IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 4 March 2024, accepted 20 March 2024, date of publication 22 March 2024, date of current version 28 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3381049

RESEARCH ARTICLE

An Advanced Hybrid Boot-LSTM-ICSO-PP Approach for Day-Ahead Probabilistic PV Power Yield Forecasting and Intra-Hour Power Fluctuation Estimation

IOANNIS K. BAZIONIS[1], MARKOS A. KOUSOUNADIS-KNOUSEN[1], VASILEIOS E. KATSIGIANNIS[1], FRANCKY CATTHOOR[2,3], (Fellow, IEEE), AND PAVLOS S. GEORGILAKIS[1], (Senior Member, IEEE)

[1]School of Electrical and Computer Engineering, National Technical University of Athens (NTUA), 15780 Athens, Greece
[2]IMEC, 3001 Leuven, Belgium
[3]KULeuven, Leuven, 3001 Heverlee, Belgium

Corresponding author: Pavlos S. Georgilakis (pgeorg@power.ece.ntua.gr)

This work was supported in part by European Regional Development Fund of European Union; and in part by the Greek National Funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the Call RESEARCH—CREATE—INNOVATE under Project T2EDK-00864.

ABSTRACT Probabilistic forecasting models have been developed over the past years in order to aid in the estimation of the uncertainty of the predictive results. A hybrid, bootstrapping long-short term memory (Boot-LSTM)-based model is proposed in this paper, in order to construct accurate prediction intervals (PIs) for short-term solar power generation. A novel approach that introduces an improved chicken swarm optimization (ICSO) algorithm along with a prey-predator (PP) mechanism is developed in order to optimize the predictive accuracy. Exploiting the ICSO's ability to optimize the position of the swarm's particles as well as the PP's ability to further improve the particles' searching possibilities, the weights and biases of the neurons of the neural network (NN) of the model are optimized and the predictive accuracy is further improved. The accuracy of the PIs is evaluated by minimizing the coverage width criterion (CWC) cost function. The efficiency and the accuracy of the proposed hybrid Boot-LSTM-ICSO-PP model is confirmed via comparing the predictive outputs with state-of-the-art methodologies considering probabilistic evaluation metrics. The proposed model was applied on two datasets of existing solar parks and was further analyzed from a seasonal perspective, in order to prove its efficiency with real-life cases. In terms of CWC minimization, the proposed model, for the first PV park, achieves a 60.3% and 46.94% average improvement compared to the base BELM and LSTM models, respectively, while for the second PV park, achieves a 50.64% and 37.87% average improvement to the respective base models as well.

INDEX TERMS Solar power forecasting, long short-term memory, bootstrap, prediction intervals, chicken swarm optimizer, seasonal analysis.

NOMENCLATURE

ACRONYMS & ABBREVIATIONS

BELM	Bootstrap Extreme Learning Machine.
CSO	Chicken Swarm Optimization.
CWC	Coverage Width Criterion.
ELM	Extreme Learning Machine.
ICSO	Improved Chicken Swarm Optimization.
LSTM	Long Short-Term Memory.
NN	Neural Network.
PI	Prediction Interval.
PICP	Prediction Interval Coverage Probability.

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

PINAW	Prediction Interval Normalized Average Width.
PINC	Prediction Interval Nominal Confidence.
PP	Prey-Predator.
PV	Photovoltaic
RES	Renewable Energy Sources.
RNN	Recurrent Neural Network.
SPF	Solar Power Forecasting
SPPF	Solar Power Probabilistic Forecasting.

PARAMETERS

G	Update time of swarm.
N_C	Number of chicks in swarm.
$N_{chickens}$	Number of chickens in swarm.
N_D	Dataset length.
N_h	Number of hens in swarm.
N_m	Number of mother hens in swarm.
N_r	Number of roosters in swarm.
m	number of output layer neurons.
n	number of input layer neurons.
T	maximum number of generations.

FUNCTIONS

f	Activation function of hidden layer.

VARIABLES

b	biases of hidden neurons.
$\hat{q}^{\bar{a}}/\hat{q}^{\underline{a}}$	Upper/Lower bound of PI 1-γ.
t_j	Real-targeted output.
w	weights that connect input with hidden layer.
$\hat{y}\left(x_j\right)$	Estimated output.
β	weights that connect hidden with output layer.

I. INTRODUCTION

A. STATE OF PV POWER FORECASTING OVERVIEW AND LITERATURE REVIEW

Energy needs on a global level keep increasing daily. Nowadays, power supply and management has become a twofold problem. Apart from satisfying the continuously increasing global energy needs efficiently, an environment-friendly way of power production has become mandatory. Aiming to decrease the dependance on conventional sources of energy, Renewable Energy Sources (RES) have been used widely over the last decade [1]. Therefore, efficient exploitation of RES as well as management of their penetration into power systems has been the focus of researchers over the previous years.

Solar power is one of the most widely exploited RES. Considering the fact that sunlight can be an abundant energy source, solar power is one of the most important and highly implemented sources of energy into power systems [2]. Thanks to technological advances in recent years, solar power can be efficiently exploited via solar photovoltaic (PV) systems.

The continuously rising exploitation of solar power, while providing a ''green'' alternative to conventional energy sources, can create various problems in terms of smooth penetration into energy systems, grid stability and energy management. The amount of incoming solar radiation reaching photovoltaic panels has a direct impact on how much solar power can be produced [3]. However, a number of variables can limit the specific solar irradiance at a given moment [4]. Since solar energy can only be produced during the day and also due to the presence of passing clouds, time is the primary factor. Furthermore, both on a global and local scale, the distribution of solar irradiance cannot be uniform. Additionally, a number of meteorological parameters, including the ambient temperature as well as the temperature on the PV modules, the wind's direction and speed and humidity levels, highly affect solar power generation [5].

Maintaining the stability of the power grid and proper management of electricity markets are matters of core importance that are directly affected by the penetration of solar power into power systems [6]. In order to address those issues, different Solar Power Forecasting (SPF) models have been introduced over the past years and keep developing. Numerous SPF approaches have been studied in an effort to regulate the rising solar power penetration in the world's power grids. Aiming to further improve the control of energy markets, such forecasting models also assist in the scheduling of power systems and in preserving their stability and reliability [7]. Given their significance in the organization and management of daily electricity markets and power systems, short-term SPF models, with a focus of one hour to day-ahead forecasting horizon, have been the main interest of researchers.

Such forecasting models, mainly based on Neural Network (NN) methodologies and architectures, are able to generate point forecasts in order to provide the user with an estimated output series. However, because such models do not include information over the uncertainty of the prediction, they cannot always be as efficient for real-life cases. Therefore, probabilistic forecasting models have been researched and developed over the last years [8]. Contrary to point forecasting models, probabilistic forecasting models offer a wider perspective of the predictive outcome since the predictive information is provided in the form of prediction intervals (PIs), quantile distributions or predictive scenarios [9].

Several different solar power probabilistic forecasting (SPPF) models have been developed over the years. In [10], an SPPF model based on weather scenario generation was proposed and predictive quantiles were constructed through the empirical distribution of the scenarios. In [11], a non-parametric density forecasting model was proposed, which applied an Extreme Learning Machine (ELM) model for the regression process. The work [12] proposed a quantile regression based SPPF model, which was further optimized via a Bayesian bootstrapping technique. In [13], a Long Short-Term Memory (LSTM)-based model was proposed for solar irradiance forecasting, along with residual modeling. In work [14], probabilistic forecasts were obtained

TABLE 1. Summary of the main characteristics of the literature review.

Reference	Methodology	Main Characteristics	Forecasting Type
[10]	Weather Scenario Generation	- Considers inherent correlation among different weather variables. - Use of Copula for model correlation among weather variables.	Probabilistic
[11]	BELM	- Proposes a nonparametric density forecasting method based on ELM as a regression tool to avoid restrictive assumptions on the shape of the forecast densities. - Provides reliable and sharp predictive densities.	Deterministic/Probabilistic
[12]	Quantile Regression	- Focuses on the interaction between quantile regression and Bayesian bootstrap.	Probabilistic
[13]	LSTM	- Combines deep neural networks with residual modeling. - Captures the dependencies among the consecutive samples in time series.	Probabilistic
[14]	Ensemble	- Focuses on extending the data used in analog ensemble models. - Avoids using time consuming optimizers. - Applies Kernel Density Estimation to produce probabilistic forecasts from the analog ensemble.	Probabilistic
[15]	VMD-LSTM	- Proposes a MVD-based data pre-processing method. - Eliminates seasonal factors.	Deterministic
[16]	NP-CNN-LSTM	- Focuses on providing a hybrid model more robust and with more powerful fitting capability. - Effectively captures the seasonality in PV power data. - Efficiently exploits the abilities of deep learning LSTM model.	Deterministic
[17]	ISt-LSTM-Informer	- Aims in achieving accurate multi-timescale PV power forecasts. - Improves k-fold cross validation to avoid information loss during time-series integration.	Deterministic
[18]	Quantile Regression Forests	- Investigates the benefits to forecasting accuracy from co-locating wind and PV parks. - Analyzes the smoothing effect on the probabilistic predictive accuracy and the electricity market participation.	Probabilistic
[19]	Hybrid	- Proposes a hybrid model that combines deterministic forecasting along with uncertainty analysis. - Includes a multivariate forecasting model for extending the dimensionality of the available data.	Deterministic/Probabilistic
[20]	PSO-BP, GA-BP	- Focuses on comparative forecasting by implementing different optimization algorithms.	Deterministic

via a kernel density estimation-based model. The work [15] proposed a hybrid LSTM-based model for accurate load forecasting considering complex consumption data. In [16], a hybrid LSTM-based model is proposed for day-ahead PV power prediction in large-scale systems. The work [17] proposed an improved Stacking ensemble algorithm-based LSTM-Informer model (ISt-LSTM-Informer), aiming to achieve accurate multi-timescale PV power forecasting. The work [18] analyzes the effect of aggregating varying shares of installed wind power and PV power in a co-located park using calibrated probabilistic forecasts generated using a fixed modelling framework. In [19], a hybrid PV power forecasting model was proposed, which combined deterministic forecasting and uncertainty analysis for complete and accurate PV power forecasts. The work [20], evaluates the efficiency of three backpropagation models in PV power forecasting. Table 1 summarizes the main characteristics of the above reviewed works as well as limitations and spaces for improvement.

