Applied Energy xxx (xxxx) 123215

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Solar irradiance time series forecasting using auto-regressive and extreme learning methods: Influence of transfer learning and clustering

Milan Despotovic [a,*], Cyril Voyant [b], Luis Garcia-Gutierrez [c], Javier Almorox [d], Gilles Notton [b]

[a] Faculty of Engineering, University of Kragujevac, 34000 Kragujevac, Serbia
[b] Sciences for Environment Laboratory, UMR CNRS 6134, University of Corsica Pasquale Paoli, 20000 Ajaccio, France
[c] Optical Materials, Photonics and Systems Laboratory, LMOPS, University of Lorraine & CentraleSupelec, 57000 Metz, France
[d] Universidad Politécnica de Madrid, 28040 Madrid, Spain

ARTICLE INFO

Keywords:
Extreme learning machine
Solar irradiance forecasting
Transfer learning
Artificial intelligence

ABSTRACT

Solar resource forecasting is essential for an optimal energy management in smart grids using photovoltaic (PV) production. For many sites and for short time horizons (nowcasting), approaches based on the use of time series and statistical or Artificial Intelligence methods are often preferred. These methods require a long historical time series of solar radiation not always available. A practical question often arises: what happens if we want to install a PV plant in an area not covered by a consistent historical data set? Is it possible then to "delocalize" the learning using data of another location? The objective is to compare the performances of two forecasting methods (Auto-Regressive and Extreme Learning) using alternatively for the learning process solar data measured on-site, data measured on stations with similar characteristics (from the same cluster) and data collected on any other stations. 70 Spanish meteorological stations are used (with a maximal distance between stations of 2000 km and with an altitude varying between 18 and 950 m). Using transfer Learning when no solar data are available conduces to obtain results in the same range of performances than using Direct Learning. This work shows that transfer learning (based on extreme learning forecasting) is fully acceptable (about 1 percentage point higher for classical error metric nRMSE) whatever the stations used for learning, and this deviation is reduced to 0.5 percentage point when using a station related to the same cluster. Thus, the clustering seems to be not efficient to improve the reliability of the solar irradiance forecasting.

List of abbreviations including units and nomenclature

AI	Artificial Intelligence
AR	AutoRegressive
ClusterSkill	ClusterSkill metric (unitless)
DL	Direct Learning
ELM	Extreme Learning Machine
F	Forecastability (unitless)
GHI	Global Horizontal Irradiance (W/m^2)
GHIcs	Global Horizontal Irradiance by clear sky calculated by Mc Clear model (W/m^2)
Lat	Latitude (°)
Long	Longitude (°)
MAE	Mean Absolute Error (W/m^2)
nMAE	Normalized Mean Absolute Error (unitless)
nRMSE	Normalized Root Mean Square Error (unitless)
NWP	Numerical Weather Prediction
PV	Photovoltaic
RMSE	Root Mean Square Error (W/m^2)
SiAR	Agroclimatic Information System for Irrigation
$\text{SkillScore}_{i,j,h}$	SkillScore metrics (unitless)
TL	Transfer Learning

1. Introduction

Today, more than 50% of planet's inhabitants live in cities consuming two-thirds of global energy consumption and being responsible for >70% of global carbon emissions [1]. These percentages are expected to grow significantly in the coming decades, which makes the cities very crowded and complex spaces with an inevitable need for implementing new intelligent solutions to efficiently manage resources. This concept is known as the smart city, and it implies a wide range of assets and services like transportation, energy supply [2], waste management, security, information systems, etc. A very important concept in smart

* Corresponding author.
E-mail address: mdespotovic@kg.ac.rs (M. Despotovic).

https://doi.org/10.1016/j.apenergy.2024.123215
Received 28 November 2023; Received in revised form 24 February 2024; Accepted 11 April 2024
0306-2619/© 20XX

cities is the smart electrical grid, which has an ability to better integrate renewable energy systems into the network, to improve the energy management, to reduce the electricity bill (both for producer and consumer) and the carbon footprint of the cities [3]. The main problem associated with the optimization of Energy Management Systems (EMS) in an energy system using intermittent energy sources is due to their uncertain and intermittent nature and their dependency on the weather and environmental conditions [4,5]. An imbalance between the fluctuating generated power and varying consumer demands may cause a blackout, an unstable operation of the utility grid and at least, a distribution of electric power of poor quality.

Such microgrids, using more and more intermittent and stochastic renewable energy productions, need to incorporate energy storage means to store the excess production (comparing to the load) and to release it when the load supply requires it. Adding energy storage means is not sufficient; it is important and crucial to know when this storage must be used (charged or discharged) and how the different sources must occur in the production at each instant.

Decisions to charge or discharge the storage or to use an electricity production mean rather than another one at various time horizons must be taken according to:

- the availability of the production means;
- the state of the storage means;
- the cost of the produced, sold or exchanged electricity;
- the future electrical needs and the future of the solar energy production;
- other influential parameters.

In other words, anticipating an event or an availability is essential to manage efficiently an EMS [6].

Numerous researchers have developed short-term and long-term demand forecasting models for power systems, considering the interdependence of the demand patterns with the spatiotemporal behavior of the consumers and weather conditions [7–10]. The strong connection between demand forecasting and smart grid was well described in Refs. [11–13]. Applications of solar forecasting in energy management of smart grid were investigated in detail in [14,15]. Tripathy and Prusty focused on the probabilistic forecasting of intermittent types of renewable energy generation for different smart grid applications [16].

The forecasting methods can be divided into three categories: sky image-based methods [17–19], time series-based methods [20,21] and numerical Weather Prediction (NWP) methods [22,23]. The choice of the appropriate forecasting methods depends on the forecasting time horizon and the temporal resolution, i.e. on the application. The most appropriate forecasting methods for intra-day grid operation are time series-base methods that are suitable for short time horizon from some minutes to some hours with a time step from 10 min to 1 h [24,25]. These methods use historical data of solar irradiance measurement and they are divided into statistical models [25], which derive relations between future solar irradiance and input variables (endogenous and/or exogenous) from statistical analysis, and models based on artificial intelligence (AI) [26–31].

Both statistical and AI models provide effective solutions for solar irradiance forecasting to be used in energy management system of smart grid with photovoltaic (PV) production with or without storage. However, newly-built photovoltaic plants often lack historical data, making it very difficult to predict future solar irradiance accurately. The small number of solar radiation measurement stations is particularly due to high investment and maintenance costs, and even in industrialized countries, the national network often comprises relatively small number of solar radiation stations. The exact number of solar radiations measuring stations through the world is difficult to determine [32]. In about 98% of cases, the stations are too far to deliver accurate information to users because it is considered that the solar data are valid in an area of 30 km around the solar measuring station and in regions of great altitude variation, this distance logically decreases sharply [33].

Often, when no solar radiation data are available on the future site of installation, the PV project manager tries to find data measured on the nearest meteorological station since a sufficient long time to be statistically representative.

The obvious question that arises is: Is it possible to use with an acceptable accuracy the model that has been developed and trained on data collected on one site for forecasting solar irradiance in sites where existing time series are either not available or not long enough? This concept is known as Transfer Learning (**TL**) [34,35]; it is a machine learning technique where a pre-trained model is reused for a new problem. it adapts the knowledge acquired from solving a previous task (source domain) to improve generalization on a different task (target domain). In our case, it consists to train a solar forecasting model using data from a meteorological station called source station and to apply this learned knowledge to predict the solar radiation in another meteorological station called target station.

The objective of this work is to answer to the above question; the structure of this paper is:

- some useful information about solar radiation forecasting are firstly given;
- then, the concept of transfer learning applied to our problem is presented;
- in a third paragraph, a literature review about the application of transfer learning methods to solar radiation prediction is then realized; this bibliographical study will highlight the originality of this work and how it differed from the existing literature;
- the two chosen forecasting methods are briefly described: Auto-Regressive and Deep Learning;
- the 70 meteorological stations spread over Spain are situated; these stations have been clustered into 3 or 5 groups according to various meteorological and statistical parameters (such as variability, which will be described briefly) and using different clustering techniques that have been described in a earlier work [36]; the distribution of meteorological stations by class and by clustering method will be given; at last, the metrics used to estimate the accuracy of the models will be described;
- In a last paragraph, the performances of the forecasting using as input data solar irradiances measured on-site, measured on sites with similar characteristics (belonging to the same cluster) or collected in any of the 69 other sites will be commented.

1.1. Some useful reminders about solar radiation time-series forecasting

We thought it was useful, even essential, to give some information on the utilization of the clearness index (or clear sky index) k_t, as the replacement of *GHI* and on the calculation of the clear-sky solar irradiance. More details are given in a previous work [36] whose results are to be used in this paper.

The forecasting models based on time series are accurate for stationary time series. Non-stationarity influences the performance of the prediction models [37–40]. But, solar radiation behavior does not follow a stationary process because it has a double periodicity (a yearly cycle and a diurnal one) [41,42]. It is then necessary to convert this non-stationary solar data time series into a stationary one in transforming *GHI* in a dimensionless variable called clear sky index (k_t) defined as the ratio of *GHI* to the global solar irradiation estimated under clear-sky conditions (GHI_{cs}) at the same time t:

$$k_t(t) = \frac{GHI(t)}{GHI_{cs}(t)} \quad (1)$$

Several clear sky models based on parameters like sun position, latitude, altitude and atmospheric variables are existing and a relevant review of these models can be found in [43]. In this work, the McClear model [44] was applied using the New Heliosat-4 method [45] which is more detailed in [36].

1.2. Transfer learning principle

To understand the basic concept of transfer learning, consider that a model **X** is successfully trained to perform task **A**. If the size of the dataset for task **B** is too small, preventing a model **Y** from training effectively or causing data overfitting, a part of model **X** can be used as a basis for building model **Y** to perform task **B** (pre-learning). This is even more interesting when there is no historical data related to task **B** - the model **X**, after appropriate tuning (scaling), can be directly used to perform task **B** (full-learning).

In this study, we deployed transductive transfer learning, when the source and the target tasks are similar, but the corresponding domains are different either in terms of data or marginal probability distributions [46,47]. This method is explained in Fig. 1.

In this paper, the objective of the **TL** consists in using, in the learning process of the forecasting model **X**, the solar data set collected in a site **A** (with a large number of data); once developed, this model **X** is applied and tested to another set of solar radiations measured in a site **B** after a rescaling.

As said previously, the forecasting method is not applied directly to the *GHI* value but to the clearness index k_t; the k_t dataset, used for the learning step, is constructed using measured *GHI* data and calculated GHI_{cs} of the site **A**; when the forecasting model is constructed, it is tested on solar data of the site **B**, using this time the k_t dataset of the site **B** (using measured *GHI* and calculated GHI_{sc} of the site **B**) and predicting a future value of k_t which will be transformed in *GHI* in multiplying it by GHI_{cs} (calculated in the site **B**); it is the rescaling.

During the transfer of the learning (when training and testing sites are different) a scaling method is involved because adjustment of the forecasting are needed, some locations being more than 1000 km distant. Actually, the *GHI* which are used as input of the models are divided by the values of the GHI_{cs} concerning the training site. Then, when the prediction is performed, the output is rescaled, multiplying the predicted k_t by the GHI_{cs} computed under clear sky referring to the testing or target station. Fig. 2 illustrates the process for predicting solar radiation in Domain **2** using a forecasting model trained using data of Domain **1** (at the top) i.e. with **TL** and re-scaling and using a forecasting model trained using data from Domain **2** (at the bottom).

1.3. State of the art of Transfer Learning applied to solar irradiance

A complete and recent review, with about 300 cited references, about the role of **TL** methods applied to sustainable smarts cities applications such as load forecasting, thermal comfort control, energy disaggregation, smart grid and energy trading, etc. was written by Himeur et al. [47]. A special paragraph is dedicated to renewable energy and particularly on PV systems.

