APPLIED SOLAR ENERGY

Generalized global solar radiation forecasting model via cyber-secure deep federated learning

Arash Moradzadeh[1] · Hamed Moayyed[2] · Behnam Mohammadi-Ivatloo[1,3] · António Pedro Aguiar[4] · Amjad Anvari-Moghaddam[5] · Zulkurnain Abdul-Malek[6]

Abstract

Recently, the increasing prevalence of solar energy in power and energy systems around the world has dramatically increased the importance of accurately predicting solar irradiance. However, the lack of access to data in many regions and the privacy concerns that can arise when collecting and transmitting data from distributed points to a central server pose challenges to current predictive techniques. This study proposes a global solar radiation forecasting approach based on federated learning (FL) and convolutional neural network (CNN). In addition to maintaining input data privacy, the proposed procedure can also be used as a global supermodel. In this paper, data related to eight regions of Iran with different climatic features are considered as CNN input for network training in each client. To test the effectiveness of the global supermodel, data related to three new regions of Iran named Abadeh, Jarqavieh, and Arak are used. It can be seen that the global forecasting supermodel was able to forecast solar radiation for Abadeh, Jarqavieh, and Arak regions with 95%, 92%, and 90% accuracy coefficients, respectively. Finally, in a comparative scenario, various conventional machine learning and deep learning models are employed to forecast solar radiation in each of the study regions. The results of the above approaches are compared and evaluated with the results of the proposed FL-based method. The results show that, since no training data were available from regions of Abadeh, Jarqavieh, and Arak, the conventional methods were not able to forecast solar radiation in these regions. This evaluation confirms the high ability of the presented FL approach to make acceptable predictions while preserving privacy and eliminating model reliance on training data.

Responsible Editor: Philippe Garrigues

✉ Behnam Mohammadi-Ivatloo
mohammadi@ieee.org

Arash Moradzadeh
arash.moradzadeh@tabrizu.ac.ir

Hamed Moayyed
mha@isep.ipp.pt

António Pedro Aguiar
pedro.aguiar@fe.up.pt

Amjad Anvari-Moghaddam
aam@energy.aau.dk

Zulkurnain Abdul-Malek
zulkurnain@utm.my

1   Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz 5166616471, Iran

2   GECAD - Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development,

    LASI - Intelligent Systems Associate Laboratory, Polytechnic of Porto, P-4200-072 Porto, Portugal

3   School of Energy Systems, LUT University, Lappeenranta, Finland

4   SYSTEC-ARISE Research Center for Systems and Technologies, Electrical and Computer Enginnering Department, Faculty of Engineering, University of Porto, P-4200 465 Porto, Portugal

5   Department of Energy (AAU Energy), Aalborg University, 9220 Aalborg, Denmark

6   Institute of High Voltage & High Current, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, Johor Bahru, Malaysia

Introduction

Increasing energy demand and environmental issues have led to significant penetration of renewable energy sources (RESs) in the power systems and promoting the transition of energy structure (Dadashi et al. 2022). Solar energy has been introduced as a clean and permanent energy among RESs, so photovoltaic (PV) energy production is becoming an important part of the energy portfolio worldwide (Nasirpour et al. 2021; Rezazadeh et al. 2022). According to a statistical report through the International Energy Agency, 2.5% of the world's total primary energy and 18.5% of RESs in 2019 have been supplied by solar energy (International Energy Agency (IEA) 2019). As indicated in another section of these statistics, the average annual growth rate of solar energy is 37%, which has created new possibilities and challenges for integration into existing power and energy systems (International Renewable Energy Agency 2020). However, the dependence of solar radiation on the synoptic meteorological pattern has caused the power generated by solar energy under PV energy technology to always have high variability and fluctuation (Wang et al. 2019a; Riahi et al. 2021). In addition, it causes some problems and inconsistencies related to the stabilization of energy resources and demand in electric power systems, particularly when the storage capacity of PV systems is low. Accordingly, solar radiation is a critical and effective parameter for solar energy applications. Awareness of the amount of solar radiation in different time periods can play an important role in the scheduling and planning of power network operators, residential solar energy operators, and industrial projects operators to manage their energy production and consumption (Bosman and Darling 2018; Sharda et al. 2021). Thus, the ability to forecast solar radiation is extremely important in terms of reliability and performance and is considered a critical factor in the penetration of solar energy into the power and energy networks. Also, accurate forecasts support the safe operation of the grid and allow maximum access to solar energy. To date, solar radiation forecasting based on various methods such as physical model (Caldas and Alonso-Suárez 2019; Kakimoto et al. 2019; Marzouq et al. 2020), statistical approach (Scolari et al. 2018; Van Der Meer et al. 2018; Louzazni et al. 2020), data mining-based solutions such as artificial neural network (ANN) (Khan et al. 2020; Wang et al. 2020a, b; Zambrano and Giraldo 2020), machine learning (Prasad et al. 2019; Deo et al. 2019; Yagli et al. 2019; Feng et al. 2019; Belmahdi et al. 2022), and deep learning (Wang et al. 2019b; Khodayar et al. 2020; Wen et al. 2021; Abdel-Nasser et al. 2021; Liu et al. 2022) techniques has been investigated in various studies.

Physical models forecast the solar radiation based on the principles of PV cell production and numerical weather prediction (NWP). In Caldas and Alonso-Suárez (2019), a novel model for short-term solar radiation forecasting via sky measurements and online imaging has been presented. In Marzouq et al. (2020), a multi-model evolutionary framework based on a physical model for solar radiation forecasting has been developed. A probabilistic approach to solar radiation prediction based on the joint probability distribution function (PDF) of the radiation calculated by NWP and the observed radiation was proposed in (Kakimoto et al. (2019). In the same study, the probabilistic forecast has been achieved by deriving a conditional PDF given a current NWP via the Bayes rule. The use of these methods is not cost-effective and requires additional peripherals such as sky cameras to collect input data. In addition, the collection of input data that determines the amount of solar radiation requires experienced people.

Statistical methodologies are usually suitable for time-series estimating of solar radiation. In Louzazni et al. (2020), the autoregressive moving average method is considered a statistical technique for forecasting solar radiation and PV output. In another similar work (Van Der Meer et al. 2018), an autoregressive integrated moving average model has been employed to forecast PV power production based on solar radiation. Two different mathematical methods and the Kalman filter technique are also proposed in (Scolari et al. (2018) as predictors of solar radiation. Despite their low cost, these methods have high mathematical computational complexity, and in most cases, the results of forecasts are not very accurate.

Data mining-based solutions make the necessary forecasts based on data affecting the amount of solar radiation. Each of these techniques has strengths and weaknesses in dealing with different datasets. One of the ANN techniques called direct explainable neural network, a forward-forward neural network, has been employed in Wang et al. (2020a) to forecast solar radiation at various time-horizons in Lyon, France. In Wang et al. (2020b), various ANN-based techniques for solar radiation prediction have been reviewed. In Zambrano and Giraldo (2020), an ANN-based procedure has been trained to forecast solar radiation with a horizon altering from 1 to 48 h utilizing full on-site data. A hybrid model based on multi-stage multivariate empirical mode disintegration combined with a random forest and ant colony optimization was created in Prasad et al. (2019) for solar radiation prediction. The developed algorithm performs the forecasting process in three steps including feature extraction, determining the best features, and finally forecasting the monthly solar radiation. Belmahdi et al. (2022) provide a complete assessment of the machine learning and time

series methods used for solar radiation forecasting. In Yagli et al. (2019), the hourly solar radiation forecasting for 7 locations and 5 climatic zones in the continental United States has been done using 68 machine learning algorithms in a comparative approach. In a valuable study (Deo et al. 2019), another machine learning application called extreme learning machine has been selected to forecast long-term solar radiation in Australia. An unsupervised clustering-based solar forecasting procedure has been developed in Feng et al. (2019) for short-term global horizontal solar radiation forecasting. The suggested method in this study consists of three parts: pattern recognition, global horizontal radiation time series unsupervised clustering, and unsupervised clustering-based forecasting. The ANN-based and machine learning methods have been able to significantly address the shortcomings of prior techniques. However, these solutions suffer from the processing of high-dimensional and time-series data and do not have the ability to model the behavioral pattern of this data, which reduces the accuracy of forecasts.

In Abdel-Nasser et al. (2021), solar radiation forecasting was performed with a deep-learning application called long short-term memory (LSTM) and an aggregation function based on the Choquet integral employing six realistic datasets gathered at different locations in Finland. A multi-step solar forecasting scheme based on a convolutional neural network (CNN) has been introduced in Wen et al. (2021) for PV power ramp-rate control. In that study, in addition to meteorological data, stacked sky images that combine temporal-spatial information of cloud motions were utilized to improve forecasting performance. In Khodayar et al. (2020), a hybrid model called convolutional graph autoencoder, which combines two deep learning techniques, namely, the use of a CNN and autoencoders, has been presented to forecast solar radiation in multiple sites in a wide area in the northern states of the USA. An integrated precise short-term solar radiation forecasting model has been suggested in Wang et al. (2019b) based on the least absolute shrinkage and selection operator and LSTM. A novel solar radiation forecasting model based on the LSTM, variational Bayesian inference, and federated learning (FL) has been introduced in Zhang et al. (2020). In recent years, deep learning techniques with influence on power and energy system applications could solve problems related to machine learning methods in this field. However, these techniques often have difficulty processing real-time data that involve noise and require pre-processing. In addition, all reviewed forecasting methods have the ability to provide a local forecasting model and require training data from that area to provide forecasting in a region. To date, none of the studies has presented a global model with the ability to maintain data security and generalizability that can forecast solar radiation without the need for training data from a region.

