IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 6 June 2024, accepted 12 July 2024, date of publication 16 July 2024, date of current version 29 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3429073

TOPICAL REVIEW

Metaheuristic Algorithms for Solar Radiation Prediction: A Systematic Analysis

SERGIO A. PÉREZ-RODRÍGUEZ[1], JOSÉ M. ÁLVAREZ-ALVARADO[2], JULIO-ALEJANDRO ROMERO-GONZÁLEZ[1], MARCOS AVILES[2], (Member, IEEE), AMERICA EILEEN MENDOZA-ROJAS[3], CARLOS FUENTES-SILVA[3], AND JUVENAL RODRÍGUEZ-RESÉNDIZ[2], (Senior Member, IEEE)

[1]Facultad de Informática, Universidad Autónoma de Querétaro, Querétaro 76230, Mexico
[2]Facultad de Ingeniería, Universidad Autónoma de Querétaro, Querétaro 76010, Mexico
[3]Universidad Politécnica de Querétaro, El Marqués, Querétaro 76240, Mexico

Corresponding author: José M. Álvarez-Alvarado (jmalvarez@uaq.edu.mx)

ABSTRACT In the contemporary world, where the escalating demand for energy and the imperative for sustainable sources, notably solar energy, have taken precedence, the investigation into solar radiation (SR) has become indispensable. Characterized by its intermittency and volatility, SR may experience considerable fluctuations, exerting a significant influence on energy supply security. Consequently, the precise prediction of SR has become imperative, particularly in the context of the potential proliferation of photovoltaic panels and the need for optimized energy management. Several works in the existing literature review the state of the art in SR prediction, focusing on trends identified using machine learning (ML) or deep learning (DL) techniques. However, there is a gap in the literature regarding the integration of optimization algorithms with ML and DL techniques for SR prediction. This systematic review addresses this gap by studying prediction models for SR that leverage metaheuristic optimization algorithms alongside artificial intelligence (AI) techniques, aiming primarily for maximum prediction accuracy. Metaheuristic algorithms such as Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) have featured in 29% and 12.1% of the analyzed articles, respectively, while intelligent approaches like Convolutional Neural Networks (CNN), Extreme Learning Machine (ELM), and Multilayer Perceptron (MLP) have emerged as the predominant choices, collectively accounting for 43.9% of the studies. Analysis has encompassed studies examining SR across hourly, daily, and monthly intervals, with daily intervals representing 48.7% of the focus. Noteworthy variables including temperature, humidity, wind speed, and atmospheric pressure have surfaced, capturing proportions of 90%, 68.2%, 56%, and 41.4%, respectively, within the reviewed literature.

INDEX TERMS Forecasting, metaheuristics, optimization, solar radiation, time-series.

I. INTRODUCTION

Energy have played a crucial role in a growing society, with solar energy having emerged as a promising and sustainable alternative [1]. The increasing dependence on technology and the scarcity of traditional energy resources have driven the search for solutions, underscoring the importance of accurate prediction of global solar radiation (GSR) for efficient solar energy capture [2]. In many countries, especially those in development, the high cost of GSR measurement instruments have hindered obtaining precise data. To address this challenge, various SR prediction models have been developed employing a variety of techniques, such as Machine Learning (ML), Deep Learning (DL), and optimization algorithms, utilizing meteorological data and satellite images. While ML models have demonstrated remarkable performance, it have been crucial to highlight that parametric search have become essential in constructing the underlying structure of any ML model. This had been achieved by incorporating optimization methods that had allowed fine-tuning and refining the model parameters, playing a crucial role in its effectiveness and accuracy [3].

The associate editor coordinating the review of this manuscript and approving it for publication was Bo Pu.

Metaheuristic algorithms have significantly contributed to enhancing various machine learning models, such as random forest (RF) [4] and Support Vector Regression (SVR) [5], by facilitating the attainment of optimal parameters for each model. Similarly, they have positively impacted deep learning models, such as convolutional neural networks (CNN) [6] or multilayer perceptron (MLP) [7], substantially improving their efficacy in SR prediction.

The aim of this article is to conduct an examination of the latest methods utilized in the research and development in the field of SR forecasting, especially through the implementation of hybrid models based on ML, DL, and metaheuristic algorithms. These approaches have been designed to mitigate the computational burden required, constituting a key contribution in terms of efficiency and accuracy [8]. The fundamental purpose of conducting a systematic review have lain in obtaining a broader and more detailed view of the current state of model optimization through the use of metaheuristic algorithms for SR prediction. This analysis have encompassed both ML and DL techniques, in addition to considering the most commonly used meteorological variables in these models. Furthermore, it have examined the prediction time horizon adopted by such models, as well as the evaluation metrics employed to demonstrate their effectiveness. This approach have allowed discerning both the effectiveness of the models and some of the key configurations used in their development. Additionally, it have sought to understand the underlying purpose of the optimization algorithms employed in the various reviewed research works.

According to the existing literature, several works review the state of the art in solar radiation prediction. Some of these works specialize in identifying trends using machine learning or deep learning techniques [9], [10], [11], [12]. Other works analyze specific techniques such as SVM with search algorithms [13]. However, ML and DL techniques for prediction have not been addressed from the perspective of the integration of optimization algorithms. Therefore, the main objective of this article is to study prediction techniques in time series with optimization algorithms.

This article also endeavors to address the following pivotal inquiries:

- Which metaheuristic algorithms and AI models are frequently employed in SR prediction?
- What variables and time horizons are predominantly utilized in SR prediction?
- How do metaheuristic algorithms enhance the accuracy of SR prediction when integrated with AI models?
- What future research directions can be inferred from the current trends in SR prediction methodologies?

The article is organized as follows. Section I provides an introduction to the topic, highlighting its significance and background, while also outlining the intended contributions of this article. Section II offers a concise overview of time-series forecasting, bioinspired optimization algorithms, artificial neural networks (ANN), and performance metrics. In Section III, the methodological framework utilized for conducting the systematic review of the literature is elaborated. Section IV presents a summary of articles focusing on ML or DL prediction of SR with metaheuristic optimization techniques. Section V outlines the key characteristics employed in SR prediction, including commonly used variables, metaheuristic algorithms, AI models, and time horizons. Section VI offers conclusions drawn from this article, suggestions for future research endeavors, clarifications of terminologies, and, finally, a list of bibliographic references.

II. THEORETICAL BACKGROUND

This section has focused on exploring the application of DL and ML models optimized by metaheuristic algorithms for predicting SR using time series data. The research has delved into concepts related to time series analysis, SR prediction, bioinspired optimization algorithms, and DL. Additionally, the study has aimed to examine the interaction between these elements, including the role of bioinspired optimization algorithms in optimizing DL and ML models to achieve accurate predictions of SR. It has also investigated the impact of these algorithms on the performance and reliability of the predictive model.

A. TIME-SERIES FORECASTING

Temporal series methods, a cornerstone of statistical analysis, stand as one of the primary techniques employed in estimating SR. Typically, temporal series models have shown considerable efficacy in SR estimation, using historical SR data to make predictions [14].

A temporal series consists of observations, each point recorded at specific moments. Each temporal series can be identified by important components such as trend, indicating the overall movement upward or downward, cyclically, implying a repetitive pattern without a fixed period, and seasonality, referring to periodic fluctuations in the temporal series [15].

A temporal series is considered stationary when its statistical properties, such as mean and standard deviation, remain constant over time. This aspect is of great importance as stationary suggests a high probability of the temporal series reproducing its behavior in the future. Consequently, predicting a stationary series is considerably facilitated [16].

The field of predicting temporal series encompasses a wide spectrum of methodologies. These range from automated parametric model selection to conventional ML techniques such as kernel regression and support vector regression. Additionally, Gaussian processes have been widely employed in temporal series prediction. Recent advancements in this area include the emergence of deep Gaussian processes and the incorporation of parallelization in ML through neural processes [17].

In addition to metaheuristic optimization, AI techniques are often utilized due to the efficacy of cutting-edge models, such as recurrent neural networks (RNN), including long short-term memory (LSTM) networks, which excel at capturing intricate temporal dependencies [17], [18], [19], [20]. These techniques are valuable in determining the appropriate hyper-parameters and selecting the most suitable frameworks [21], [22], particularly time series Transformers, for tasks such as forecasting, anomaly detection, and classification [23]. Their effectiveness is largely attributed to their ability to capture long-range dependencies and interactions.

1) SOLAR RADIATION PREDICTION

SR is highly intermittent and chaotic; even small fluctuations in solar radiation can significantly impact the security of energy supply. This requires increasingly accurate predictive models, especially given the potential rise in the use of solar energy supplied by photovoltaic panels [24]. Therefore, creating a predictive model for global solar radiation is crucial for ensuring optimal energy dispatch and management practices [6]. According to the state of the art, various models have been developed using machine learning techniques such as SVR [25], [26], [27] and RF [4], [28], as well as deep learning techniques like RNN [29], [30], MLP [6], [31], [32],which have gained significant relevance in recent years [33].

B. BIOINSPIRED OPTIMIZATION ALGORITHMS

Optimization is the process of obtaining optimal values for a set of variables regarding the objective of the function, which can be constrained or unconstrained. Any set of values assigned to the variables always produces an output, but the optimal set produces the optimal output. The problem objectives and constraints can be formulated in terms of functions or mathematical equations. The mathematical expression representing the optimization objective is called the objective function. The objective function may be defined for maximization or minimization depending on the problem domain.

Bioinspired algorithms are metaheuristic algorithms developed based on principles of biological evolution, swarm behavior, and physical and chemical processes. They are a bioinspired computational intelligence technique as they incorporate intelligence into algorithms. Popular bioinspired algorithms used for time series forecasting include the Jaya Algorithm (JAYA), Fire Hawk Optimizer (FHO), Whale Optimization Algorithm (WOA), Grey Wolf Optimizer (GWO) [34], Social Spider Optimization (SSO), Bat Algorithm (BA), Genetic Algorithm (GA) [35], Differential Evolution (DE), Simulated Annealing (SA), and Particle Swarm Optimization (PSO) [36]. Some fields where bioinspired algorithms are used include engineering, especially computer science, economics, and mechanical design [37].

C. ARTIFICIAL NEURAL NETWORK (ANN)

Artificial Neural Networks (ANNs) represent parallel networks characterized by strong interconnections among neurons (nodes). These networks are structured hierarchically and designed to interact with real-world objects, mirroring the functionality of biological nervous systems [38]. Within the realm of computer science, ANNs are mathematical models governed by various parameters. The model comprises multiple nested functions, incorporating equation 1 iteratively. Effective algorithms for learning in neural networks are typically supported by mathematical demonstrations. A classic example of a neural network is the perceptron, a binary classifier composed of two layers of neurons. The input layer serves as the recipient of external signals, forwarding them to the output layer housing an M-P neuron, commonly known as a threshold logical unit [39].

$$y_j = f(\sum_i W_i X_i - \theta) \tag{1}$$

Various types of artificial neural networks (ANN) utilized in solar radiation (SR) prediction encompass a diverse range. Notably, the Extreme Learning Machine (ELM) stands out as a variant of feed-forward neural networks characterized by a solitary hidden layer [40]. Additionally, recurrent neural networks (RNN), convolutional neural networks (CNN), and multilayer perceptron (MLP) represent other prominent examples in this domain [41].

D. HYPERPARAMETER OPTIMIZATION

Hyperparameter optimization is a critical process in ML and DL that involves selecting the best set of hyperparameters for a given model [42]. As shown in Fig. 1 illustrates the common process for hyperparameter optimization using metaheuristic algorithms. Hyperparameters are the adjustable parameters that control the training process and the structure of the model, such as learning rate, number of hidden layers, and batch size for artificial neural networks, and other hyperparameters for specific ML models like C for SVR and the maximum number of levels in each decision tree for RF [43].

Unlike model parameters, which are learned during training, hyperparameters must be set before the learning process begins [44]. Effective hyperparameter optimization enhances model performance by finding the optimal configuration that minimizes error and improves predictive accuracy. This process often employs various techniques, including grid search, random search, and advanced methods like Bayesian optimization and metaheuristic algorithms, to efficiently navigate the vast hyperparameter space and identify the most effective settings [45].

E. PERFORMANCE METRICS

When creating an artificial intelligence model, it is necessary to evaluate the quality of the model, which is how well the model was trained with the training data with which it was built and how accurate it is in predicting new test observations [46].

