

Article

Solar Irradiance Forecasting Using Dynamic Ensemble Selection

Domingos S. de O. Santos, Jr.¹, Paulo S. G. de Mattos Neto^{1,†}, João F. L. de Oliveira^{2,†}, Hugo Valadares Siqueira^{3,†}, Tathiana Mikamura Barchi^{3,†}, Aranildo R. Lima^{4,†}, Francisco Madeiro^{2,†}, Douglas A. P. Dantas^{2,†}, Attilio Converti^{5,*,†}, Alex C. Pereira^{6,†}, José Bione de Melo Filho^{6,†} and Manoel H. N. Marinho^{2,†}

¹ Centro de Informática, Universidade Federal de Pernambuco, Recife 50740-560, PE, Brazil; dsosj@cin.ufpe.br (D.S.d.O.S.J.); psgmn@cin.ufpe.br (P.S.G.d.M.N.)

² Escola Politécnica de Pernambuco, Universidade de Pernambuco, Recife 50720-001, PE, Brazil; fausto.lorenzato@upe.br (J.F.L.d.O.); madeiro@poli.br (F.M.); douglasazevedo839@gmail.com (D.A.P.D.); marinho75@poli.br (M.H.N.M.)

³ Graduate Program in Computer Sciences, Federal University of Technology—Paraná, Ponta Grossa 84017-220, PR, Brazil; hugosiqueira@utfpr.edu.br (H.V.S.); tathianabarchi@alunos.utfpr.edu.br (T.M.B.)

⁴ Aquatic Informatics, Vancouver, BC V6E 4M3, Canada; proj.brain@gmail.com

⁵ Department of Civil, Chemical and Environmental Engineering, University of Genoa, Via Balbi, 5, 16126 Genoa, Italy

⁶ CHESF, (DEGS), Recife 50761-901, PE, Brazil; alexcp@chesf.gov.br (A.C.P.); jbonef@chesf.gov.br (J.B.d.M.F.)

* Correspondence: converti@unige.it

† These authors contributed equally to this work.



Citation: de O. Santos, D.S., Jr.; de Mattos Neto, P.S.G.; de Oliveira, J.F.L.; Siqueira, H.V.; Barchi, T.M.; Lima, A.R.; Madeiro, F.; Dantas, D.A.P.; Converti, A.; Pereira, A.C.; et al. Solar Irradiance Forecasting Using Dynamic Ensemble Selection. *Appl. Sci.* **2022**, *12*, 3510. <https://doi.org/10.3390/app12073510>

Academic Editor: Alireza Dehghanisanij

Received: 6 March 2022

Accepted: 27 March 2022

Published: 30 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Solar irradiance forecasting has been an essential topic in renewable energy generation. Forecasting is an important task because it can improve the planning and operation of photovoltaic systems, resulting in economic advantages. Traditionally, single models are employed in this task. However, issues regarding the selection of an inappropriate model, misspecification, or the presence of random fluctuations in the solar irradiance series can result in this approach underperforming. This paper proposes a heterogeneous ensemble dynamic selection model, named HetDS, to forecast solar irradiance. For each unseen test pattern, HetDS chooses the most suitable forecasting model based on a pool of seven well-known literature methods: ARIMA, support vector regression (SVR), multilayer perceptron neural network (MLP), extreme learning machine (ELM), deep belief network (DBN), random forest (RF), and gradient boosting (GB). The experimental evaluation was performed with four data sets of hourly solar irradiance measurements in Brazil. The proposed model attained an overall accuracy that is superior to the single models in terms of five well-known error metrics.

Keywords: solar irradiance forecasting; heterogeneous ensemble dynamic selection model; neural networks

1. Introduction

Energy production plays a major role in modern societies, impacting the economy and the development of several countries [1]. Over the past years, energy production has primarily relied on the employment of fossil fuels, which are still the main energy production source [2]. However, fossil fuels are one of the primary emission sources of CO₂, which also contributes to greenhouse effects and global warming [3]. Moreover, it is a non-renewable energy source, and it is being consumed more than it is produced.

Alternatively, solar energy is an important renewable energy source that produces on average 1.74×10^{17} W in 1 year, which meets energy demands worldwide [4]. Solar energy is considered a promising alternative to fossil fuels. However, it has intermittent and volatile characteristics due to factors, such as precipitation, temperature, wind velocity, and

atmospheric pressure [5,6]. These volatile characteristics can produce voltage fluctuations, possibly leading to instability in the power grid if they are not taken into account [7]. Moreover, the integration of traditional power systems with renewable power sources requires a precise balance between the demand and supply of power [6,8]. If the energy demand surpasses the energy supply, destabilization of the power grid can occur, resulting in power quality degradation or even grid damage.

Conversely, if the demand is lower than the supply, energy will eventually be lost, and wastage can occur because of the possible high costs associated with the storage of solar energy [7]. Considering the prediction of the demand and supply, the strategic section of energy plants decides whether energy is purchased. Prediction errors can lead to extra charges for the final consumer [9]. Therefore, accurate forecasting of solar irradiance plays a major role in power grid systems, estimation of reserves, scheduling, congestion management, use of produced energy in the energy market, and reduction of energy production costs [5,9,10].

Traditional statistical methods, such as the autoregressive integrated moving average (ARIMA) and seasonal ARIMA (SARIMA), are popular due to their simplicity and a well-defined Box and Jenkins methodology is used to select a suitable configuration of the model [11]. However, ARIMA models assume a linear correlation structure between forecasts and past data, which may underperform in the presence of nonlinear patterns. Considering that real-world data is often composed of a combination of linear and nonlinear patterns [12], and solar irradiance data present intermittent and volatile characteristics, the ARIMA models may perform worse than other methods in the literature [5,13].

In contrast to linear statistical models, forecasting methods based on machine learning are more suitable for mapping nonlinear patterns [13] in data. Machine learning methods, such as artificial neural networks (ANNs) and support vector machines (SVMs), are flexible data-driven models and have been used in the irradiation forecasting context [13]. However, machine learning-based models can also present problems related to model misspecification, overfitting, and underfitting [14–16], which leads to a poor generalization capacity [17].

To overcome the limitations of single and machine learning models, one possible strategy is to combine the forecasts of several models through the employment of ensembles to improve the system's accuracy [18]. This strategy is promising because it reduces the risk of selecting an inappropriate model. In addition, it has the flexibility of the use of different combination strategies and different model generation methodologies [19]. Ensemble learning strategies are often composed of three stages: generation, pruning, and combination [19,20]. The ensembles can be classified according to the base learners being homogeneous or heterogeneous. Homogeneous ensembles are composed of the same learning algorithm while heterogeneous ones are often composed of different ones [19]. The main issue is that ensembles must be as accurate and diverse as possible, and the diversity control must be performed in the generation stage. In this sense, heterogeneous ensemble approaches are expected to produce more diverse models, considering the different learning algorithms of the base models [19,21].

In the context of solar irradiation forecasting, several ensemble learning approaches have employed a static ensemble of forecasters through time [13,22–25]; however, considering the volatile and dynamic characteristics of solar irradiance, static ensemble approaches can present inaccurate results since the best model could change over time. This dynamic characteristic is expected because solar irradiance may be affected by weather conditions [4]. Therefore, a dynamic model selection strategy [14] could improve the results over static ones.

In this work, an ensemble based on dynamic selection is proposed for solar irradiance forecasting. The proposed model, referred to as heterogeneous ensemble dynamic selection (HetDS), chooses the most appropriate set of models from a pool of forecasters comprising ARIMA, support vector regression (SVR), multilayer perceptron neural network (MLP), extreme learning machine (ELM), deep belief network (DBN), random forest (RF), and gradient boosting (GB). The best model in the pool is selected based on a local accuracy

procedure, which defines a region of competence of size k . Experiments were conducted using the base models and static ensemble approaches, and the results demonstrate that the proposed method achieves the overall best results, considering the root mean square error (RMSE) and mean absolute percentage error (MAPE), mean absolute error (MAE), average relative variance (ARV), and index of agreement (IA), based on evidence provided by Friedman–Nemenyi [26] hypothesis testing. Moreover, the proposed model presents the following advantages:

- Reduces the risk of selecting an inappropriate model;
- Dynamically searches for the most suitable forecaster to predict a given local pattern in a solar irradiance series;
- It is an agnostic model since other forecasting models can be explored in the pool;
- Increases the generalization capacity of the system.

The rest of the paper is divided as follows: the related works are discussed in Section 2; Section 3 describes the models used to perform solar irradiance forecasting; Section 4 presents our proposal to solve the same problem; Section 5 shows the details of the database used, computational results, and a discussion; and finally, Section 6 presents the main conclusions and future works.

2. Related Works

Considering the volatility of irradiation data and the importance of obtaining accurate results in power grid systems, several models have been employed to improve the accuracy of forecasts. In general, forecasting models for solar irradiance are classified into physical models, empirical, statistical, and machine learning models [27]; however, hybridization between these classes of models is also possible and tends to improve the results when compared with single model approaches [28,29].

