An Integrated GA-PSO-ANFIS Based Hybrid Approach for a Day-ahead PV Power Generation Forecasting in Microgrids

Yordanos K. Semero and J.H. Zhang, Member, IEEE
School of Electrical and Electronic Engineering
North China Electric Power University
Beijing, China
yordanos.kassa@yahoo.com

D.H. Zheng, Senior Member, IEEE
Goldwind Science and Technology Co., Ltd.
Beijing, China

***Abstract*—This paper presents a hybrid approach for forecasting of electricity production in microgrids with solar photovoltaic (PV) installations. An accurate PV power generation forecasting tool essentially addresses the issues resulting from the intermittent and uncertain nature of solar power to ensure efficient and reliable system operation. A day-ahead, hourly mean PV power generation forecasting method based on a combination of genetic algorithm (GA), particle swarm optimization (PSO) and adaptive neuro-fuzzy inference systems (ANFIS) is presented in this study. Binary GA is used to determine important input parameters that significantly influence the amount of output power of a PV generation plant; and an integrated hybrid algorithm combining GA and PSO is used to optimize an ANFIS based PV power forecasting model for the plant. The proposed approach is tested based on practical information of PV power generation data of a practical case study microgrid in Beijing. Evaluation of performance is made with the persistence method as a reference model, and forecasting results are compared with actual scenario. The results validate effectiveness of the proposed methodology as compared with commonly used forecasting approaches. The proposed approach outperformed artificial neural network (ANN), linear regression (LR), and persistence based forecasting models, demonstrating its effectiveness and favorable accuracy.**

***Keywords*: PV Power Forecasting; Feature Selection; Binary Genetic Algorithm; Particle Swarm Optimization; Hybrid Method; ANFIS.**

1. INTRODUCTION

Solar energy is one of the most promising energy sources considered free, clean and abundantly available. For these reasons, it keeps extending its share in electric power generation in the face of diminishing conventional fossil fuel energy sources and rising environmental protection concerns [1]. While solar irradiation is an inexhaustible source of energy, its variability poses operational difficulties in management of electricity supply systems [2]. Balance between energy generation and demand is a critical requirement in operation of electric power systems. High penetration of weather-dependent renewable energy resources like PV into such systems makes power regulation more challenging. PV power generation forecasting plays an important role in mitigating the challenges arising from resource uncertainty in power system networks in general and microgrid systems in particular. Microgrids typically aggregate distributed generation systems and energy storage units with local controllable loads. Day-ahead PV power production forecasting in microgrids is crucial to ensure system stability and to plan optimal unit commitment, economic generation scheduling, energy storage dispatch, and load shedding.

The importance of the issue of solar resource forecasting has drawn the focus of many studies worldwide. Substantial body of studies in the field is mainly concerned with forecasting solar radiation [3-5], which is the single most important parameter in solar power production. Other studies have focused on forecasting of solar energy production directly [6-8]. Several forecasting techniques targeting different forecasting time horizons have been reported in the literature [9-11]. Application of time series modeling techniques for solar radiation and PV power generation forecasting has been demonstrated in references [7-8, 12]. Artificial neural networks based PV power and solar radiation forecasting methodologies have been presented in references [13] and [14] respectively. Other techniques reported in the literature include forecast modeling approaches based on support vector machines (SVM) [15], adaptive neuro-fuzzy (NF) networks [16, 17], evolutionary optimization and other hybrid methods [18-22]. References [23-25] provide comprehensive review of various photovoltaic power forecasting techniques.

One of the most important tasks in data preparation for forecast modeling is variable or feature selection. Feature selection is an optimization process with the aim of choosing a subset of available candidate features to improve prediction performance, by eliminating features with little or no predictive information and also redundant features that are strongly correlated [26]. There are a number of feature selection techniques available in the literature. Feature selection techniques based on evolutionary algorithms such as particle swarm optimization (PSO) and genetic algorithms have been reported to be efficient and flexible in various classification and pattern recognition problems [27-30]. Reference [31] employs binary genetic algorithm for dimensionality reduction to enhance modeling performance for classification problems. A binary genetic algorithm based feature selection technique for regression problems is implemented in this work to enhance the performance of a PV power generation forecasting model.

This paper proposes a GA-PSO-ANFIS based hybrid approach for a day-ahead hourly PV power generation forecasting. The study initially considers forecasted hourly representations of nine weather condition variables as candidate inputs. Binary genetic algorithm is used to

determine a suitable set of predictive variables to improve the efficiency and accuracy of the PV power forecaster. An integrated optimization algorithm that combines genetic algorithm and particle swarm optimization is then used to optimize the ANFIS-based forecast model. The proposed approach benefits from the simplicity and effectiveness of the particle swarm optimization algorithm and the strong global searching capability of the genetic algorithm to iteratively optimize the relatively complex ANFIS structure. The superior forecasting capability of the proposed approach is demonstrated by comparing its performance with three benchmark approaches; namely, persistence, linear regression model, and BP-NN based forecasting methodologies.

The remainder of this paper is organized in three sections. Section two presents overview of the case study microgrid and source of modeling data, followed by the forecasting methodology applied in this work. Section three shows evaluation results and discusses the comparative advantage of the proposed method over other approaches. Finally, section four presents conclusions drawn from training and test results.