Because solar energy monitoring systems require the ability to analyze complex atmospheric datasets, optimization algorithms are employed to selectively choose the necessary

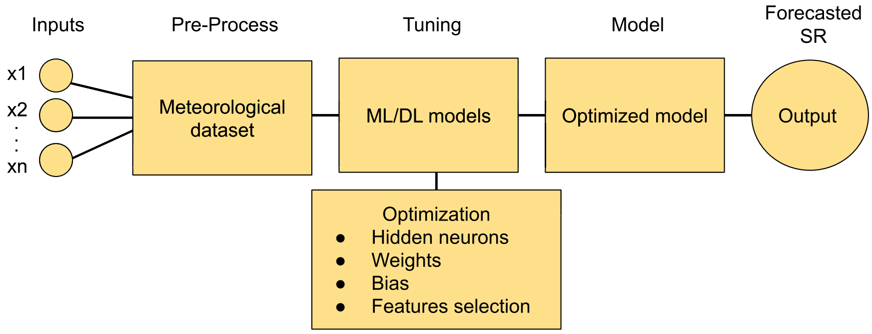

FIGURE 1. Common methodologies for ML/DL model hyperparameters tuning using metaheuristic algorithms, where x1,x2 to xn are input variables.

information. This selection is carried out by eliminating irrelevant features to prevent model performance degradation [6].

For this, different model evaluation metrics are needed, some of them are the following:

- R^2 is a measure representing the proportion of variation in the outcome that can be attributed to the predictor variables. In multiple regression models, R^2 indicates the squared correlation between the observed output values and the predictor values utilized in the model. A higher R^2 value suggests better performance of the model.
- Root Mean Squared Error (RMSE) evaluates the typical error of the model in forecasting observation outcomes. It is mathematically derived as the square root of the Mean Squared Error (MSE), representing the variance between the actual observed values and the model's predicted values. A lower RMSE indicates higher model accuracy.
- Mean Absolute Error (MAE) quantifies prediction accuracy, similar to RMSE. It computes the average absolute disparity between observed and predicted values. Its normalized version (nMAE or NMAE) is commonly used to facilitate comparison across different scales [47]. MAE exhibits lower sensitivity to outliers compared to RMSE.
- Mean Absolute Percentage Error (MAPE) obtained by 2, [7].

$$\text{MAPE} = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{G_{\text{Sim},i} - G_{\text{Act},i}}{X_{\text{Act},i}}\right| \tag{2}$$

- Normalized Mean Absolute Percentage Error (NMAPE) caculated with equation 3 [48]:

$$\text{NMAPE} = \frac{\frac{1}{N}\sum_{i=1}^{N}\left|\hat{X}_{\text{season},i} - X_{\text{season},i}\right|}{\frac{1}{N}\sum_{i=1}^{N} X_{\text{season},i}} \tag{3}$$

- Relative Root Mean Squared Error (rRMSE) obtained as 4 [7]:

$$rRMSE = \frac{1}{\overline{G}_{\text{Act}}}\sqrt{\frac{\sum_{i=1}^{N}(G_{\text{Sim},i} - G_{\text{Act},i})^2}{N}} \tag{4}$$

- Mean Bias Error (MBE) [49]
- Nash-Sutcliffe Efficiency (NSE) based on 5 [50].

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{N}\left[(SR_0)_i - (SR_c)_i\right]^2}{\sum_{i=1}^{N}\left[(SR_0)_i - \overline{SR_0}\right]^2}, \tag{5}$$

$$-\infty < \text{NSE} \leq 1 \tag{6}$$

- Normalized Root Mean Squared Error (NRMSE) obtained by 7

$$\text{NRMSE} = \frac{100}{\overline{x}}\sqrt{\frac{1}{T}\sum_{t=1}^{T}(x_t - \hat{x}_t)^2}, \tag{7}$$

- Root Mean Square Deviation (RMSD) calculed by 8, [51].

$$\text{RMSD} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(Y_{i,o} - Y_{i,p})^2} \tag{8}$$

III. MATERIALS AND METHODS

In this section the methodological framework employed to conduct the systematic literature review is detailed. The process of searching for relevant literature, the inclusion and exclusion criteria applied, as well as the databases and analytical tools used are outlined. Fig. 2 shows an analysis of the number of documents published annually between 2018 and 2024. This timeframe was chosen to ensure the inclusion of the most recent and relevant studies in the field of solar radiation prediction using metaheuristic optimization algorithms and artificial intelligence. This graph is essential for understanding the evolution and growing interest in the search for better hybrid models for predicting SR, highlighting how the quantity of research has increased over the years.

A systematic review was carried out using the methodological framework established by Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [52]. This rigorous approach ensures transparency and quality in conducting the review. To determine the eligibility of articles, inclusion criteria were applied focusing on the presence of specific keywords such as "solar radiation",

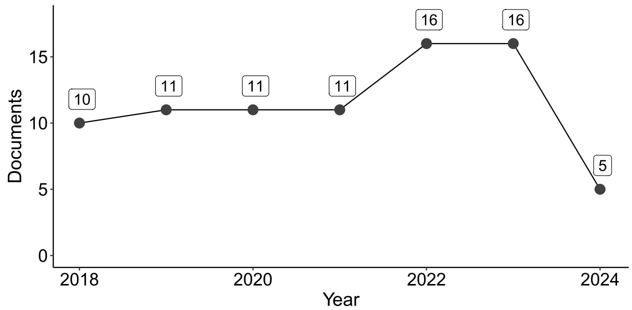

FIGURE 2. Publication report from 2018 to 2024, presenting the documents published per year on Scopus.

"metaheuristics", "optimization algorithms" and the prediction of weather variables using metaheuristic algorithms. This strategic approach ensured the inclusion of relevant and high-quality studies in the analysis.

Selected works were extracted from reputable publishers in the academic and scientific fields, including Elsevier, IEEE, Springer, Wiley, and Taylor and Francis. These publishers are recognized for their excellence in publishing cutting-edge research and their contribution to advancing knowledge in various disciplines. The choice of high-quality sources ensured the reliability and validity of the data collected in this systematic review.

In addition, exclusion criteria were established to further refine the scope of the review. These criteria included restricting the language to Spanish and English, thus ensuring linguistic coherence and accessibility of the reviewed materials. Additionally, the inclusion of articles published between 2018 and 2024 was limited, allowing a focus on recent and relevant research for the topic at hand.

To conduct the search for relevant articles, the Scopus platform was used, known for its extensive coverage and comprehensive academic database. This search yielded a total of 79 relevant publications, carefully selected to address specific aspects related to the prediction of SR. Articles exploring topics such as ML, and DL were prioritized, while those dealing with unrelated areas such as medicine, business, arts, economics, chemistry, mathematics, and sociology were excluded.

Moreover, specific keywords such as "solar radiation," "forecasting," "prediction," "solar radiation predictions," as well as names of ML and ML models, were incorporated to ensure the comprehensiveness of the search and the inclusion of studies relevant to the review objectives. This refined search strategy ensured the collection of a representative and relevant selection of available scientific literature on the topic.

The methodology used to identify relevant articles was designed as follows: TITLE((solar AND radiation) OR (irradiance)) AND TITLE-ABS-KEY((prediction OR forecasting) AND ((machine AND learning) OR prediction) OR (neural AND networks) OR (deep and learning)) AND PUBYEAR > 2017 AND PUBYEAR < 2025 AND NOT (power OR ultraviolet) AND PUBYEAR > 2017 AND PUBYEAR < 2025 AND (LIMIT-TO (SRCTYPE,"j")) AND (LIMIT-TO (DOCTYPE,"ar")).

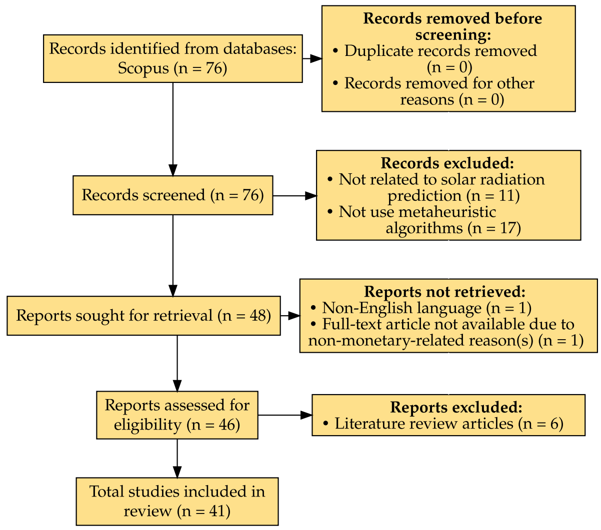

FIGURE 3. Adapted PRISMA 2020 flow diagram.

To filter the obtained articles, the diagram represented in Fig. 3 was applied. Out of the 76 articles initially gathered, only 41 met the established criteria. Among the reasons for exclusion were the lack of information on SR acquisition, the absence of any metaheuristic optimization algorithm, non-use of the English language, unavailability of an article, and the review nature of other articles.

The visual representation of keywords in the authors' bibliometric network reveals five recognized clusters, as depicted in Fig. 4. This visualization encapsulates the relationships and clustering of keywords, providing an intuitive understanding of the thematic structure and interrelations of the research topics covered in this review. The red cluster pertains to algorithms and machine learning related to solar energy. The yellow cluster consists of nodes related to the precision of solar radiation model predictions with time series. The purple cluster encompasses nodes associated with solar irradiation, regression analysis, and temperature. The green cluster links machine learning models to optimization and evaluation metrics such as MSE, along with climate models. Lastly, the most significant cluster, depicted in blue, includes primary nodes such as solar radiation and prediction, along with keywords like climate models and daily global solar radiation.

IV. RESULTS

This section describes various studies in the literature for predicting SR using ML or DL models optimized with an optimization algorithm, such as MLP, CNN, RNN, RF, SVR, ELM, from which their results, the variables used, their

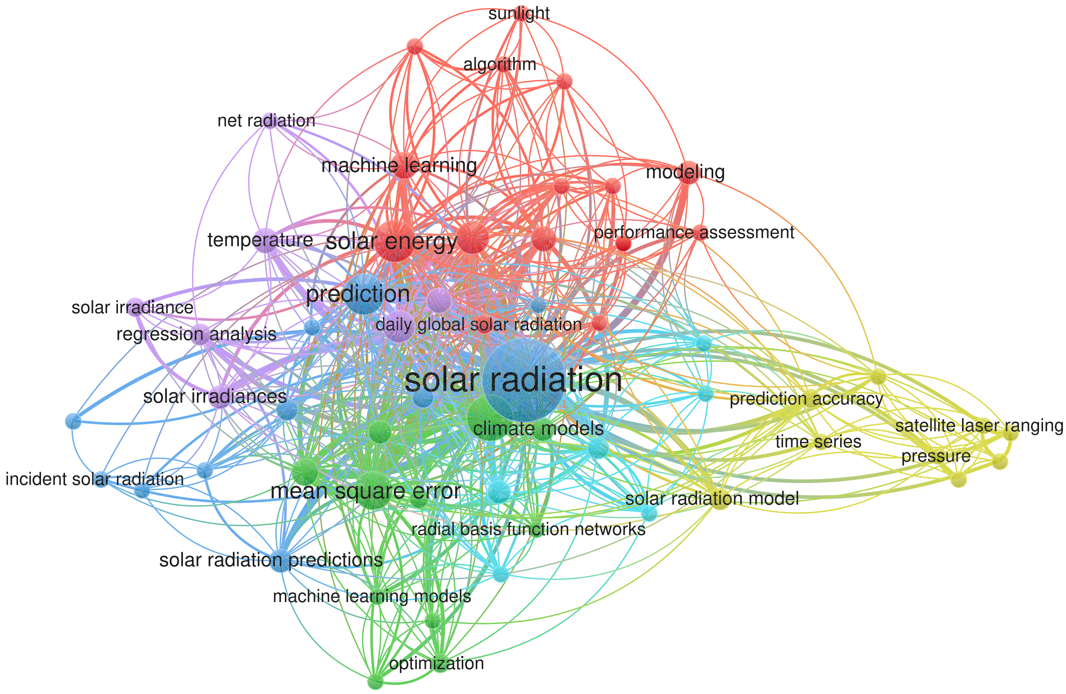

FIGURE 4. Bibliometric network of authors' keywords in solar radiation prediction literature.

prediction time horizon, and some characteristics of their models will be obtained.

A. MLP

Multiple studies have been investigated that have implemented MLP along with various metaheuristic optimization algorithms to enhance the accuracy and efficiency of predictions, as illustrated in Table 1. Each article selected for this review has applied a combination of metaheuristic techniques with MLP, aiming to optimize from feature selection to model parameter tuning. The methodology, results obtained, and relevance of each approach have been analyzed, thus providing a comprehensive insight into how these hybrid methods can be employed to achieve more robust and reliable predictions.