Quiereshi et al. [46] showed that deep neural network based meta-regressor trained on one wind farm can be used with a good reliability to predict the wind power for the four other wind farms. Such transfer

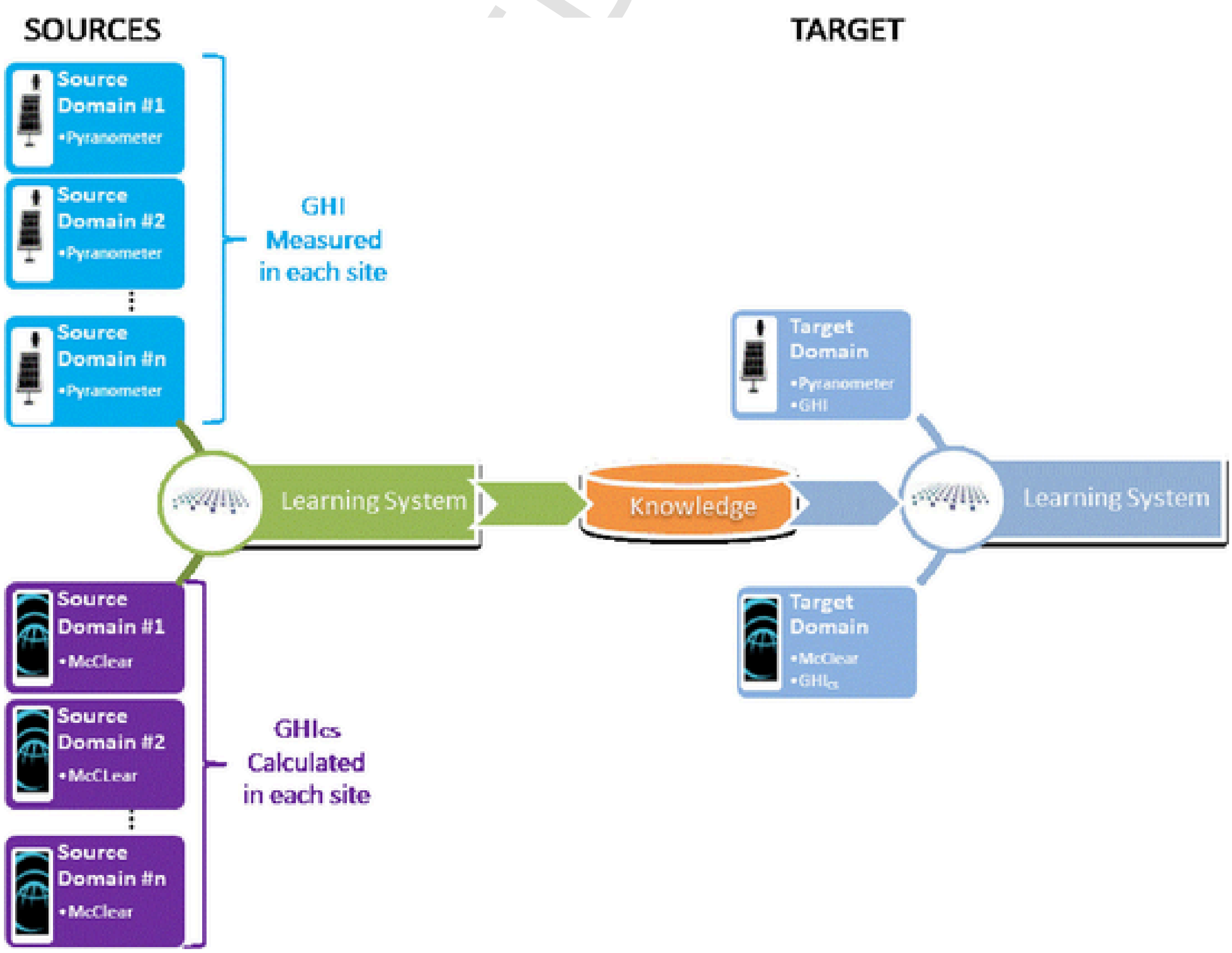

Fig. 1. Transfer learning principle.

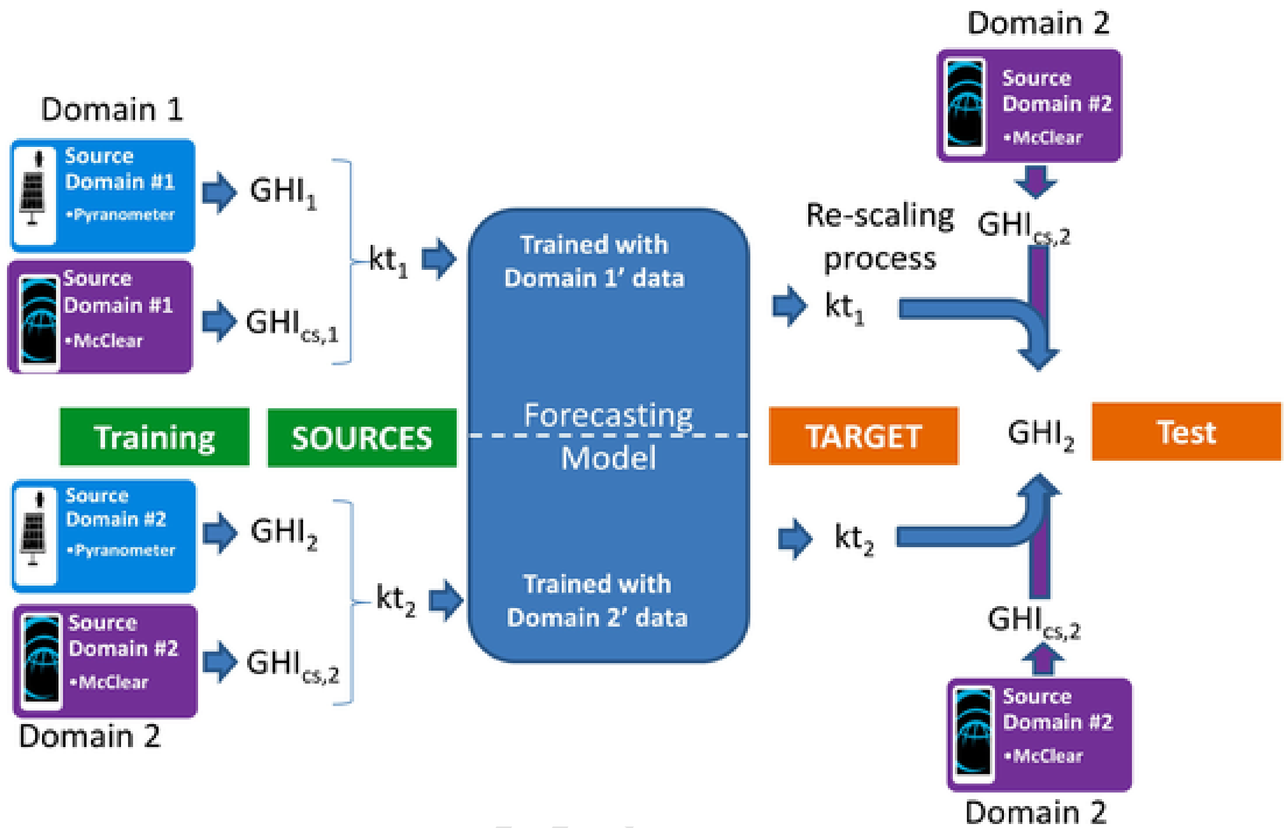

Fig. 2. Forecasting with and without transfer learning (respectively at the top and the bottom of the figure).

learning has added advantage of saving extra time needed to train from scratch a new model on each of the wind farms. The authors evaluated the accuracy from five Wind Farms located at different climate regions in Europe using three years of data.

Zhang et al. [49] showed that **TL** is not always an effective way to improve the performances; they wrote a survey showing that in some cases, **TL** can cause negative transfer. They concluded that to be efficient, its assumptions must be satisfied: 1) the learning tasks in the two domains are related/similar; 2) the source and target domains data distributions are not too different; and, 3) a suitable model can be applied to both domains.

In [50], the authors developed a Feed-forward (FF) and Long Short-Term Memory (LSTM) neural networks to forecast hourly solar irradiance over a multi-step horizon; they wanted to observe the generalization capability of these models which were trained on the solar irradiance data collected in Como station in Italy (source domain) and to apply, without retraining, to three different locations (target domain) situated at about 1° latitude away, with a different environment and with a difference of yearly average solar irradiation of about 25% with the source domain. The results showed a small loss of precision (about 3% difference for a 6-h time horizon in term of R^2) in using a learning made in Como and without a retraining for both the two models. In this paper, the maximum distance between station is about 160 km.

To forecast global horizontal solar irradiance up to 6 h ahead in sites with no solar irradiance data, a multi-model framework based on evolutionary artificial neural networks has been proposed [51]. This framework was tested on hourly global horizontal solar irradiance data acquired from Meteonorm software for 28 Moroccan cities (maximum distance between cities is around 1800 km). These cities were first partitioned into six zones according to climate factors in Morocco, and then three scenarios were used to evaluate the framework: 1) for each city, the train phase and the test phase were performed using the dataset of that city, 2) datasets of cities of a particular zone were used for training to predict solar irradiance of other cities belonging to the same zone, 3) a single train dataset was built from different cities belonging to different zones, and a single test dataset was built with the remaining cities. The average values of error metrics from t + 1 h to t + 6 h according to each scenario are given in Table 1. The decrease of performance levels passing from scenario 1 to scenario 3 is very small and probably not significant.

A constrained long short-term memory (C-LSTM) was used in [52], to predict the hourly day-ahead photovoltaic power generation in Australia, by using the historical data of eight weather variables during 24 months in the source domain, and testing different transfer learning strategies on the data of 10 days in the target domain. The utilization of two **TL** strategies to forecast the PV production in the target domain improves the MSE between 18.42% and 24.77% which indicates clearly the improvement due to the application of a transfer learning method. No information is given on the distance between the two PV stations.

To address the forecasting problem for newly built PV sites, a hybrid PV power forecasting model incorporating transfer learning strategy that fine-tunes the pre-trained model is proposed in [53]. This model is capable of forecasting PV generation from historical data of active power along with meteorological factors including wind speed, atmospheric temperature, relative humidity, global horizontal radiation, diffuse horizontal radiation, etc. For validating the **TL** strategies, the data

Table 1
Average values (from t + 1 h to t + 6 h) of nMAE and nRMSE (minimum-maximum) [51].

Scenario	1	2	3
nMAE	4.41%–8.12%	4.49%–8.20%	4.48%–8.37%
nRMSE	7.59%–12.49%	8.38%–13.27%	8.44%–13.76%

from one site within the Alice Springs plant was used as the source domain, and two different sites with different specifications from the same plant are chosen as two independent target domains. The improvement due to the **TL** application was estimated for various days and the authors concluded that the "pre-trained model parameter fine-tuning" pattern not only addresses the insufficient data but also speeds up the model training on modelling for new sites. The maximum distance between source and targets sites is about 500 m.

In [54], a hybrid deep learning framework with a **TL** module is proposed for multi-step ahead global horizontal irradiance forecasting. The framework was tested by two simulation experiments that were carried out using two pairs of datasets belonging to the source domain and the corresponding target domain from California. The latitude difference between the locations of the studied datasets was less than one-degree latitude in both experiments. One of the major findings is that there is evidence that the introduction of the **TL** step in the forecasting process is an effective way to enhance the accuracy of multi-step ahead *GHI* forecasting.

In [55], a transfer learning approach for PV production forecasting in the case of data scarcity, with predictive LSTM model was used to train the model in PV plants with data adequacy and then transfer knowledge to PV plants with small sample of available data. The authors used a Long Short-Term Memory (LSTM) model coupled with 3 **TL** strategies. **TL** is used both for weight initialization of the LSTM model and for feature extraction and the performances are compared to the non-**TL** model. For testing the approach, data collected from seven PV plants located in four Portuguese cities (with a maximal distance of 500 km) were used. **TL** models allows to reach 12.6% accuracy improvement in terms of RMSE and 16.3% in terms of forecast skill index.

Gao et al. [56] used two sets of solar radiation data (Tokyo and Okinawa, 1600 km distant) and interchanged source and target domains to validate the efficacy and robustness of the proposed model. They showed that compared with the method of directly using the source domain model (DL), transfer learning (TL) can enhance the accuracy of the test set by at least 14%.

Abubakr et al. [57] used a **TL** for training one-hour-ahead forecasters of hourly *GHI* with a neural network-based models. Cairo (Egypt) was the source site and the target sites were five other locations in Tunisia, Morocco, and Jordan. The approach was found valid for all targeted locations, with MAEs and RMSEs lower than 46.4 and 73.4 W/m^2.

TL was used for short-term solar radiation forecast using total sky imager (TSI) [19] and it has been shown that **TL** approach significantly reduces the time and resources required for modelling solar radiation while enhancing the effectiveness of the method.

This short literature review shows that transfer learning is a well-known concept often realized when no data are available. Excepted in some particular cases, **TL** improves the performances of the models when it is added in the forecasting model. When the solar data used for training the model are not measured on the site of forecasting, the decrease of the efficiency of the forecasting method is generally small.

