Energy and AI 15 (2024) 100314

Contents lists available at ScienceDirect

Energy and AI

journal homepage: www.sciencedirect.com/journal/energy-and-ai

Development and assessment of artificial neural network models for direct normal solar irradiance forecasting using operational numerical weather prediction data

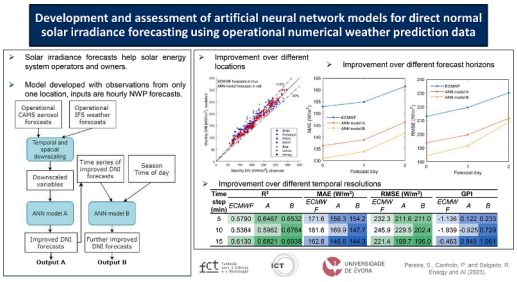
Sara Pereira [a,*], Paulo Canhoto [a,b], Rui Salgado [a,c]

[a] Institute of Earth Sciences, University of Évora, Rua Romão Ramalho 59, 7000-671 Évora, Portugal
[b] Department of Mechatronics Engineering, University of Évora, Rua Romão Ramalho 59, 7000-671 Évora, Portugal
[c] Physics Department, University of Évora, Rua Romão Ramalho 59, 7000-671 Évora, Portugal

HIGHLIGHTS

- Accurate short-term DNI forecasts benefit solar energy systems operation.
- ANNs model non-linear relations between solar irradiance and atmospheric variables.
- ANN models designed based on operational NWP and aerosol forecast data.
- Improved DNI forecasts for different locations, temporal resolutions and horizons.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Solar radiation
Solar energy
Numerical weather prediction
Artificial neural network
Operational forecasting

ABSTRACT

Accurate operational solar irradiance forecasts are crucial for better decision making by solar energy system operators due to the variability of resource and energy demand. Although numerical weather prediction (NWP) models can forecast solar radiation variables, they often have significant errors, particularly in the direct normal irradiance (DNI), which is especially affected by the type and concentration of aerosols and clouds. This paper presents a method based on artificial neural networks (ANN) for generating operational DNI forecasts using weather and aerosol forecasts from the European Center for Medium-range Weather Forecasts (ECMWF) and the Copernicus Atmospheric Monitoring Service (CAMS), respectively. Two ANN models were designed: one uses as input the predicted weather and aerosol variables for a given instant, while the other uses a period of the improved DNI forecasts before the forecasted instant. The models were developed using observations for the location of Évora, Portugal, resulting in 10 min DNI forecasts that for day 1 of forecast horizon showed an improvement over the downscaled original forecasts regarding R^2, MAE and RMSE of 0.0646, 21.1 W/m^2 and 27.9 W/m^2, respectively. The model was also evaluated for different timesteps and locations in southern Portugal, providing good agreement with experimental data.

* Corresponding author.
E-mail address: spereira@uevora.pt (S. Pereira).

https://doi.org/10.1016/j.egyai.2023.100314

Available online 10 November 2023

Nomenclature

AOD469	Aerosol optical depth at 469 nm (–)
AOD550	Aerosol optical depth at 550 nm (–)
AOD670	Aerosol optical depth at 670 nm (–)
AOD865	Aerosol optical depth at 865 nm (–)
AOD1240	Aerosol optical depth at 1240 nm (–)
avgCF	Vertical average cloud fraction (–)
DHI	Diffuse horizontal irradiance (W/m^2)
DNI	Direct normal irradiance (W/m^2)
DUAOD	Dust aerosol optical depth at 550 nm (–)
GHI	Global horizontal irradiance (W/m^2)
maxCF	Vertical maximum cloud fraction (–)
OMAOD	Organic matter aerosol optical depth at 550 nm (–)
SSAOD	Sea salt aerosol optical depth at 550 nm (–)
T	Air temperature (°C)
WD	Wind direction (°)
WS	Wind speed (m/s)
Zen	Solar zenith angle (°)

Acronyms

ANN	Artificial Neural Network
BSRN	Baseline Surface Radiation Network
CAMS	Copernicus Atmosphere Monitoring Service
ECMWF	European Center for Medium-range Weather Forecasts
FS	Forecast Skill
GFS	Global Forecasting System
NWP	Numerical Weather Prediction
MAE	Mean Absolute Error
ML	Machine Learning
MSE	Mean Squared Error
RF	Random Forest
rMSE	Relative Mean Squared Error
RMSE	Root Mean Squared Error
RRTM	Rapid Radiation Transfer Model
SVM	Support Vector Machines
WRF	Weather Research Forecasting

1. Introduction

Economic growth and an increase in energy demand are usually co-dependent, yet they can be partially decoupled by improving energy efficiency, electrification and wise use of energy. Energy generation, distribution and consumption are becoming more electrified, efficient, interconnected and clean with the increase of the use of renewable energy sources such as solar photovoltaic and wind power and the increasing production and use of electric vehicles [1]. However, renewable resources show a strong spatial and temporal variability that affect the power generation, which makes finding an optimum balance between electric generation and consumption at any moment challenging since direct and reliable large-scale storage systems of electric energy are not available.

The solar photovoltaic energy, for example, is mainly dependent on the solar irradiance, namely direct beam irradiance (usually measured in the normal plane, known as Direct Normal Irradiance, DNI) and diffuse irradiance (usually measured in the horizontal plane, known as Diffuse Horizontal Irradiance, DHI), which means that having accurate forecasts of these variables allows for a more accurate estimation of electric power generation. Forecasting solar irradiance can be done through various models and methods, from the physical based methods such as numerical weather prediction (NWP) models, models based on sky/shadow imagery or satellite imagery to the more data-driven and machine-learning based methods such as regression, artificial neural networks (ANN), support vector machines (SVM) and Kalman filtering, or even a combination of these, which are usually known as hybrid methods.

NWP models provide the evolution of the atmosphere by integrating constitutive and state equations describing physical phenomena in the atmosphere which are subject to boundary and initial conditions. Its main goal is weather forecast with a focus on meteorologic variables such as air temperature, humidity, wind or precipitation. An accurate prediction of the partition between the two components of solar global radiation is not so relevant in that case, i.e., DNI and DHI, with global horizontal irradiance (GHI) being used for the closure of the energy balance at the Earth surface. The Integrated Forecasting System (IFS) developed at the European Center for Medium-range Weather Forecast (ECMWF) is the most widely used global NWP model in Europe being its performance attested by various studies such as in [2], where 24-hour forecasts of global solar irradiation from IFS and the American Global Forecasting System (GFS) were compared with observations made at four stations in Morocco. The authors show that the IFS/ECMWF model performs better than the GFS model for all-sky conditions based on the mean bias error and correlation coefficient. In [3], hourly GHI forecasts made by the IFS/ECMWF global model and GFS-driven Weather Research Forecasting (WRF) mesoscale model (run with different configurations by various forecast providers) were compared with observations made in the US, Canada and Europe showing that the model from the ECMWF performs significantly better for all locations and different climatic conditions. Perdigão et al. [4] analyzed one year of hourly and daily direct normal irradiation forecasts of the ECMWF against observations made at Évora, Portugal, for different temporal forecast horizons (0 to 3 days ahead) showing that the model reproduces hourly and daily experimental values with a RMSE of 210.6 W/m^2 and 68.5 W/m^2, respectively, for the first day ahead and that the performance of the model tends to decrease with higher forecast horizon.

The research on improving accuracy of solar radiation variables in NWP models' output has been recently brought to light due to the need for development of solar energy systems [5,6]. DNI is especially difficult to forecast due to its strong dependency on the presence of clouds and aerosol type and concentration in the atmosphere. In the case of aerosols, NWP models such as the IFS/ECMWF use monthly-mean aerosol climatologies instead of more detailed aerosol forecasts to reduce computation time. In the comprehensive review made by Yang et al. [7], the impact of aerosols in the solar resource is discussed. The importance of including aerosols in the NWP forecasts is mentioned where the global models Copernicus Atmospheric Monitoring Service (CAMS) and Goddard Earth Observing System Version 5 (GEOS-5) are shown to be of particular interest. Breitkreuz et al. [8] used libRadtran [9] and achieved a decrease in the relative mean squared error (rMSE) of 4.3 % in the case of hourly GHI from IFS/ECMWF forecasts using aerosol predictions and experimental data. Other studies on aerosol modeling and prediction and its use with NWP models can be found in [10], here the WRF model coupled with Chemistry (WRF-Chem) was used to model aerosol and radiation data which were compared against observations showing 2 to 5 times higher shortwave radiative forcing during a dust storm relative to values during non-dust days, and in [11], where, using the RRTM_SW (Rapid Radiation Transfer Model for Shortwave radiation) widely used in atmospheric models, the propagation of small uncertainties on the aerosol microphysical parameterization on the simulated direct radiative effects is demonstrated.

The knowledge of all the complex phenomena that occur in the atmosphere of the Earth, including the interaction between solar radiation and the atmosphere, and between these and the surface, is still challenging and so is its representation, even in the most complex and detailed NWP models, which are also constrained by data availability

[12] and computational limitations. A trade-off between increasing complexity of deterministic physical models, with the consequent higher computational effort, and the development of additional tools that, based on actual physical models output and experimental data, allow for better solar radiation forecasts without such effort, must be considered. Thus, several techniques have been developed and evaluated to further improve solar radiation estimations made by deterministic NWP models. These are divided into classic statistical methods and Machine Learning (ML) methods. Classic statistical methods can be as simple as interpolation to improve temporal and/or spatial resolution or more complex methods based on stepwise linear regression to select the variables that best represent the errors, so they can be incorporated in a multi-variate regression model that gives an estimate of the forecast error as a linear function of the variables that have been selected in the process [13]. ML methods have been reported extensively in the literature for solar forecasts [14], such as k-nearest neighbor, SVMs, random forest (RF) or ANNs, and tend to focus on the prediction of solar global irradiation. [15] Developed a novel deep learning-based auto-selective tool that allows for the determination of the best GHI forecasting model from four different ML approaches with 81% accuracy. These models can capture the relation between systematic errors of the outputs of NWP models and relevant variables by comparing historical databases of forecasts and observations and thus provide improved values in a fraction of the time it would take with more detailed physical model approaches [16]. In the work by Alfadda et al. [17], hourly GHI, DNI and DHI forecasts made by various machine learning models and using as inputs solar irradiance and aerosol observations and wind and aerosol forecasts in Saudi Arabia were evaluated resulting in a better performance of the ANN model. Pang et al. [18] compared the use of feedforward artificial neural networks with recurrent neural networks for global solar irradiance prediction using onsite measurements showing that the later moderately improved the prediction performance but with additional computational cost. For a short-range forecast of solar irradiance (15 to 180 min), McCandless et al. [19] developed a regime-dependent artificial neural network forecasting model that showed improvements over a global ANN and the persistence. Fonseca et al. [20] obtained day-ahead solar irradiation forecasts through SVMs using NWP data from the mesoscale model of the Japan Meteorological Agency as input, which resulted in a reduction of the root mean squared error (RMSE) of 16 % in a regional scale. Lima et al. [21] used artificial neural networks for the post-processing of NWP solar irradiation forecasts from the WRF model for the Brazilian Northeastern region and obtained a reduction of the model bias, MSE and RMSE. The same approach can be used to improve solar resource assessment since, once machine learning models were trained, they can generate more accurate spatial distribution of solar irradiation in a given location or region based on historical data of NWP forecasts. For example, Pereira et al. [22] developed and optimized an ANN model in order to create a method for solar resource assessment from the meso-scale Meso-NH NWP model outputs for a typical meteorological year. In this case, the method was extended and validated for the South of Portugal, showing important improvements regarding the direct normal and global horizontal irradiation mapping with a horizontal resolution of 1.25 km. Comprehensive reviews on machine learning methods for solar irradiance forecasting can be found in [23–29]. Although various ML models have been developed for solar irradiation forecasting, most of them focus only on GHI or cannot be readily used to obtain useful power output forecasts of solar energy systems since the temporal resolution and forecast time horizon are not synchronized with the real-time market forecasting requirements [30, 31].