In this study, a global solar radiation forecasting model with preserving data privacy is presented for the first time. The developed forecasting model is based on the FL approach, which uses the CNN technique in each client to accomplish the training cycle and extract features from the training data to form the main server, which is considered a global model. Each CNN network designed on each client of FL structure, while preserving the data security, extracting data-related features, and recognizing the pattern of solar radiation behavior at various time intervals. Then, based on the features extracted in each client, a global supermodel is generated. The proposed method is trained and tested using meteorological data related to different regions of Iran such as Birjand, Seydabad, Abhar, Mahan, Eqlid, Khaf, Meybod, and Tabriz that have solar power plants to produce the main server or global supermodel. To express the effectiveness of the proposed method and the performance of the global supermodel generated in solar radiation forecasting for regions whose data had no role in the formation of the server or global supermodel, meteorological data of new regions such as Abadeh, Jarqavieh, and Arak are used to test the global supermodel. In addition, in order to express the performance of the developed novel procedure against traditional models, conventional forecasting models such as decision tree (DT), back-propagation neural network (BPNN), LSTM, and bidirectional LSTM (Bi-LSTM) were also utilized to estimate solar radiation in the proposed regions. Figure 1 illustrates the overview of the solar radiation forecasting in this work.

In general, the most important contributions and novelty of this study are listed as follows:

(i) For the first time, an FL-based solar radiation forecasting model is suggested to offer a secure procedure for protecting data privacy by training the forecasting models locally and avoiding the exchange of raw data across various solar power plants.

(ii) The suggested procedure enables forecasting to benefit from the improved performance provided by global supermodel aggregation in the absence of data exchange.

(iii) Provide a generalizable model that is capable of forecasting solar radiation in various regions with different climatic conditions for which no training data are available. The higher generalizability of the proposed model is also substantiated while the privacy of the underlying data is preserved.

(iv) Various artificial intelligence-based techniques such as DT, BPNN, LSTM, and Bi-LSTM were used to forecast solar radiation in the study regions, in order to appraise the performance of the developed model compared to conventional models in a comparative approach.
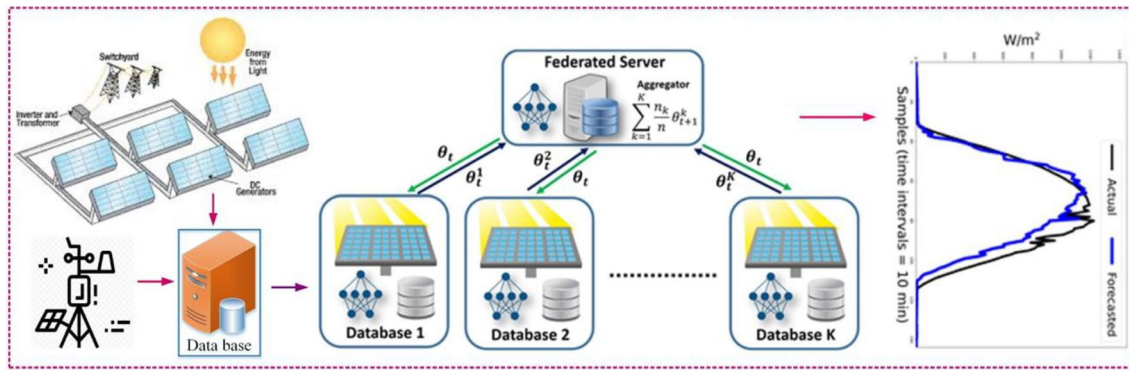
Fig. 1 The overview of the proposed approach for solar radiation forecasting

(v)   The evaluation of the results obtained is done for the proposed technique based on various performance evaluation indicators that indicate the high accuracy of the proposed FL-based solar radiation forecasting using real-world data sets.

The remainder of this paper is structured as follows: the methodologies used such as FL and CNN are described in the "Methodologies" section. The "Case studies" section introduces the study regions under the title of case studies. The "Solar radiation forecasting results" section presents the results of solar radiation forecasting. Finally, the paper is concluded in the "Conclusion" section.

Methodologies

In this paper, the FL technique is proposed as one of the novel machine learning techniques to generate a global supermodel of solar radiation forecasting. The process of training and extracting features in each client from the federated network is done by one of the most well-known deep learning techniques called CNN. Conventional models of DT, BPNN, LSTM, and Bi-LSTM also are employed for solar radiation forecasting in the studied regions to provide a comparative approach for evaluating the results. The rest of this section introduces each of the developed forecasting procedures.

Federated learning (FL)

Machine learning (ML) techniques including deep learning (DL) models are becoming increasingly complex as datasets grow, and there is a strong need to distribute the analysis process and optimization of model parameters across different machines. Moreover, current ML and DL models are trained for a specific dataset and in most

cases they fail when analyzing observations that are not part of the data sample and have a different distribution. The generalization property is a desired feature for ML and DL models. Considering these properties, Google recently introduced the FL technique (Konečný et al. 2016), which proposes a communicated global model to link local ML models without seeing their location and, more significantly, without compromising privacy (Lu et al. 2021; Wang et al. 2021).

The main goal of the FL technique is to perform a global model with a partnership of several local parties (clients) that maintains the protection of the data while each local party is typically large and has Internet networks. This technique is called a decentralized privacy technique that allows clients to train their ML model without losing the privacy of the data. Furthermore, they do not need a central server to upload their local data. According to each iteration of the training, a distributed model is taken from the global server. Later, it is trained based on the local data, and eventually, the server gets the updated weights or gradients. The uploaded models by the clients must be aggregated to obtain a new overall model on the server.

FL has more robust local computational capabilities and more communication resources compared to traditional centralized learning. The local edges are taken into account when distributing the learning data, but the global server cloud does not have access to them. Nevertheless, the learned model is distributed to the clients and the server. Moreover, the local device is considered as the location for model training rather than the server. The server's task is to aggregate the local models uploaded by the clients. The goal is to obtain a distributed global model and send it back to the clients (Lu et al. 2021; Moayyed et al. 2022).

A typical FL process is made up of one server and K clients, as shown in Fig. 2, and Fig. 3 shows the general process of implementing the developed FL method in the

form of a flowchart. The FL algorithm is similar to the training algorithm of the multilayer perceptron neural network and intends to minimize the loss function $l(\theta)$ while working in a distributed scheme (Moayyed et al. 2022):

$$l(\theta) = \sum_{k=1}^{K} \frac{n_k}{n} L_k(\theta) \qquad (1)$$

where K denotes the number of total clients' index and $L_k(\theta)$ represents the loss function of the kth local client and can be formulated as:

$$L_k(\theta) = \frac{1}{n_k} \sum_{i \in P_k} l_i(\theta) \qquad (2)$$

where P_k is considered a set of data indexes with the length of n_k, i.e., $n_k = |P_k|$.

As shown in Fig. 3, the implementation process of the developed model is possible in 4 steps as follows (Moradzadeh et al. 2021a):

- Step 1 (local training): All clients receive the parameters θ_t of the global model from the central server for local training. Then, the clients' exclusive local models are trained by employing their own data.
- Step 2 (model aggregating): When the local parameters are trained, the server receives these parameters, such as θ_t^1, from the local devices to perform secure aggregation across the local parties. The federated averaging algorithm (FedAvg) is the most common method for server aggregation (Kapur et al. 2011).
- Step 3 (parameter distribution): An updated global model θ_{t+1} is broadcasted for the next iteration's training;
- Step 4 (model updating): All clients update their local ML models based on the parameters received from the server and analyze the efficacy of the current models.
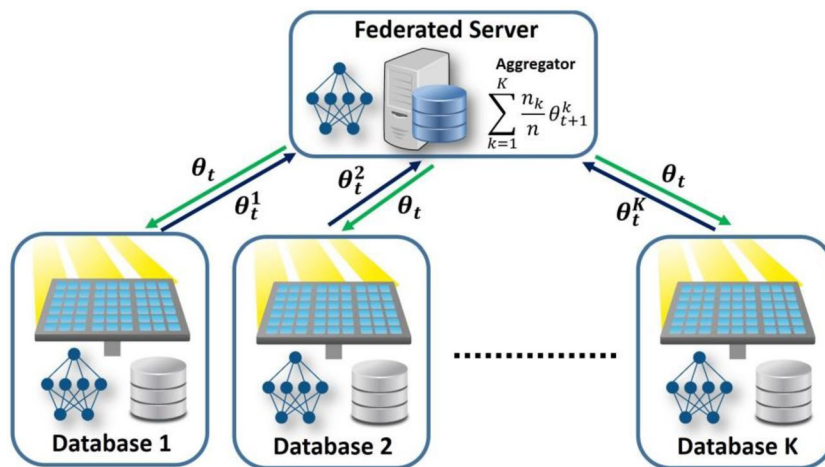

