Huertas-Tato, J., Aler, R., Rodríguez-Benítez, F.J., Arbizu-Barrena, C., Pozo-Vázquez, D. y Galván, I.M. (2018). Predicting Global Irradiance Combining Forecasting Models Through Machine Learning. In *HAIS 2018: Hybrid Artificial Intelligent Systems,* 10870, pp. 622-633.

DOI: https://doi.org/10.1007/978-3-319-92639-1_52

# Predicting Global Irradiance Combining Forecasting Models Through Machine Learning

J. Huertas-Tato[1(✉)], R. Aler[1], F. J. Rodríguez-Benítez[2], C. Arbizu-Barrena[2], D. Pozo-Vázquez[2], and I. M. Galván[1]

[1] Computer Science Departament, Universidad Carlos III de Madrid, Madrid, Spain
jahuerta@inf.uc3m.es
[2] Department of Physics, Universidad de Jaén, Jaén, Spain

**Abstract.** Predicting solar irradiance is an active research problem, with many physical models having being designed to accurately predict Global Horizontal Irradiance. However, some of the models are better at short time horizons, while others are more accurate for medium and long horizons. The aim of this research is to automatically combine the predictions of four different models (Smart Persistence, Satellite, Cloud Index Advection and Diffusion, and Solar Weather Research and Forecasting) by means of a state-of-the-art machine learning method (Extreme Gradient Boosting). With this purpose, the four models are used as inputs to the machine learning model, so that the output is an improved Global Irradiance forecast. A 2-year dataset of predictions and measures at one radiometric station in Seville has been gathered to validate the method proposed. Three approaches are studied: a general model, a model for each horizon, and models for groups of horizons. Experimental results show that the machine learning combination of predictors is, on average, more accurate than the predictors themselves.

**Keywords:** Global irradiance forecasting · Machine learning
Combining forecasting models

## 1 Introduction

A key issue to increase the competitiveness of the solar energy and to increase their share in the electric systems is the improvement of the reliability of the solar energy forecasts. In the last years, a wide range of forecasting methodologies has been developed, with very different characteristics, such as the spatial and temporal resolution or their forecasting horizon [1].

Machine Learning [2] has played an important role on improving solar energy forecasting [3,4]. Nevertheless, there is still room for improvement. In this regard, there have been some efforts to combine different sources of information (observations, camera, satellite, ...) and take advantage of the possible synergies.

For example, real measures and camera have been combined using an arti-ficial neural network optimized through a genetic algorithm [5]. [6] proposes a combination of NAM (North American Mesoscale Model) with cloudiness information obtained from satellite images. This model improves spatial resolution of the NAM, while improving intra-day and 1-day predictions. A system based in extreme learning machines optimized through evolutionary computation (coral reef algorithm) combines direct measures, radiosondes and NWP to obtain the daily prediction of solar irradiance [7]. In [8], a combined prediction of cloud cover derived from a sky-camera and satellite offers a forecast of up to three hours. In a similar way, a combination of cloudiness estimation from satellite and the NWP from European Center for Medium range Weather Forecasting (ECMWF) is proposed [9]. Further results [10] show that the combination of statistical models and NWP is able to reduce the forecasting error at one hour horizons. In [11] a machine learning blending of irradiance forecasts using a Ran-dom Forest is used. This approach combines three models: the NAM model, the SREF (Short Range Ensemble Forecast) model and the GFS model. A recent work combines satellite-derived, ground data, solar radiation, and total cloud cover to improve solar radiation forecasting for horizons between 1 h and 6 h [12].

Similar combination approaches have also been applied for wind energy. [13] combines different predictions at several horizons using an adaptive weighting, dependent on the error yielded by the past sources.

The novelty of this research is to use machine learning to combine Global Horizontal Irradiance (GHI) forecasts obtained from four sources, in order to output an improved GHI forecast. The sources are: Smart Persistence [14], Satellite [15], Cloud Index Advection and Diffusion (CIADCast) [16], and Solar Weather Research and Forecasting (WRF-Solar) [17]. These perform differently under different situations and forecasting horizons. The aim of our approach is to use machine learning to combine them automatically, so as to take advantage of their synergies and to improve the performance of sources used separately.