Reference [31] developed a model for predicting global horizontal solar irradiance (GHI) in North Carolina, USA. They used genetic algorithms (GA) and particle swarm optimization (PSO). It was observed that GA demonstrated better performance in optimizing a neural network with 15 hidden layers of MLP type, with an initial population and maximum iterations of 50. The identified optimal activation function was the hyperbolic tangent sigmoidal. The model achieved a root mean square error (RMSE) of 47.6955 Wh/m^2 and a mean absolute error (MAE) of 24.7772 Wh/m^2. Additionally, model validation tests were conducted for periods of 24 hours, 4 days, and one week.

Reference [30] compared MLP, CNN, and RNN models for solar irradiance (SI) prediction, aiming to minimize the error between the actual and predicted values. In this study, a hybrid approach was employed that combines the Deer Hunting Optimization Algorithm (DHOA) and Grey Wolf Optimization (GWO) to optimize the number of hidden neurons in the neural networks. The results showed an RMSE of 212.27 Wh/m^2 and a MAE of 136.51 Wh/m^2.

Reference [55] conducted a study to improve the accuracy in daily prediction of GSR in the arid northwest of China using a Backpropagation Neural Network (BP) and a BP optimized by PSO algorithm (PSO-BP), in order to optimize critical network parameters such as weights and biases. It was found that PSO-BP models (with a coefficient of determination R^2 of 0.7649-0.9678) were more accurate than BP models (R^2 of 0.7215-0.9632) for daily GSR prediction in the four subzones of the arid northwest of China.

Reference [7] implemented an MLP with 5 hidden layers, using the Tansig activation function in the hidden layers and Pureline in the output layer. This model was used to predict daily GSR in Algeria. The firefly algorithm was employed with a population of 10 and a maximum of 50 generations to select the best features for neural network training. The obtained results were $R = 0.9321$ W/m^2, $rRMSE\% = 11.96$, and $MAPE\% = 8.26$.

Reference [47] developed a model to predict monthly SR in semi-arid and humid regions of Iran. In their study, they compared MLP and ANFIS models, which were optimized using the Grasshopper algorithm, with a population of 300 individuals. This approach resulted in the best objective function value of 3.12 (KWh/m^2/day), with an nMAE of 0.54.

Reference [54] utilizes a Feed-Forward Artificial Neural Network (FFANN) wich is a type of MLP, optimized with various metaheuristic algorithms, including Equilibrium Optimization (EO), Optical Inspired Optimization

TABLE 1. Works where MLP are used for SR prediction.

Author	Algorithm	Comparison	Time horizon	Tuning	Results
[31]	GA	-	Daily	Feature selection	RMSE 47.6955 Wh/m^2 and MAE 24.7772 Wh/m^2
[30]	GU-DHOA	CNN, RNN	Intra-hourly	Hidden neurons	RMSE 212.27 Wh/m^2 and MAE 136.51 Wh/m^2
[53]	PSO	-	Intra-hourly	Weights and biases	R^2 0.7649–0.9678
[7]	FFA	SVM	Daily	Feature selection	R 0.9321 W/m^2, rRMSE% 11.96 and MAPE% 8.26
[47]	GOA	ANFIS	Monthly	Not mentioned	nMAE 0.545
[54]	EO	-	Hourly	Weights and biases	RMSE 141.61 W/m^2, MAE 108.07 W/m^2 and R^2 0.91

(OIO), Wind-Driven Optimization (WDO), and Social Spider Algorithm (SOSA). The FFANN has a configuration of (8,6,1), meaning it has 8 neurons in the input layer, 6 in the hidden layer, and 1 in the output layer, using Tansig and Purelin as activation functions. The results showed that the EO-FFANN outperformed the other models in terms of accuracy, with an RMSE of 142.38 W/m^2 in training and 141.61 W/m^2 in testing, an MAE of 110.09 W/m^2 in training and 108.07 W/m^2 in testing, and an R^2 of 0.91. The EO optimizes the weights and biases of the FFANN to minimize the prediction error.

To summarize, the implementation of MLP models combined with various metaheuristic optimization algorithms has significantly improved the accuracy and efficiency of solar radiation predictions. Studies reviewed in this section demonstrate the effective use of genetic algorithms, particle swarm optimization, deer hunting optimization, gray wolf optimization, firefly algorithm, and grasshopper algorithm to optimize different aspects of MLP models, from feature selection to parameter tuning. The results highlight the robustness of these hybrid methods, with notable improvements in key metrics such as RMSE, MAE, and R^2 across different geographical regions and time horizons, reinforcing the potential of these advanced techniques to improve predictive performance in SR forecasting.

B. CNN

Studies have focused on integrating CNN with various metaheuristic optimization algorithms to refine feature selection and optimize model parameters, aiming to improve the accuracy and efficiency of predictions, as demonstrated in Table 2. Through a detailed analysis, it has been discussed how different research has implemented these techniques to address specific challenges in solar prediction, evaluating their methodologies, results, and key contributions.

[6] developed an advanced hybrid model for GSR prediction integrating CNN and utilizing Harris Hawks Optimization (HHO) for effective feature selection, notable for its optimal performance with a population of 250, which proved to be the most effective. This model was tested on various solar farms across Australia, demonstrating superior predictive capability compared to conventional ML such as MLP, ELM, Deep Belief Network (DBN), Deep Neural Network (DNN), and Multivariate Adaptive Regression Splines (MARS). In the Dunblane solar farm, the model achieved an R of 0.960, an RMSE of 2.118 MJ m^2/day, and an MAE of 1.587 MJ m^2/day.

In the study by [56], a hybrid model for GSR prediction was developed combining CNN, LSTM, and MLP, where ReLU was used as the activation function, along with Slime Mould Algorithm (SMA) optimization for feature selection. This approach was applied to solar farms in Queensland, Australia, demonstrating superior performance on all selected farms compared to ML, including ANN, RF, and SADE-ELM. With an R value between 0.880 and 0.935, along with an RMSE between 2.207 and 2.792 MJ/m^2/day across different farms.

In the study conducted by [57], a hybrid model for GSR prediction in Queensland, Australia, was proposed by integrating a CNN with three convolutional layers, using ReLU as the activation function. For optimal feature selection, the Atom Search Optimization (ASO) algorithm was used, and hyperparameters were optimized using the HyperOpt method. The model's performance was evaluated on six solar farms, demonstrating superiority over various alternative models in both ML and DL such as LSTM, RBF, DBN, M5TREE, and MARS. Specifically, the model achieved an RMSE ranging from 2.172 to 3.305 MJ m^2/day and an MAE ranging from 1.624 to 2.370 MJ m^2/day.

Reference [30] developed a hybrid model for solar irradiation prediction using MLP, CNN, and RNN, optimizing the number of hidden neurons using a hybrid algorithm that combines DHOA and GWO, called GU-DHOA. The GU-DHOA-CNN model obtained solar irradiation prediction with an RMSE of 2.11×10^2 and MAE of 166.53.

The study by [58] introduced a SR prediction model based on Hybrid Deep Learning Arithmetic Optimization (AOHDL-SRP), utilizing a CNN with attention-oriented long-term memory (ALSTM). Hyperparameter optimization was performed using the Arithmetic Optimization Algorithm (AOA), highlighting its approach in three stages: preprocessing, prediction, and hyperparameter optimization, with an R^2 of 100% and MSE, RMSE, and MAE values of 0.18, 0.43, and 0.32 respectively.

TABLE 2. Works where CNN are used for SR prediction.

Author	Algorithm	Comparison	Time horizon	Tuning	Results
[6]	HHO	ML(ELM)	Daily	Feature selection	R 0.960, RMSE 2.118 MJ/m^2/day, MAE 1.587 MJ m^2/day
[56]	SMA	ANN, RF and SADE-ELM	Daily	Feature selection	R 0.880-0.935 and RMSE: 2.207-2.792 MJ/m^2/day
[57]	ASO	LSTM, RBF, DBN, M5TREE and MARS	Intra-hourly, daily	Feature selection	RMSE 2.172-3.305 MJ m^2/day and MAE 1.624-2.370 MJ m^2/day.
[30]	GU-DHOA	MLP, RNN	Intra-hourly	Hidden neurons	RMSE 2.11 × 10² and MAE 166.53.
[58]	AOA	-	Intra-hourly	Not specified	MSE 0.18, RMSE 0.48 and MAE 0.32
[59]	SWD	-	Hourly	Time series decomposition, feature selection	NRMSE 0.75% - 3.4%

The article by [59] analyzes the improvement in Direct Normal Irradiance (DNI) prediction for concentrated solar power systems, through a hybrid model combining Variational Mode Decomposition (VMD), Swarm Decomposition Algorithm (SWD) used for time series decomposition, feature selection with Random Forest (RF), and Deep Convolutional Neural Networks (DCNN). This model was tested in four regions with different climatic conditions, demonstrating outstanding accuracy with low prediction errors in terms of NRMSE ranging from 0.75% to 3.4% for one-hour-ahead forecasts.

To summarize, the integration of CNN with various metaheuristic optimization algorithms has shown significant advancements in SR prediction accuracy and efficiency. The studies reviewed illustrate the successful application of optimization techniques such as Harris Hawks Optimization, Slime Mould Algorithm, Atom Search Optimization, and Arithmetic Optimization Algorithm to enhance feature selection and model parameter tuning. These hybrid approaches, applied across diverse geographical locations and conditions, have consistently outperformed traditional machine learning models, achieving impressive results in key metrics such as RMSE, MAE, and R^2. This highlights the robustness and potential of these advanced CNN-based methods in improving predictive performance for SR forecasting.

C. RECURRENT NEURAL NETWORKS

The exploration has focused on how RNNs, combined with various metaheuristic optimization techniques, have been applied to enhance the accuracy of SR predictions, which are presented in Table 3. Each reviewed study contributes to understanding the effectiveness of RNNs in this field, providing insights into the most effective model configurations and optimization strategies that have yielded the best predictive results.

The study conducted by [60] presented an approach to improving the efficiency of SR prediction using a hybrid model that integrates an Adaptive Dynamic Search Squirrel Optimization Algorithm (ADSSOA) with LSTM. This hybrid model, called ADSSOA-LSTM, was designed to optimize SR prediction using climatic factors, notable for its ability to handle complex temporal dependencies in SR data. Specifically, the ADSSOA-LSTM model achieved an RMSE of 0.000388 compared to the RMSE of 0.001221 for the standard LSTM model.

The article by [61] presents an advanced method for SR prediction using Satin Bowerbird Optimization (SBO) with Modified Deep Learning, specifically a hybrid model of Modified Bidirectional Gated Recurrent Unit (BGRU) and Online Sequential Extreme Learning Machine (OSELM), called OMBGRU-SRP. This model incorporates attention mechanisms and skip connections to optimize SR prediction, achieving an MSE of 10.231, RMSE of 3.199, MAE of 1.748, and an R-squared coefficient of 0.974, along with a Mean Absolute Percentage Error (MAPE) of 11.856

Reference [30] developed a hybrid model for solar irradiation prediction using MLP, CNN, and RNN, optimizing the number of hidden neurons through a hybrid algorithm combining DHOA and GWO, called GU-DHOA. The GU-DHOA-RNN model obtained solar irradiation prediction with an RMSE of 3.15×10^2 and MAE of 223.32.

The study by [62] developed a solar irradiation prediction model using a Long Short-Term Memory Neural Network (LSTM) optimized by a hybrid Chicken Swarm Optimization (CSO) and GWO algorithm, called LSTM-CSO-GWO. This model focuses on predicting solar irradiation at different time scales: hourly, daily, and weekly, employing a multitask learning approach to effectively share resources between each task. In terms of statistical metrics, for the hourly time scale, the model achieved an MSE of 0.2392 and an MAE of 0.3562.

[63] analyzes solar irradiation forecasts using hybrid models that combine PSO with three independent models: XGboost (PSO-XGboost), PSO-LSTM, and Gradient Boosting Regression Algorithm (PSO-GBRT), applied to a combined cold and electricity generation microgrid in Tahiti. The hybrid models were compared with other independent models such as ANN, CNN, Random Forests (RF), LSTM, GBRT, and XGboost. This model achieved an MAE of 99.37 W/m^2, an RMSE of 154.84 W/m^2, and R^2 of 0.82.

TABLE 3. The following articles employ some optimized recurrent neural networks models.

