Renewable Energy 140 (2019) 367–379

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

Short-term PV power forecasting using hybrid GASVM technique

William VanDeventer [a], Elmira Jamei [b], Gokul Sidarth Thirunavukkarasu [c], Mehdi Seyedmahmoudian [c, *], Tey Kok Soon [d], Ben Horan [a], Saad Mekhilef [c, e], Alex Stojcevski [c]

[a] School of Engineering, Deakin University, VIC 3216, Victoria, Australia
[b] College of Engineering and Science, Victoria University, VIC 3011, Victoria, Australia
[c] School of Software and Electrical Engineering, Swinburne University of Technology, VIC 3122, Victoria, Australia
[d] Department of Computer System and Technology, Faculty of Computer Science and Information Technology, University of Malaya, 50603 Kuala Lumpur, Malaysia
[e] Department of Electrical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia

ARTICLE INFO

Article history:
Received 16 June 2018
Received in revised form
6 February 2019
Accepted 16 February 2019
Available online 20 February 2019

Keywords:
Genetic algorithm (GA)
Genetic algorithm based support vector machine (GASVM)
Photovoltaic (PV)
Short-term forecasting
Support vector machine (SVM)

ABSTRACT

The static, clean and movement free characteristics of solar energy along with its contribution towards global warming mitigation, enhanced stability and increased efficiency advocates solar power systems as one of the most feasible energy generation resources. Considering the influence of stochastic weather conditions over the output power of photovoltaic (PV) systems, the necessity of a sophisticated forecasting model is increased rapidly. A genetic algorithm-based support vector machine (GASVM) model for short-term power forecasting of residential scale PV system is proposed in this manuscript. The GASVM model classifies the historical weather data using an SVM classifier initially and later it is optimized by the genetic algorithm using an ensemble technique. In this research, a local weather station was installed along with the PV system at Deakin University for accurately monitoring the immediate surrounding environment avoiding the inaccuracy caused by the remote collection of weather parameters (Bureau of Meteorology). The forecasting accuracy of the proposed GASVM model is evaluated based on the root mean square error (RMSE) and mean absolute percentage error (MAPE). Experimental results demonstrated that the proposed GASVM model outperforms the conventional SVM model by the difference of about 669.624 W in the RMSE value and 98.7648% of the MAPE error.

1. Introduction

The rapidly advancing evolution of technology and the increase in population density creates an increasing demand for energy resources, both renewable and non-renewable. The use of non-renewable energy sources, such as coal or other fossil fuels, not only cause a significant amount of pollution but also ultimately result in the depletion of the limited resources. Therefore, the use of renewable energy (RE) sources is further encouraged across the world. Among diverse RE sources, photovoltaic (PV) systems are considered as one of the fastest growing technologies that address the energy demand subsequently [1]. Studies showed that there is a 3% increase in the level of carbon emissions within the last decade [2,3] when observed under extreme weather conditions, such as heat waves, droughts, thunderstorms, and flooding [2]. Global warming phenomenon and the increasing rate of emissions from the electrical and heating sectors directly affect the environment and advocates the widespread use of RE in the energy generation sector.

In a few south Asian countries like India, the government has encouraged the use of the RE sector by setting the deadline for 80% emission-free target for the year 2030 [4]. These policies have resulted in the exponential growth of PV installations in the residential-, commercial-, and utility-scale sectors [5]. Despite the large investments and technological advancements administered by the measure taken by the governments across the world the cost

* Corresponding author.
E-mail addresses: william.vandeventer@crossmuller.com.au (W. VanDeventer), elmira.jamei@vu.edu.au (E. Jamei), gthirunavukkarasu@swin.edu.au (G.S. Thirunavukkarasu), mseyedmahmoudian@swin.edu.au (M. Seyedmahmoudian), koksoon@um.edu.my (T.K. Soon), ben.horan@deakin.edu.au (B. Horan), saad@um.edu.my (S. Mekhilef), astojcevski@swin.edu.au (A. Stojcevski).

https://doi.org/10.1016/j.renene.2019.02.087
0960-1481/

of the PV system is comparatively expensive [6] emphasizing the inverse relationship between the cost of the PV systems and the increase in demand. Thus, to reduce costs, the demand should significantly increase. Therefore, companies with government support provide neutral cost solutions that are tailored to the household income and circumstances [7]. A competitive marketing campaign within the National Energy Market encourages independent power-producing companies to use mathematical tools, such as forecasting strategies and meta-heuristic energy management methods [8].

Sophisticated stochastic systems are modelled considering the influence of external environmental factors affecting the forecasting of PV systems sensitivity [9]. The two primary factors that influence the PV forecasting include solar irradiance and air temperature. In addition, the climate pattern also plays a major role in PV forecasting. The climate pattern varies depending on the geographical location, thereby directly influencing the forecasting of the PV systems. Therefore, customized PV models for new and desired locations are essential for an optimum outcome.

