Received: 22 September 2021 | Revised: 9 December 2021 | Accepted: 21 December 2021

IET Renewable Power Generation

DOI: 10.1049/rpg2.12389

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

# Execution of synthetic Bayesian model average for solar energy forecasting

**Oveis Abedinia** | **Mehdi Bagheri**

Department of Electrical and Computer Engineering, Nazarbayev University, Nursultan, Kazakhstan

**Correspondence**
Oveis Abedinia, Department of Electrical and Computer Engineering, Nazarbayev University, Nursultan 010000, Kazakhstan.
Email: oveis.abedinia@gmail.com, oveis.abedinia@nu.edu.kz

**Funding information**
Collaborative Research Project (CRP) Grant of Nazarbayev University, Grant/Award Number: 021220CRP0322

**Abstract**
Accurate photovoltaic (PV) forecasting is quite crucial in planning and in the regular operation of power system. Stochastic habit along with the high risks in PV signal uncertainty and a probabilistic forecasting model is required to address the numerical weather prediction (NWP) underdispersion. In this study, a new synthetic prediction process based on Bayesian model averaging (BMA) and Ensemble Learning is developed. The proposed model is initiated by the improved self-organizing map (ISOM) clustering K-fold cross-validation for the training process. To provide desirable learning model for different input samples, three learners including long short-term memory (LSTM) network, general regression neural network (GRNN), and non-linear auto-regressive eXogenous NN (NARXNN) are employed. The proposed BMA approach is combined with the output of the learners to obtain accurate and desirable outcomes. Different models are precisely compared with the obtained numerical results over real-world engineering test site, that is, Arta-Solar case study. The numerical analysis and recorded results validate the performance and superiority of the proposed model.

## 1 | INTRODUCTION

### 1.1 | Problem definition

Due to increasing the renewable energies in power network, that is, wind power and solar energy, an excessive alarm for power system operators as well as planners have been proposed and created [1]. Penetration of stochastic energy generators will potentially threaten the power system's efficiency and reliability. To address this issue and overcome the challenge, it is crucial to employ an accurate forecasting approach [2]. A powerful method to handle this issue is development of an accurate solar power forecasting system [3]. For this purpose, different research works have been published in recent years (deterministic models and probabilistic approaches) however, the probabilistic approaches are able to support and cover the prediction uncertainty entirely.

### 1.2 | Literature review

Several prediction methods have been implemented over solar signal, that is, autoregressive moving average (ARMA) [4], autoregressive integrated moving average (ARIMA) [5], and autoregressive moving average with eXogenous inputs (ARMAX) [6]. Due to nonlinear behaviour of solar signal, the linear methods are unable to support the uncertainty of the input signal and will terminate to low accuracy in forecasting model. To tackle this drawback, different machine learning approaches have been considered such as neural network (NN) [7–9], support vector machine (SVM) [10] etc.; even tough, such models might trick in an unrealistic optimal point for training the solar energy mapping function in multi-modal input/output model.

In recent years, different methods have been examined for numerical weather prediction (NWP) which mitigates the bias and underdispersion of the solar signal. In [11], two step post-processing approach has been presented for solar energy prediction. In [12], ensemble model output statistics (EMOS) with variance deficit is provided through fitting the normal distribution of parametric truncated. This approach has been utilized for wind speed [13], as well as for the electricity price prediction [14]. The combination of EMOS with Bayesian model averaging (BMA) is discussed in [15] over aggregated solar irradiance prediction.

The BMA has been applied inn different forecasting approaches, that is, wind speed [16], visibility [17], and precipitation [18]. This approach is a kernel-dressing model based on dressed ensemble members by probability density function (PDF). This model is a non-parametric approach same as kernel density estimation (KDE) which is discussed in [19–22] for wind and solar signals. In [22], a general framework is offered for renewable energies probabilistic forecasts based on k-nearest neighbour and KDE method. In [19], a new time-adaptive quantile-copula estimator is presented based on KDE to pick the adequate kernels for modelling various variables of the problem. In [21], new ensemble algorithm based on KDE is proposed to predict distributed generation from renewable energy resources. In [23], a new short-term PV generation prediction model based on ensemble adaptive boosting random forests is proposed. In [24], a novel pattern-recognition-based transient stability assessment model is offered based on an ensemble of OS-extreme learning machine through the binary Jaya-based feature selection by considering the phasor measurement units data.

The BMA make the KDE better through offering enhanced customization. In this process, each kernel will not be centred in NWP data point as a biased correction step. The input members are separately specified for the weights and bandwidth. In this model, more dependable members are marked as higher weight which cause the spreading of uncertainty. To remedy the ensemble bias and underdispersion problems, each member is weighted and shaped by BMA. Additionally, this approach employs accurate method at longer time horizons as inputs to a mixture-model, which are unable to present through a single parametric distribution.

As a remedy of the highlighted problems, in this study a novel synthetic forecasting method based on BMA and Ensemble learning is provided and discussed technically for the solar energy signal. In this line, the input data is divided for training by the improved Self organizing map clustering with K-fold cross validation (K-FCV) method. This problem provides suitable diversity for input samples of the learners. Afterwards, the proposed hybrid forecast engine including the long short-term memory (LSTM) network, general regression neural network (GRNN), and non-linear auto-regressive eXogenous NN (NARXNN) are utilized in training the signal based on multiple ensemble learning combination. Hence, to reach a better performance and reliability, combination approach of BMA is examined based on the mentioned learners' output precision. In view of this, the contribution and novelties of this work can be listed as:

### 1.3 | Contributions

a. Combination of ISOM clustering with K-FCV for training approach. This synthetic model increases the diversity of different learners.
b. Different leaners are developed in this study as LSTM, GRNN and NARXNN. The outputs of these learners are different while, the Ensemble learning provides suitable and accurate performance.
c. A new model of BMA is discussed through theoretical constraints on applying a beta kernel. This model is quantified based on different probabilistic metrics.
d. Suitable precision and high stability have been provided by proposed approach in prediction performance based on suggested BMA and Ensemble learning methods.

### 1.4 | Paper structure

Section 2 provides the new probabilistic forecasting approach based on Ensemble learners and training approach. In Section 3 the suggested BMA model is developed. Section 4, provides the numerical analysis and Section 5 will conclude the results and study.

## 2 | PROPOSED PV POWER FORECASTING APPROACH

In PV energy forecasting, the historical signal values (the auto-regression part), the irradiation and weather condition (exogenous variables) are completely critical factors for the forecasting engine development. The weather parameters, that is, temperature, cell temperature etc., are considered as exogenous data based on their correlation with PV signal.

### 2.1 | The data of Arta-Solar

In this study, the Arta-Solar power plant as a real practical implemented and operational test case is taken into consideration which is located in North West of Iran, that is, Latitude 38.5576 and Longitude 48.1432. This power plant has been constructed as the first part of 6 MW power plant includes of 2060 modules, that is, 700 kW. The average annual output of this site is about 1081 MWh with 1873 kWh/$m^2$ per year with the title angle 34° horizontally. It is able to support around 450 households load demand. For this plant, 12 Huawei inverters (60 kW per each) have been utilized. This plant started energy generation in May 5, 2020. This power plant is planned to be extended to 2 MW at the end of 2021 and grow to 6 MW within 2 years. The data have been recorded from 5 May 2020 till 5 May 2021. The average hourly profile for total PV power output (kWh) is provided in Figure 1.

### 2.2 | Training procedure

Due to application of different forecasting engines, the Ensemble learning approach is considered to be able to combine the learners' output. In addition, to improve the input candidates' differences, the diversity of the training subset is increased.