Physical models, such as numerical weather prediction (NWP), are complex structures that consider several aspects of the environment and the irradiance data [30]. In contrast, empirical models are less complex and often employ linear and nonlinear regressions to perform estimates [31]. However, empirical models may present limited accuracy.

Statistical models, such as the autoregressive integrated and moving average model (ARIMA) perform temporal mappings over past data and produce forecasts, and have been used in several irradiance forecasting applications. Shadab et al. [32] employed a multiplicative seasonal ARIMA model to forecast monthly average isolation data considering different sky conditions. Voyant et al. [33] employed an ARIMA to predict global irradiation. In this work, the model is used in conjunction with MLP networks, and the proposed approach is compared with single models.

Machine learning models are not only capable of performing nonlinear mappings in the data but also present some flexibility regarding noisy data and tend to achieve an improved performance since real-world data often composed of nonlinear patterns. The employment of SVM models in irradiation tasks has achieved promising results. Chen and Li [34] used the SVM models, considering exogenous variables, such as sunshine ratio, maximum and minimum air temperature, relative humidity, and atmospheric water vapor pressure. Moreover, Bendiek et al. [35] employed an optimization method for several models, such as SVM and MLP, and then performed experimental comparisons. The results indicated that SVM with the proposed optimization achieved the best results.

Random forests (RFs) are part of the ensemble class of models, where several regression trees are employed and combined to achieve the final result. This approach is interesting to solar radiation forecasting since the data may present high volatility, and it mitigates the chance of selecting an inappropriate model. Srivastava et al. performed several comparisons among machine learning models, such as classification and regression trees (CART) and random forests, using data from India, and the RF model obtained the best results. Huang et al. [36] compared several methods using exogenous data and concluded that the RF method achieved the smallest errors. Another algorithm employed in irradiation forecasting tasks is gradient boosting (GB), which is based on a tree-based ensemble learning

technique. Park et al. [37] employed LightGBM to perform multi-step ahead forecasting in irradiation data. Moreover, Fan et al. [38] used an extreme gradient boosting algorithm and SVM to predict daily global solar radiation using temperature and precipitation in humid subtropical climates.

The work of Elminir et al. [39] was one of the pioneering works that applied neural networks when considering solar irradiance components. In this investigation, the authors predicted infrared, ultraviolet, and global insolation using an MLP trained by the Levenberg–Marquardt algorithm. The database addressed was from Egypt. They also applied the trained network considering a database from Helwan and Aswan. In both cases, the accuracy was up to 90%.

MLP was also applied to generate synthetic daily solar radiation series using five exogenous variables as inputs: daily clear sky global radiation, cloud cover, temperature, water vapor, and ozone. They proposed that the model can generate good predictions for locations in which there are no ground measurements. The results proved the generalization capability of MLPs.

The paper from Salcedo-Sanz et al. [40] considers the outputs of the WRF meso-scale model as inputs of an ELM to predict solar irradiation. They proposed the application of the coral reefs optimization algorithm with species (CRO-SP) as a feature selection procedure to minimize the number of input variables. The same author proposed the use of coral reefs optimization (CRO) hybridized with an ELM to predict solar irradiation [41]. A binary-encoded CRO is used as a variable selection model. A comparison was performed considering the performances of the genetic algorithm replacing the CRO, and the SVR, MARS, and MLR as the predictor. The results revealed the viability of the proposed method.

The deep belief network was applied to perform the monthly solar forecasting task considering Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data in the work of Ghimire et al. [42]. The authors also evaluated the use of 15 feature selection algorithms among the filters, wrappers, and bio-inspired approaches. The computational results revealed that the model can overcome MLP, decision trees, RF, and GB. A similar architecture (functional DBN) was proposed together with embedding clustering for daily global solar radiation forecasting by Zang et al. [43]. The model could overcome the stand-alone DBN, functional DBN, and other ML-based models.

3. Background

This section describes the models used to perform solar irradiance forecasting.

3.1. Autoregressive and Moving Average Model

The autoregressive and moving average models (ARMA) belongs to the Box and Jenkins family of linear models [11]. While the AR part considers the lags of the series, the MA creates the output response addressing random shocks a_{t-p-j} [44], which are weighted by θ_j coefficients, as in Equation (1):

$$\hat{x}_t = \phi_1 x_{t-p} + \dots + \phi_p x_{t-p-p+1} - \theta_1 a_{t-p} - \dots - \theta_q a_{t-p-q+1} + a_t, \quad (1)$$

where ϕ_i and θ_j are the autorregressive and moving average coefficients, respectively; P is the number of steps ahead considering the direct approach; and $i = 1, 2, \dots, p$ and $j = 1, 2, \dots, q$, are free coefficients [44].

The most usual way to apply the model considers the random shocks a_t as equivalent to the residual set of the previous samples [45].

Finally, to adjust the free coefficients of the ARMA model, we used the maximum likelihood estimator (MLE), which is the usual way [11]. Note that this model can also be named the autoregressive and moving average model (ARMA) when no differentiations are performed in the time series [11].

3.2. Random Forest

The random forest (RF) algorithm was introduced by Breiman in 2001 [46]. It is a learning method used for classification and regression that operates by building multiple decision trees during the training and production of the final prediction.

The name came from random decision forests that were first proposed by Ho [47]. The method combines Breiman's idea of bagging and the random selection of features, independently introduced by Ho [47] and Amit and Geman [48], to build a collection of decision trees with a controlled variation.

The trees used in RF are based on binary partitioning trees. These trees partition the predictor space using a sequence of binary divisions on individual variables. The root node of a tree comprises the entire predictor space. The nodes that are not split are called terminals and form the final partition of this space. Therefore, each non-terminal node is divided into two descending nodes, according to the value of the predictor variables. A division is determined by a division point or a continuous predictor variable. When the points for the predictor are less than the split point, the decision is made to go to the left, and the remainder goes to the right [49].

3.3. Gradient Boosting

Decision trees have several important characteristics for forecasting tasks. However, decision trees' individual predictive power can be enhanced by forming a prediction committee, which is the central idea of boosting techniques.

Boosting is a generic strategy to improve the performance of any learning algorithm. The method was proposed to deal with pattern classification problems, with the introduction of the AdaBoost [50]. Subsequently, several generalizations emerged from the original strategy, among them, the gradient boosting algorithm [51], which can be applied to both classification and regression problems, using any differentiable objective function.

The so-called boosting methods are based on the principle of minimizing a cost function through the aggregation of multiple weak models (weak learners) to build a more robust model using the gradient method as a procedure, which comprises a systematic strategy for the construction of forecast committees [51].

In a boosting method, the performance evaluation of the models is performed sequentially. The criteria in the next step can be changed according to the evaluation in the previous step: if the first evaluator indicates a low score for a criterion, this score will be considered by the next evaluator. Thus, the final assessment is more accurate, and the process is more dynamic [52].

In the gradient boosting method, errors are minimized by the gradient descent algorithm. The proposal avoids local minimum points so that the second-best criterion receives the best rating by adjusting the criteria to capture the best performance [53].

3.4. Support Vector Regression

Support vector regression (SVR) [54] is based on the structural risk minimization, characterized by a convex optimization problem, generating a single global minimum. The objective is to find a function in the form of 2:

$$\left\{ f \mid f(x) = w^T x + b, w \in \mathbb{R}^d, b \in \mathbb{R} \right\} \quad (2)$$

where w is a weight vector with components determined by the regularized risk function, x is the input vector, and b is the bias.

In regression problems, the model provides continuous outputs to the training data, considering a maximum deviation of ε of the expected value [55]. However, the cost function is not computed for values inside the region limited by ε . In addition, the use of kernel functions makes it possible to perform nonlinear mappings in the data to a feature space in which a linear function is found.

3.5. Multilayer Perceptron

Multilayer perceptron (MLP) is a famous architecture of neural networks. It is a feedforward approach since the information present in the inputs flows in just one direction, from the input layer to the output layer. The MLP is a universal function approximator because it can map any continuous, nonlinear, limited, and differentiable function with a predetermined error threshold.

The most usual way to train the weights of MLP is the backpropagation algorithm [56], which uses the information of the gradient function to modify the values of the weights in order to reduce the output error. In this work, we address an MLP with three layers (input, hidden, and output layers), and a sigmoid function is employed as an activation function [57].

3.6. Extreme Learning Machines

The extreme learning machines (ELMs) are feedforward neural networks but contain just a single hidden layer [58]. The main difference between them lies in the training process. While the MLP uses an iterative process to adjust the weights of the network, the ELM just modifies the weights of the hidden layer.

To perform the training, a linear regression problem is considered. The proposers of the architecture, Huang et al. [59], used the Moore–Penrose generalized inverse operation since it simultaneously minimizes the norm of the output weight vector and the mean square error of the output response [60]. The authors also proved that the insertion of a new randomly generated neuron in the hidden layer led to a decrease in the output error.