This work proposes a method for operational DNI forecasts which includes: (i) the spatial and temporal downscaling of the IFS/ECMWF and aerosol forecasts for a specific location and time step; (ii) an ANN model for generating improved DNI forecasts using only the predicted weather and aerosol data for a given instant of a specific time step; (iii) a second ANN model that takes as inputs a period of improved DNI forecasts immediately before the time step being forecasted as well as the season and time of day of that time step. The model was developed considering the location of Évora, Portugal (38.567811, -7.911459), a temporal resolution of 10 min and a forecast horizon of 25 to 48 h (day 1 of forecast). After development, evaluation against observations show that this method can be generalized and applied to different locations, temporal resolutions and forecast time horizons up to 72 h.

Table 1
Forecast variables retrieved from the IFS/ECMWF database.

Variable	Symbol	Units	Range / Comments
Longitude	*Lon*	° East	0° to 360°
Latitude	*Lat*	° North	-90° to 90°
Time step	*Step*	h	1 to 240 h
Date	*Date*	Days	Days since 1900–01–01 00:00:00
Low cloud cover	*LCC*	0–1	–
Medium cloud cover	*MCC*	0–1	–
High cloud cover	*HCC*	0–1	–
Total cloud cover	*TCC*	0–1	–
10 m *U* wind component	*u10*	m/s	–
10 m *V* wind component	*v10*	m/s	–
2 m temperature	*T*	K	Air temperature at 2 m
Solar zenith angle	*Zen*	°	–
Surface solar radiation downwards	*SSRD*	J/m^2	Irradiation since forecast issuance
Direct solar radiation	*DSRP*	J/m^2	Irradiation since forecast issuance

The paper is organized as follows: in Section 2 the different data sets used in this work are presented and the treatment of this data is explained including the methods for spatial and temporal downscaling of the forecast data; in Section 3 an evaluation of the downscaled DNI forecasts made by the ECMWF is performed; Section 4 presents the development of the ANN models used for the improvement of the DNI forecasts obtained from the ECMWF; in Section 5 is performed an evaluation of the models for different time steps while in Section 6 this is done for different sites located in the region surrounding the one used for model development; Section 7 presents an analysis of the ANN models in an operational forecast setting and finally, in Section 8, the conclusions of this work are presented.

2. Forecast and experimental data

The method was developed using forecast data from the IFS/ECMWF and CAMS models and observed solar radiation data from a network of measuring stations scattered in the South of Portugal. These models are operational and issued everyday which means that the developed model can also be used operationally. The period and area of the data used was from December 2016 to the end of May 2021 and between latitudes 37.0° and 39.3° and longitudes -7.4° and -9.2°, respectively. A description of the data, quality check, filtering and pre-processing is presented in the following subsections.

2.1. Weather forecast data

In this work, various weather forecast variables from the IFS/ECMWF model [32] were used. This NWP model includes the radiative scheme ecRad [33] which solves the 1D radiative transfer equation both for small and long wavelength ranges, considering the vertical profiles of temperature, moisture, cloud droplet and ice cloud effective radius, average monthly climatologies of aerosols, carbon dioxide, ozone and trace gases, and also ground surface temperature, albedo and emissivity for different spectral bands and solar zenith angle. The code is based on the RRTM model (Rapid Radiative Transfer Model) using the McICA method (Monte Carlo Independent Column Approximation), which allows for the parameterization of the interactions between radiation and clouds. The ECMWF versions RRTMLW and RRTMSW describe the

Table 2
Forecast variables retrieved from the CAMS database.

Variable	Symbol	Units	Range / Comments
Longitude	*Lon*	° East	0° to 360°
Latitude	*Lat*	° North	-90° to 90°
Time step	*Step*	h	1 to 120
Date	*Date*	Days	Days since 1900–01–01 00:00:00
Total aerosol optical depth at 469 nm	*AOD469*	–	–
Total aerosol optical depth at 550 nm	*AOD550*	–	–
Total aerosol optical depth at 670 nm	*AOD670*	–	–
Total aerosol optical depth at 865 nm	*AOD865*	–	–
Total aerosol optical depth at 1240 nm	*AOD1240*	–	–
Sea salt aerosol optical depth at 550 nm	*SSAOD*	–	–
Organic matter aerosol optical depth at 550 nm	*OMAOD*	–	–
Dust aerosol optical depth at 550 nm	*DUAOD*	–	–

radiative transfer in the long wavelength range, with 16 spectral bands, and in the short wavelength range, with 14 spectral bands, respectively [33].

The operational deterministic model of the ECMWF is run every day at 00UTC and 12UTC, providing hourly forecast values up to 90 h ahead (and then 3 and 6-hourly values up to 144 and 240 h ahead, respectively) at discrete points of a global grid with horizontal spatial resolution of 0.125° × 0.125° The variables retrieved from the ECMWF database are presented in Table 1.

Data were retrieved for the maximum grid point density and with a temporal forecast horizon up to 72 h. The retrieved GRIB files were converted to netCDF format and processed with a MATLAB® routine to obtain hourly mean values of global horizontal and direct normal irradiance in W/m^2, converting air temperature (*T*) to °C and computing wind speed (*WS*) in m/s and direction (*WD*) in degrees from North using the 10-meter *U* and *V* wind speed components.

2.2. Aerosol forecast data

CAMS developed a global atmospheric composition forecast based on the IFS model but with additional modules enabled for aerosols, reactive gases and greenhouse gases taking into consideration phenomena such as the emissions and transport of trace gases and aerosols, uptake and release by vegetation and land and sea surface, removal by dry deposition at the surface and scavenging in precipitation, chemical conversion and aerosol microphysics. It generates atmospheric composition variables, including aerosol optical depth at different wavelengths, in a three-dimensional grid with approximately 40 km of horizontal spatial resolution and 1 h time step [34]. Hourly mean total aerosol optical depth forecasts at surface level for various discrete wavelengths are computed every day at 00UTC and 12UTC with a temporal forecast horizon of 5 days. The variables retrieved from the CAMS database are shown in Table 2, for the same area and period as the variables retrieved from the IFS/ECMWF.

2.3. Spatial and temporal downscaling of forecast data

To obtain forecast values for a specific location and with different time steps, spatial and temporal downscaling techniques were employed to all simulated variables. The spatial downscaling is conducted using bilinear interpolation of the values of the four grid points surrounding the desired location. This allows for the development and validation of the

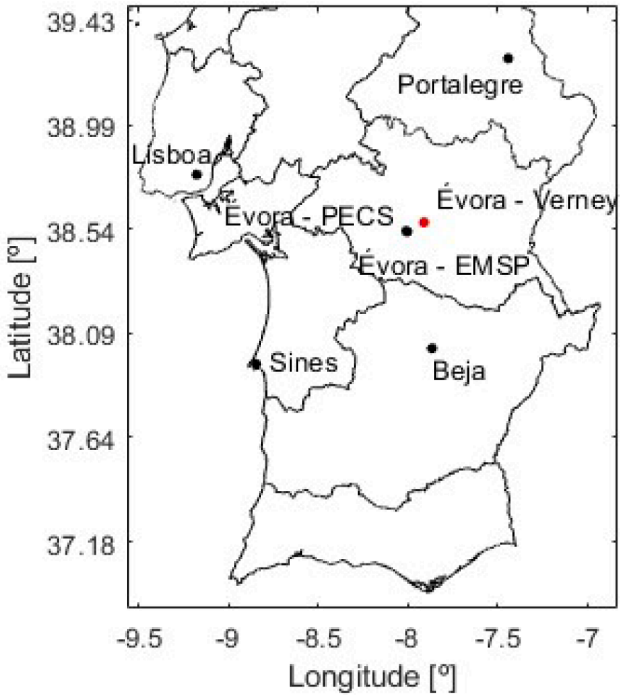


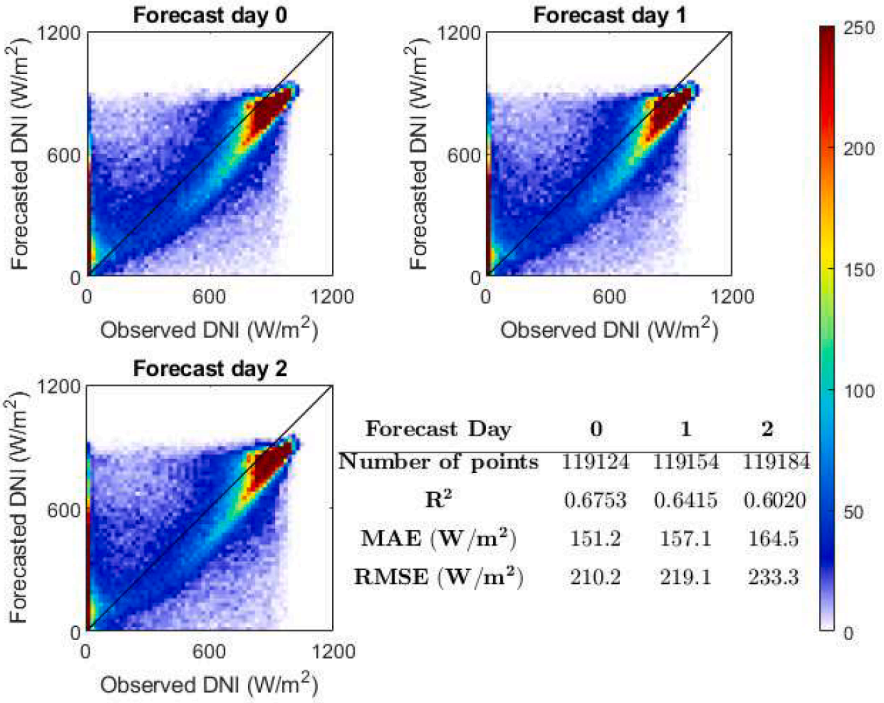
Fig. 1. Location of solar radiation measuring stations from the DNI-ALENTEJO network [22].

presented method against observations made at various solar radiation measuring stations which are not exactly located at a grid point, but also, in an operational setting, the inclusion of this technique allows for the forecasting of solar irradiance at any point of the domain. The temporal downscaling is computed using piecewise cubic hermite interpolation of the hourly mean irradiance. This method might not allow for the conservation of energy in each hour (forecast values) however, the goal is simply to obtain data in shorter timesteps which can be used as input for the more complex machine learning models that will perform the improvement of DNI forecasts. This is a compromise between developing a more elaborated physical downscaling method that preserves the hourly energy predictions (but preserves the error associated with those forecasts) and feeding the machine learning models directly with the original hourly irradiation values, and then having to develop a model for each desired timestep, with the consequent loss of model generalization. In this way, the deviation introduced by this downscaling method in the input values (which, at different instants, can either increase or decrease the error of the energy predictions, taking the ground-based measurements as reference) will also be assimilated by the machine learning model.

