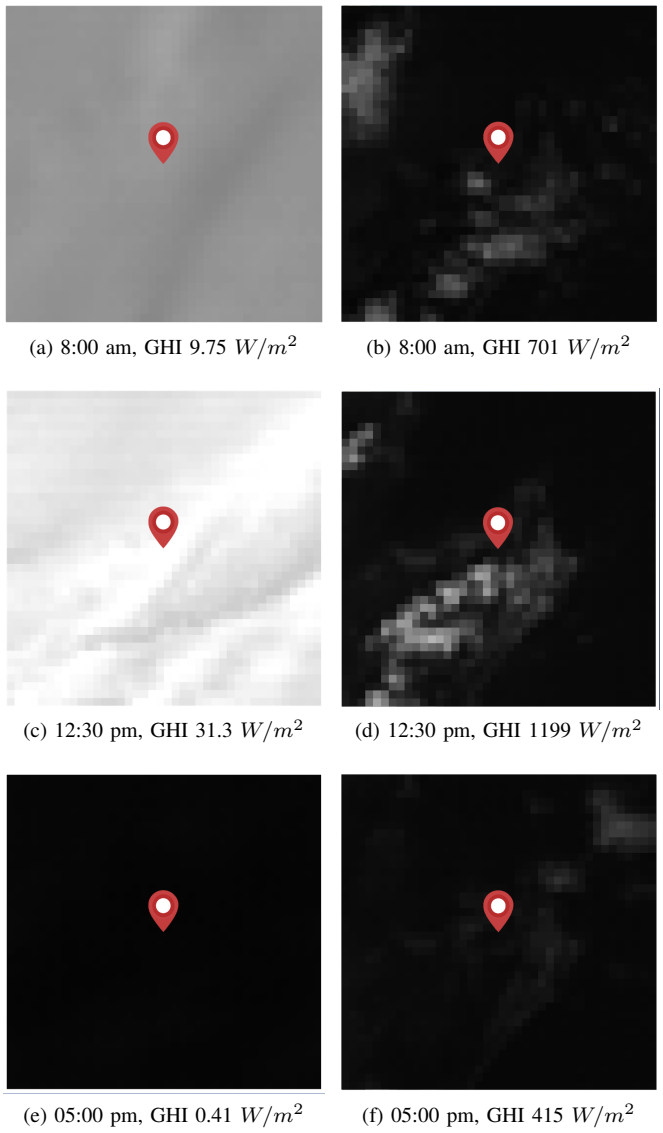


Fig. 2. Outer image provides a broad view of the location of interest. The crop shown in the bottom-right corner corresponds to actual area used as input to the deep learning models. Images shown here are obtained from Google Maps. Models are built on the basis of GOES-16 images from channel 2.

Actual images, as provided to the models, are shown in Fig. 3. Each line of the figure accounts for a different hour of the day: early-day, mid-day, and end-day. For each given time, we selected the images with the lowest and highest GHI observed for the years of 2018 and 2019 (time period of the database). Note that, for images in which sunlight is present, low GHI ground measurements are an indicative of high reflectance. This is actually expected and can be observed in the subfigures (a) and (c), in which bright pixels indicate a high reflectance (which can be caused, for instance by clouds). A low GHI is not observed only in images with high reflectance though. This can be noticed in subfigure (e), an image from June 12, 2018, for which at 17h there is limited sunlight, resulting in a dark image and low GHI. Finally, a cloud free image is depicted in (d), which allows for a visual comparison of the island contours with Fig. 2.

Having made such considerations regarding the pre-processing steps, which are mostly related to image cropping, and considering the initial time window between 9:00am and 5:30pm, the satellite image database consists of 7,136 images for the year 2018 and 7,193 images for the year 2019.



(a) 8:00 am, GHI 9.75 W/m^2 (b) 8:00 am, GHI 701 W/m^2

(c) 12:30 pm, GHI 31.3 W/m^2 (d) 12:30 pm, GHI 1199 W/m^2

(e) 05:00 pm, GHI 0.41 W/m^2 (f) 05:00 pm, GHI 415 W/m^2

Fig. 3. Actual satellite images as provided to the models evaluated in this work. Each image is a crop with 40×40 pixels in dimension, centered at the PV Power Plant. Images were obtained from the GOES-16 satellite, channel 2, at different time instants. Below each image we depict its timestamp and correspondent GHI, measured at ground level at the solarimetric station.