3.7. Deep Belief Network

The deep belief network (DBN) is a deep learning model that was introduced by Hinton et al. in 2006 [61]. It is a probabilistic generative graphical model created to overcome the limitations of traditional models, such as slow convergence and convergence to a local minimum, among others.

A DBN is composed of restricted Boltzmann machines (RBMs), in which each one presents two layers, visible and hidden. In this case, each hidden layer acts as a visible layer to the next RBM. The first two layers present indirect connections that create an associative memory. On the other hand, the lower layers present direct connections [62].

The training process follows a greedy approach so that each layer is adjusted until the global optimum is found, starting from the visible layer of the RBM. Then, the hidden layer is tuned, and its outputs are used as inputs to the next RBM layer until the end of the network [63].

4. Proposed Method

The proposed model comprises two stages: (i) model generation and (ii) model selection. The model generation stage performs the training of all forecasting methods (ARIMA, GB, RF, SVR, DBN, ELM, MLP) to produce a pool of models. The model generation stage is crucial to the ensemble, considering that the models should be as accurate and diverse as possible. The accuracy is achieved through a search methodology over the space of the possible parameter configurations of all models, and diversity is achieved because different models with distinct learning algorithms are used in the pool. The model generation architecture is presented in Figure 1.

The dynamic model selection stage occurs as new test instances arrive at the proposed model, $HetDS(m,k)$ (source code available at: https://github.com/domingos108/solar_forecasting, accessed on 5 March 2022), standing for a heterogeneous ensemble dynamic selection. Based on the local performance, the proposed method selects the best set models with a size m from the pool for each new test instance. It defines a region of competence of a size k , composed of the k nearest instances in the validation set. Therefore, all models in the pool are evaluated in this region employing the RMSE metric, and the m models with the lowest RMSE are selected to perform the forecast of the test instance.

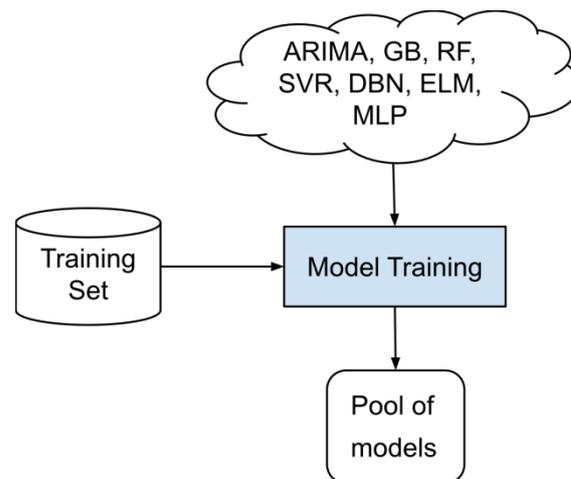


Figure 1. Model generation stage.

The Euclidean distance metric is employed to determine the k nearest instances on the validation set. Therefore, different models can be selected for each test instance to improve the forecast accuracy. The architecture of the dynamic model selection is presented in Figure 2. The configuration of the number of models m used in the dynamic ensemble and the number of nearest instances k plays a major role in the performance of the proposed system.

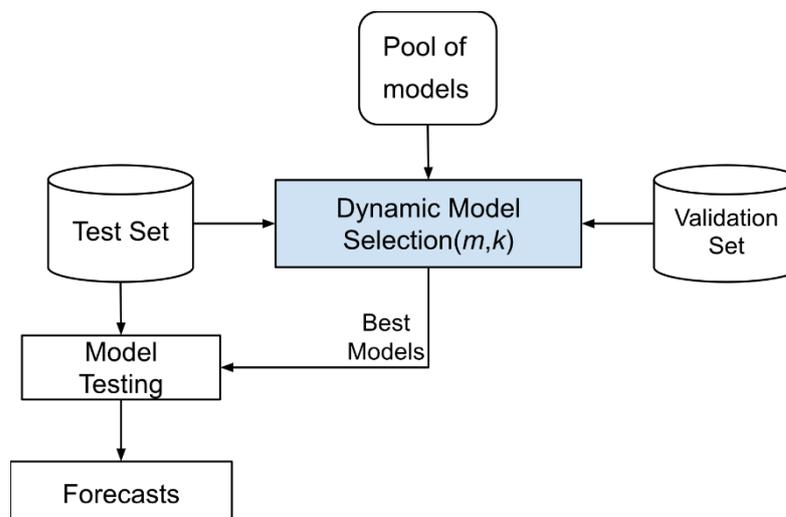


Figure 2. Dynamic model selection stage.

Lower and higher values of m increase the chance of selecting inappropriate model(s) since there is no guarantee that the best model(s) in the region of competence (validation set) will also demonstrate an improved generalization capacity on the test set. In the first scenario, when the value of m is near 1, the selected model may not achieve an improved performance on the test set, whereas in the second situation, when the value of m is near the size of the pool, it increases the chance of selecting inaccurate models for the ensemble combination.

Moreover, solar irradiation data presents dynamic characteristics due to exogenous variables. In addition, the value of k defines the size of the region of competence, which should represent the dynamics of the system. The models are combined through a median operator [64] since it is more robust to outliers. It is essential to mention that if the value of $m = 1$, then no combination is performed.

5. Experiments

The experiments were conducted on four datasets of solar plants from four cities in Brazil. All data correspond to hourly measures of solar irradiance, considering sunlight times, which is generally between 6:00 and 18:00 in Brazil. The data was split into three disjoint sets to perform training and model selection, and the last set was used for testing. To select the model parameters, a grid search procedure was conducted in the parameter space defined in Table 1. Regarding the ARIMA, p and q are the AR order and the MA order, respectively. Concerning the SVR, γ is the kernel parameter, C is the regularization factor, and ε is the maximum deviation of the observed value.

Table 1. Parameter configuration of the models and algorithms used.

Model	Parameters	Option
ARIMA	p,d,q	Hyndman Method [65]
MLP	Algorithm	Backpropagation
	Activation Function	Sigmoid
	Number of Hidden Layer Nodes	20, 50, 100
ELM	Algorithm	Moore-Penrose pseudo-inverse
	Activation Function	Hyperbolic Tangent
	Number of Hidden Layer Nodes	20, 50, 100, 200, 500
SVR	Kernel	RBF
	γ	0.1, 0.01, 0.001
	C	10, 100, 1000
	ε	0.1, 0.01, 0.001
GB	Number of estimators	50, 100, 200
	Max depth	5, 10, 15
	Max features	0.6, 0.8, 1
	Sub sample	0.6, 0.8, 1
	Learning rate	0.1, 0.3, 0.5
RF	Number of estimators	50, 100, 200
	Max depth	5, 10, 15
	Max features	0.6, 0.8, 1
DBN	Number of Hidden Layer Node	100, 200
	Learning rate RBM	0.01, 0.001
	Learning rate	0.01, 0.001

The ARIMA model employed a stepwise methodology to select an appropriate model [65]. To perform the predictions, we used 4758 hourly samples of the solar irradiation series so that 60% were used for training (the first 2856 samples), 20% for validation (951 samples), and 20% for testing (the last 951 samples). Moreover, the data sets used in the experiments were scaled to the range [0.1, 0.9] [66].

Several configurations of the proposed model were tested, using the m best models varying in the set 1, 3, 5 and the size of the region of competence varying in the set 5, 10, 20. Moreover, static ensemble strategies were also tested using mean and median combination operators using the same models from the pool.

5.1. Data Description

The databases are related to four major cities in Brazil: Florianópolis, Fortaleza, Salvador, and São Paulo.

Florianópolis is the capital of Santa Catarina state, located in the south of the country. It is a coastal town. Fortaleza is the capital of Ceará state, in the Brazilian northeast. It is also a coastal town relatively close to the equator line. Salvador is the capital of Bahia state, located in the Brazilian northeast region. São Paulo city is the major city in Brazil. It is almost on the Tropic of Capricorn. The red marks in Figure 3 represent the locations of these cities on a map of Brazil. Considering the cyclic behavior of solar irradiation, where maximum values are often achieved at 12:00 and minimum values at 6:00 and 18:00, all inputs were analyzed in a window size of 12. Moreover, this value can also be achieved through an autocorrelation plot.



Figure 3. Locations of the solar irradiation meters in Fortaleza, Florianópolis, Salvador, and São Paulo. The satellite map is from Google Maps (Map data©2020 Google; <https://www.google.com/maps/place/Brazil/>, accessed on 15 January 2022); the satellite is from Google Earth Pro (Mapdata©2020 Google; <https://www.google.com/maps/@-23.6815315,-46.8754814,10z>, accessed on 15 January 2022). The maps were edited with Microsoft Power Point (version 16.28-19081202).