Solar power forecasting constitutes a dynamic environment for researchers. As can be seen from Table 1, various methodologies focus on efficient deterministic and probabilistic forecasting. However, the understanding of the prediction process of the uncertainty and its implementation to real life problems is becoming more and more relevant. Therefore, the need to evaluate models that combine deterministic and probabilistic forecasting is a sector with high potential for improvement and understanding. Furthermore, optimization algorithms that focus on improving the forecasting errors are usually applied on a single-dimension framework, where they either focus on the preprocessing part of the forecast or on the recalculation of the model's parameters. Considering swarm type optimizers, while able to broaden the searching abilities of the forecasting models, they tend to be based on simple single population swarms, easily trapped in local optima. Specifically, for solar power probabilistic forecasting, the application of advanced swarm-type optimizers is extremely limited. Moreover, the use of the state-of-the-art LSTM network has become widely popular over the last years, owing to its great ability to handle time series data. However, LSTM itself, having a training process based on backpropagation with gradient descent may even be more prone to getting trapped in local optima. Implementing resampling techniques for probabilistic power forecasting along with LSTM has not been actively researched, considering that such techniques tend to mix the timeseries and therefore their sequentiality

is lost. However, in the case of PV power generation where we have a daily recurrence, the timeseries can be ''broken'' into daily samples and we can apply resampling on them. Therefore, combining the advantages of resampling techniques such as bootstrap, along the LSTM model becomes a viable option.

B. CONTRIBUTIONS OF THE PRESENTED WORK

Aiming to cover the limitations presented above, this paper focuses on constructing accurate PIs through developing an SPF model under the probabilistic spectrum, for a short-term, day-ahead forecasting horizon. Considering the strong time-series nature of the problem evaluated in this paper, LSTM as a base model is a clear choice, combined with a resampling technique in order to be applied for probabilistic forecasting. Aiming to further improve the searching capabilities of deep learning, and therefore the selected LSTM model, by ensuring global searching without compromising the computational efficiency, an advanced, metaheuristic, improved optimization algorithm is proposed and applied. To further expand the global searching ability of the proposed algorithm and to eliminate as much as possible the possibility of it being trapped in local optima, another meta-heuristic mechanic is applied. Furthermore, considering the increasing need to evaluate and implement the uncertainty of the forecasting process in real life problems, while also following recent research trends, deterministic forecasts as well as uncertainty predictions are being taken into account, to provide a complete image of the predictive results to the reader.

The evaluation of the proposed hybrid model is determined by the minimization of the Coverage Width Criterion (CWC) cost function, which simultaneously optimizes the sharpness and the calibration of the produced PIs. Intra-hour deterministic forecasts are computed and qualitatively evaluated by the final proposed probabilistic model, in order to estimate possible power fluctuations. The provided results are further analyzed from a seasonal perspective. The presented work successfully verifies the application of the improved optimization mechanisms for probabilistic solar power forecasting. Furthermore, along with the proposed application of the PP mechanism, it manages to ''challenge'' state-of-the-art forecasting methodologies.

The contributions of this paper are manifold:

a) A hybrid probabilistic Boot-LSTM model is constructed, taking advantage of the Bootstrap method's resampling ability and the LSTM model's architecture, which allows it to capture sequential dependencies and complex relationships within the time-series data. The proposed forecasting model is developed for accurate PI construction and efficient day-ahead PV power yield estimation. An advanced metaheuristics algorithm is further applied for the improvement of the hybrid model's performance.

b) An Improved version of the Chicken Swarm Optimizer (ICSO) algorithm is further introduced in order to optimize the weights and biases of the NN. To the best of our knowledge, the CSO, let alone and improved extension of the CSO, has not been applied to probabilistic SPF before. By including an improved way to update the position of the swarm's particles, the swarm's best particles' local and global search ability strengthens while avoiding getting caught in a local minimum. A prey-predator algorithm is adjusted and carefully co-optimized to further improve the CSO. The prey-predator algorithm serves as an escaping mechanism that encourages the swarm's particles to exploit promising areas or new regions of the search space, which allows better convergence and higher quality predictions. The final proposed hybrid Boot-LSTM model, with the implementation of the adjusted PP, manages to: i) provide accurate PIs of high quality, and ii) outperform state-of-the-art methodologies in terms of predictive errors and accuracy.

c) It considers intra-hour spot forecasts of PV power in order to provide the Boot-LSTM-ICSO-PP with a qualitative evaluation tool of the possible day-ahead power fluctuations. In turn, the Boot-LSTM-ICSO-PP model is able to efficiently use the above information to construct more accurate PIs.

d) The efficiency of the model is tested on two different datasets from two real PV parks. The results of the proposed model are evaluated and compared to state-of-the-art methodologies applied on the same datasets. The accuracy of the proposed model is further confirmed by a seasonality analysis of the results.

The rest of the paper is organized as follows. Section II presents an overview of the proposed methodology, Section III describes the architecture and the innovation of the proposed model, Section IV includes the evaluation metrics used to estimate the efficiency of the proposed model, Section V presents the case study used in this paper, Section VI analyses the results and Section VII concludes the paper.

II. INDIVIDUAL METHODS

This section describes the individual methods that will be used in Section III, in order to obtain the proposed hybrid forecasting methodology of this paper.

A. CO-OPTIMIZED HYBRID PROBABILISTIC WPF MODEL

The model used in our work is LSTM, which is a type of Recurrent Neural Network (RNN) that can learn order dependence in sequence prediction problems by preserving previous information and establishing temporal correlations between sequential data with internal self-looped repeating networks [21]. LSTM contains four connected layers, three gate layers and a tanh layer. Cell state is a core variable that can run straight down the architecture, carrying information of previous steps [21]. The architecture of a typical LSTM cell is presented in Figure 1. The LSTM is capable of removing or adding information to the cell state, regulated by gates.

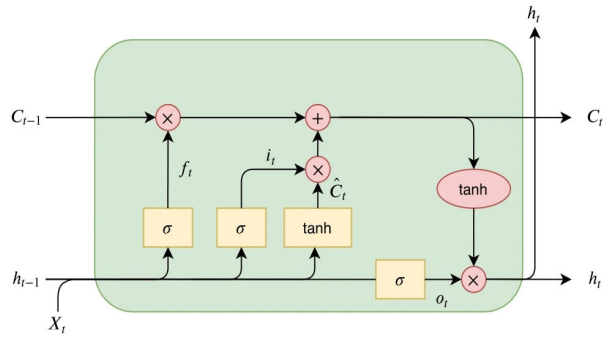


FIGURE 1. Typical LSTM cell architecture.

The first layer is the forget layer that decides which information of previous steps to forget.

The mathematical equation for the output of the forget gate (f_t) is given in (1):

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, X_t] + b_f\right) \tag{1}$$

where σ is the sigmoid activation function, W_f is the weight of the forget gate, b_f is the bias of the forget gate, X_t is the input at time t and h_{t-1} is the hidden layer output at time t-1.

The second layer is the input gate (i_t), which decides what new information will be stored in the cell state by using:

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, X_t] + b_i\right) \tag{2}$$

where W_i is the weight of the input gate and b_i is the bias of the input gate.

The third layer is the tanh layer or cell state layer ($\breve{C}_t$). It generates a vector with new candidate values, defined in (3):

$$\breve{C}_t = \varphi\left(W_C \cdot [h_{t-1}, X_t] + b_C\right) \tag{3}$$

where φ is the tanh function, W_C is the weight of the cell and b_C is the bias of the cell.

After the first three layers, the old cell state C_{t-1} is updated by C_t. The update comes from the combination of the output of forget gate and input gate, the first one determines what to forget and the second one determines what to add to the new cell state, as shown in (4):

$$C_t = f_t C_{t-1} + i_t \breve{C}_t \tag{4}$$

The final layer is the output gate, which generates the final output according to the updated cell. The process of the output gate is:

$$o_t = \sigma\left(W_o \cdot [h_{t-1}, X_t] + b_o\right) \cdot \varphi\left(C_t\right) \tag{5}$$

where W_o is the weight of the output gate and b_o is the bias of the output gate.

B. BOOTSTRAP

Bootstrap is regarded as a general approach of statistical inference based on building a sampling distribution by uniform sampling with replacements from the original data [22]. It is widely applied as a robust alternative to statistical inference based on the parametric assumptions, which can be unreliable and even impossible due to the sophistications involved in computing the standard errors in some conditions.

Three different bootstrap algorithms can be applied for regression analysis [23], including the pairs bootstrap, the standard residuals bootstrap, and the wild residuals bootstrap (wild bootstrap). Compared to the other bootstrap techniques, the pairs bootstrap technique is selected because it is more suitable for problems, such as SPF, where observations are correlated, and because it is further able to preserve the dependence structure between observations. For N arbitrary distinct samples $\{(x_i, t_i)\}_{i=1}^N$, where $x_i \in \boldsymbol{R}^n$ with $x_i = [x_{i1}, x_{i2}, \ldots, x_{in}]^T$, $t_i \in \boldsymbol{R}^m$ with $t_i = [t_{i1}, t_{i2}, \ldots, x_{im}]^T$ and $\hat{y}_l = (x_i)$ the prediction value of the input samples generated by the lth bootstrapped dataset, the pairs bootstrap can be applied according to the following steps:

Step 1) Obtain the training samples $\{(x_i, t_i)\}_{i=1}^N$.

Step 2) Generate bootstrapped pairs $\left\{\left(x_i^*, t_i^*\right)\right\}_{i=1}^N$ by uniform sampling with replacement from the original training data $\{(x_i, t_i)\}_{i=1}^N$.

Step 3) Estimate the model's $\hat{y}_l = \left(x_i^*\right)$ from the lth bootstrapped dataset $\left\{\left(x_i^*, t_i^*\right)\right\}_{i=1}^N$.

Step 4) Repeat steps 2) and 3) to obtain bootstrap replicates.

C. CHICKEN SWARM OPTIMIZATION (CSO)

The CSO optimizer is a group intelligent optimizer that simulates the hierarchical system and the foraging behavior of the chicken swarm [24]. The CSO is a metaheuristic optimization algorithm inspired by the social behaviour of chickens. Compared to other optimization algorithms, CSO appears to be more flexible, since it can be applied to various optimization problems, both single and multi-objective. The optimizer divides chickens into groups. Every group includes a rooster, several hens and a few chicks. In the CSO optimizer, the following rules are used to simulate the behaviors of the flock [25]:

(1) There are several subpopulations in the whole chicken population. Every subpopulation includes a rooster, multiple hens and chicks. Roosters have the strongest foraging ability and thus are dominant of the flock. Hens have the second-best foraging ability, and chicks have the worst foraging ability.

(2) The flocks are classified based on their fitness values S. A few of the chickens with good fitness values are selected as roosters and a few of the chickens with poor fitness are selected as chicks. The remaining chickens are selected as mother hens. The hens are arbitrarily added to a subgroup.