It is this last point that this work aims to verify in a new sample of meteorological stations, by comparing the performances the machine learning trained with solar irradiance data measured at stations with similar "variation similarity" (belonging to the same group) and data measured anywhere in the sample of weather stations.

In this literature review, it appears that generally:

- The number of stations used for TL is small, the maximum number of sites was 28 in Morocco [51];
- The maximum distance between the site is relatively short, from 500 m to some hundred kilometres, excepted for Morocco with a maximum distance of 1800 km but using a small data set;
- One-hour time step of solar radiation or PV production the data are generally used;
- Excepted in ref. [51], only few studies have been done to date on the influence of the belonging of meteorological stations to a same family (same intermittent variation characteristics).

The originality of this paper is due to two points:

- In order to make the results more relevant and generalizable, in this paper, 4 years of half-hourly solar irradiance data from 70 meteorological stations will be used with a maximum distance between two stations equal to about 2000 km;
- Furthermore, the influence of the belonging or not of the two stations (target and source) to the same cluster was not studied.

1.4. Data and clustering

1.4.1. Data description

The data of Global Horizontal Irradiance (*GHI*) used in this study have been downloaded from Agroclimatic Information System for Irrigation (SiAR) which was launched by the Spanish Ministry of Agriculture, Fisheries, Food and the Environment [58]. To test the effectiveness of the proposed model, *GHI* data from 70 meteorological stations, spanned all over Spain, were used. The complete dataset used contains data series recorded from 1 January 2017 to 31 December 2020 (4 years), with a 30 min time step (4,908,960 data in total). All the data were cleaned and verified according an ad-hoc process described in [36].

As said in the first section, the forecasting will be realized on the clearness index k_t, what imposes to know the clear sky solar irradiance corresponding to each half-hour values of *GHI*. A clear-sky model estimates the solar irradiance is reaches the ground surface when the sky is clear without clouds. Such models have been developed and differ from one another mainly in the input data required. The clear sky solar irradiance was calculated here using the McClear model [45] which is an efficient model with a fast implementation and available via a web service (https://www.oie.minesparis.psl.eu/Valorisation/Outils/McClear/). This model was tested and verified, in a previous work, realized with the same data set [36]. More information of the cleaning methods and on the calculation of kt are given in reference [36].

The position of the 70 stations is shown in Fig. 3 and the name and geographical coordinates with altitudes are given in Table 2.

We note that the study area is large:

- o from 13.94°W for Antigua (Canary Islands) to 1.44°E for Santa Eulia del Rio in longitude i.e. a difference equal to 16.38° of longitude; if the station in the Canary Islands is not taken into account, the westernmost town in Spain city is Moraleja (6.69°W). The maximum longitude difference between Moraleja and Santa Eulia del Rio is therefore 8.13°
- o from 28.32°N for Antigua to 42.69°N for Adios in latitude i.e. a gap of 14.37° in latitude; if Antigua is not considered then the far Southern city is Antequera (37.03°N) i.e. a variation in latitude of 5.66°.

Concerning the altitude, some stations are at the sea level as Algemesi (19 m) and some stations are at medium altitude as Monreal del Campo at 950 m above the sea level.

In distance, the longest distance as the crow flies is between Antigua and Santa Celia with 1974 km and in the Spanish mainland between Antequera and Santa Celia with about 700 km.

In short, 4 years of half-hourly solar irradiance data from 70 meteorological stations are used with a maximum distance between two stations equal to about 2000 km; to our knowledge, no study on transfer learning influence has ever been carried out on such a large number of stations with such a large distance between them.

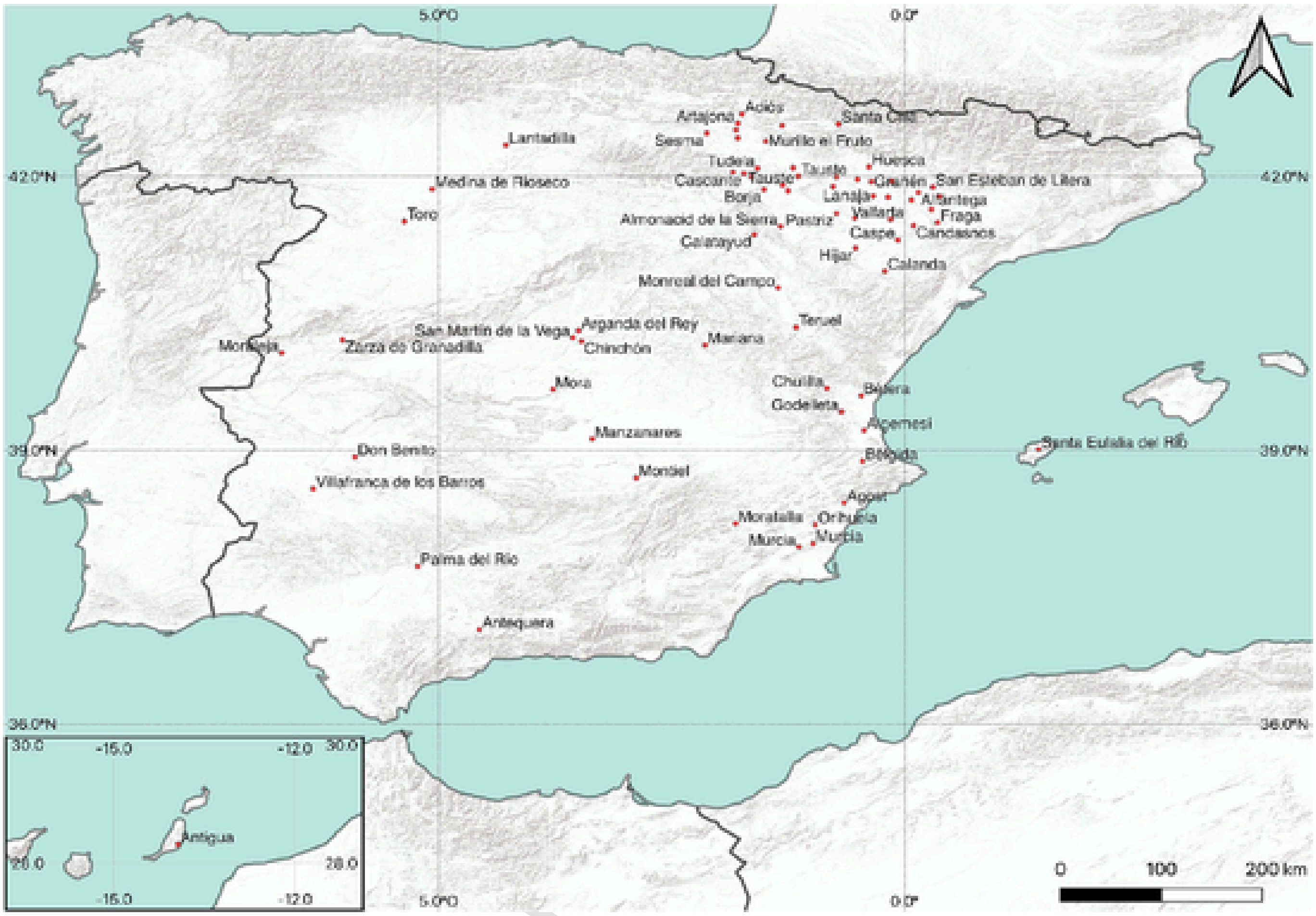

Fig. 3. Spatial distribution of locations of the stations used in the study.

1.4.2. Clustering methods: main results of a previous works

In a previous paper, Garcia-Gutierrez et al. [36] used five clustering methods in view to classify the 70 previous meteorological stations. The objective was to sort the stations not only from a quantitative point of view (available energy) but also and mainly from a qualitative point of view (variability of the solar radiation).

Only the main results are given here:

- each station was characterized by three parameters (defined below): the mean, the Hurst exponent and the forecastability index;
- four well-known clustering methods were used: hierarchical clustering, k-medoids clustering, spectral clustering and k-means clustering.
- a renowned classification for climate is the Köppen classification which were applied to our stations; it considers five main climate types, subdivided into a total of thirty classes with a series of letters that indicate the behavior of temperatures and precipitations. In Spain, 5 Köppen classes are existing; the second letter of this classification is relative to the precipitation regime which is the meteorological the most in relation with the solar radiation GHI; in Spain, 3 Köppen precipitation classes are present. Even if the solar potential is not considered in this classification, it seems that it can be used in our study as a reference. Thus, it was decided to choose a number of classes equal three and five.

As previously noted, each station was defined by three parameters described below; they were chosen after a redundancy study using the Spearman's Rank Correlation Coefficient between 6 descriptors:

- Forecastability index F [59] which characterizes how time series of solar radiation measurements at a particular site can be forecasted. This parameter is estimated from a Monte Carlo method to determine $RMSE_{max}$ (the maximum Root Mean Square Error related to a site where the *GHI* randomly oscillates between 0 and *GHI* under clear sky conditions and the RMSE of the smart persistence predictor ($RMSE_p$). F is calculated by Eq. (2):

$$F = 100\% \times \left(1 - RMSE_p / RMSE_{\max}\right) \tag{2}$$

The smaller the value F, the greater the difficulty to predict the solar time series. The value of F for each station is given in Table 2, it varies between 67.18% for Teruel (high-altitude town (≈900 m) in Aragon area, located in eastern Spain) and 76.32% for Zaidin (located in North-East of Spain in the Huesca province);

- Hurst exponent H: H is in relation with various fractal dimensions; it was used to quantify the variability (the intermittence) of the solar source [60–62]; the values of H for a time series with n components are in the range of 0 and 1 and are computed from the expected value (E[x]) such as:

Table 2
Locations used in the study and their partitioning using different clustering methods.