A few different approaches are used to track down the environmental parameters such as solar radiation, air temperature, etc., which is then utilized by the forecasting model of the PV system. Remote sensing is the most commonly used method for trapping environmental data [10]. In addition to this, few hybrid machine learning technique were used for prediction of weather parameters like solar radiation based on historical meteorological data monitored involving techniques like linear auto-regression [11], support vector machine-firefly model [12], etc.,

In this paper, the proposed method for the forecasting of PV Power uses a hybrid ensemble method that utilizes both GA and SVM to form a hybrid GASVM model. The proposed model has been developed to address the forecasting of the rapid and multifaceted changing weather patterns. The use of real-world environmental data from a city like Geelong with diversified weather patterns helps in evaluating and improving the accuracy of modelling the GASVM ensemble technique. Real-world experimental analysis greatly increases the accuracy of the results by removing unforeseen influences that may be present otherwise. In addition to this, the inclusion of the PV power data gathered allows us to validate the results under real-world conditions.

The rest of the paper is organized as follows. Section 2 presents a review on the existing works related to the objectives of the manuscript, followed by Section 3 which gives a brief overview of the support vector machines model used for classification. Section 4 describes the theory of a genetic algorithm. Section 5 presents the methodology of the proposed GASVM technique. Followed by which, Section 6 deliberately explains the results of the experimental analysis done on evaluating the proposed system, and finally, Section 7 elaborates on the advantage of using the proposed GASVM technique for forecasting of PV power output.

2. Related works

Observations from the extensive literature review indicated that a significant number of models used in forecasting of PV solar output power lagged in terms of the quality of environmental data used in the process of predicting the PV output power. It was evident that the two common classifiers used for PV forecasting included neural network (NN) and support vector machine (SVM). In the absence of the preliminary assessment made on the quality of the data used to train the system, the performance indicators of the classifier used in the forecasting region cannot be easily justified. The following section critically evaluates the performance of the various other techniques used for forecasting and explains why the proposed GASVM approach was considered.

Many hybrid algorithms integrating the predominant classifiers with other evolution algorithms were proposed for optimizing the performance of the classifiers. Both NN and SVM classifier have their own advantages and disadvantages which makes them distinctly different from each other. Nevertheless, despite their differences, they are both considered strong classifiers, which are often paired up with weaker classifiers like k-nearest neighbor (k-NN), decision tree, and nave bayes (NB) to increase the accuracy of the classification.

In addition, many other approaches can be used to strengthen the performance of the classifier. A potentially stronger classifier is obtained using an ensemble of classifiers. Multiple classifiers are used to combine their classification attributes to obtain a complementary pair of classifiers. Some examples of popular ensemble methods include bagging and AdaBoosting [13]. These techniques are generally used with classifiers, such as discriminant or decision tree classifiers.

Optimization techniques like Bayes optimization [14] and genetic algorithms (GA's) [15] were used to increase the accuracy of the classifiers/ensembles used for forecasting tasks. Both methods follow a similar principle, in which they iteratively test each parameter associated with the classifiers until the fitness value converges into a single optimized classification model. However, GA is prone to issues associated with local minima and therefore produce different results under multiple iterations.

The persistence model commonly known as the nave predictor model is widely considered as the benchmark for other considerably complex forecasting models [16]. The main shortcoming of the model is that the accuracy of the model decreases strongly with an increase in cloud cover and the forecasting tasks which is carried out for a longer duration greater than one hour [17] which rules out its preference to be used as a forecasting model. However, it provides a benchmark that can be used to standardize forecasting performance with the other methods.

In general, forecasting methods which predict the PV output one hour in advance with minimal error are achieved by learning from 20 years of historical environmental data as highlighted by Gandoman in 2016 [18]. The most significant parameter for PV output is solar irradiance; thus, the Oktas scale is used to predict cloud cover to forecast solar irradiance. The Oktas scale can be combined with fuzzy logic (FL) to allow inputs, which are bound to uncertainty, to be classified into the predefined framework of the Oktas scale. The importance of considering the cloud cover effect and the forecasted climate type is emphasized in this study [18].

Another common method used is Autoregressive moving average (ARMA) which is a linear method for statistical modelling to correlate time series data. This method can incorporate different types of time series data and extract their statistical properties. However, the time series data has to be stationary. The ARMA method was improved with the introduction of the autoregressive integrated moving average (ARIMA) method. In this method, an extension of ARMA with an added capability to handle non-stationary time series data and capture sharp transitions in irradiance with greater accuracy was proposed [17,19]. The capability to capture irregular data is favourable when the climate pattern highly fluctuates, such as in Melbourne and Geelong. The drawback of ARIMA, contrary to ARMA, is that it is more computationally intensive due to the inclusion of a summation/integration function in the method. In addition, ARIMA does not outperform the SVM models with regard to time series data, although both models are extremely powerful and popular in various fields [20].