2. DATA AND METHODOLOGY

2.1 Modeling Data

The data used to develop the models in this study were collected from a numerical weather prediction model (NWP) and from three PV installation units in a case study microgrid system located in Beijing. Usually, forecast horizons larger than 6 hours exploit outputs from numerical weather prediction models to generate accurate results [6]. Nine weather variables' forecast data obtained from Weather Research and Forecasting (WRF) model were initially considered as candidate predictors. The candidate inputs include temperature (at the height where the PV panels are installed), humidity, air pressure, air density, surface temperature, shortwave solar radiation, low cloud cover, middle cloud cover and high cloud cover at the selected site with 15 minutes resolution covering 1 square kilometer area. The recorded power data correspond to observations of 10 minutes average PV output power of three PV power generation units installed at Goldwind Smart Microgrid system for a period of one year from January 01, 2015 to December 31, 2015. The microgrid system is a megawatt level demonstration project consisting of a 2.5 MW permanent magnet direct drive wind turbine, PV installations totaling 480kWp, two 65kW micro turbine units, two diesel generator units of 200kW and 300kW capacity, and two energy storage units rated at 200kW. Each of the three PV generation units considered in this study has peak installed power generation capacity of 100kW. The weather forecast and power generation data acquired from WRF model and the microgrid respectively were converted into hourly averaged representations of each variable for the specified period of time to form the initial training dataset. The final modeling dataset consisting of the most influential input features and the training targets is prepared using GA based feature selection technique. Binary genetic algorithm with a cost function associated with a Gaussian process regression model as an objective function was formulated to evaluate the relative predictive importance of the weather parameters. Another set of WRF weather forecast data representing the five consecutive days of the first week of January 2016 and the corresponding on-site recorded power data of the three power generation units was used for model testing and evaluation of prediction capability.

2.2 Adaptive Neuro Fuzzy Inference Systems (ANFIS)

ANFIS are a class of adaptive networks which are functionally equivalent to fuzzy inference systems (FIS). They are hybrid intelligent systems which integrate the principles of fuzzy logic and neural networks. These systems are therefore functionally able to benefit from the advantages of both neural networks and fuzzy logic in a single integrated framework [33, 33].

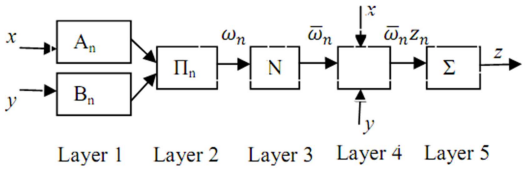

Fig. 1: ANFIS architecture [34]

As depicted in Fig. 1, the ANFIS structure is generally composed of five layers that carryout distinct function. They are referred to as *fuzzification layer*, *rule layer*, *normalization layer*, *defuzzification layer* and *summation layer* respectively from layer one to layer five. The most fundamental parameters of the ANFIS system are called premise and consequent parameters. The premise parameters $\{p_i, q_i, r_i\}$ belong to fuzzy membership functions (MFs) of the *fuzzification layer* used by the ANFIS system to create input spaces by looking for patterns within the input data. The consequent parameters $\{a_i, b_i, c_i\}$ correspond to the MFs of the *defuzzification layer*. Details of mathematical description of operation of the five layers of an ANFIS system can be found in [34]. The premise and consequent parameters are tuned through training to achieve the best values that make up the rules capable of modeling the target system represented by the training data provided. The training algorithm employed to optimize the parameters of the ANFIS system in the proposed approach is a GA-PSO hybrid algorithm as elaborated in the next section.

2.3 The Proposed GA-PSO-ANFIS Hybrid Method

The proposed hybrid method employs binary genetic algorithm in the first stage to select the most relevant subset of input features set; and a combination of genetic algorithm and particle swarm optimization algorithm in the second stage to tune the parameters of an ANFIS model. Genetic Algorithm is an evolutionary search and optimization technique inspired by natural selection. It involves major steps including random

generation of an initial population, ranking and selection of individuals from the current population to produce the next generation, modification of individual solutions by means of crossover and mutation, and evaluation of fitness function [35, 36]. The algorithm iteratively minimizes an objective function, also called a fitness function over successive generations and it is known to have a good global searching capability [37].

Particle swarm optimization is an evolutionary computational search method drawing its inspiration from the social behavior of birds in a flock searching for food. It was first proposed by Kennedy and Eberhart in 1995 [38]. Like genetic algorithm, PSO starts with a randomly initialized set of individuals (particles) and gradually evolves towards the optimal solution through a series of generations involving velocity and position update operations.

$$v_i(t) = \omega v_i(t-1) + \rho_1\big(P_{best} - x_i(t)\big) + \rho_2(G_{best} - x_i(t)) \quad (1)$$

$$x_i(t) = x_i(t-1) + v_i(t) \quad (2)$$

The random variables ρ_1 and ρ_2 are defined as $\rho_1 = r_1 C_1$ and $\rho_2 = r_2 C_2$, where r_1 and r_2 are random numbers generated from a uniform distribution in the range [0, 1], C_1 and C_2 are positive acceleration constants, and ω is an inertia weight. C_1 and C_2 are called cognitive acceleration constant and social acceleration constant respectively. It has been suggested that $C_1 + C_2 \leq 4$ guarantees stability of the PSO algorithm [39]. In this study, C_1 and C_2 are set to 1 and 2 respectively. The inertia weight ω is updated at each iteration according to:

$$\omega = \omega_{damp} * \omega \quad (3)$$

where ω_{damp} is an inertia weight damping ratio. Higher values of inertia weight encourage exploration of the entire search space, while lower values facilitate convergence to a local optimum value. The inertia weight damping ratio plays a key role in balancing global and local search in PSO.

I. Feature Selection

The original raw dataset of weather inputs and corresponding PV power was preprocessed to obtain a dataset suitable for model training. Erroneous measurements and inconsistent data points that could affect modeling efficiency were removed. The remaining dataset which represents weather forecast data with a time-step of 15 minutes was converted into hourly average representations. The recorded power data corresponding to the same time period was similarly prepared. The initial set of input features and corresponding targets was formed in this way and subsequently normalized column-wise between -1 and 1. Binary genetic algorithm was applied at this stage to determine the optimal combination of input variables which provide the best forecasting accuracy. The GA representation of the problem is formulated by presenting a string of binary inputs to the objective function as encoding parameters of the solution. The length of each candidate solution (chromosome) is equal to the number of elements in the input features set, hence nine in this case. Therefore, a '1' in the chromosome depicts the corresponding input variable is selected, while a '0' in the string indicates the corresponding variable is not selected.