In this work, several prediction horizons have been tested from 15 to 360 min, in steps of 15 min. Three different approaches are proposed for integrating the four sources: general, horizon-individual and horizon-group. The approaches are different ways to treat the horizon information either by using a single gen-eral model valid for all horizons, by making specialized models for each horizon (horizon-individual), or a compromise between both (horizon-group).

The machine learning method chosen for combining them is extreme gradient boosting [18]. Gradient boosting has been used before in solar forecasting. For instance, [19,20] use meteorological variables to predict GHI. However, the aim of our work is different, because our inputs are not meteorological variables but GHI forecasts themselves.

The structure of the paper is the following: first, the predictors used as inputs of machine learning method are presented in Sect. 2. In Sect. 3, the dataset used to the evaluation is described. Section 4 explains the different machine learning approaches to be studied in this work. In Sect. 5, the experimental methodology

is explained and the experimental results are presented. The final conclusions of this research and future lines of work are presented in Sect. 6.

## 2 Description of the Predictors

This section describes the four forecasting models (or predictors) that will be combined by a machine learning model.

### 2.1 Smart Persistence

This model is computed with the actual measured irradiance $I_0$ and corrected with the variation of the clear-sky (cs) irradiances $I_{cs}$ from the initial time to a future time $t$. The relation between actual irradiance and cs irradiance at a certain time 0 is kept constant and multiplied by the clear-sky irradiances in future $t$. The European Solar Radiation Atlas [14] cs model is used (Eq. 1).

$$I(t) = \frac{I_0}{I_{cs}(0)} * I_{cs}(t) \tag{1}$$

### 2.2 Satellite-Based Model

In this method, satellite images are first processed to derive the so-called cloud index images, an intermediate step to retrieve the clear-sky index images and then the solar radiation maps [21]. Secondly, a statistical comparison of various consecutive cloud index images allows deriving the cloud motion vector field. In this case OpenPIV is used (http://www.openpiv.net/openpiv-python/). The discrete cloud motion vector field is transformed into a continuous flow computing the streamlines, i.e., a family of curves tangent to this wind field [15]. The streamline passing through the station location is used to obtain the future cloud index values, then the clear-sky index values and, finally, the GHI forecast

### 2.3 CIADCast

The CIADCast model [16] for short-term solar radiation forecasting is based on the advection and diffusion of cloud index estimates derived from satellite using the Weather Research and Forecasting [22] NWP models. Cloud index values are inserted in the WRF cell which corresponds to the cloud top height provided by the EUMETSAT product. Then, WRF is used to advect and diffuse the cloud index values as dynamical tracers both horizontally and vertically.

After the model run, the sum of each column of cloud index values is computed to obtain again a two-dimensional cloud index map. The cloud index values at the station location are used finally to derive the GHI forecast, similarly to the satellite-based model. CIADCast was run with the standard WRF model version 3.7.1, configured with 37 vertical levels and three nested domains of 27, 9 and 3 km spatial resolution. The cloud index maps were ingested in the inner domain, which has similar resolution to the satellite images. 18-h simulations were run discarding the first 6 simulated hours as spin-up.

### 2.4 WRF-Solar

NWP uses mathematical models based on physical principles of the atmo-sphere and oceans to predict the weather based on current weather conditions. WRF-Solar [17] is a particular physical configuration of the WRF numerical weather prediction model version 3.6 devised for solar energy applications. It has improved parameterizations for the interactions of solar radiation with clouds and aerosols. The model configuration used here consisted of two nested domains with 9 and 3 km spatial resolution and 37 vertical levels. As with CIADCast, simulations with 18 h of forecasting horizon and 6 h of spin-up were run.

## 3 Data

The evaluation is conducted at one radiometric station in Seville (southern Spain) where GHI (the total amount of shortwave radiation received from above by a surface horizontal to the ground) has been measured. GHI has been acquired with a Kipp & Zonen CMP6 pyranometer with a 1 min sample rate. The maintenance of radiometric stations follows World Meteorological Organization recommendations and the quality control of the data is applied following Long and Dutton [23]. The observations cover from March 2015 to March 2017.