Author	Algorithm	Comparison	Time horizon	Recurrent Layers	Tuning	Results
[60]	ADSSOA	-	Hourly	LSTM	Feature selection	RMSE 0.000388
[61]	SBO	-	Hourly	GRU	Hyperparameter tuning	MSE 10.231, RMSE 3.199, MAE 1.748, R^2 0.974, MAPE 11.856%.
[30]	GU-DHOA	MLP, CNN	Intra-hourly	Standard RNN	Number of hidden neurons	RMSE 3.15×10² and MAE 223.32
[62]	CSO	-	Hourly, Daily	LSTM	Hyperparameter tuning	MSE 0.2392 and MAE 0.3562
[63]	PSO	-	Daily	LSTM	Feature selection	MAE 99.37 W/m^2, RMSE 154.84 W/m^2 and R^2 0.82.

To summarize, the application of RNNs combined with various metaheuristic optimization techniques has significantly enhanced the accuracy of SR predictions. The studies reviewed demonstrate the effectiveness of hybrid models, such as ADSSOA-LSTM, OMBGRU-SRP, and LSTM-CSO-GWO, in optimizing temporal dependencies and feature selection for SR data. These advanced methods have consistently outperformed standard models, achieving impressive results in key metrics such as RMSE, MAE, MSE, and R^2. This highlights the potential of RNN-based hybrid models in providing robust and accurate SR predictions across different temporal horizons and geographical regions.

D. OTHER ML MODELS

Beyond conventional neural network models, various other ML approaches have been successfully applied in SR prediction. Several studies have examined the utilization of various models, such as Support Vector Machines (SVM), Random Forest (RF), among others, all optimized with metaheuristic techniques to enhance their predictive performance, as summarized in Table 4.

Each of these models has offered unique advantages in terms of handling different types of data and complexities of the problem. It has been explored how researchers implemented these alternative models, integrating optimization algorithms to fine-tune their structures and parameters, with the aim of obtaining more accurate and effective predictions of SR. Additionally, model configurations and optimization strategies that have been shown to be most effective in the specific context of SR prediction have been discussed.

The study by [55] presents a hybrid model for daily SR prediction, combining ELM optimized by Bat Algorithm (BA), with the use of Wavelet Transform (WT) and Principal Component Analysis (PCA). Evaluated in various cities such as Beijing, New York, Melbourne, and Sao Paulo, it particularly excelled in Beijing for its accuracy. This model includes a network with one hidden layer and 10 hidden neurons, using a sigmoidal activation function. Optimization using BA was configured with an initial population of 10 and up to 500 iterations to adjust the input weight and bias of the hidden layer. WT helps decompose and clean the SR time series, while PCA reduces the dimensionality of the input data. The model yielded results of MAE 0.14, MAPE 2.2%, RMSE 0.028, and an R^2 coefficient of 0.99.

The article by [64] introduces an Elephant Herding Optimization with Deep Extreme Learning Machine (EHO-DELM) model for SR prediction using weather forecasts. The neural network structure includes an input layer with 4 neurons, 6 hidden layers with 10 neurons each, and an output layer with one neuron, using the logistic sigmoid activation function. The Elephant Herding Optimization algorithm (EHO) is employed to optimally adjust the weights and biases of the DELM model, improving predictive accuracy, with an MSE of 265.8546 W/m^2 and an RMSE of 16.30 W/m^2. Comparison with other models highlights the effectiveness of EHO-DELM, outperforming techniques such as SVR, RFR, GBR, and XGB in SR prediction.

The study by [49] proposes a SR prediction model based on the Sine Cosine Algorithm (ASCA). This optimized ensemble model consists of preprocessing and training phases, using K-Nearest Neighbors (KNN) regression for ensemble model training. It is evaluated compared to various algorithms, including PSO, Whale Optimization Algorithm (WOA), GA, with an RMSE of 0.00175482, MAE of 0.00161235, and MBE of -0.00036521.

Reference [40] developed a SR prediction model using WT, Complete Ensemble Empirical Decomposition with Adaptive Noise (CEEMDAN), Improved Atom Search Optimization (IASO), and Outlier-Robust Extreme Learning Machine (ORELM) with a hyperbolic tangent sigmoidal activation function. The model uses WT for noise removal and CEEMDAN to decompose data into intrinsic modes. Then, IASO optimizes the weights and thresholds of ORELM, with a population of 50 and a maximum of 30 iterations. This hybrid approach achieved an RMSE of 26.1022 W/m^2, MAE of 13.9023 W/m^2, an R of 0.9939, and a MAPE of 0.2209.

The study presented in the article by [50] focuses on SR prediction using an enhanced version of the Multi-Verse Optimizer Algorithm (IMVO), integrated with Least Squares Support Vector Machine (LSSVM). This combination is compared with other configurations using different optimization algorithms with LSSVM, including GA, PSO, and others.

TABLE 4. The following articles employ some optimized ML models.

Author	Algorithm	Comparison	Time horizon	Model	Tuning	Results
[55]	BA	LSSVM, BPNN	Hourly	ELM	Weights and biases	RMSE 0.028, MAE 0.14, MAPE 2.2 and R^2 0.99
[64]	EHO	SVR, RFR, GBR, XGB	Intra-hourly	ELM	Weights and biases	RMSE 16.30 W/m^2 and MSE 265.8546 W/m^2
[49]	ASCA	PSO, WOA, GA	Intra-hourly	KNN	Weights and biases	RMSE 0.00175482, MAE 0.00161235 and MBE -0.00036521
[40]	WT, IASO	-	Intra-hourly	ELM	Weights and thresholds	RMSE 26.1022 W/m^2, MAE 13.9023 W/m^2, R 0.9939 and MAPE 0.2209
[50]	IMVO	GA, PSO	-	LSSVR	Hyperparameter tuning	RMSE 0.748, MAE 0.574, R^2 0.969, and NSE 0.964
[4]	ACO	-	Monthly	RF	Feature selection	RMSE 0.89 MJ/m^2, MAE 0.69 MJ/m^2 and R 0.984
[65]	FA	SVM	Daily	RS	Weights assigned to individual forecasts	RMSE 13.11-18.79 W/m^2 and MAE 0.044-0.086
[66]	SSA	LSTM	Intra-hourly	ELM	Hyperparameters	RMSE 23.153 w/m^2, MAE 13.225 w/m^2 and R 0.9923
[5]	IPSO	M5T, GP, MARS	Daily, Monthly	SVR	Feature selection	RMSE 9.01 and MAE 4.10
[67]	ACO, CS, GWO	DT, BP	Daily	SVM	Parameter tunning	RMSE 1.341 MJ/m^2/day, R^2 0.959, and NSE 0.956
[68]	PSO	ELM, BPNN, BPNN	Daily	LSSVM	Parameter tunning	RMSE 0.7524, MAE 0.7257 and R 0.9773
[27]	PSO	-	Intra-hourly, Hourly	SVM	Hyperparameter tuning	R 0.982-0.995
[69]	GSA	-	Daily	LSSVR	Weights and biases	NRMSE 2.96%, MAPE 2.83% and DS 88.24%
[70]	ABC	LSSVM, CS	Hourly	LSSVR	Parameter tunning	RMSE 116.22 Wh/m^2 and R 94.3%
[71]	GA	-	Daily	ELM	Not specified	Improves by 3.64% to 4.49%
[72]	GA, PSO, GWO, HHO	-	Monthly	ANFIS	Parameter selection	RMSE 1.489 MJ/m^2, R^2 0.959 and MAE 1.248 MJ/m^2
[3]	MFO	-	Intra-hourly	XGBoost	Refine model hyperparameters	RMSE 76.29 W/m^2 and R^2 0.9337
[48]	SSA	-	Daily	VMD	Parameter optimization	NRMSE 3.87% and NMAPE 1.58%
[73]	GSA	-	Intra-hourly, Hourly	ELM	Optimizing kernel variables	MAPE 0.71%
[51]	BAT, PSO, WOA	MARS, XGBoost	Daily	SVM	Inputs parameters	RMSD 11.1%, 10.0% and 10.4%
[74]	PSO	-	Daily, Monthly	GEM	Parameter optimization	RMSE 1.045-1.719 MJ m^2 d$^-$1, rRMSE 7.6-12.7%, MAE 0.801-1.283 MJ m^2 d$^-$1, R^2 0.953-0.981
[3]	MFO	MLP, SVR	Daily	RFR	Hyperparameter tuning	RMSE 75.8613 W/m^2 and R^2 0.9402
[75]	CA	PSO	Daily	RF	Hyperparameter tuning	RMSE 0.0366 and MAPE 2.98%

The research was conducted using data from two stations located in the southeast region of China. The main features of the neural network in this study include the use of LSSVM as the primary model for prediction, with IMVO being used to fine-tune the hyperparameters of the LSSVM model. In one of the stations and a data split scenario of 75% for training and 25% for testing, the LSSVM-IMVO model achieved an RMSE of 0.748, MAE of 0.574, R^2 of 0.969, and NSE of 0.964, standing out as the most effective approach among those analyzed.

The article by [4] analyzes SR forecast using a hybrid model, integrating Multivariate Empirical Mode Decomposition (MEMD), Ant Colony Optimization (ACO), and Random Forests (RF). Significant features were selected using ACO, employing 1000 trees and 10 predictors for splitting in the RF model, achieving an RMSE of 0.89 MJ/m^2, MAE of 0.69 MJ/m^2, and an R of approximately 0.984 at the Springfield site.

For the article [65], the primary focus is on accurately predicting hourly GSR using datasets from eight locations in Xinjiang province, China. The method is built upon joint learning with a Random Subspace (RS) method to split the original data into several covariate subspaces. A new covariate selection method called Smoothly Trimmed Absolute Deviation Square Root (SRSCAD) is introduced, applied to each subspace for relevant covariate extraction. To combine predictions obtained through RS and SRSCAD, a Firefly Algorithm (FA) is used to estimate weights assigned to individual forecasts, with RMSE ranging between 13.11 and 18.79 W/m^2 for 24-hour ahead prediction and between 17.79 and 21.12 W/m^2 for 48-hour ahead prediction, and MAE ranging from 0.044 to 0.086, outperforming other compared methods such as Elman neural networks, SVM, and Angstrom-Prescott empirical models.

The study by [66] proposes a hybrid model for SR prediction, named OVMD-PACF-ISSA-DBN-OSELM, which incorporates Optimal Variational Modal Decomposition (OVMD), Partial Autocorrelation Function (PACF), and a hybrid model of Deep Belief Neural Network (DBN) with Online Sequential Extreme Learning Machine (OSELM), optimized by an Improved Sparrow Search Algorithm (ISSA). The model achieves an RMSE of 23.153 W/m^2, MAE of 13.225 W/m^2, and R of 0.9923.

Reference [5] developed a model to predict SR using Support Vector Regression (SVR) with radial kernel function and an Improved Particle Swarm Optimization (IPSO) algorithm with a population of 40 particles. This study focused on enhancing the global search capability of the PSO algorithm and integrating SVR with IPSO to obtain optimal SVR parameter values. Two SR stations in Turkey were considered to evaluate the model, along with various comparative models, including M5 tree model (M5T), Genetic Programming (GP), SVR integrated with other optimization algorithms, and Multivariate Adaptive Regression Splines (MARS), with an RMSE of 9.01 and MAE of 4.10 for the Antakya station.

The study by [67] optimized SVM models for GSR prediction in different climatic zones of China, using ACO, CS, and GWO algorithms to tune SVM's internal parameters. Among independent models, SVM demonstrated the highest accuracy in all zones, followed by Decision Tree (DT) model and BP. Hybrid models, especially GWO-SVM and CS-SVM, outperformed the standalone SVM model in all climatic zones, where GWO-SVM emerged as the most effective model, achieving an R^2 of up to 0.959, RMSE from 1.341 MJ/m^2/day, and NSE up to 0.956 in daily GSR prediction.

In the study by [68], four AI technologies were developed to estimate GSR in Zhengzhou, China. These technologies include ELM, a hybrid model combining Least Squares Support Vector Machine with PSO (PSO-LSSVM), where PSO is used to optimize LSSVM parameters with radial basis kernel function, Back Propagation Neural Network (BPNN), and Generalized Regression Neural Network (GRNN). Six meteorological variables were selected as evaluation indices, with daily GSR as output. For the PSO-LSSVM model, the R for training, testing datasets, and all samples is 0.9812, 0.9773, and 0.9793, respectively, RMSE of 0.6324, 0.7524, and 0.7425, respectively, and for MAE are 0.5128, 0.7257, and 0.6228. These results demonstrate that the PSO-LSSVM model performs superiorly and is ranked in order of performance as PSO-LSSVM > ELM > BPNN > GRNN. Thus, the hybrid PSO-LSSVM model can be used to predict daily GSR with high accuracy in Zhengzhou, China.