2.4. Solar radiation measurements and data quality check

The experimental data used for the model development were 1 min DNI and GHI ground-based observations made at Évora–Verney (38.567811°, -7.911459°) with a Kipp & Zonen CHP1 pyrheliometer in the case of DNI and a Kipp & Zonen CMP11 Pyranometer in the case of GHI. Experimental data from this location have been widely validated and used in other works in this field, e.g. [22,35–37].

For the model validation, 1 min DNI and GHI ground-based observations obtained from the network of radiometric stations of the DNI-ALENTEJO project [38] were also used. This is a solar radiation measurement network in the south of Portugal which comprised 13 stations scattered in the region, each station being typically equipped with a Kipp & Zonen Solys2 sun tracker, one CHP1 pyrheliometer and two CMP11 pyranometers. All instruments are periodically calibrated in accordance with ISO 9059:1990 and ISO 9847:1992 and the



Forecast Day	0	1	2
Number of points	119124	119154	119184
R^2	0.6753	0.6415	0.6020
MAE (W/m^2)	151.2	157.1	164.5
RMSE (W/m^2)	210.2	219.1	233.3

Fig. 2. Comparison between downscaled IFS/ECMWF forecasts and DNI observations (10 min) at Évora – Verney for forecast days 0, 1 and 2 (the colormap represents the number of data points in each bin. Bin size: 20 × 20 W/m^2).

observations are corrected regarding the zero offset of sensors, filtered according to the Baseline Surface Radiation Network (BSRN) quality control procedure [39] and gaps filled according to the method developed in [38].

The experimental data used for the model validation were retrieved from the following stations of the network: Évora–PECS (38.5306°, -8.0112°), Évora–EMSP (38.5289°, -8.0053°), Portalegre (39.2692°, -7.4428°), Beja (38.0249°, -7.8672°), Lisboa (38.7734°, -9.1779°) and Sines (37.9576°, -8.8473°). The location of these stations is shown in Fig. 1 including the Évora–Verney station which data were used for the model development.

Original solar irradiance records from the radiometric stations are 1 min average, maximum, minimum and standard deviation from which the mean irradiance values for the different temporal resolutions used in this work were calculated (10 min for the development stage). The same period defined for the forecast data was used.

3. Analysis of NWP direct normal irradiance forecasts

To understand the accuracy of the original solar irradiance predictions from the IFS/ECMWF model, the forecast data of DNI and other atmospheric variables issued at 00UTC from 1st of December 2016 to the 31st of May 2021 were analyzed and compared against observations made at Évora–Verney station for each of the first three days of forecast time horizon. All forecast data was spatially downscaled to the location of the radiometric station and 10 min mean values were determined as described in Section 2.3. Experimental data were also averaged for the same 10 min time step as described in Section 2.4.

3.1. Comparison of NWP direct normal irradiance forecasts with experimental data

Firstly, a direct comparison between the 10 min DNI forecasts after spatial and temporal downscaling and observations at Évora–Verney station was conducted for each of the 3 days of forecast horizon as presented in Fig. 2.

These results show a better performance (lower mean absolute error, MAE, and root mean squared error, RMSE) of the model for day 0 of forecast which decreases for higher forecast time horizons which was also verified by other works in the literature [4]. The correction of these errors can be done with many different tools, from a simple bias correction to more complex machine learning models, as mentioned in Section 1. Simple artificial neural networks offer a middle ground where the computational effort is not too high but the influence of the different variables used as inputs is taken into consideration. Day 1 of forecast (second day of forecast time horizon, or day-ahead) was used for the model development since, even though forecast issuance corresponds to the 00 UTC, data can take until 6:55 UTC to be available, which may turn the forecast unhelpful for that day depending on the location of interest.

3.2. Correlations between forecast variables and DNI observations

The various meteorological variables are to some extent all correlated with each other and, in a system so complex as the atmosphere, a way to identify the degree of correlation between two variables is to compute their linear correlation coefficients. Fig. 3 was generated with this purpose, comparing forecast values of each variable after spatial and temporal downscaling (10 min) with observed DNI at Évora–Verney station for the three forecast days.

The Pearson's linear correlation coefficients (the values shown in each graph of Fig. 3) for each forecast variable are similar but tend to show a decrease in correlation across the three days of forecast time horizon. As expected, the highest absolute values of linear correlation occur with the forecasted DNI followed by GHI, cloud cover variables and solar zenith angle. While the aerosol variables show lower values of linear correlation, this does not mean that they are necessarily more independent since they might have a nonlinear relationship.

It is difficult to find any physical world phenomenon which follows linearity straightforwardly. Thus, a non-linear model that can approximate the non-linear phenomenon is needed. Representing these kinds of relationships using classical methods is known to be difficult. In that

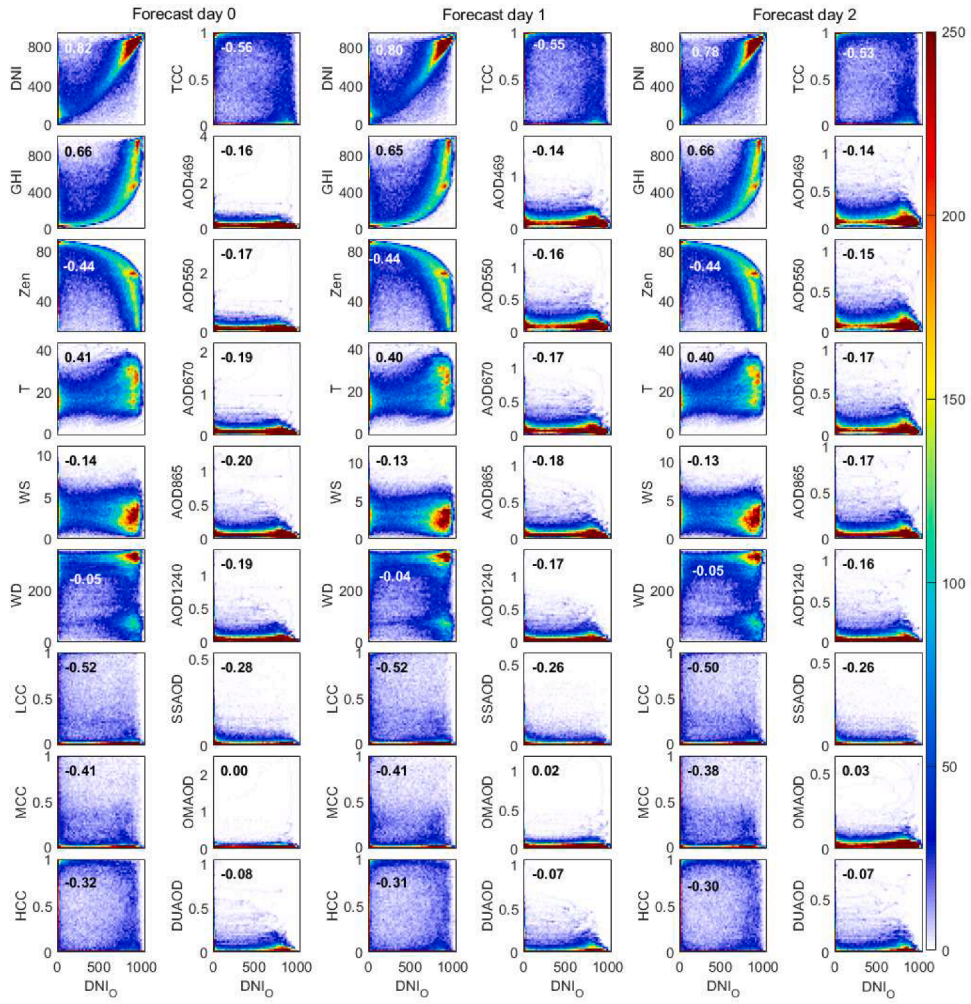


Fig. 3. Comparison and Pearson's linear correlation coefficients of the forecast values of each meteorological variable (y axis) with the 10 min observed DNI at Évora–Verney station (x axis) for each of the forecast days. The colormap represents the number of points (total number of data points for each graph: 236,592); Bins grid (50 × 50 bins).

sense, machine learning models like artificial neural networks with non-linear activation functions will allow the model to create complex mappings between the inputs and outputs of the network. The non-linear layers enable ANNs to learn making conditional decisions for controlling the computational flow which makes them better suited to provide accurate data fitting and forecast.

4. Artificial neural network model development

Artificial neural networks (ANN) are constituted of connected artificial neurons forming a network. Since the relevance of each input given to the neuron is not equal, different weights are assigned to each of the inputs, then a linear net function is used to aggregate a bias and the weighted inputs after which a transfer function is applied to obtain the output of the neuron that will then be passed on to the next neuron.