C. Exploratory Analysis

With both databases ready for further analyses, we first verified the linear correlation (Pearson correlation — ρ) between satellite data and ground GHI data, as measured by the solarimetric station. All timestamps presented from now consider the GMT-3 time zone. Given that satellite images cover a large region and actual GHI ground measurements are located in the center of the images, we obtained the reflectance of each image by the average of the closest 4 points (pixels) w.r.t. the solatrimetric station, which corresponds to an area of $1km^2$ centered at it. Once this averaged reflectance was obtained, it was then correlated with ground GHI data.

When correlating GHI with reflectance in the original time window considered for analysis (9:00am to 5:30pm), a weak correlation was observed ($\rho = -0.34$). Note that the negative sign indicates that higher reflectances result in lower GHI measurements, which is expected. Si et al. [9] emphasizes the influence of the zenith (which accounts for the angle os sunlight incidence) angle when estimating the GHI using satellite images. To investigate its influence in our results, we correlated the same data, within a different time window, from 11:00am to 1:00pm, for which the zenith angle influence should be diminished. Indeed, for this interval, the correlation we observed was higher ($\rho = -0.77$). Given that we were not aiming for zenith correction at this point of investigation (future work), we made a compromise with the window of predictions our model would deal, and shortened it to consider measurements obtained between 9:00am and 03:00pm, for

which a moderate correlation was observed ($\rho = -0.66$). The correlation between GHI and satellite reflectance for each point of the cropped region of interest for this time frame is depicted in Fig. 4.

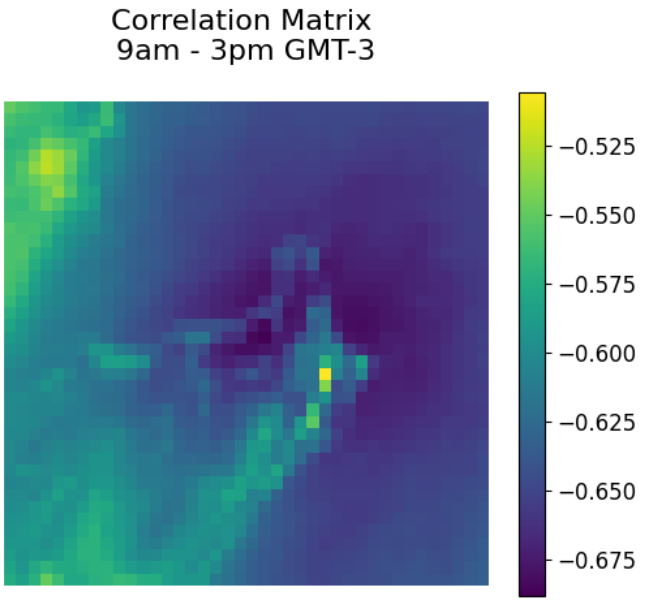


Fig. 4. Correlation between ground GHI measurements and reflectance, as obtained from satellite data, w.r.t. the 9am to 3pm (GMT-3) time window.

With the time window for the study defined, we briefly discuss how input data was organized before being fed to the models. This is discussed in the next section, given that data organization is heavily influenced by the model topology.

IV. Deep Learning Based Approach

Before defining models and input data organization, it is worth recalling that we have two different forecast settings, in which we consider 30min and 60min ahead predictions. The Convolution Neural Networks (CNNs) investigated in this paper follow closely the proposal by Si et al. [9], receiving as input three satellite images from different timestamps, namely: t, $t - s$, and $t - 2s$, where t accounts for the actual time and s accounts for the time between consecutive satellite image measurements, which is 30min. Therefore, each model receives as inputs images from t, $t - 30$min, and $t - 60$min. Examples of this arrangements are provided in Table II, for 30min and 60min forecasting horizons. Considering that each image is 40×40 pixels in size, the actual input of the Deep Learning approach is composed of 4800 values.