(3) The relationship between mother hens and chicks as well as the leadership relationship, remain unchanged under a specific hierarchy. But as the chicks grow, these states are updated every time G (G is the update time, which has a certain value).

(4) Hens follow the roosters in their group to forage and can steal food from other chicks. In their turn, the chicks follow the hens in order to look for food around them.

During the optimization process, the position of each chicken represents a feasible solution. Since each chicken

has different foraging capabilities, different chickens have different update strategies. Assuming that the search space of chickens is d, there are $N_{chickens}$ in total. There are N_c chicks, N_h hens and N_r roosters. The equation for updating the i-th rooster is:

$$Z_{i,j}^{t+1} = Z_{i,j}^{t} + Z_{i,j}^{t} randn\left(0, \sigma^2\right) \tag{6}$$

where,

$$\sigma^2 = \begin{cases} 1, & S_e \geq S_i \\ exp\left(\dfrac{S_e - S_i}{|S_i - \varepsilon|}\right), & S_e < S_i \end{cases} \tag{7}$$

where $randn(0,\sigma^2)$ represents the Gaussian distribution, S_e ($e \in [1, Nr]$, $i \neq e$) indicates a rooster different than the i-*th* rooster, ε is an infinitesimal value, which provides that the denominator of (7) is not zero.

The equation for updating the i-th hen is:

$$Z_{i,j}^{t+1} = Z_{i,j}^{t} + p1 \cdot randm\left(Z_{r1,j}^{t} - Z_{i,j}^{t}\right) + p2 \cdot randm\left(Z_{r2,j}^{t} - Z_{i,j}^{t}\right) \tag{8}$$

where,

$$\begin{cases} p1 = exp\left(\dfrac{S_i - S_{r1}}{ads\,(S_i + \varepsilon)}\right) \\ p2 = (S_{r2} - S_i) \end{cases} \tag{9}$$

where $randm$ is a random number between 0 and 1, $r1$ represents the rooster in the group where the i-*th* hen is located, r2 ($r1 \neq r2$) is the rooster in the other group.

In order to update the chick particles, the following equation is used:

$$Z_{i,j}^{t+1} = Z_{i,j}^{t} + GL\,(i) \cdot \left(Z_{mj}^{t} - Z_{i,j}^{t}\right) \tag{10}$$

where Z_{mj}^{t} is the mother chicken followed by the i-*th* chick and GL represents a random number between 0 and 2 [26].

III. PROPOSED HYBRID BOOT-LSTM-PP-ICSO METHOD AND INNOVATIONS

A. BOOTSTRAP LSTM (BOOT-LSTM)

In the proposed Boot-LSTM, bootstrapping helps improve the robustness of LSTM models by training them on multiple bootstrapped datasets. This approach is particularly useful in handling the variability of solar power generation data caused by changing weather conditions. In the specific case, the resampling process takes place on daily datasets, which function as bootstrapped samples. The reason for that is to preserve the temporal autocorrelation of the PV output observed during the day.

Furthermore, bootstrapping allows exploration of the behavior of the LSTM model on different resampled sequences of daily datasets, with different weather conditions. Through the continuous issue of different forecasts during the resampling process and the temporal interdependencies identified by the LSTM, quality PIs are generated that incorporate the inherent uncertainty of the day-ahead estimations. The model's weights are then re-adjusted via the ICSO-PP algorithm in order to improve the accuracy of the final constructed PI. The overall hybrid procedure can lead to better model performance and adaptability to the available data.

B. IMPROVED CSO (ICSO)

In order to further improve and optimize the accuracy of the LSTM-based methodology, a CSO was included.

Considering the use of the CSO, the total population of chickens is split into three different subpopulations of roosters, hens and chicks. It is noted that the relationships between the chickens of the swarm are updated every G number of generations. The above parameters are chosen based on the research and optimization process focused on the minimization of the CWC, which is defined in Equation (20).

Since the traditional CSO algorithm is able to find a local minimum, an improved CSO (ICSO) is implemented to expand the algorithm's searching capabilities. The rooster functions as the swarm's leader. When it gets stuck in a local minimum, it causes the swarm to miss the global one. Therefore, a cosine factor is included in the rooster's position update equation to strengthen the rooster's local and global search ability:

$$Z_{i,j}^{t+1} = Cip \cdot Z_{i,j}^{t} + Z_{i,j}^{t} \cdot randn\left(0, \sigma^2\right) \tag{11}$$

$$Cip = Cip_{min} + (Cip_{max} - Cip_{min}) \cdot \cos\left(\pi \cdot \frac{t}{T}\right) \tag{12}$$

where $Cip_{min} = 0.3$, $Cip_{max} = 0.8$ and T is the maximum number of generations.

Furthermore, the chick particles instead of following only the hens, they are affected by the swarm's best particle, so they can avoid getting caught in a local minimum. After the improvement, the position update equation of the ith chick is:

$$Z_{i,j}^{t+1} = Z_{i,j}^{t} + GL\,(i) \cdot \left(Z_{m,j}^{t} - Z_{i,j}^{t}\right) + BL\,(i) \cdot \left(Z_{best,j}^{t} - Z_{i,j}^{t}\right) \tag{13}$$

where BL is the learning coefficient and is computed as follows:

$$BL(i) = \exp\left(S_{best} - S_i\right) \tag{14}$$

Additionally, a Gaussian mutation operator was inserted, which shifts the chicken's position using the normal distribution around zero, and the variance by 0.1, as presented below:

$$Z_{i,j}^{t*} = Z_{i,j}^{t} + randn\left(\mu = 0, \sigma^2 = 0.1\right) \tag{15}$$

where $Z_{i,j}^{t}$ is the pre-mutated chicken, $Z_{i,j}^{t*}$ is the mutated chicken, and $randn\left(\mu = 0, \sigma^2 = 0.1\right)$ represents the normal distribution with zero mean value, and variance of 0.1.

Therefore, the ICSO algorithm is able to enhance the prediction performance of the LSTM model in PV power prediction, since, along with the PP mechanism, the exploitation is promoted by encouraging the swarm to explore promising regions of the search space that have already been explored,

i.e., it helps the LSTM fine-tune its parameters, and optimize the hidden nodes, weights and biases. Also, it helps the LSTM achieve better balance between overfitting and underfitting, since it maintains the diversity within the swarm, and thus it prevents premature convergences to solutions that correspond to local minimums.

C. PREY-PREDATOR (PP)

The prey-predator (PP) mechanism is included in order to further improve the algorithm's searching ability. The PP mechanism can help balance the exploration and exploitation abilities of the search process, which can lead to better convergence and higher quality predictions [27]. Specifically, the method encourages the swarm's particles to exploit promising areas of the search space that have already been explored, by moving a chicken to the position of one of the best fitness chickens (escaping mechanism).

The escaping mechanism of PP shifts the search focus of the swarm towards promising and already explored areas, while the breeding mechanism of PP releases particles that are trapped in local optima. By employing these mechanisms, PP achieves a tradeoff between convergence rate and global search. Also, the swarm is encouraged to explore new regions of the search space by initializing the chickens' positions that get caught by the predator. The above process can help prevent premature convergence and improve the robustness of the solutions, since it maintains diversity within the swarm, by preventing individuals from crowding too closely together and encouraging them to explore different regions of the search space.

PP proved to be ideal to further co-optimize the ICSO components, since it enhances global search while maintaining the convergence rate at high levels. As a swarmbased optimization algorithm, ICSO can still suffer from premature convergence. On the other hand, ICSO has extreme convergence rate capabilities. Furthermore, it is characterized by high complexity. Therefore, PP manages to achieve a trade-off between convergence rate and global search, while decreasing the required training time.

As presented in Figure 2, the prey-predator mechanism is applied right after the end of each ICSO algorithm generation. New members join the swarm, which is updated before the next ICSO iteration, and the searching field of the optimization algorithm is expanded. Also, current local optima are taken into account through the replacement of a chicken with a random best mechanism. This way, the algorithm avoids premature convergence to local optima, while increasing the chances of reaching the global optimum.

D. COMPLETE PROPOSED MODEL

Aiming to exploit the advantages of the aforementioned mechanisms, the architecture of the final model carefully implements their characteristics gradually. The proposed model's structure can be presented in the form of stages that deal with different parts of the problem. The deterministic forecasting process precedes the rest of the model's functions due to it being used as the initial set point for the model's parameters. The bootstrap mechanism is implemented afterwards, for the resampling of the input data in order for the probabilistic forecasting process to start, taking into account the parameters of the deterministic forecasting stage. Finally, the optimization algorithm consisting of the ICSO and the PP mechanism is implemented. During this stage, each chicken ($Z_{i,j}^{t}$) represents an array consisting of the weights and biases of each LSTM. Figure 3 presents the complete proposed model where the interaction between the different stages is presented.

Stage 1: In the initialization of the process, the model functions as a deterministic forecasting model in order to generate detailed day-ahead point predictions, as well as to estimate initial weights and biases values that will act as good starting points for the probabilistic forecasting model.

By introducing a deterministic stage in the whole model, the user is able to use the optimized weights and biases of the model in order to initially train the probabilistic forecasting model proposed in stages 2 and 3. Furthermore, the deterministic estimations can be further evaluated by the probabilistic model in order to improve the final prediction intervals. Moreover, a deterministic point of view, along with the probabilistic model, could provide to the user a more complete image of the predictive ability and credibility of the model.

Stage 2: Before the second stage, the input data are subjected to bootstrapping, where the resampling process is applied on daily datasets. The weights and biases estimated by the deterministic model are used to construct the initial PIs set.

This stage aims to take advantage of the bootstrapping mechanism's ability to locate and capture the sequential dependencies and different relationships of the input data. Furthermore, this stage takes advantage of the previously trained deterministic model and uses its optimized weights and biases for the initialization of the probabilistic model. The above results in faster convergence and training of the proposed model as well as improved predictive results.

Stage 3: During the final stage, the parameters of the LSTM model are re-optimized using the ICSO-PP algorithm, in order to improve the initial probabilistic predictions and construct an optimal final set of day-ahead PIs.

This stage exploits the advantages of the CSO algorithm and its function with the implementation of the prey predator mechanism. The ICSO algorithm introduced in the proposed methodology is an advanced optimized version of the simple CSO algorithm. As a swarm-type algorithm, the ICSO considers more than one population for the training data, which, while it makes the training process more complex, greatly improves the particles' global search ability. The implementation of the prey-predator mechanism enhances the searching capabilities of the ICSO since it allows the swarm's particles to increase the explored search space and to avoid getting trapped in local optima, thus contributing the best possible to the final predictive results.