The other method widely presented in the literature is the Artificial neural network (ANN) which was developed to shift from linear algorithm to nonlinear solution. This method is extremely popular due to its capability to learn on its own from the training

data set, thereby obtaining greater accuracy. ANN has self-adaptation, fault tolerance, robustness, and strong inference capability [17]. ANN-based models outperform conventional mathematical models in terms of accuracy and adaptability [8]. ANN-based forecasting model that forecasts 48 h in advance; under a clear sky model was proposed in this research which highlighted the use of training phase to adjust weights [21]. The study showed that clear sky index is random in nature and therefore adds no significant value to the training of ANN [17]. The drawback of ANN is that it greatly increases the complexity of the system due to the nature of using multi-layered network architecture. In addition, NNs require a random initial data set that may reduce the reliability of the output.

An integrated model of ANN creates the wavelet neural network (WNN) which let the time-frequency localization and focal features that help in more accurate forecasting. The extension of WNN which integrated the recurrent neural networks into the WNN model to form a diagonal recurrent wavelet neural network (DRWNN) [17]. Therefore, fuzzification of parameters can be used with DRWNN, which lead to a better back-propagation model (BPN) compared to the other BPN models [22]. Similar approaches combining WNN with radial basis function neural network (RBFNN) aiming to capture the relationship among PV, solar radiation, and temperature to forecast PV power output are also highlighted in the literature [23]. RBFNN is a two-layered network that optimizes the weights found in the second layer. The use of the wavelet method increases the computational intensity and the accuracy of prediction models. However, this increase in accuracy is still rivalled by methods, such as SVM, which outperforms the wavelet method in relation to the computational intensity involved [24]. Decisively, wavelet transform method can be used as one of the kernel computing an SVM model; which combines the advantages of both the SVM and wavelet transform methods to classify between data sets.

Few other methods highlighted in the literature include the use of national weather prediction (NWP) models to utilize the non-exogenous input data to predict the weather up to 15 days ahead [23]. NWPs are used by companies, such as global forecast System and the European centre for medium-range weather forecasts [17]. The NWP models perform better when predicting in a lower scale, say five hours ahead due to the inclusion of the cloud motion vectors, cloud motions, formations, and dissolutions [25].

Inference rule-based fuzzy logic (FL) models are one of the other commonly used predictive approaches. FL provides a robust and advantageous method of modelling parameters, and it has been used for medium-term models because it handles the uncertainty of real-world data. Fuzzy rules are linguistic-based and allow a statistical approach in the process of decision making. However, the other models use crisp values, which are detracted from the capability to portray a real-world representation [26]. In one of the other study, the FL method was combined with NN to produce an adaptive neuro-fuzzy inference method [25]. The FL is combined with the self-adaptive training of NN, resulting in a 95% improvement in the accuracy level, compared to the standard FL method. However, the FL method requires long execution time, in contrast to SVM models. Also, FL based forecasting models have numerous test nodes that limit its performance of the system [27].

SVM which is a simple, powerful, highly flexible, nonlinear and computationally less expensive model is mainly used in wind and PV output power forecasting systems. SVMs can learn without being heavily dependent on prior knowledge, contrary to ANN [28]. In contrast with NNs, SVMs can overcome the issues regarding the local minima. SVM simplifies the complex mathematical problems of PV forecasting [29] and is also used in conjunction with methods such as wavelet, radial basis function and firefly algorithm for more optimized accurate results. However, SVM is highly sensitive to the parameters, which are used to classify the data. Therefore, in many cases, optimization techniques are used to find the optimal balance between the different parameters [27].

SVM method aims to solve the optimal hyper-plane that differentiates between classification areas. It also allows for the greatest distance between points, which increases the reliability of the classifier. This mathematical problem is solved using Equations (1)–(3), where α represents the weighting factors in the definition of the final hyper-plane used [30]:

$$f(\alpha) = \frac{1}{2}\alpha^T Q\alpha - e^T\alpha, \tag{1}$$

$$0 \le \alpha_i \le C, i = 1 \to N, \tag{2}$$

$$y^T\alpha = 0, \tag{3}$$

GA is a type of evolution algorithm that can be used to optimize weights and other features within a base classifier. This algorithm is used to optimize NNs, such as convolution neural networks, deep neural networks, and SVM [27,31]. GA mimics the use of chromosomes to develop new DNA strings that combine elements from two different chromosome strands [32]. GA can be highly versatile in their application given that the chromosomes reflect a data array, which can be used to represent as many attributes/weights as required by the base classifier. GA is susceptible to local minima and therefore may need to be tested iteratively for better and reliable results. Thus, in the proposed GSVM model the GA optimizer is programmed to run on the ensemble classifier multiple times in parallel and arithmetic mean is taken to determine the weights of the classifier.