A Gaussian process regression-based fitness function is defined to evaluate the predictive capability of different subsets of the original features set. Gaussian process regression (GPR) is a non-parametric probabilistic modeling approach recently widely being used for real supervised learning applications. GPR permits prior probability distribution to be defined over latent functions directly [40, 41]. A Gaussian process $\{f(x), x \in \mathbb{R}^d\}$ can be fully described by its mean function $m(x)$ and covariance (kernel) function $k(x, x')$ [42]. The mean function is often assumed to be zero, and the covariance function defines proximity or likeness between input data points. That is, for points x_1 and x_2 in the input space that are similar, the corresponding output values $f(x_1)$ and $f(x_2)$ will be similar too. Thus, training points near a particular test point are informative for the prediction at that point. The kernel function is therefore a crucial component in the Gaussian process predictors as it encodes the prior assumptions on the latent function, such as the smoothness and scale of the variation [43]. The covariance function that relates two functional values evaluated at set points x_1 and x_2 is given as

$$k(x_1, x_2) = E[(f(x_1) - m(x_1))(f(x_2) - m(x_2))] \quad (4)$$

where $m(x)$ is the mean function which can be defined as

$$m(x) = E\big(f(x)\big) \quad (5)$$

The Gaussian process can then be expressed in terms of the mean function and kernel function as

$$f(x_1) \sim GP[m(x_1), k(x_1, x_2)] \quad (6)$$

While there are many different types of covariance functions, we have used the squared exponential covariance function in this study, which is one of the most commonly used kernel functions. The squared exponential covariance function is given as

$$k\big(x_i, x_j|\theta\big) = \sigma_f^2 \exp\left[-\frac{1}{2}\sum_{k=1}^{d}\left(\frac{\big(x_{i,k} - x_{j,k}\big)^2}{l_k^2}\right)\right] \quad (7)$$

where σ_f^2 is the noise signal variance linked to the overall function variance, σ_l is length-scale parameter, d is input dimension, and $\theta = \{\sigma_f^2, l_1 \ldots l_k\}$ is a vector of hyper-parameters. The values of the noise variance σ_f^2 and hyper-parameters θ of the kernel function are estimated from the modeling data during training. Further details on Gaussian process regression modeling could be found in [40-43]. The fitness of different feature subsets (presented to the GA in the form of chromosomes with binary elements) is evaluated

based the MSE (mean squared error) of prediction residuals introduced by GPR model $f(x)$ fitted for each subset of features.

$$L_{fit} = \frac{1}{n}\sum_{i=1}^{n}(f_i - t_i)^2 \qquad (8)$$

where t is a vector of training targets and n is number of instances in the training data.

The genetic algorithm starts by randomly generating an initial generation of individuals from which it starts evaluation of fitness of each possible solution candidate. The objective of the GA is to minimize the fitness function (MSE) by selecting a combination of input variables with the best fitness over successive iterations. A section of the population with best fitness, called *elite children*, will be copied and directly passed to the next generation. The algorithm is set to copy the best 10 percent of the current population (2 chromosomes) and pass to the next generation. New individuals are created by combining the genetic characteristics of parent individuals through *crossover* operators; and by making random changes to the genes of the individual parents through *mutation* operation. *Scattered* crossover function, which creates a random binary vector of the same length as the solution chromosomes and picks the genes of the first parent where the vector is a 1 and the genes from the second parent where the vector is a 0, is used.

TABLE I: DESCRIPTION OF FEATURE SELECTION ALGORITHM

Step 1: Dataset Preparation and normalization
- Convert all variables in to hourly average representations
- Normalize dataset column wise between -1 and 1.

Step2: Initialization
- Randomly generate initial population of N (*population size*) chromosomes of length l (*number of features*). $N = 20$ and $l = 9$.
- Convert chromosomes in to binary bit strings.

Step 3: Develop GPR models for feature subsets
- Create feature subsets using binary chromosomes
- Model GPR for new dataset

Step 4: Fitness evaluation
- Evaluate GPR model for each feature subset
- Calculate *MSE* (*fitness function*) for each chromosome

Step 5: Generate new population
- Rank chromosomes based on their fitness value
- Select *elite children*– first 2 best individuals
- Perform *crossover* (14 crossover children)and *mutation*(4 mutation children)
- Get a new population for next generation

Step 6: Check whether termination criteria are met
- Generations (*MaxGen = 100*)
- Stall generations (*MaxStallGen = 50*)
- If neither criterion is satisfied, go back to Step 3. Otherwise proceed to Step 7.

Step 7: Terminate the algorithm and output the results.

Selection of parent chromosomes for crossover operation is carried out using the *Tournament* selection method. In this method, four candidate chromosomes are randomly chosen from the population for parentship and the individual with the best fitness is picked to be a parent. This process is repeated to get the second parent for crossover operation. The *crossover fraction*, a parameter that specifies the fraction of the next generation apart from the elite children that will be produced by crossover operation, is set to 0.8. *Uniform mutation* operation is applied to produce the remaining portion of chromosomes of the next generation. In uniform mutation, the algorithm generates a vector of random numbers from a uniform distribution. The value of each number is then compared to a mutation probability *rate*. If the value of the mutation probability *rate* is greater, the corresponding gene in the chromosome is flipped (0 to 1, or 1 to 0), otherwise it is left unchanged. Mutation provides genetic diversity in the population and ensures broader search space for the algorithm [37]. The detailed binary GA-based feature selection algorithm is described in Table I. The parameters of the GA applied implement the feature selection algorithm are summarized in Table II.

TABLE II: PARAMETERS OF GA ALGORITHM

Parameter	Value
Population size (*N*)	20
Length of chromosome (*l*)	9
Population Type	bitstring
Max. number of generations (*MaxGen*)	100
Stall generations (*MaxStallGen*)	50
Number of elite chromosomes	2
Selection method	Tournament
Tournament size	2
Crossover	Arithmetic
Crossover fraction	0.8
Mutation	Uniform
Mutation *rate*	0.01

II. Forecast Modeling

An initial fuzzy inference structure (FIS) is generated from the training data using fuzzy c-means (FCM) clustering with all its premise and consequent parameters randomly initialized. The premise and consequent parameters of the generated ANFIS model are iteratively obtained to determine the size of each chromosome/particle for setting up the optimization problem. These parameters constitute the set of variables to be optimized by the hybrid GA-PSO algorithm. The ANFIS model is evaluated at all candidate solutions to determine the strength of each chromosome/particle. The RMSE of the residuals produced by the ANFIS model is used as a criterion to define the fitness function for the optimization problem.