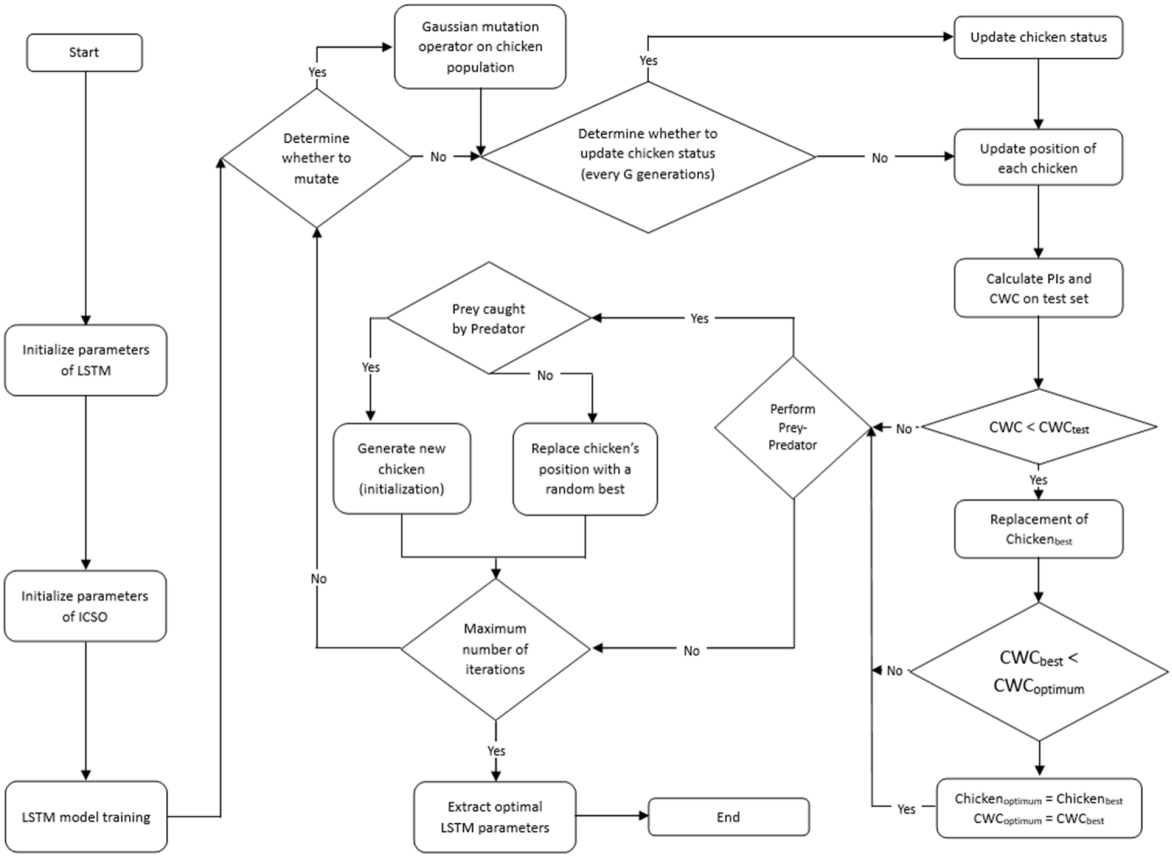


FIGURE 2. ICSO model functionality and prey-predator implementation.

TABLE 2. CWC optimization for the comparative models for PV Park 1.

Test day	CL (%)	CWC (%)						
		BELM	LSTM	Boot-LSTM	Boot-LSTM-CSO	Boot-LSTM-ICSO	BELM-ICSO-PP	Boot-LSTM-ICSO-PP
W/Sun	90	0.7234	0.4973	0.3682	0.3219	0.3086	0.2957	**0.2549**
	95	0.8692	0.5102	0.4537	0.4136	0.3618	0.3753	**0.3276**
	99	1.2931	0.5934	0.4928	0.4538	0.4021	0.4901	**0.3971**
W/LC	90	0.5937	0.4365	0.3216	0.2846	0.2812	0.2585	**0.2401**
	95	0.3889	0.4618	0.3925	0.3596	0.3269	0.3163	**0.3137**
	99	0.5532	0.4731	0.4481	0.4001	0.3781	0.4149	**0.3497**
W/HC	90	0.4293	0.3299	0.2821	0.2543	0.2103	0.2315	**0.1218**
	95	0.4690	0.3518	0.2850	0.2621	0.2276	0.2537	**0.1537**
	99	0.6173	0.3574	0.2953	0.2743	0.2453	0.2943	**0.2134**
S/Sun	90	0.5224	0.3532	0.2927	0.2437	0.2398	**0.1954**	0.2021
	95	0.5509	0.3945	0.3267	0.2814	0.2586	0.2407	**0.2348**
	99	0.6241	0.4474	0.4406	0.3608	0.3124	**0.2606**	0.2891
S/LC	90	0.4964	0.4220	0.3895	0.3491	0.2634	0.2143	**0.1369**
	95	0.5043	0.4881	0.4628	0.3939	0.2941	0.2389	**0.1713**
	99	0.7805	0.5007	0.4934	0.3987	0.3165	0.3205	**0.2894**
S/HC	90	0.4108	0.3598	0.2318	0.2024	0.1847	0.2109	**0.1246**
	95	0.4897	0.3718	0.3259	0.2238	0.2196	0.2481	**0.1482**
	99	0.5075	0.4107	0.3711	0.3187	0.2781	0.2882	**0.2234**

IV. FORECAST EVALUATION METRICS

The evaluation of the constructed PIs is a result of the estimation of its calibration and its sharpness [28]. The following metrics are commonly used to compute the calibration and sharpness of the PI.

PI Coverage Probability (PICP) is the measure related to the quality of the constructed PIs. It represents the percentage of observations (yt) found between the upper bounds and lower bounds of all observations. The larger the PICP, the more targets are supposed to be found in the corresponding

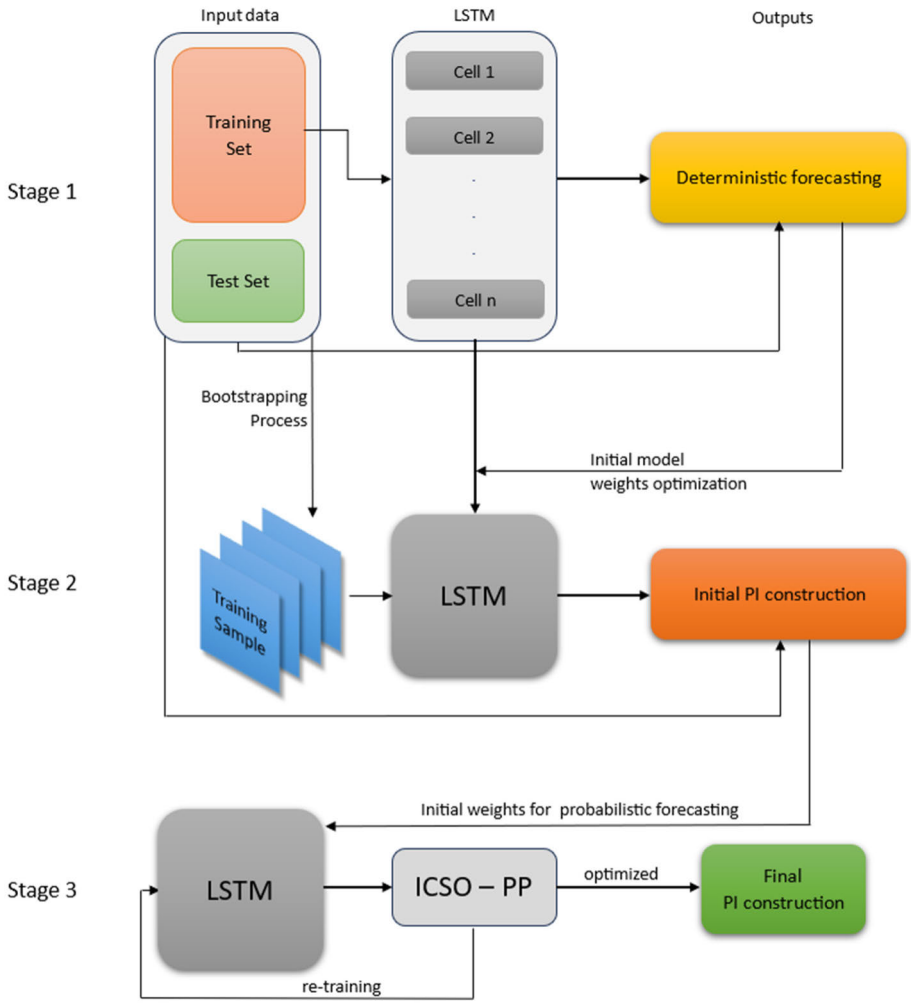


FIGURE 3. Proposed model structure and interaction between each processing stage. Stage 1 represents the deterministic stage of the proposed model, stage 2 represents the Bootstrapping application on the data and the initialization of the probabilistic forecasting process and stage 3 represents the ICSO-PP application and finalization of the probabilistic forecasting process.

PI. The PICP is computed by:

$$PICP = \frac{1}{N}\sum_{t=1}^{N} c_t \tag{16}$$

where N is the total number of samples and c_t is defined as:

$$c_t = \begin{cases} 1, & y_t \in [L_t, U_t] \\ 0, & y_t \notin [L_t, U_t] \end{cases} \tag{17}$$

To maximize the reliability of the PI, the PICP should be as close as possible to the nominal confidence level. It should be noted that for the PICP to be valid, it has to be greater than or equal to the predetermined level of confidence, which is the Prediction Interval Nominal Confidence (PINC):

$$PINC = 100\,(1-\alpha)\,\% \tag{18}$$

where (1-α) expresses the nominal coverage probability of the prediction intervals. The closer the PICP is to the PINC, the more reliable is the PI.

Another metric used in order to estimate the width of a PI is the Prediction Interval Normalized Average Width (PINAW). Since excessive PIs may cause difficulty in the predictive estimations and thus in decision making problems, metrics such as PINAW play a vital role in improving the reliability of a PI. PINAW is defined by:

$$PINAW = \frac{1}{NR}\sum_{t=1}^{N}(U_t - L_t) \tag{19}$$

where R is the range of the underlying targets that are used for normalizing the PIs.

The use of the median values for the PICP and PINAW metrics aims to highlight the improvement of the PI accuracy

TABLE 3. CWC optimization for the comparative models for PV Park 2.