Fig. 2 Architecture of FL model

Fig. 3 Implementation flow-chart of the solar radiation forecasting process based on the developed FL method
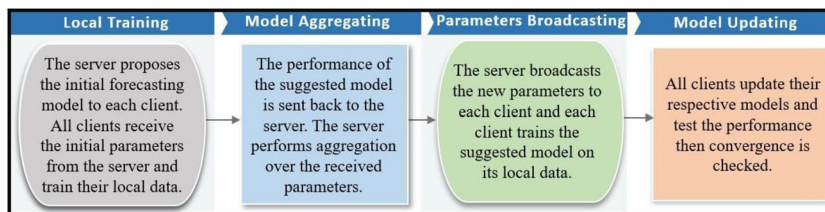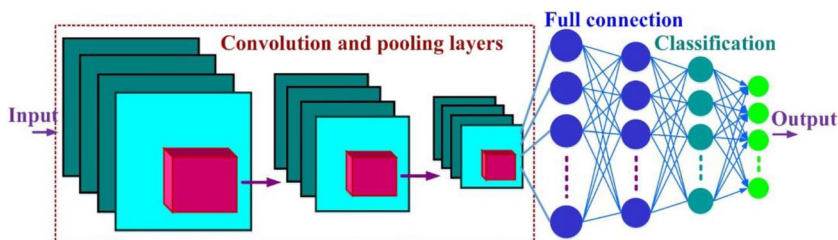

Fig. 4 General architecture of CNN

Convolutional neural network (CNN)

The CNN is a subset of powerful deep learning techniques that have been used in recent years mainly in scientific and industrial applications associated with energy and power systems. Feature extraction, image processing, pattern recognition, categorization, and forecasting are some of the obvious capabilities of this scheme (Moradzadeh et al. 2021d). As Fig. 4 shows, the CNN structure consists of convolution layers, pooling layers, fully connected layers, and finally classification layers. The deep and layer-by-layer structure of CNN makes it possible to extract local features in the lower layers and then combine them into more abstract features in the higher layers (Moradzadeh et al. 2021b).

The convolution layer consists of a trainable bias and a set of learnable kernels that act as a filter to find the main characteristics from the input data. Each kernel, in proportion to its size, extracts the features from the data to eventually form a feature map in each convolution layer. The output of neurons that are connected to the input volume is obtained by calculating the product of the point between their weight and the small area. In each convolution layer, a nonlinear activation function called rectified linear unit (ReLU) is employed as $(0, x)$ (Moradzadeh et al. 2021d). The l-layer convolved feature maps in a CNN structure with L-layers can be calculated as follows (Moradzadeh et al. 2021b):

$$y_{l,j}^{conv} = \sum_{i}^{k} w_{i,j}^{l} \, y_{l-1,i}^{pool} + b_{j}^{l} \tag{3}$$

$$y_{l,j}^{ReLU} = f\left(y_{l,j}^{conv}\right) = \max\left[0, y_{l-1,j}^{conv}\right] \tag{4}$$

where $y_{l,j}^{conv}$ and $w_{i,j}^{l}$ are the output and convolutional kernel related to the lth layer, respectively. k shows the number of kernels and b_{j}^{l} represents the bias. $f(\cdot)$ is the activation function.

After feature extraction by different kernels, a pooling layer is added to each convolution layer to perform the pooling operation to collect the features and generate the feature map. Pooling is a type of nonlinear down-sampling that can reduce each map size and achieve spatial invariance. Pooling strategy can be done in two ways, average and max pooling. The max pooling is mainly used due to advantages such as leading to faster convergence, selection of superior invariant features, and improved generalization. The max pooling operation is obtained as follows (Moradzadeh et al. 2021d):

$$y_{l,j}^{pool} = \max_{M \times N}\left(p\left(S_1, S_2\right) y_{l-1,j}^{ReLU}\right) \tag{5}$$

where $p(S_1, S_2)$ shows the window function used for the stored data and S_1 and S_2 depict the window size, which can

be of arbitrary size and overlapping. $M \times N$ corresponds to the size related to the l-layer feature map output.

After the last convolution layer in the CNN structure, there are fully connected layers that act like a feed-forward neural network. Output of the last convolution layer, the last feature map formed, is applied as the input of fully connected layers. At this stage, the extracted features, after determining the weight and bias, are transferred to the last layer to estimate the final output of the CNN network. The last layer is a Softmax layer that forecasts the final output of the network based on the specified targets (Moradzadeh et al. 2021b).

Decision tree (DT)

The decision tree is one of the most commonly applied techniques of machine learning, which is based on inductive learning and is introduced in the form of a systematic method called recursive binary partition. Linear regression, pattern recognition, classification, and forecasting are obvious applications of this method (Moradzadeh et al. 2021e; Latif 2021). Like the model shown in Fig. 5, tree-shaped diagrams form the composition of decision trees. This architecture consists of branches and three types of nodes named root node, inner node, and leaf node (Moradzadeh et al. 2021e). In the implementation of the training process, the input data is divided into small categories or a number of subsets with the help of some dichotomous classifications. Nodes, which are points in a tree, are responsible for processing features. Branches are also test results and are responsible for forming subsequent nodes in this architecture. The root node is the highest node, the inner nodes are in the middle of the DT structure, and the leaf nodes are considered the end nodes. Reaching a node to a predetermined category purity level and achieving only a single output per node completes the training and test process of the DT (Namazkhan et al. 2020).

Back-propagation neural network (BPNN)

BPNN is one of the most well-known applications of ANNs, first introduced in 1972 by Parker. This algorithm is structurally and functionally slightly different from a feed-forward neural network (FFNN) and has significant applications in
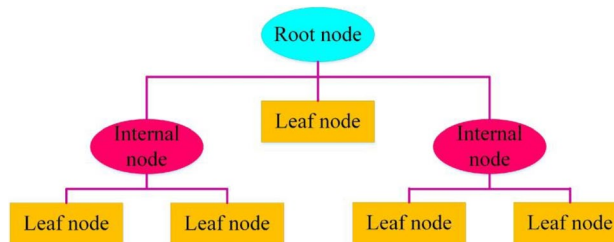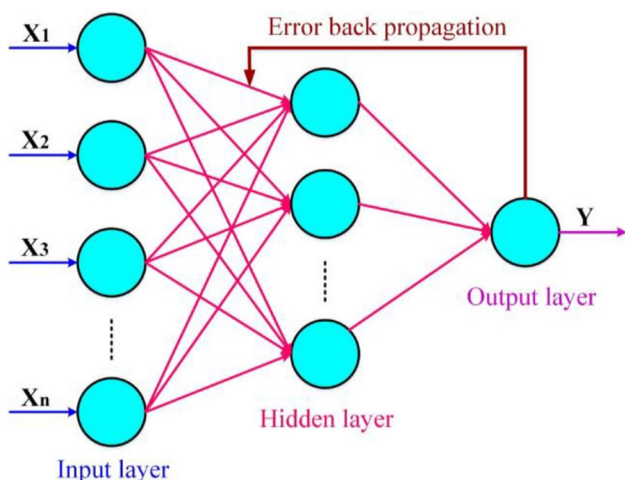

Fig. 5 The architecture of DT

Fig. 6 The architecture of BPNN

Fig. 7 Structural schematic of an LSTM unit

various sciences and industries (Yokoyama et al. 2009). Like the structure presented in Fig. 6, the BPNN consists of an input layer, a hidden layer, and an output layer in its structure which the existence of a back propagation has caused the structural difference between this network and FFNNs (Moradzadeh et al. 2022a).