As can be seen, the longest distance is between Fortaleza and Florianópolis, almost 3400 km. Due to this, it is clear that the sunlight may present different magnitudes along the year considering all cities. All databases are composed of 4758 hourly samples of solar irradiation (kJ/m^2). We considered the measures from 1 January 2020 to 31 December 2020, the whole year. However, we selected just the samples from 6:00 to 18:00 during each day.

Table 2 presents the geographical location of the solar irradiation stations and the descriptive statistics of the corresponding series. Note the difference in the altitude of São Paulo among the others, the only city that is not a coastal town. The acronyms STD and CV represent the standard deviation and coefficient of variation, respectively.

Table 2. Location of the solar irradiation stations and descriptive statistics regarding the mean and deviation of solar irradiation.

Station	Coordinates	Altitude	Mean	STD	CV
Florianopolis	−27.0253; −48.620096	4.87	1230.7	1101.2	0.895
Fortaleza	−3.815701; −38.537792	29.55	1223.9	883.18	0.722
Salvador	−13.551500; −8.505760	47.56	1348.2	1122.6	0.833
São Paulo	−23.496294; −6.620088	785.64	1355.4	1204.6	0.889

5.2. Evaluation Metrics

In this work, we consider the following metrics to perform a comparative analysis among the forecasting models presented, the root mean squared error (RMSE), the mean absolute percentage error (MAPE), the mean absolute error (MAE) the average relative variance (ARV), and the index of agreement (IA) [67,68], which are given by Equations (3)–(7), respectively:

$$\text{RMSE} = \frac{1}{N} \sum_{t=1}^N \left(\sqrt{(\hat{x}_t - x_t)^2} \right), \quad (3)$$

$$\text{MAPE} = \frac{100}{N} \sum_{t=1}^N \left| \frac{\hat{x}_t - x_t}{x_t} \right|, \quad (4)$$

$$\text{MAE} = \frac{1}{N} \sum_{t=1}^N |\hat{x}_t - x_t|, \quad (5)$$

$$\text{ARV} = \frac{\sum_{t=1}^N (\hat{x}_t - x_t)^2}{\sum_{t=1}^N (\hat{x}_t - \bar{x})^2}, \quad (6)$$

$$\text{IA} = 1 - \frac{\sum_{t=1}^N (\hat{x}_t - x_t)^2}{\sum_{t=1}^N (|\hat{x}_t - \bar{x}| - |x_t - \bar{x}|)^2}, \quad (7)$$

where N is the number of samples, x_t is the observed value in time t , \hat{x}_t is the predicted value, and \bar{x} is the mean of the series. The percentage difference (PD), given by Equation (8), is calculated to compare the performance of HetDS with the other approaches:

$$\text{PD} = \frac{\text{Metric}_{\text{model}} - \text{Metric}_{\text{HetDS}}}{\text{Metric}_{\text{model}}} \cdot 100, \quad (8)$$

where $\text{Metric}_{\text{HetDS}}$ and $\text{Metric}_{\text{model}}$ are the RMSE values attained by HetDS and the comparative models used in the experimental evaluation, respectively. The higher the PD value, the better the RMSE value obtained by HetDS with respect to the model under comparison.

5.3. Results

The performance of the forecasting models was assessed based on the test set (the last 951 samples) of each series. For the sake of simplicity, the proposed method is henceforth referred to as HetDS(m,k), standing for a heterogeneous ensemble dynamic selection with parameters m (best-selected models) and k (size of the region of competence). Table 3 presents the results achieved by the models described in Section 2, ARIMA, DBN, ELM, GB, MLP, RF, and SVR, and the variations of the proposed models. The values presented are the average of 10 independent simulations, and the best results are highlighted in bold.

Table 3. Performance of HetDS(m,k) and literature models in terms of RMSE, MAPE, MAE, ARV and IA for all data sets of solar irradiance. The best values attained in each data set are highlighted in bold.

Series	MODEL	RMSE	MAPE	MAE	ARV	IA
Fortaleza	ARIMA [32,33]	0.0824	21.21	0.0639	0.1181	0.9705
	RF [69,70]	0.0742	13.06	0.0563	0.1170	0.9739
	GB [37,38]	0.0746	13.63	0.0558	0.1155	0.9739
	SVR [34,71]	0.0629	12.87	0.0457	0.0695	0.9830
	MLP [39,72]	0.0676	14.65	0.0516	0.0853	0.9797
	ELM [40,41]	0.0718	12.95	0.0528	0.1037	0.9762
	DBN [42,43]	0.0660	11.89	0.0481	0.0839	0.9803
	HetDS(1,5)	0.0626	11.07	0.0447	0.0689	0.9831
	HetDS(1,10)	0.0617	10.73	0.0436	0.0672	0.9836
	HetDS(1,20)	0.0612	10.67	0.0434	0.0660	0.9838
	HetDS(3,5)	0.0607	10.77	0.0435	0.0665	0.9839
	HetDS(3,10)	0.0605	10.61	0.0429	0.0659	0.9841
	HetDS(3,20)	0.0600	10.57	0.0426	0.0646	0.9844
	HetDS(5,5)	0.0618	11.31	0.0448	0.0715	0.9830
	HetDS(5,10)	0.0616	11.23	0.0445	0.0710	0.9832
HetDS(5,20)	0.0616	11.22	0.0445	0.0708	0.9832	
Hetmean	0.0644	12.80	0.0479	0.0807	0.9812	
Hetmedian	0.0647	12.30	0.0477	0.0807	0.9811	
Florianópolis	ARIMA [32,33]	0.1024	24.74	0.0746	0.2135	0.9478
	RF [69,70]	0.0938	21.28	0.0671	0.1944	0.9547
	GB [37,38]	0.0942	20.36	0.0663	0.1876	0.9553
	SVR [34,71]	0.0962	20.43	0.0655	0.1754	0.9558
	MLP [39,72]	0.0956	20.76	0.0667	0.1850	0.9548
	ELM [40,41]	0.1002	22.44	0.0713	0.2067	0.9499
	DBN [42,43]	0.0962	22.00	0.0693	0.1970	0.9531
	HetDS(1,5)	0.0955	20.68	0.0665	0.1840	0.9550
	HetDS(1,10)	0.0955	20.82	0.0670	0.1839	0.9550
	HetDS(1,20)	0.0938	20.52	0.0655	0.1791	0.9565
	HetDS(3,5)	0.0937	20.03	0.0648	0.1800	0.9564
	HetDS(3,10)	0.0934	19.87	0.0645	0.1788	0.9567
	HetDS(3,20)	0.0931	19.87	0.0643	0.1780	0.9570
	HetDS(5,5)	0.0934	19.97	0.0646	0.1802	0.9566
	HetDS(5,10)	0.0933	19.97	0.0645	0.1801	0.9566
HetDS(5,20)	0.0933	19.95	0.0646	0.1805	0.9566	
Hetmean	0.0933	20.28	0.0648	0.1819	0.9564	
Hetmedian	0.0936	20.30	0.0650	0.1820	0.9563	
Salvador	ARIMA [32,33]	0.1051	27.15	0.0788	0.2739	0.9398
	RF [69,70]	0.0954	19.48	0.0657	0.2102	0.9524
	GB [37,38]	0.0973	20.18	0.0668	0.2190	0.9504
	SVR [34,71]	0.0902	17.12	0.0579	0.1581	0.9610
	MLP [39,72]	0.0900	17.95	0.0601	0.1678	0.9599
	ELM [40,41]	0.0972	20.48	0.0666	0.1866	0.9541
	DBN [42,43]	0.0918	19.38	0.0632	0.1822	0.9573
	HetDS(1,5)	0.0935	18.28	0.0619	0.1804	0.9568
	HetDS(1,10)	0.0945	18.37	0.0626	0.1855	0.9557
	HetDS(1,20)	0.0943	18.03	0.0619	0.1842	0.9559
	HetDS(3,5)	0.0907	17.60	0.0599	0.1743	0.9589
	HetDS(3,10)	0.0903	17.55	0.0597	0.1730	0.9592
	HetDS(3,20)	0.0902	17.30	0.0592	0.1714	0.9594
	HetDS(5,5)	0.0896	17.74	0.0596	0.1720	0.9597
	HetDS(5,10)	0.0894	17.64	0.0593	0.1705	0.9600
HetDS(5,20)	0.0895	17.64	0.0593	0.1706	0.9599	
Hetmean	0.0903	18.53	0.0609	0.1792	0.9585	
Hetmedian	0.0898	18.05	0.0602	0.1740	0.9593	

Table 3. Cont.