To ensure the quality of the data, a preprocessing of the dataset has been made. Only predictions taken when the zenith is less than 75° are included in the dataset, because the hours selected by this filter are the most relevant to global irradiance in the day. Forecasts of GHI up to 6 h ahead, with a time step of 15 min, are obtained based on four different models: Smart Persistence, Satellite-based, CIADCast, and WRF-Solar. An example of the structure of the dataset is shown in Table 1. There are four different numerical inputs (four predictors), a target (measure column). On average, each horizon contains 2400 instances,

**Table 1.** Dataset example

| Date | Hour | Horizon | SmartPers | Satellite | CIADCast | WRF-Solar | Measure |
|---|---|---|---|---|---|---|---|
| 2015-03-03 | 10:15 | 15 | 299.72 | 620.9 | 230.3 | 226.29 | 283.1 |
| 2015-03-03 | 10:15 | 30 | 427.23 | 649.3 | 240.42 | 283.04 | 254.55 |
| 2015-03-03 | 10:15 | 45 | 627.45 | 674.16 | 249.23 | 326.29 | 303.18 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 2015-03-03 | 10:15 | 360 | 71.80 | 312.25 | 119.01 | 31.23 | 280.22 |
| 2015-03-03 | 10:30 | 15 | 417.26 | 649.01 | 295.54 | 296.93 | 254.55 |
| 2015-03-03 | 10:30 | 30 | 666.79 | 673.98 | 306.37 | 401.20 | 303.18 |
| 2015-03-03 | 10:30 | 45 | 636.63 | 619.67 | 445.28 | 599.42 | 347.69 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 2015-03-03 | 10:30 | 360 | 471.29 | 556.22 | 298.87 | 411.25 | 546.71 |
| ... | ... | ... | ... | ... | ... | ... | ... |

although the distribution of the number of instances per horizon is not uniform (short horizons have more instances than long horizons).

## 4 Methods

The approach to predict GHI is to combine a set of $n$ predictors by means of machine learning models, which the aim to improve the final prediction for every forecast horizon. Therefore, the GHI can be described by Eq. 2, where machine learning model $f$ is used to combine several predictors $P_i$.:

$$ghi = f(P_1, P_2, ..., P_n) \tag{2}$$

In this work there are four predictors available which are used as inputs for the machine learning algorithm. The predictors combined in this work have been described in detail in Sect. 2. The machine learning method for finding $f$ is the extreme gradient boosting tree ensemble (*xgbtree*) [18]. This decision has been taken after comparing preliminary results with random forests and support vector machines. *Xgbtree* displayed a good performance, while at the same time it is a very fast and efficient implementation. In any case, other methods could have been used within the schema proposed in this article.

Given that each predictor performance depends on the horizon, three different approaches for dealing with horizons have been studied: general, horizon-individual, and horizon-group. They are described in the following subsections.

### 4.1 General Model

The first approach constructs a model that minimizes error for all horizons considered together. This is achieved by combining data from the different horizons (i.e. excluding the horizon column in Table 1), and training a single model $f$ from the joint dataset. Equation 3 shows how model $f$ can be used for forecasting GHI at time $t$ for horizon $h$. It can be seen that $f$ is common for all horizons.

$$ghi(t+h) = f(P_1(t,h), P_2(t,h), ..., P_n(t,h)) \tag{3}$$

### 4.2 Horizon-Individual Model

The second approach builds a different machine learning model $f_h$ for each horizon $h$. Each $f_h$ is trained using data from each horizon $h$ only. The end result will be a set of 25 machine learning models specialized in predicting GHI for every single horizon. There is a model for horizon 15, able to predict GHI at 15 min forward in time, another model for horizon 30, able to predict GHI at 30 min forward in time, and so on. Equation 4 shows how to use models $f_h$ for making GHI forecasts at time $t$ for each horizon $h$.

$$ghi(t+h) = \begin{cases} f_{15}(P_1(t,15), ..P_n(t,15)), & h \equiv 15 \\ f_{30}(P_1(t,30), ..P_n(t,30)), & h \equiv 30 \\ ... & \\ f_{360}(P_1(t,360), ..P_n(t,360)), & h \equiv 360 \end{cases} \tag{4}$$

### 4.3 Horizon-Group Model

This last approach builds a set of models, this time by using groups of horizons instead of individual horizons (as in the previous approach). Given that some predictors work better for close horizons (Smart Persistence and Satellite) and others for medium or long horizons (WRF-Solar), the aim is to identify groups of horizons for which some predictors are better than others. The advantage over the horizon-individual approach is that now each group of horizons have more data for training. The end result will be $g$ machine learning models, each one specialized in predicting GHI for a horizon group, where $g$ is the number of groups. Each group model is trained using data from horizons belonging to that group only. This model is represented by Eq. 5, where the $p_i$'s represent the partition points in the horizon range and the $15 \leq h \leq p_1$, $p_1 \leq h \leq p_2$, $\ldots$, $p_{g-1} \leq h \leq 360$ are the $g$ horizon groups.