In this structure, the letters m, T, and n represent the number of nodes in the input layer, hidden layer, and output layer, respectively. w_{ij} and w_{jk} denote the connection weight in the layers, and X and Y represent the input values and the forecasting values, respectively. The input data is received as a signal by the input layer and after processing is transferred to the hidden layer. The hidden layer, as an internal processing layer, has the task of processing and sending signals to the output layer. The training process of ANN-based algorithms requires an activation function, which in this paper, $f(x)$ is used as a sigmoid activation function (Chaudhuri et al. 2019):

$$f(x) = \frac{1}{(1 + e^{-x})} \tag{6}$$

In the input layer, the output related to each neuron is computed as follows (Moradzadeh et al. 2022a):

$$O_{ik} = f\left(\sum_{j=0}^{l} W_{kj}O_{ij}\right), k = 1, 2, \ldots, n \tag{7}$$

where the connection weight from the jth input neuron to kth hidden neuron is denoted by W_{kj}. Finally, the network output for n hidden neurons is forecasted according to the following equation (Moradzadeh et al. 2022a):

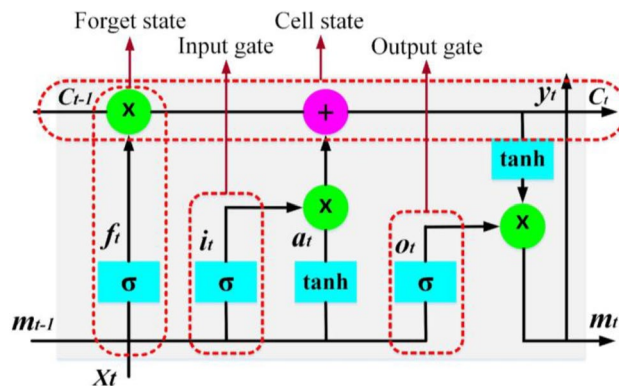$$Y_i = f\left(\sum_{k=0}^{n} W_k O_{ik}\right), k = 1, 2, \ldots, n \tag{8}$$

Long short-term memory (LSTM)

LSTM network is one of the deep learning techniques that has been proposed to mitigate the vanishing/exploding gradient problems related to RNN networks. The LSTM uses a memory unit to store dependent information in consecutive data for a long horizon time (Moradzadeh et al. 2020; Fu et al. 2021). Nonlinear data analysis, regression, estimation of the relationship between input and output variables, and classification are the salient applications of LSTM. Figure 7 shows the structure of an LSTM unit, which consists of a forget gate, input gate and output gate, and a memory cell (Moradzadeh et al. 2022d).

The presence of an intrinsic memory function in this structure allows the network to preserve pre-extracted information. The cell state update in each LSTM unit is done based on the results of the forget gate and input gate. The training process of this network is mainly done using algorithms such as gradient descent and back-propagation. Mathematical modeling of LSTM architecture is as follows (Moradzadeh et al. 2022c):

$$f_t = \sigma\left(W_{lf}l_t + W_{mf}m_{t-1} + b_f\right) \tag{9}$$

$$i_t = \sigma\left(W_{li}l_t + W_{mi}m_{t-1} + b_i\right) \tag{10}$$

$$o_t = \sigma\left(W_{lo}l_t + W_{mo}m_{t-1} + b_o\right) \tag{11}$$

$$a_t = \tanh\left(W_{la}l_t + W_{ma}m_{t-1} + b_a\right) \tag{12}$$

$$c_t = c_{t-1}\phi f_t + i_t\phi a_t \tag{13}$$

$$m_t = o_t\phi\tanh c_t \tag{14}$$

where σ is the activation function called logistic sigmoid. Each of the forget, input, and output gates are denoted by

f_t, i_t, and o_t, respectively. c_t shows the memory cell and a_t demonstrates the hidden vector. $W_{l*} = \{W_{lf}, W_{li}, W_{la}, W_{lo}\}$ and $W_{m*} = \{W_{mf}, W_{mi}, W_{ma}, W_{mo}\}$ represent the weights related to corresponding gates. b_f, b_i, b_o, and b_a are output biases. ϕ is the Hadamard product operator.

Bidirectional long short-term memory (Bi-LSTM)

As an advanced type of LSTM, the Bi-LSTM is developed to model time-series state of data and process large volumes of data. Like the architecture presented for the Bi-LSTM network in Fig. 8, the two-way training process, which includes forward and backward layers, distinguishes the Bi-LSTM architecture from LSTM (Moradzadeh et al. 2021c). Thus, the transfer of information during the training process of this network is done as a two-way process forward (blue line) and reverse (purple line) that can use the past and current information associated with the data to train the next data. The two-direction calculation of the hidden layer and the output layer of Bi-LSTM at time t is as follows (Das et al. 2020; Moradzadeh et al. 2022b):

$$\vec{m}_t = \sigma\left(\vec{W}_i x_t + \vec{V}_i \vec{m}_{t-1} + \vec{b}\right) \tag{15}$$

$$\overleftarrow{m}_t = \sigma\left(\overleftarrow{W}_i x_t + \overleftarrow{V}_i m_{t+1} + \overleftarrow{b}\right) \tag{16}$$

$$y_t = \sigma\left(U\left[\vec{m}_t; m_t\right] + c\right) \tag{17}$$

Evaluating the performance of each of the Bi-LSTM and LSTM networks in various studies has emphasized the high capability of Bi-LSTM, citing some of its structural advantages, such as the law of learning forward and backward. Today, the development of power systems and the increasing use of measuring equipment and communication devices

have resulted in the production of large volumes of data that include certain behavioral patterns of the power system over long periods of time. Hence, the Bi-LSTM network, due to its capabilities in processing large volumes of data and modeling the time-series state of data, can be significantly used in power system applications.

Case studies

In this paper, solar radiation forecasting is done for regions of Iran that have solar power plants. As shown in Fig. 9, the regions studied for training and generating supermodel, as well as new regions for testing the generated supermodel, are selected from various areas of Iran with various climatic and geographical features. In each client of the federated network, the CNN is trained using data from regions of Birjand, Seydabad, Abhar, Mahan, Eqlid, Khaf, Meybod, and Tabriz and extracts the pattern of solar radiation behavior for different regions. Then, based on the extracted pattern, it forms the global supermodel. Data for regions Abadeh, Jarqavieh, and Arak that did not exist in each client training dataset are utilized as anonymous data to test the global supermodel. In each region, data collected by meteorological stations consisting of humidity, temperature, wind speed and direction, and solar radiation with the sampling intervals of 10 min were used.

Solar radiation forecasting results

The solar radiation forecasting in this study is done based on the FL and CNN schemes. It is a technique that, while preserving the privacy of the data, by training with dataset related to 8 scattered regions of Iran, provides a global supermodel that can respond to forecast the solar radiation
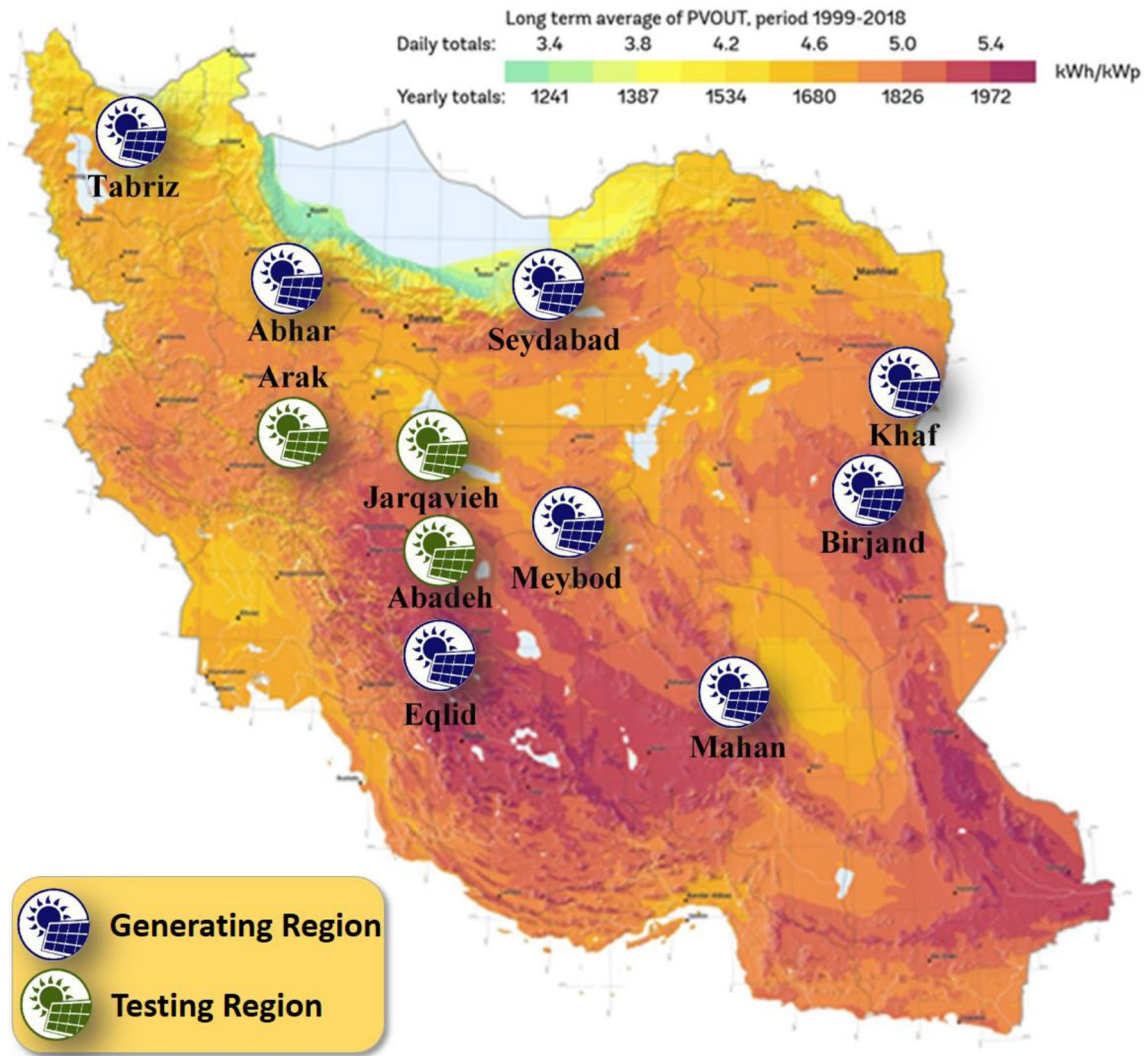
Fig. 8 The structure of Bi-LSTM
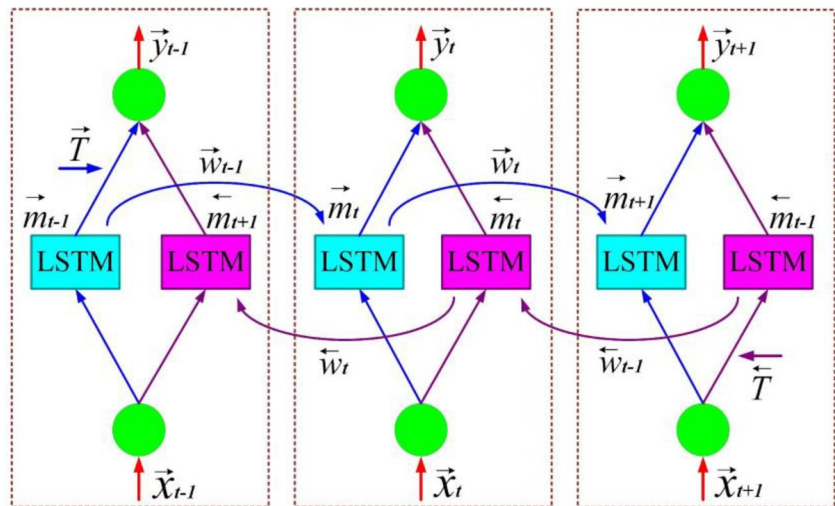
Fig. 9 Locations of considered meteorological stations within Iran's solar radiation map