Test day	CL (%)	CWC (%)						
		BELM	LSTM	Boot-LSTM	Boot-LSTM-CSO	Boot-LSTM-ICSO	BELM-ICSO-PP	Boot-LSTM-ICSO-PP
W/Sun	90	0.4423	0.3624	0.2425	0.2397	0.2031	**0.1943**	0.2006
	95	0.6481	0.3812	0.3318	0.3006	0.2845	0.2342	**0.2212**
	99	0.5627	0.4039	0.3648	0.3394	0.3036	0.2610	**0.2118**
W/LC	90	0.5608	0.5326	0.5123	0.4936	0.4730	0.4355	**0.3847**
	95	0.8893	0.5536	0.4982	0.4637	0.4249	0.3972	**0.3514**
	99	2.8967	0.5782	0.5279	0.4863	0.4312	0.3832	**0.3177**
W/HC	90	0.4372	0.3932	0.3862	0.3698	0.3701	0.3742	**0.3699**
	95	0.4699	0.4039	0.3211	0.2965	0.2742	0.2512	**0.2075**
	99	0.7074	0.4428	0.4216	0.3986	0.3608	0.3260	**0.2907**
S/Sun	90	0.3001	0.2813	0.2568	0.2412	0.2143	0.1759	**0.1714**
	95	0.3474	0.3287	0.2989	0.2741	0.2408	0.2177	**0.1869**
	99	0.5075	0.3308	0.3133	0.2813	0.2589	0.2269	**0.1930**
S/LC	90	0.3856	0.3791	0.3574	0.3411	0.3126	0.3012	**0.2765**
	95	0.3933	0.4017	0.3593	0.3048	0.3153	0.2769	**0.2281**
	99	0.4517	0.4396	0.3965	0.3436	0.3244	0.2996	**0.2573**
S/HC	90	0.4485	0.4296	0.3826	0.3566	0.3278	0.2987	**0.1988**
	95	0.5043	0.4642	0.4199	0.4024	0.3961	0.3408	**0.2986**
	99	0.8261	0.4719	0.4573	0.4126	0.4055	0.3802	**0.3574**

of the proposed model compared to the other models after the implementation of the prey-predator algorithm.

Since, during the training process, the bounds of the PIs are not yet constructed, an indirect training method is proposed in order to concurrently consider the calibration and sharpness aspects of the PIs. As a result, the training process is performed by using a PI-based cost function called Coverage Width Criterion (CWC) that is defined by:

$$CWC = PINAW + \gamma\left(PICP, \mu\right) e^{-\eta(PICP-\mu)} \tag{20}$$

where, μ *i*s the confidence level for PIs, η is a penalty coefficient used when PICP is smaller than μ in order to increase the difference between PICP and μ. Furthermore, $\gamma(PICP, \mu)$ is defined as:

$$\gamma\left(PICP, \mu\right) = \begin{cases} 0, & PICP \geq \mu \\ 1, & PICP < \mu \end{cases} \tag{21}$$

V. CASE STUDIES

In order to sufficiently evaluate its efficiency, the proposed forecasting method is tested on two different case studies (two different PV parks).

The first case study refers to a PV park of 1551 kWp (PV park 1). Two years of historical data are available dating from January 1st, 2020, to December 31st, 2021, with a timestep of one min. The available dataset includes historical power production data as well as meteorological data, including the inclined irradiance, the module temperature, the ambient temperature, the relative humidity, the precipitation as well as the wind speed and direction.

The second case study refers to a PV park of 11900 kWp (PV park 2). Two years of historical data are available dating from January 1st 2019, to December 31st 2020, with a timestep of 15 min. The available dataset includes historical power production data and meteorological data, including the inclined irradiance, the module temperature, the ambient temperature and the relative humidity.

Considering that for the majority of the data it was observed that the PV power output reached the maximum at times close to 13:00, for the optimization of the whole process, it is considered that the PV power output reaches the maximum at 13:00. Considering the PV power production curve for each day, the days were further subcategorized into sunny days, days with light clouds and days with heavy clouds. An example of the respective PV power production curves of the above three subcategories is shown in Figure 4. The power output curves of PV power generation systems fail to demonstrate a cycle pattern under light and heavy cloud conditions. The PV power output during light and heavy cloud weather fluctuates more than in sunny weather, and there is no law to estimate the power output in these conditions. Considering that decrease of one degree of the PV module's temperature could lead to 0.5% power production loss, the solar panels' dependency to the temperature is highly affected by the cloud fluctuations. Therefore, estimating the power generation in light and heavy cloud conditions is in general very challenging. That is also visible from the published literature for the papers that explicitly report the results for such days [29], [30].

VI. RESULTS AND DISCUSSION

A. EXPERIMENTAL PARAMETRIZATION OF THE PROPOSED MODEL

The parametrization of the proposed model has demanded experimentation on the optimal combination of the

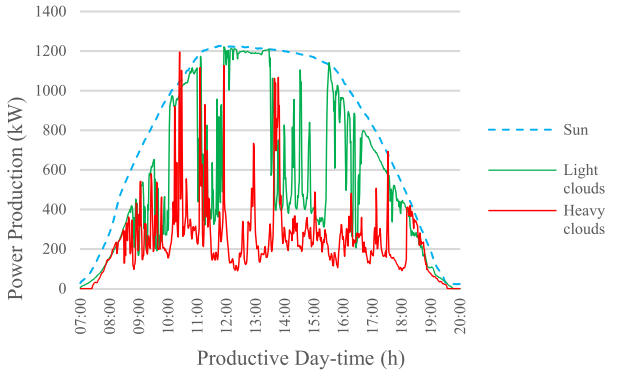


FIGURE 4. Illustration of typical PV power production curves for days with sun, light clouds and heavy clouds. Example from PV park1.

parameters of the LSTM and the ICSO. The number of neurons of the hidden layer is determined via the k-fold cross validation technique. The optimization of the hyperparameters was rather a complicated process, considering the fact that, in comparison to other swarm type optimization algorithms, in the CSO there is more than one population to be optimized.

During the data pre-processing step, all the samples are normalized in the range [0,1] using the min-max method, so the algorithm does not give biases results, because input features are normalized in the same scale. Then, the hyperparameters of the model are determined. Regarding the LSTM model, the confidence levels used for the forecasting are 90%, 95% and 99%, and the model's hidden layers are set equal to 30. The optimization of the proposed model is performed in two stages. First, the parameters of the LSTM are configured to generate optimal point forecasts using the Adam optimizer. The training set is then split into daily subsets, to maintain the temporal autocorrelation of PV power generation after sampling with replacement. Initial PIs are formed using the bootstrap technique. The parameters of the LSTM are then re-configured using the proposed ICSO-PP algorithm, to optimize the generated PIs. The day-ahead predictions are generated using the aforementioned PVrelated measurements of the previous five days. The sigmoid activation function is used at the output layer of the LSTM model.

Considering the implementations of the CSO, the number of chickens is set equal to 40, 30% of which are roosters, 50% are hens, and 40% of the hens are mothers. It is noted that the relationships between the chickens of the swarm are updated every G = 5 generations, and the maximum number of generations is set equal to 20. The mutation operator is applied with a probability that linearly decreases for the first 70% of the generations. The final selected parameters were the result of fine-tuning based on the minimization of the forecasting error. This way, the final swarm configuration effectively avoids getting stuck in local minima. Table 5 presents the applied hyperparameters of the combined ICSO-PP algorithm.

The probability of a prey encountering a predator is set to 20%, and the prey-predator mechanism is applied for the first 70% of the generations, when a chicken's fitness value is 15% higher than the swarm's optimum. Then, a probabilistic tournament is performed, on whether the particle will escape the predator, or it will get caught, with probabilities 75% and 25%, respectively. The above parameters are chosen based on the minimization of the CWC that is defined in Equation (20).

Aiming to confirm the efficiency and the accuracy of the proposed model, its performance is compared to several methodologies of SPPF. The selected benchmark models are the simple LSTM model as well as the BELM model which has been the center of attention for PV forecasting over the past years. The simple CSO and the improved CSO algorithm were afterwards implemented to the simple LSTM. Finally, the prey-predator mechanism was added to the CSO and the ICSO algorithms. Seven models are compared: 1) LSTM, 2) Boot-LSTM, 3) Boot-LSTM-CSO, 4) Boot-LSTM-ICSO, 5) BELM, 6) BELM-CSO-PP and 7) the proposed Boot-LSTM -ICSO-PP model.

The above comparisons suffice to:

a. prove that the application of an advanced optimization algorithm such as the Improved CSO, combined with the use of the PP algorithm, can successfully improve the state-of-the-art LSTM model, and
b. explain the different optimization stages that were used to reach the final proposed model that significantly improves the forecasting accuracy.

TABLE 4. Comparative results of PI construction.

Case	*PINC*	*Error Metric*	*Comparative model*						
			BELM	LSTM	Boot-LSTM	Boot-LSTM-CSO	Boot-LSTM-ICSO	BELM-ICSO-PP	Boot-LSTM-ICSO-PP
1	0.90	PICP	0.9672	0.9203	0.9431	0.9395	0.9141	**0.9132**	0.9089
	0.95		0.9701	0.9662	0.9739	0.9702	0.9685	**0.9602**	0.9514
	0.99		0.9987	0.9953	0.9968	0.9927	0.9936	0.9927	**0.9935**
	0.90	PINAW	0.5094	0.3909	0.2014	0.1659	0.1762	0.1627	**0.1237**
	0.95		0.4970	0.4282	0.2751	0.2203	0.2156	0.2093	**0.1948**
	0.99		0.6207	0.4603	0.3636	0.3394	0.3461	0.2890	**0.2671**
2	0.90	PICP	0.9364	0.9087	0.9217	0.9199	0.9123	0.8813	**0.9067**
	0.95		0.9762	0.9598	0.9731	0.9688	0.9714	0.9610	**0.9533**
	0.99		1.0000	0.9962	1.0000	1.0000	1.0000	**1.0000**	**1.0000**
	0.90	PINAW	0.4398	0.3862	0.3233	0.2945	0.2682	0.2547	**0.2317**
	0.95		0.4871	0.4028	0.3479	0.3428	0.2987	0.2901	**0.2763**
	0.99		0.6351	0.4412	0.4123	0.3962	0.4018	0.3218	**0.3132**

TABLE 5. Hyperparameter values of the ICSO-PP algorithm.

Hyperparameter	*Value*
G	5
$N_{chickens}$	40
N_r	12
N_h	20
N_m	8
N_c	8
Cip_{min}	0.3
Cip_{max}	0.8

B. EXPERIMENTAL RESULTS

Tables 2 and 3 present the comparative results for the minimization of the CWC for PV park 1 and PV park 2, respectively. The results are presented from a seasonal perspective, including results from a sunny day (Sun), a light cloudy day (LC) and a heavily cloudy day (HC) for winter (W) and summer (S) seasons. The selected confidence levels (CL) are 90%, 95% and 99%.