					3 clusters					5 clusters				
Station	Lat (°)	Lon (°)	Alt (m)	F (%)	a	b	c	d	e	a	b	c	d	e
Ablitas	42.00	−1.64	338	70.05	1	1	3	1	1	3	3	4	4	2
Adiós	42.69	−1.75	443	68.79	2	3	1	1	3	2	4	3	4	4
Agost	38.42	−0.65	941	71.91	3	2	2	1	1	1	5	1	4	2
Aibar/Oibar	42.56	−1.32	420	69.71	2	3	1	1	3	2	4	3	4	4
Alcolea de Cinca	41.74	0.07	225	75.72	1	1	2	1	1	5	1	2	4	2
Alfántega	41.82	0.15	249	75.33	1	1	2	1	1	5	1	2	4	2
Algemesí	39.22	−0.44	19	73.92	3	2	2	1	3	1	2	1	4	4
Almonacid de la Sierra	41.45	−1.33	384	70.62	1	1	3	1	1	3	3	4	4	2
Antequera	37.03	−4.56	457	73.46	3	2	2	1	1	1	5	1	4	2
Antigua	28.32	−13.94	72	73.83	3	2	2	1	1	1	2	1	1	2
Arganda del Rey	40.31	−3.50	531	73.43	1	1	2	1	1	5	1	2	4	2
Artajona	42.58	−1.79	360	68.89	2	3	1	1	2	2	4	3	4	3
Bèlgida	38.88	−0.45	281	73.01	3	2	2	1	3	1	2	1	4	4
Bétera	39.60	−0.47	97	73.82	3	2	2	1	1	1	5	1	4	2
Boquiñeni	41.84	−1.25	227	71.32	1	1	3	1	1	5	1	2	4	2
Borja	41.85	−1.51	378	70.33	1	1	3	1	3	3	3	4	4	4
Calanda	40.96	−0.21	439	71.79	1	1	2	1	1	5	1	2	4	2
Calatayud	41.36	−1.61	523	69.08	2	3	3	1	1	3	3	4	4	2
Candasnos	41.46	0.09	307	75.48	1	1	2	1	1	5	1	2	4	2
Cascante	42.03	−1.72	346	69.68	2	3	3	1	1	3	3	4	4	2
Caspe	41.30	−0.07	175	74.69	1	1	3	1	1	3	3	4	4	2
Chinchón	40.19	−3.47	534	72.70	1	1	3	1	1	3	3	4	4	2
Chulilla	39.68	−0.83	378	73.14	3	2	2	1	1	1	2	1	4	2
Don Benito	38.93	−5.90	268	75.35	3	2	2	1	1	1	2	1	4	2
Ejea de los Caballeros	42.10	−1.20	316	73.04	1	1	3	1	2	5	1	2	4	3
Falces	42.42	−1.79	292	69.31	2	3	3	1	3	3	4	4	4	4
Fitero	42.05	−1.84	436	68.78	2	3	1	1	3	2	4	3	4	4
Fraga	41.49	0.35	98	73.75	1	1	3	1	1	3	3	4	4	2
Godelleta	39.42	−0.68	270	71.87	3	2	2	1	3	1	2	1	4	4
Grañén	41.94	−0.36	323	74.71	1	1	2	1	1	5	1	2	4	2
Gurrea de Gállego	41.99	−0.73	364	73.83	1	1	2	1	1	5	1	2	4	2
Híjar	41.21	−0.53	306	73.90	1	1	2	1	1	5	1	2	4	2
Huerto	41.95	−0.14	415	75.65	3	2	2	1	1	1	2	1	4	2
Huesca	42.10	−0.38	432	74.17	2	3	3	1	1	3	3	4	4	2
Lanaja	41.79	−0.34	361	74.02	1	1	2	1	1	5	1	2	4	2
Lantadilla	42.35	−4.28	793	69.95	2	3	1	1	3	3	4	3	4	5
Manzanares	39.12	−3.35	645	74.59	3	2	2	1	1	1	2	1	4	2
Mariana	40.15	−2.14	941	67.25	2	3	1	1	1	2	4	3	4	2
Medina de Rioseco	41.86	−5.07	727	71.24	2	3	3	1	3	3	3	4	4	5
Miranda de Arga	42.51	−1.81	345	69.05	2	3	1	1	3	2	4	3	4	4
Monreal del Campo	40.78	−1.36	950	68.26	2	3	3	1	1	3	3	4	4	2
Montiel	38.70	−2.88	887	72.42	1	1	2	1	1	5	1	2	4	2
Mora	39.66	−3.77	735	72.49	3	2	2	1	1	1	2	1	4	2
Moraleja	40.07	−6.69	272	73.74	1	1	2	1	1	5	1	2	4	2
Moratalla	38.20	−1.81	458	73.22	3	2	2	1	1	1	5	1	4	2
Murcia	37.98	−0.98	128	74.75	3	2	2	2	1	4	5	5	3	1
Murcia	37.94	−1.13	54	73.97	3	2	2	1	1	1	2	1	4	1
Murillo el Fruto	42.38	−1.49	348	70.52	1	1	3	1	3	3	3	4	4	4
Orihuela	38.18	−0.96	96	74.25	3	2	2	2	1	4	5	5	3	2
Osera de Ebro	41.54	−0.54	251	74.63	3	2	2	1	1	1	2	1	4	2
Palma del Río	37.73	−5.23	57	76.79	3	2	2	1	1	1	5	1	4	2
Pastriz	41.59	−0.73	182	74.29	1	1	2	1	1	5	1	2	4	2
San Esteban de Litera	41.88	0.30	316	76.01	1	1	2	1	1	5	1	2	4	2
San Martín de la Vega	40.23	−3.56	516	73.65	1	1	2	1	1	5	1	2	4	2
Santa Cilia	42.58	−0.71	733	71.23	1	3	3	1	3	3	3	4	4	4
Santa Eulalia del Río	39.01	1.44	122	70.25	3	2	2	1	1	1	2	1	4	2
Sariñena	41.77	−0.18	291	74.90	1	1	2	1	1	5	1	2	4	2
Sesma	42.47	−2.13	456	67.75	2	3	1	1	3	2	4	3	4	4
Tamarite de Litera	41.78	0.38	218	75.41	1	1	3	1	1	5	1	2	4	2
Tardienta	41.97	−0.51	366	74.69	1	1	2	1	1	5	1	2	4	2
Tauste	42.00	−1.14	353	72.31	1	1	2	1	1	5	1	2	4	2
Tauste	41.90	−1.31	237	72.36	1	1	3	1	1	5	1	2	4	2
Teruel	40.35	−1.17	914	67.18	2	3	1	1	1	2	4	3	4	2
Toro	41.51	−5.37	652	71.39	1	1	3	1	3	3	3	4	4	4
Tudela	42.09	−1.58	243	71.60	1	3	3	1	1	3	3	4	4	2
Valfarta	41.53	−0.15	359	74.65	3	2	2	1	1	1	2	1	4	2
Villafranca de los Barros	38.57	−6.35	406	73.11	3	2	2	1	1	1	2	1	4	2
Zaidín	41.64	0.29	182	76.32	1	1	2	1	1	5	1	2	4	2
Zarza de Granadilla	40.21	−6.03	354	72.02	1	1	3	1	1	3	3	4	4	2
Zuera	41.89	−0.77	288	73.66	1	1	2	1	1	5	1	2	4	2

Note: a - k-means; b - spectral; c - k-medoids; d - hierarchical; e – Köppen.

$$E\left[\frac{R(n)}{S(n)}\right] = \alpha n^{H} \; as \; n \to \infty \tag{3}$$

Where *R(n)* is the range of the n cumulative deviations and *S(n)* is the sum of the first *n* standard deviations. *H* can be computed according to various methods [63];

- Mean value of *GHI*.

The results for these four clustering methodologies and for the Köppen classification according to three or five clusters are given in Table 2. These classifications will be used later in this paper.

As an example, Fig. 4 shows the results of the spectral clustering considering 5 clusters.

1.5. Forecasting methods

In this work, two models are used for forecasting: Extreme Learning Machine (ELM) and autoregressive (AR) model we will briefly present. The AR model is well known and is often used, it is applied in this work as a reference model.

Extreme Learning Machine (ELM) is a learning algorithm for single-hidden-layer neural networks that converges much faster (converges more rapidly than classical learning methods, making it appealing for real-time learning tasks) than traditional methods and shows promising performance (in areas such as classification, clustering, and regression). ELM is a powerful tool for solving machine learning problems, and its applications continue to diversify.

Due to all these advantages, and as all the methods for forecasting time series shows approbatively the same performances, we thought it would be interesting to apply this method to our problematic.

1.5.1. Autoregressive model

In statistics, econometrics and signal processing, an autoregressive (AR) model is a representation of a type of random process; as such, it is used to describe certain time-varying processes in nature, economics, behavior and so on. The autoregressive scheme specifies that the output variable depends linearly on its own past values and a stochastic term (an imperfectly predictable term) [25]. The model thus takes the form of a stochastic difference equation (or recurrence relation), which should not be confused with a differential equation. It is a special case and a key element of the more general Autoregressive Moving Average (ARMA) and Autoregressive Integrated Moving Average (ARIMA) models of time which have a more complicated stochastic structure. Unlike the moving average (MA) model, the autoregressive model is not al-

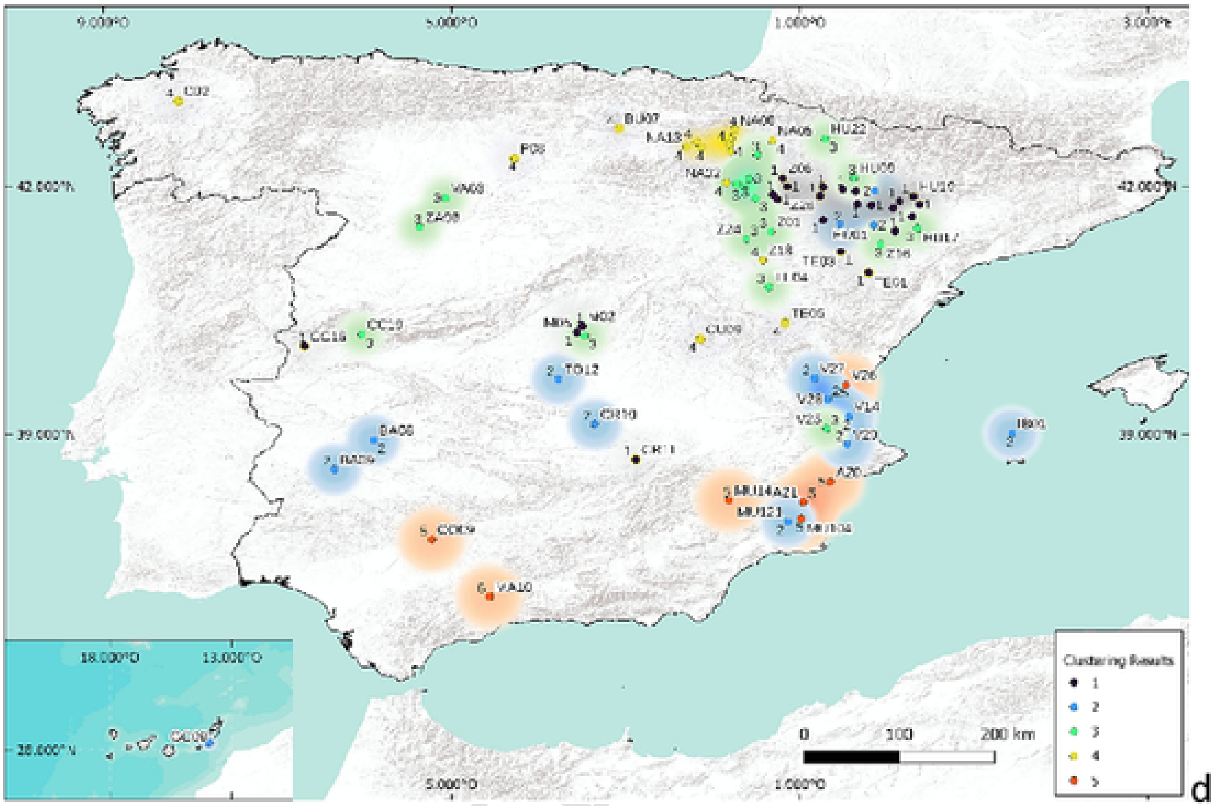

Fig. 4. Spectral Clustering results with 5 classes [36].

ways stationary, as it may contain a unit root [28]. This is the model that has been most widely used in *GHI* forecasting, because even though its intrinsic linearity may be a limitation to the quality of forecasts for sites with low forecastability, it is a simple model to implement and very often gives very good results [25]. The notation AR(p) indicates an autoregressive model of order p. The AR(p) model is defined as $X_t = \sum_{i=1}^{p} \varphi_i X_{t-i} + \epsilon_t$ where $\varphi_1 \cdots \varphi_p$ are the parameters of the model, and ϵ_t is a white noise.

1.5.2. Extreme learning Machine (ELM)

An Extreme Learning Machine (ELM) is a particular type of machine learning configuration [64]. An ELM comprises a number of hidden neurons where the input weights are randomly assigned. ELMs use the concept of random projection and early perceptron models to perform specific types of problem solving. This is a feedforward neural network, which means that data only flows in one direction through the series of layers. The learning time of an ELM is mainly spent on calculating the Moore-Penrose inverse matrices of the hidden layer output matrix. Proponents of ELMs argue that these feedforward networks are, in many ways, able to outperform networks using backpropagation, where information flows back through the network. ELMs can be useful for classification tasks, logical regression and clustering. Many recent studies applied this new forecasting method to solar radiation forecasting, we can mention in particular [65] for the coupling between ELM and the particle swarm optimization or [66] for the coupling with persistent model. Some authors improved ELM as in the case in [67] with the adaptative ELM, while others have used the speed of learning to offer online modes [68]. Given a single hidden layer of ELM, suppose that the output function of the i-th hidden node is $h_i(x) = G(a_i, b_i, x)$, where a_i and $b_i$are the parameters of the i-th hidden node, G an activation function and x a p-vector such as $x = [X_{t-p}, \ldots, X_{t-1}]$. The output function of the ELM for single hidden layer feedforward networks (this is the case for this study) with L hidden nodes is $X_t = \sum_{i=1}^{L} \varphi_i h_i(x) + \epsilon_t$. As a special case in this study, a simplest ELM training algorithm learns a model of the form (for single hidden layer sigmoid neural networks) $\hat{X}_t = W_2 \sigma(W_1 x)$, where W_1 is the matrix of input-to-hidden-layer weights, σ is an activation function, and W_2 is the matrix of hidden-to-output-layer weights. The algorithm proceeds as follows:

- Fill W_1 with random values (e.g., Gaussian random noise);
- Estimate W_2 by ordinary least-squares method to a matrix of response variables X_t, computed using the pseudoinverse +, given a design matrix x such as $W_2 = \sigma(W_1 x)^{+} X_t$.

1.6. Experiments and performance metrics

The first study pertains to the forecasting of *GHI* in all 70 selected locations using time series collected at these locations for the learning and test processes. This study refers to direct learning (DL). For training the model, data from the first three years of recorded values for the corresponding location (75% of the data for that location) have been used, and the data from the last year (25% of the data for that location) have been used for testing the models.

The second study refers to Transfer Learning (**TL**) and it pertains to forecast *GHI* in all selected locations using **DL** models from other locations (see Fig. 2). An analysis of the results will be then realized in distinguishing if the station used for the learning (Source) belongs or not to the same class that the test station (Target).