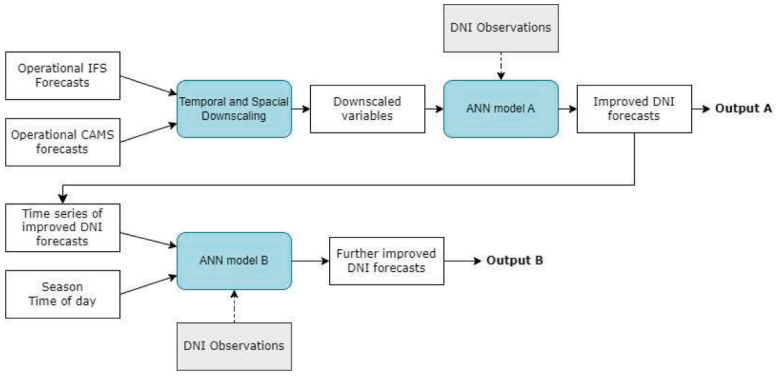


Fig. 4. Flowchart of the developed model. DNI observations are only used for the development and evaluation of the ANN models they are not required as inputs in an operational setting.

Table 3
ANN model A training and validation results (five best configurations in descending order).

Training function	Number of neurons	Number of inputs	R^2	MAE (W/m^2)	RMSE (W/m^2)	FS (%)	GPI
trainbr	25	18	0.7535	122.9	180.7	22.0	1.159
trainbr	24	18	0.7528	123.0	181.0	21.9	1.144
trainbr	23	18	0.7525	122.9	181.1	22.0	1.141
trainbr	25	17	0.7511	123.6	181.6	21.5	1.099
trainbr	24	17	0.7508	123.5	181.7	21.6	1.097

The ANNs developed in this work are multi-layer feed-forward networks with back propagation learning, which are some of the most established ANN architectures due to their ability to perform arbitrary non-linear mappings. These are usually composed of an input and an output layer and one or more hidden layers of neurons. When training, for a given input dataset, the information flows forward through the network until it reaches the output layer and errors are calculated using the given desired output dataset (targets). These errors are then propagated backwards through the network and the weights of each neuron are adjusted so that the next iteration results in outputs with smaller error. This allows solving complex problems that can be stochastic, poorly defined, non-linear, non-analytic and/or non-stationary with low or no intervention in the program itself.

After the spatial and temporal downscaling of the forecast data shown in Section 2.3, two ANN models coupled in series are applied in this work according to Fig. 4 for the generation of improved DNI forecasts from NWP data. Observations of DNI made at Évora - Verney described in Section 2.4 were used for the development of the two ANNs: (i) development of an ANN using only the predicted weather and aerosol variables for a given time step in the temporal horizon of forecasts (Section 4.1); (ii) development of an ANN using the predicted DNI data over a period of time before the forecasted instant (Section 4.2).

For the evaluation and comparison of the different models, the coefficient of determination (R^2), mean absolute error (MAE) and root mean squared error (RMSE) were used as metrics along with the forecast skill (FS), which represents the improvement in terms of MAE over the original ECMWF forecasts, and a global performance index (GPI) based on the three statistical indicators, for model configurations comparison. The coefficient of determination is intuitively informative as it provides a measure of how well observed outcomes are replicated by the model, based on the proportion of total variation of outcomes explained by the model. According to the literature, the MAE and RMSE metrics are suitable [40] and the most commonly used indicators to assess the performance of machine learning regression algorithms [41]. Each error contributes to MAE in proportion to the absolute value of the error while the RMSE involves squaring the differences, so that a few major differences will increase the RMSE to a greater degree than the MAE.

The used definition of GPI results in a statistical tool that combines various metrics allowing for the performance comparison of different models and has been widely used in many works in several fields of study, for example in [42–44]. To compute the GPI, the n metrics need to be normalized into values ranging from 0 to 1 and then Eq. (1) is used, where the GPI value for the i^{th} model configuration is determined using the median of the normalized values, $\tilde{y}_j$, of the indicator j, the normalized value of indicator j for model configuration i, y_{ij}, and a factor α_j that has a value of 1 for all indicators except the coefficient of determination for which it is -1.

$$GPI_i = \sum_{j=1}^{n} \alpha_j \left(\tilde{y}_j - y_{ij}\right) \quad (1)$$

In this work, the metrics used for the computation of GPI are the previously mentioned R^2, MAE and RMSE. Since all the metrics used in this definition take the ground-based measurements as reference, a higher value of GPI represents a better performance of the respective model configuration.

4.1. ANN model with weather and aerosol forecasts as inputs (ANN model A)

The forecast data included as inputs in the development of this ANN model are the various weather and aerosol variables from the IFS/ECMWF and CAMS global NWP models (see Tables 1 and 2) for forecast day 1 (25 to 48 h ahead), after temporal and spatial downscaling as described in Section 2.3. The 10 min ground-based observations of DNI at Évora–Verney station as described in Section 2.4, are used as targets.

To obtain an ANN configuration that generates improved DNI forecasts while having a good generalization capability, datasets were divided in three subsets: the first subset for training and validation of various internal configurations of the tested ANNs (78,854 data points from 1st December 2016 to 30th November 2019), a second subset for testing and selection of the final ANN configuration (23,654 data points from 1st December 2019 to 30th November 2020) and finally a third subset of data for a blind test so the performance and generalization capability of the selected ANN configuration can be demonstrated (12,470 data points from 1st December 2020 to 31st May 2021). For training and validation, the input data is divided randomly using 80 % of the available data points in the first subset for training and the other 20% for validation.

An ANN is defined by numerous parameters and specifications, some of which are pre-established such as using a feedforward ANN with one hidden layer and a linear layer output (*fitnet*) with an initialization function that initializes the weights and biases of the layers according to the Nguyen–Widrow initialization algorithm (*initnw*), the hyperbolic tangent sigmoid transfer function (*tansig*) and the mean squared error as performance function (*mse*). The input and output data are treated by removing rows with constant values (*removeconstantrows*) and scaling the mean of each row to 0 and deviations to 1 (*mapstd*).

Other parameters and configurations were specifically evaluated in this work, namely the training function, the number of neurons and the input variables. The training functions assessed were the Levenberg-Marquardt backpropagation (*trainlm*) and the Bayesian regularization backpropagation (*trainbr*) and, for each of these, the number of neurons was varied from 1 to 25 and the input forecasted variables were added one by one according to the sequence DNI, GHI, TCC, Zen, T, SS, AOD1240, WS, LCC, MCC, HCC, AOD865, AOD670, AOD550, AOD469, DU, WD, OM, for each combination of training function and number of neurons.

Additionally, for better results, ten randomly initialized ANNs were trained and validated for each configuration combo mentioned above, being the average output of those ten ANNs considered the result of the corresponding ANN configuration, as already tested in other works in this field [22].

The values of R^2, MAE, RMSE were computed and the performance of the different ANN configurations is compared using the FS and the GPI. The five ANN configurations with highest GPI at the training and validation stage (among the 900 cases evaluated with GPI values ranging between 1.159 and -1.840) are presented in Table 3.

The configurations using the training function *trainbr* with all input variables and high number of neurons show more accurate DNI forecasts when considering observations as reference. For the best performing configuration, there is an improvement over the original ECMWF forecasts revealed by a forecast skill of 22 % and an increase of 0.1118 in the

Table 4
ANN model A testing results (five best configurations in descending order).

Training function	Number of neurons	Number of inputs	R^2	MAE (W/m^2)	RMSE (W/m^2)	FS (%)	GPI
trainbr	7	13	0.7127	136.7	197.8	10.2	0.876
trainlm	7	13	0.7132	137.1	197.7	9.9	0.867
trainlm	9	13	0.7121	136.8	198.0	10.1	0.844
trainlm	5	13	0.7112	136.5	198.3	10.3	0.832
trainbr	5	13	0.7115	136.9	198.2	10.0	0.818

Table 5
ANN model B training and validation results using a period of 2 h of predicted DNI from ANN model A as input (five best configurations in descending order).

Training function	Number of neurons	R^2	MAE (W/m^2)	RMSE (W/m^2)	FS (%)	GPI
trainbr	25	0.7332	126.1	188.0	20.0	1.126
trainbr	24	0.7337	126.3	188.1	19.8	1.019
trainbr	23	0.7335	126.4	188.2	19.8	0.977
trainbr	22	0.7321	126.4	188.3	19.8	0.917
trainbr	21	0.7319	126.6	188.4	19.6	0.852

Table 6
ANN model B training and validation results using a period of 24 h of predicted DNI from ANN model A as input (five best configurations in descending order).

Training function	Number of neurons	R^2	MAE (W/m^2)	RMSE (W/m^2)	FS (%)	GPI
trainbr	25	0.7518	123.0	181.2	21.9	1.990
trainbr	24	0.7496	123.5	182.0	21.6	1.788
trainbr	23	0.7485	123.9	182.5	21.4	1.672
trainbr	22	0.7479	123.8	182. 7	21.4	1.642
trainbr	21	0.7456	124.4	183.5	21.1	1.418

Table 7
ANN model B test results using a period of 2 h of predicted DNI from ANN model A as input (five best configurations in descending order).

Training function	Number of neurons	R^2	MAE (W/m^2)	RMSE (W/m^2)	FS (%)	GPI
trainlm	8	0.7197	133.3	195.3	12.4	0.709
trainlm	7	0.7194	133.3	195.4	12.4	0.630
trainbr	7	0.7191	133.2	195.5	12.5	0.572
trainlm	10	0.7191	133.3	195.5	12.4	0.563
trainbr	5	0.7190	133.3	195.5	12.4	0.512

Table 8
ANN model B test results using a period of 24 h of predicted DNI from ANN model A as input (five best configurations in descending order).

Training function	Number of neurons	R^2	MAE (W/m^2)	RMSE (W/m^2)	FS (%)	GPI
trainbr	3	0.7184	133.1	195.7	12.4	0.451
trainlm	12	0.7190	133.4	195.5	12.2	0.439
trainlm	11	0.7183	133.2	195.7	12.4	0.429
trainlm	5	0.7191	133.5	195.4	12.2	0.429
trainbr	4	0.7180	133.3	195.8	12.3	0.299

R^2 and a decrease of 34.2 W/m^2 and 38.4 W/m^2 in the MAE and RMSE, respectively.

However, regarding the final purpose of the developed model, that is, an algorithm that can generalize well and improve newly generated forecasts, the 900 combos were run for the test data set to select the best configuration. The five ANN configurations with highest GPI are presented in Table 4.

Thus, the ten ANNs using the *trainbr* training function, 7 neurons and 13 inputs (DNI, GHI, TCC, Zen, T, SS, AOD1240, WS, LCC, MCC, HCC, AOD865 and AOD670) were selected as the base configuration for this ANN model (ANN model A). This configuration showed an improvement over the original ECMWF forecasts (values in Fig. 2) with a forecast skill of 10.2 % and an increase of 0.0712 in the R^2 and a decrease of 20.4 W/m^2 and 21.3 W/m^2 in the MAE and RMSE, respectively. The performance and good generalization of the selected ANN configuration was verified by using a new set of data (blind test), which resulted in values of R^2, MAE and RMSE of 0.6874, 147.2 W/m^2 and 207.6 W/m^2, respectively, which is better than all statistical indicators of the original DNI forecasts (Fig. 2).