$$ghi(t+h) = \begin{cases} f_1(P_1(t,h),..P_n(t,h)), & 15 \leq h \leq p_1 \\ f_2(P_1(t,h),..P_n(t,h)), & p_1 \leq h \leq p_2 \\ ... \\ f_g(P_1(t,h),..P_n(t,h)), & p_{g-1} \leq h \leq 360 \end{cases} \tag{5}$$

After visual analysis and taking into account the performance of the forecasting models for the different horizons, three groups ($g = 3$) have been used, although larger values could be considered at the expense of computational cost and diminishing the number of data for each group.

In order to decide the actual location of $p_1$ and $p_2$, a greedy search has been implemented. Starting from some initial values for $p_1$ and $p_2$, all combinations of neighboring points are explored. The set of neighbors of ($p_1$, $p_2$) is considered to be ($p_1 \pm 0$, 15, 30, $p_2 \pm 0$, 15, 30). Table 2 shows those neighboring points. For each partition explored, three different models are obtained (one per group), each one trained with data from each horizon group and evaluated on validation sets. Out of all the neighbors, the four combinations with lowest errors are kept.

The reason for keeping more than one ($p_1$, $p_2$) combination is to avoid falling into local minima. The combination with lowest error, and still unvisited by the algorithm, is the next explored combination of points. $p_1$ and $p_2$ are then updated to those locations that minimize the average validation error. This process is repeated until there is an empty list of combinations or the 4 best possible combinations have already been explored. At the end of the search, the algorithm chooses the partition (i.e. combination of $p_1$ and $p_2$) with the best error found throughout all the search.

The method is detailed in Algorithm 1. Line 1 creates a table (*VisitedTable*) that stores information about all combinations ($p_1$, $p_2$) explored by the algorithm. That is, *Accuracy* is the average validation error for each particular combination; *visited* informs whether this combination has already been explored; and *lastVisit* marks one of the pairs as the one that should be selected for expanding the neighbors.

**Table 2.** Neighbors for any $p_1$ and $p_2$

| p1 | p2 | p1 | p2 |
|---|---|---|---|
| $p_1 - 30$ | $p_2 - 30$ | $p_1 - 15$ | $p_2 - 15$ |
| $p_1 - 30$ | $p_2 - 15$ | ... | ... |
| $p_1 - 30$ | $p_2$ | $p_1$ | $p_2 - 15$ |
| $p_1 - 30$ | $p_2 + 15$ | $p_1$ | $p_2 + 15$ |
| $p_1 - 30$ | $p_2 + 30$ | ... | ... |
| $p_1 + 15$ | $p_2 - 30$ | $p_1 + 30$ | $p_2 + 30$ |

Loop in lines 2–17 runs while it is possible to find combinations that improve the error. In line 3 the last state is retrieved from *VisitedTable* (i.e. *lastVisit* == TRUE). After being retrieved, *lastVisit* will be set to FALSE. In line 4, the $(p_1, p_2)$ combination is retrieved from *LastState* and all possible neighbors are calculated in line 5 (*AllNeighbours*$(p_1, p_2)$). The loop in lines 6–12 checks every pair of points from the list of neighbors (see Table 2) previously expanded (*NewPointList*). If the pair has already been visited, it can be extracted out of the *VisitedTable*. Otherwise, its performance is computed (*Accuracy*$(p_1, p_2)$) in line 10. At the end of the loop, *PairErrors* contains the performance of all neighbors. Line 13 selects the best four combinations (*BestErrors*), out of which the best unvisited pair is finally selected (line 14). This best pair is marked with *lastVisit* = TRUE, so that it will be selected in the next iteration for computing neighbors. All information regarding new explored pairs and their respective errors is included into *VisitedTable* (line 15). Exploration will continue, as far as at least one of the four best pairs was unvisited (*BestError* not empty, line 18). Once the termination condition is satisfied, the best pair $(p_1, p_2)$ from *VisitedTable* is returned.