For this study a total of $70 \times 69 = 4830$ experiments have been conducted. All experiments have been carried out for twelve different horizons – from 30 min to 360 min, with a time step of 30 min; the forecasts are done every 30 min.

For both AR and ELM, it is required to determine the optimal number of the inputs corresponding to lagged values of *GHI*. For example, considering 10 inputs, the values of $GHI(t)$, $GHI(t-1)$...$GHI(t-9)$ are considered to predict $GHI(t+h)$ with h the horizon. Note that with 30 min time step, $h = 12$ corresponds to a prediction 360 min ahead, equivalent to 6 h ahead.

Prior to conducting the experiments, the optimum number of inputs (between 1 and 192) both for AR and ELM and the number of hidden neurons (between 24 and 600) for each horizon for ELM have been found by surrogate optimization (to find a global minimum of an objective function using few objective function evaluations). For reason of clarity, the optimized number of neurons and inputs will be not given for each station, each forecasting horizon and each source station; for example, these optimum values (considering prediction errors) are given in Table 3 for Agost station and ELM forecasting method.

Actually, as the aim of the study is to propose predictions on sites with no measurement history, it is not possible to a priori establish this optimization without measurements. It is important to mention that the number of inputs refers to the data collected over the previous 1.5–3 days (considering the horizon) and that the number of hidden nodes is about 4 times higher than the number of inputs. During this study, the simulations were performed using MATLAB (96 workers) running on a computer cluster with 28 Intel nodes (2 x Xeon 6230R (26 cores@2.1GHz) - 192Go de RAM DDR-4).

The performance of the proposed framework has been evaluated by Normalized Mean Absolute Error (nMAE), Normalized Root Mean Square Error (nRMSE), and Cluster Skill metrics that are described below. Night hours were not used during the calculation of these metrics, a filtration on the solar height (5°) has been made.

The mean absolute error (MAE) (in W/m^2) is defined by:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|\widehat{GHI}_{t_i} - GHI_{t_i}\right| \tag{4}$$

where N is the number of observations, $\widehat{GHI}_{t_i}$ is forecasted value of *GHI* at time t_i, and GHI_{t_i} is measured value of *GHI* at the same time. The normalized mean absolute error (nMAE) is obtained by dividing the MAE by the mean value of measured *GHI* calculated on N data:

$$nMAE = \frac{MAE}{\frac{1}{N}\sum_{i=1}^{N} GHI_{t_i}} \tag{5}$$

The Root Mean Square Error (RMSE) is a frequently used metric of the differences between the observed values and values predicted by a model. RMSE is more sensitive to important forecast errors, and hence is suitable for applications where small errors are more tolerable and larger errors lead to costs that are disproportionate, as in the case of utility applications, for example. It is defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{GHI}_{t_i} - GHI_{t_i}\right)^2} \tag{6}$$

Normalized Root Mean Square Error (nRMSE) is given as:

$$nRMSE = \frac{RMSE}{\frac{1}{N}\sum_{i=1}^{N} GHI_{t_i}} \tag{7}$$

In order to compare the performance of **TL** model against the reference scenario (**DL** model), i.e. in order to assess the effectiveness of using the forecasting method developed for a different site as compared to using the forecasting method developed for a given site, we proposed an extension of the classical skill score metric [28], which is defined as:

$$\mathrm{SkillScore}_{i,j,h} = 1 - \frac{\mathrm{Error\_TL}_{i,j,h}}{\mathrm{Error\_DL}_{j,h}} \tag{8}$$

where $Error\_TL$ is the nRMSE or nMAE error metric for Transfer Learning, and $Error\_DL$ is the nRMSE or nMAE error metric for Direct Learning (with that in mind, there are two SkillScore metrics: one that pertains to nRMSE, and the other that pertains to nMAE). Index i refers to testing station, index j refers to training station, and index h refers to horizon. A negative value of SkillScore metric indicates that the estimator (TL) is less efficient than the reference model (DL), zero indicates that the performance is similar, and positive values of that score indicate that the performance is better compared with the reference model.

In order to estimate the influence of clustering on forecast accuracy, we introduced a new metric called ClusterSkill, with an intention to define a single metric to quantify the forecast error for all sites and all horizons by averaging the $\mathrm{SkillScore}_{i,j,h}$ as follows:

$$\mathrm{ClusterSkill} = \frac{1}{HN(N-1)}\sum_{i=1}^{N}\sum_{\substack{j=1 \\ i\neq j}}^{N}\sum_{h=1}^{H} SkillScore_{i,j,h} \tag{9}$$

where H is the maximum horizon considered ($H = 12$; for 30-min data these is equivalent to prediction between 30 min and 360 min) and N is the number of considered stations. Consequently, there are two ClusterSkill metrics - one that pertains to nRMSE, and the other that pertains to nMAE.

The ClusterSkill metric for the cases where the training and testing stations are in the same cluster is labelled as InSample, while the OutSample label denotes that the training and testing stations are in the different clusters.

A ClusterSkill equal to 0 indicates that **TL** is as reliable as **DL** and that the lack of information concerning the testing site is not detrimental and the training is fully generalizable. A value of ClusterSkill higher than 0 means that **TL** is more efficient than DL, which makes no logical sense and can only occur due to numerical calculations and high variability in the data. It can't be significant. Finally, a negative ClusterSkill value indicates that **TL** induces an increase in terms of forecast error (expressed either by nRMSE or nMAE), concerning all the horizons and all the stations.

Table 3
Number of inputs and hidden neurons for Agost station and ELM model.

Horizon (min)	Number of inputs	Number of hidden neurons
30	78	472
60	110	485
90	108	465
120	105	419
150	117	373
180	135	491
210	141	478
240	136	443
270	136	463
300	138	450
330	142	475
360	133	468

1.7. Results and discussion

1.7.1. Prediction with direct learning

nRMSE and nMAE for Direct Learning (training and testing data sets referring to the same station) for horizons up to six hours are shown in Fig. 5. The circles are the average values for all the stations, each point represents the result for a given station and the probability density function for each horizon is also shown. These results show that both AR and ELM are good predictors of Global Horizontal Irradiance (GHI), but ELM outperforms AR model for all horizons. As expected, both er-

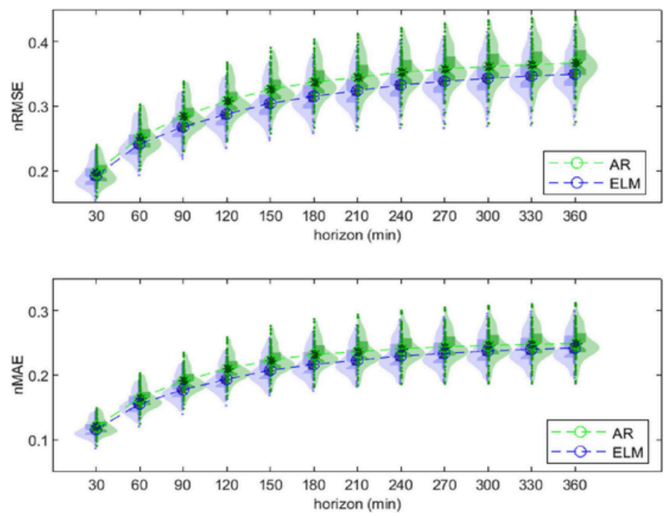

Fig. 5. nRMSE (top) and nMAE (bottom) errors for Direct Learning (DL) for horizons in the range of half-an-hour to six hours (values equal 0.0039 and 0.0098 for *nRMSE* and nMAE respectively).

ror metrics increase when horizon increases but the errors tend to stabilize and tend toward an asymptote beyond a forecasting horizon of about 270 min i.e. about 4.5 h. The largest mean nRMSE (circles) for horizons of six hours is 0.3483 for ELM and 0.3639 for AR. As for the nMAE, the largest mean values are 0.2405 and 0.2473 for ELM and AR respectively.

In addition, on two sets of errors obtained using ELM and AR models, we conducted the Mann-Whitney *U* test [66] with the null hypothesis that there is no significant difference in distribution of errors in these two models. The Mann-Whitney U test is a non-parametric test for comparing the means of two samples (paired in the context of the study). A *p*-value of <0.05 indicates that the difference between the two means is significant. The mean *p*-values (averaged over all twelve horizons) equal 0.0039 for nRMSE and 0.0098 for nMAE and they are both lower than the significance level $\alpha = 0.05$, meaning that the null hypothesis is rejected, i.e., that there is a significant difference between these two models. Because ELM was shown to be better predictor, in our further analysis we used only ELM.

1.7.2. Transfer learning: 2 sites validations

Forecasted solar irradiance profiles for the station Antequera, for four randomly picked days from different seasons, obtained using the **TL** are given in Figure . For this simulation, the model was trained using data from the station Adiós located 674 km North-east of Antequera (measured using Google Maps tools). Antequera has a warm Mediterranean climate with dry summers. The subtype for this climate is ‘Csa’ according to the Köppen-Geiger classification. Over the year, the average temperature in Antequera is 18 °C and the average rainfall is 584.7 mm. Adiós has a warm oceanic climate without a dry season (‘Cfb’ according to the Köppen-Geiger classification). Over the year, the average temperature in Adiós is 13.3 °C and the average rainfall is 654.7 mm.

Although simulations include all possible combinations of learning stations and testing stations, we decided to present results of the simulations for these two stations, because we wanted to show results for **TL** between the stations that meet following criteria: 1) they are relatively far away from each other, and 2) one station belongs to Andalusia, and the other is near Pyrenees, thus stations are in the areas with very different climate. These results are given along with the results obtained using the DL, and they are both compared with measured values. To save space, only results referring to horizons of 30 min, 180 min, and 360 min are shown. Fig. 6 indicates that the errors associated with both **DL** and TL, increase when horizon increases: $nRMSE_{DL}$ equals 0.174, 0.267 and 0.293 and $nRMSE_{TL}$ equals 0.211 (+0.037/+21.3%), 0.269 (0.002/+0.75%) and 0.294 (0.001/0.34%) for horizon equals 30 min, 3 h and 12 h respectively. The $Skill\ Score_{nRMSE}$ is then −0.213, −0.007 and − 0.003 for these three horizons. The gap in nRMSE between **DL** and **TL** decreases when the time horizon increases. The other thing that can be concluded from Fig. 6 is that the error mainly pertains to cloudy days (winter and autumn), while for sunny days (summer and spring), there is a very good match between the predicted values and measured values for all horizons. In addition, even for cloudy days, the predicted

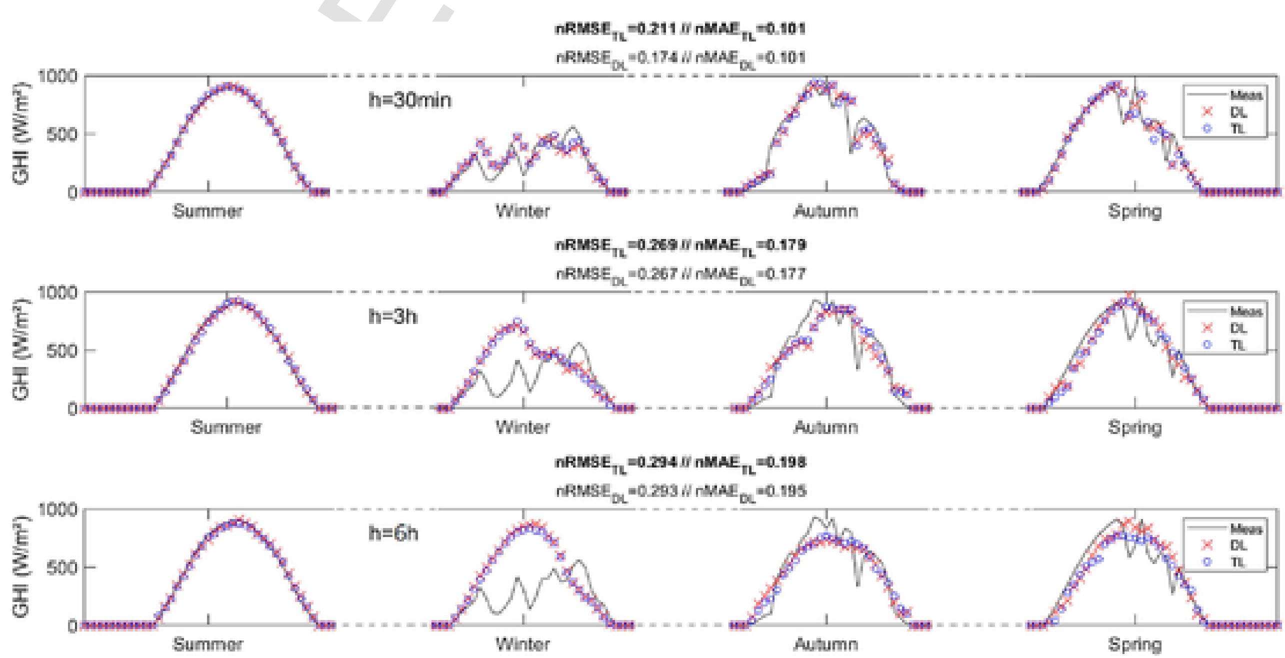

Fig. 6. Measurement vs Prediction for the case of **DL** vs Prediction for the case of **TL** for station (Antequera) for horizons of 30 min, 180 min, and 360 min.

values are very close to measured values for smaller horizons. Finally, and most importantly when it comes to transfer learning, this figure indicates that there is no large difference in predictions that are obtained using **DL** and TL. Note that this conclusion can be generalized to the reverse case: the learning and testing stations are reversed. Under those conditions and according to all horizons, nRMSE are between 0.226 and 0.415 for **DL** and 0.290 and 0.412 for **TL** while concerning nMAE values are between 0.145 and 0.296 for **TL** and between 0.139 and 0.300 for DL.