4.2. ANN model with forecasted DNI time series and date/time as inputs (ANN model B)

A different approach for an ANN configuration that considers the temporal variation of solar irradiance as input was also addressed and evaluated. This configuration uses the same internal parameters as the one described above, being the evaluated parameters the training function (*trainlm, trainbr*) and the number of neurons (from 1 to 25), while the inputs are the season and time of day of forecast and the existing time series of DNI values for a given period before the forecast computation. This time series can be either the original IFS/ECMWF predictions or experimental values but, to further improve the solar irradiance forecasts and to make the model independent of ground-based measurements, the corrected DNI forecast data from the ANN model A was used instead in this case. Regarding the length of the time series, two cases were tested: (i) a period corresponding to the previous 2 h (12 time steps of 10 min data); and (ii) a period of 24 h (144 time steps of 10 min data). The same data sets of experimental data were used for training and validation, testing and blind testing as in the Section 4.1, as well as the procedure of averaging the results of ten randomly initialized ANNs for each configuration. The resulting metrics for the five configurations with highest GPI are shown in Tables 5 and 6 for these two cases.

Similar to the case of ANN model A, the configurations that show better performance tend to have more neurons and the more complex training function *trainbr*. However, the best models that arise from the training process might not be the best at generalization, thus all configurations were also evaluated using the test data set to select the configuration that achieves good performance on new data. Tables 7 and 8 show the five best configurations using the previous 2 and 24 h of predicted DNI from ANN model A as inputs, respectively.

Selecting the best configurations for each case and applying these models to the blind test data set, values of R^2, MAE and RMSE of 0.6934, 142.4 W/m^2 and 205.1 W/m^2 for the model using a period of 2 h and 0.6916, 142.6 W/m^2 and 205.5 W/m^2 for the model using a period of 24 h were obtained, respectively. These metrics result in a forecast skill of 12.4 % for both models and an improvement over the predicted values from the ANN model A with an increase in R^2 of 0.0060 and a decrease of MAE and RMSE of 4.8 W/m^2 and 2.5 W/m^2 for the 2 h model and an increase in R^2 of 0.0042 and a decrease of MAE and RMSE of 4.6 W/m^2 and 2.1 W/m^2 for the 24 h model, respectively.

Although the differences between the two cases are small, the selected configuration of the ANN model B was the one using as input a period of 2 h of DNI predictions due to its slightly better performance and simplicity. The resulting composite ANN model (ANN model B) is thus the combination of an ANN model based on the input of a complete

Table 9
Metrics for the original downscaled ECMWF predictions, ANN model A and ANN model B with different temporal resolution and for the different data sets used in the development of the models (for all data and each statistical indicator, the best performing model is represented in bold for each time step, the best performing time step is underlined for each model, and the best combination of time step and model is marked with *).

Time step (min)	Model	R^2				MAE (W/m^2)				RMSE (W/m^2)			
		Train	*Test*	*Blind test*	*All*	*Train*	*Test*	*Blind test*	*All*	*Train*	*Test*	*Blind test*	*All*
5	ECMWF	0.6290	0.6489	0.6168	0.6341	160.7	155.9	168.7	160.4	225.2	222.0	236.3	225.7
	A	0.7088	0.7017	0.6776	0.7058	135.9	140.9	151.3	138.7	199.4	204.5	213.7	202.1
	B	0.7129	0.7053	0.6802	**0.7096**	131.1	135.2	146.3	**133.6**	198.2	203.2	213.4	**201.0**
10	ECMWF	0.6353	0.6579	0.6273	0.6415	157.5	152.1	165.1	157.1	220.0	216.1	230.5	219.1
	A	0.7176	0.7130	0.6874	0.7148	132.3	136.7	147.2	135.2	193.4	197.8	207.6	196.1
	B	0.7265	0.7197	0.6934	**0.7234**	128.3	133.3	142.4	**130.9**	190.3	195.3	205.1	**193.0**
15	ECMWF	0.6515	0.6739	0.6422	0.6574	153.0	147.3	159.7	152.5	214.5	210.0	224.3	214.6
	A	0.7339	0.7284	0.7053	0.7314	127.8	131.9	141.7	130.2	187.3	191.5	200.4	189.7
	B	0.7411	0.7355	0.7117	**0.7386***	127.4	130.4	138.4	**129.3***	185.1	189.2	197.1	**187.3***
60	ECMWF	0.5527	0.5877	0.5519	0.5624	176.3	169.0	179.3	175.0	239.9	231.8	245.0	238.7
	A	0.6677	0.6705	0.6505	0.6684	153.3	154.1	157.7	**153.9**	203.5	204.2	209.0	**204.3**
	B	0.6946	0.6904	0.6685	**0.6929**	174.4	171.5	164.3	172.7	217.7	217.8	213.3	217.3

set of meteorological and aerosol forecast data for a given instant (ANN model A) and an ANN model based on the input of a time series of DNI for a given period before the forecast computation (ANN model B configuration), provided that the input values of the second model are the predicted DNI output values of the first model.

Using two separate ANN models can result in better forecasts than a single model which considers both the actual and temporal variations of weather variables on DNI forecasts since it allows for better model specialization. This means that each ANN can be specialized for a specific task. The first ANN was developed for improving DNI forecasts from NWP data at a specific time step being designed and optimized for representing the relationship between the different atmospheric variables and DNI, being optimized to extract relevant features and relationships specific to the current conditions. The second ANN was developed to model the temporal tendencies and patterns in DNI which can be complex and non-linear. By using a separate ANN model to capture these tendencies, it allows the model to focus solely on learning the temporal dependencies and trends, which might be distinct from the relationships governing the immediate forecast.

5. Assessment of the developed ANN models using different temporal resolutions

The proposed model was developed using 10 min mean values, but it

Table 10
Results of hourly mean DNI forecasts with different temporal resolutions and from different models and data sets. Color comparison within each metric and data set where darker color means better performance.

Data set	Time step (min)	R^2			MAE (W/m^2)			RMSE (W/m^2)			GPI		
		ECMWF	*A*	*B*	*ECMWF*	*A*	*B*	*ECMWF*	*A*	*B*	*ECMWF*	*A*	*B*
Train	5	0.5678	0.6429	0.6493	173.5	156.4	154.5	234.0	211.8	211.1	-1.177	0.140	0.240
	10	0.5280	0.5915	0.6738	183.2	170.2	147.5	247.5	230.3	202.2	-1.926	-0.880	0.740
	15	0.6029	0.6798	0.6916	164.3	146.2	144.2	222.7	199.3	195.6	-0.515	0.864	1.057
	60	0.5527	0.6677	0.6946	176.3	153.3	174.4	239.9	203.5	217.7	-1.456	0.528	-0.126
Test	5	0.5994	0.6505	0.6560	166.6	155.3	153.0	226.9	210.6	210.2	-1.449	-0.111	0.035
	10	0.6164	0.6629	0.6777	162.1	152.6	147.7	221.3	206.5	202.2	-0.974	0.219	0.652
	15	0.6330	0.6848	0.6952	157.6	145.9	143.9	216.0	199.1	196.1	-0.507	0.870	1.125
	60	0.5877	0.6705	0.6904	169.0	154.1	171.5	231.8	204.2	217.8	-1.787	0.298	-0.528
Blind test	5	0.5883	0.6424	0.6507	170.4	158.2	154.8	232.8	212.7	212.0	-1.272	-0.076	0.103
	10	0.6027	0.6533	0.6695	167.0	155.4	149.1	228.4	209.2	204.2	-0.971	0.163	0.577
	15	0.6166	0.6720	0.6854	163.6	151.1	145.6	224.2	203.3	198.1	-0.674	0.555	0.932
	60	0.5519	0.6505	0.6685	179.3	157.7	164.3	245.0	209.0	213.3	-2.068	0.078	-0.076
All	5	0.5790	0.6467	0.6532	171.6	156.3	154.2	232.3	211.6	211.0	-1.136	0.122	0.233
	10	0.5384	0.5962	0.6764	181.6	169.9	147.7	245.9	229.5	202.4	-1.939	-0.925	0.729
	15	0.6130	0.6821	0.6938	162.8	146.6	144.3	221.4	199.7	196.0	-0.463	0.848	1.061
	60	0.5624	0.6684	0.6929	175.0	153.9	172.7	238.7	204.3	217.3	-1.463	0.473	-0.132

Table 11
Metrics for the original downscaled ECMWF DNI forecasts and ANN models B predictions (10 min) for the different stations located in the south of Portugal.

Station	R^2		MAE (W/m^2)		RMSE (W/m^2)		FS (%)	N
	ECMWF	ANN	ECMWF	ANN	ECMWF	ANN		
Verney	0.6630	0.7237	151.3	131.8	215.1	194.4	12.9	28,101
EMSP	0.6528	0.7222	153.2	130.5	218.1	194.2	14.8	96,323
PECS	0.6348	0.7128	161.7	135.5	226.9	200.4	16.2	113,300
Portalegre	0.6329	0.7083	157.5	134.8	225.9	200.8	14.4	115,997
Beja	0.6393	0.6996	156.8	135.0	219.8	199.7	13.9	97,109
Lisboa	0.5970	0.6659	170.9	147.6	238.3	212.2	13.6	61,942
Sines	0.5459	0.6184	175.8	153.3	244.8	220.9	12.8	53,192
All	0.6269	0.6976	160.4	137.2	226.5	202.4	14.4	478,574

could easily be adapted to different time steps as required by solar system operators. Also, since mean irradiance values are used instead of irradiation data in the case of DNI, inputs with different time steps can also be directly used with the already developed model.

In order to evaluate the performance of this feature, input data with temporal resolutions of 5 and 15 min and the original hourly (60 min) forecasts were fed to the model, generating the results shown in Table 9. It is important to note that, as described in Section 4.2, the ANN model B was developed using as inputs 12 records of 10 min DNI predictions, corresponding to the period of 2 h before the forecast time step, while for different temporal resolutions, the number of inputs used is still 12 records but instead of corresponding to a period of 2 h they are equivalent to periods of 1, 3 and 12 h for 5, 15 and 60 min temporal resolutions, respectively.

Comparing the results for the different temporal resolutions, the model seems to perform better for a time step of 15 min, followed by 10, 5 and finally 60 min. However, these results are also affected by the number of data points available which differs with temporal resolution (more data for shorter temporal resolution). Thus, to take this aspect into consideration, the results of the models for each temporal resolution were converted to hourly means for comparison and the same metrics were computed, as shown in Table 10.