The study by [27] employed an approach combining PSO for feature selection, with 500 maximum iterations, a population range of 100 to 1000, and Maximal Overlap Discrete Wavelet Transform (moDWT) for data decomposition in three Australian solar cities: Alice Springs, Coburg, and Perth. This method allowed tuning SVR model hyperparameters with Grid Search, achieving a notable correlation between predictions and observations for R from 0.982 to 0.995.

In this study, [69] developed a learning approach for SR prediction using a Decomposition-Clustering-Ensemble (DCE) method in Beijing, employing Ensemble Empirical Mode Decomposition (EEMD) to decompose original SR data. For individual prediction and component assembly, Least Squares Support Vector Regression (LSSVR) was used, optimized by Gravitational Search Algorithm (GSA). Clustering was performed with the Kmeans method, applying a second LSSVR-GSA method for final assembly, based on cluster assembly weights, achieving a Normalized Root Mean Square Error (NRMSE) of 2.96%, MAPE of 2.83%, and DS of 88.24% for one-day-ahead prediction.

In the study conducted by [70], a new hybrid model named ABC-LS-SVM was proposed for global solar radiation (GHI) prediction several hours in advance, using Artificial Bee Colony Algorithm (ABC) to optimize free parameters of LS-SVM model. This approach was validated with data measured over five years, from 2013 to 2017, at the Renewable Energies Applied Research Unit (URAER) in Ghardaïa, southern Algeria. The prediction results of GHI with 12 hours in advance using the ABC-LS-SVM model led to an RMSE of 116.22 Wh/m^2 and an R of 94.3%. In comparison, the classical LS-SVM model obtained an RMSE of 117.73 Wh/m^2 and an R of 92.42%, while the combination of CS with LS-SVM resulted in an RMSE of 116.89 Wh/m^2 and an R of 93.78%.

The study [71] presents a deep generative model, the Convolutional Graph Autoencoder (CGAE), to predict spatiotemporal probabilistic solar irradiance using a deep neural network. It captures spatiotemporal features of historical SR observations to generate samples estimating future solar irradiaerms once. The CGAE model demonstrated superior performance in terms of reliability, sharpness, and Continuous Ranked Probability Score (CRPS) compared to both temporal and spatiotemporal reference methods. CGAE achieved a 3.64% to 4.49% improvement in reliability over the best temporal model (ELM) for one-hour and six-hour ahead predictions, respectively.

The article [72] develops an approach to improve SR predictions using global climate models and advanced ML methods, specifically ANFIS combined with metaheuristic algorithms such as GA, PSO, GWO, and HHO. Initial populations of 500 and maximum iterations of 1000 are used to efficiently adjust their internal parameters. Using temperature data as the sole input variable, various data split configurations for training and testing were compared, finding that the enhanced ANFIS with optimization algorithms showed superior accuracy in SR prediction. In particular, ANFIS combined with GWO optimization (ANFIS-GWO) proved to be the most effective, with an R^2 of 0.959 and the lowest RMSE and MAE errors of 1.489 MJ/m^2 and 1.248 MJ/m^2, respectively, for the Adana station with an 80%-20% training-testing data split.

The article [3] analyzes a solar irradiance prediction model using Extreme Gradient Boosting (XGB) technique with 100 to 1000 trees and a maximum depth of 5, optimized through metaheuristic algorithms such as Moth Flame Optimization (MFO) and GWO. Five population sizes (50, 100, 150, 200, and 500) and 100 maximum iterations were experimented with to refine the XGB model hyperparameters, aiming to maximize prediction accuracy of solar irradiance at a specific location. Highlighted results include an R^2 of 0.9337 for GHI, using the XGB-GWO model, with an RMSE of 76.29 W/m^2.

The article by [48] proposes a model for predicting solar irradiance over a 24-hour horizon. For this prediction, Variational Mode Decomposition (VMD) techniques and the STACK algorithm with 200 neurons in the first hidden layer and 500 in the second are used, along with VMD parameter optimization using the Salp Swarm Algorithm (SSA). It is emphasized that seasonal data partitioning and proper VMD parameter selection through SSA are crucial for improving prediction accuracy. The results show a Normalized Root Mean Square Error (NRMSE) of 3.87% and a Normalized Mean Absolute Percentage Error (NMAPE) of 1.58% for cloudy days.

In the article by [73], a prediction model based on Kernel Extreme Learning Machine (KELM) optimized using Chaotic Gravitational Search Algorithm (CGSA) is proposed to forecast solar irradiance, crucial for photovoltaic energy management. The optimized model, called OKELM, significantly reduces parameter selection process by optimizing kernel variables, achieving more precise and efficient prediction. It is validated using data from a 1 MW photovoltaic plant in India, with outstanding results in solar irradiance prediction under various weather conditions. The KELM model, before optimization, does not require defining the number of hidden nodes, reducing result variability. CGSA implementation improves model generalization and stability, demonstrated by metrics such as MAPE, which in the optimized case reaches a value of 0.71% for the summer season.

The article [51] analyzes the prediction of daily Diffuse Solar Radiation (DSR) in polluted regions using hybrid support vector machines with heuristic algorithms: PSO, BA, and WOA. These models are compared with standard SVM models, Multivariate Adaptive Regression Splines (MARS), and Extreme Gradient Boosting (XGBoost). Results indicate that incorporating air pollution variables (PM2.5, PM10, O3) significantly improves model accuracy in predicting daily DSR, with average relative decreases in Root Mean Square Deviation (RMSD) of 11.1%, 10.0%, and 10.4%. Moreover, the SVM-BAT model outperformed SVM-PSO and SVM-WOA models, as well as standard SVM, MARS, and XGBoost models, proving to be the most effective for predicting daily DSR, especially in regions with high atmospheric pollution.

The study by [74] focuses on improving daily and monthly GSR predictions in Northeast China by optimizing PSO algorithms integrated with a Gaussian exponential model (GEM). The PSO-GEM model is compared with four other ML models and two empirical models using daily weather data from 1997 to 2016 from four stations in Northeast China. It achieved an RMSE between 1.045 and 1.719 MJ m^2 d^-1, rRMSE from 7.6 to 12.7%, MAE from 0.801 to 1.283 MJ m^2 d^-1, R^2 from 0.953 to 0.981 for daily estimations. For monthly estimations, PSO-GEM also showed the highest accuracy with RMSE from 0.197 to 0.575 MJ m^2 d^-1, rRMSE from 1.5 to 7.0%, MAE from 0.137 to 0.499 MJ m^2 d^-1, R^2 from 0.999 to 1.

Reference [3] developed an advanced model for solar irradiance prediction using ML models like SVR, MLP, and RFR, optimized with MFO, GWO, and Evolve Class Topper Optimization (ECTO). For SVR, optimal parameters include an RBF kernel function and a regularization parameter C optimized to 0.01. For MLP, 100 hidden layers were identified as optimal with ReLu activation function. RFR showed superior performance with 100 trees and a maximum depth of 5. It was observed that the RFR-ECTO model outperformed others, offering the highest R^2 values of 0.9402 with the lowest RMSE value of 75.8613 W/m^2.

Reference [75] developed a prediction technique for SR and wind speed using a deep learning-based random forest approach, optimized by a novel Coot algorithm (CA). This model outperformed the PSO algorithm in terms of accuracy. The study managed to reduce RMSE to 0.0366 for solar irradiance and 0.0602 for wind speed, while MAPE stood at 2.98% and 4.78% respectively.

In summary, various ML models, including SVM, RF, and others, have been successfully applied to SR prediction, each enhanced by metaheuristic optimization techniques. The studies reviewed demonstrate the unique advantages of these models in handling diverse data types and problem complexities. Through optimization algorithms such as Bat Algorithm, Elephant Herding Optimization, Sine Cosine Algorithm, and more, these models have achieved significant improvements in predictive accuracy and efficiency. The results highlight the most effective configurations and strategies, showcasing their potential to provide robust and precise SR predictions across different geographical locations and time horizons.

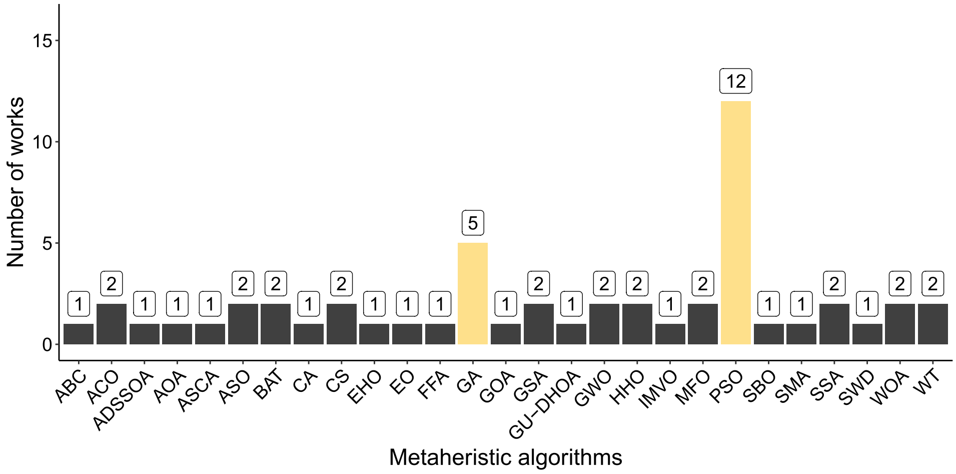

FIGURE 5. The number of works utilizing the different optimization algorithms.

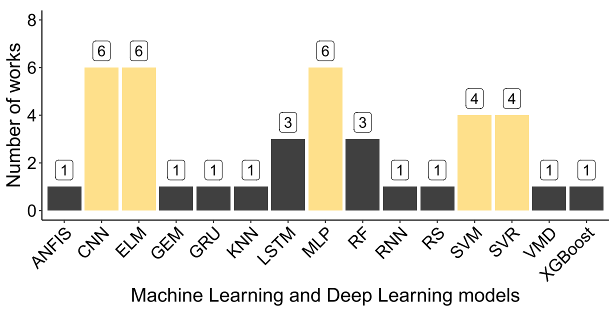

FIGURE 6. The number of studies utilizing the different AI models reviewed in the literature.

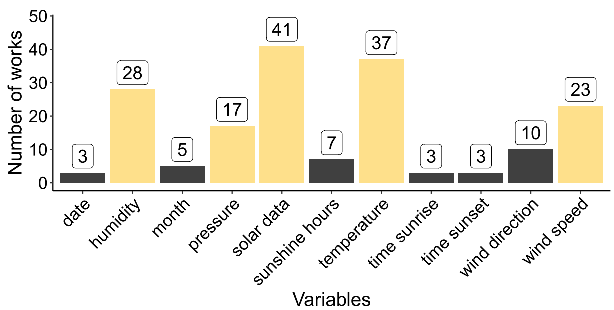

FIGURE 8. The number of studies utilizing the different meteorological variables.

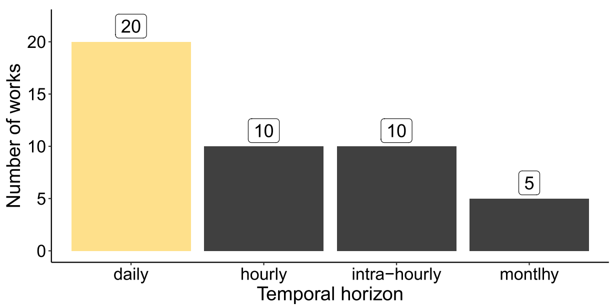

FIGURE 7. The number of works utilizing the different temporal horizons reviewed in the literature.

V. DISCUSSION

As shown in Fig. 5, the PSO algorithm has stood out as the most widely used optimizer in the scientific literature, having appeared in a total of 12 studies, representing 29% of the reviewed literature. This highlights its importance and effectiveness in solving a wide range of problems. Close behind is the GA, which has been present in 5 occasions, representing 12.1%, confirming its popularity and applicability in various contexts.

Other optimization methods, such as ACO, ASO, BAT, CS, GSA, GWO, HHO, MFO, SSA, WOA, and WT, have also been employed in 2 works each, although some of them were only mentioned in their respective articles. This array of approaches reflects the diversity of available tools and the ongoing quest for effective methods to address complex problems.