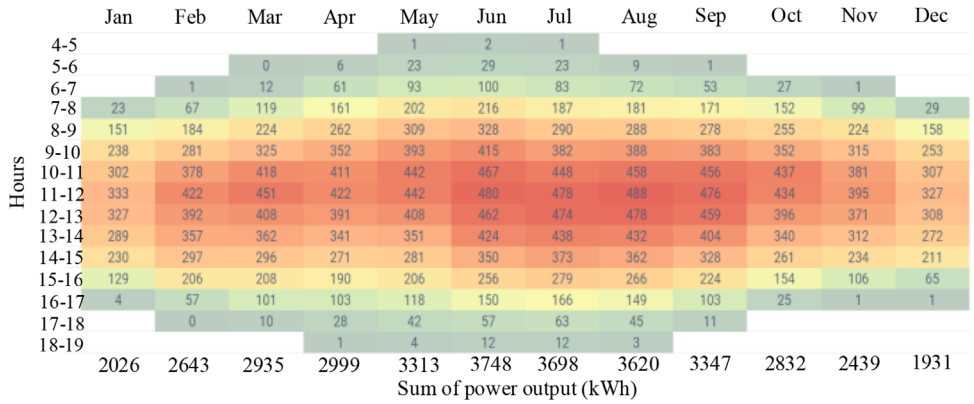


**FIGURE 1** The average hourly profile of Arta-Solar test case

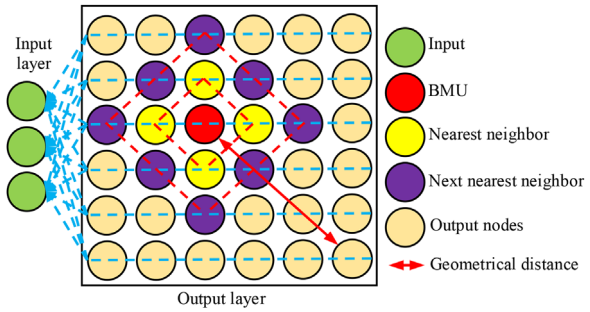


**FIGURE 2** The SOM diagram and key definition

This issue is considered due to important time periodicity of weather and seasonal data. Over the classic models, a bulk quantity of redundant data are observed in training due to uneven distribution of candidates for various weather conditions. Hence, to rise the diversity of the training subset excluding data rising, the improved SOM clustering model based K-FCV is offered in this study. In this way, the input signal will be clustered into sample sets before considering as the input of learners. Subsequently, the K-FCV is performed to decrease the data redundancy in training process.

### 2.2.1 | Proposed self organizing map

In this section, the proposed ISOM is discussed while the main construction of this approach is depicted in Figure 2. According to Figure 2, the model is a two-layer, non-supervised NN which evaluates the output based on the input data nature. The output layer includes of weighted nodes organized on a 2D map. The learning approach of this model is working as [25];

1. Initialization of output nodes weight by accidental real value.
2. Selection one input arbitrarily from the training data
3. Evaluation the Euclidean space (ES) among the chosen node's weight and input vector and (1).
4. Selection of lower ES from nodes as the best matching unit (BMU).
5. Evaluation of BMU neighbourhood function (2).
6. Updating the node's weight using (2) and (3).
7. Repeating till convergence criteria, that is, best ES value between the node's weight and input vector:

$$Dist = \sqrt{\sum_{i=0}^{i=n} (V_i - U_i)^2} \tag{1}$$

In (1), the input vector is defined by $V_i$ and node's weight is presented by $U_i$. to evaluate the BMU's neighbourhood function, it can be written as:

$$W_{bi}(t) = \alpha(t) e^{\frac{sd(b,i)}{2\delta^2(t)}} \tag{2}$$

where, the iteration depicted by $t$, the BMU defined by $b$, the learning rate is shown by $\alpha(t)$, $\delta(t)$ defines the neighbourhood radius, and the square distance of BMU and output plane ($i$) is defined by $sd(b,i)$. The node's weight function update, is given by:

$$U_i(t) = U_i + W_{bi}(t)[V_i - U_i] \tag{3}$$

This process will be continuing till the convergence criteria is satisfied, that is, providing the feature map while the nodes are clustered based on high similarities. To improve the SOM model, the diagnosis process is separated into two steps. The first step is the training of SOM and the second one is testing (validation process) of ISOM. The proposed model for this improvement is presented in Figure 3.

### 2.2.2 | Proposed learners

The proposed learners, that is, LSTM, GRNN and NARXNN are discussed separately in this section.

- General regression neural network: As a first learner, the GRNN is discussed to provide detailed information on the proposed method. This model is topologically similar to NN with lower variables in comparison to SVR and NN

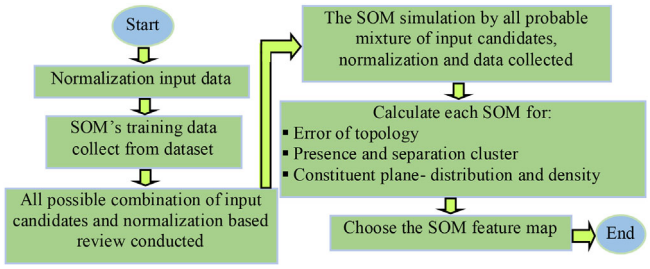


**FIGURE 3** The main structure of ISOM feature map

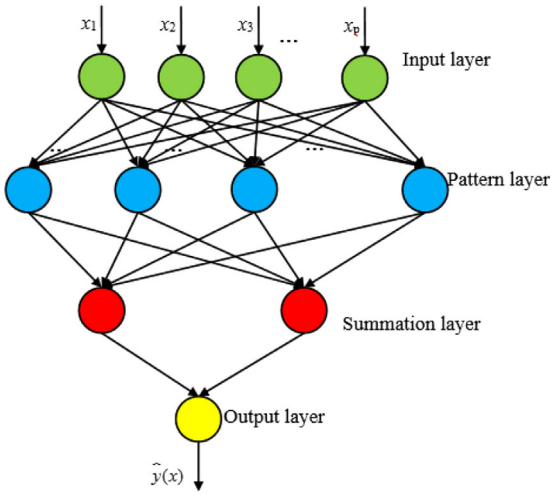


**FIGURE 4** GRNN architecture

which potentially make it simpler for implementation and practice. The main approach of the proposed method is depicted in Figure 4 where four level layers with a special function for each one was taken into consideration. Moreover, GRNN does not simply trap in local points which increase its performance through the large-scale training datasets and well process with high volatility signals [26]. The mathematical modelling of this method is working with PDF $v$ and $u$ as $f(v,u)$. This regression can be presented as [26]:

$$\widehat{u}(v_r) = E\left[u\,|v_r\right] = \frac{\int_{-\infty}^{\infty} u f(v_r, u)du}{\int_{-\infty}^{\infty} u f(v_r, u)du} \tag{4}$$

In this equation, once the $f(v,u)$ is unpredictable it will be evaluated from a sample of $y$ and $v$. In addition, to have the $f(vr; u)$ function [26], the Parzen distribution free is applied on (4) so:

$$f(v_r, u) = \frac{1}{n(2\pi)^{\frac{g+1}{2}} w_1 w_2 \ldots w_g w_u} \sum_{i=1}^{n} e^{-D(v_r, v_i)} e^{-D(u, u_i)} \tag{5}$$

$$D(v_r, v_i) = \frac{(v_r - v_i)^2}{2\sigma^2}, D(u, u_i) = \frac{(u - u_i)^2}{2\sigma^2} \tag{6}$$

Over this model, the $\sigma$ is one of the critical variables which named smooth and will have impact on GRNN's forecasting accuracy. The large and low value of this parameter can influence the input signal and then accuracy of forecasting. This parameter is enhanced by an optimization algorithm through the forecasting error minimization. By substituting (5) in (4), it can be obtained that [26]:

$$\widehat{u}(v_r) = \frac{\sum_{i=1}^{n}\left(e^{-D(v_r, v_i)} \int_{-\infty}^{\infty} u e^{-D(u, u_i)} du\right)}{\sum_{i=1}^{n}\left(e^{-D(v_r, v_i)} \int_{-\infty}^{\infty} e^{-D(u, u_i)} du\right)} \tag{7}$$

where, the $\widehat{u}(v_r)$ is the weighted mean of all $u_i$ parameters.

By the solving the two integrals in this equation it can be written: $\widehat{u}(v_r) = \frac{\sum_{i=1}^{n} u_i e^{-D(v_r, v_i)}}{\sum_{i=1}^{n} e^{-D(v_r, v_i)}}$