**Algorithm 1.** Horizon-group greedy search process

1: $VisitedTable \leftarrow$ Table($p_1, p_2$,Accuracy($p_1, p_2$), $visited, lastVisit$)
2: **while** *continue* **do**
3: $LastState \leftarrow lastVisit$ in $VisitedTable$ is $TRUE$
4: $p_1, p_2 \leftarrow p_1$ and $p_2$ in $LastState$
5: $NewPointList \leftarrow$ AllNeighbours($p_1, p_2$)
6: **for** each $PointPair$ in $NewPointList$ **do**
7: **if** $p_1, p_2$ exists in $VisitedTable$ **then**
8: $PairError \leftarrow error$ in $VisitedTable$
9: **else**
10: $PairError \leftarrow$ Accuracy($p_1, p_2$)
11: **end if**
12: **end for**
13: $BestErrors \leftarrow$ select 4 lowest from $PairErrors$
14: $BestError \leftarrow$ select lowest and !$visited$ from $BestErrors$. Mark this pair with $lastVisit \leftarrow TRUE$
15: update $VisitedTable$ with $NewPointList$ and $PairErrors$
16: $continue \leftarrow$ !($BestError$ is empty)
17: **end while**
18: $BestSeparation \leftarrow$ min $error$ from $VisitedTable$

## 5 Experimentation

The aim of the experimentation is to compare the skill of each method pre-sented here to predict GHI at each horizon from $h = 15$ to $h = 360$. There are seven different methods to compare, four of them are the predictors (WRF-Solar, CIADCast, Smart Persistence and Satellite) and the other three are the different machine learning approaches proposed in this work (General, Horizon-individual and Horizon-group).

### 5.1 Methodology

The dataset is divided into two subsets: training and test. The former is made up of the 21 first days of each month (3 weeks of data), the latter is for the test set, that will be used to evaluate the trained model. The training set itself is divided into a model-training set and a validation set. The first one contains the 14 first days of the month and it is used for training the models. The validation set is used for hyper-parameter tuning and to guide the search process for the horizon-group approach and to select the best horizon groups (see Sect. 4.3). The metric used for comparison purposes is the normalized root mean square error, which is calculated in Eq. 6.

$$nRMSE = \frac{\sqrt{\sum (x_i - o_i)^2 / N}}{\sum (o_i) / N} \tag{6}$$

where $x_i$ is a prediction, $o_i$ is an observation and $N$ is the number of samples. nRMSE is calculated for each horizon. The global nRMSE is the mean of all horizon nRMSE values.

### 5.2 Results

In Table 3 the global test nRMSE for each model is shown. The first four rows refer to the predictors. The fifth row displays the error that would be obtained if for each horizon, the best predictor would be selected (called optimal selection in Table 3). Given that this selection is done using the test set, it could not be applied in practice. It is provided only for comparison purposes with the machine learning approaches. It can be seen that all machine learning approaches have better global error than any of the predictors or even their optimal selection. On average, WRF-Solar is the most reliable predictor (0.2603 global nRMSE), with Smart Persistence being the second best (0.2837 nRMSE). Observing the machine learning blending approaches, the most accurate prediction method is the horizon-group approach with a global nRMSE of 0.227. After applying Algorithm 1, horizon groups are 15–60, 75–270 and 285–360.

Table 4 shows the nRMSE broken down by horizon (from $h = 15$ to $h = 360$) for each of the different predictors and approaches. This information is also displayed in Figs. 1 and 2. Figure 1 compares the predictors. Under an hour, CIADCast is the best predictor available, then Smart Persistence is best until

**Table 3.** Global nRMSE

| Method | nRMSE | Method | nRMSE |
|---|---|---|---|
| Smart persistence | 0.2837 | General | 0.2291 |
| Satellite | 0.3117 | Horizon-individual | 0.2312 |
| CIADCast | 0.3039 | **Horizon-group** | **0.227** |
| WRF-Solar | 0.2603 | | |
| Optimal selection | 0.2543 | | |

90 min, when WRF-Solar starts being the best model from that point onwards. There are a couple of times where WRF-Solar is worse, at 105 and 150 min. Interestingly, for WRF-Solar, the nRMSE decreases as $h$ increases, although after 285 min the error starts increasing again.