In the field of SR prediction, as illustrated in Fig. 6, it has been observed that CNN, ELM, and MLP have emerged as the most used models, representing 43.9% of the works, suggesting their ability to capture complex patterns in solar data.

Regarding the time horizons utilized in the studies shown in Fig. 7, it is noteworthy that the daily period has been the most frequently referenced, with a total of 20 mentions, constituting 48.7% of the reviewed literature. This indicates a significant emphasis on short-term prediction and the optimization of energy management for production, consumption, and network operations. This emphasis underscores the importance of addressing immediate operational needs. Subsequently, the intra-hourly and hourly periods have garnered 10 mentions, accounting for 24.3% of the references. This suggests a keen interest in predictions across various time scales for effective power management and system safety considerations. Conversely, the monthly period has received comparatively less attention, with only

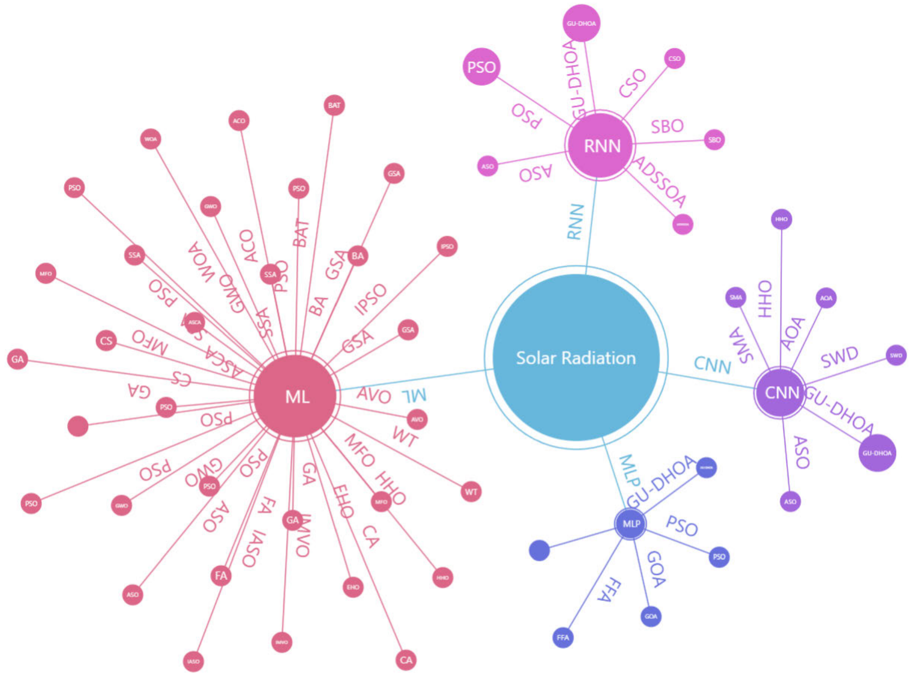

FIGURE 9. Main techniques of AI for predicting solar radiation with optimization algorithms.

5 mentions. This may imply its diminished relevance in certain contexts within the field of solar radiation research.

The distribution of works based on the utilization of various meteorological variables, as reviewed in the literature, has been depicted in Fig. 8. An analysis of the most frequently employed variables reveals the consistent inclusion of solar data across all studies, reflecting its pivotal role in SR prediction. Temperature has emerged as a widely studied variable, mentioned 37 times, representing 90% of the cases, underscoring its significance in prediction models. Additionally, humidity, wind speed, and atmospheric pressure have emerged prominently, with 28, 23, and 17 mentions, respectively, constituting 68.2%, 56%, and 41.4% of the cases. These findings highlight the importance of these variables in the modeling and prediction of SR.

According to the review conducted, algorithms such as MLP, CNN, and RNN, including LSTM and GRU, are widely used due to their effectiveness when integrated with specific optimization strategies such as PSO, FFA, and GWO. In this context, MLP shows adaptability when combined with PSO and FFA, due to its ability to optimize solutions in complex predictions. Meanwhile, CNN, when integrated with AOA and GWO, facilitates the identification of relevant spatial and temporal characteristics, enabling the discovery of complex patterns in solar radiation prediction. It is also important to highlight the effectiveness of the ADSSOA-LSTM model, which achieves an RMSE of 0.000388, making it one of the best-performing models in the literature. Furthermore, methods like GWO, known for their robustness in optimization, strengthen adaptation in dynamic environments, Fig. 9 was created to summarize this information.

Additionally, RNN, when combined with algorithms such as ADSSOA and PSO, allows for the precise adjustment of pattern-based event predictions to the models, thus improving the model's adaptability to new conditions without compromising accuracy. The use of transformation techniques such as the Wavelet Transform (WT) facilitates the decomposition of solar radiation time series into more manageable components, enhancing the accuracy of predictive models. According to these approaches, the diversity of models not only provides adaptability and efficiency in data processing but also improves the accuracy of predictions, which significantly impacts the management of renewable energy resources.

Among the most commonly used AI techniques for time series prediction were MLP, LSTM, and, more recently, Transformer, which typically included an attention layer in the neural model. MLP stood out in several reviewed studies, highlighting its effectiveness in SR prediction when combined with metaheuristic algorithms. However, LSTM and Transformer were not featured as frequently. This suggests that while MLP had been widely adopted, there is an opportunity to further explore the potential of LSTM and Transformer in this specific field.

VI. CONCLUSION

In this study, the application of metaheuristic algorithms for the prediction of SR has been extensively explored. The results have highlighted the predominant efficacy of the PSO algorithm, which has been demonstrated to be the optimizer most used, representing 29% in the reviewed literature, closely followed by the GA with 12.1%. This preference indicates the robustness and adaptability of these algorithms against a diverse range of scenarios and challenges in solar prediction and optimization, particularly emphasizing the ADSSOA-LSTM model, which achieves an RMSE of 0.000388. This outcome not only underscores the effectiveness of combining metaheuristic optimization algorithms with ML techniques, but also emphasizes the importance of properly selecting hyperparameters and features in solar prediction models [76].

Furthermore, it has been observed that models such as PSO-LSSVM and SVM-BAT have also shown promising results, highlighting the relevance of hybrid strategies in improving predictive accuracy under different environmental and temporal conditions. The variability in the effectiveness of these models suggests that model selection may have been critically dependent on the specific context and characteristics of the dataset used.

Moreover, comparative evaluation with advanced ML models has revealed that techniques such as CNN, ELM and MLP have emerged as promising tools due to their ability to capture complex patterns in SR data. These models have been identified as the most widely used AI models in this review, representing 43.9% of the works, and are crucial for improving the accuracy of short- and long-term predictions. This reflects a significant advancement in prediction methodology, opening new directions for future research that could focus on the integration of these models.

Based on the findings and discussion in the study on the use of metaheuristic algorithms and ML techniques in SR prediction, several directions for future research have been suggested that could expand and deepen the impact and application of these technologies:

- Investigate the integration of advanced hybrid predictive models, such as CNN and ELM, optimized through PSO and GA. Integrated with real-time intelligent control systems, these models would not only improve SR prediction but also allow automatic and real-time adjustments of system parameters to maximize energy efficiency.
- Explore the development of new hybrid algorithms that merge the advantages of other unexplored metaheuristic algorithms with ML techniques. This approach would focus on refining accuracy in SR predictions and real-time optimization efficiency, opening possibilities for practical applications in solar energy systems.
- Investigate how hybrid algorithms such as PSO-CNN, PSO-ELM, GA-CNN, and GA-ELM could be adapted to optimize energy management and distribution in microgrids. This study would be particularly relevant for remote areas or regions with limited access to centralized electrical infrastructure, improving the resilience and efficiency of these systems.
- Include studies using attention mechanisms in DL models. Attention mechanisms have shown great promise in enhancing the performance of DL models by allowing the model to focus on the most relevant parts of the input data, potentially improving the accuracy and robustness of SR predictions.

This analysis underscores the importance of continued development and adaptation of metaheuristic algorithms and ML models to effectively address the challenges of solar energy management [77]. The integration of these hybrid models into practical applications could potentially revolutionize the accuracy of SR predictions, which is vital for optimizing the performance of solar energy systems and efficient energy management [78].

ACKNOWLEDGMENT

The authors sincerely thank the Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT) for the invaluable support provided through the scholarship. Acknowledgments are extended to the for funding the research project.

REFERENCES

[1] D. Sharma and P. Khurana, "Solar power: Challenges, mission and potential of solar power in India," *Target*, vol. 2010, no. 3, p. 13, Mar. 2018.

[2] J. Fan, L. Wu, F. Zhang, H. Cai, X. Wang, X. Lu, and Y. Xiang, "Evaluating the effect of air pollution on global and diffuse solar radiation prediction using support vector machine modeling based on sunshine duration and air temperature," *Renew. Sustain. Energy Rev.*, vol. 94, pp. 732–747, Oct. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032118304659

[3] K. Namrata, M. Kumar, and N. Kumar, "Data-driven hyperparameter optimized extreme gradient boosting machine learning model for solar radiation forecasting," *Adv. Electr. Electron. Eng.*, vol. 20, no. 4, pp. 549–559, Feb. 2023. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85162649473&doi=10.15598%2faeee.v20i4.4650&partnerID=40&md5=61f12c105a82d1c3717d6e9065e96160

[4] R. Prasad, M. Ali, P. Kwan, and H. Khan, "Designing a multi-stage multivariate empirical mode decomposition coupled with ant colony optimization and random forest model to forecast monthly solar radiation," *Appl. Energy*, vol. 236, pp. 778–792, Feb. 2019.

[5] H. Ghazvinian, S.-F. Mousavi, H. Karami, S. Farzin, M. Ehteram, M. S. Hossain, C. M. Fai, H. B. Hashim, V. P. Singh, F. C. Ros, A. N. Ahmed, H. A. Afan, S. H. Lai, and A. El-Shafie, "Integrated support vector regression and an improved particle swarm optimization-based model for solar radiation prediction," *PLoS ONE*, vol. 14, no. 5, May 2019, Art. no. e0217634. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85066505795&doi=10.1371%2fjournal.pone.0217634&partnerID=40&md5=261a81c04082f7656b172db8fbb7804c

[6] S. Ghimire, R. C. Deo, D. Casillas-Pérez, and S. Salcedo-Sanz, "Boosting solar radiation predictions with global climate models, observational predictors and hybrid deep-machine learning algorithms," *Appl. Energy*, vol. 316, Jun. 2022, Art. no. 119063.

[7] H. Djeldjli, D. Benatiallah, C. Tanougast, and A. Benatiallah, "Solar radiation forecasting based on ANN, SVM and a novel hybrid FFA-ANN model: A case study of six cities south of Algeria," *AIMS Energy*, vol. 12, no. 1, pp. 62–83, 2024.

[8] C. S. W. Ng, M. N. Amar, A. J. Ghahfarokhi, and L. S. Imsland, "A survey on the application of machine learning and metaheuristic algorithms for intelligent proxy modeling in reservoir simulation," *Comput. Chem. Eng.*, vol. 170, Feb. 2023, Art. no. 108107. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0098135422004409

[9] Y. Zhou, Y. Liu, D. Wang, X. Liu, and Y. Wang, "A review on global solar radiation prediction with machine learning models in a comprehensive perspective," *Energy Convers. Manage.*, vol. 235, May 2021, Art. no. 113960. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85102358454&doi=10.1016%2fj.enconman.2021.113960&partnerID=40&md5=c1de4beb8d33c6e6ee0efeeb826bf6b4

[10] C. Voyant, G. Notton, S. Kalogirou, M.-L. Nivet, C. Paoli, F. Motte, and A. Fouilloy, "Machine learning methods for solar radiation forecasting: A review," *Renew. Energy*, vol. 105, pp. 569–582, May 2017.

[11] A. Sharma and A. Kakkar, "Forecasting daily global solar irradiance generation using machine learning," *Renew. Sustain. Energy Rev.*, vol. 82, pp. 2254–2269, Feb. 2018. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85028350144&doi=10.1016%2fj.rser.2017.08.066&partnerID=40&md5=aa422b6636550ff2b60f9bcfe3346759

[12] T. Rajasundrapandiyanleebanon, K. Kumaresan, S. Murugan, M. S. P. Subathra, and M. Sivakumar, "Solar energy forecasting using machine learning and deep learning techniques," *Arch. Comput. Methods Eng.*, vol. 30, no. 5, pp. 3059–3079, Jun. 2023. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85148607782&doi=10.1007%2fs11831-023-09893-1&partnerID=40&md5=e92aeeb6456ac01a8934f2d3702b7b97

[13] J. M. Álvarez-Alvarado, J. G. Ríos-Moreno, S. A. Obregón-Biosca, G. Ronquillo-Lomelí, E. Ventura-Ramos, and M. Trejo-Perea, "Hybrid techniques to predict solar radiation using support vector machine and search optimization algorithms: A review," *Appl. Sci.*, vol. 11, no. 3, p. 1044, Jan. 2021.