Regarding the temporal downscaling, although the metrics for the ECMWF forecasts can, for some instances, deteriorate with downscaling (metrics for hourly values better than for the remaining temporal resolutions) due to a smoothing effect that does not capture the rapid and nonlinear variations of real data, these metrics are improved when using the developed ANN models to the downscaled data when compared with the original hourly ECMWF forecasts. This means that the developed models can not only counteract the possible deviations induced by the temporal downscaling but also further improve the DNI forecasts. As previously discussed, the ANN model B tends to perform better than the ANN model A and ECMWF models except for hourly temporal resolutions.

Finally, the performance (measured by the GPI) of the developed model with a temporal resolution of 15 min is better than the one for 10, 5 and 60 min for the data used in this work which is in accordance with the results shown in Table 9.

6. Application of the ANN models for different locations

The ANN models described in the previous sections were developed based on data for a specific location (Évora - Verney). In this section, the hypothesis of using the same models in other locations in the region surrounding the reference station of Évora-Verney (south of Portugal) is analyzed. The ECMWF and CAMS forecast data obtained for six different sites where solar radiation measuring stations are installed (DNI-A network, see Fig. 1) were used as inputs. The original 10 min downscaled ECMWF forecasts and the DNI forecasts obtained from the ANN model B were compared with the available observed DNI data for each station. The results and the number of data points (N) for each station are shown in Table 11, where the data used for evaluating Verney station is comprised of the test and blind test datasets.

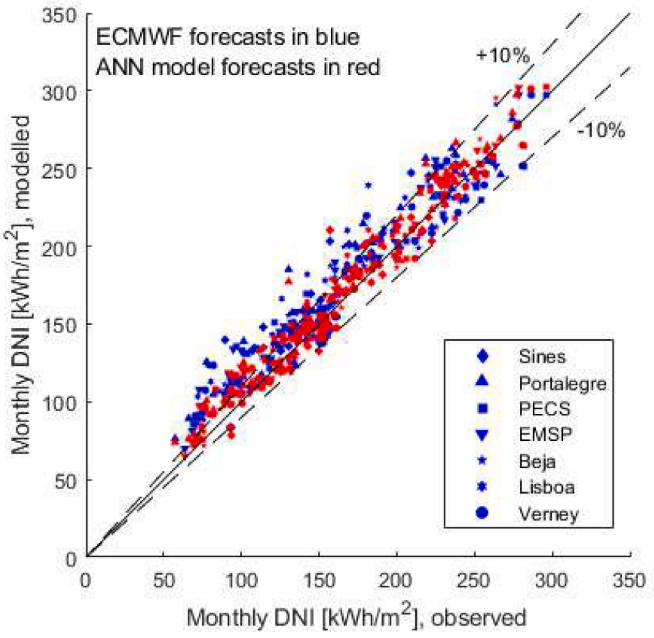


Fig. 5. Comparison between monthly DNI forecasts (irradiation) from the ECMWF (blue) and from ANN model (red) against the available experimental data for each station in the South of Portugal.

The results show that the developed model can improve the overall metrics for all stations being best at generating improved DNI forecasts for the station of Verney which data was used in the development of the algorithm, as expected. The stations for which the improvement of DNI forecasts is lower are Lisboa and Sines, both located near the coast of Portugal where the climate is more disparate from that observed at the station of Évora–Verney. Considering all stations and data available, the use of the developed models result in a forecast skill of 14.4 % and improve the original downscaled DNI forecats by 0.0707 in R^2, 23.1 W/m^2 in MAE and 24.1 W/m^2 in RMSE, taking experimental data as reference showing that the application of this model for locations in the surrounding area of the location for which it was developed, can generate improved DNI forecasts.

Fig. 5 shows a direct comparison between the monthly forecasted DNI (irradiation) by the ECMWF and the developed ANN models (based on 10 min data) and the available monthly observed DNI for the different stations, while Table 12 presents the metrics resulting from this data. For the Verney station only the data from test and blind test datasets was considered.

Considering the monthly irradiation values for all stations, an improvement over the original DNI forecasts from the ECMWF of 0.0331, 5.9 kWh/m^2 and 6.6 kWh/m^2 for R^2, MAE and RMSE, respectively, is achieved by the developed algorithm with a forecast skill of 38.1 %.

Table 12
Metrics for the monthly original spatially downscaled ECMWF DNI (irradiation) forecasts and ANN model B predictions for the different stations located in the south of Portugal.

Station	R^2		MAE (kWh/m^2)		RMSE (kWh/m^2)		FS (%)
	ECMWF	ANN	ECMWF	ANN	ECMWF	ANN	
Verney	0.9431	0.9771	13.0	7.4	16.3	9.4	43.0
EMSP	0.9483	0.9758	15.0	9.8	17.6	12.1	34.9
PECS	0.9518	0.9803	13.6	7.8	16.8	9.7	43.1
Portalegre	0.9402	0.9717	15.9	11.9	19.7	14.9	25.5
Beja	0.9472	0.9689	14.0	9.6	16.2	12.3	31.0
Lisboa	0.8948	0.9598	22.2	11.4	25.7	13.6	48.6
Sines	0.8244	0.8742	18.3	9.6	22.1	13.9	47.6
All	0.9313	0.9644	15.5	9.6	18.9	12.3	38.1

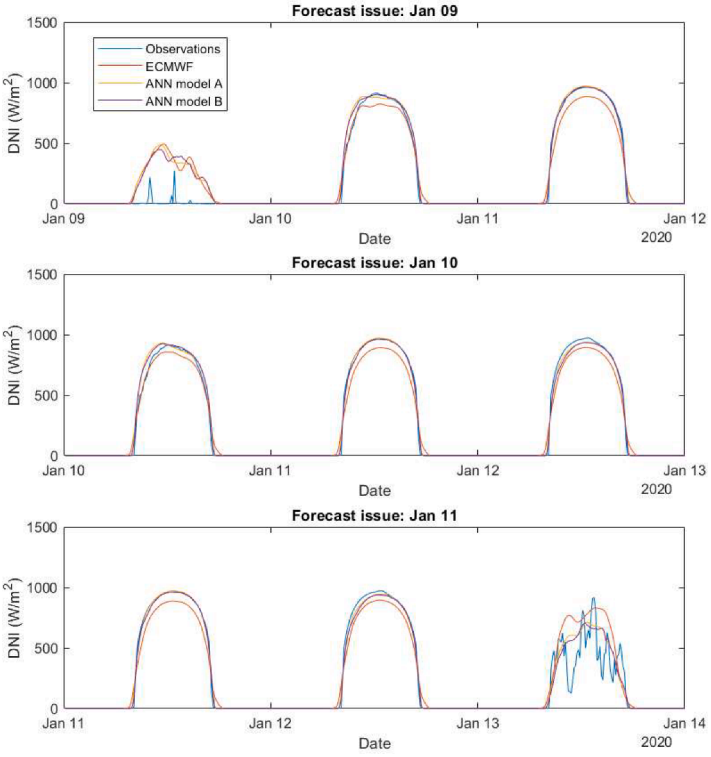


Fig. 6. Example of operational use of the developed model for 3 consecutive days.

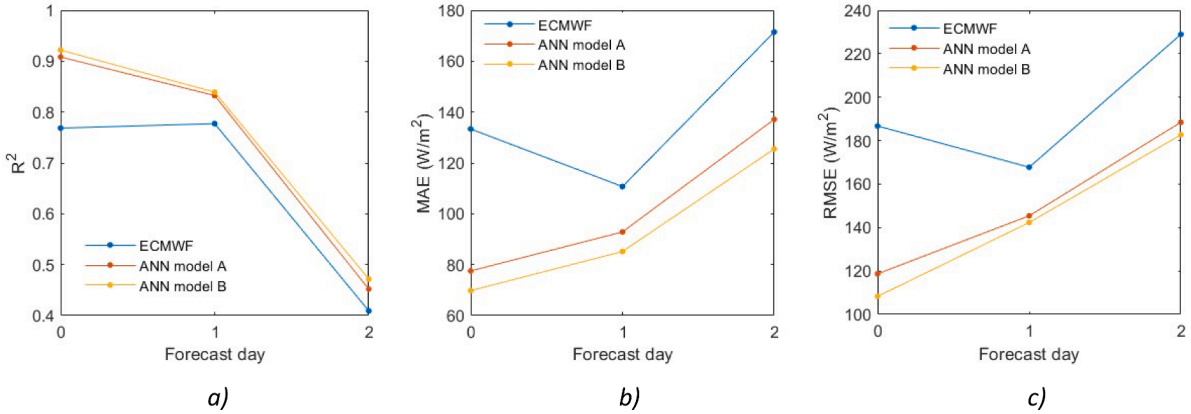
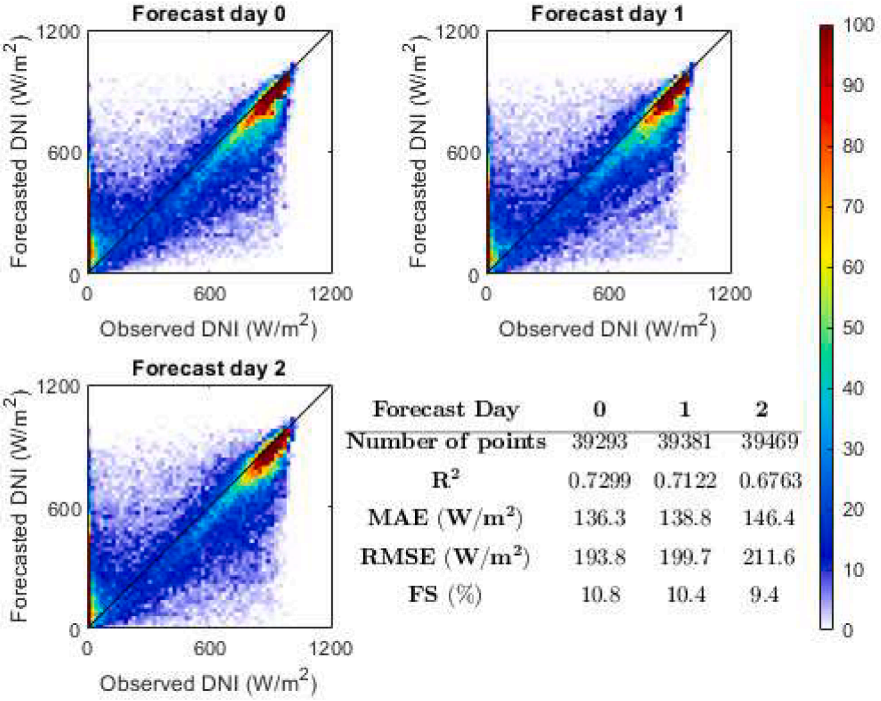
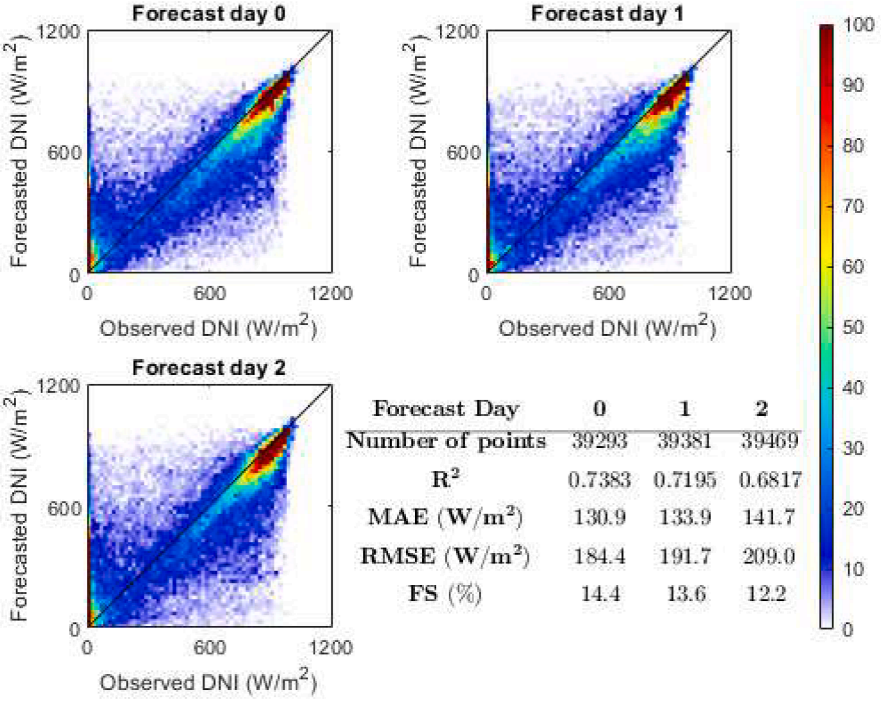