**Table 4.** nRMSE by horizon

| h= | 15 | 30 | 45 | 60 | 75 | 90 | 105 | 120 | 135 | 150 | 165 | 180 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| General | 0.197 | 0.205 | 0.211 | 0.217 | 0.226 | 0.229 | 0.22 | 0.217 | 0.219 | 0.215 | 0.216 | 0.225 |
| h-individual | 0.208 | 0.21 | 0.213 | 0.221 | 0.234 | 0.229 | 0.22 | 0.217 | 0.225 | 0.223 | 0.22 | 0.227 |
| h-group | 0.203 | 0.208 | 0.209 | 0.212 | 0.228 | 0.23 | 0.222 | 0.218 | 0.218 | 0.217 | 0.216 | 0.226 |
| CIADCast | 0.229 | 0.244 | 0.252 | 0.263 | 0.283 | 0.287 | 0.281 | 0.29 | 0.297 | 0.282 | 0.273 | 0.270 |
| Satellite | 0.229 | 0.248 | 0.255 | 0.26 | 0.288 | 0.298 | 0.291 | 0.28 | 0.268 | 0.267 | 0.275 | 0.289 |
| SmartPer | 0.245 | 0.252 | 0.255 | 0.259 | 0.274 | 0.28 | 0.265 | 0.271 | 0.281 | 0.281 | 0.276 | 0.293 |
| WRFSolar | 0.284 | 0.272 | 0.279 | 0.277 | 0.284 | 0.274 | 0.269 | 0.264 | 0.262 | 0.268 | 0.266 | 0.265 |
| h= | 195 | 210 | 225 | 240 | 255 | 270 | 285 | 300 | 315 | 33 | 345 | 360 |
| General | 0.22 | 0.221 | 0.22 | 0.233 | 0.236 | 0.234 | 0.232 | 0.246 | 0.261 | 0.275 | 0.264 | 0.260 |
| h-individual | 0.226 | 0.227 | 0.223 | 0.226 | 0.241 | 0.234 | 0.225 | 0.241 | 0.264 | 0.271 | 0.256 | 0.269 |
| h-group | 0.22 | 0.224 | 0.217 | 0.231 | 0.234 | 0.23 | 0.227 | 0.244 | 0.256 | 0.257 | 0.255 | 0.248 |
| CIADCast | 0.264 | 0.291 | 0.301 | 0.312 | 0.356 | 0.348 | 0.337 | 0.353 | 0.358 | 0.372 | 0.37 | 0.381 |
| Satellite | 0.298 | 0.306 | 0.311 | 0.346 | 0.332 | 0.345 | 0.348 | 0.353 | 0.382 | 0.391 | 0.422 | 0.4 |
| SmartPer | 0.288 | 0.279 | 0.275 | 0.294 | 0.296 | 0.288 | 0.275 | 0.304 | 0.318 | 0.333 | 0.314 | 0.313 |
| WRFSolar | 0.263 | 0.253 | 0.251 | 0.237 | 0.229 | 0.242 | 0.231 | 0.237 | 0.254 | 0.26 | 0.264 | 0.263 |

In Fig. 2 the machine learning approaches (General, Horizon-individual and Horizon-group) are compared to the optimal selection of predictors mentioned above. All machine learning combination of predictors are better than the original predictors up to 255 min. At that point, it is difficult to observe a difference respect to the predictors (WRF-Solar being the most accurate one at those horizons, as shown in Fig. 1). The machine learning approaches show minor differences. First, the horizon-individual model is consistently worse during the early horizons, while both the general and horizon-group models are similar in their performance. However when the 255 min horizon is reached, they become