for all regions of Iran. A CNN procedure has been employed to train the server on each client. To build the FL algorithm, the authors did not use any prepared library based on FL and constructed the model to drive the server, the aggregation process, and broadcasting models to the clients. The CNN was developed employing the TensorFlow deep learning

Table 1 The total stored data was dedicated to training and testing from each region

	Location	No. of samples	No. of training	No. of tests	Start	End
1	Birjand	13,392	10,714	2678	22/06/2007	22/09/2007
2	Seydabad	13,392	10,714	2678	22/06/2007	22/09/2007
3	Abhar	13,324	10,660	2664	07/07/2015	07/10/2015
4	Mahan	13,392	10,714	2678	01/07/2015	01/10/2015
5	Eqlid	13,073	10,459	2614	24/06/2006	22/09/2006
6	Khaf	13,104	10,484	2620	09/07/2007	07/10/2007
7	Meybod	13,392	10,714	2678	22/06/2015	22/09/2015
8	Tabriz	4649	3719	930	22/06/2011	11/09/2011
9	Abadeh	5721	–	5721	22/06/2006	31/07/2006
10	Jarqavieh	5691	–	5691	24/06/2006	02/08/2006
11	Arak	8281	–	8281	07/05/2006	31/08/2006

Table 2 Performance of server for test data of 8 clients

Region	MSE	RMSE	MAE	MAPE	R^2
Birjand	0.0016	0.040	0.018	1.33	0.98
Seydabad	0.0047	0.069	0.036	2.54	0.95
Abhar	0.0043	0.066	0.038	2.92	0.96
Mahan	0.0067	0.082	0.039	2.75	0.94
Eqlid	0.0047	0.069	0.041	3.33	0.96
Khaf	0.0022	0.047	0.020	1.30	0.98
Meybod	0.0019	0.044	0.020	1.36	0.98
Tabriz	0.0079	0.089	0.042	3.11	0.94

framework and trained to operate the Adam optimization algorithm and the mean squared error loss function. The CNN network consists of 2 convolution layers with 64 and 128 filters and kernel size 3. Three dense layers complete the network. These layers are equipped with activation functions of ReLU. The data for Birjand, Seydabad, Abhar, Mahan, Eqlid, Khaf, Meybod, and Tabriz regions in each client is used to train and test the CNN as an input dataset. In each dataset, time variables, such as date and hour of the day, and climatic factors including wind speed and direction, humidity, and temperature with the sampling intervals of 10 min are considered input variables and the solar radiation parameters as output or target variables. Eighty percent of each dataset is considered for network training and recognition of solar radiation behavior patterns and 20% of

the data is selected to test the CNN in each client. Following the completion of the network training and validation process, the federated network will generate a global supermodel based on the features extracted by the CNN. This supermodel, which is also the main objective of this study, is able to forecast the solar radiation for new regions from anywhere in Iran. Accordingly, to test the global supermodel and illustrate its effectiveness, data from new regions of Iran named Abadeh, Jarqavieh, and Arak are utilized to forecast solar radiation. Table 1 shows the number of samples in each region and the number of samples intended for training and testing the CNN in each client and testing the global supermodel generated.

At each stage of training and test, the CNN results are evaluated for each client based on a variety of performance evaluation indicators. These indicators include mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and correlation coefficient (R^2) each of which expresses a specific concept of network performance and forecasts of solar radiation. These metrics were also employed to appraise the performance of the supermodel. Each of the above indicators is computed based on the following equations (Moradzadeh et al. 2022c):