The accuracy of the proposed model in terms of PI is further investigated. The median values of the PICP and PINAW metrics for the 0.90, 0.95 and 0.99 confidence levels for both case studies are presented in Table 4.

C. DISCUSSION

As it can be derived from Tables 2 and 3, the proposed model surpasses the comparative models in terms of CWC minimization. Considering the implementation of the prey-predator algorithm, it can be seen that it improves the results of CWC minimization in the majority of cases. Moreover, it can be seen that in cases of cloudy days (light or heavy), the improvement of the CWC results is significantly better than in cases of sunny days where the improvement is slightly smaller. That is to be expected because, as already discussed at the end of Section V, it is much easier to forecast the PV power on sunny days, and, as a result, simpler methods will already do quite well, without deep optimization.

Moreover, while not focused on the deterministic stage of the proposed methodology, a representation of the deterministic forecasting stage should further prove the validity of the probabilistic forecasting model's input data as well as the quality of the deterministic forecasts overall. Figure 5 presents different regression plots of the deterministic predictions for six different cases, one for a sunny day, one for a light clouded day and one for a heavy clouded day for park1 and park2. It can be seen that the predicted deterministic values present high correlation to the real values, which except from the plots is also confirmed by the MAE (<5%) and RMSE (<7%) metric values for park1 while the respective values for park2 are <6.5% and <7.5%. Slightly better results are to be expected for PV park1 since the larger amount of information provided by PV park1, due to its one-minute timestep, provides better training possibilities to the forecasting model. It can also be concluded that the correlated values are better aligned in the case of sunny days than in the case of cloudy days. Furthermore, in the case of light cloudy days, the correlated values show lower dispersity for values closer to zero PV production as well as the nominal power production value and higher dispersity for values in between; while for heavy cloudy days the correlated values show lower dispersity for values closer to zero PV production. The above are to be expected considering the effect of the volatility of the clouds during the day, since predictions tend to be more accurate in cases of production closer to the nominal (clear sky) or cases of very low or even no production (nighttime or time closer to dusk or dawn). This is further confirmed by the regression plot of a sunny day where the correlated values show minimum dispersity. The above conclusions can be better seen from Figures 5a, 5b, and 5c, since park1 has a oneminute time step and therefore the existing scatter plot represents a bigger number of regression data. However, in the case of PV park2, a similar image is observed, even though in a smaller scale. The scatter plotted regression data are better aligned in the case of the sunny day while for cloudy days, the correlated values show lower dispersity for values closer to zero PV production and nominal production and higher dispersity for values in between.

It can further be derived from Tables 2 and 3 that the proposed model functions better in cases of light and heavy cloudy days. The LSTM-based model is able to capture the sequential dependencies in the data. Furthermore, thanks to the resampling ability and the sequential-based architecture of the proposed model, the Boot-LSTM-ICSO-PP model is able to locate and exploit the relationship between the input data and the solar power data that are highly nonlinear and complex. Moreover, there are cases where the BELM-ICSO-PP model is equal to or slightly better than the proposed model, as can be seen in Table 2 for the summer sunny day and more clearly in both sunny day cases in Table 3. Such cases are noticed mostly in summer examples where the training data are more stable and present low variability. It should further be noted that the Boot-LSTM is able to take advantage of the intra-day spot forecasts of the proposed model. In cases where the intra-hour spot forecasts do not follow an expected PV-production curve (such as cases of bad weather), the proposed model is capable of utilizing them in a more efficient way, in order to better evaluate the day-ahead fluctuations. This is clearly presented in Tables 2 and 3, where the results for the CWC of the proposed model are significantly better than the comparative ones. On the other hand, BELM does not consider previous values in the final estimation and therefore fails to take advantage of the intra-hour forecasts to further improve the final forecasting error.

Furthermore, the proposed model seems to respond better to heavy cloud days than to light cloud days. This could be the result of the high fluctuation of the light cloud days, in comparison to heavy cloud days, and the impact on the solar modules' dependency to the continuous temperature changes during light days. It should be noted that similar tests have been run for more days of the same year, where the results were similar to the presented ones.

The results of Table 4 show the superiority of the proposed model in terms of PI construction accuracy. While the results for the other models are close to the PINC, in the case of the proposed model, it can be seen that the PICP values are extremely close to the PINC values for both winter and summer cases and for all confidence levels. Furthermore, the better PINAW results for the proposed model in all test cases presented indicate the superiority of the proposed model in terms of accuracy.

A visualization of the results that include the constructed PIs as well as an intra-hour estimated power yield for different

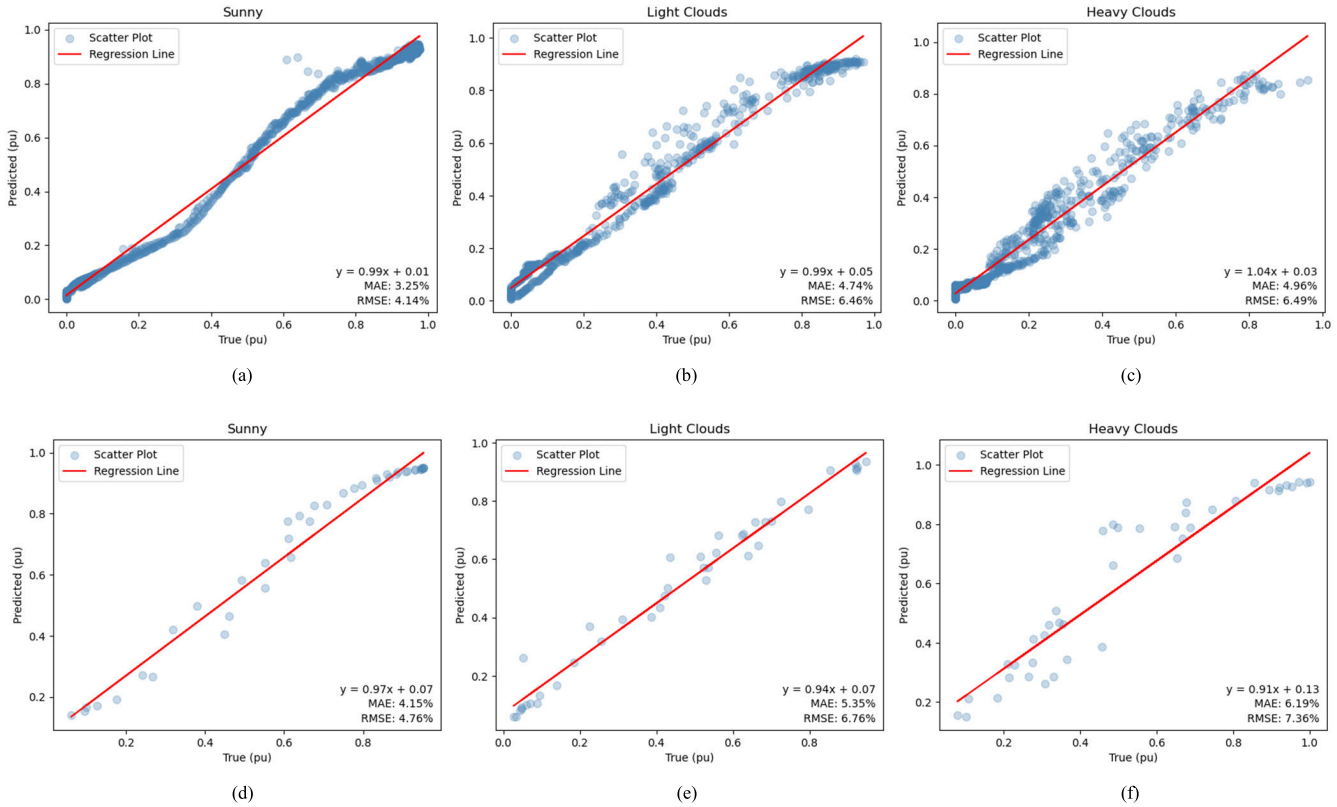


FIGURE 5. Deterministic forecasting regression representation for: (a) a sunny day, (b) a light cloudy day and (c) a heavy cloudy day, for PV park1 and (d) a sunny day, (e) a light cloudy day and (f) a heavy cloudy day, for PV park2.

TABLE 6. Average training and run times for PV parks 1 and 2 and scalability test results.

PV Park 1		PV Park 2			Optimization Parameters		Total Training Parameters
Training time (sec)	Run time (sec)	Training time (sec)	Run time (sec)	Hidden Units	Weights	Biases	
949		742		5	280	20	12000
1514		1067		10	760	40	32000
2839		**1958**		**20**	2320	80	96000
3934	**10.2**	3019	**8.6**	30	4680	120	192000
6109		5318		50	11800	200	480000
8613		7427		70	22240	280	900800
11785		9823		100	43600	400	1760000

meteorological cases for both park1 and park2 is presented in Figure 6. Such combined result could be incredibly useful to decision makers. Deterministic forecasting is valuable for short-term operational planning, while probabilistic forecasting is able to cover cases of a dynamic and uncertain context. PIs are simple to interpret and offer qualitative uncertainty information to decision makers. Furthermore, generating day-ahead PIs enables interval optimization for power system tasks such as dayahead economic dispatch or unit commitment, leading to higher flexibility and several more or less conservative choices. Moreover, the consideration of intra hour deterministic estimations could further facilitate decision makers, not only to have a more in-depth view of the development of the forecasting model's results, but also to understand and self-validate the precision of the prediction intervals constructed by the probabilistic estimations.