To solve the proposed optimization algorithm (minimization the forecasting error as objective function), improved bear smell search algorithm (IBSSA) is considered in this stage. This algorithm is proposed recently in [27]. In this model, the bear absorbs various odours particle to find the pray (optimal solution). The matrix of initial solution is $OM = [O_i]_{n\times k} = [oc^j_i]_{n\times k}$, therefore the odour components can be obtained as:

$$DS_i^j = \begin{cases} MG_i(t - t_{inhale}) + DS_i^{t_{inhale}}, & t_{inhale} \le t \le t_{exhale} \\ DS_i^{t_{exhale}} \exp\left(\dfrac{t_{exhale} - t}{\eta_{exhale}}\right), & t_{exhale} \le \mathrm{t} \end{cases} \tag{8}$$

where, exhalation time defined by $t_{exhale}$, inhalation time by $t_{inhale}$ and constant value for exhalation time defined by $\eta_{exhale}$. Furthermore, receptor sympathies and the odour absorption as well as individuality (see Figure 5) presented by $MG = \{MG_1, MG_2, \ldots, MG_i, \ldots, MG_n\}$ while, the non-negative $MG$ is evaluated as:

$$MG_i(O_i) = \frac{1}{k}\sum_{j=1}^{k} f(oc_i^j), f(oc_i^j) = \begin{cases} 1, & \mathrm{T}_1 \le oc_i^j \\ 0, & \mathrm{T}_1 > oc_i^j \end{cases} \tag{9}$$

where, the odour length in $i$th odour presented by $k$, and threshold variable is depicted by $T_1$. All data are shifted to mitral to obtain the best response by implementation of Li–Hopfield and the Erdi as [27]:

$$\begin{aligned} \dot{X} &= -H_0\Psi_y(Y) - \alpha_x X + \sum L_0\Psi_x(X) + DS \\ \dot{Y} &= W_0\Psi_x(X) - \alpha_y Y + DS_c \end{aligned} \tag{10}$$

where, $X = \{x_1, x_2, \ldots., x_n\}$ and $Y = \{y_1, y_2, \ldots., y_n\}$ are the activities of mitral and granule cells. In addition, $DS = \{ds_1, ds_2, \ldots, ds_n\}$ and $DS_c = \{ds_{c1}, ds_{c2}, \ldots, ds_{cn}\}$ are the outside input to the mitral and the dominant input to the granule cells. The output of the mitral and granule cells are obtained by $\Psi_x(X) = \{f_x(x_1), f_x(x_2), \ldots., f_x(x_n)\}$, and

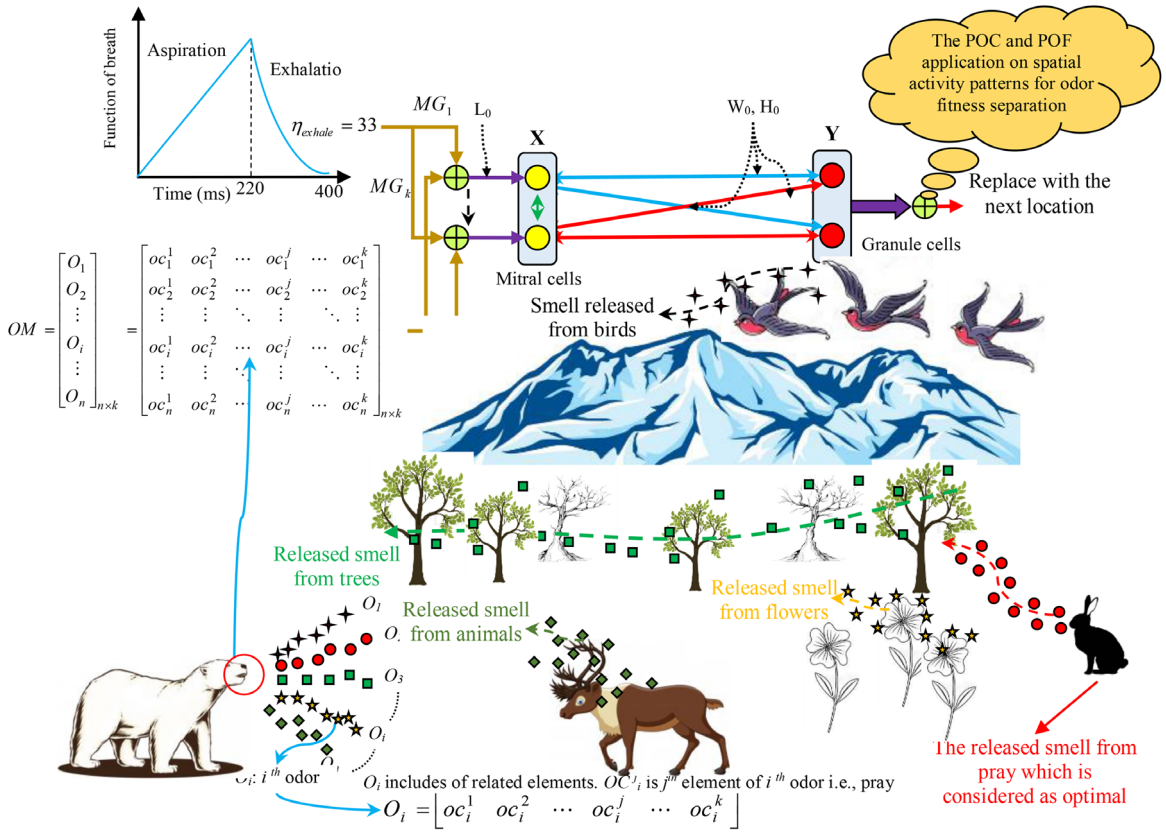


**FIGURE 5** The bear smelling system to find the pray as optimal answer in search environment

$\Psi_y(Y) = \{f_y(y_1), f_y(y_2), \ldots, f_y(y_n)\}$, respectively. Therefore,

$$f_x(x) = \begin{cases} 0.14 + 0.14\tanh\left(\dfrac{x-\varphi}{0.14}\right), & x < \varphi \\ 0.14 + 1.4\tanh\left(\dfrac{x-\varphi}{1.4}\right), & x \geq \varphi \end{cases} \tag{11}$$

$$f_y(y) = \begin{cases} 0.29 + 0.29\tanh\left(\dfrac{x-\phi}{0.29}\right), & x < \phi \\ 0.29 + 2.9\tanh\left(\dfrac{x-\phi}{2.9}\right), & x \geq \phi \end{cases} \tag{12}$$

where, the threshold value presented by $\varphi$, the synaptic-strength connection matrixes are presented by $H_0$, $W_0$ and $L_0$ that shows the formula of granular and mitral cells as well as mitral cells and calculate which presented as:

$$H_{0_i}^j = \begin{cases} \dfrac{rand()}{T_h}, & 0 < d_i^j < T_h \\ 0, & T_h < d_i^j \end{cases} \tag{13}$$

$$W_{0_i}^j = \begin{cases} \dfrac{rand()}{T_w}, & 0 < d_i^j < T_w \\ 0, & T_w < d_i^j \end{cases} \tag{14}$$

$$L_{0_i}^j = \begin{cases} \dfrac{rand()}{T_l}, & 0 < d_i^j < T_l \\ 0, & T_l < d_i^j \end{cases} \tag{15}$$

The connection constants are depicted by $T_h$, $T_w$ and $T_l$. The random value in this equation is generated by Chaotic formula [28] and shown by *rand*(). Also, the formulation of probability odour fitness (POF), probability odour components (POC), and odour fitness (OF) are given by:

$$POC_i = \frac{O_i}{\max(O_i)} \tag{16}$$

$$POF_i = \frac{OF_i}{\max(OF_i)} \tag{17}$$

Moreover, the expected odour fitness (EOF) and distance odour components (DOC) are given by:

$$DOC_i = 1 - \frac{\sum_{i=1}^{k}(POC_j^1 - POC_j^2)}{\sqrt{\sum_{j=1}^{k}(POC_j^1 - POC_j^2)^2}} \tag{18}$$

$$EOF_i = |POF_i - POF^g| \tag{19}$$

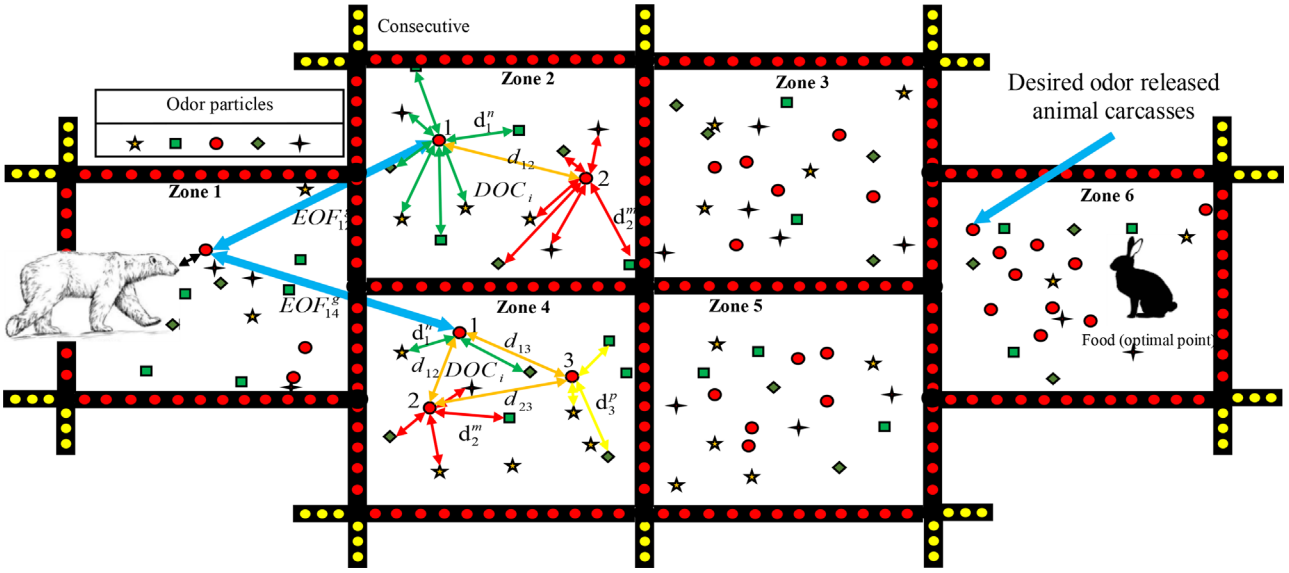


**FIGURE 6** Mechanism of position improvement based on mesh model

The global solution in equations are defined by *g*. Thus, the relationships between odour particle as a source and bear is modelled mathematically and shown in Figure 6. It can be highlighted that the outputs of the brain choose an appropriate method for the next location as well. Each mesh zone, the space of all odours are evaluated by two thresholds as; $\zeta_1, \zeta_2$, so, the next odours are generated with:

$$O_{k+1} = \begin{cases} C_{1,i} \times O_k - rand \times C_{2,i} \times (O_k - O_{best}), \\ DOC_i \le \zeta_1 \& EOF_i \le \zeta_2 \\ C_{3,i} \times O_k - rand \times C_{4,i} \times (O_k - O_{best}), \quad otherwise \end{cases} \tag{20}$$

$$\begin{aligned} C_{1,i} &= -EOF_i \frac{2 - DOC_i}{\zeta_1}, \\ C_{2,i} &= -EOF_i \frac{2 - DOC_i}{\zeta_2} \\ C_{3,i} &= EOF_i \frac{2 - DOC_i}{\zeta_1}, \\ C_{4,i} &= EOF_i \frac{2 - DOC_i}{\zeta_2} \end{aligned} \tag{21}$$

- Long short-term memory: This system is a kind of recurrent NN, see Figure 7, including memory cell and three gates as; forget, input and output. The main conceptual of LSTM is based on cell state same as conveyor belt. In this model, the output of sigmoid layer is set among 0 and 1 while, zero defines "nothing can pass" and one defines "everything can pass" [29].

In addition, a forget gate takes away the data from the cell state as indicated in Figure 7. In this gate, the value of zero

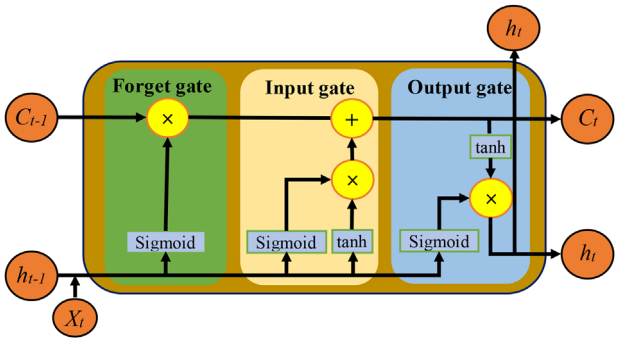


**FIGURE 7** The main model of LSTM network

means "totally forget it" and one means "totally hold it". This gat can be obtained by:

$$f_t = \sigma\left(W_f . \left[h_{t-1}, x_t\right] + b_f\right) \tag{22}$$

The most significant part of LSTM is the input gate which decides whether to kept data in the cell stat. Once this gate is closed, nothing is able to pass to the memory.

This gate includes of two parts as "input gate layer" where the updating of data is done in this part and "hyperbolic tangent layer" which makes a new candidate vector. This gate is modelled as [29]:

$$i_t = \sigma\left(W_i . \left[h_{t-1}, x_t\right] + b_i\right) \tag{23}$$

$$\widetilde{C}_t = \tanh(W_C . \left[h_{t-1}, x_t\right] + b_C) \tag{24}$$

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C}_t \tag{25}$$

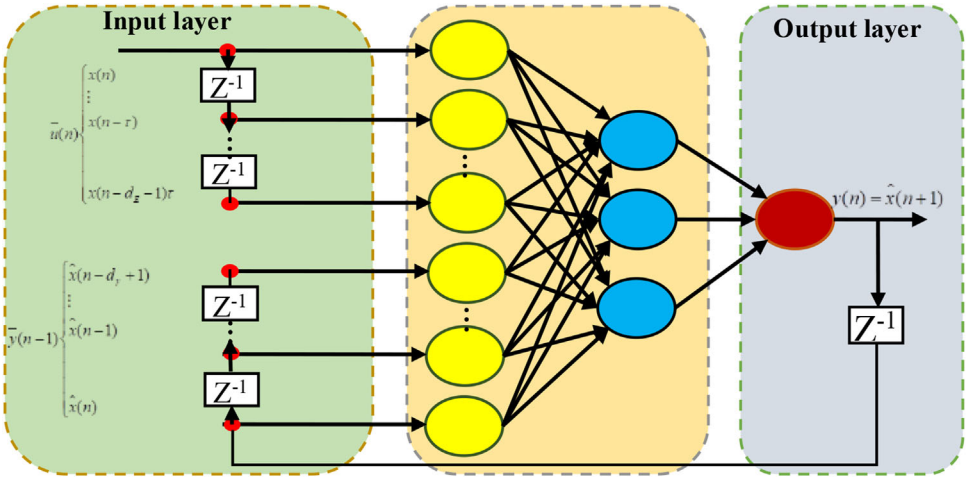


**FIGURE 8** The main structure of NARXNN

In the last part, decision of output data will be made by output gate. This gate is modelled as [29]:

$$O_t = \sigma(W_O.[h_{t-1}, x_t] + b_O) \tag{26}$$

$$h_t = O_t * \tanh(C_t) \tag{27}$$

- Non-linear auto-regressive exogenous NN: This network consists of single delayed feedback loop (which improves the sensitivity in forecaster approach of historical data) based on output regressor with two tapped-delay lines across input–output signals. In this model, the exogenous signal aggregated as:

$$y(n) = f[x(n); \bar{u}(n); \bar{y}(n-1)] \tag{28}$$

$$y(n) = f[x(n), \ldots, x(n - d_E + 1); \hat{x}(n), \ldots, \hat{x}(n - d_y + 1)] \tag{29}$$

In these equations, $\bar{u}(n) \in \mathbb{R}$ and $\bar{y}(n) \in \mathbb{R}$ define the output and input method at $n$ as discrete timestep. Additionally, $d_E \geq 1$ and $d_y \geq 1$define the memory in output and input by $\{d_E, d_y\} \in \mathbb{N} *$ [30]. The main construction of NARXNN is obtainable in Figure 8, which includes a two layer feed-forward network based on linear transfer function. The sigmoid function is evaluated as following for hidden layer:

$$\sigma(x) = \frac{1}{1 + \exp(-\sigma)} \tag{30}$$

This model is able to store the old values of $u(n)$ and $y(n)$ sequences which consists of tap delay lines. Also, in main structure a feedback from output, that is, $y(n)$ is fed the input with a delay.

## 3 | PROPOSED BAYESIAN MODEL AVERAGING

In this strategy, the $y^s{}_t$ is considered as a combination of restricted PDFs for each predictor $f^s{}_{k,t}$ in an ensemble of $K$ members. In PDF process the summation of weights is 1 as a non-negative variable. The forecaster PDF is evaluated by BMA as:

$$p(y \,|\, f_1, \ldots, f_K) = \sum_{k=1}^{K} \omega_k h_k(y \,|\, f_K) \tag{31}$$

In (31), the appropriate selection of kernel, that is, $h_k(y|f_K)$ depends on the problem. In this study, the mentioned PDF divided into two parts for discrete-continuous mixture to solve the clipping issue. To evaluate the clipping probability, the following logistic regression is considered as:

$$\text{logit}(P(y \geq \lambda\rho \,|\, f_K)) \equiv \log \frac{P(y \geq \lambda\rho \,|\, f_K)}{P(y < \lambda\rho \,|\, f_K)} = a_0 k + a_{1k} f_k \tag{32}$$

where, the conditional probability is presented by $P(y < \lambda\rho|f_K)$ where, the solar energy is not clipped while the $f_k$ is the best ensemble member [31]. The continuous kernel is considered as the second part of mixture model to prevent the clipping. This part considered the beta and truncated normal kernel while, the PDF for standard distribution evaluated as:

$$p_\phi(z, \mu, \sigma) = \frac{\phi\left(\frac{z-\mu}{\sigma}\right)}{\sigma\left(\Phi\left(\frac{\rho-\mu}{\sigma}\right) - \Phi\left(\frac{0-\mu}{\sigma}\right)\right)} \tag{33}$$