[14] A. E. Gürel, Ü. Ağbulut, H. Bakır, A. Ergün, and G. Yıldız, "A state of art review on estimation of solar radiation with various models," *Heliyon*, vol. 9, no. 2, Feb. 2023, Art. no. e13167.

[15] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*. New York, NY, USA: Springer, 2002.

[16] Y. Ensafi, S. H. Amin, G. Zhang, and B. Shah, "Time-series forecasting of seasonal items sales using machine learning—A comparative analysis," *Int. J. Inf. Manage. Data Insights*, vol. 2, no. 1, Apr. 2022, Art. no. 100058. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2667096822000027

[17] B. Lim and S. Zohren, "Time-series forecasting with deep learning: A survey," *Philos. Trans. Roy. Soc. A*, vol. 379, no. 2194, 2021, Art. no. 20200209.

[18] Z. Han, J. Zhao, H. Leung, K. F. Ma, and W. Wang, "A review of deep learning models for time series prediction," *IEEE Sensors J.*, vol. 21, no. 6, pp. 7833–7848, Mar. 2021.

[19] P. Lara-Benítez, M. Carranza-García, and J. C. Riquelme, "An experimental review on deep learning architectures for time series forecasting," *Int. J. Neural Syst.*, vol. 31, no. 3, Mar. 2021, Art. no. 2130001, doi: 10.1142/s0129065721300011.

[20] Y.-Y. Hong and J. J. F. Martinez, "Forecasting solar irradiation using convolutional long short-term memory and feature selection of data from neighboring locations," *Sustain. Energy, Grids Netw.*, vol. 38, Jun. 2024, Art. no. 101271. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352467723002795

[21] J. F. Torres, D. Hadjout, A. Sebaa, F. Martínez-Álvarez, and A. Troncoso, "Deep learning for time series forecasting: A survey," *Big Data*, vol. 9, no. 1, pp. 3–21, Feb. 2021.

[22] M. K. Sandhya Arora, "Forecasting the future: A comprehensive review of time series prediction techniques," *J. Electr. Syst.*, vol. 20, no. 2s, pp. 575–586, Apr. 2024.

[23] A. Zeng, M. Chen, L. Zhang, and Q. Xu, "Are transformers effective for time series forecasting?" in *Proc. AAAI Conf. Artif. Intell.*, vol. 37, no. 9, 2023, pp. 11121–11128.

[24] R. A. Rajagukguk, R. A. A. Ramadhan, and H.-J. Lee, "A review on deep learning models for forecasting time series data of solar irradiance and photovoltaic power," *Energies*, vol. 13, no. 24, p. 6623, Dec. 2020.

[25] A. Dahmani, Y. Ammi, and S. Hanini, "A novel non-linear model based on bootstrapped aggregated support vector machine for the prediction of hourly global solar radiation," *Smart Grids Sustain. Energy*, vol. 9, no. 1, 2024, Art. no. 3. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85178247865&doi=10.1007%2fs40866-023-00179-w&partnerID=40&md5=855887600dbced3a534a5a803f998269

[26] X. Yu, "Evaluating parallelized support vector regression and nearest neighbor regression with different input variations for estimating daily global solar radiation of the humid subtropical region in China," *Int. J. Low-Carbon Technol.*, vol. 18, pp. 95–110, Feb. 2023. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85160210613&doi=10.1093%2fijlct%2fctad005&partnerID=40&md5=f3944dc43b876a00ea9f7ed24f69f908

[27] S. Ghimire, R. C. Deo, N. Raj, and J. Mi, "Wavelet-based 3-phase hybrid SVR model trained with satellite-derived predictors, particle swarm optimization and maximum overlap discrete wavelet transform for solar radiation prediction," *Renew. Sustain. Energy Rev.*, vol. 113, Oct. 2019, Art. no. 109247. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85068217665&doi=10.1016%2fj.rser.2019.109247&partnerID=40&md5=082ff30f958637d84649a37100a18442

[28] M. Chaibi, E. M. Benghoulam, L. Tarik, M. Berrada, and A. El Hmaidi, "Machine learning models based on random forest feature selection and Bayesian optimization for predicting daily global solar radiation," *Int. J. Renew. Energy Develop.*, vol. 11, no. 1, pp. 309–323, Feb. 2022. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85123176455&doi=10.14710%2fIJRED.2022.41451&partnerID=40&md5=f02e1446ccfaad5a99a6fb5f18a2de0d

[29] O. Bamisile, A. Oluwasanmi, C. Ejiyi, N. Yimen, S. Obiora, and Q. Huang, "Comparison of machine learning and deep learning algorithms for hourly global/diffuse solar radiation predictions," *Int. J. Energy Res.*, vol. 46, no. 8, pp. 10052–10073, Jun. 2022.

[30] T. Vaisakh and R. Jayabarathi, "Analysis on intelligent machine learning enabled with meta-heuristic algorithms for solar irradiance prediction," *Evol. Intell.*, vol. 15, no. 1, pp. 235–254, Mar. 2022.

[31] A. Jadidi, R. Menezes, N. De Souza, and A. De Castro Lima, "A hybrid GA–MLPNN model for one-hour-ahead forecasting of the global horizontal irradiance in Elizabeth city, North Carolina," *Energies*, vol. 11, no. 10, p. 2641, Oct. 2018.

[32] I. M. Mfetoum, S. K. Ngoh, R. J. J. Molu, B. F. Nde Kenfack, R. Onguene, S. R. D. Naoussi, J. G. Tamba, M. Bajaj, and M. Berhanu, "A multilayer perceptron neural network approach for optimizing solar irradiance forecasting in central Africa with meteorological insights," *Sci. Rep.*, vol. 14, no. 1, p. 3572, Feb. 2024.

[33] P. Kumari and D. Toshniwal, "Deep learning models for solar irradiance forecasting: A comprehensive review," *J. Cleaner Prod.*, vol. 318, Oct. 2021, Art. no. 128566.

[34] F. K. Karim, D. S. Khafaga, M. M. Eid, S. K. Towfek, and H. K. Alkahtani, "A novel bio-inspired optimization algorithm design for wind power engineering applications time-series forecasting," *Biomimetics*, vol. 8, no. 3, p. 321, Jul. 2023.

[35] M. Shahvaroughi Farahani and S. H. Razavi Hajiagha, "Forecasting stock price using integrated artificial neural network and metaheuristic algorithms compared to time series models," *Soft Comput.*, vol. 25, no. 13, pp. 8483–8513, Jul. 2021.

[36] J. Behnamian and S. M. T. Fatemi Ghomi, "Development of a PSO–SA hybrid metaheuristic for a new comprehensive regression model to time-series forecasting," *Expert Syst. Appl.*, vol. 37, no. 2, pp. 974–984, Mar. 2010.

[37] A. Vasuki, *Nature-Inspired Optimization Algorithms*. Boca Raton, FL, USA: CRC Press, 2020.

[38] T. Kohonen, *Self-Organization and Associative Memory*, vol. 8. Berlin, Germany: Springer, 2012.

[39] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: Analysis, applications, and prospects," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 12, pp. 6999–7019, Dec. 2022.

[40] C. Zhang, L. Hua, C. Ji, M. S. Nazir, and T. Peng, "An evolutionary robust solar radiation prediction model based on WT-CEEMDAN and IASO-optimized outlier robust extreme learning machine," *Appl. Energy*, vol. 322, Sep. 2022, Art. no. 119518. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85132880447&doi=10.1016%2fj.apenergy.2022.119518&partnerID=40&md5=f18af0b02be8593d76c2cc945ae7e2f4

[41] J. A. Bellido-Jiménez, J. E. Gualda, and A. P. García-Marín, "Assessing new intra-daily temperature-based machine learning models to outperform solar radiation predictions in different conditions," *Appl. Energy*, vol. 298, Sep. 2021, Art. no. 117211. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85110378076&doi=10.1016%2fj.apenergy.2021.117211&partnerID=40&md5=b2d81901f9c3fbd621b1f4dbfd3c2d72

[42] B. Bischl, M. Binder, M. Lang, T. Pielok, J. Richter, S. Coors, J. Thomas, T. Ullmann, M. Becker, and A.-L. Boulesteix, "Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges," *Wiley Interdiscipl. Rev., Data Mining Knowl. Discovery*, vol. 13, no. 2, p. e1484, 2023.

[43] L. Yang and A. Shami, "On hyperparameter optimization of machine learning algorithms: Theory and practice," *Neurocomputing*, vol. 415, pp. 295–316, Nov. 2020.

[44] E. Elgeldawi, A. Sayed, A. R. Galal, and A. M. Zaki, "Hyperparameter tuning for machine learning algorithms used for Arabic sentiment analysis," *Informatics*, vol. 8, no. 4, p. 79, Nov. 2021.

[45] F. Abbas, F. Zhang, M. Ismail, G. Khan, J. Iqbal, A. F. Alrefaei, and M. F. Albeshr, "Optimizing machine learning algorithms for landslide susceptibility mapping along the Karakoram highway, Gilgit Baltistan, Pakistan: A comparative study of baseline, Bayesian, and metaheuristic hyperparameter optimization techniques," *Sensors*, vol. 23, no. 15, p. 6843, Aug. 2023.

[46] A. Kassambara, *Machine Learning Essentials: Practical Guide in R*. Marseille, France: STHDA, 2018.

[47] H. Huang, S. S. Band, H. Karami, M. Ehteram, K.-W. Chau, and Q. Zhang, "Solar radiation prediction using improved soft computing models for semi-arid, slightly-arid and humid climates," *Alexandria Eng. J.*, vol. 61, no. 12, pp. 10631–10657, Dec. 2022.

[48] S. El Bakali, O. Hamid, and S. Gheouany, "Day-ahead seasonal solar radiation prediction, combining VMD and STACK algorithms," *Clean Energy*, vol. 7, no. 4, pp. 911–925, Aug. 2023.

[49] E. M. El-Kenawy, S. Mirjalili, S. S. M. Ghoneim, M. M. Eid, M. El-Said, Z. S. Khan, and A. Ibrahim, "Advanced ensemble model for solar radiation forecasting using sine cosine algorithm and Newton's laws," *IEEE Access*, vol. 9, pp. 115750–115765, 2021, doi: 10.1109/ACCESS.2021.3106233.

[50] R. M. A. Ikram, H.-L. Dai, A. A. Ewees, J. Shiri, O. Kisi, and M. Zounemat-Kermani, "Application of improved version of multi verse optimizer algorithm for modeling solar radiation," *Energy Rep.*, vol. 8, pp. 12063–12080, Nov. 2022. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85141872972&doi=10.1016%2fj.egyr.2022.09.015&partnerID=40&md5=cf0f6e915d9ae883084e34733ac4a275

[51] J. Fan, L. Wu, X. Ma, H. Zhou, and F. Zhang, "Hybrid support vector machines with heuristic algorithms for prediction of daily diffuse solar radiation in air-polluted regions," *Renew. Energy*, vol. 145, pp. 2034–2045, Jan. 2020.

[52] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, and S. E. Brennan, "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *Int. J. surgery*, vol. 88, Mar. 2020, Art. no. 105906.

[53] Y. Zhang, N. Cui, Y. Feng, D. Gong, and X. Hu, "Comparison of BP, PSO-BP and statistical models for predicting daily global solar radiation in arid Northwest China," *Comput. Electron. Agricult.*, vol. 164, Sep. 2019, Art. no. 104905.

[54] T. Xu, M. H. Sabzalian, A. Hammoud, H. Tahami, A. Gholami, and S. Lee, "An innovative machine learning based on feed-forward artificial neural network and equilibrium optimization for predicting solar irradiance," *Sci. Rep.*, vol. 14, no. 1, p. 2170, Jan. 2024.