Given the input images, convolutional layers are then employed for processing satellite images, with each layer consisting of 2D convolution, normalization, and pooling, as shown in Fig. 5. Convolution was based on ReLU activation functions. As for pooling, 2D *average pooling* was applied, which uses the same concept as MaxPooling but averaging the kernel, with a 2x2 step. This topology was adapted from Si et al. [9] for our case study. The number of convolutional layers,

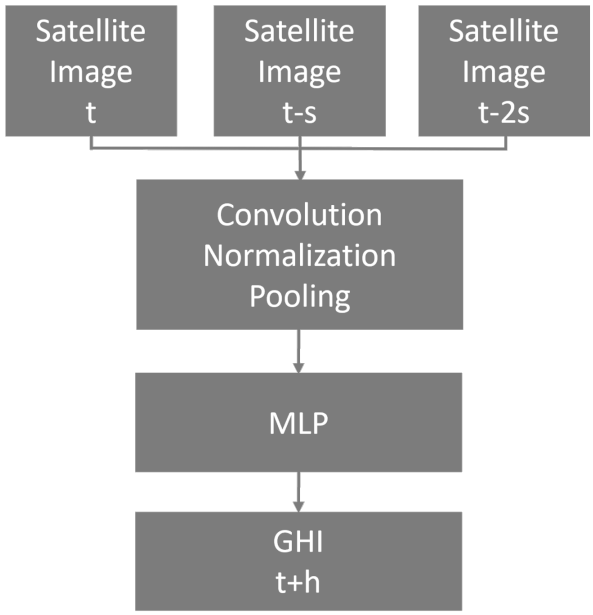


Fig. 5. Topology of the Deep Learning approach employed in our case study (based on the work from Si et al. [9]). Note that a model takes three satellite images as input, in different time instants: t (actual time), $t - s$, and $t - 2s$, with s equals the time difference between satellite readings, which is 30min. The data goes through a CNN and a MLP, in order to obtain an estimated GHI for $t + h$, with h (horizon) equals to 30min or 60min.

the amount of convolutional filters and the kernel size are objects of this study in this work, discussed in the following.

In order to train the models, in addition to the satellite images, the expected GHI values at $t + h$ (30min or 60min, according to the model) were also considered, normalized between 0 and 1. For each forecast horizon we tested different model settings, aiming to verify how each hyperparameters affects the actual forecasting. Different models were built considering the number of convolutional layers, the number of filters, the size of the kernel, the number of hidden layers (for MLP), the dropout rate and the number of neurons per layer. The possible values for each one of this hyperparameters is given below:

- Convolutions: {1, 2, 3, 4}
- Filters: {16, 32}
- Kernel: {2, 3, 4, 5}
- Layers: {1, 2, 3, 4, 5}
- Dropout: {0.00, 0.15, 0.30, 0.45}
- Neurons: {64, 128, 256, 512, 1024}

Considering these hyperparameters and values, a total of 6,400 different models were trained and evaluated, with 3,200 models for each forecasting horizon (30min and 60min).

The database was split into 3 subsets, namely: train, validation, and test. The first one was employed to train each one of the models. This subset is composed of 70% of the data from 2018, with a balanced representation of each month, that is, the training set has 70% of January, 70% of February, and

TABLE II
ILLUSTRATIVE EXAMPLES OF HOW SATELLITE IMAGES WERE CONSIDERED AS INPUT FOR THE DL APPROACH.

Forecasting Horizon	Current Time	Model Inputs			Model Output
		Image t	Image t+30	Image t+60	
30	10/03/2018 12:00 am	Image 12:00 am	Image 11:30 am	Image 11:00 am	GHI 12:30 am
	10/03/2018 12:30 am	Image 12:30 am	Image 12:00 am	Image 11:30 am	GHI 01:00 pm
	10/03/2018 01:00 pm	Image 01:00 pm	Image 12:30 am	Image 12:00 am	GHI 01:30 pm
	10/03/2018 01:30 pm	Image 01:30 pm	Image 01:00 pm	Image 12:30 am	GHI 02:00 pm
60	10/03/2018 12:00 am	Image 12:00 am	Image 11:30 am	Image 11:00 am	GHI 01:00 pm
	10/03/2018 12:30 am	Image 12:30 am	Image 12:00 am	Image 11:30 am	GHI 01:30 pm
	10/03/2018 01:00 pm	Image 01:00 pm	Image 12:30 am	Image 12:00 am	GHI 02:00 pm
	10/03/2018 01:30 pm	Image 01:30 pm	Image 01:00 pm	Image 12:30 am	GHI 02:30 pm

so on. The second subset is composed of the remaining 30% of the 2018 data, and is used for validation, that is, parameter selection. Finally, the third subset comprises the data from 2019 (not employed before), used to test the selected models.