Figure 6 efficiently captures the robustness and adaptability of the proposed model. It is clear that cases 6(a), 6(b) and 6(c) contain a larger amount of data than 6(d), 6(e) and 6(f), since the predictive result for PV park1 has a timestep of one minute while for PV park2 has a timestep of 15 minutes. However, it is obvious that the constructed PIs efficiently include the forecasted timeseries as well as the real one. Moreover, it can be seen from the forecasted lines in Figure 6 that the model is able to efficiently capture and estimate the shifts in the slope of the productive curve, in cases of sunny weather (normal conditions, easy to forecast), as well

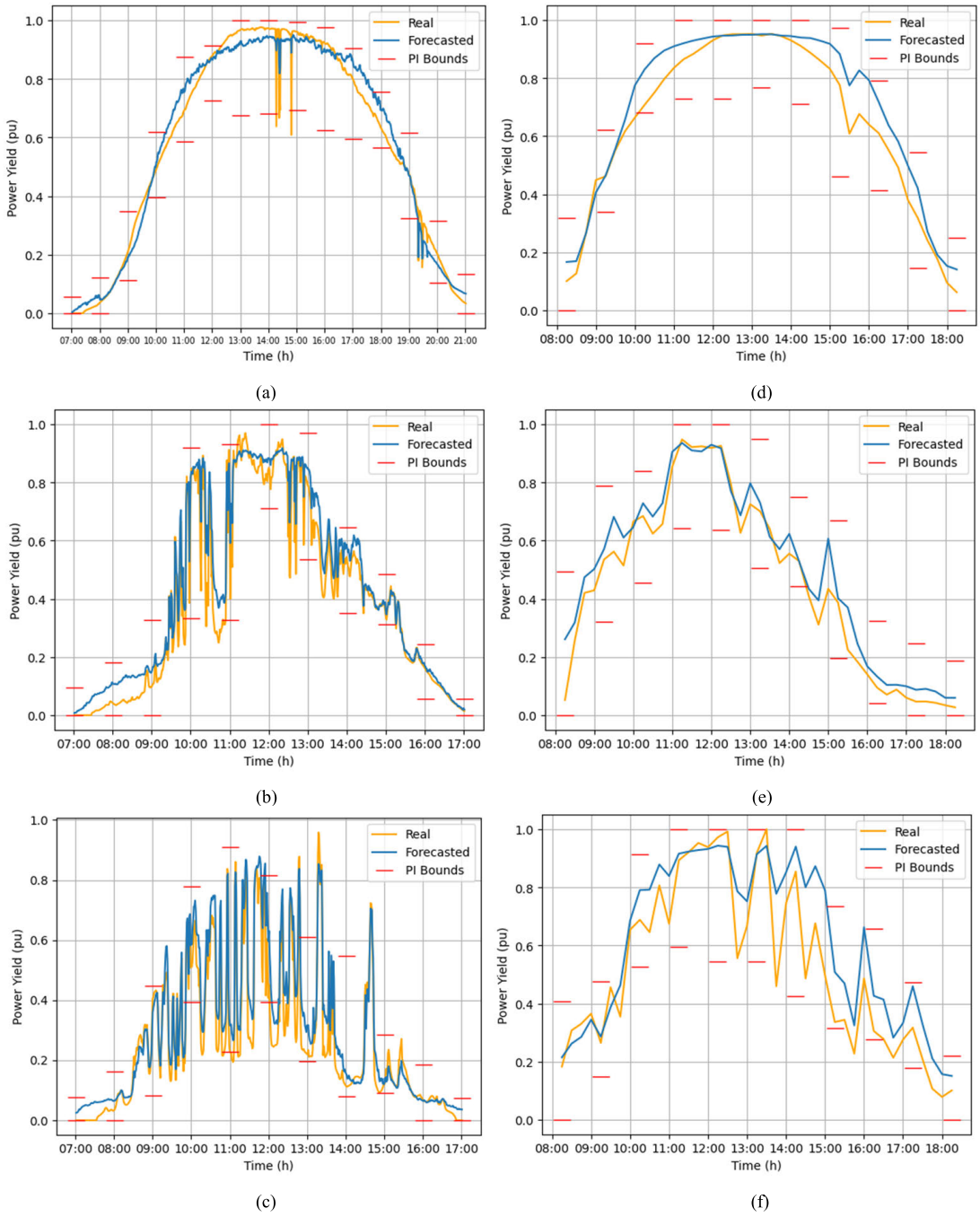


FIGURE 6. Visualization of constructed PIs and day-ahead forecasted power yield for: (a) a sunny day, (b) a light cloudy day and (c) a heavy cloudy day, for PV park1 and (d) a sunny day, (e) a light cloudy day and (f) a heavy cloudy day, for PV park2. Comparison with real power curves.

as in cases of cloudy weather (high fluctuation, more difficult to forecast).

From a time-efficiency point of view, as it can be derived from Table 6, the proposed model is able to perform in efficient time caps in terms of training as well as prediction run-time. At this point, it should be noted that the forecasting model was developed and run on a computer system with the following specifications: Windows 10 Pro, 2H22, 16 GB RAM and an AMD Ryzen 5 3600X 6-Core Processor of 3.80 GHz, with a memory usage of around 900MB for the model to run. Considering their application to real time data, forecasting models' training time varies depending on the dataset length, the number of input variables and the complexity of the model's architecture.

To establish an optimal architecture for the proposed model in terms of predictive accuracy and low computational cost,

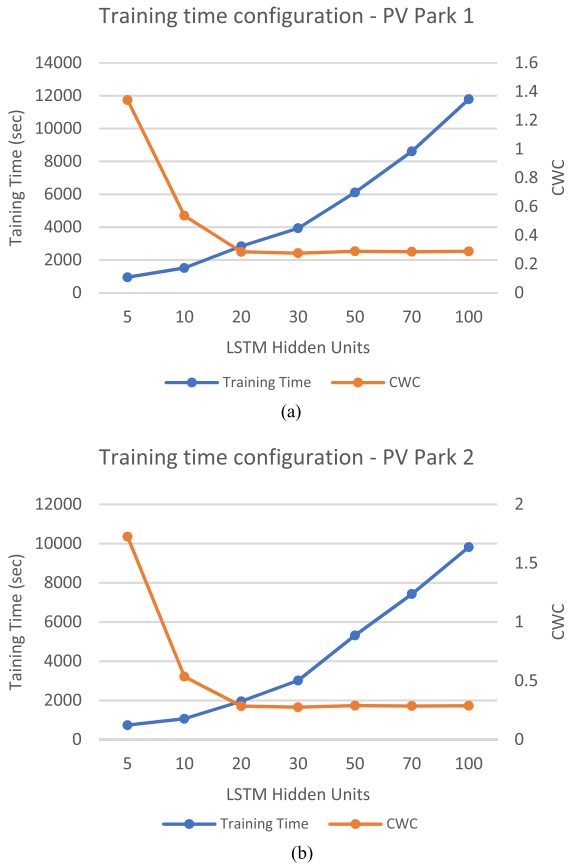


FIGURE 7. Illustration of the training time configuration process in terms of computational efficiency and CWC minimization for: (a) PV Park 1 and (b) PV Park 2.

the complexity of the model was evaluated by the response of the model to different hidden units. As shown in Figure 7, for both PV parks, with the increase of the LSTM's hidden units, the training time greatly increases. This is to be expected since, as presented in Table 6, by increasing the hidden units, the LSTM parameters increase proportionally, for more than 50%. Therefore, a significant increase in the training time is inevitable. However, it is obvious from the plots in Figure 7, that the proposed model manages to efficiently converge to the optimal CWC value really fast, at 20 LSTM hidden units, avoiding to resort to higher training times and therefore, increased computational burden.

The proposed model, optimized to function in 20 LSTM hidden units, is trained in an average time of 47 minutes and 19 seconds (2839 sec) for PV park1 and of 32 minutes and 38 seconds (1958 sec) for PV park2, as seen in Table 6. It should be noted that between 20 and 30 hidden units, a 50increase in the training parameters leads to a 38.54% increase in training time for park1 and a 52.1% for park2, without achieving a significant improvement in CWC minimization. The difference to the training time is to be expected considering that the dataset of PV park1 has a timestep of 1 minute, while the dataset of PV park2 has a timestep of 15minutes. In general, while highly accurate, the model seems computationally intensive, considering the application of the advanced ICSO under a probabilistic spectrum.

However, it still remains in permissible running limits, thanks to the contribution of the PP mechanism.

D. FUTURE RESEARCH

The sector of solar power probabilistic forecasting has gained a lot of attention over the past years. The proposed methodology, while it is able to provide results of high accuracy in the probabilistic spectrum, it could be further evaluated under different conditions and considering different problems. Its application to a more complete electrical energy system could further evaluate its performance in larger scale systems and also investigate the usefulness of the simultaneous production of deterministic and probabilistic forecasts. Furthermore, the application of the proposed methodology to a dataset that includes data from satellite images could further improve the total accuracy of the proposed model. While open access to such data can be limited, they have been included more and more into solar power forecasting research over the past years. Moreover, developing the proposed methodology from a spatio-temporal viewpoint could further prove its stability to real-life conditions and problems. With meteorological variables (solar irradiance, cloud coverage, wind conditions, temperatures) tending to show high correlation is case of close locations [31], using data from different PV parks in close locations could further improve the model's accuracy as well as provide useful results as to the power production dependencies between different PV parks. Finally, considering that green computing is becoming more and more important, exploring methods to reduce the complexity of the model while retaining its high accuracy standards would also be of research interest.

VII. CONCLUSION

In this paper, a hybrid Boot-LSTM-based probabilistic forecasting model was developed in order to efficiently construct PIs. The Boot-LSTM's resampling ability and its advantage to capture the sequential dependencies were successfully exploited for the construction of accurate PIs. The model was trained and further optimized via an improved version of the CSO algorithm. The inclusion of the ICSO allowed for a faster and more accurate convergence of the model and therefore for higher accuracy overall.

A prey-predator mechanism was included and adjusted to the model's architecture in order to facilitate the ICSO particles' searching ability and improve the predictive accuracy along with the computational cost.

Two different datasets were used in order to prove the proposed model's superiority compared to six state-of-theart models. Results presented in Section VI indicate that the proposed model is indeed superior in the majority of the cases in terms of the evaluation metrics. In comparison to BELM and LSTM base models, our proposed model achieves a 60.3% and 46.94% average improvement for PV park1, and a 50.64% and 37.87% average improvement for PV park2.

The integration of the ICSO to the proposed bootstrapped version of the LSTM model and the further implementation of the prey-predator mechanism, are proven necessary in terms of improvement in the forecasting accuracy, especially in cases of difficult forecasted days (light and heavy cloud days). As a result, the proposed model could contribute to the optimal integration of solar power into power systems and its application on real life problems.

REFERENCES

[1] J. Jia, G. Yang, and A. H. Nielsen, "A review on grid-connected converter control for short-circuit power provision under grid unbalanced faults," *IEEE Trans. Power Del.*, vol. 33, no. 2, pp. 649–661, Apr. 2018.

[2] M. Z. Jacobson and M. A. Delucchi, "Providing all global energy with wind, water, and solar power, part I: Technologies, energy resources, quantities and areas of infrastructure, and materials," *Energy Policy*, vol. 39, no. 3, pp. 1154–1169, Mar. 2011.