TABLE III: PARAMETERS OF GA ALGORITHM

Parameter	Value
Population size (*N*)	50
Max. number of generations (*MaxIt*)	1000
Number of elite chromosomes	2
Selection method	Roulette wheel
Crossover function	Scattered
Crossover fraction	0.8
Mutation function	Uniform
Mutation *rate*	0.1

TABLE IV: PARAMETERS FOR PSO ALGORITHM

Parameter	Value
Swarm size ($PopSize$)	50
Max. number of iterations ($MaxIt$)	1000
Cognitive acceleration constant (C_1)	1
Social acceleration constant (C_2)	2
Inertia Weight (ω)	1
Inertia weight damping ratio (ω_{damp})	0.99

TABLE V: DESCRIPTION OF HYBRID GA-PSO ALGORITHM

Step 1: FIS initialization
- Arrange predictor and target variables column-wise
- Generate initial FIS structure from the modelling dataset

Step 2: Generation of initial solutions
- Obtain parameters of initial FIS structure and determine number of variables $VarSize$
- Randomly generate $PopSize$ particles P of length $VarSize$ each and with positions $p_{k,l} \in [P_{min}, P_{max}]$ where k = 1, 2, 3.... $PopSize$ and k = 1, 2, 3.... $VarSize$
- Randomly generate $PopSize$ chromosomes C of length $VarSize$ each and with genes $C_{k,l} \in [P_{min}, P_{max}]$ where k = 1, 2, 3.... $PopSize$ and k = 1, 2, 3.... $VarSize$
- Initialize velocity of each particle to zero
- Initialize best cost of PSO to infinity
- Initialize global best cost to infinity

Step 3: Create next generation of solutions
- Update velocity, position and inertia weight of particles'
- Perform *elite selection, crossover* and *mutation*

Step 4: Assign parameters of solution to ANFIS structure
- Get vector of variables of each particle/chromosome *sol*
- For each membership function of each predictor variable, assign corresponding values of *sol* to the input MF parameters
- For each membership function of the target variable, assign corresponding values of *sol* to the output MF parameters

Step 5: Evaluate cost of candidate solutions
- Evaluate cost of each particle p_i and update personal best $PPbest$ and global best $PGbest$ of PSO
  If $Cost_i < Cbest_i$:
  $PPbest \leftarrow p_i$
  $Cbest_i \leftarrow Cost_i$, and
  If $Cbest_i < Cbest_G$:
  $Cbest_G \leftarrow Cbest_i$
  $PGbest \leftarrow PPbest$
- Evaluate fitness of each chromosome c_i and find best GA solution $GAbest$ and its fitness $fbest_i$
- Apply global solution update rules

Step 6: Check whether termination criterion is met
- Iterations (*MaxIt = 2000*)
- If the criterion is not satisfied, go back to Step 3. Otherwise proceed to Step 7.

Step 7: Terminate the algorithm and output optimized model

The parameters corresponding to the GA and PSO algorithm used to train the ANFIS structure in this study are as shown in Table III and Table IV respectively. The parameters of the GA and PSO were selected after a large number of trial simulation runs. The PSO algorithm updates the velocity and position of each particle in the swarm according to update rules (1) and (2), whereas the genetic algorithm performs elite selection, crossover and mutation operations over successive generations until a preset maximum number of iterations is attained. The hybrid training algorithm takes in to account all variables of a chromosome/particle (i.e. all antecedent and consequent parameters of the ANFIS structure) at each iteration. All chromosomes and particles of the same generation are evaluated and ranked in terms of their fitness. The algorithm works by comparing the fitness of the best solutions achieved by the GA and PSO at each iteration and choosing the better as the global best solution. If the best particle of the PSO has achieved a better fitness than the best chromosome from the GA population, parameters of the best chromosome are updated to assume the variables of the best particle. On the other hand, if the GA has achieved better solution, variables of the best particle are replaced by the genes of the best chromosome. In this way, the algorithm which has produced the best solution sets the global best solution. This process is repeated until the convergence criterion is met. The variables optimized using the hybrid algorithm are finally assigned to the ANFIS model as antecedent and consequent parameters to form the final ANFIS model and the training process is concluded. The summary of the steps involved in implementing the second stage of the proposed PV power prediction approach are described in Table V.

3 RESULTS AND DISCUSSION

3.1 Forecasting Accuracy Evaluation Metrics

Several standard error metrics were used to evaluate the proposed PV power forecasting strategy. For a target sequence t, forecast sequence f with N time steps and maximum recorded power P_m, root mean square error (*RMSE*), mean absolute error (*MAE*), and normalized mean absolute error ($nMAE$) criteria are expressed as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(t_i - f_i)^2} \quad (9)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|f_i - t_i| \quad (10)$$

$$nMAE = \frac{100}{N}\sum_{i=1}^{N}\frac{|f_i - t_i|}{P_m} \quad (11)$$

The *skill* criterion evaluates the improvement in prediction capability of models over that of the benchmark persistence prediction model.

$$s = \left(1 - \frac{RMSE}{RMSE_P}\right) * 100 \quad (12)$$

3.2 Selection of Predictors

The binary GA based feature selection method was implemented and applied to the datasets of each of the three PV generation units in the system. Description of the datasets in terms of sample size and feature dimensions is presented in Table VI. The feature selection algorithm was applied to each dataset and the experimental results obtained for the three normalized datasets are provided in Table VII.

TABLE VI: DESCRIPTION OF DATASETS

DATASET	SAMPLES	DIMENSION
PV1	468	9
PV2	711	9
PV3	1466	9

Considering the fact that all the input parameters are weather condition variables which could exhibit significant interrelationships, it can be inferred from the results that the number of features selected by the binary GA-based FS algorithm is significantly smaller than the number of features in the original dataset in each case. Moreover, the closeness between the mean and best fitness values indicates the effectiveness of the employed feature selection methodology. For dataset of PV1, the improvement in MSE when only the selected variables are used to model the data using GPR approach is over 19 percent. Similarly, selected subset datasets of PV2 and PV3 have provided improvements of roughly 8.5 and 5.1 percent respectively over the original dataset with respect to the MSE criterion.