Series	MODEL	RMSE	MAPE	MAE	ARV	IA
São Paulo	ARIMA [32,33]	0.2138	61.59	0.1456	1.2968	0.7355
	RF [69,70]	0.1746	55.63	0.1173	0.9224	0.8215
	GB [37,38]	0.1772	53.20	0.1161	0.8342	0.8260
	SVR [34,71]	0.1942	55.88	0.1211	0.8148	0.8083
	MLP [39,72]	0.1849	58.32	0.1257	0.8780	0.8115
	ELM [40,41]	0.1927	59.09	0.1304	0.9796	0.7919
	DBN [42,43]	0.1892	57.10	0.1263	1.0258	0.7917
	HetDS(1,5)	0.1864	54.57	0.1194	0.8251	0.8163
	HetDS(1,10)	0.1808	52.70	0.1156	0.7878	0.8267
	HetDS(1,20)	0.1797	52.28	0.1154	0.8031	0.8262
	HetDS(3,5)	0.1767	52.40	0.1155	0.8335	0.8261
	HetDS(3,10)	0.1726	51.90	0.1137	0.8118	0.8327
	HetDS(3,20)	0.1720	51.51	0.1132	0.8144	0.8332
	HetDS(5,5)	0.1777	53.80	0.1176	0.8814	0.8201
	HetDS(5,10)	0.1764	53.84	0.1171	0.8748	0.8224
HetDS(5,20)	0.1762	53.85	0.1169	0.8730	0.8230	
Hetmean	0.1802	55.30	0.1205	0.9473	0.8113	
Hetmedian	0.1832	55.11	0.1208	0.9539	0.8070	

In addition, Table 4 presents the percentage difference (PD) regarding RMSE between single and ensemble models with the best proposed version of HetDS. It is possible to note that HetDS obtained a superior performance in terms of RMSE when compared to the other statistical and ML models used.

Table 4. Percentage difference (%) regarding RMSE between single and ensemble models with the best proposed version of HetDS.

MODEL	SERIES			
	Fortaleza	Florianópolis	Salvador	São Paulo
ARIMA [32,33]	14.97	27.28	19.55	9.11
RF [69,70]	6.38	19.16	1.5	0.72
GB [37,38]	8.21	19.67	2.96	1.18
SVR [34,71]	0.89	4.67	11.45	3.18
MLP [39,72]	0.75	11.29	6.97	2.62
ELM [40,41]	8.11	16.55	10.74	7.03
DBN [42,43]	2.72	9.16	9.08	3.23
Hetmean	4.43	4.23	7.76	2.46
Hetmedian	5.43	2.88	4.86	2.47

Figures 4–7 show the dispersion of the results considering the 10 independent simulations by means of a boxplot graphic.

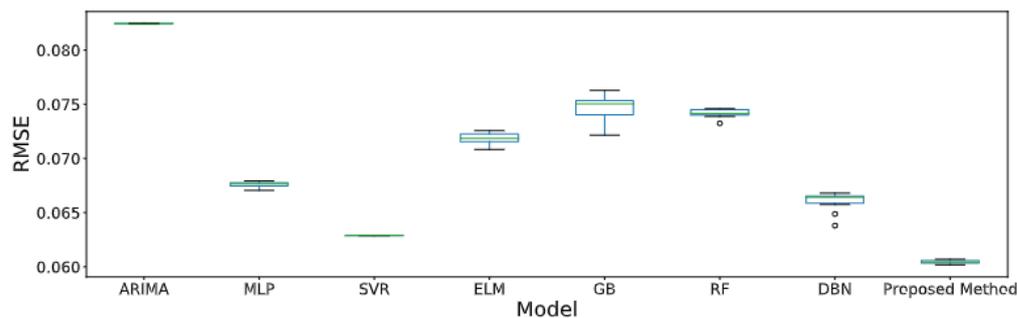


Figure 4. Boxplot graphic regarding Fortaleza.

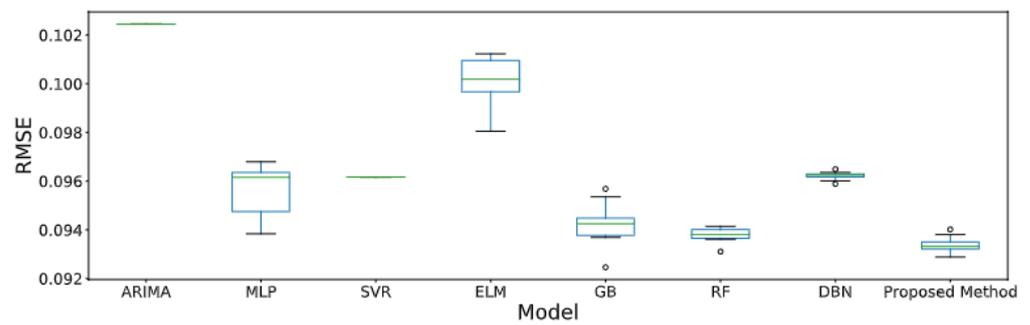


Figure 5. Boxplot graphic regarding Florianópolis.

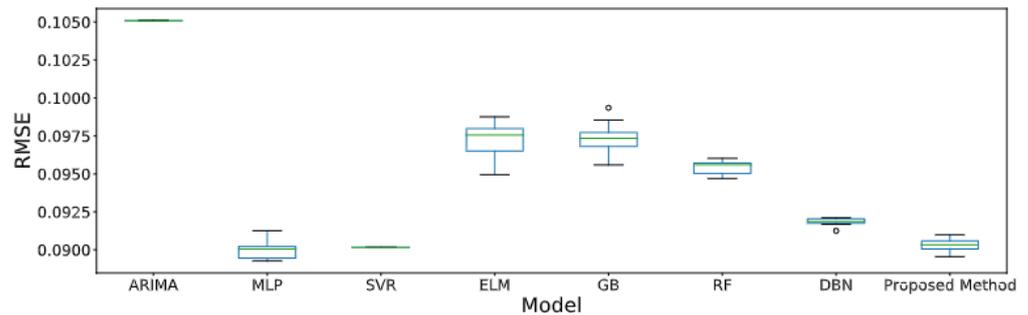


Figure 6. Boxplot graphic regarding Salvador.

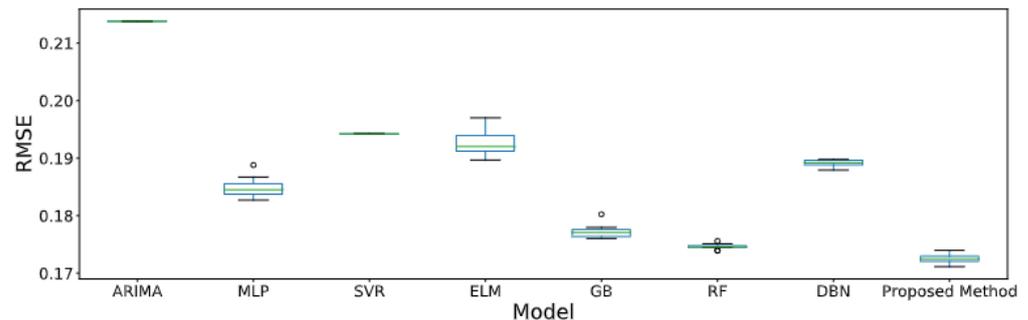


Figure 7. Boxplot graphic regarding São Paulo.

5.4. Discussion

Due to the number of computational results presented in Table 3, it is possible to discuss many aspects regarding the performance of the models. One can observe that the best RMSE value does not necessarily correspond to the best values for the other error metrics for a given method [73]. Previous works on time series forecasting have indicated such a behavior, which can occur when using distinct error metrics [56,74].

In two scenarios, a single model achieved the smallest error. Precisely, the smallest MAPE for Salvador, and the smallest MAPE, MAE, ARV, and IA for Florianópolis were achieved by SVR. In all other cases, the proposed model stood out. In terms of RMSE, Table 4 reveals the proposed approach overcame the single propositions, and Het_{mean} and Het_{median} .

After the application of the Friedman–Nemenyi [26] statistical test using 95% confidence, it is observed that the proposed method ($HetDS(m,k)$) achieved the best overall results in almost all comparisons, outperforming the single methods considering the f4our datasets, as highlighted in Figure 8.

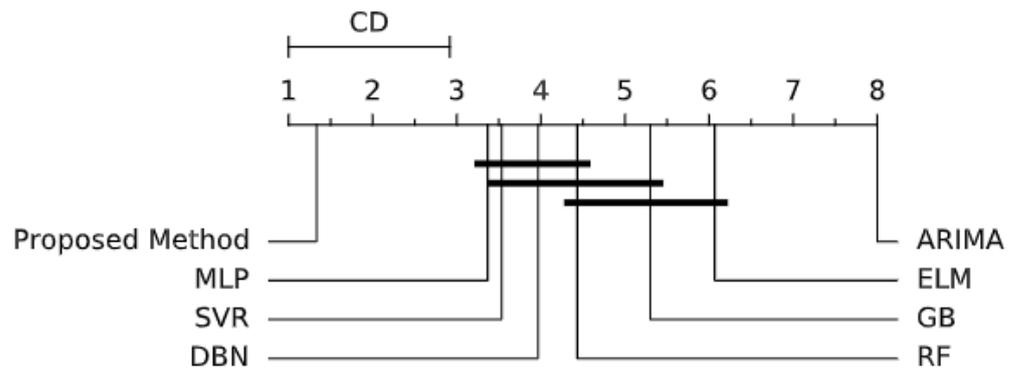


Figure 8. Friedman–Nemenyi test over all datasets and methods.