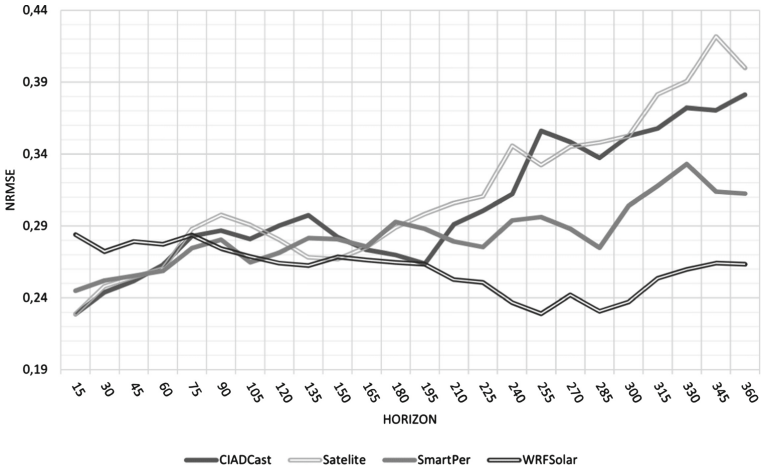


**Fig. 1.** Predictor performance at different horizons.

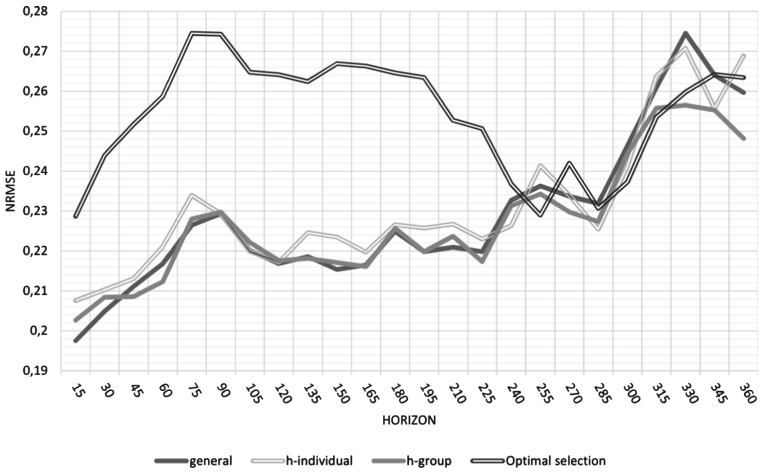


**Fig. 2.** Performance of the three machine learning approaches at different horizons.

harder to differentiate and behave similarly. At long horizons, starting at the 330 horizon, the horizon-group approach outperforms the other approaches and predictors. All machine learning models consistently increase in nRMSE as the horizon increases.

## 6 Conclusions

In this paper machine learning methods have been tested in order to combine GHI forecasting models (Smart Persistence, Satellite, CIADCast and WRF-Solar) as inputs to Xgboost, with the aim of improving predictions in horizons

from 15 to 360 min. Three approaches have been studied: a general approach that disregards horizon information, a horizon-individual approach that builds a Xgboost model for each horizon, and a horizon-group approach that separates horizons in three groups and builds a Xgboost model for each group. Experimental results show a great accuracy improvement over the predictors on short time horizons, and equivalent performance to the best predictor on further horizons. The general, horizon-individual and horizon-group models display similar performance, although for far horizons, the latter displays a better performance.

Overall, the final results are satisfactory, showing that there is a lot of margin for improvement in the field of solar forecasting using machine learning. In the future, we would like to improve results further by including additional features, such as other predictors or historical solar radiation data. It can be also interesting training different models for different seasons or different weather regimes. Another interesting research work would be automatically selecting subsets of predictors for each horizons or groups of horizons.

**Acknowledgments.** The authors are supported by the Spanish Ministry of Economy and Competitiveness, projects ENE2014-56126-C2-1-R and ENE2014-56126-C2-2-R and FEDER funds. Some of the authors are also funded by the Junta de Andalucía (research group TEP-220).

## References

1. Inman, R.H., Pedro, H.T.C., Coimbra, C.F.M.: Solar forecasting methods for renewable energy integration. Prog. Energy Combust. Sci. **39**(6), 535–576 (2013)
2. Witten, I.H., Frank, E., Hall, M.A., Pal, C.J.: Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann (2016)
3. Voyant, C., Notton, G., Kalogirou, S., Nivet, M.-L., Paoli, C., Motte, F., Fouilloy, A.: Machine learning methods for solar radiation forecasting: a review. Renew. Energy **105**, 569–582 (2017)
4. Terren-Serrano, G.: Machine learning approach to forecast global solar radiation time series (2016)
5. Chu, Y., Pedro, H.T.C., Coimbra, C.F.M.: Hybrid intra-hour DNI forecasts with sky image processing enhanced by stochastic learning. Sol. Energy **98**, 592–603 (2013)
6. Mathiesen, P., Collier, C., Kleissl, J.: A high-resolution, cloud-assimilating numerical weather prediction model for solar irradiance forecasting. Sol. Energy **92**, 47–61 (2013)
7. Salcedo-Sanz, S., Casanova-Mateo, C., Pastor-Sánchez, A., Sánchez-Girón, M.: Daily global solar radiation prediction based on a hybrid coral reefs optimization-extreme learning machine approach. Sol. Energy **105**, 91–98 (2014)
8. Alonso, J., Batlles, F.J.: Short and medium-term cloudiness forecasting using remote sensing techniques and sky camera imagery. Energy **73**, 890–897 (2014)
9. Lorenz, E., Kühnert, J., Heinemann, D.: Short term forecasting of solar irradiance by combining satellite data and numerical weather predictions. In: Proceedings of 27th European Photovoltaic Solar Energy Conference, Valencia, Spain, pp. 4401–440 (2012)