TABLE VII: DESCRIPTION OF DATASETS

PV UNIT	SELECTED FEATURES	MEAN FITNESS ($\times 10^{-3}$)	BEST FITNESS ($\times 10^{-3}$)	WITHOUT FS ($\times 10^{-3}$)
PV1	2, 3, 4, 5, 6, 7, 8	4.12	3.76	4.7
PV2	1, 4, 5, 6, 7, 8	0.13	0.12	0.13
PV3	4, 5, 6, 7, 8, 9	0.265	0.259	0.273
ALL	4, 5, 6, 7, 8			

Features 4, 5, 6, 7, and 8, which stand for surface temperature, irradiance, low cloud cover, middle cloud cover, and high cloud cover respectively, have been selected in all the three cases. Therefore, for mere purpose of consistency, these five variables have been selected to constitute the input dataset for PV power generation forecast modeling for all the three generation units.

3.3 Experimental Results

In this study a combination of genetic algorithm and particle swarm optimization is designed to effectively optimize an ANFIS model for the intended PV power forecasting task. Historical weather data of the selected variables for the aforementioned period were given as the inputs along with the corresponding PV output power data as the training targets for all the modeling approaches. The performance of the proposed forecasting technique as applied to the three PV power generation units is evaluated and compared with the benchmark forecasting methods. Fig. 2 shows comparison of actual values and forecasted values for PV1, PV2 and PV3 during training stage considering the first 50 data points. Table VIII summarizes evaluation of the proposed hybrid approach and that of an ANN and LRM for individual units during model training stage using three different error metrics. For all PV units, the proposed GA-PSO-ANFIS based hybrid strategy has more efficiently captured the trends of the data. With respect to the $nMAE$ criterion, relative percentage improvements over ANN approach for PV1, PV2, and PV3 are 23.36, 5.16 and 12.66 respectively. Taking into account the RMSE criterion, the corresponding figures are 24.81, 8.52 and 12.18 percent respectively.

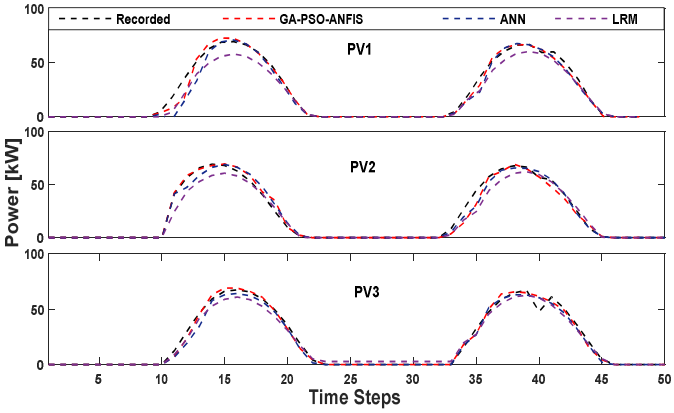

Fig. 2: Comparison of models during training stage

TABLE VIII: EVALUATION OF HYBRID MODEL AGAINST BP-NN MODEL

PV UNIT	METHOD	RMSE	MAE	$nMAE$
PV1	PSO-ANFIS	5.09	3.47	4.43
	BP NN	6.77	4.52	5.78
	LRM	7.54	5.27	6.74
PV2	PSO-ANFIS	5.15	3.49	4.58
	BP NN	5.63	3.68	4.83
	LRM	7.35	5.49	7.20
PV3	PSO-ANFIS	6.20	4.65	6.0
	BP NN	7.06	5.32	6.87
	LRM	7.77	5.88	7.59

The proposed hybrid forecasting approach and the three benchmark techniques were applied on a test dataset representing hourly weather forecast data of five consecutive days from January 01, 2016 to January 05, 2016. The results of 24 hours ahead PV power generation forecasting during the test days further demonstrated the superiority of the proposed hybrid forecasting technique over the other methods. The proposed hybrid PV power generation forecasting method has consistently outperformed the ANN, LRM and persistence based methods regarding RMSE, MAE and NMAE criteria as demonstrated in Table IX through Table XI. Table IX and Table X show comparison of forecasting performance with regard to $RMSE$ and MAE criteria as the forecasting methods are applied to individual generation units. The corresponding results with regard to $nMAE$ are similarly tabulated in Table XI. The proposed hybrid approach has yielded an average RMSE performance improvement of 25.98 percent over ANN method, the second best performing approach, in PV1 over the test period. In contrast, PV3 has seen the least average performance improvement among the three units, where the proposed hybrid approach has reduced the RMSE obtained using the ANN method by about 11.3 percent. Fig. 3 displays

measured power and predicted power of the combined system (combining all three generation units) using the hybrid GA-PSO-ANFIS model and the benchmark approaches during the five days test period. Comparison of daily forecasting error distribution of the combined plant introduced by different approaches during the test period is illustrated in Fig. 4. It was observed that the forecasting techniques produce lower prediction errors during the hours on both sides of noon, where the solar radiation, which is the most important factor for power production among the input parameters, and surface temperature remain essentially constant.