$$MSE = \frac{1}{N} \sum_{i=1}^{N} \left(X_i - Y_i \right)^2 \tag{18}$$
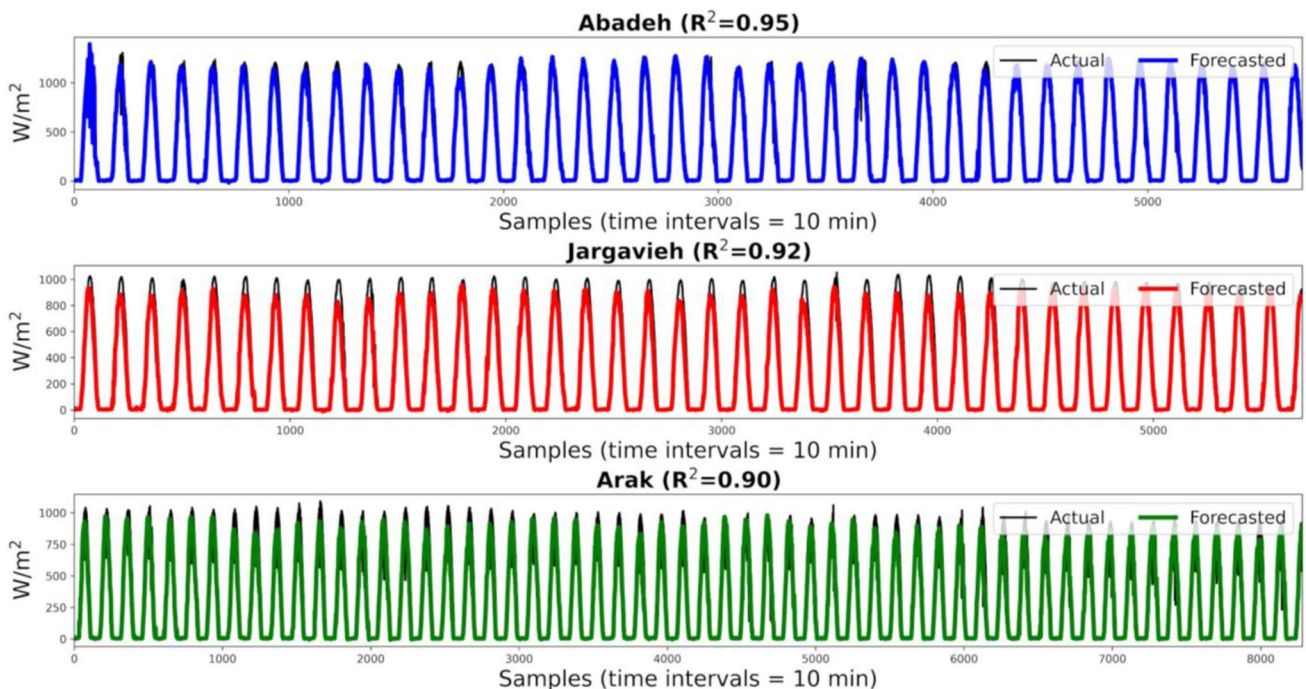
Fig. 10 Comparison of actual and forecasted solar radiation values by the supermodel for the three new regions based on the R^2 evaluation index
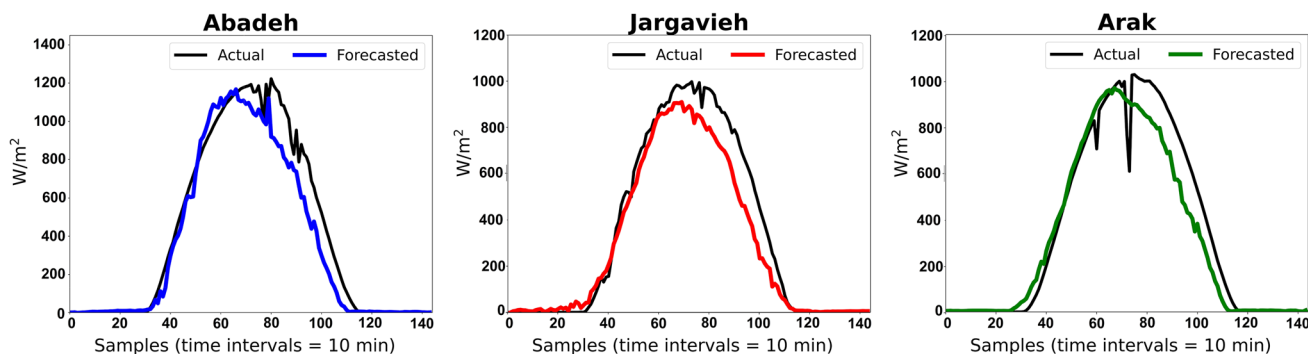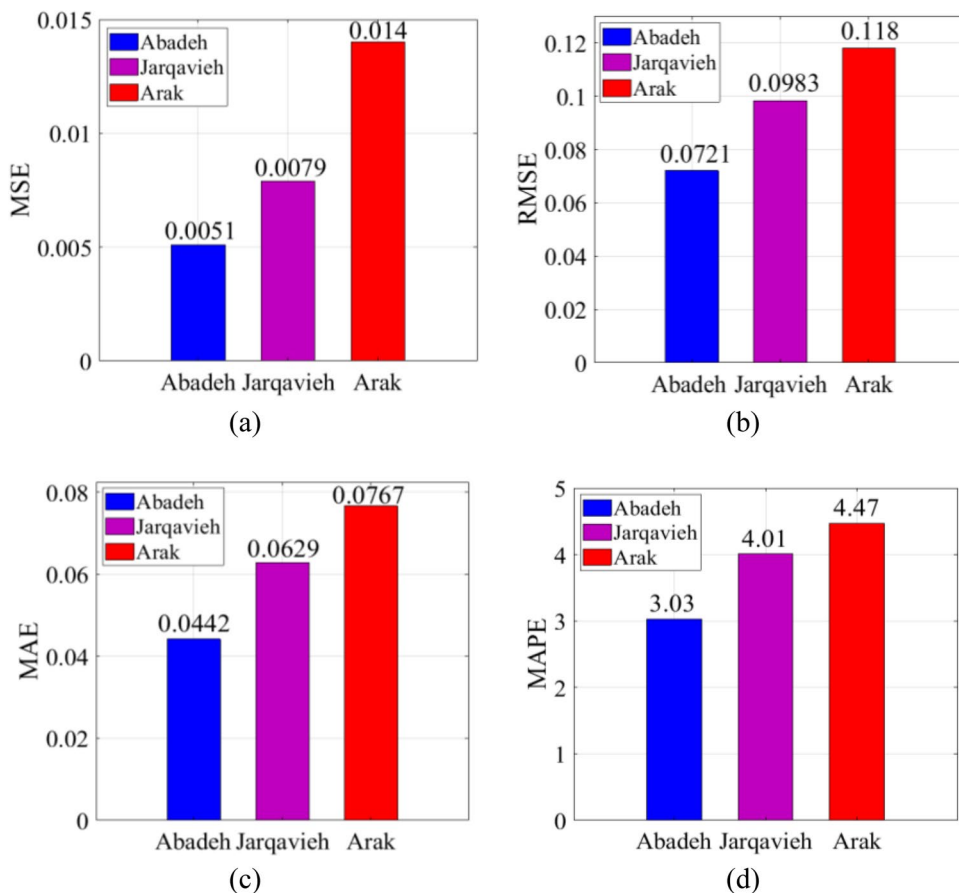
Fig. 11 Examples of 1-day (24 h) solar radiation forecasts from each new region by the global supermodel

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(X_i - Y_i\right)^2} \qquad (19)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left|\frac{X_i - Y_i}{X_i}\right| \times 100 \qquad (21)$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left|\left(X_i - Y_i\right)\right| \qquad (20)$$

$$R^2 = \frac{\sum_{i=1}^{N} \left(X_i - \underline{X}\right)\left(Y_i - \underline{Y}\right)}{\sqrt{\sum_{i=1}^{N} \left(X_i - \underline{X}\right)^2 \sum_{i=1}^{N} \left(Y_i - \underline{Y}\right)^2}} \qquad (22)$$

Fig. 12 Performance evaluation of global supermodel in forecasting solar radiation for new regions: **a** MSE, **b** RMSE, **c** MAE, **d** MAPE

where X_i and Y_i are the real solar radiation parameters and forecasted solar radiation parameters, respectively. $\underline{X}$ and $\underline{Y}$ denote the average of real solar radiation parameters and average of forecasted solar radiation parameters, respectively.

Each of the mentioned statistical indicators evaluates the performance of the desired network with a specific concept. In general, indicators MSE, RMSE, MAE, and MAPE are related to forecast error, and in these indicators, values close to zero are ideal. The R^2 index indicates the accuracy of the network and it is better to have its values close to 1. Each CNN network is trained in each client based on data from the intended regions. At the end of the training process, each network is saved as a toolbox to be used to forecast the test data. As mentioned, at this stage, each trained CNN network is tested by 20% of the data from the 8 regions mentioned. Table 2 presents the results of solar radiation forecasting for the regions intended for supermodel generation in the test stage according to the mentioned evaluation indicators. The high accuracy and excellent performance of the CNN model in training and the test process of data related to each region can be seen from the results presented in Table 2.

The results show that the CNN in each client has been able to train the behavioral pattern of solar radiation well and to forecast its values based on input variables from each region with high correlation. Based on this training and test process, the federated network generates a global supermodel that contains features extracted in each client. The supermodel produced should be able to forecast solar radiation for regions that have not been included in the training process. To do this, data from regions of Abadeh, Jarqavieh, and Arak are fully utilized (as shown in Table 1) to test the generated global supermodel. Figure 10 shows the real values of solar radiation and related results of forecasting solar radiation using the supermodel generated for the three new regions. In this figure, the correlation between the actual parameters and the forecasted parameters of solar radiation based on the R^2 index is presented. In order to express the effectiveness of the global supermodel, predictive samples for 1 day (24 h) are randomly provided from each region with high resolution in Fig. 11. It can be seen that although no data from these three regions is used in the process of server training in any client, the supermodel is able to provide an acceptable forecast of solar radiation in these regions, namely, Abadeh, Jarqavieh, and Arak, with R^2 values of 0.95, 0.92, and 0.90, respectively.

Fig. 12 evaluates the forecasting results presented by the global supermodel based on statistical indices MSE, RMSE, MAE, and MAPE, respectively. The evaluation of the results in Fig. 12 has been performed based on various statistical indicators and it can be seen that the supermodel is able to perform the prediction process for each of the new regions with low and acceptable error values. Among the new

Table 3 Analyzing and comparing the results of methods DT, BPNN, LSTM, and Bi-LSTM in solar radiation forecasting

Region	DT					BPNN					LSTM					Bi-LSTM				
	MSE	RMSE	MAE	MAPE	R^2	MSE	RMSE	MAE	MAPE	R^2	MSE	RMSE	MAE	MAPE	R^2	MSE	RMSE	MAE	MAPE	R^2
Birjand	0.0027	0.051	0.024	1.47	0.97	0.0033	0.057	0.029	1.52	0.97	0.0019	0.043	0.020	1.39	0.97	0.0013	0.036	0.017	1.31	0.98
Seydabad	0.0051	0.071	0.039	2.60	0.94	0.0054	0.073	0.042	2.65	0.93	0.0046	0.067	0.034	2.52	0.95	0.0043	0.065	0.034	2.51	0.96
Abhar	0.0054	0.073	0.043	2.92	0.94	0.0051	0.071	0.039	2.58	0.94	0.0046	0.067	0.042	2.97	0.95	0.0044	0.066	0.040	2.96	0.96
Mahan	0.0070	0.083	0.044	2.93	0.93	0.0072	0.084	0.045	2.90	0.94	0.0075	0.086	0.047	2.84	0.93	0.0060	0.077	0.031	2.72	0.95
Eqlid	0.0063	0.079	0.058	3.40	0.93	0.0065	0.080	0.060	3.39	0.93	0.0058	0.076	0.053	3.49	0.94	0.0046	0.067	0.040	3.29	0.96
Khaf	0.0037	0.060	0.034	1.54	0.97	0.0044	0.066	0.039	1.61	0.96	0.0031	0.055	0.029	1.47	0.96	0.0020	0.044	0.019	1.24	0.98
Meybod	0.0040	0.063	0.039	1.61	0.97	0.0042	0.064	0.043	1.67	0.97	0.0027	0.051	0.027	1.43	0.98	0.0024	0.048	0.023	1.41	0.97
Abadeh	Failed					Failed					Failed					Failed				
Jarqavieh	Failed					Failed					Failed					Failed				
Arak	Failed					Failed					Failed					Failed				

regions, the forecasts made for the Abadeh region, which also has a higher solar energy potential (as shown in Fig. 9) than two other regions, have the highest accuracy score and the lowest forecasting error values. The supermodel is also able to forecast the solar radiation for the region of Jarqavieh with a higher accuracy coefficient than the Arak. Figure 9 also shows that the region of Jarqavieh has a higher solar energy potential than the Arak region. The results show that the proposed approach can be used as a powerful tool for power grid and renewable energy operators where privacy is a high priority.