The analysis of the boxplot graphics in Figures 4–7 leads to some interesting remarks. As expected, the ARMA and SVR do not present dispersion since their training processes are performed using closed-form solutions. In general, the proposed approach achieved the best solutions and relatively small dispersion. Moreover, it did not find the best general solution for Salvador.

It is clear that the linear ARIMA model did not overcome any machine learning approach. One possibility of enhancing its prediction capability is to use bio-inspired metaheuristics [75]. Regarding the other single models, we observe high variability in terms of the performance or dispersion. It indicates how difficult it is to solve solar irradiation forecasting problems.

Finally, Figures 9–12 present the predictions provided by the best models in comparison with the original series in the test set. It is important to observe the performance on the peaks of the solar irradiance, where most of the deviation occurs.

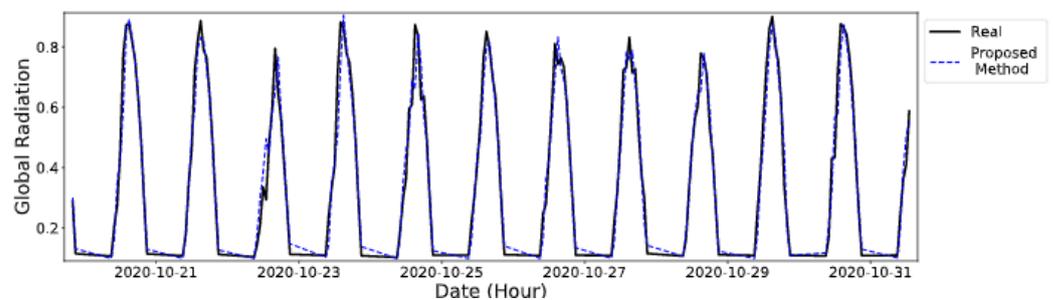


Figure 9. Solar irradiation forecasting obtained by the proposed model for Fortaleza series (first 150 test set points).

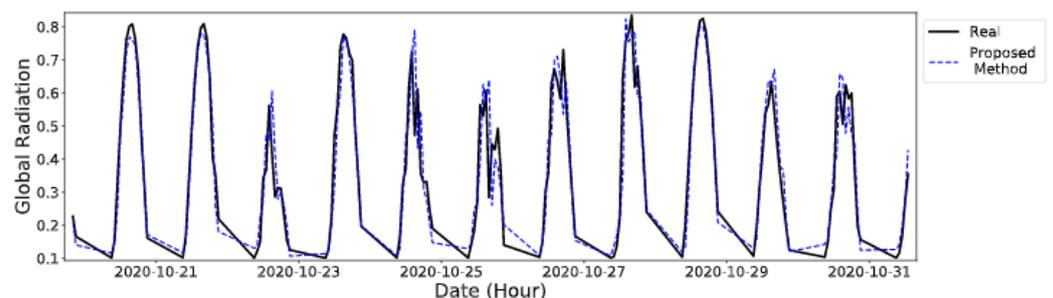


Figure 10. Solar irradiation forecasting obtained by the proposed model for Florianópolis series (first 150 test set points).

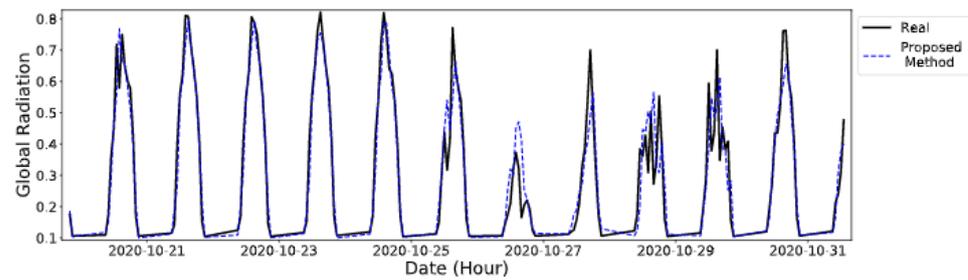


Figure 11. Solar irradiation forecasting obtained by the proposed model for Salvador series (first 150 test set points).

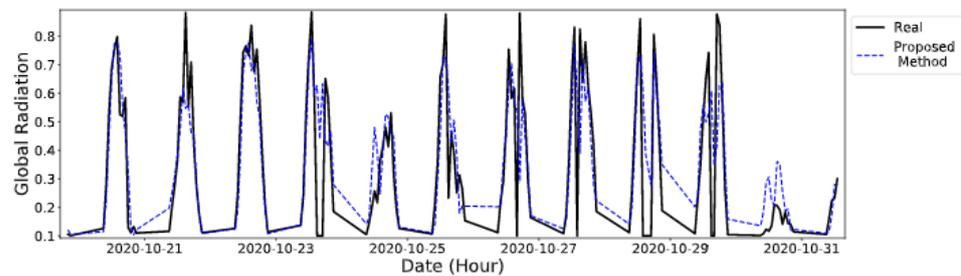


Figure 12. Solar irradiation forecasting obtained by the proposed model for São Paulo series (first 150 test set points).

6. Conclusions

At present, solar energy has a significant role due to the current concerns about green energy around the world. Thus, to improve the quality of energy generation processes, accurate forecasting systems are desirable. Solar irradiation data can present dynamic characteristics due to weather conditions. Therefore, single models (linear or nonlinear) may not achieve the best results. In this work, a heterogeneous dynamic ensemble (HetDS) approach was used to perform solar irradiation forecasting. Static ensemble approaches employ a group of models to predict future values. However, since the characteristics of the data may change over time, a dynamic approach could improve the results. In this sense, a dynamic ensemble composed of a pool of heterogeneous models was selected for each test pattern.

As stand-alone approaches, we considered the linear ARIMA from the Box and Jenkins methodology, and a pool of machine learning models: DBN, ELM, GB, MLP, RF, and SVR. The databases considered were from four major Brazilian cities. The experimental evaluation showed that the proposal can overcome the single models in almost all scenarios with a small dispersion, revealing its high approximation capability. The dynamic selection ensemble of the forecasting models increased the accuracy of the system, showing its competitiveness.

Regarding future work, we highlight the possibility of using recurrent neural networks and deep learning approaches. In addition, deeper analysis of the selection of the best lags can be considered, and the use of exogenous variables, such as precipitation, pressure, and temperature, or synthetic data, aiming to attain more accurate forecasts.

Author Contributions: Conceptualization, D.S.d.O.S.J., P.S.G.d.M.N. and J.F.L.d.O.; methodology, D.S.d.O.S.J., P.S.G.d.M.N., J.F.L.d.O. and H.V.S.; software, D.A.P.D., D.S.d.O.S.J. and T.M.B.; validation, A.R.L., M.H.N.M. and F.M.; formal analysis, P.S.G.d.M.N., J.F.L.d.O. and A.C.; investigation, D.A.P.D., D.S.d.O.S.J. and T.M.B.; resources, D.S.d.O.S.J., P.S.G.d.M.N. and J.F.L.d.O.; data curation, P.S.G.d.M.N., J.F.L.d.O. and A.C.P.; writing—original draft preparation, D.A.P.D., D.S.d.O.S.J., T.M.B., P.S.G.d.M.N., J.F.L.d.O., A.R.L., H.V.S. and A.C.P.; visualization, D.A.P.D., D.S.d.O.S.J., T.M.B., P.S.G.d.M.N. and J.F.L.d.O.; supervision, A.C.; project administration, M.H.N.M. and J.B.d.M.F.; funding acquisition, M.H.N.M. and J.B.d.M.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received external funding from Companhia Hidro Elétrica do São Francisco (CHESF) P&D + I N°02/2019, Coordination for the Improvement of Higher Education Personnel (CAPES)- Financing Code 001, Brazilian National Council for Scientific and Technological Development (CNPq), processes number 40558/2018-5, 315298/2020-0, Araucaria Foundation, process number 51497, and Foundation for support of Science and Technology of the State of Pernambuco (FACEPE) process number APQ-1252-1.03/21.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data are publicly available at the National Institute of Meteorology (INMET) <https://portal.inmet.gov.br/>, accessed on 5 March 2022.