3. Support vector machines

SVM is based on the principle of creating a hyperplane that has the greatest margin of separation between the two types of classifications being made over the dataset. This margin of separation is governed by the support vectors that are used to classify the data. A pictorial representation of the SVM model is clearly explained in Fig. 1 which helps the reader to get a clear idea of the hyper-planes in the search space. The primary relationship governing the mathematical relationship of the hyperplane is expressed in Equation (4).

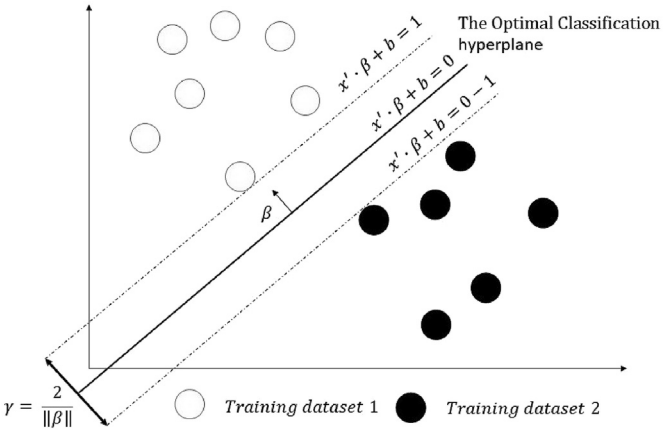


Fig. 1. The pictorial representation of the SVM model.

$$f(x) = x\cdot\beta + b = 0,\\ \beta \in R^d\left(d = n^{th}Dimensions\right), b \in R, \tag{4}$$

The margin of the separation can be written as a function of the difference between the positive and negative support vectors which is expressed in Equation (5).

$$margin = \gamma = (x'_+ - x'_-)\cdot\frac{\beta}{\|\beta\|} \tag{5}$$

β is a non-linear function which maps the input space into a higher dimensional space. We obtain the largest margin of separation between the planes using the structural risk minimization principle, this risk bound is minimized by the use of the optimization techniques resulting in Equation (6).

$$margin = \gamma = \frac{2}{\|\beta\|} \tag{6}$$

The Lagrange multipliers are used to rewrite the equations inclusive of the constraints that define the property of the hyperplane which classifies the data which is expressed in Equation (7). In order to find the local minimum with respect to β, differentiation is applied to the Lagrangian multiplier expressed in Equation (7).

$$L = \frac{1}{2}(\|\beta\|)^2 - \sum_j \alpha_j\left(y_j(x'_j\beta + b) - 1\right) \tag{7}$$

The box constraints are used to limit for α_j that changes in order to increase the boundary size. This is seen below in Equation (8).

$$\frac{\partial L}{\partial \beta} = \sum_j \alpha_j y_j = 0, C \geq \alpha_j \geq 0 \tag{8}$$

The Gaussian kernel that is used in the proposed GASVM is defined as expressed in Equation (9).

$$K_G(x_j, x_i) = e^{-x_j - x_i{}^2} \tag{9}$$

Using Equation (9), the final equation for obtaining the optimal hyperplane of the Gaussian kernel is derived as expressed in Equation (10).

$$\sum_j \alpha_j y_j . e^{-||x_j - x_i||^2} + b \geq 0, then +'ve\ sample \tag{10}$$

To adapt the standard SVM handling non-separable data, constants are injected into the formula to construct a soft margin. This relates to the use of a constant such as the box constraint (C) as expressed in Equation (11).

$$\min_{\beta,b,\xi}\left(\frac{1}{2}||\beta||^2 + C\sum_j \xi_j\right) \tag{11}$$

4. Genetic algorithm

Generic Algorithm is based on the theory of evolution which considers the fitness level of elements with regards to survival of those elements into the next generation. This algorithm includes three main processes such as generation, mutation and crossover. The generic overview of the genetic algorithm is expressed in Fig. 2 shown below.

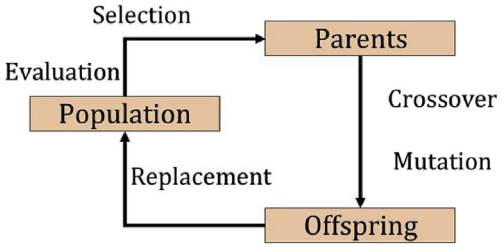


Fig. 2. The generic overview of the genetic algorithm.