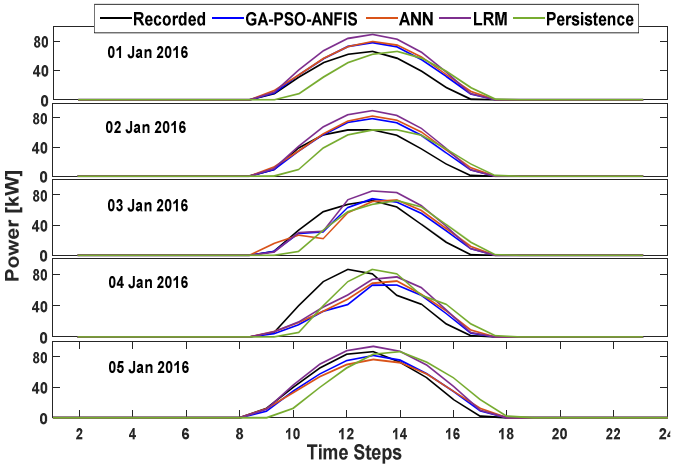

Fig. 3: Comparison of prediction using different approaches

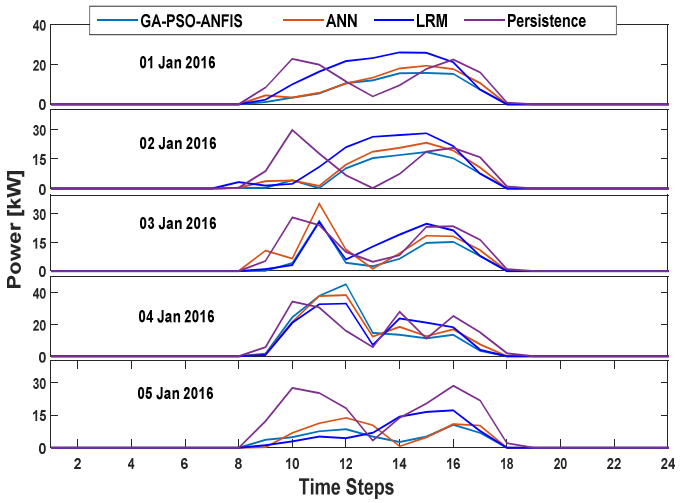

Fig. 4: Comparison of prediction errors over the test period

The test results demonstrate that the proposed hybrid approach outperformed the three benchmark approaches throughout the test period with respect to all performance evaluation metrics. This was validated across all three PV generation units considered in the study. Similarly, the forecasting performance of the models considering the combined generation plant was evaluated. The results are consistent with the findings considering individual PV units. The proposed hybrid technique has provided the best predictive capability with respect to four different performance evaluation criteria over the test period as presented in Table XII. With respect to the RMSE criterion, the hybrid approach was found to have achieved an average reduction of 10.95, 30.3 and 25.5 percent relative to BP-NN, persistence and LRM methods respectively over the whole test period. The average NMAE experienced by the proposed approach over the five days test period is 5.31 percent. This accounts for 82.45%, 63.74% and 60.68% of the corresponding values obtained from the BP-NN, persistence and LRM based models respectively over similar period.

TABLE IX: EVALUATION OF DIFFERENT APPROACHES W.R.T RMSE

		DAY1	DAY2	DAY3	DAY4	DAY5	AVERAGE
PV1	PSO-ANFIS	1.16	1.43	3.38	2.79	1.38	2.08
	BP NN	2.86	2.95	3.03	3.72	1.49	2.81
	PERSISTENCE	3.05	3.09	3.47	3.93	3.89	3.49
	LRM	3.52	3.43	3.66	3.78	2.36	3.35
PV2	PSO-ANFIS	1.62	1.85	4.20	6.86	3.97	3.70
	BP NN	1.52	1.79	7.55	6.91	3.52	4.26
	PERSISTENCE	3.45	3.48	3.87	4.86	4.53	4.04
	LRM	3.5	3.66	3.47	5.57	2.32	3.70
PV3	PSO-ANFIS	4.41	4.63	3.84	4.65	2.10	3.92
	BP NN	4.75	4.89	4.75	5.16	2.57	4.42
	PERSISTENCE	3.29	3.48	3.78	4.5	4.22	3.85
	LRM	5.74	5.71	4.83	5.01	4.04	5.07

TABLE X: EVALUATION OF DIFFERENT APPROACHES W.R.T MAE

		DAY1	DAY2	DAY3	DAY4	DAY5	AVERAGE
PV1	PSO-ANFIS	0.62	0.72	1.91	1.39	0.68	1.06
	BP NN	1.65	1.57	1.71	1.85	0.9	1.53
	PERSISTENCE	1.75	1.68	1.90	2.18	2.25	1.95
	LRM	3.09	2.63	2.88	2.98	2.28	2.77
PV2	PSO-ANFIS	0.70	0.91	1.63	3.10	1.47	1.56
	BP NN	0.74	0.95	3.74	3.43	1.83	2.14
	PERSISTENCE	1.96	1.89	2.14	2.65	2.58	2.24
	LRM	2.43	2.52	2.82	3.33	1.69	2.56
PV3	PSO-ANFIS	2.37	2.42	2.04	2.42	1.11	2.07
	BP NN	2.78	2.75	2.61	2.86	1.37	2.47
	PERSISTENCE	1.82	1.75	1.99	2.38	2.38	2.06
	LRM	4.79	4.74	3.7	4.17	3.65	4.21

TABLE XI: EVALUATION OF DIFFERENT APPROACHES W.R.T NMAE

		DAY1	DAY2	DAY3	DAY4	DAY5	AVERAGE
PV1	PSO-ANFIS	2.95	3.59	8.07	5.32	2.53	4.49
	BP NN	7.82	7.78	7.49	7.06	3.33	6.7
	PERSISTENCE	8.33	8.33	8.33	8.33	8.33	8.33
	LRM	14.65	13.02	12.61	11.41	8.44	12.03
PV2	PSO-ANFIS	2.99	4.02	6.35	9.76	4.76	5.58
	BP NN	3.15	4.18	14.61	10.78	5.93	7.73
	PERSISTENCE	8.33	8.33	8.33	8.33	8.33	8.33
	LRM	10.36	11.09	10.99	10.45	5.45	9.67
PV3	PSO-ANFIS	10.86	11.51	8.54	8.464	3.90	8.65
	BP NN	12.73	13.05	10.93	10.01	4.78	10.3
	PERSISTENCE	8.33	8.33	8.33	8.33	8.33	8.33
	LRM	21.93	22.55	15.51	14.6	12.77	17.47