As stated in the literature, various conventional methods such as DT, BPNN, LSTM, and Bi-LSTM were also employed to forecast solar radiation in the study regions to make a comparative assessment of the performance of the developed FL method. Table 3 evaluates and compares the results of DT, BPNN, LSTM, and Bi-LSTM methods in solar radiation forecasting.

The presented results in Table 3 show the unique ability of each method to forecast solar radiation in different regions. It can be seen that the Bi-LSTM method was superior to DT, BPNN, and LSTM models and had the least forecasting error values compared to other methods. However, the lack of generalizability of conventional methods has made none of the DT, BPNN, LSTM, and Bi-LSTM models able to provide forecasts in regions Abadeh, Jarqavieh, and Arak. Meanwhile, the developed FL method has been able to forecast the solar radiation in regions Abadeh, Jarqavieh, and Arak, with the accuracy of 95%, 92%, and 90%, respectively, under the same data distribution conditions as conventional methods. Accordingly, due to the fact that meteorological stations are not available in all parts of the country, the proposed method can be used effectively for industrial projects that are in the direction of operation of RESs, especially in remote areas. In addition, the FL scheme can be employed for all forecasting problems in power and energy systems.

Conclusion

Optimal operation and management of renewable energy sources (RESs), especially solar energy, need a comprehensive database to construct precise models for forecasting the behavior of solar radiation in diverse regions and time spans. Because environmental data and solar radiation are based on in-site and satellite measurements, the problem of lack of meteorological stations in most places and data privacy in most industrial projects is one of the most important challenges of using RESs. Providing a reliable forecasting model of solar radiation that can also protect data privacy can play a key role in allowing high PV penetration without compromising network performance. In this paper, after reviewing the various aspects of solar radiation forecasting and

the methods presented in various studies, a global model of solar radiation forecasting with data privacy was presented. The proposed method is one of the newest approaches to machine learning called federated learning (FL). Independence of the required data in each region and high ability to protect existing data against cyber-attacks and security issues are the most prominent features of the proposed FL procedure. To achieve a global prediction supermodel, the proposed method was trained based on the data of eight scattered regions of Iran with different geographical and meteorological features. The training and test stage for each client in the federated network was performed based on data from each region by one of the deep learning models named convolutional neural network (CNN). The CNN extracted the behavioral pattern of solar radiation related to each region in each client and produced a global supermodel based on the recognized patterns. In order to test and evaluate the performance of the global supermodel produced, which contains the behavioral pattern of solar radiation in various conditions, data related to three new regions of Iran named Abadeh, Jarqavieh, and Arak were used as input dataset for the supermodel. These three regions had no role in the training and generation of the supermodel and were fully used to test the supermodel and express its effectiveness. The presented results were analyzed with various performance evaluation indicators. The result showed that the developed forecasting model was able to forecast solar radiation with a correlation coefficient of 95%, 92%, and 90% for regions Abadeh, Jarqavieh, and Arak, respectively. Finally, in order to evaluate the performance of the developed FL model, four conventional models such as decision tree (DT), back-propagation neural network (BPNN), long short-term memory (LSTM), and bidirectional LSTM (Bi-LSTM) were employed to forecast solar radiation in each of the study regions. It was detected that since no training data were accessible from regions of Abadeh, Jarqavieh, and Arak, the conventional DT, BPNN, LSTM, and Bi-LSTM methods cannot forecast solar radiation in these regions. However, the proposed FL method under the same conditions can present an acceptable solar radiation forecasting of these regions. It should be noted that the developed FL model can be employed as a robust and powerful tool for solving forecasting problems in industrial projects and power or energy systems where data protection is of utmost importance.

Author contributions Arash Moradzadeh: software, writing—original draft, visualization. Hamed Moayyed: formal analysis, investigation, writing—original draft, validation. Behnam Mohammadi-Ivatloo: conceptualization, methodology, writing—review and editing. António Pedro Aguiar: writing—review and editing, supervision, validation. Amjad Anvari-Moghaddam: writing—review and editing, investigation, validation. Zulkurnain Abdul-Malek: writing—review and editing, supervision, validation. All authors contributed to the study conception and design. Material preparation, data collection, and analysis

Data Availability All of the used data in this manuscript are available by request from the corresponding author.

Declarations

Ethical approval The authors declare that the manuscript does not include human or animal research. No ethical approval is applicable.

Consent to participate This research does not include human research. All authors agreed with the content and all gave explicit consent to submit the paper.

Consent for publication All authors agreed with the content and all gave explicit consent to publish the paper.

Competing interests The authors declare no competing interests.

References

Abdel-Nasser M, Mahmoud K, Lehtonen M (2021) Reliable solar irradiance forecasting approach based on Choquet integral and deep LSTMs. IEEE Trans Industr Inform 17:1873–1881. https://doi.org/10.1109/TII.2020.2996235

Belmahdi B, Louzazni M, El Bouardi A (2022) Comparative optimization of global solar radiation forecasting using machine learning and time series models. Environ Sci Pollut Res 29:14871–14888. https://doi.org/10.1007/s11356-021-16760-8

Bosman LB, Darling SB (2018) Performance modeling and valuation of snow-covered PV systems: examination of a simplified approach to decrease forecasting error. Environ Sci Pollut Res 25:15484–15491. https://doi.org/10.1007/s11356-018-1748-1

Caldas M, Alonso-Suárez R (2019) Very short-term solar irradiance forecast using all-sky imaging and real-time irradiance measurements. Renew Energy 143:1643–1658. https://doi.org/10.1016/j.renene.2019.05.069

Chaudhuri T, Soh YC, Li H, Xie L (2019) A feedforward neural network based indoor-climate control framework for thermal comfort and energy saving in buildings. Appl Energy 248:44–53. https://doi.org/10.1016/j.apenergy.2019.04.065

Dadashi Z, Mahmoudi A, Rashidi S (2022) Capacity and strategies of energy production from renewable sources in Arab countries until 2030: a review from renewable energy potentials to environmental issues. Environ Sci Pollut Res 29:47837–47866. https://doi.org/10.1007/s11356-022-20544-z

Das A, Annaqeeb MK, Azar E et al (2020) Occupant-centric miscellaneous electric loads prediction in buildings using state-of-the-art deep learning methods. Appl Energy 269:115135. https://doi.org/10.1016/j.apenergy.2020.115135

Deo RC, Şahin M, Adamowski JF, Mi J (2019) Universally deployable extreme learning machines integrated with remotely sensed MODIS satellite predictors over Australia to forecast global solar radiation: a new approach. Renew Sustain Energy Rev 104:235–261. https://doi.org/10.1016/j.rser.2019.01.009

Feng C, Cui M, Hodge B-M et al (2019) Unsupervised clustering-based short-term solar forecasting. IEEE Trans Sustain Energy 10:2174–2185. https://doi.org/10.1109/TSTE.2018.2881531

Fu M, Le C, Fan T et al (2021) Integration of complete ensemble empirical mode decomposition with deep long short-term memory model for particulate matter concentration prediction. Environ Sci Pollut Res 28:64818–64829. https://doi.org/10.1007/s11356-021-15574-y

International Energy Agency (IEA) (2019) Renewables information 2019 – analysis - IEA. IEA static reports

International Renewable Energy Agency (2020) https://irena.org/solar. https://irena.org/solar

Kakimoto M, Endoh Y, Shin H et al (2019) Probabilistic solar irradiance forecasting by conditioning joint probability method and its application to electric power trading. IEEE Trans Sustain Energy 10:983–993. https://doi.org/10.1109/TSTE.2018.2858777

Kapur A, Bhatia S, Goyal A (2011) Talon cusp in a primary incisor: a rare entity. J Indian Soc Pedod Prev Dent 29:248. https://doi.org/10.4103/0970-4388.85835

Khan AN, Nadeem MA, Hussain MS et al (2020) A forecasting model approach of sustainable electricity management by developing adaptive neuro-fuzzy inference system. Environ Sci Pollut Res 27:17607–17618. https://doi.org/10.1007/s11356-019-06626-5