The network was trained to a limit of 1000 epochs, using an early stopping function with 50 epochs of persistence, with a batch size of 512. Models were built using Keras [14].

V. RESULTS

After performing training and validation for each one of the models with 2018 data, these were then tested with 2019 data and evaluated using relative Root Mean Squared Error (rRMSE) as metric [15].

Given the number of different models under evaluation (6,400) we produce different plots in order to better evaluate/understand the contribution of each one of their hyperparameters to the observed values of rRMSE. These plots are shown in Fig. 6. We produce a subplot — from (a) to (f) — for each combination of hyperparameter and forecasting horizon. Within each of these plots, a boxplot accounts for each one of the individual hyperparameter values available/considered in our work. An evaluation of the number of convolutional layers, shown (a), indicates that, for both horizons, the adoption of 2 convolutional layers led to results with lowest values of rRMSE and decreased variance. Regarding the number of filters, according to (b), one can observe that a number of 16 filters appears to be the the best choice for the two forecasting horizons. Regarding kernel size, shown in (c), values 3 and 4 have a very similar distribution, with slightly better results.

As for the number of hidden layers, shown in (d), a single hidden layer provides better results than the remaining confifurations (up to 5 hidden layers were considered). Indeed, it is clear from the boxplots that the error dispersion increases together with the number of layers, which may indicate that complex networks had difficulties to convergence. Regarding the distribution of the number of neurons, presented in (e), the results obtained for 256 and 512 neurons are very close for the 30 minute horizon, whereas for the 60 minute horizon the results for 512 neurons had a lower dispersion and a lower rRMSE. Finally, considering the dropout rate presented in (f), which contributes to model generalization, models achieved the best result with 0.15. However, it is interesting to note that increasing the dropout rate seems to produce more unstable ANNs, which can be seen by an increasing error dispersion.

By evaluating each of the hyperparameters individually, it is possible to find those that lead to better results, serving as a reference for choosing a combination for the final ANN model. We select for further inspection the top five models for each one of the forecasting horizons (30min. and 60min.). The results of these methods, alongside with their specific hyperparameters is shown in Table III. The best models achieved good results with rRMSE of 15.6% and 17.2% for the 30-minute and 60-minute forecasting horizons, respectively. A careful observation of the best results shown in Table III, allows one to conclude that, in general, fewer hidden layers led to better models. Although in general both horizons performed better with a small number of neurons per layer, the 30-minute horizon, with one exception, has an even smaller number of neurons per layer when compared to the 60-minute horizon.

TABLE III
CONFIGURATION OF THE TOP FIVE METHODS W.R.T. THEIR RRMSE (%) VALUES CONSIDERING 30 AND 60 MINUTE FORECASTING HORIZONS.

Horizon (min)	*# Convolutions*	*# Filters*	*Kernel Size*	*# Layers*	*Dropout*	*# Neurons*	*rRMSE (%)*	*Model #*
30	2	16	5	1	0.45	64	15.6	M01
	2	32	4	1	0.45	1024	15.8	M02
	2	32	5	1	0.30	64	15.8	M03
	2	16	3	2	0.45	64	15.8	M04
	2	16	3	2	0.30	64	15.8	M05
60	2	32	5	1	0.45	256	17.2	M06
	2	32	3	2	0.15	256	17.2	M07
	1	32	3	5	0.15	256	17.5	M08
	1	16	5	1	0.30	64	17.5	M09
	1	32	5	1	0.45	512	17.5	M10