Similarly, the PDF can be rewritten for beta kernel through considering the beta and alpha upper than zero and

obtained as:

$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)} z^{(\alpha-1)}(1-z)^{(\beta-1)} \tag{34}$$

Using $\mu \equiv \alpha/(\alpha+\beta), 0<\mu<1$ and $\gamma \equiv \alpha+\beta, \gamma>0$, PDF can be given as:

$$p_\beta(z,\mu,\gamma) = \frac{\Gamma(\gamma)}{\Gamma(\gamma\mu)\Gamma(\gamma(1-\mu))} z^{\mu\gamma-1}(1-z)^{(1-\mu)\gamma-1} \tag{35}$$

Three parameters of $\sigma, \mu$, and $\gamma$ are evaluated similarly for the mentioned kernels. Evaluation of kernel mean by scaling factor of $b_k$, and $k$ member is given by:

$$\mu_k = \begin{cases} b_k \frac{f_k}{\rho}, & \text{if betakernel} \\ b_k f_k, & \text{if truncatednormalkernel} \end{cases} \tag{36}$$

To evaluate the $\gamma_k$ and $\sigma_k$, as the shape variables, the distribution standard deviation is considered. In this method, to have positive values of beta and alpha, the variance should be $\sigma^2 < \mu(1-\mu)$ based on maximum quadratic domain value as 0.25. To evaluate the variance of observation based quadratic trend by the height $c_k$, (37) needs to be taken into consideration.

$$\sigma_k^2 = -\frac{c_k}{0.25}(\mu_k - 0.5)^2 + c_k \tag{37}$$

In (37), the large uncertainty provided by large $c_k$ value and vice versa. Having this parameter restricted between 0 and 0.25 guarantees the boundary on $\sigma^2$ for the beta kernel which is suitable for truncated normal distribution. This value can be also evaluated directly with some further steps of beta's shape parameter based on (37) and given by:

$$\gamma_k = \frac{0.25 - c_k}{0.25} \tag{38}$$

For additional simplification, the parameter's number can be reduced which potentially decrease the computational time. After tuning the *a, b* and *c*, fine setting of $\gamma$ is critical for this model while different values can provide different shapes. Specially, zero and maximum power highlight the likelihoods in "∪-shape" distributions. This distribution is occurring once the alpha and beta are lower than 1. To simplify, the predicted distribution mean is given by (39) in minimum value:

$$\underline{\gamma_k} = \begin{cases} \frac{1}{1-\mu_k} & \text{if } \mu_k \le 0.5 \\ \frac{1}{\mu_k} & \text{if } \mu_k > 0.5 \end{cases} \tag{39}$$

This is in fact decreasing the variance amounts as fell in shaded grey zone onto the boundary line. Consequently, the conditional PDF can be presented for each ensemble member $f_k$, that is, greatest prediction and given as:

$$\begin{aligned} h_k(y\,|\,f_k) &= \frac{P(y \ge \lambda\rho\,|\,f_k)}{(1-\lambda)\rho}\mathbf{1}[y \ge \lambda\rho] \\ &+ \frac{P(y < \lambda\rho\,|\,f_k)}{G_k(y\,|\,f_k)\,|_\lambda} g_k(y\,|\,f_k)\mathbf{1}[y < \lambda\rho] \end{aligned} \tag{40}$$

where, the indicator function is presented by **1** and $g_k(y|f_k)$ presents the truncated normal kernel:

$$g_k(y\,|\,f_k) = \begin{cases} p_\beta(\frac{y}{\rho}, \mu_k, \gamma_k) \\ p_\phi(y, \mu_k, \sigma_k) \end{cases} \tag{41}$$

The cumulative distribution function is given by $G_k(y|f_k)$ which is added to balance approximating the discrete element unceasingly over a non-zero width, that is, $(1-\lambda)\rho$. Accordingly, the weighted model can be estimated in (17) by these PDFs conditional members.

## 4 | OBTAINED NUMERICAL RESULTS

The suggested prediction approach is applied over various case studies, in this section. Initially, this strategy is applied on a famous case study and compared with other recent works to proof the strength of suggested approach based on different error criteria. Then, the proposed Arta-Solar test case will be considered to evaluate the proposed forecasting approach. In the following, the results of each case study are separated as:

### 4.1 | Verification of proposed approach

In this section, Site 1 of PV power plant located in China, Tibet is considered with the diverse parallel of 4326.99 m. To have similar and fair comparison, the condition should be considered for all the prediction models, the input data will be composed of global horizontal irradiance (W/m$^2$), solar peak angle (degree), hotness (degree centigrade), comparative moisture percent (%), pressure (mbar), rain (cm), wind acceleration m/s) and wind direction (unit of measurement: degree) that are quoted from National Solar Radiation Database (NSRDB) [25]. In this test case, five error criteria have been taken in to consideration as the prediction intervals coverage probability (PICP), middling coverage error named (ACE), PI normalized average width (PINAW), cover width-based yardstick named (CWC), and Winkler score (WS). The PICP is given by:

$$PICP = \frac{1}{L}\sum_{l=1}^{L} I_l, \quad \text{and } I_l = \begin{cases} 1, & \text{if } R_p^l \in [L_p^l, U_p^l] \\ 0, & \text{otherwise} \end{cases} \tag{42}$$

In which, the *l*th PI lower and upper boundary are presented by $L_p{}^l$ and $U_p{}^l$, respectively. High value of PICP built PIs by

solar data. The ACE is calculated from difference of PICP with PI nominal confidence level (PINC) as:

$$ACE = PICP - PINC \tag{43}$$

To check the PI breadth, the PI normalized average width (PINAW) is considered as [3]:

$$PINAW = \frac{1}{L(Max_p - Min_p)} \sum_{l=1}^{L} (U_p^l - L_p^l) \tag{44}$$

where, the extreme and least value of forecasting are given by $Max_p$ and $Min_p$, respectively. To check the simultaneous evaluation, cover probability and the breadth, the CWC is given by [3]:

$$CWC = PINAW[1 + I(PICP)e^{-\rho(PICP-\nu)}] \tag{45}$$

where, length of the test set is defined by $L$, fixed volume is defined by $C$, and the forecasted and real value of PV signal are defined by $R_p^l$ and $F_p^l$, respectively. Also, the $I$ ($PICP$) is given by:

$$I(PICP) = \begin{cases} 1, & \text{if } PICP < \nu \\ 0, & \text{otherwise} \end{cases} \tag{46}$$

Finally, the WS is formulated as:

$$WS = \frac{1}{L} \sum_{l=1}^{L} S_l \tag{47}$$

$$S_i = \begin{cases} -2(1 - PINC)W_p^l - 4(L_p^l - R_p^l), & \text{if } R_p^l < L_p^l \\ -2(1 - PINC)W_p^l, & \text{if } R_p^l \in W_p^l \\ -2(1 - PINC)W_p^l - 4(R_p^l - U_p^l), & \text{if } R_p^l > U_p^l \end{cases} \tag{48}$$

In (48), the breadth of the $i$th built PI is presented by $W_p^l = U_p^l - L_p^l$. And the PV value in $t$ (time) defined by $R_p^l$.

The suggested approach is compared with [32], in this test case which depicted in Table 1. According to Table 1, the persistence model, autoregressive integrated moving average (ARIMA), quantile regression (QR), persistence, and hybrid model are based on Bayesian method in [32], and compared with the proposed forecasting approach.

In Table 1, the recommended approach could outperform all the prediction models in this test case. Accordingly, performance confirmation of the proposed method is correct and it can be employed in a new test cases. Thus, in the next subsection, a new real-world Engineering test case, that is, Arta-Solar is considered to display the effectiveness of recommended forecast strategy.