Acknowledgments: This work is a result of the research and development project entitled “Technical arrangement to increase reliability and electrical safety by applying energy storage by batteries and photovoltaic systems to the auxiliary service of 230/500 kV substations”, from the Public Call—P&D + I N°02/2019 with financing from Companhia Hidro Elétrica do São Francisco (CHESF). This is an initiative within the scope of ANEEL’s (Agência Nacional de Energia Elétrica) Research and Technological Development Program for the Electric Energy Sector, under execution by the University of Pernambuco (UPE), Edson Mororó Moura Technology Institute (ITEMM) and Itaipu Technological Park Foundation (PTI). The authors thank CHESF and “Superintendência de Pesquisa e Desenvolvimento e Eficiência Energética” (SPE)/ANEEL for their support in making available the resources that allowed the preparation of this work. The authors also thank the Brazilian agencies Coordination for the Improvement of Higher Education Personnel (CAPES)- Financing Code 001, Brazilian National Council for Scientific and Technological Development (CNPq), processes number 40558/2018-5, 315298/2020-0, Araucaria Foundation, process number 51497, and Foundation for support of Science and Technology of the State of Pernambuco (FACEPE) process number APQ-1252-1.03/21 for their financial support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Khan, A.M.; Osińska, M. How to predict energy consumption in BRICS countries? *Energies* **2021**, *14*, 2749. [CrossRef]
2. Jackson, R.B.; Friedlingstein, P.; Andrew, R.M.; Canadell, J.G.; Quéré, C.L.; Peters, G.P. Persistent fossil fuel growth threatens the Paris Agreement and planetary health. *Environ. Res. Lett.* **2019**, *14*, 121001. [CrossRef]
3. Eyring, V.; Gillett, N.; Rao, K.A.; Barimalala, R.; Parrillo, M.B.; Bellouin, N.; Cassou, C.; Durack, P.; Kosaka, Y.; McGregor, S.; et al. Human Influence on the Climate System. In *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK, 2021. (in press)
4. Kumari, P.; Toshniwal, D. Deep learning models for solar irradiance forecasting: A comprehensive review. *J. Clean.* **2021**, *318*, 128566. [CrossRef]
5. 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* **2017**, *105*, 569–582. [CrossRef]
6. Espinar, B.; Aznarte, J.L.; Girard, R.; Moussa, A.M.; Kariniotakis, G. Photovoltaic Forecasting: A state of the art. In Proceedings of the 5th European PV-Hybrid and Mini-Grid Conference, Tarragona, Spain, 29–30 April 2010; OTTI-Ostbayerisches Technologie-Transfer-Institut: Tarragona, Spain, 2010; p. 250.
7. Perera, K.S.; Aung, Z.; Woon, W.L. Machine Learning Techniques for Supporting Renewable Energy Generation and Integration: A Survey. In Proceedings of the Second International Conference on Data Analytics for Renewable Energy Integration, DARE’14, Nancy, France, 19 September 2014; pp. 81–96. [CrossRef]
8. Frías-Paredes, L.; Mallor, F.; Gastón-Romeo, M.; León, T. Assessing energy forecasting inaccuracy by simultaneously considering temporal and absolute errors. *Energy Convers. Manag.* **2017**, *142*, 533–546. [CrossRef]
9. Moreno-Munoz, A.; de la Rosa, J.J.G.; Posadillo, R.; Bellido, F. Very short term forecasting of solar radiation. In Proceedings of the 2008 33rd IEEE Photovoltaic Specialists Conference, San Diego, CA, USA, 11–16 May 2008; pp. 1–5. [CrossRef]
10. Diagne, H.M.; Lauret, P.; David, M. Solar irradiance forecasting: State-of-the-art and proposition for future developments for small-scale insular grids. In Proceedings of the WREF 2012-World Renewable Energy Forum, Denver, CO, USA, 13–17 May 2012.
11. Box, G.E.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. *Time Series Analysis: Forecasting and Control*; John Wiley & Sons: Hoboken, NJ, USA, 2015.
12. Zhang, G.P. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing* **2003**, *50*, 159–175. [CrossRef]

13. Guermoui, M.; Melgani, F.; Gairaa, K.; Mekhalfi, M.L. A comprehensive review of hybrid models for solar radiation forecasting. *J. Clean. Prod.* **2020**, *258*, 120357. [[CrossRef](#)]
14. de Oliveira, J.F.L.; Silva, E.G.; de Mattos Neto, P.S.G. A Hybrid System Based on Dynamic Selection for Time Series Forecasting. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, 1–13. [[CrossRef](#)] [[PubMed](#)]
15. Izidio, D.M.; de Mattos Neto, P.S.; Barbosa, L.; de Oliveira, J.F.; Marinho, M.H.d.N.; Rissi, G.F. Evolutionary Hybrid System for Energy Consumption Forecasting for Smart Meters. *Energies* **2021**, *14*, 1794. [[CrossRef](#)]
16. Campos, D.S.; de Souza Tadano, Y.; Alves, T.A.; Siqueira, H.V.; de Nóbrega Marinho, M.H. Unorganized machines and linear multivariate regression model applied to atmospheric pollutant forecasting. *Acta Sci. Technol.* **2020**, *42*, e48203. [[CrossRef](#)]
17. Geman, S.; Bienenstock, E.; Doursat, R. Neural networks and the bias/variance dilemma. *Neural Comput.* **1992**, *4*, 1–58. [[CrossRef](#)]
18. Mendes-Moreira, J.; Soares, C.; Jorge, A.M.; Sousa, J.F.D. Ensemble approaches for regression: A survey. *ACM Comput. Surv. (Csur)* **2012**, *45*, 1–40. [[CrossRef](#)]
19. Ren, Y.; Zhang, L.; Suganthan, P.N. Ensemble classification and regression-recent developments, applications and future directions. *IEEE Comput. Intell. Mag.* **2016**, *11*, 41–53. [[CrossRef](#)]
20. Brown, G.; Wyatt, J.L.; Tino, P.; Bengio, Y. Managing diversity in regression ensembles. *J. Mach. Learn. Res.* **2005**, *6*, 41–53.
21. Webb, G.; Zheng, Z. Multistrategy ensemble learning: Reducing error by combining ensemble learning techniques. *IEEE Trans. Knowl. Data Eng.* **2004**, *16*, 980–991. [[CrossRef](#)]
22. Jiang, H.; Dong, Y.; Xiao, L. A multi-stage intelligent approach based on an ensemble of two-way interaction model for forecasting the global horizontal radiation of India. *Energy Convers. Manag.* **2017**, *137*, 142–154. [[CrossRef](#)]
23. Jovanovic, R.; Pomares, L.M.; Mohieldeen, Y.E.; Perez-Astudillo, D.; Bachour, D. An evolutionary method for creating ensembles with adaptive size neural networks for predicting hourly solar irradiance. In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017; pp. 1962–1967.
24. Sun, S.; Wang, S.; Zhang, G.; Zheng, J. A decomposition-clustering-ensemble learning approach for solar radiation forecasting. *Sol. Energy* **2018**, *163*, 189–199. [[CrossRef](#)]
25. Rodríguez, F.; Martín, F.; Fontán, L.; Galarza, A. Ensemble of machine learning and spatiotemporal parameters to forecast very short-term solar irradiation to compute photovoltaic generators' output power. *Energy* **2021**, *229*, 120647. [[CrossRef](#)]
26. Demšar, J. Statistical comparisons of classifiers over multiple data sets. *J. Mach. Learn. Res.* **2006**, *7*, 1–30.
27. Zhou, Y.; Liu, Y.; Wang, D.; Liu, X.; Wang, Y. A review on global solar radiation prediction with machine learning models in a comprehensive perspective. *Energy Convers. Manag.* **2021**, *235*, 113960. [[CrossRef](#)]
28. Rajagukguk, R.A.; Ramadhan, R.A.; Lee, H.J. A review on deep learning models for forecasting time series data of solar irradiance and photovoltaic power. *Energies* **2020**, *13*, 6623. [[CrossRef](#)]
29. Akhter, M.N.; Mekhilef, S.; Mokhlis, H.; Shah, N.M. Review on forecasting of photovoltaic power generation based on machine learning and metaheuristic techniques. *IET Renew. Power Gener.* **2019**, *13*, 1009–1023. [[CrossRef](#)]
30. Rigollier, C.; Lefèvre, M.; Wald, L. The method Heliosat-2 for deriving shortwave solar radiation from satellite images. *Sol. Energy* **2004**, *77*, 159–169. [[CrossRef](#)]
31. Jiang, Y. Computation of monthly mean daily global solar radiation in China using artificial neural networks and comparison with other empirical models. *Energy* **2009**, *34*, 1276–1283. [[CrossRef](#)]
32. Shadab, A.; Said, S.; Ahmad, S. Box-Jenkins multiplicative ARIMA modeling for prediction of solar radiation: A case study. *Int. J. Energy Water Resour.* **2019**, *3*, 305–318. [[CrossRef](#)]
33. Voyant, C.; Muselli, M.; Paoli, C.; Nivet, M.L. Numerical weather prediction (NWP) and hybrid ARMA/ANN model to predict global radiation. *Energy* **2012**, *39*, 341–355. [[CrossRef](#)]
34. Chen, J.L.; Li, G.S. Evaluation of support vector machine for estimation of solar radiation from measured meteorological variables. *Theor. Appl. Climatol.* **2014**, *115*, 627–638. [[CrossRef](#)]
35. Bendiek, P.; Taha, A.; Abbasi, Q.H.; Barakat, B. Solar irradiance forecasting using a data-driven algorithm and contextual optimization. *Appl. Sci.* **2021**, *12*, 134. [[CrossRef](#)]
36. Huang, J.; Troccoli, A.; Coppin, P. An analytical comparison of four approaches to modelling the daily variability of solar irradiance using meteorological records. *Renew. Energy* **2014**, *72*, 195–202. [[CrossRef](#)]
37. Park, J.; Moon, J.; Jung, S.; Hwang, E. Multistep-ahead solar radiation forecasting scheme based on the light gradient boosting machine: A case study of Jeju Island. *Remote Sens.* **2020**, *12*, 2271. [[CrossRef](#)]
38. Fan, J.; Wang, X.; Wu, L.; Zhou, H.; Zhang, F.; Yu, X.; Lu, X.; 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 Convers. Manag.* **2018**, *164*, 102–111. [[CrossRef](#)]
39. Elminir, H.K.; Areed, F.F.; Elsayed, T.S. Estimation of solar radiation components incident on Helwan site using neural networks. *Sol. Energy* **2005**, *79*, 270–279. [[CrossRef](#)]
40. Salcedo-Sanz, S.; Jiménez-Fernández, S.; Aybar-Ruíz, A.; Casanova-Mateo, C.; Sanz-Justo, J.; García-Herrera, R. A CRO-species optimization scheme for robust global solar radiation statistical downscaling. *Renew. Energy* **2017**, *111*, 63–76. [[CrossRef](#)]
41. Salcedo-Sanz, S.; Deo, R.C.; Cornejo-Bueno, L.; Camacho-Gómez, C.; Ghimire, S. An efficient neuro-evolutionary hybrid modelling mechanism for the estimation of daily global solar radiation in the Sunshine State of Australia. *Appl. Energy* **2018**, *209*, 79–94. [[CrossRef](#)]