[3] Y. Fu, H. Chai, Z. Zhen, F. Wang, X. Xu, K. Li, M. Shafie-Khah, P. Dehghanian, and J. P. S. Catalão, "Sky image prediction model based on convolutional auto-encoder for minutely solar PV power forecasting," *IEEE Trans. Ind. Appl.*, vol. 57, no. 4, pp. 3272–3281, Jul. 2021.

[4] A. Strzalka, N. Alam, E. Duminil, V. Coors, and U. Eicker, "Large scale integration of photovoltaics in cities," *Appl. Energy*, vol. 93, pp. 413–421, May 2012.

[5] I. K. Bazionis, M. A. Kousounadis-Knousen, P. S. Georgilakis, E. Shirazi, D. Soudris, and F. Catthoor, "A taxonomy of short-term solar power forecasting: Classifications focused on climatic conditions and input data," *IET Renew. Power Gener.*, vol. 17, no. 9, pp. 2411–2432, Jul. 2023.

[6] A. G. Tsikalakis, N. D. Hatziargyriou, Y. A. Katsigiannis, and P. S. Georgilakis, "Impact of wind power forecasting error bias on the economic operation of autonomous power systems," *Wind Energy*, vol. 12, no. 4, pp. 315–331, May 2009.

[7] J.-F. Toubeau, J. Bottieau, Y. Wang, and F. Vallée, "Interpretable probabilistic forecasting of imbalances in renewable-dominated electricity systems," *IEEE Trans. Sustain. Energy*, vol. 13, no. 2, pp. 1267–1277, Apr. 2022.

[8] B. Li and J. Zhang, "A review on the integration of probabilistic solar forecasting in power systems," *Sol. Energy*, vol. 210, pp. 68–86, Nov. 2020.

[9] I. K. Bazionis, P. A. Karafotis, and P. S. Georgilakis, "A review of short-term wind power probabilistic forecasting and a taxonomy focused on input data," *IET Renew. Power Gener.*, vol. 16, no. 1, pp. 77–91, Jan. 2022.

[10] M. Sun, C. Feng, and J. Zhang, "Probabilistic solar power forecasting based on weather scenario generation," *Appl. Energy*, vol. 266, May 2020, Art. no. 114823.

[11] F. Golestaneh, P. Pinson, and H. B. Gooi, "Very short-term nonparametric probabilistic forecasting of renewable energy generation—With application to solar energy," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3850–3863, Sep. 2016.

[12] M. Bozorg, A. Bracale, P. Caramia, G. Carpinelli, M. Carpita, and P. De Falco, "Bayesian bootstrap quantile regression for probabilistic photovoltaic power forecasting," *Protection Control Modern Power Syst.*, vol. 5, no. 1, pp. 1–12, Dec. 2020.

[13] H. He, N. Lu, Y. Jie, B. Chen, and R. Jiao, "Probabilistic solar irradiance forecasting via a deep learning-based hybrid approach," *IEEJ Trans. Electr. Electron. Eng.*, vol. 15, no. 11, pp. 1604–1612, Nov. 2020.

[14] T. Carriere, C. Vernay, S. Pitaval, and G. Kariniotakis, "A novel approach for seamless probabilistic photovoltaic power forecasting covering multiple time frames," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2281–2292, May 2020.

[15] L. Lv, Z. Wu, J. Zhang, L. Zhang, Z. Tan, and Z. Tian, "A VMD and LSTM based hybrid model of load forecasting for power grid security," *IEEE Trans. Ind. Informat.*, vol. 18, no. 9, pp. 6474–6482, Sep. 2022.

[16] Z. Xiao, X. Huang, J. Liu, C. Li, and Y. Tai, "A novel method based on time series ensemble model for hourly photovoltaic power prediction," *Energy*, vol. 276, Aug. 2023, Art. no. 127542.

[17] Y. Cao, G. Liu, D. Luo, D. P. Bavirisetti, and G. Xiao, "Multi-timescale photovoltaic power forecasting using an improved stacking ensemble algorithm based LSTM-informer model," *Energy*, vol. 283, Nov. 2023, Art. no. 128669.

[18] O. Lindberg, D. Lingfors, J. Arnqvist, D. van der Meer, and J. Munkhammar, "Day-ahead probabilistic forecasting at a co-located wind and solar power park in Sweden: Trading and forecast verification," *Adv. Appl. Energy*, vol. 9, Feb. 2023, Art. no. 100120.

[19] J. Wang, Y. Yu, B. Zeng, and H. Lu, "Hybrid ultra-short-term PV power forecasting system for deterministic forecasting and uncertainty analysis," *Energy*, vol. 276, Feb. 2023, Art. no. 127542.

[20] Y. Li, L. Zhou, P. Gao, B. Yang, Y. Han, and C. Lian, "Short-term power generation forecasting of a photovoltaic plant based on PSO-BP and GA-BP neural networks," *Frontiers Energy Res.*, vol. 9, pp. pp. 958–965, Jan. 2022.

[21] X. Luo, D. Zhang, and X. Zhu, "Deep learning based forecasting of photovoltaic power generation by incorporating domain knowledge," *Energy*, vol. 225, Jun. 2021, Art. no. 120240.

[22] B. Efron, "Bootstrap methods: Another look at the jackknife," *Ann. Statist.*, vol. 7, no. 1, pp. 1–26, Jan. 1979.

[23] R. Tibshirani, "A comparison of some error estimates for neural network models," *Neural Comput.*, vol. 8, no. 1, pp. 152–163, Jan. 1996.

[24] X. B. Meng, Y. Liu, X. Z. Gao, and H. Z. Zhang, "A new bio-inspired algorithm: Chicken swarm optimization," in *Advances in Swarm Intelligence*, vol. 8794. Cham, Switzerland: Springer, 2014, pp. 86–94.

[25] W. Shi, Y. Guo, S. Yan, Y. Yu, P. Luo, and J. Li, "Optimizing directional reader antennas deployment in UHF RFID localization system by using a MPCSO algorithm," *IEEE Sensors J.*, vol. 18, no. 12, pp. 5035–5048, Jun. 2018.

[26] C. Fu, G.-Q. Li, K.-P. Lin, and H.-J. Zhang, "Short-term wind power prediction based on improved chicken algorithm optimization support vector machine," *Sustainability*, vol. 11, no. 2, p. 512, Jan. 2019.

[27] H. Zhang, M. Yuan, Y. Liang, and Q. Liao, "A novel particle swarm optimization based on prey–predator relationship," *Appl. Soft Comput.*, vol. 68, pp. 202–218, Jul. 2018.

[28] Y. Wu, P. Su, T. Wu, J. Hong, and M. Y. Hassan, "Probabilistic wind power forecasting using weather ensemble models," in *Proc. IEEE/IAS 54th I&CPS*, Niagara Falls, ON, Canada, May 2018, pp. 1–8.

[29] M. Chai, F. Xia, S. Hao, D. Peng, C. Cui, and W. Liu, "PV power prediction based on LSTM with adaptive hyperparameter adjustment," *IEEE Access*, vol. 7, pp. 115473–115486, 2019.

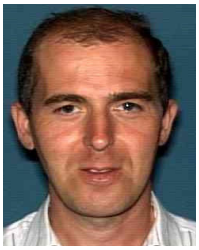
[30] M. Gao, J. Li, F. Hong, and D. Long, "Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using LSTM," *Energy*, vol. 187, Nov. 2019, Art. no. 115838.

[31] A. Tascikaraoglu, B. M. Sanandaji, G. Chicco, V. Cocina, F. Spertino, O. Erdinc, N. G. Paterakis, and J. P. S. Catalão, "Compressive spatio-temporal forecasting of meteorological quantities and photovoltaic power," *IEEE Trans. Sustain. Energy*, vol. 7, no. 3, pp. 1295–1305, Jul. 2016.

IOANNIS K. BAZIONIS received the Diploma degree in geology and geo-environment from the National and Kapodistrian University of Athens, Greece, in 2015, and the M.Sc. degree in energy production and management from the National Technical University of Athens (NTUA), Greece, in 2018, where he is currently pursuing the Ph.D. degree with the School of Electrical and Computer Engineering. His research interests include power forecasting and day-ahead scheduling of power distribution networks in electricity market environment considering the forecasting errors in the production of renewable energy sources.

MARKOS A. KOUSOUNADIS-KNOUSEN received the Diploma degree in electrical and computer engineering from the National Technical University of Athens (NTUA), Greece, in 2021, where he is currently pursuing the Ph.D. degree with the School of Electrical and Computer Engineering. His research interests include power systems management optimization, renewable power generation forecasting, and photovoltaic power integration into electric power systems. He is a member of the Technical Chamber of Greece.

VASILEIOS E. KATSIGIANNIS received the Diploma degree in mechanical engineering from the National Technical University of Athens (NTUA), Greece, in 2022, specializing in energy engineering and the M.Sc. degree in data science and machine learning from the School of Electrical and Computer Engineering, NTUA, in 2024, where he is currently pursuing the Ph.D. degree with the Biomedical Engineering Laboratory, School of ECE. He is also a Researcher. His research interests include managing and processing large volumes of data, machine learning algorithms optimization, and developing AI-oriented medical decision support systems in healthcare applications.



FRANCKY CATTHOOR (Fellow, IEEE) received the Ph.D. degree in EE from Katholieke Universiteit Leuven (KULeuven), Belgium, in 1987. Between 1987 and 2000, he has headed several research domains in the area of synthesis techniques and architectural methodologies. Since 2000, he has been strongly involved in other activities with IMEC, Leuven, Belgium, including co-exploration of application, computer architecture and deep submicron technology aspects, biomedical systems and IoT sensor nodes, and photo-voltaic modules combined with renewable energy systems with IMEC. Currently, he is an IMEC Senior Fellow. He is also a part-time Full Professor with the EE Department, KULeuven. He has been an associate editor of several IEEE and ACM journals.

PAVLOS S. GEORGILAKIS (Senior Member, IEEE) received the Diploma and Ph.D. degrees in electrical and computer engineering from the National Technical University of Athens (NTUA), Athens, Greece, in 1990 and 2000, respectively. In September 2009, he joined as a Faculty Member of the School of Electrical and Computer Engineering, NTUA, where he is currently a Full Professor. From 2004 to 2009, he was an Assistant Professor with the School of Production Engineering and Management, Technical University of Crete, Greece. From 1994 to 2003, he was with Schneider Electric AE, the Greek Subsidiary of Schneider Electric, where he was a Quality Control Engineer, a Transformer Design Engineer, the Research and Development Manager, and the Low Voltage Products Marketing Manager. His current research interests include optimization algorithms and computational intelligence techniques for the optimal operation and planning of smart distribution systems. He is a member of the Technical Chamber of Greece.