**TABLE 1** Comparison of proposed method with [32], in Site 1 based on different error criteria

| | Site 1 | | | | | |
|---|---|---|---|---|---|---|
| **Prediction Models** | **PINC (%)** | **PICP (%)** | **PINAW (%)** | **WS** | **ACE (%)** | **CWC** |
| ARIMA [32] | 90 | 84.35 | 27.68 | −80.15 | −5.65 | 0.34 |
| | 96 | 87.65 | 41.57 | −45.36 | −8.35 | 0.87 |
| | 99 | 92.15 | 52.65 | −15.24 | −6.85 | 1.35 |
| Persistence [32] | 90 | 86.35 | 29.87 | −75.65 | −3.65 | 0.38 |
| | 96 | 87.48 | 37.65 | −44.25 | −8.52 | 0.76 |
| | 99 | 94.98 | 52.21 | −13.24 | −4.02 | 1.42 |
| QR [32] | 90 | 82.68 | 26.45 | −80.22 | −7.32 | 0.37 |
| | 96 | 87.16 | 35.98 | −41.26 | −8.84 | 0.55 |
| | 99 | 93.43 | 54.17 | −12.35 | −5.57 | 0.98 |
| Hybrid Bayesian model [32] | 90 | 85.76 | 30.43 | −72.67 | −4.24 | 0.28 |
| | 96 | 89.76 | 44.23 | −40.68 | −6.24 | 0.36 |
| | 99 | 95.65 | 58.98 | −11.21 | −3.35 | 1.11 |
| Proposed | 90 | 86.89 | 31.98 | −70.12 | −3.11 | 0.14 |
| | 96 | 90.56 | 46.63 | −38.47 | −5.44 | 0.29 |
| | 99 | 86.25 | 60.14 | −10.17 | −3.75 | 0.92 |

## 4.2 | Arta-Solar test case analysis

This part describes the applied method on the new Arta-Solar power plant as described earlier in Section 2.1. To present the accuracy of recommended model, various error criteria were formulated with different behaviours. This comparison will support the validity and well performance of the proposed approach with different test cases and error criteria.

To show the concentration ability of the proposed model, sharpness measure is considered in this stage which can be evaluated on $T$ period-based averaging interval width, that is, $\delta$, for central interval interest ($100\% \times (1 - \rho)$) as:

$$\overline{\delta} = \frac{1}{T} \sum_{t=1}^{T} F_t^{-1}\left(1 - \frac{\rho}{2}\right) - F_t^{-1}\left(\frac{\rho}{2}\right) \tag{49}$$

In (49), the prediction cumulative distribution function is given by $F_t$ at time $t$. The average continuous ranked probability score ($\overline{CRPS}$) is presented as follows which can capture the reliability and sharpness in a metric:

$$\overline{CRPS} = \int_0^1 \frac{1}{T} \sum_{t=1}^{T} QS_\xi(F_t^{-1}(\xi), y_t) d\xi \tag{50}$$

The $\overline{CRPS}$ can be decomposed over all quantiles based on integral of quantile score (QS). The QS of $F_t^{-1}(\xi)$ can be evaluated through (51) at level $\xi \in (0, 1)$:

$$QS_\xi = 2(\mathbf{1}\{y_t \le F_t^{-1}(\xi)\} - \xi)(F_t^{-1}(\xi) - y_t), \tag{51}$$

**TABLE 2** Obtained results of rolling 4-hour forecasting horizons based $\overline{CRPS}$ and SS

| $\overline{CRPS}(\%p)$ | | | | | $SS_{PeEn}(\%)$ | | $SS_{RaEn}(\%)$ | |
|---|---|---|---|---|---|---|---|---|
| **PeEn** | **RaEn** | **EMOS** | **BMA** | **Proposed** | **BMA** | **Proposed** | **BMA** | **Proposed** |
| 9.12 | 6.72 | 7.44 | 6.45 | 5.22 | 36.17 | 44.72 | −6.75 | 8.14 |

* $SS_{PeEn}$ defines SS with forecast reference of PeEn, $SS_{RaEn}$ defines SS with forecast reference of RaEn.

Since the distribution of lower tail is very concerned, which is occurred once solar energy is quite low and will affect the system reliability. For this purpose, three different quantile-weighting functions as considered as; $\omega_1(\xi) = (1-\xi)^2$, $\omega_c(\xi) = \xi(1-\xi)$, and $\omega_r(\xi) = \xi^2$ as left, centre and right tail functions, respectively. Using the weighted QS on (36), the weighted average CRPS can be calculated ($w\overline{CRPS}$). For the last criterion, the skill score of CRPS can be evaluated though:

$$SS = \frac{\overline{CRPS} - \overline{CRPS}_{ref}}{\overline{CRPS}_{ideal} - \overline{CRPS}_{ref}} = 1 - \frac{\overline{CRPS}}{\overline{CRPS}_{ref}} \tag{52}$$

In (52), the average CRPS is defined by $\overline{CRPS}_{ref}$ and ideal prediction is defined by $\overline{CRPS}_{ideal}$.

The calculated forecasting values for the Arta-Solar test case is provided in Table 2 for average performance comparison based rolling 4-h ahead horizon through CRPS. In Table 2, the proposed model is compared by famous forecasting approaches, that is, persistence ensemble (PeEn), Raw ensemble (RaEn), ensemble model output statistics (EMOS), and simple BMA model. In addition, the recommended approach could outperform all other strategies. To have more precise view over this forecasting model, the sharpness and calibration tension is presented in Figure 9. In Figure 9, the interval width and reliability diagrams have been plotted for different comparison models.

Additionally, the graphical presentation of hourly probabilistic prediction is shown in Figure 10, through the validation by a solo signal of the plant average hourly power. In Figure 10, all four models have been compared over 18th April 2021.

As shown in Figure 10, the sharpness of PeEn model is poor in comparison with other methods steadily huge intervals, where this factor for RaEn is unreliable through overrating power. The proposed approach and simple BMA could be evaluated for a better trend, while the simple BMA provides better outcome in comparison with RaEn, PeEn and EMOS. However, the sharpness is missed which leads to loss the accuracy. The proposed prediction approach, evaluates better broader prediction intervals which makes suitable reliability.

Additionally, to present the effectiveness of suggested model Table 3., is presented based on unweighted SS on RaEn for various prediction horizons, that is, 1, 4, 12 h and ultimately one-day.

In this study, the proposed strategy is compared with simple BMA which is able to provide the closest results for the suggested model. The proposed strategy could outperform simple BMA in this table, over all forecasting horizons. Generated forecasting results shows the accuracy and validity of recommended model in test cases with different error criteria and forecasting horizons.

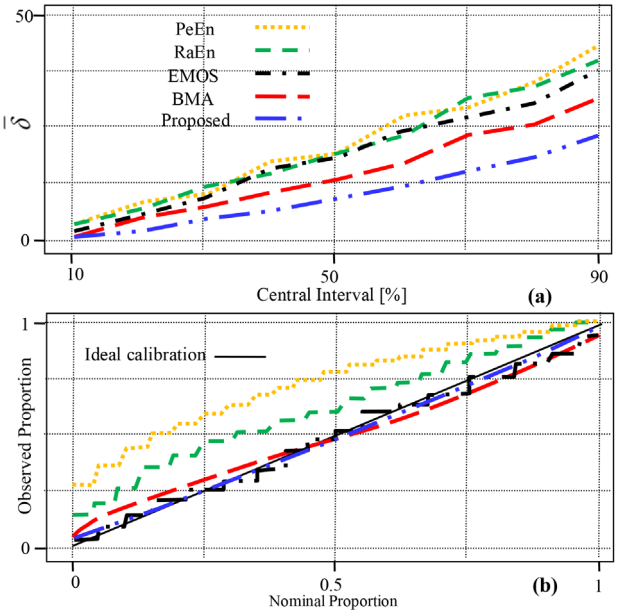


**FIGURE 9** (a) The central interval's average width from 10% to 90%, (b) diagram of reliability for 1th to 99th prediction percentile