TABLE XII: ERROR EVALUATION FOR THE WHOLE PLANT

METRIC	METHOD	DAY1	DAY2	DAY3	DAY4	DAY5	AVERAGE
RMSE	PSO-ANFIS	6.68	7.33	7.33	14.10	4.01	7.89
	BP NN	7.95	8.37	8.12	15.42	4.46	8.86
	PERSISTENCE	9.76	10	11.09	13.21	12.54	11.32
	LRM	11.6	11.95	10.21	12.94	6.27	10.59
MAE	PSO-ANFIS	3.59	3.72	3.41	6.86	2.29	3.98
	BP NN	4.56	4.41	4.77	7.94	2.25	4.79
	PERSISTENCE	5.53	5.3	6.03	7.21	7.21	6.26
	LRM	6.85	7.33	6.47	7.53	3.69	6.37
NMAE	PSO-ANFIS	5.41	5.86	4.71	7.93	2.65	5.31
	BP NN	6.87	6.94	6.59	9.18	2.61	6.44
	PERSISTENCE	8.33	8.33	8.33	8.33	8.33	8.33
	LRM	10.32	11.52	8.95	8.7	4.26	8.75
SKILL	PSO-ANFIS	31.56	26.7	33.9	-6.74	68.02	30.3
	BP NN	18.55	16.3	26.78	-16.7	64.43	21.73
	LRM	-18.8	-19.5	7.94	2.04	50	6.45
	PERSISTENCE	-	-	-	-	-	-

In terms of the *Skill (s)* criterion, averaged over the test period, the LRM model has performed slightly better than the persistence approach, despite falling behind on the first and second test days. Both the proposed approach and the NN method provided significantly improved forecasting skills averaging to 30.3 and 21.73 respectively. The proposed approach has easily beaten both NN and LRM models throughout the test period. The results summarized in Table IX through Table XII validate the superior PV power generation forecasting capability of the integrated GA-PSO-ANFIS based hybrid technique by employing daily weather conditions data at a reasonable accuracy with improved precision compared to other forecasting approaches.

4 CONCLUSIONS

With the rapid growth of photovoltaic energy generation, reliable and accurate short-term PV power forecasting tools are needed. This paper proposes an integrated GA-PSO-ANFIS based hybrid technique for short term photovoltaic power generation forecasting. The proposed method implements binary GA based feature selection strategy to eliminate insignificant variables and applies a combination of GA and PSO to optimize a forecasting model. A GPR model based fitness function was implemented to enable the binary GA to significantly reduce the number of input features required to achieve improved forecast modeling. An integrated GA-PSO algorithm is then used to optimize the relatively complex ANFIS structure for forecast modeling. Performance of the proposed technique is compared with ANN, LRM and persistence methods. Results show that the proposed method has the capability of accurately forecasting day ahead hourly PV power generation with substantial performance improvement over other techniques. Model testing over a five days test period returned daily average normalized forecast errors essentially lower than 8 percent; demonstrating the effectiveness of the proposed approach for short term PV power forecasting.

REFERENCES

[1] M. Oliver and T. Jackson, "The market for solar photovoltaics", *Energy Policy*, Vol. 27, pp. 371-385, 1999.

[2] M. Zamo, O. Mestre, P. Arbogast and O. Pannekoucke, "A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production, part I: Deterministic forecast of hourly production", *Solar Energy*, Vol. 105, pp. 792-803, 2014.

[3] A. Azadeh, A. Maghsoudi, and S. Sohrabkhani, "An integrated artificial neural networks approach for predicting global radiation", *Energy Conversion and Management*, 50(6), pp.1497-1505, 2009.

[4] R. Marquez and C.F.M. Coimbra, "Forecasting of global and direct solar irradiance using stochastic learning methods, ground experiments and the NWS database", *Solar Energy*, 85(5), pp. 746-756, 2011.

[5] A. Mellit and A. MassiPavan, "A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy", *Solar Energy*, 84(5), pp. 807-821, 2010.

[6] D. P. Larson, L. Nonnenmacher and C. F.M. Coimbra, "Day-ahead forecasting of solar power output from photovoltaic plants in the American Southwest", *Renewable Energy*, Vol. 91, pp. 11-20, 2016.

[7] P. Bacher, H. Madsen and H. Nielsen, "Online short-term solar power forecasting", *Solar Energy*, Vol. 83, pp. 1772–1783, 2009.

[8] Y. Li, Y. Su and L. Shu, "An ARMAX model for forecasting the power output of a grid connected photovoltaic system", *Renewable Energy*, Vol. 66, pp. 78–89, 2014.

[9] R. Perez, S. Kivalov, J. Schlemmer, K. Hemker Jr., D. Renné and T.E. Hoff, "Validation of short and medium term operation solar radiation forecasts in the US", *Solar Energy*, Vol. 84, pp. 2161-2172, 2010.

[10] F. Wang, Z. Mi, S. Su and H. Zhao, "Short-term solar irradiance forecasting model based on artificial neural network using statistical feature parameters", *Energies*, Vol. 5, pp. 1355-1370, 2012.

[11] C.W. Chow, B. Urquhart, M. Lave, A. Dominguez, J. Kleissl, J. Shields and B. Washom, "Intra-hour forecasting with a total sky imager at the UC San Diego solar energy tested", *Solar Energy*, 85 (11) , pp. 2881-2893, 2011.

[12] Y. Chu., B. Urguhart, S. Gohari, H. Pedro, J. Kleissl and C. Coimbra, "Short-term reforecasting of power output from a 48MWe solar PV plant", *Solar Energy*, Vol. 112, pp. 68–77, 2015.

[13] F. Almonacid, P. Perez-Higueras, E. Fernandez and L. Hontoria, "A methodology based on dynamic artificial neural network for short-term forecasting of the power output of a PV generator", *Energy Conversion and Management*, Vol. 85, pp. 389–398, 2014.

[14] Y. Kashyap, A. Bansal and A. K. Sao, "Solar radiation forecasting with multiple parameters neural networks", *Renewable and Sustainable Energy Reviews*, Vol 49, pp. 825-835, 2015.

[15] M. Rana, I. Koprinska and V. Agelidis, "2D-interval forecasts for solar powerproduction." *Solar Energy*, Vol. 122, pp. 191–203, 2015.

[16] A. Mellit and S. A. Kalogirou, "ANFIS-based modelling for photovoltaic power supply system: A case study", *Renewable Energy*, 36(1), 250-256, 2015.