4.1. Generation

Each generation consists of a population size that represents a configuration of a parameter set which determines the architecture or weight matrix of the proposed system. At each new Generation, an elite group of previous generation candidate (parents) are chosen into the next generation. The mathematical representation of this stage is indicated in Equation (12).

$$G_i(p\{x_0, \ldots, x_b\}) = \begin{Bmatrix} p_0 \\ \vdots \\ p_n \end{Bmatrix} = \begin{Bmatrix} x_0 & \cdots & x_{0b} \\ \vdots & \ddots & \vdots \\ x_{n0} & \cdots & x_{nb} \end{Bmatrix} \tag{12}$$

where *i* represents the number of iterations, *p* representing the person or candidate in that current generation, *n* representing the total population of the current generation and the *b* indicating the parameter size of the person.

4.2. Mutation

The mutation occurs in each new generation and creates genetic diversity within the children of each parent. Each of the children is mutated based on a Gaussian distribution with a predefined to the mean of zero. The standard deviation, however, is determined by the shrink factor and the number of generations. As the generations increase the deviations within the children decrease causing the algorithm to comprehend to a single parent configuration. This process is repeated for each parameter within each child, and the mathematical representation is expressed in Equation (13). Fig. 3 indicates the visual representation of the mutation operation in GA.

$$\sigma_{bk} = \sigma_{bk-1}\left(1 - Shrink.\frac{k}{Generations}\right) \tag{13}$$

4.3. Crossover

The crossover causes a swap of parameter values to construct a hybrid child of the parents. This is achieved by choosing random

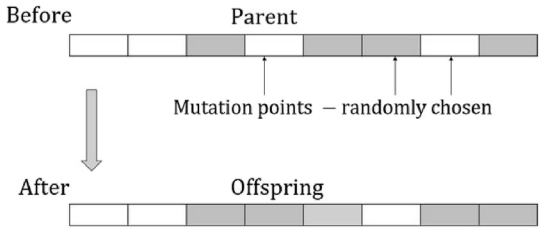


Fig. 3. Mutation operation of the GA.

parents and generating a random series of binary points that equal the number of parameters within each parent. This binary matrix is then used to swap all parameters that correlate to the binary matrix as expressed in Equation (14). Fig. 4 indicates the visual representation of the crossover operation in GA.

$$\begin{aligned} parent_a &= \{x_{a1}, x_{a2}, x_{a3}, x_{a4}\} \\ parent_b &= \{x_{b1}, x_{b2}, x_{b3}, x_{b4}\} \\ Bin_{rnd} &= [0\ 1\ 0\ 1] \\ child_{ab} &= \{x_{a1}, x_{b2}, x_{a3}, x_{b4}\} \\ child_{ba} &= \{x_{b1}, x_{a2}, x_{b3}, x_{a4}\} \end{aligned} \tag{14}$$

The classifier is trained using the training data set, and then the validation set is used to predict and formulate an accuracy for that given GA configuration. The fitness value of the generation after the evaluation of the validation results is expressed in Equation (15). The number of parents need not to equate to the population size since there is the possibility that the children will be derived from the parents to make up the difference.

$$f_i = MAPE_i = \frac{100}{N}\sum_{j=1}^{N}\frac{|A_j - F_j|}{A_j} \tag{15}$$

Thus, the general overview of the SVM model used for classification and the GA algorithm used for optimization in the proposed concept is explained. The following section explains the operational flow of the proposed methodology for performing short-term forecasting of PV output power using the GASVM technique.

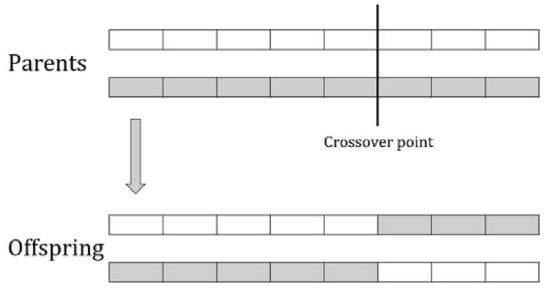


Fig. 4. Crossover operation of the GA.

5. Methodology

Fig. 5 depicts the operational flow of the proposed methodology for performing short-term forecasting of PV output power using the GASVM technique. The method involves the process of data collection, data preprocessing and optimization of the SVM classifier using GA that will lead to the formation of the proposed GASVM.

One of the most important drawbacks of many forecasting algorithms proposed in the literature is the lack of reliable and trustable data collected on-site for training the forecasting model. The accurate recording of meteorological data from the location close to the PV system helps in enhancing the performance of the forecasting algorithm. For this reason, the data collection was conducted using a local weather station installed close to the PV system situated on the rooftop of the School of Engineering, Deakin University, Victoria, Australia. Fig. 6 indicates the setup in the rooftop of Deakin University. The recorded climatic data include air temperature and solar irradiance.

The data collection included the hourly samples obtained from the Envirodata WeatherMaster 2000 weather station and the output power obtained from the PV system. The PV system has about 12 PV modules, and the peak power is near to 3 kW, and the individual specification of the PV modules is expressed in Table 1.