## 4.3 | Deterministic forecasting analysis

To presents the efficacy of the suggested approach, deterministic forecasting analysis is also discussed in this stage. For this purpose, Table 4 is presented for comparison of suggested approach with support vector machine (SVM), least squares SVM (LSSVM), persistence model, autoregressive integrated moving average (ARIMA), quantile regression (QR), wavelet transform plus SVM (WT+SVM), and particle swarm optimization plus ANFIS (PSO+ANFIS) models. In this table, three different forecasting criteria have been taken into consideration as normalized mean absolute error (NMAE), normalized root mean square error (NRMSE) and mean absolute percentage error (MMAPE) [1]:

$$\frac{1}{LC}\sum_{l=1}^{L}\left|R_p^l - F_p^l\right| \times 100\% \tag{53}$$

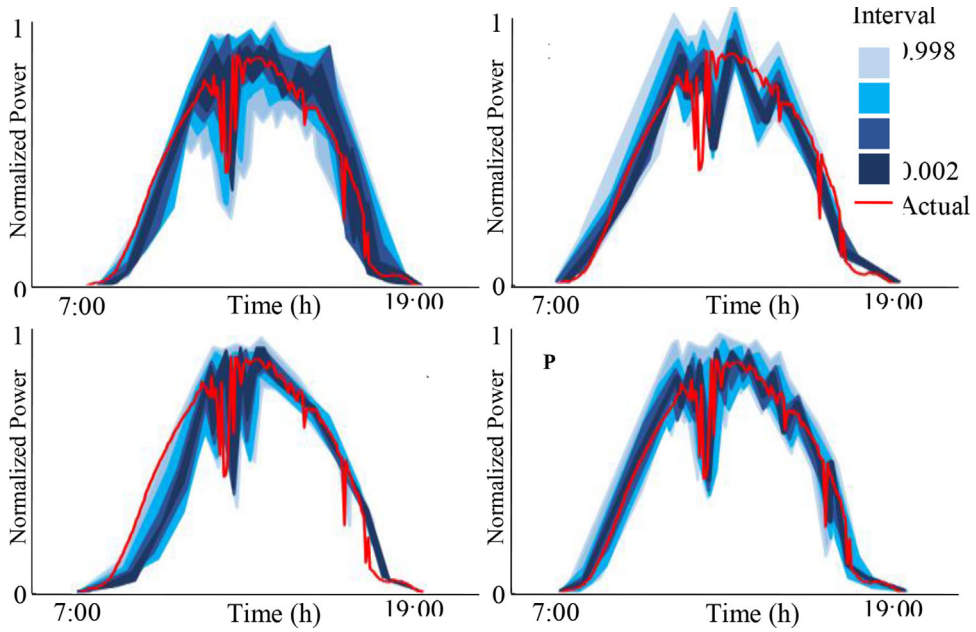


**FIGURE 10** Comparison of suggested approach with other probabilistic methods for intervals in 4-h ahead from 99.8% to 0.2% coverage

**TABLE 3** Obtained results of $\overline{CRPS}$ SS (%) for BMA and proposed model on RaEn at varying lead time

| 1-h | | 4-h | | 12-h | | 24-h | |
|---|---|---|---|---|---|---|---|
| **BMA** | **Proposed** | **BMA** | **Proposed** | **BMA** | **Proposed** | **BMA** | **Proposed** |
| 3.51 | 8.45 | −3.5 | 12.52 | −5.1 | 14.52 | 0.41 | 10.47 |

**TABLE 4** Obtained results of forecasting model comparison with other deterministic approaches

| Prediction Approaches | Arta-Solar test case | | |
|---|---|---|---|
| | **NMAE** | **NRMSE** | **MMAPE** |
| SVM | 8.35 | 11.43 | 14.66 |
| ARIMA | 7.78 | 11.42 | 14.62 |
| Persistence | 7.33 | 10.21 | 13.87 |
| QR | 7.03 | 10.32 | 12.64 |
| PSO+ANFIS | 6.98 | 10.12 | 11.87 |
| WT+SVM | 6.65 | 9.74 | 11.21 |
| LSSVM | 5.87 | 9.12 | 10.15 |
| Proposed | 2.14 | 3.72 | 4.79 |

$$\sqrt{\frac{1}{L}\sum_{l=1}^{L}\left(\frac{R_p^l - F_p^l}{C}\right)^2} \times 100\% \tag{54}$$

$$\sum_{l=1}^{L}\left(\frac{R_p^l - F_p^l}{\sum_{l=1}^{L} R_p^l}\right) \times 100\% \tag{55}$$

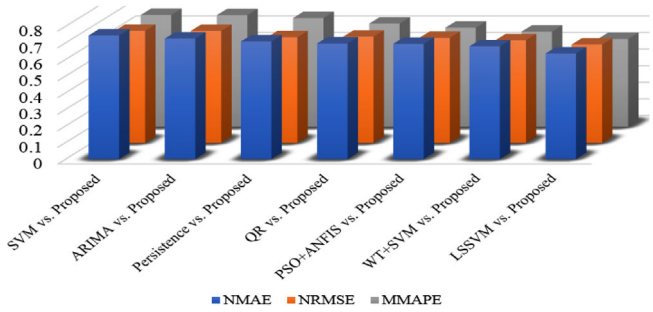


**FIGURE 11** The improvement rate of proposed method in comparison with other forecasting approaches

In these equations, the real signal is defined with $R$ and forecasted value is depicted by $F$. The recommended process could outperform all forecasting approaches on all error criteria. The enhancement of suggested approach in comparison with LSSVM is around (5.87–2.14)/5.87 = 63% for NMAE error criterion.

Similarly, this value for NRMSE and MMAPE are obtained as 59% and 53%, respectively. To have again more precise comparison over the recommended method with other models and providing an improvement rate for all methods, Figure 11 is provided to show the improvements rates for all approaches.

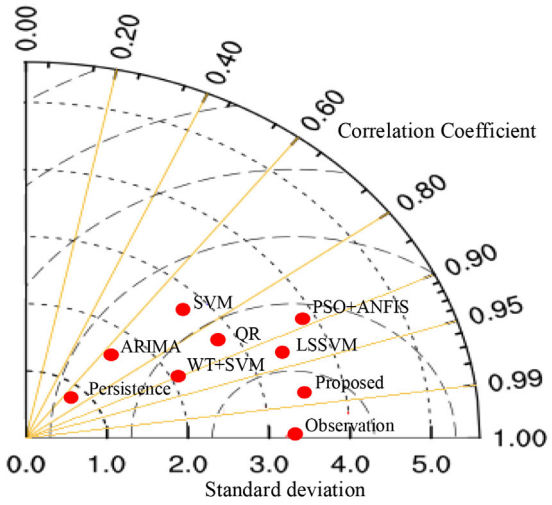


**FIGURE 12** Comparison of proposed model in Taylor diagram

To validate the superiority of the suggested approach, the Taylor diagram, see Figure 12, of the proposed model with other methods is also provided over the Arta-Solar test case, where the presented numerical analysis, the error criteria, standard deviations and outline connections with "observation" for each approach are illustrated. Additionally, due to small value in forecasting error, it is clear that the correlation coefficient of proposed method is superior to the other compared models. This concept is quoted from nearby results to the "observation" as the most accurate strategy and worst model provides farthest point in this figure.

Technically speaking, a forecasting ability in a short time can potentially highlight a better capability of the method; however, computational time is always dependent significantly on the system memory and processor. In this study, a personal desktop with specifications of 1 terabyte of internal memory, 4 GB of RAM, 2.53 CPU of three cores was used. Accordingly, the average duration is less than 2 min, which is quite reasonable performance as compared to other studies and literature. It is worth noting that since other literature have employed different systems, it may not be feasible to fairly compare their performance accurately. However, an approximate comparison is always possible.

## 4.4 | Scalability analysis

In this section, a new analysis is provided to demonstrate the scalability of the suggested approach. As the performance review, the amount of different input data volume (from −40% to 40%) were taken into consideration. Hence, the MAPE criterion was used to evaluate the performance of the proposed method. Table 5 shows the simulation results of proposed analysis. As shown in the second row of Table 5, the proposed method was able to obtain an almost constant MAPE by changing the amount of input data. As another measure of scalability, we have considered the duration of the forecast. In Table 5, in the fourth row, by increasing the forecast period, it still shows acceptable performance according to the MAPE criteria.

## 4.5 | The sensitivity analysis

The proposed optimization algorithms is based on a random initial population, such as determining the initial population number and the number of iterations, while, the self-regulating coefficients is one of the advantages in this algorithm.

Table 6 shows the results of the sensitivity analysis based on presented description. It should be noted that the number of input data is selected according to the 0% mode and 24-h time horizon in Table 5. As shown in Table 6, it can be seen that

**TABLE 5** The scalability analysis in forecasting approach

| Increase the volume of input data | 40% | 30% | 20% | 10% | 0% | 10-% | 20-% | 30-% | 40-% |
|---|---|---|---|---|---|---|---|---|---|
| The MAPE for 24 forecasting horizon | 2.15 | 2.16 | 2.15 | 2.14 | 2.14 | 2.14 | 2.15 | 2.14 | 2.15 |
| Predicted time horizon changes (hours) | 192 | 168 | 144 | 120 | 96 | 72 | 48 | 24 | 12 |
| MAPE | 3.45 | 3.25 | 3.13 | 3.01 | 2.93 | 2.87 | 2.64 | 2.43 | 2.14 |

**TABLE 6** The result of sensitive analysis in prediction based on changes in control parameters of the algorithm

| Number of initial population | 20 | 40 | 60 | 80 | 100 | 120 | 140 | 160 |
|---|---|---|---|---|---|---|---|---|
| MAPE for 24-h horizon | 2.14 | 2.14 | 2.15 | 2.14 | 2.139 | 2.14 | 2.14 | 2.135 |
| Iteration | 100 | 200 | 300 | 400 | 500 | 600 | 700 | 800 |
| MAPE for 24-h horizon | 2.14 | 2.15 | 2.14 | 2.145 | 2.138 | 2.14 | 2.14 | 2.14 |

despite changes in population and iteration, the algorithm was still able to show the best performance.