1.7.3. Transfer learning: Statistical validation

The heat maps of nRMSE that are calculated after simulations of all possible combinations of training stations and testing stations ($70 \times 69 = 4830$) for three different horizons (30 min, 180 min and 360 min) are shown in Fig. 7. Consequently, the values that are on the main diagonal of the heat maps pertain to **DL** (same target and source stations), while all other values pertain to TL. These heat maps clearly indicate that there is no large variation in values of nRMSE for the same testing (target) station even if this small variation tends to increase with the forecasting horizon. Such a result was previously shown in Italy [50] but only for four stations geographically very close.

The SkillScore metrics for horizons 30 min, 180 min and 360 min are shown in Fig. 8. Consequently, the values that are on the main diagonal are equal to zero, because they refer to DL. Fig. 8 visually indicates that the distribution of SkillScore metrics is quite uniform, and the ranges of SkillScore metrics, that are [−0.0598, 0.0195] for horizon of 30 min and [−0.0829, 0.0499] for horizon of 12 h, show no large differences in prediction errors regarding the reference (DL) model. As some Skill Score are positive.

These SkillScore metrics are statistically presented for the station Antequera and for all twelve horizons in Fig. 9, a special feature is observed because for particular horizons (such as h = 300 min), **DL** seems globally less interesting than **TL** (median SkillScore > 0) which may seem a little unsettling. There don't seem to be any physical or mathematical explanations for this phenomenon; it's probably linked to statistical fluctuations related to sampling and perhaps also linked to the quality of the data in some stations (as for station 11 in Figs. 7 and 8). It seems marginal and insignificant.

The variation of nRMSE when the target station changes for a same source station is high (high variation of colours by line) whereas the variation of nRMSE when the source station changes for the same target source is low (low variation of colours by column). What influences the error metric (nMAE or nRMSE) is essentially the variability (or forecastability) of the target station. If a data series of a target station has a high variability, whatever the source data are used, the performances of the forecasting method will be low.

1.7.4. Transfer learning: Clustering impact

The ClusterSkill metrics for different clustering algorithms and for the case of three classes are shown in Fig. 10. InSample refers to a transfer learning for which the source and target data stations belongs to the same cluster; OutSample refers to a transfer learning for which the

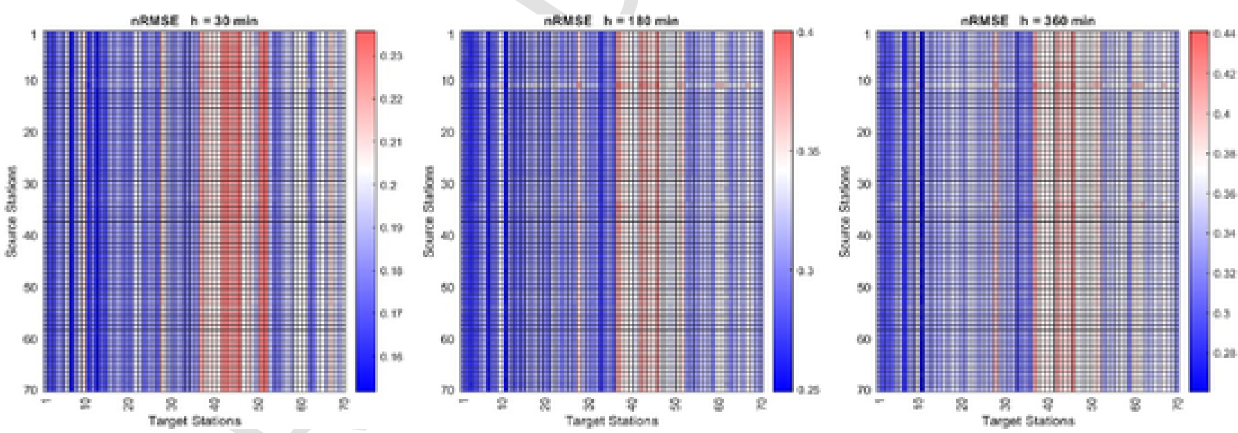

Fig. 7. Heat maps of nRMSE for horizons 30 min (left), 180 min (middle) and 360 min (right).

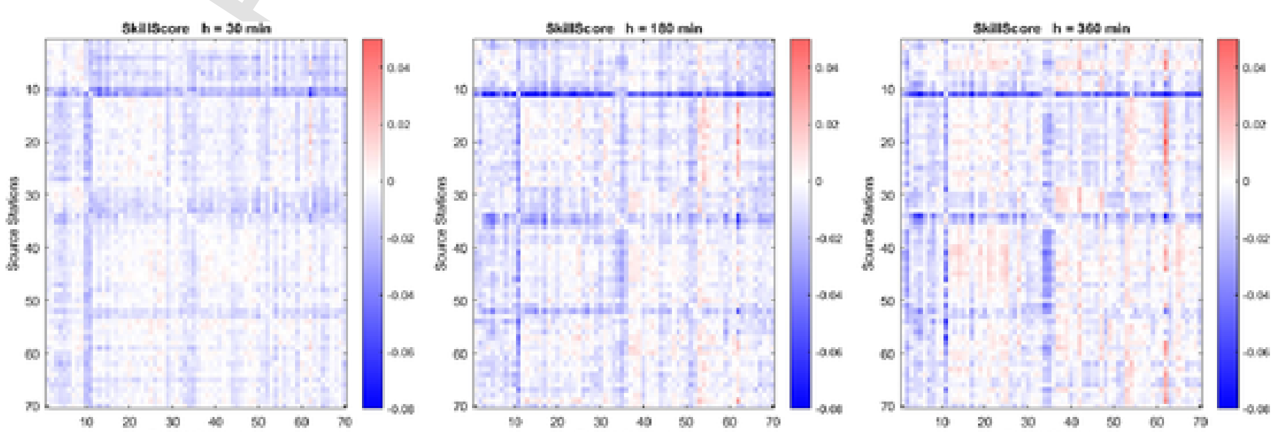

Fig. 8. Heat maps of SkillScore for horizons 30 min (left), 180 min (middle) and 360 min (right).

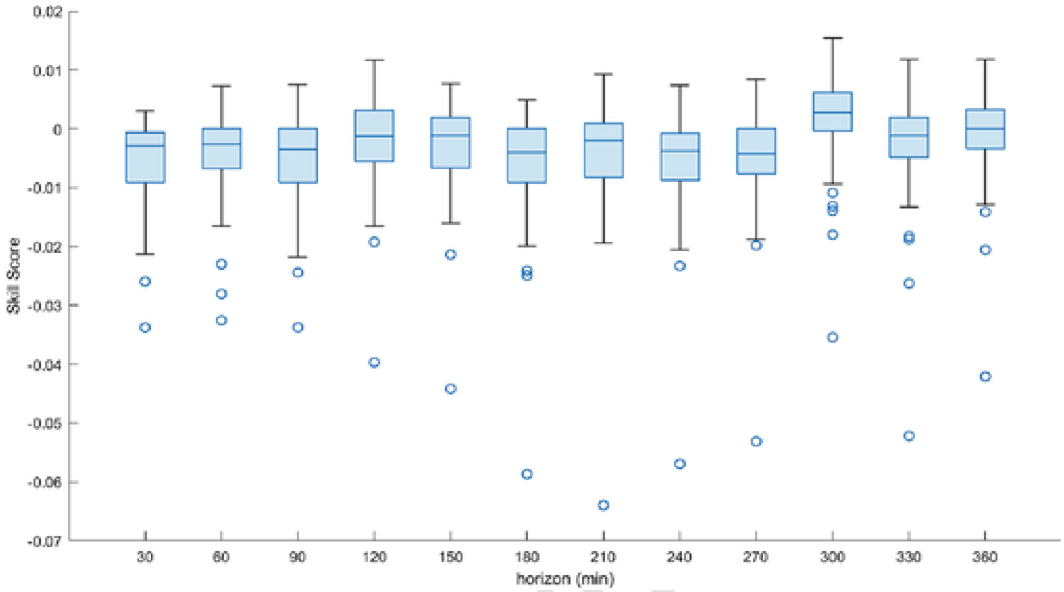

Fig. 9. SkillScores for Antequera for different horizons. Box plot concerning minimum and maximum (excluding any outlier), median, first and third quartiles (respectively Q1 and Q3) and outliers defined by points outside 1.5 times the interquartile range (Q3-Q1).

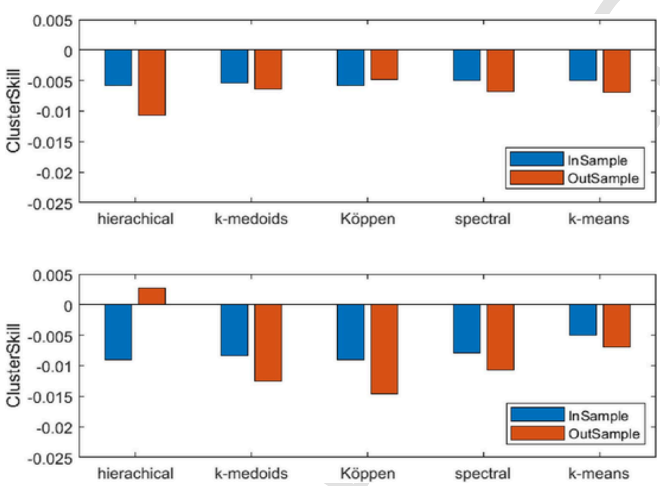

Fig. 10. ClusterSkill metrics for different clustering algorithms for the case of three classes (nRMSE – top; nMAE - bottom).

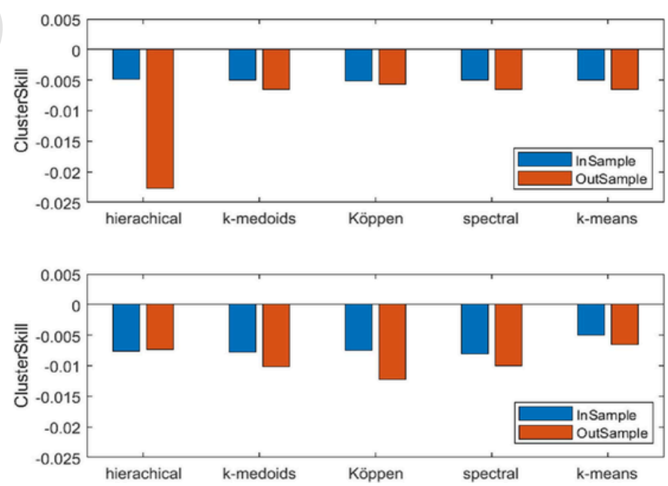

Fig. 11. ClusterSkill metrics for different clustering algorithms for the case of five classes (nRMSE - top; nMAE - bottom).

source and target data stations does not belong to the same cluster. Remember that more the ClusterSkill is equal to a high negative value, more the performance of the forecasting using a **TL** is less efficient. When nRMSE is considered, the forecasting using a **TL** with data measured in a station belonging to the same cluster performs better than with data measured in station outside the same cluster except when the Köppen classification is applied; as underlined in [36], Köppen classification was developed to characterize the climate of a station (Climate type, rainfall regime and temperature variation), but not the solar radiation variation, it is then not well adapted to our objectives.