[55] X. Zhang and Z. Wei, "A hybrid model based on principal component analysis, wavelet transform, and extreme learning machine optimized by bat algorithm for daily solar radiation forecasting," *Sustainability*, vol. 11, no. 15, p. 4138, 2019. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85070458783&doi=10.3390%2fsu11154138&partnerID=40&md5=9b80eab2fe1626afe9739ac173f5f829

[56] S. Ghimire, R. C. Deo, D. Casillas-Pérez, S. Salcedo-Sanz, E. Sharma, and M. Ali, "Deep learning CNN-LSTM-MLP hybrid fusion model for feature optimizations and daily solar radiation prediction," *Measurement*, vol. 202, Oct. 2022, Art. no. 111759. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85136630745&doi=10.1016%2fj.measurement.2022.111759&partnerID=40&md5=46708f4a08dcaf6d4ec21fea82f19e27

[57] S. Ghimire, B. Bhandari, D. Casillas-Pérez, R. C. Deo, and S. Salcedo-Sanz, "Hybrid deep CNN-SVR algorithm for solar radiation prediction problems in queensland, Australia," *Eng. Appl. Artif. Intell.*, vol. 112, Jun. 2022, Art. no. 104860. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85128175831&doi=10.1016%2fj.engappai.2022.104860&partnerID=40&md5=2048c39a31f83c0704ed1c350876a9e0

[58] K. Irshad, N. Islam, A. A. Gari, S. Algarni, T. Alqahtani, and B. Imteyaz, "Arithmetic optimization with hybrid deep learning algorithm based solar radiation prediction model," *Sustain. Energy Technol. Assessments*, vol. 57, Jun. 2023, Art. no. 103165. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2213138823001583

[59] M. Guermoui, T. Arrif, A. Belaid, S. Hassani, and N. Bailek, "Enhancing direct normal solar irradiation forecasting for heliostat field applications through a novel hybrid model," *Energy Convers. Manage.*, vol. 304, Mar. 2024, Art. no. 118189.

[60] D. S. Khafaga, A. A. Alhussan, M. M. Eid, and E.-S. M. El-kenawy, "Improving solar radiation source efficiency using adaptive dynamic squirrel search optimization algorithm and long short-term memory," *Frontiers Energy Res.*, vol. 11, Jul. 2023, Art. no. 1164528. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85165185758&doi=10.3389%2ffenrg.2023.1164528&partnerID=40&md5=ff9967569ad49cbfa93d033f1d9f048a

[61] S. Sadiq Hasan, Z. Salih Agee, B. Shamsaldeen Tahir, and S. R. M. Zeebaree, "Solar radiation prediction using satin bowerbird optimization with modified deep learning," *Comput. Syst. Sci. Eng.*, vol. 46, no. 3, pp. 3225–3238, 2023. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85158834421&doi=10.32604%2fcsse.2023.037434&partnerID=40&md5=d2d5982a0eb9a1af1bd349cad70404ce

[62] N. Y. Jayalakshmi, R. Shankar, U. Subramaniam, I. Baranilingesan, A. Karthick, B. Stalin, R. Rahim, and A. Ghosh, "Novel multi-time scale deep learning algorithm for solar irradiance forecasting," *Energies*, vol. 14, no. 9, p. 2404, Apr. 2021.

[63] V. Sansine, P. Ortega, D. Hissel, and M. Hopuare, "Solar irradiance probabilistic forecasting using machine learning, metaheuristic models and numerical weather predictions," *Sustainability*, vol. 14, no. 22, p. 15260, Nov. 2022.

[64] K. N. Reddy, M. Thillaikarasi, B. S. Kumar, and T. Suresh, "A novel elephant herd optimization model with a deep extreme learning machine for solar radiation prediction using weather forecasts," *J. Supercomput.*, vol. 78, no. 6, pp. 8560–8576, Apr. 2022. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85122682242&doi=10.1007%2fs11227-021-04244-y&partnerID=40&md5=be2fd363ea7f690054dca39c40e95096

[65] Y. Dong and H. Jiang, "Global solar radiation forecasting using square root regularization-based ensemble," *Math. Problems Eng.*, vol. 2019, May 2019, Art. no. 9620945. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85066324914&doi=10.1155%2f2019%2f9620945&partnerID=40&md5=123c962bd752c4dab4719714322ce2bb

[66] T. Peng, Y. Li, Z. Song, Y. Fu, M. S. Nazir, and C. Zhang, "Hybrid intelligent deep learning model for solar radiation forecasting using optimal variational mode decomposition and evolutionary deep belief network–Online sequential extreme learning machine," *J. Building Eng.*, vol. 76, Oct. 2023, Art. no. 107227. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85164216027&doi=10.1016%2fj.jobe.2023.107227&partnerID=40&md5=6e8c6600b8e81b8d0899d84d8a6c77ce

[67] Z. Wu, N. Cui, D. Gong, F. Zhu, Y. Li, L. Xing, Z. Wang, B. Zhu, X. Chen, S. Wen, and Y. Zha, "Predicting daily global solar radiation in various climatic regions of China based on hybrid support vector machines with meta-heuristic algorithms," *J. Cleaner Prod.*, vol. 385, Jan. 2023, Art. no. 135589. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85144627742&doi=10.1016%2fj.jclepro.2022.135589&partnerID=40&md5=a77f34c054f082d196a45ddc58c752f9

[68] X. Xue and H. Zhou, "Soft computing methods for predicting daily global solar radiation," *Numer. Heat Transf., B, Fundamentals*, vol. 76, no. 1, pp. 18–31, Jul. 2019. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85068799815&doi=10.1080%2f10407790.2019.1637629&partnerID=40&md5=bf24f90990766665435b87ef1ef16dcd

[69] S. Sun, S. Wang, G. Zhang, and J. Zheng, "A decomposition-clustering-ensemble learning approach for solar radiation forecasting," *Sol. Energy*, vol. 163, pp. 189–199, Mar. 2018.

[70] M. Guermoui, K. Gairaa, J. Boland, and T. Arrif, "A novel hybrid model for solar radiation forecasting using support vector machine and bee colony optimization algorithm: Review and case study," *J. Sol. Energy Eng.*, vol. 143, no. 2, pp. 20801–20817, Apr. 2021.

[71] M. Khodayar, S. Mohammadi, M. E. Khodayar, J. Wang, and G. Liu, "Convolutional graph autoencoder: A generative deep neural network for probabilistic spatio-temporal solar irradiance forecasting," *IEEE Trans. Sustain. Energy*, vol. 11, no. 2, pp. 571–583, Apr. 2020.

[72] A. Mirbolouki, S. Heddam, K. Singh Parmar, S. Trajkovic, M. Mehraein, and O. Kisi, "Comparison of the advanced machine learning methods for better prediction accuracy of solar radiation using only temperature data: A case study," *Int. J. Energy Res.*, vol. 46, no. 3, pp. 2709–2736, Mar. 2022. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85116013923&doi=10.1002%2fer.7341&partnerID=40&md5=b4fa2cb1cc12ef6ef32bf50ea488a749

[73] A. K. Pani and N. Nayak, "Forecasting solar irradiance with weather classification and chaotic gravitational search algorithm based wavelet kernel extreme learning machine," *Int. J. Renew. Energy Res.*, vol. 9, no. 4, pp. 1650–1659, 2019.

[74] Y. Jia, H. Wang, P. Li, Y. Su, F. Wang, and S. Huo, "Particle swarm optimization algorithm with Gaussian exponential model to predict daily and monthly global solar radiation in Northeast China," *Environ. Sci. Pollut. Res.*, vol. 30, no. 5, pp. 12769–12784, Sep. 2022.

[75] M. M. Islam, M. Nagrial, J. Rizk, and A. Hellany, "Solar radiation and wind speed forecasting using deep learning technique," in *Proc. IEEE Asia–Pacific Conf. Comput. Sci. Data Eng. (CSDE)*, 2021, pp. 1–6, doi: 10.1109/CSDE53843.2021.9718372.

[76] S. Y. Heng, W. M. Ridwan, P. Kumar, A. N. Ahmed, C. M. Fai, A. H. Birima, and A. El-Shafie, "Artificial neural network model with different backpropagation algorithms and meteorological data for solar radiation prediction," *Sci. Rep.*, vol. 12, no. 1, p. 10457, Jun. 2022.

[77] S. N. Langazane and A. K. Saha, "Effects of particle swarm optimization and genetic algorithm control parameters on overcurrent relay selectivity and speed," *IEEE Access*, vol. 10, pp. 4550–4567, 2022.

[78] T. Zhu, Y. Guo, Z. Li, and C. Wang, "Solar radiation prediction based on convolution neural network and long short-term memory," *Energies*, vol. 14, no. 24, p. 8498, Dec. 2021. [Online]. Available: https://www.mdpi.com/1996-1073/14/24/8498

SERGIO A. PÉREZ-RODRÍGUEZ received the B.S. degree in computer systems from Instituto Tecnológico de Querétaro, Mexico, in 2021. He is currently pursuing the master's degree in computer science with the Faculty of Informatics, Universidad Autónoma de Querétaro. He is a Professor with the Faculty of Informatics, Universidad Autónoma de Querétaro. His research interests include machine learning, deep learning, metaheuristic algorithms, data science, and cloud computing.

JOSÉ M. ÁLVAREZ-ALVARADO received the B.S. degree in mechatronics from the Instituto Tecnológico de La Laguna, Coahuila, Mexico, in 2013, and the M.Sc. degree in automatic control and the Ph.D. degree from the School of Engineering, Universidad Autónoma de Querétaro, Mexico, in 2017 and 2021, respectively. He is currently a Professor with the School of Engineering, Universidad Autónoma de Querétaro. His research interests include machine learning techniques, renewable energy prediction, energy sustainability, and energy generation systems. He is a member of the National Researcher System of CONACYT.

JULIO-ALEJANDRO ROMERO-GONZÁLEZ received the degree in automation engineering and the master's degree in instrumentation and automatic control from the Faculty of Engineering, Autonomous University of Querétaro, in 2012 and 2014, respectively, and the Ph.D. degree in computer science from the Faculty of Informatics, in 2022. He is currently a Research Professor with the Faculty of Informatics, Autonomous University of Querétaro. His main research interests include vision and artificial intelligence, embedded systems, and automatic control.

MARCOS AVILES (Member, IEEE) received the bachelor's degree in biomedical engineering, the master's degree in instrumentation and control, and the Ph.D. degree from the Doctorate in Engineering Program, Autonomous University of Querétaro (UAQ). Currently, he is a Teacher with the Biomedical Engineering Degree Program. His areas of expertise include signal processing, artificial intelligence, and using metaheuristic algorithms.

AMERICA EILEEN MENDOZA-ROJAS received the bachelor's degree in mechanical engineering from Universidad de Guanajuato and the Master of Science degree (Hons.) in nanotechnology from Universidad Autónoma de Querétaro. She is currently pursuing the Ph.D. degree in applied engineering from Colegio de Formación Educativa Tenam. She has worked with different institutes: Centro Mexicano de Energías Renovables (CEMER SC), Centro de Física Aplicada and Tecnología Avanzada (CFATA), UNAM, and also in the industrial sector with Recicla.Lo, Guanajuato, Mexico. She is a Professor with the Industrial Division and the IT Division, Universidad Politécnica de Querétaro, Mexico. She has several publications in national and international congresses focused on different areas: alternative and mathematics applied to solve engineering problems, renewable energy, optimizing, alternative fuels, nanomaterials, and nanomedicine. She has received awards in Scientific and Technological Innovation (Regional Government) and Energetic Efficiency (by the Instituto Nacional de Electricidad y Energías Limpias, INEEL), in a regional level and national level in Mexico, respectively.

CARLOS FUENTES-SILVA received the M.S. and Ph.D. degrees in intelligent control from Universidad Autónoma de Querétaro (UAQ), Mexico. He is currently a Professor with the Manufacturing Technology Engineering Educational Program, Universidad Politécnica de Querétaro (UPQ), Mexico. He is working on autonomous guided vehicles with computer vision for road, pedestrians, and traffic signal detection, generating navigation reports to a digital dashboard over the Internet of Things (IoT). His main research interests include fuzzy decision-making, neural networks, machine learning, and computer vision.

JUVENAL RODRÍGUEZ-RESÉNDIZ (Senior Member, IEEE) was with West Virginia University as a Visiting Professor, in 2012. Currently, he is the coordinator for doctorate in engineering with Querétaro State University (UAQ), Mexico. Herein, he has taught more than 200 digital signal processing and research methodology courses. He belongs to Mexican Academy of Sciences, Mexican Association of Robotics and Mechatronics, and the National Research Academy in Mexico. He has a higher membership for researchers, called SNII level three, given by Mexican Government. He has developed more than 50 industrial projects by linking UAQ and the government. His team has published 180 technical articles. He patented more than ten innovations. He has won national and international prizes because of his academic and innovation developments. He has been the advisor of more than 200 theses of undergraduate, master, and Ph.D. grades. He has been invited to give 30 conferences around the world. Because of his research, he collaborates with 30 national and international institutions.