10. Huang, J., Korolkiewicz, M., Agrawal, M., Boland, J.: Forecasting solar radiation on an hourly time scale using a coupled autoregressive and dynamical system (cards) model. Sol. Energy **87**, 136–149 (2013)
11. Lu, S., Hwang, Y., Khabibrakhmanov, I., Marianno, F.J., Shao, X., Zhang, J., Hodge, B.M., Hamann, H.F.: Machine learning based multi-physical-model blending for enhancing renewable energy forecast - improvement via situation dependent error correction. In: European Control Conference (ECC), pp. 283–290 (2015)
12. Mazorra Aguiar, L., Pereira, B., Lauret, P., Díaz, F., David, M.: Combining solar irradiance measurements, satellite-derived data and a numerical weather prediction model to improve intra-day solar forecasting. Renew. Energy **97**, 599–610 (2016)
13. Sánchez, I.: Adaptive combination of forecasts with application to wind energy. Int. J. Forecast. **24**(4), 679–693 (2008)
14. Rigollier, C., Bauer, O., Wald, L.: On the clear sky model of the ESRA-European Solar Radiation Atlas-with respect to the heliosat method. Sol. Energy **68**(1), 33–48 (2000)
15. Nonnenmacher, L., Coimbra, C.F.M.: Streamline-based method for intra-day solar forecasting through remote sensing. Sol. Energy **108**, 447–459 (2014)
16. Arbizu-Barrena, C., Ruiz-Arias, J.A., Rodríguez-Benítez, F.J., Pozo-Vézquez, D., Tovar-Pescador, J.: Short-term solar radiation forecasting by advecting and diffusing MSG cloud index. Sol. Energy **155**, 1092–1103 (2017)
17. Jimenez, P.A., Hacker, J.P., Dudhia, J., Haupt, S.E., Ruiz-Arias, J.A., Gueymard, C.A., Thompson, G., Eidhammer, T., Deng, A.: WRF-solar: description and clear-sky assessment of an augmented NWP model for solar power prediction. Bull. Am. Meteorol. Soc. **97**(7), 1249–1264 (2015)
18. Chen, T., Guestrin, C.: XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2016, pp. 785–794. ACM (2016)
19. Urraca, R., Antonanzas, J., Antonanzas-Torres, F., Martinez-de-Pison, F.J.: Estimation of daily global horizontal irradiation using extreme gradient boosting machines. In: Graña, M., López-Guede, J.M., Etxaniz, O., Herrero, Á., Quintián, H., Corchado, E. (eds.) ICEUTE/SOCO/CISIS-2016. AISC, vol. 527, pp. 105–113. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-47364-2 11
20. Fan, J., Wang, X., Lifeng, W., Zhou, H., Zhang, F., Xiang, Y., Xianghui, L., Xiang, Y.: Comparison of support vector machine and extreme gradient boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: a case study in China. Energy Conv. Manag. **164**, 102–111 (2018)
21. Rigollier, C., Lefèvre, M., Wald, L.: The method heliosat-2 for deriving shortwave solar radiation from satellite images. Sol. Energy **77**(2), 159–169 (2004)
22. Skamarock William, C., Joseph, B.K., Jimy, D., David, O.G., Dale, M.B., Michael, G.D., Huang, X.Y., Wang, W., Jordan, G.P.: A description of the advanced research WRF version 3. NCAR technical note, 126 (2008)
23. Long, C.N., Dutton, E.G.: BSRN global network recommended QC tests, v2. x. (2010)