In addition to the climatic data, the PV output power was also measured to train/verify the results obtained from the proposed GASVM technique based on the real-world data. The data used in developing the forecasting model consist of three parameters. These parameters include the PV power, solar irradiance, and air temperature. The data has been collected over a span of 278 days is used for training the proposed GASVM model. The raw data recorded is represented in a scatter plot representing all the three parameters in Fig. 7a. Fig. 7b and c represents the relationship of the air temperature (°C), and solar radiance (W/m^2) with the

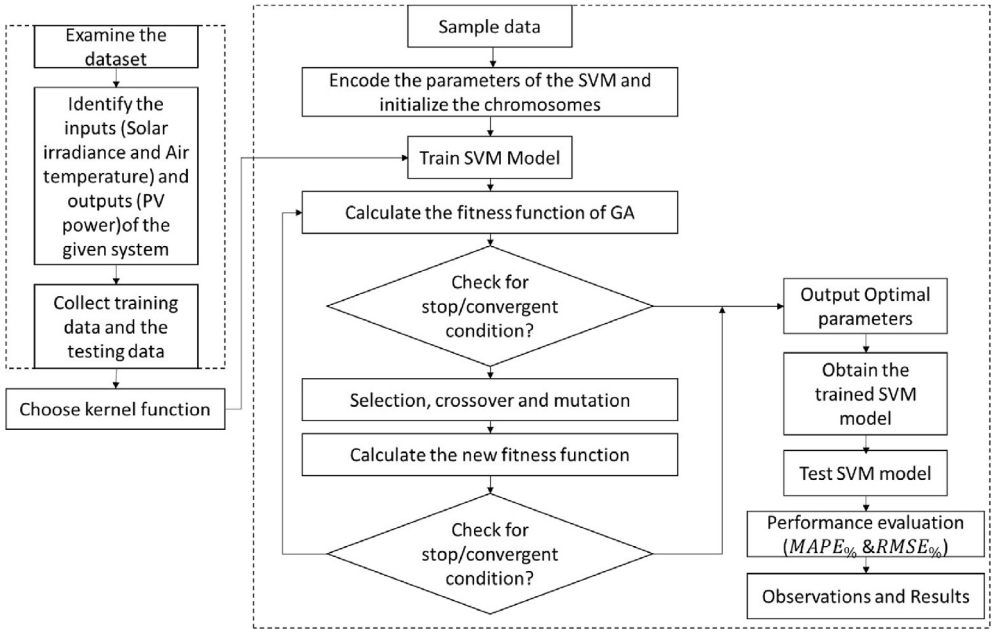


Fig. 5. The operational flow of the proposed GASVM method.

Fig. 6. Images of weather station and PV system located at Deakin University, Australia.

Table 1
PV specifications.

Parameters	Values
Solar module type	CS6P-250M
Number of modules	12
Module rated power output	250 W
Open circuit voltage (V_{oc})	37.5 V
Short circuit current	8.74 A
Optimum operating voltage (V_{mp})	30.4 V
Optimum operating current (I_{mp})	8.22 A

Note: All parameters are based on the standard testing condition (STC) in which the ambient temperature is 25°C and irradiance level is 1000 W/m^2.

output power generated.

The raw data obtained from Fig. 7 was used to develop the density plot which is shown in Fig. 8. The density plot shows that a relatively high density of data is collected for low solar irradiance and an air temperature of 14.93 °C. On the contrary, a high crossover between classification outcomes is not specific to this region. These high classification crossover regions exist in the areas of low data collection, causing exceeding difficulty in these areas to forecast PV output.

The discrepancy found between the two outputs and their variations is indicated in Fig. 9. This discrepancy can be explained by evaluating the variation in the raw data of the PV power output. The plot in Fig. 9 highlights the nonlinearity of the data variation in terms of PV power. The air temperature has a lower variation than solar irradiance. However, the low variation does not necessarily result in a better classifier given that the air temperature is significantly lagging the solar irradiance. Furthermore, given that solar irradiance is a critical factor to classifiers, this variant can greatly affect the classification of PV power.

The dataset gathered has been divided into three subsets of the initial raw data to obtain acceptable and accurate prediction data which was then preprocessed. The first subset includes the training data that are used to train the classifier. This dataset is allocated to 80% of the total data. The second subset is the validation set, which is used in conjunction with the training data to optimize and configure the classifier which used another 10% of the total data. Finally, the last subset of data is used for testing the end result of the training data. This dataset is assigned to the remaining 10% of the total data. The reason for the segmentation of the initial dataset is to obtain an accurate assessment of the classifier without risk of using test data in the training process.