With the hierarchical clustering, the difference between nRMSE values related to InSample and OutSample data are the higher but this clustering method gave a non-uniform repartition of stations by cluster [36] (out of 70 stations, 68 stations belong to one class, two stations belong to the other class, and the third class is empty). The major part of the stations belongs to InSample group. Consequently, this clustering method did not give satisfying results in [36] and the results, here, must be taken with caution.

In addition, when nRMSE is considered, the very similar values of ClusterSkill for spectral and k-means clustering (−0.503% vs. -0.499% respectively for InSample and − 0.682% vs. -0.689% respectively for OutSample) are very much in line with the previous study [36] that showed high similarity between these two clustering (Rand index = 95% and Jaccard index = 96%).

When nMAE is considered, one more time the hierarchical clustering doesn't seem to be a good method because the ClusterSkill for OutSample is higher than the ClusterSkill for InSample. Bearing in mind the other findings regarding nRMSE, the hierarchical clustering cannot be recommended for transfer learning. On the other hand, Köppen classification underperforms other clustering methods regarding transfer learning as already seen previously. In general, when it comes to transfer learning and three classes, Fig. 10 indicates that the best clustering method considering nRMSE is k-medoids followed by spectral (k-means), and when nMAE is considered, the best clustering method is k-means.

The ClusterSkill metrics for different clustering algorithms and the case of five classes are shown in Fig. 11. On the first glance, these results indicate almost identical findings for the case of three classes except that Köppen classification. For ClusterSkill calculated using nRMSE error metrics, hierarchical clustering shows very large difference between InSample and OutSample values. Similar to the case of three classes, this clustering method is not fully validated because one class includes 67 stations, the other class includes two stations, the third class includes one station, and the two classes are empty. The other clustering methods show identical InSample ClusterSkill values (−0.5%), whereas the OutSample ClusterSkill value is slightly better for Köppen classification (−0.56%) than for k-medoids, spectral or k-means (−0.64%). These findings are also in line with the aforementioned previous study [36] when it comes to similarity between k-medoids, spectral and k-means (Rand index in the range from 95% to 99%, and Jaccard index in the range from 91% to 99%), but not in line with the conclusion derived for similarity with Köppen classification. When nMAE is considered, the hierarchical clustering, similar to the case of three classes, is not recommended, and among the other clustering, the k-means again is the best clustering method for transfer learning.

Table 4 presents the average values, minimum and maximum values of nRMSE obtained for three forecasting horizons (30, 180 and 360 min) for **DL** and **TL** (using stations regardless of their class membership), then using stations belonging to the same class (InSample) or outside of the same class (OutSample) considering the five clustering methods with for each of them 3 or 5 classes.

The average value of nRMSE is slightly smaller for InSample that OutSample and the gap between the lowest and the highest values is smaller too but the difference is not significative. The difference of performances between InSample and OutSample increases with the forecasting horizon but not in a very frank way.

The Hierarchical clustering method shows the highest difference between nRMSE with InSample data and OutSample data because, as for this method there is a cluster with a very high number of stations and the others with a very small number, the results calculated with OutSample is realized with a very small number of stations with solar radiation characteristics different to the target stations.

Globally, the application of a clustering method on the stations before performing the forecasting does not allow to improve a lot the performances of the forecasting model but complicates the global method and is more time-consuming. This low influence of the clustering on the improvement of the performance of the forecasting model was confirmed by reference [51] in Morocco.

2. Conclusions

Although the forecast of global radiation has never given rise to as many studies as in recent years, there are very few studies that deal with the real problem linked to the installation of a power plant or setting up a smart grid in an area where very few measurements are available. One year of measurements is a prohibitive limit, but solar engineers using artificial intelligence techniques know that 3 years (at least) are necessary for quality forecasts. Some take the side of using NWP outputs, but this is at the risk of poor predictive quality for short and very short horizons.

Transfer learning (**TL**) or rather the transfer of the learning (perhaps a more adequate term) is feasible and shows that in many respects, this method is robust for horizons ranging from 30 min to 6 h (error lower than 1 percentage point related to direct learning). The use of an ELM model is more efficient than the classical AR model and makes it possible to respond to this problem by integrating between 1 day and 3 days of data at the network input.

Finally, concerning the use of the station clustering and its contribution to the decrease of the prediction error, the gain is very low, and as a first approximation, within the framework of the study presented here, it does not seem appropriate to complicate the methodology by adding a clustering phase. However, the user who wishes to improve the performance of these forecasts by some 0.3 percentage points can use a classic k-means method.

CRediT authorship contribution statement

Milan Despotovic: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis. **Cyril Voyant:** Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization, Visualization, Writing – original draft,

Table 4
Average, Minimum and Maximum values (in %) of nRMSE for three forecasting horizons for DL, **TL** and **TL** coupled with cluterings (The number between brackets is the number of clusters).

Horizon		30 min			180 min			360 min		
		Mean	Min	Max	Mean	Min	Max	Mean	Min	Max
DL		**19.077**	15.199	23.050	**31.333**	24.875	37.849	**34.830**	26.270	41.872
TL		**19.187**	15.200	23.576	**31.544**	24.926	40.018	**34.991**	26.524	44.120
Hierarchical 3	InS	**19.127**	15.188	23.433	**31.433**	25.088	38.311	**34.863**	26.150	42.568
	OutS	**20.214**	15.156	23.576	**33.427**	24.875	40.018	**37.169**	26.032	44.121
Hierarchical 5	InS	**19.090**	15.156	23.492	**31.395**	24.875	39.481	**34.811**	26.032	43.056
	OutS	**20.047**	21.010	21.312	**33.089**	21.010	40.018	**36.726**	21.010	44.121
k-Medoids 3	InS	**19.086**	15.156	23.521	**31.410**	24.875	40.018	**34.847**	26.032	43.013
	OutS	**19.320**	15.196	23.576	**31.719**	26.672	41.321	**35.179**	26.530	44.121
k-Medoids 5	InS	**19.181**	15.200	23.446	**31.509**	24.875	39.149	**34.952**	26.032	43.013
	OutS	**19.190**	15.156	23.446	**31.556**	25.649	40.018	**35.004**	27.949	44.121
Köppen 3	InS	**19.159**	15.156	23.446	**31.583**	24.875	40.018	**34.864**	26.032	43.013
	OutS	**19.388**	15.188	23.576	**31.817**	25.649	41.321	**35.303**	26.530	44.121
Köppen 5	InS	**19.081**	15.188	23.462	**31.408**	24.875	39.481	**34.944**	26.032	43.656
	OutS	**19.292**	15.156	23.576	**31.647**	24.926	40.018	**35.133**	26.524	44.121
Spectral 3	InS	**19.160**	15.200	23.557	**31.543**	24.875	39.149	**34.990**	26.032	43.013
	OutS	**19.378**	15.156	23.576	**31.544**	24.926	40.018	**34.991**	26.524	44.121
Spectral 5	InS	**19.168**	15.200	23.446	**31.527**	24.875	39.149	**34.970**	26.032	43.013
	OutS	**19.193**	16.429	23.576	**31.549**	26.092	40.118	**34.997**	29.083	44.121
k-Means 3	InS	**19.123**	15.200	23.557	**31.439**	24.875	39.149	**34.860**	26.032	43.013
	OutS	**19.278**	15.156	23.576	**31.606**	24.926	40.018	**35.068**	26.524	44.121
k-Means 5	InS	**19.175**	15.200	23.446	**31.492**	24.875	39.149	**34.927**	26.032	43.013
	OutS	**19.192**	15.156	23.576	**31.562**	24.956	40.018	**35.013**	26.530	44.121

Writing – review & editing. **Luis Garcia-Gutierrez:** Data curation, Software, Visualization. **Javier Almorox:** Resources. **Gilles Notton:** Writing – review & editing, Supervision, Methodology, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no shown competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors would like to thank (i) the network of agroclimatic stations of the Spanish Ministry of Agriculture, Food, and Environment, specifically the agrometeorological stations network of SIAR for the information provided on solar irradiance (https://eportal.mapa.gob.es) and (ii) the atmospheric monitoring service CAMs-McClear for the information on solar irradiance under clear sky conditions (http://www.soda-pro.com).

This work was partially supported by ANR grant SAPHIR project ANR-21-CE04-0014-03.

References

[1] IEA. Empowering cities for a net zero future: Unlocking resilient, smart, sustainable urban energy systems. Paris: OECD Publishing; 2021. https://doi.org/10.1787/7a222c8b-en.

[2] Calvillo C.F, Sánchez-Miralles A, Villar J. Energy management and planning in smart cities. Renew Sustain Energy Rev 2016;55:273–87. https://doi.org/10.1016/J.RSER.2015.10.133.

[3] WoriGHI I, Maach A, Hafid A, Hegazy O, van Mierlo J. Integrating renewable energy in smart grid system: architecture, virtualization and analysis. Sustain Energy, Grids Networks 2019;18:100226. https://doi.org/10.1016/j.segan.2019.100226.

[4] Notton G, Nivet M.L, Voyant C, Paoli C, Darras C, Motte F, et al. Intermittent and stochastic character of renewable energy sources : consequences, cost of intermittence and benefit of forecasting. Renew Sustain Energy Rev 2018;87:96–105. https://doi.org/10.1016/j.rser.2018.02.007.

[5] Ullah K, Hafeez G, Khan I, Jan S, Javaid N. A multi-objective energy optimization in smart grid with high penetration of renewable energy sources. Appl Energy 2021;299:117104. https://doi.org/10.1016/j.apenergy.2021.117104.

[6] Chaudhary P, Rizwan M. Energy management supporting high penetration of solar photovoltaic generation for smart grid using solar forecasts and pumped hydro storage system. Renew Energy 2018;118:928–46. https://doi.org/10.1016/j.renene.2017.10.113.

[7] Fan S, Methaprayoon K, Lee W.J. Multiregion load forecasting for system with large geographical area. IEEE Trans Ind Appl 2009;45:1452–9. https://doi.org/10.1109/TIA.2009.2023569.

[8] Hong T, Wilson J, Xie J. Long term probabilistic load forecasting and normalization with hourly information. IEEE Trans Smart Grid 2014;5:456–62. https://doi.org/10.1109/TSG.2013.2274373.

[9] Nose-Filho K, Lotufo A.D.P, Minussi C.R. Short-term multinodal load forecasting using a modified general regression neural network. IEEE Trans Power Deliv 2011;26:2862–9. https://doi.org/10.1109/TPWRD.2011.2166566.

[10] Fay D, Ringwood J.V. On the influence of weather forecast errors in short-term load forecasting models. IEEE Trans Power Syst 2010;25. https://doi.org/10.1109/TPWRS.2009.2038704.

[11] Khodayar M.E, Wu H. Demand forecasting in the smart grid paradigm: features and challenges. Elect J 2015;28:51–62. https://doi.org/10.1016/J.TEJ.2015.06.001.

[12] Khan A.R, Mahmood A, Safdar A, Khan Z.A, Khan N.A. Load forecasting, dynamic pricing and DSM in smart grid: a review. Renew Sustain Energy Rev 2016;54:1311–22. https://doi.org/10.1016/j.rser.2015.10.117.

[13] Vrablecová P, Bou Ezzeddine A, Rozinajová V, Šárik S, Sangaiah A.K. Smart grid load forecasting using online support vector regression. Comp Elect Eng 2018;65:102–17. https://doi.org/10.1016/J.COMPELECENG.2017.07.006.

[14] Wan C, Zhao J, Song Y, Xu Z, Lin J, Hu Z. Photovoltaic and solar power forecasting for smart grid energy management. CSEE J Power Energy Syst 2016;1:38–46. https://doi.org/10.17775/CSEEJPES.2015.00046.

[15] Manur A, Marathe M, Manur A, Ramachandra A, Subbarao S, Venkataramanan G. Smart solar home system with solar forecasting. In: 2020 IEEE international conference on power electronics, smart grid and renewable energy. PESGRE; 2020. p. 2020. https://doi.org/10.1109/PESGRE45664.2020.9070340.

[16] Tripathy D.S, Prusty B.R. Chapter 10 - forecasting of renewable generation for applications in smart grid power systems. In: Tomar Anuradha, Kandari Ritu, editors. Advances in smart grid power system: Network, control and security. Academic Press; 2021. p. 265–98. https://doi.org/10.1016/B978-0-12-824337-4.00010-2.

[17] Feng C, Zhang J. SolarNet: a sky image-based deep convolutional neural network for intra-hour solar forecasting. Solar Energy 2020;204:71–8. https://doi.org/10.1016/j.solener.2020.03.083.