## 5 | CONCLUSION

In this research study, a new probabilistic prediction process was introduced and recommended to solve and address the solar energy prediction issue. To have an accurate forecasting approach, a synthetic forecasting process based on BMA and Ensemble learning was discussed. Additionally, the proposed model was combined and evaluated with the improved ISOM clustering K-fold cross-validation for the training process and three-stage learner based LSTM, GRNN, and NARXNN. The recommended approach was employed on different real-world engineering case studies through different error criteria. The obtained numerical analysis with different test results validates the model, and the superiority and accuracy of recommended prediction approach in comparison with other models.

## ACKNOWLEDGEMENTS

This work was supported by Collaborative Research Project (CRP) Grant of Nazarbayev University under grant no. 021220CRP0322.

## CONFLICT OF INTEREST

There is no conflict of interest.

## ORCID

*Oveis Abedinia* https://orcid.org/0000-0001-9327-8738
*Mehdi Bagheri* https://orcid.org/0000-0002-8078-9173

## REFERENCES

1. Abedinia, O., Raisz, D., Amjady, N.: Effective prediction model for Hungarian small-scale solar power output. IET Renewable Power Gener. 11(13), 1648–1658 (2017).
2. Abedinia, O., Amjady, N., Ghadimi, N.: Solar energy forecasting based on hybrid neural network and improved metaheuristic algorithm. Comput. Intell. 34(1), 241–260 (2018).
3. Abedinia, O., Lotfi, M., Bagheri, M., Sobhani, B., Shafie-Khah, M., Catalão, J.P.: Improved EMD-based complex prediction model for wind power forecasting. IEEE Trans. Sustainable Energy 11(4), 2790–2802 (2020).
4. Cheng, H., Cao, W.S., Ge, P.J.: Forecasting research of long-term solar irradiance and output power for photovoltaic generation system. In: *2012 Fourth International Conference on Computational and Information Sciences*, pp. 1224–1227. Chongqing, China (2012)
5. Colak, I., Yesilbudak, M., Genc, N., Bayindir, R.: Multi-period prediction of solar radiation using ARMA and ARIMA models. In: *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*, pp. 1045–1049. Chongqing, China (2015)
6. Li, Y., Su, Y., Shu, L.: An ARMAX model for forecasting the power output of a grid connected photovoltaic system. Renewable Energy 66, 78–89 (2014).
7. Al-Alawi, S.M., Al-Hinai, H.A.: An ANN-based approach for predicting global radiation in locations with no direct measurement instrumentation. Renewable Energy 14(1-4), 199–204 (1998).
8. Reikard, G.: Predicting solar radiation at high resolutions: A comparison of time series forecasts. Sol. Energy 83(3), 342–349 (2009).
9. Zhang, N., Behera, P.K.: Solar radiation prediction based on recurrent neural networks trained by Levenberg-Marquardt backpropagation learning algorithm. In: *2012 IEEE PES Innovative Smart Grid Technologies (ISGT)*, pp. 1–7. Chongqing, China (2012)
10. Bae, K.Y., Jang, H.S., Sung, D.K.: Hourly solar irradiance prediction based on support vector machine and its error analysis. IEEE Trans. Power Syst. 32(2), 935–945 (2016)
11. Sperati, S., Alessandrini, S., Delle Monache, L.: An application of the ECMWF Ensemble Prediction System for short-term solar power forecasting. Sol. Energy 133, 437–450 (2016).
12. Gneiting, T., Raftery, A.E., Westveld, A.H. III, Goldman, T.: Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. Mon. Weather Rev. 133(5), 1098–1118 (2005).
13. Baran, S., Lerch, S.: Mixture EMOS model for calibrating ensemble forecasts of wind speed. Environmetrics 27(2), 116–130 (2016).
14. Chai, S., Xu, Z., Jia, Y.: Conditional density forecast of electricity price based on ensemble ELM and logistic EMOS. IEEE Trans. Smart Grid 10(3), 3031–3043 (2018).
15. Aryaputera, A.W., Verbois, H., Walsh, W.M.: Probabilistic accumulated irradiance forecast for Singapore using ensemble techniques. In: *2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC)*, pp. 1113–1118. Chongqing, China (2016)
16. Sloughter, J.M., Gneiting, T., Raftery, A.E.: Probabilistic wind speed forecasting using ensembles and Bayesian model averaging. J. Am. Stat. Assoc. 105(489), 25–35 (2010).
17. Chmielecki, R.M., Raftery, A.E.: Probabilistic visibility forecasting using Bayesian model averaging. Mon. Weather Rev. 139(5), 1626–1636 (2011).
18. Sloughter, J.M.L., Raftery, A.E., Gneiting, T., Fraley, C.: Probabilistic quantitative precipitation forecasting using Bayesian model averaging. Mon. Weather Rev. 135(9), 3209–3220 (2007).
19. Bessa, R.J., Miranda, V., Botterud, A., Zhou, Z., Wang, J.: Time-adaptive quantile-copula for wind power probabilistic forecasting. Renewable Energy 40(1), 29–39 (2012).
20. Zhang, Y., Wang, J.: K-nearest neighbors and a kernel density estimator for GEFCom2014 probabilistic wind power forecasting. Int. J. Forecasting 32(3), 1074–1080 (2016).
21. Lotfi, M., Javadi, M., Osório, G.J., Monteiro, C., Catalão, J.P.: A novel ensemble algorithm for solar power forecasting based on kernel density estimation. Energies 13(1), 216 (2020).
22. Zhang, Y., Wang, J.: GEFCom2014 probabilistic solar power forecasting based on k-nearest neighbor and kernel density estimator. In: *2015 IEEE Power and Energy Society General Meeting*, pp. 1–5. Chongqing, China (2015)
23. Wang, G., Yang, M., Yu, Y.: A short-term forecasting method for photovoltaic power based on ensemble adaptive boosting random forests. In: *2020 IEEE/IAS 56th Industrial and Commercial Power Systems Technical Conference (I&CPS)*, pp. 1–8. Chongqing, China (2020)
24. Li, Y., Yang, Z.: Application of EOS-ELM with binary Jaya-based feature selection to real-time transient stability assessment using PMU data. IEEE Access 5, 23092–23101 (2017).
25. Chen, Y., Ashizawa, N., Yeo, C.K., Yanai, N., Yean, S.: Multi-scale self-organizing map assisted deep autoencoding Gaussian mixture model for unsupervised intrusion detection. Knowledge-Based Syst. 224, 107086 (2021).
26. Hapugoda, J.C., Sooriyarachchi, M.R.: A simulation study for performance comparison between generalized linear mixed modeling (GLMM) and generalized regression neural networks (GRNN) in joint modeling of mixed responses. Tropical Agric. Res. 32(1), 49–57 (2021).
27. Ghasemi-Marzbali, A.: A novel nature-inspired meta-heuristic algorithm for optimization: Bear smell search algorithm. Soft Comput. 24(17), 13003–13035 (2020).
28. Abedinia, O., Amjady, N.: Short-term wind power prediction based on Hybrid Neural Network and chaotic shark smell optimization. Int. J. Precis. Eng. Manuf. Green Technol. 2(3), 245–254 (2015).
29. Shahid, F., Zameer, A., Mehmood, A., Raja, M.A.Z.: A novel wavenets long short term memory paradigm for wind power prediction. Appl. Energy 269, 115098 (2020).
30. Louzazni, M., Mosalam, H.: Dailly forecasting of photovoltaic power using non-linear auto-regressive exogenous method. In: *2020 International*

*Conference on Decision Aid Sciences and Application (DASA)*, pp. 1016–1020. Chongqing, China (2020)
31. Wang, G., Jia, R., Liu, J., Zhang, H.: A hybrid wind power forecasting approach based on Bayesian model averaging and ensemble learning. Renewable Energy 145, 2426–2434 (2020).
32. Abedinia, O., Bagheri, M., Agelidis, V.G.: Application of an adaptive Bayesian-based model for probabilistic and deterministic PV forecasting. IET Renewable Power Gener. 15, 2699–2714 (2021).

**How to cite this article:** Forecasting, S., Abedinia, O., Bagheri, M.: Execution of synthetic Bayesian model average for solar energy forecasting. IET Renew. Power Gener. 16, 1134–1147 (2022).
https://doi.org/10.1049/rpg2.12389