Fig. 7. Variation of statistical indicators (a) R^2, (b) MAE and (c) RMSE for Jan. 11, 2020 using 00:00 UTC forecasts at day 0 (forecast issue: Jan. 11), 1 day ahead (forecast issue: Jan. 10) and 2 days ahead (forecast issue: Jan. 9), based on 10 min data.



Forecast Day	0	1	2
Number of points	39293	39381	39469
R^2	0.7299	0.7122	0.6763
MAE (W/m^2)	136.3	138.8	146.4
RMSE (W/m^2)	193.8	199.7	211.6
FS (%)	10.8	10.4	9.4

Fig. 8. Results of ANN model A for test and blind test 10 min data of Évora-Verney station (the colormap represents the number of data points in each bin. Bin size: 20 × 20 W/m^2).



Forecast Day	0	1	2
Number of points	39293	39381	39469
R^2	0.7383	0.7195	0.6817
MAE (W/m^2)	130.9	133.9	141.7
RMSE (W/m^2)	184.4	191.7	209.0
FS (%)	14.4	13.6	12.2

Fig. 9. Results of ANN model B for test and blind test 10 min data of Évora-Verney station (the colormap represents the number of data points in each bin. Bin size: 20 × 20 W/m^2).

7. Application of the developed model for operational DNI forecasts

The model developed in this work can be used operationally to obtain more accurate DNI forecasts which can help solar energy systems and power grid operators to estimate the energy generation and make better informed decisions. After retrieving the operational forecasts of ECMWF of the 00UTC run, the model is applied resulting in 10 min DNI forecasts for the present day and the next 2 days. As an example, Fig. 6 was generated, where the model was used to generate DNI forecasts for 3 consecutive days (forecast issue time at 00:00 UTC of 9, 10 and 11 of January 2020, as example) at the location of Évora - Verney station. Here, the model results show very good agreement with observations for clear sky conditions, with larger errors for overcast or partially cloudy

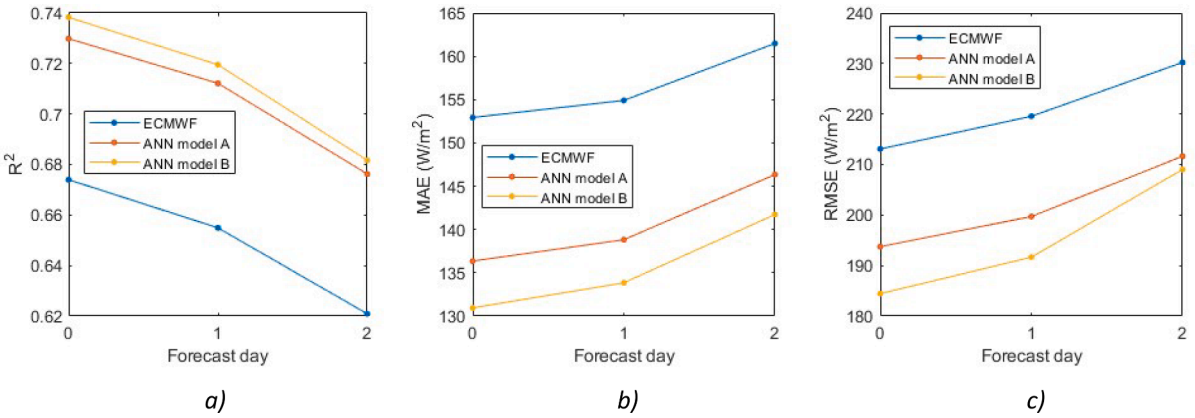


Fig. 10. Variation of statistical indicators (a) R^2, (b) MAE and (c) RMSE results for Évora–Verney station for each forecast day ahead, based on 10 min data.

skies, as expected. The day January 11 is forecasted for all presented runs of the model and, in order to assess the variation of model performance along the successive forecast issues, the variation of statistical indicators R^2, MAE and RMSE for this day is shown in Fig. 7.

In this case, the performance of both ANN models is better for smaller forecast time horizons where the ANN model B achieved the best results followed by the ANN model A and finally the ECMWF model. It should be noted that this is not a typical behavior. As previously indicated in Figs. 2 and 3, the irradiation forecast tends to worsen with the horizon. However, although, in general, the predictability of solar irradiance decreases with the time horizon, in particular cases the forecast for two days ahead can be better than the forecast for the next day. A recent and good discussion about the predictability of the numerical prediction of solar irradiation and its variation with the horizon can be found in [45].

As this is just an example for a specific day, the same operational test was carried out for the test and blind test datasets of Évora–Verney station and the results can be seen in Figs. 8 and 9 for the ANN model A and ANN model B, respectively. Comparing these figures with the one obtained for the original downscaled ECMWF forecasts (Fig. 2), a higher density of data points near the bisection line is found for the forecasts of the ANN models.

A similar way to visualize these results to that of Fig. 7 is shown in Fig. 10 for each forecasting model and each of the forecast days, now for all data and not for a specific day. Again, all models show an overall increase in performance for shorter forecast time horizons and all metrics are improved by the models developed in this work for all days of the forecast time horizon.

8. Conclusions

This work presents the development of a model that can successfully generate improved DNI forecasts based on data from NWP models in an operational setting. This model uses forecast data from the IFS/ECMWF and CAMS of several atmospheric variables including aerosol data which affects the transport of solar radiation through the atmosphere. These forecasts are spatially and temporally downscaled to the location and time resolution desired using bi-linear interpolation of the values of the four surrounding grid points and piecewise cubic interpolation of the hourly mean variables, respectively. Then, two different models based on artificial neural networks were designed and optimized to generate improved DNI forecasts with the desired temporal resolution and for a forecast time horizon of 72 h. Different configurations of these models were tested and the selected configuration uses two feedforward ANNs in series, that is, the second ANN uses the output of the first as input. In an operational setting, the model is run every day when the ECMWF and CAMS operational forecasts are available for retrieval (maximum dissemination time 06:55 UTC) and results for the next two days can be used by solar energy producers and power grid operators to estimate the energy production and make better informed decisions.

Comparing the final model results using only test and blind test datasets with a temporal resolution of 10 min and for the location of Évora–Verney (the location used for model development) with the ground-based observations at the same location, values of R^2, MAE and RMSE of 0.7195, 133.9 W/m^2 and 191.7 W/m^2 for forecast day 1 and of 0.6812, 141.7 W/m^2 and 209.0 W/m^2 for forecast day 2 were achieved, respectively. This results in an improvement over the original downscaled DNI forecast from ECMWF of 0.0646, 21.1 W/m^2 and 27.9 W/m^2 for forecast day 1 and of 0.0608, 19.8 W/m^2 and 21.2 W/m^2 for forecast day 2 for R^2, MAE and RMSE, respectively.

The model was also applied to other locations scattered in the region surrounding the site for which it was developed showing improvements of the DNI forecasts of 0.0713, 23.3 W/m^2 and 24.2 W/m^2 for day 1 of forecast and for R^2, MAE and RMSE respectively, when compared to the ECMWF forecasts and taking the ground-based measurements in each location as reference.

It was also shown that the already developed algorithm can be applied to similar temporal resolutions, such as 5 or 15 min, achieving good results which can be extremely helpful when forecasts with different time steps are needed so that they are in accordance with the real-time market forecasting requirements.

With these improved forecasts, more accurate estimations of energy generation in solar energy systems can be achieved which is extremely important for solar power plant and energy market operators.

Declaration of Competing Interest

The authors declare that they have no known 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

This work was funded by National funds through FCT - Fundação para a Ciência e Tecnologia, I.P. (projects UIDB/04683/2020 and UIDP/04683/2020). S. Pereira acknowledges the support of FCT- Fundação para a Ciência e Tecnologia through the grant with reference SFRH/BD/145378/2019. The authors are thankful to Afonso Cavaco for the scientific and technical support and for maintaining the measuring stations and instruments from the DNI-Alentejo network.