We select for further inspection the best models evaluated in each one of the forecasting horizons under consideration. A comparison between the predicted (obtained from the model) and expected (obtained from the solarimetric station) GHI

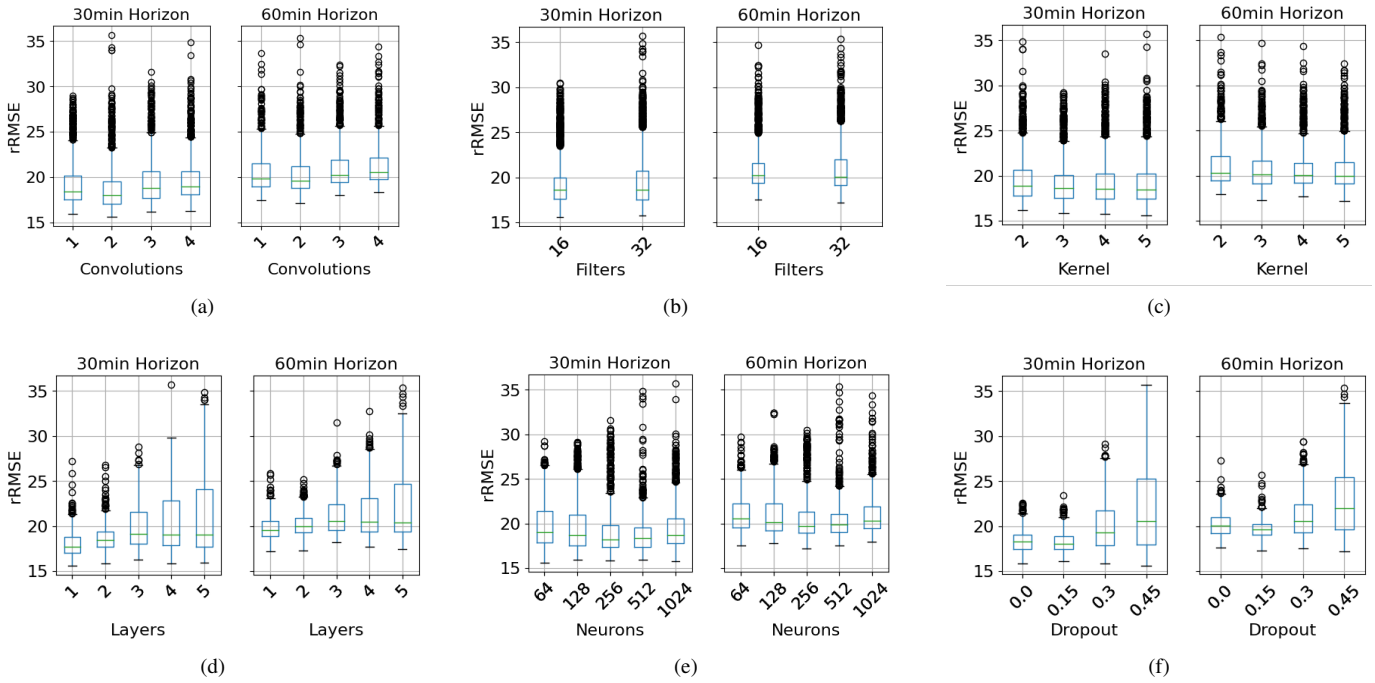


Fig. 6. Summary of the relative Root Mean Squared Error (rRMSE) for the 6,400 different model configurations evaluated in this work (3,200 for each one of the forecasting horizons, that is, 30min and 60min). Each subplot — indexed from (a) to (f) — presents boxplots indicating model performance factored by a specific hyperparameter and its possible values, in order to allow for a better inspection of their individual influences. In clockwise order, we depict: (a) number of convolutions; (b) number of filters; (c) kernel size; (d) number of layers; (e) number of hidden units/neurons per hidden layer; and (f) dropout rate.

values for all instances of the database, is presented through the scatter plot, as shown in Fig. 7. By analyzing the graph it is possible to infer that, in general, the dispersion is symmetric for the different GHI values, except for the small values, where the predicted values are overestimated in relation to the actual values. No reasons were found throughout this work to justify this particular behavior, and it is therefore a subject of investigation for future work.