[17] R. Chauvin, J. Nou, S. Thil and S. Grieu, "Intra-day DNI forecasting under clear sky conditions using ANFIS", Proceedings of the 19th World Congress The International Federation of Automatic Control Cape Town, South Africa, August, 2014.

[18] M. Bouzerdoum, A. Mellit and M. A. Pavan, "A hybrid model (SARIMA-SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant" *Solar Energy*, Vol. 98, pp. 226–235, 2013.

[19] A. Vaz, B. Elsinga, W. V. Sark and M. Brito, "An artificial neural network to assess the impact of neighbouring photovoltaic systems in power forecasting in Utrecht, the Netherlands", *Renewable Energy*, Vol. 85, pp. 631–641, 2016.

[20] K.-P. Lin and P.-F. Pai, "Solar power output forecasting using evolutionary seasonal decomposition least-square support vector regression", *Journal of Cleaner Production* (2015), http://dx.doi.org/10.1016/j.jclepro.2015.08.099

[21] J. L. S. García, E. E. Juárez and J. J. Flores, "Short term photovoltaic power production using a hybrid of nearest neighbor and artificial neural networks", IEEE PES Transmission & Distribution Conference and Exposition - Latin America (PES T&D-LA), Morelia, Mexico, 2016.

[22] A. Dolara, F. Grimaccia, S. Leva, M. Mussetta and E. Ogliari, "A physical hybrid artificial neural network for short term forecasting of PV plant power output", *Energies*, vol. 8, pp. 1138 – 1153, 2015.

[23] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F.J. Martinez-de-Pison and F. Antonanzas-Torres, "Review of photovoltaic power forecasting", *Solar Energy*, Vol. 136, pp. 78-111, 2016.

[24] Ye Ren, P.N. Suganthan and N.Srikanth, “Ensemble methods for wind and solar power forecasting-A state-of-the-art review”, *Renewable and Sustainable Energy Reviews*, Vol. 50, pp. 82-91, 2015.
[25] R. H. Inman, H. T.C. Pedro and C. F.M. Coimbra, “Solar forecasting methods for renewable energy integration”, *Progress in Energy and Combustion Science*, Vol. 39, pp. 535-576, 2013.
[26] I. Guyon and A. Elisseeff, “An introduction to variable and feature selection”, *Journal of Machine Learning Research*, vol. 3, pp. 1157-1182, 2003.
[27] R. A. Welikala*et al.*, “Genetic algorithm based feature selection combined with dual classification for the automated detection of proliferative diabetic retinopathy”, *Computerized Medical Imaging and Graphics*, vol. 43, pp. 64-77, 2015.
[28] A. Pereira *et al.*, “Feature selection for disruption prediction from scratch in JET by using genetic algorithms and probabilistic predictors”, *Fusion Engineering and Design*, vol. 96-97, pp. 907-911, 2015.
[29] T. Khadhraoui*et al.*, “Features selection based on modified PSO algorithm for 2D face recognition” Proceedings of the 13th International Conference Computer Graphics, Imaging and Visualization, pp. 99-104, 2016.
[30] H. Y. Markid, B. Z. Dadaneh and M. E. Moghaddam, “Sequence based feature selection using ant colony optimization”, Proceedings of the 5thInternational Conference on Computer and Knowledge Engineering, pp. 100-105, 2015.
[31] B. Oluleye*et al.*, “A genetic algorithm-based feature selection”, *International Journal of Electronics Communication and Computer Engineering*, **5**(4) pp. 899-905, 2014.
[32] J.-S. R. JANG, C.T. Sun and E. Mizutani, Neuro-fuzzy and soft computing, A computational approach to learning and machine intelligence, New Jersey: Prentice Hall, 1997, pp. 73,74, 86, 95-97,86-87, 26-28, 74-85.
[33] J.-S.R. JANG, “ANFIS: adaptive network-based fuzzy inference systems”, *IEEE Transactions on Systems, Man and Cybernetics*, **23**(3) pp. 665–685, May/June 1993.
[34] Y. Kassa, J.H. Zhang, D.H. Zheng and D. wei, “Short term wind power prediction using ANFIS”, IEEE International Conference on Power and Renewable Energy, pp. 388-393, 2016.
[35] M. K. Deshmukh and C. B. Moorthy, “Application of genetic algorithm to neural network model for estimation of wind power potential”, *Journal of Engineering, Science and Management Education*, Vol. **2,** pp. 42-48, 2010.
[36] Mitsuo Gen and Runwei Cheng, “Foundations of genetic algorithm,” *Genetic algorithms and engineering application*, New York, John Wiley and Sons Inc., 2000.
[37] Y. Kassa, J.H. Zhang, D.H. Zheng and D. wei, “A GA–BP hybrid algorithm based ANN model for wind power prediction”, Proceedings of the 4th IEEE International Conference on Smart Energy Grid Engineering, pp. 158-163, 2016.
[38] J. Kennedy, R.C. Ebenhart, “Particle swarm optimization”, Proceedings of the IEEE International Conference on neural networks, Vol. 4, pp. 1942 – 1948, 1995.
[39] J. Kennedy. The behavior of particles. In V. W. Porto, N. Saravanan, D. Waagen, and A. E. Eiben, Eds. *Evolutionary Programming VII*: Proceedings of 7th Annual Conference on Evolutionary Programming Conf., San Diego, CA, 581–589. Berlin: Springer-Verlag, 1998.
[40] C. Zhang *et al.*, “A Gaussian process regression based hybrid approach for short-term wind speed prediction”, *Energy Conversion and Management*, vol. 126, pp. 1084-1092, 2016.
[41] S.A.Aye and P.S.Heyns, “An integrated Gaussian process regression for prediction of remaining useful life of slow speed bearings based on acoustic emission”, *Mechanical Systems and Signal Processing*, vol. 84 pp. 485–498, 2017.
[42] P. Poggi*et al.*, “Forecasting and simulating wind speed in Corsica by using an autoregressive model”, *Energy Conversion and Management*; 44(20), pp. 3177–3196, 2003.
[43] J. Hu and J. Wang, “Short-term wind speed prediction using empirical wavelet transform and Gaussian process regression”, *Energy*, vol. 93, pp. 1456-1466, 2015.