[18] Paletta Q, Arbod G, Lasenby J. Benchmarking of deep learning irradiance forecasting models from sky images – an in-depth analysis. Solar Energy 2021;224:855–67. https://doi.org/10.1016/j.solener.2021.05.056.

[19] Manandhar P, Temimi M, Aung Z. Short-term solar radiation forecast using total sky imager via transfer learning. Energy Rep 2023;9:819–28. https://doi.org/10.1016/J.EGYR.2022.11.087.

[20] Voyant C, Notton G, Duchaud J.L, Gutiérrez L.A.G, Bright J.M, Yang D. Benchmarks for solar radiation time series forecasting. Renew Energy 2022;191:747–62. https://doi.org/10.1016/j.renene.2022.04.065.

[21] Sharadga H, Hajimirza S, Balog R.S. Time series forecasting of solar power generation for large-scale photovoltaic plants. Renew Energy 2020;150:797–807. https://doi.org/10.1016/j.renene.2019.12.131.

[22] Verbois H, Huva R, Rusydi A, Walsh W. Solar irradiance forecasting in the tropics using numerical weather prediction and statistical learning. Solar Energy 2018;162:265–77. https://doi.org/10.1016/j.solener.2018.01.007.

[23] Zhang G, Yang D, Galanis G, Androulakis E. Solar forecasting with hourly updated numerical weather prediction. Renew Sustain Energy Rev 2022;154:111768. https://doi.org/10.1016/j.rser.2021.111768.

[24] Yang D, Wang W, Bright J.M, Voyant C, Notton G, Zhang G, et al. Verifying operational intra-day solar forecasts from ECMWF and NOAA. Solar Energy 2022;236:743–55. https://doi.org/10.1016/j.solener.2022.03.004.

[25] Ahmed R, Sreeram V, Mishra Y, Arif M.D. A review and evaluation of the state-of-the-art in PV solar power forecasting: techniques and optimization. Renew Sustain Energy Rev 2020;124:109792. https://doi.org/10.1016/j.rser.2020.109792.

[26] Kumari P, Toshniwal D. Deep learning models for solar irradiance forecasting: a comprehensive review. J Clean Prod 2021;318. https://doi.org/10.1016/j.jclepro.2021.128566.

[27] Kumari P, Toshniwal D. Long short-term memory–convolutional neural network based deep hybrid approach for solar irradiance forecasting. Appl Energy 2021;295:117061. https://doi.org/10.1016/j.apenergy.2021.117061.

[28] Voyant C, Notton G, Kalogirou S, Nivet M.L, Paoli C, Motte F, et al. Machine learning methods for solar radiation forecasting: a review. Renew Energy 2017;105:569–82. https://doi.org/10.1016/j.renene.2016.12.095.

[29] Hossain M, Mekhilef S, Danesh M, Olatomiwa L, Shamshirband S. Application of extreme learning machine for short term output power forecasting of three grid-connected PV systems. J Clean Prod 2017;167:395–405. https://doi.org/10.1016/j.jclepro.2017.08.081.

[30] Keshtegar B, Bouchouicha K, Bailek N, Hassan M.A, Kolahchi R, Despotovic M. Solar irradiance short-term prediction under meteorological uncertainties: survey hybrid artificial intelligent basis music-inspired optimization models. Eur Phys J Plus 2022;137:362. https://doi.org/10.1140/epjp/s13360-022-02371-w.

[31] Haider S.A, Sajid M, Sajid H, Uddin E, Ayaz Y. Deep learning and statistical methods for short- and long-term solar irradiance forecasting for Islamabad. Renew Energy 2022;198:51–60. https://doi.org/10.1016/j.renene.2022.07.136.

[32] Paulescu M, Paulescu E, Gravila P, Badescu V. Weather Modeling and Forecasting of PV Systems Operation, Green Energy and Technology. vol. 103. London: Springer Verlag; 2013. https://doi.org/10.1007/978-1-4471-4649-0_2.

[33] Cros S, Wald L. Survey of the main databases providing solar radiation data at ground level. In: Goossens R.I, editor. 23rd EARSeL annual symposium "remote sensing in transition", Ghent, Belgium; 2003. p. 491–7.

[34] Pan S.J, Yang Q. A survey on transfer learning. IEEE Trans Knowl Data Eng 2010;22:1345–59. https://doi.org/10.1109/TKDE.2009.191.

[35] Ben-David S, Blitzer J, Crammer K, Kulesza A, Pereira F, Vaughan J.W. A theory of learning from different domains. Mach Learn 2010;79:151–75. https://doi.org/10.1007/s10994-009-5152-4.

[36] Garcia-Gutierrez L, Voyant C, Notton G, Almorox J. Evaluation and comparison of spatial clustering for solar irradiance time series. Appl Sci (Switzerland) 2022;12. https://doi.org/10.3390/app12178529.

[37] Franses P.H. Seasonality, non-stationarity and the forecasting of monthly time series. Int J Forecast 1991;7–2:199–208. https://doi.org/10.1016/0169-2070(91)90054-Y.

[38] Livieris I.E, Stavroyiannis S, Pintelas E, Pintelas P. A novel validation framework to enhance deep learning models in time-series forecasting. Neural Comput Applic 2020;32(23):17149–67. https://doi.org/10.1007/s00521-020-05169-y.

[39] Livieris I.E. A novel forecasting strategy for improving the performance of deep learning models. Exp Syst Appl 2023;230:120632. https://doi.org/10.1016/j.eswa.2023.120632.

[40] Lauret P, Alonso-Suárez R, Le Gal La Salle J, David M. Solar forecasts based on the clear sky index or the clearness index: which is better? Solar 2022;2:432–44. https://doi.org/10.3390/solar2040026.

[41] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. Neural Netw 1989;2:359–66. https://doi.org/10.1016/0893-6080(89)90020-8.

[42] Trapletti A, Leisch F, Hornik K. Stationary and integrated autoregressive neural network processes. Neural Comput 2000;12:2427–50. https://doi.org/10.1162/089976600300015006.

[43] Yang D. Choice of clear-sky model in solar forecasting. J Renew Sustain Energy

2020;12:026101. https://doi.org/10.1063/5.0003495.
[44] Gschwind B, Wald L, Blanc P, Lefèvre M, Schroedter-Homscheidt M, Arola A. Improving the McClear model estimating the downwelling solar radiation at ground level in cloud-free conditions-McClear-v3. Meteorol Z 2019;28:147–63. https://doi.org/10.1127/metz/2019/0946.
[45] Qu Z, Oumbe A, Blanc P, Espinar B, Gesell G, Gschwind B, et al. Fast radiative transfer parameterisation for assessing the surface solar irradiance: the Heliosat-4 method. Meteorol Z 2017;26:33–57. https://doi.org/10.1127/metz/2016/0781.
[46] Rezaei S, Tahmoresnezhad J, Solouk V. A transductive transfer learning approach for image classification. Int J Mach Learn Cyber 2021;12:747–62. https://doi.org/10.1007/s13042-020-01200-9.
[47] Himeur Y, Elnour M, Fadli F, Meskin N, Petri I, Rezgui Y, et al. Next-generation energy systems for sustainable smart cities: roles of transfer learning. Sustain Cities Soc 2022;85:104059. https://doi.org/10.1016/j.scs.2022.104059.
[48] Qureshi A.S, Khan A, Zameer A, Usman A. Wind power prediction using deep neural network based meta regression and transfer learning. Appl Soft Comput 2017;58:742–55. https://doi.org/10.1016/j.asoc.2017.05.031.
[49] Zhang W, Deng L, Zhang L, Wu D. A survey on negative transfer. IEEE/CAA J Automat Sinica 2022;1–25. https://doi.org/10.1109/JAS.2022.106004.
[50] Guariso G, Nunnari G, Sangiorgio M. Multi-step solar irradiance forecasting and domain adaptation of deep neural networks. Energies 2020;13. https://doi.org/10.3390/en13153987.
[51] Marzouq M, el Fadili H, Zenkouar K, Lakhliai Z, Amouzg M. Short term solar irradiance forecasting via a novel evolutionary multi-model framework and performance assessment for sites with no solar irradiance data. Renew Energy 2020;157:214–31. https://doi.org/10.1016/j.renene.2020.04.133.
[52] Luo X, Zhang D, Zhu X. Combining transfer learning and constrained long short-term memory for power generation forecasting of newly-constructed photovoltaic plants. Renew Energy 2022;185:1062–77. https://doi.org/10.1016/j.renene.2021.12.104.
[53] Tang Y, Yang K, Zhang S, Zhang Z. Photovoltaic power forecasting: a hybrid deep learning model incorporating transfer learning strategy. Renew Sustain Energy Rev 2022;162:112473. https://doi.org/10.1016/j.rser.2022.112473.
[54] Niu T, Li J, Wei W, Yue H. A hybrid deep learning framework integrating feature selection and transfer learning for multi-step global horizontal irradiation forecasting. Appl Energy 2022;326:119964. https://doi.org/10.1016/j.apenergy.2022.119964.
[55] Sarmas E, Dimitropoulos N, Marinakis V, Mylona Z, Doukas H. Transfer learning strategies for solar power forecasting under data scarcity. Sci Rep 2022;12:1–13. https://doi.org/10.1038/s41598-022-18516-x.
[56] Gao Y, Hu Z, Shi S, Chen W.A, Liu M. Adversarial discriminative domain adaptation for solar radiation prediction: a cross-regional study for zero-label transfer learning in Japan. Appl Energy 2024;359:122685. https://doi.org/10.1016/j.apenergy.2024.122685.
[57] Abubakr M, Akoush B, Khalil A, Hassan M.A. Unleashing deep neural network full potential for solar radiation forecasting in a new geographic location with historical data scarcity: a transfer learning approach. Eur Phys J Plus 2022;137–4: 474. https://doi.org/10.1140/epjp/s13360-022-02666-y.
[58] SiAR - The Spanish Ministry of Agriculture, Fisheries, Food and the Environment, through the General Sub-directorate for Irrigation and Rural Infrastructures n.d. https://eportal.mapa.gob.es/websiar/Inicio.aspx (accessed November 18, 2022).
[59] Voyant C, Lauret P, Notton G, Duchaud J.L, Fouilloy A, David M, et al. A Monte Carlo based solar radiation forecastability estimation. J Renew Sustain Energy 2021;13:026501. https://doi.org/10.1063/5.0042710.
[60] Cadenas E, Campos-Amezcua R, Rivera W, Espinosa-Medina M.A, Méndez-Gordillo A.R, Rangel E, et al. Wind speed variability study based on the Hurst coefficient and fractal dimensional analysis. Energy Sci Eng 2019;7:361–78. https://doi.org/10.1002/ese3.277.
[61] Barbulescu A, Serban C, Maftei C. Statistical analysis and evaluation of Hurst coefficient for annual and monthly precipitation time series. WSEAS Trans Math 2010;10–9:791–800. https://doi.org/10.5555/2230756.2230763.
[62] Shesh S.R, Koirala R, Gentry R, Perfect E, Mulholland P.J, Schwartz J.S. Hurst analysis of hydrologic and water quality time series. J Hydrol Eng 2011;16–9: 717–24. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000357.
[63] Mielniczuk J, Wojdyłło P. Estimation of Hurst exponent revisited. Comput Stat Data Anal 2007;51–9:4510–25. https://doi.org/10.1016/j.csda.2006.07.033.
[64] Bin Huang G, Zhu Q.Y, Siew C.K. Extreme learning machine: theory and applications. Neurocomputing 2006;70:489–501. https://doi.org/10.1016/J.NEUCOM.2005.12.126.
[65] Mansoury I, el Bourakadi D, Yahyaouy A, Boumhidi J. Hourly solar power forecasting using optimized extreme learning machine. Lecture Notes Networks Syst 2022:455. LNNS:629–37. https://doi.org/10.1007/978-3-031-02447-4_65/COVER.
[66] Xiang X, Sun Y, Deng X. Short time solar power forecasting using persistence extreme learning machine approach. E3S Web Conf 2021;294:1002. https://doi.org/10.1051/E3SCONF/202129401002.
[67] Alzahrani A. Short-term solar irradiance prediction based on adaptive extreme learning machine and weather data. Sensors 2022;22:8218. https://doi.org/10.3390/S22218218.
[68] Zhang J, Xu Y, Xue J, Xiao W. Real-time prediction of solar radiation based on online sequential extreme learning machine. In: Proceedings of the 13th IEEE Conference on Industrial Electronics and Applications, ICIEA 2018; 2018. p. 53–7. https://doi.org/10.1109/ICIEA.2018.8397688.