The training data used in the design of the classifiers are plotted over a seven-day period as shown in Fig. 10. Fig. 11 is a normalized plot where the percentage of a max of the recorded parameters illustrates that the solar radiance is a primary indicator for predicting the PV output power. These results show that the training data are not consistent in shape or intensity when compared, because of the unknown environmental factors that skewed the results and diminished the overall PV power. However, these factors are not considered destructive to the classifier given that the classifier does not regard the prior PV outcome. Instead, it classifies based solely on the two factors of air temperature and solar irradiance.

In order to increase the amount of data used by the classifier, other aspects of the initial data have been exposed to the data preprocessing. The main features extracted from the dataset included a differential and a moving average feature representing the pattern of the data captured. This data, however, was carefully shifted whenever it was out of phase to not expose the outfalls of the prediction that is required by the classifier.

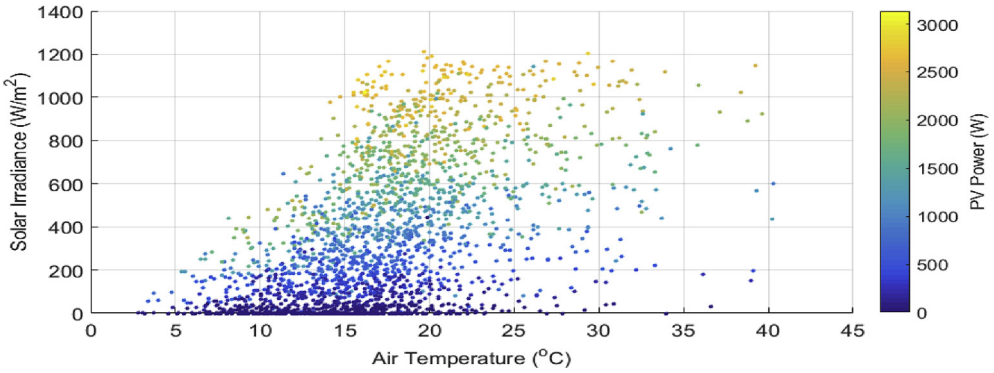
Further data manipulation was conducted by removing the period in between, where the PV power and the solar irradiance are equal to zero and low PV values (below 20 W). This measure was required to remove bias from the evaluation of accuracy using MAPE and provide greater focus on the primary concern. The systems hourly weather data and power output during this period were recorded using the self-energized data logger system that added the above-mentioned bias in the system.

5.1. Proposed GASVM into PV solar forecasting

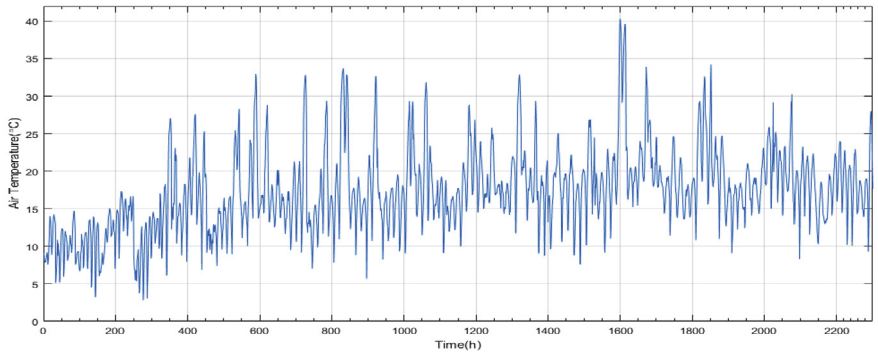
The proposed GASVM consists of two core techniques namely, classification and optimization. The classification technique uses SVM whereas the optimization technique is carried out using GA. The GA is used to optimize the SVM core architecture, therefore, defining the scaffolding of the SVM classifier. GA is then used again in defining the weight/cost matrix which enables the optimization of the support vectors to get the best fit of the validation data that is used in the process.

5.1.1. Parameter optimization of GASVM

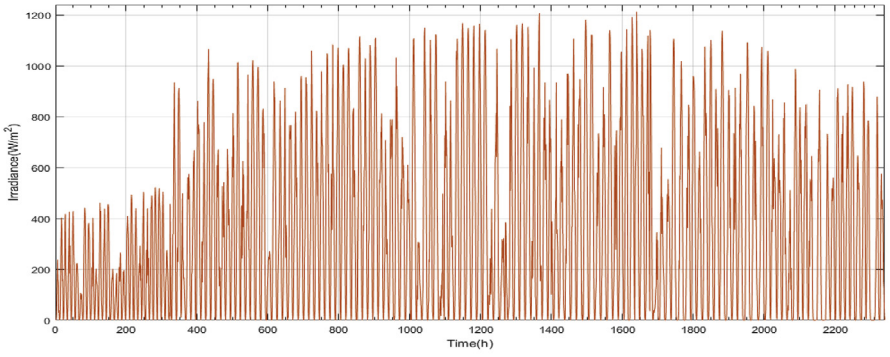
The parameter optimization was conducted on the SVM classifier that included the kernel function, kernel scale and box



(a) Solar irradiance vs air temperature vs PV power



(b) Raw air temperature



(c) Raw solar irradiance

Fig. 7. Graphical representation of raw data.

constraint. The optimization of the parameters was carried out to reduce the MAPE error of the classifier. This optimization is used to find the best structure for the classifier by varying the parameters, that dictate how the internal classification method will learn and predict the test data. Fig. 5 highlights this process and shows how GA's are used to optimize the parameters iteratively until the error/fitness value is at its lowest. This method of optimization of the internal parameters provides variations within the classifier responses with regards to the training process. The parameters have been tabulated in Table 2.