Khodayar M, Mohammadi S, Khodayar ME et al (2020) Convolutional graph autoencoder: a generative deep neural network for probabilistic spatio-temporal solar irradiance forecasting. IEEE Trans Sustain Energy 11:571–583. https://doi.org/10.1109/TSTE.2019.2897688

Konečný J, McMahan HB, Ramage D, Richtárik P (2016) Federated optimization: distributed machine learning for on-device intelligence

Latif SD (2021) Developing a boosted decision tree regression prediction model as a sustainable tool for compressive strength of environmentally friendly concrete. Environ Sci Pollut Res 28:65935–65944. https://doi.org/10.1007/s11356-021-15662-z

Liu B, Song C, Wang Q, Wang Y (2022) Forecasting of China's solar PV industry installed capacity and analyzing of employment effect: based on GRA-BiLSTM model. Environ Sci Pollut Res 29:4557–4573. https://doi.org/10.1007/s11356-021-15957-1

Louzazni M, Mosalam H, Khouya A (2020) A non-linear auto-regressive exogenous method to forecast the photovoltaic power output. Sustain Energy Technol Assess 38:100670. https://doi.org/10.1016/j.seta.2020.100670

Lu Y, Huang X, Zhang K et al (2021) Low-latency federated learning and blockchain for edge association in digital twin empowered 6G networks. IEEE Trans Industr Inform 17:5098–5107. https://doi.org/10.1109/TII.2020.3017668

Marzouq M, El Fadili H, Zenkouar K et al (2020) Short term solar irradiance forecasting via a novel evolutionary multi-model framework and performance assessment for sites with no solar

irradiance data. Renew Energy 157:214–231. https://doi.org/10.1016/j.renene.2020.04.133

Moayyed H, Moradzadeh A, Mohammadi-Ivatloo B et al (2022) A cyber-secure generalized supermodel for wind power forecasting based on deep federated learning and image processing. Energ Conver Manage 267:115852. https://doi.org/10.1016/j.enconman.2022.115852

Moradzadeh A, Moayyed H, Mohammadi-Ivatloo B et al (2021a) A secure federated deep learning-based approach for heating load demand forecasting in building environment. IEEE Access 10:5037–5050. https://doi.org/10.1109/access.2021.3139529

Moradzadeh A, Moayyed H, Mohammadi-Ivatloo B et al (2021b) Turn-to-turn short circuit fault localization in transformer winding via image processing and deep learning method. IEEE Trans Industr Inform 18:4417–4426. https://doi.org/10.1109/TII.2021.3105932

Moradzadeh A, Moayyed H, Zakeri S et al (2021c) Deep learning-assisted short-term load forecasting for sustainable management of energy in microgrid. Inventions 6:15. https://doi.org/10.3390/inventions6010015

Moradzadeh A, Mohammadi-Ivatloo B, Abapour M et al (2021d) A practical solution based on convolutional neural network for non-intrusive load monitoring. J Ambient Intell Humaniz Comput 12:9775–9789. https://doi.org/10.1007/s12652-020-02720-6

Moradzadeh A, Mohammadi-Ivatloo B, Abapour M et al (2022a) Heating and cooling loads forecasting for residential buildings based on hybrid machine learning applications: a comprehensive review and comparative analysis. IEEE Access 10:2196–2215. https://doi.org/10.1109/ACCESS.2021.3136091

Moradzadeh A, Mohammadi-ivatloo B, Pourhossein K, Anvari-Moghaddam A (2021e) Data mining applications to fault diagnosis in power electronic systems: a systematic review. IEEE Trans Power Electron 37(5):6026–6050. https://doi.org/10.1109/tpel.2021.3131293

Moradzadeh A, Mohammadpourfard M, Genc I et al (2022b) Deep learning-based cyber resilient dynamic line rating forecasting. Int J Electr Power Energy Syst 142:108257. https://doi.org/10.1016/j.ijepes.2022.108257

Moradzadeh A, Mohammadpourfard M, Konstantinou C et al (2022c) Electric load forecasting under false data injection attacks using deep learning. Energy Rep 8:9933–9945. https://doi.org/10.1016/j.egyr.2022.08.004

Moradzadeh A, Teimourzadeh H, Mohammadi-Ivatloo B, Pourhossein K (2022d) Hybrid CNN-LSTM approaches for identification of type and locations of transmission line faults. Int J Electr Power Energy Syst 135:107563. https://doi.org/10.1016/j.ijepes.2021.107563

Moradzadeh A, Zakeri S, Shoaran M et al (2020) Short-term load forecasting of microgrid via hybrid support vector regression and long short-term memory algorithms. Sustainability (Switzerland) 12:7076. https://doi.org/10.3390/su12177076

Namazkhan M, Albers C, Steg L (2020) A decision tree method for explaining household gas consumption: the role of building characteristics, socio-demographic variables, psychological factors and household behaviour. Renew Sustain Energy Rev 119:109542. https://doi.org/10.1016/j.rser.2019.109542

Nasirpour MH, Sharifi A, Ahmadi M, Jafarzadeh Ghoushchi S (2021) Revealing the relationship between solar activity and COVID-19 and forecasting of possible future viruses using multi-step autoregression (MSAR). Environ Sci Pollut Res 28:38074–38084. https://doi.org/10.1007/s11356-021-13249-2

Prasad R, Ali M, Kwan P, Khan H (2019) Designing a multi-stage multivariate empirical mode decomposition coupled with ant colony optimization and random forest model to forecast monthly solar radiation. Appl Energy 236:778–792. https://doi.org/10.1016/j.apenergy.2018.12.034

Rezazadeh S, Moradzadeh A, Pourhossein K et al (2022) Photovoltaic array reconfiguration under partial shading conditions for maximum power extraction: a state-of-the-art review and new solution method. Energ Conver Manage 258:115468. https://doi.org/10.1016/j.enconman.2022.115468

Riahi S, Abedini E, Vakili M, Riahi M (2021) Providing an accurate global model for monthly solar radiation forecasting using artificial intelligence based on air quality index and meteorological data of different cities worldwide. Environ Sci Pollut Res 28:49697–49724. https://doi.org/10.1007/s11356-021-14126-8

Scolari E, Sossan F, Paolone M (2018) Photovoltaic-model-based solar irradiance estimators: performance comparison and application to maximum power forecasting. IEEE Trans Sustain Energy 9:35–44. https://doi.org/10.1109/TSTE.2017.2714690

Sharda S, Singh M, Sharma K (2021) RSAM: robust self-attention based multi-horizon model for solar irradiance forecasting. IEEE Trans Sustain Energy 12:1394–1405. https://doi.org/10.1109/TSTE.2020.3046098

Van Der Meer D, Mouli GRC, Mouli GME et al (2018) Energy management system with PV power forecast to optimally charge EVs at the workplace. IEEE Trans Industr Inform 14:311–320. https://doi.org/10.1109/TII.2016.2634624

Wang H, Cai R, Zhou B et al (2020a) Solar irradiance forecasting based on direct explainable neural network. Energ Conver Manage 226:113487. https://doi.org/10.1016/j.enconman.2020.113487

Wang H, Liu Y, Zhou B et al (2020b) Taxonomy research of artificial intelligence for deterministic solar power forecasting. Energ Conver Manage 214:112909. https://doi.org/10.1016/j.enconman.2020.112909

Wang J, Zhong H, Lai X et al (2019a) Exploring key weather factors from analytical modeling toward improved solar power forecasting. IEEE Trans Smart Grid 10:1417–1427. https://doi.org/10.1109/TSG.2017.2766022

Wang Y, Bennani IL, Liu X et al (2021) Electricity consumer characteristics identification: a federated learning approach. IEEE Trans Smart Grid 12:3637–3647. https://doi.org/10.1109/TSG.2021.3066577

Wang Y, Shen Y, Mao S et al (2019b) LASSO and LSTM integrated temporal model for short-term solar intensity forecasting. IEEE Internet Things J 6:2933–2944. https://doi.org/10.1109/JIOT.2018.2877510

Wen H, Du Y, Chen X et al (2021) Deep learning based multistep solar forecasting for PV ramp-rate control using sky images. IEEE Trans Industr Inform 17:1397–1406. https://doi.org/10.1109/TII.2020.2987916

Yagli GM, Yang D, Srinivasan D (2019) Automatic hourly solar forecasting using machine learning models. Renew Sustain Energy Rev 105:487–498. https://doi.org/10.1016/j.rser.2019.02.006

Yokoyama R, Wakui T, Satake R (2009) Prediction of energy demands using neural network with model identification by global optimization. Energ Conver Manage 50:319–327. https://doi.org/10.1016/j.enconman.2008.09.017

Zambrano AF, Giraldo LF (2020) Solar irradiance forecasting models without on-site training measurements. Renew Energy 152:557–566. https://doi.org/10.1016/j.renene.2020.01.092

Zhang X, Fang F, Wang J (2020) Probabilistic solar irradiation forecasting based on variational bayesian inference with secure federated learning. IEEE Trans Industr Inform 17(11):7849–7859. https://doi.org/10.1109/TII.2020.3035807

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.