References

[1] IEA. World energy outlook 2021 - revised version october 2021. 2021.
[2] El Alani O, Ghennioui A, Ghennioui H, Saint-Drenan Y-M, Blanc P. Evaluation of 24-hours forecasts of global solar irradiation from IFS, GFS and McClear models. AIP Conf Proc 2020. https://doi.org/10.1063/5.0032744.
[3] Perez R, Lorenz E, Pelland S, Beauharnois M, Van Knowe G, Hemker K, Heinemann D, Remund J, Müller SC, Traunmüller W, Steinmauer G, Pozo D, Ruiz-Arias JA, Lara-Fanego V, Ramirez-Santigosa L, Gaston-Romero M, Pomares LM. Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe. Solar Energy 2013;94:305–26. https://doi.org/10.1016/J.SOLENER.2013.05.005.
[4] Perdigão J, Canhoto P, Salgado R, Costa MJ. Assessment of direct normal irradiance forecasts based on IFS/ECMWF data and observations in the south of Portugal. Forecasting 2020;2:130–50. https://doi.org/10.3390/forecast2020007.
[5] de Araujo JMS. Improvement of coding for solar radiation forecasting in Dili Timor Leste—a WRF case study. J Power Energy Eng 2021;09:7–20. https://doi.org/10.4236/jpee.2021.92002.
[6] Hogan R. Radiation in the next generation of weather forecast models : workshop report. 2018. p. 1–3.
[7] Yang D, Wang W, Gueymard CA, Hong T, Kleissl J, Huang J, Perez MJ, Perez R, Bright JM, Xia X, van der Meer D, Peters IM. A review of solar forecasting, its dependence on atmospheric sciences and implications for grid integration: towards carbon neutrality. Renew Sustain Energy Rev 2022;161:112348. https://doi.org/10.1016/j.rser.2022.112348.
[8] Breitkreuz H, Schroedter-Homscheidt M, Holzer-Popp T, Dech S. Short-range direct and diffuse irradiance forecasts for solar energy applications based on aerosol chemical transport and numerical weather modeling. J Appl Meteorol Climatol 2009;48:1766–79. https://doi.org/10.1175/2009JAMC2090.1.
[9] Mayer B, Kylling A, Emde C. libRadtran user ' s guide. 2012.
[10] Bran SH, Jose S, Srivastava R. Investigation of optical and radiative properties of aerosols during an intense dust storm: a regional climate modeling approach. J Atmos Sol Terr Phys 2018;168:21–31. https://doi.org/10.1016/J.JASTP.2018.01.003.
[11] Obiso V, Jorba O. Aerosol-radiation interaction in atmospheric models: idealized sensitivity study of simulated short-wave direct radiative effects to particle microphysical properties. J Aerosol Sci 2018;115:46–61. https://doi.org/10.1016/j.jaerosci.2017.10.004.
[12] Larson VE. Forecasting solar irradiance with numerical weather prediction models. Solar Energy Forecast Resour Assess 2013:299–318. https://doi.org/10.1016/B978-0-12-397177-7.00012-7.
[13] Verzijlbergh RA, Heijnen PW, de Roode SR, Los A, Jonker HJJ. Improved model output statistics of numerical weather prediction based irradiance forecasts for solar power applications. Solar Energy 2015;118:634–45. https://doi.org/10.1016/J.SOLENER.2015.06.005.
[14] Alkhayat G, Mehmood R. A review and taxonomy of wind and solar energy forecasting methods based on deep learning. Energy AI 2021;4:100060. https://doi.org/10.1016/J.EGYAI.2021.100060.
[15] Alkhayat G, Hasan SH, Mehmood R. SENERGY: a novel deep learning-based auto-selective approach and tool for solar energy forecasting. Energies (Basel) 2022:15. https://doi.org/10.3390/en15186659.
[16] Delle Monache L, Eckel FA, Rife DL, Nagarajan B, Searight K. Probabilistic weather prediction with an analog ensemble. Mon Weather Rev 2013;141:3498–516. https://doi.org/10.1175/MWR-D-12-00281.1.
[17] Alfadda A, Rahman S, Pipattanasomporn M. Solar irradiance forecast using aerosols measurements: a data driven approach. Solar Energy 2018;170:924–39. https://doi.org/10.1016/j.solener.2018.05.089.
[18] Pang Z, Niu F, O'Neill Z. Solar radiation prediction using recurrent neural network and artificial neural network: a case study with comparisons. Renew Energy 2020;156:279–89. https://doi.org/10.1016/J.RENENE.2020.04.042.
[19] McCandless TC, Young GS, Haupt SE, Hinkelman LM. Regime-dependent short-range solar irradiance forecasting. J Appl Meteorol Climatol 2016;55:1599–613. https://doi.org/10.1175/JAMC-D-15-0354.1.
[20] Fonseca JGDS, Uno F, Ohtake H, Oozeki T, Ogimoto K. Enhancements in day-ahead forecasts of solar irradiation with machine learning: a novel analysis with the Japanese mesoscale model. J Appl Meteorol Climatol 2020;59:1011–28. https://doi.org/10.1175/JAMC-D-19-0240.1.
[21] Lima FJL, Martins FR, Pereira EB, Lorenz E, Heinemann D. Forecast for surface solar irradiance at the Brazilian Northeastern region using NWP model and artificial neural networks. Renew Energy 2016;87:807–18. https://doi.org/10.1016/j.renene.2015.11.005.
[22] Pereira S, Abreu EFM, Iakunin M, Cavaco A, Salgado R, Canhoto P. Method for solar resource assessment using numerical weather prediction and artificial neural network models based on typical meteorological data: application to the south of Portugal. Solar Energy 2022;236:225–38. https://doi.org/10.1016/j.solener.2022.03.003.
[23] Voyant C, Notton G, Kalogirou S, Nivet M, Paoli C, Motte F, Fouilloy A. 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.
[24] 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. https://doi.org/10.1016/J.ENCONMAN.2021.113960.
[25] Yang B, Zhu T, Cao P, Guo Z, Zeng C, Li D, Chen Y, Ye H, Shao R, Shu H, Yu T. Classification and summarization of solar irradiance and power forecasting methods: a thorough review. CSEE J Power Energy Syst 2021:1–19. https://doi.org/10.17775/CSEEJPES.2020.04930.
[26] Liu W, Liu Y, Zhou X, Xie Y, Han Y, Yoo S, Sengupta M. Use of physics to improve solar forecast: physics-informed persistence models for simultaneously forecasting GHI, DNI, and DHI. Solar Energy 2021;215:252–65. https://doi.org/10.1016/j.solener.2020.12.045.
[27] Liu W, Liu Y, Zhang T, Han Y, Zhou X, Xie Y, Yoo S. Use of physics to improve solar forecast: part II, machine learning and model interpretability. Solar Energy 2022;244:362–78. https://doi.org/10.1016/j.solener.2022.08.040.
[28] Voyant C, Notton G, Duchaud J, Gutiérrez LAG, Bright JM, 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.
[29] Visser L, Alskaif T, Hu J, Louwen A, Van Sark W. On the value of expert knowledge in estimation and forecasting of solar photovoltaic power generation. Solar Energy 2023;251:86–105. https://doi.org/10.1016/j.solener.2023.01.019.
[30] hyun Jung A, Lee D-H, Kim J-Y, Kim CK, Kim H-G, Lee Y-S. Spatial and temporal downscaling of solar radiation using statistical techniques. J Korean Solar Energy Soc 2020;40:89–94. https://doi.org/10.7836/kses.2020.40.6.089.
[31] Yang D, Wu E, Kleissl J. Operational solar forecasting for the real-time market. Int J Forecast 2019;35:1499–519. https://doi.org/10.1016/j.ijforecast.2019.03.009.
[32] ECMWF. IFS documentation. 2021. https://www.ecmwf.int/en/publications/ifs-documentation (accessed December 10, 2021).
[33] Hogan RJ, Bozzo A. A flexible and efficient radiation scheme for the ECMWF model. J Adv Model Earth Syst 2018;10:1990–2008. https://doi.org/10.1029/2018MS001364.
[34] ECMWF. CAMS: global atmospheric composition forecast data documentation. 2021. https://confluence.ecmwf.int/display/CKB/CAMS%3A+Global+atmospheric+composition+forecast+data+documentation (accessed December 10, 2021).
[35] Abreu EFM, Gueymard CA, Canhoto P, Costa MJ. Performance assessment of clear-sky solar irradiance predictions using state-of-the-art radiation models and input atmospheric data from reanalysis or ground measurements. Solar Energy 2023;252:309–21. https://doi.org/10.1016/J.SOLENER.2023.01.051.
[36] Iakunin M, Abreu EFM, Canhoto P, Pereira S, Salgado R. Impact of a large artificial lake on regional climate: a typical meteorological year Meso-NH simulation results. Int J Climatol 2021:1–22. https://doi.org/10.1002/joc.7299.
[37] Lopes FM, Conceição R, Silva HG, Fasquelle T, Salgado R, Canhoto P, Collares-Pereira M. Short-term forecasts of DNI from an integrated forecasting system (ECMWF) for optimized operational strategies of a central receiver system. Energies (Basel) 2019;12. https://doi.org/10.3390/en12071368.
[38] Cavaco A, Canhoto P, Pereira MCollares. Procedures for solar radiation data gathering and processing and their application to DNI assessment in southern Portugal. Renew Energy 2021;163:2208–19. https://doi.org/10.1016/J.RENENE.2020.10.075.
[39] Long CN, Dutton EG. BSRN Global Network recommended QC tests, V2. J Clim 2010;25:8542–67. https://doi.org/10.1175/JCLI-D-11-00618.1.
[40] Shcherbakov M, Brebels A, Tyukov A. A survey of forecast error measures. World Appl Sci J 2013;24:171–6. https://doi.org/10.5829/idosi.wasj.2013.24.itmies.80032.
[41] Botchkarev A. A new typology design of performance metrics to measure errors in machine learning regression algorithms. Interdisc J Inform Knowl Manage 2019;14:045–76. https://doi.org/10.28945/4184.
[42] Jamil B, Akhtar N. Comparative analysis of diffuse solar radiation models based on sky-clearness index and sunshine period for humid-subtropical climatic region of India: a case study. Renew Sustain Energy Rev 2017;78:329–55. https://doi.org/10.1016/j.rser.2017.04.073.
[43] Sharifi SS, Rezaverdinejad V, Nourani V. Estimation of daily global solar radiation using wavelet regression, ANN, GEP and empirical models: a comparative study of selected temperature-based approaches. J Atmos Sol Terr Phys 2016;149:131–45. https://doi.org/10.1016/j.jastp.2016.10.008.
[44] Pereira S, Canhoto P, Salgado R, Costa MJ. Development of an ANN based corrective algorithm of the operational ECMWF global horizontal irradiation forecasts. Solar Energy 2019;185:387–405. https://doi.org/10.1016/j.solener.2019.04.070.
[45] Liu B, Wang J, Chen J, Li B, Sun D, Zhang G. A probabilistic perspective on predictability of solar irradiance using bootstrapped correlograms and ensemble predictability error growth. Solar Energy 2023;260:17–24. https://doi.org/10.1016/J.SOLENER.2023.05.053.