42. Ghimire, S.; Deo, R.C.; Raj, N.; Mi, J. Deep learning neural networks trained with MODIS satellite-derived predictors for long-term global solar radiation prediction. *Energies* **2019**, *12*, 2407. [CrossRef]
43. Zang, H.; Cheng, L.; Ding, T.; Cheung, K.W.; Wang, M.; Wei, Z.; Sun, G. Application of functional deep belief network for estimating daily global solar radiation: A case study in China. *Energy* **2020**, *191*, 116502. [CrossRef]
44. Siqueira, H.; Luna, I.; Alves, T.A.; de Souza Tadano, Y. The direct connection between box & Jenkins methodology and adaptive filtering theory. *Math. Eng. Sci. Aerosp. (MESA)* **2019**, *10*. Available online: <http://nonlinearstudies.com/index.php/ mesa/article/view/1868> (accessed on 5 March 2022).
45. de Mattos Neto, P.S.; Ferreira, T.A.; Lima, A.R.; Vasconcelos, G.C.; Cavalcanti, G.D. A perturbative approach for enhancing the performance of time series forecasting. *Neural Netw.* **2017**, *88*, 114–124. [CrossRef]
46. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
47. Ho, T.K. Random decision forests. In Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, Canada, 14–16 August 1995; Volume 1, pp. 278–282.
48. Amit, Y.; Geman, D. Shape quantization and recognition with randomized trees. *Neural Comput.* **1997**, *9*, 1545–1588. [CrossRef]
49. Cutler, A.; Cutler, D.R.; Stevens, J.R. Random Forests. In *Ensemble Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 157–175.
50. Freund, Y.; Schapire, R.E. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139. [CrossRef]
51. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
52. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 26.
53. 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, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
54. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013.
55. de Mattos Neto, P.S.; de Oliveira, J.F.; Domingos, S.d.O.; Siqueira, H.V.; Marinho, M.H.; Madeiro, F. An adaptive hybrid system using deep learning for wind speed forecasting. *Inf. Sci.* **2021**, *581*, 495–514. [CrossRef]
56. Belotti, J.; Siqueira, H.; Araujo, L.; Stevan, S.L.; de Mattos Neto, P.S.; Marinho, M.H.; de Oliveira, J.F.L.; Usberti, F.; Leone Filho, M.d.A.; Converti, A.; et al. Neural-Based ensembles and unorganized machines to predict streamflow series from hydroelectric plants. *Energies* **2020**, *13*, 4769. [CrossRef]
57. Siqueira, H.; Luna, I. Performance comparison of feedforward neural networks applied to stream flow series forecasting. *Math. Eng. Sci. Aerosp. (MESA)* **2019**, *10*, 41–53.
58. Ribeiro, V.H.A.; Reynoso-Meza, G.; Siqueira, H.V. Multi-objective ensembles of echo state networks and extreme learning machines for streamflow series forecasting. *Eng. Appl. Artif. Intell.* **2020**, *95*, 103910. [CrossRef]
59. Huang, G.B.; Zhu, Q.Y.; Siew, C.K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501. [CrossRef]
60. de Souza Tadano, Y.; Siqueira, H.V.; Alves, T.A. Unorganized machines to predict hospital admissions for respiratory diseases. In Proceedings of the 2016 IEEE Latin American Conference on Computational Intelligence (LA-CCI), Cartagena, Colombia, 2–4 November 2016; pp. 1–6.
61. Hinton, G.E.; Osindero, S.; Teh, Y.W. A fast learning algorithm for deep belief nets. *Neural Comput.* **2006**, *18*, 1527–1554. [CrossRef]
62. Yang, B.; Zhu, T.; Cao, P.; Guo, Z.; Zeng, C.; Li, D.; Chen, Y.; Ye, H.; Shao, R.; Shu, H.; et al. Classification and summarization of solar irradiance and power forecasting methods: A thorough review. *CSEE J. Power Energy Syst.* **2021**, 1–19. [CrossRef]
63. Le Roux, N.; Bengio, Y. Representational power of restricted Boltzmann machines and deep belief networks. *Neural Comput.* **2008**, *20*, 1631–1649. [CrossRef]
64. Kourentzes, N.; Barrow, D.K.; Crone, S.F. Neural network ensemble operators for time series forecasting. *Expert Syst.* **2014**, *41*, 4235–4244.
65. Hyndman, R.; Khandakar, Y. Automatic Time Series Forecasting: The forecast package for R. *J. Stat. Softw. Artic.* **2008**, *27*, 1–22.
66. de Mattos Neto, P.S.; Cavalcanti, G.D.; de O Santos Júnior, D.S.; Silva, E.G. Hybrid systems using residual modeling for sea surface temperature forecasting. *Sci. Rep.* **2022**, *12*, 487. [CrossRef] [PubMed]
67. Rodrigues, A.L.J.; Silva, D.A.; de Mattos Neto, P.S.G.; Ferreira, T.A.E. An experimental study of fitness function and timeseries forecasting using artificial neural networks. In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2010), ACM, Portland, OR, USA, 7–11 July 2010; pp. 2015–2018.
68. de Mattos Neto, P.S.G.; Rodrigues, A.L.J.; Ferreira, T.A.E.; Cavalcanti, G.D. An intelligent perturbative approach for the time series forecasting problem. In Proceedings of the IEEE World Congress on Computational Intelligence (WCCI 2010), Barcelona, Spain, 15–17 July 2010; pp. 1–8.
69. Bhola, P.; Bhardwaj, S. Solar energy estimation techniques: A review. In Proceedings of the 2016 7th India International Conference on Power Electronics (IICPE), Patiala, India, 17–19 November 2016; pp. 1–5.
70. Srivastava, R.; Tiwari, A.; Giri, V. Solar radiation forecasting using MARS, CART, M5, and random forest model: A case study for India. *Heliyon* **2019**, *5*, e02692. [CrossRef] [PubMed]
71. Hocaoglu, F.O. Novel analytical hourly solar radiation models for photovoltaic based system sizing algorithms. *Energy Convers. Manag.* **2010**, *51*, 2921–2929. [CrossRef]

72. Linares-Rodríguez, A.; Ruiz-Arias, J.A.; Pozo-Vázquez, D.; Tovar-Pescador, J. Generation of synthetic daily global solar radiation data based on ERA-Interim reanalysis and artificial neural networks. *Energy* **2011**, *36*, 5356–5365. [[CrossRef](#)]
73. Hyndman, R.J.; Koehler, A.B. Another look at measures of forecast accuracy. *Int. J. Forecast.* **2006**, *22*, 679–688. [[CrossRef](#)]
74. Siqueira, H.; Boccatto, L.; Attux, R.; Lyra, C. Unorganized machines for seasonal streamflow series forecasting. *Int. J. Neural Syst.* **2014**, *24*, 1430009. [[CrossRef](#)]
75. Siqueira, H.; Belotti, J.T.; Boccatto, L.; Luna, I.; Attux, R.; Lyra, C. Recursive linear models optimized by bioinspired metaheuristics to streamflow time series prediction. *Int. Trans. Oper. Res.* **2021**. [[CrossRef](#)]