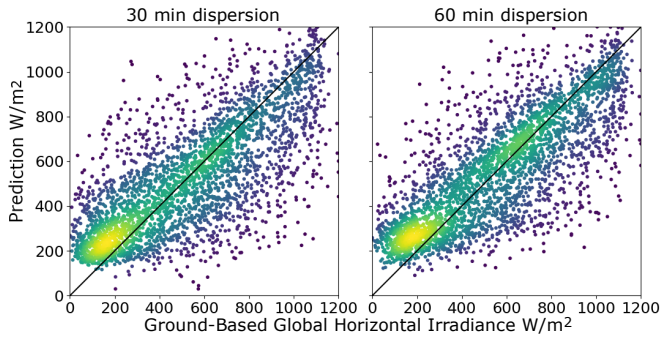


Fig. 7. Distribution of CNN's GHI prediction vs. Ground-Based GHI for the best result for the 30-minute and 60-minute prediction horizons.

VI. Conclusions

In this work we tackled the problem of photovoltaic power generation forecasting. Due to the intermittent nature of this energy resource, its forecasting is crucial not only for wide adoption, but also for grid integration with other resources.

Model forecasts were based solely on GOES-16 satellite images from channel 2, which provide important information regarding reflectance. Aiming to obtain accurate short-term forecasts for the next 30 and 60 minutes, we provided to the models a series of three consecutive satellite images. This approach was designed to supply the models with sufficient data to detect trends and subsequently produce reliable forecasts. Given that forecasts aim at Global Horizontal Irradiance (GHI) levels which, in turn, is highly correlated with PV Power Output, they are independent of the actual plant itself.

Our results support that different models were capable of providing good results for GHI estimation, with relative Root Mean Squared Errors (rRMSE) in the order of 15%. From the hyperparameters evaluated in this work, the ones that seem to mostly affect model performance are: number of network layers, dropout, and number of neurons per layer. Even though the remaining hyperparameters cause some changes in the observed model outcome, they do not seem critical enough to justify in-depth exploration, at least in our evaluation scenario.

Despite the promising results, we believe there is potential for further improvements and future work. First and foremost, we have to consider that the models evaluated in this work received as input only raw satellite images. We believe that improvements could be obtained by considering and correcting for the solar zenith angle. Another line of investigation that we conjecture would provide improvements relies on fusing temporal data, such current time, day of the year, month (or season), with satellite images as input to the models. Finally,

we believe that the fusion of satellite images with ground data from solarimetric stations is also a promising line of research.

REFERENCES

[1] L. I. E. A. Cozzi and T. I. E. A. Gould, "World energy outlook 2021," pp. 1–386, 2021. [Online]. Available: www.iea.org/weo

[2] 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, p. 109792, 2020. [Online]. Available: https://doi.org/10.1016/j.rser.2020.109792

[3] 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. [Online]. Available: http://dx.doi.org/10.1016/j.solener.2016.06.069

[4] 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. [Online]. Available: https://doi.org/10.1016/j.enconman.2017.11.019

[5] 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. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352484722025975

[6] 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, p. 112680, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032122005718

[7] N. N. STAR, "Goes abi (advanced baseline imager) realtime imagery," 2017. [Online]. Available: http://cimss.ssec.wisc.edu/goes/goesdata.html

[8] 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. [Online]. Available: http://dx.doi.org/10.1016/j.renene.2016.12.095

[9] 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.

[10] 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. [Online]. Available: http://dx.doi.org/10.1016/j.solener.2013.02.023

[11] 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. [Online]. Available: http://dx.doi.org/10.1016/j.solener.2012.12.008

[12] A. Pelisson, T. Covoes, A. Spengler, and P. Jaskowiak, "Comparative study of photovoltaic power forecasting methods," pp. 555–566, 2020.

[13] 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.

[14] F. Chollet *et al.*, "Keras," https://keras.io, 2015.

[15] M. Despotovic, V. Nedic, D. Despotovic, and S. Cvetanovic, "Evaluation of empirical models for predicting monthly mean horizontal diffuse solar radiation," *Renewable and Sustainable Energy Reviews*, vol. 56, pp. 246–260, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032115013258