After the optimization of the classifier architecture, the weight matrix needs to be optimized next. GASVM has a weight matrix that determines the amount of bias within the model that exists for each PV power region separately. This is achieved by GASVM through its one vs one model that allows for an ensemble of SVM classifiers that exists as a supervisory model. The ensemble creates a separate classifier for each PV Range and then polls all the classifiers at once with regards to their respective designated PV Ranges.

A single supervisory SVM classifier learns this behaviour amongst the other ensemble of classifiers and then reduces the polling to a single predicted result which is the aggregated mean of the outputs of the individual classifiers as illustrated in the Fig. 13. The weight matrix allows for a bias term to be assigned to each classifier giving it precedence over other classifiers. The optimization of the weight matrix is important in reducing error because it allows the increase of attention to certain points with greater impact on the accuracy of the classifier. Fig. 12 indicates the architecture of the optimization method and Fig. 13 depicts the stage

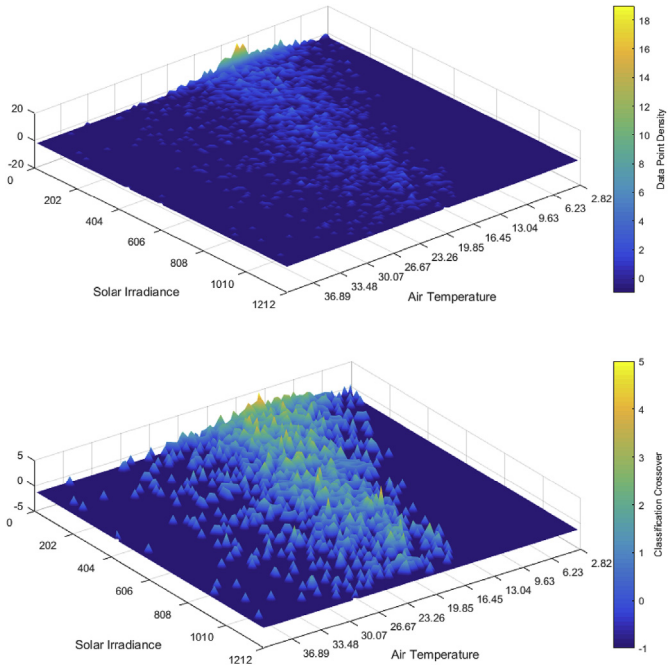


Fig. 8. 3D Density Plot (top) 3D Classification crossover per grid region (bottom).

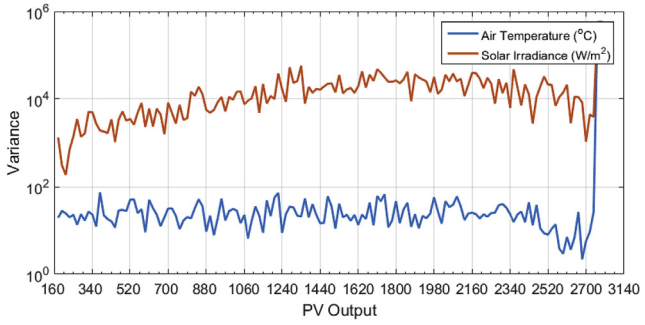


Fig. 9. Variance plot of measured data.

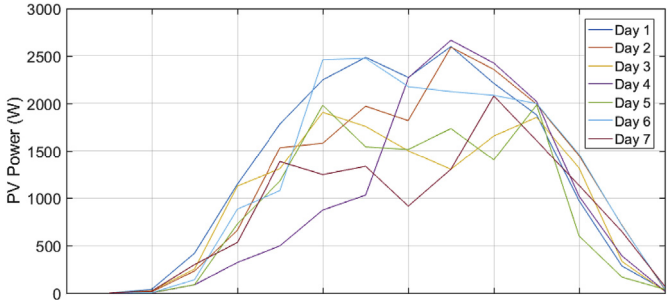


Fig. 10. Training data over seven-day period.

2 of the optimization method illustrating the supervisory SVM classifier.

5.2. Performance evaluation methods

In this section, the strength of the optimization techniques employed throughout the process is explained. The measure used in evaluating the classification model includes the root mean

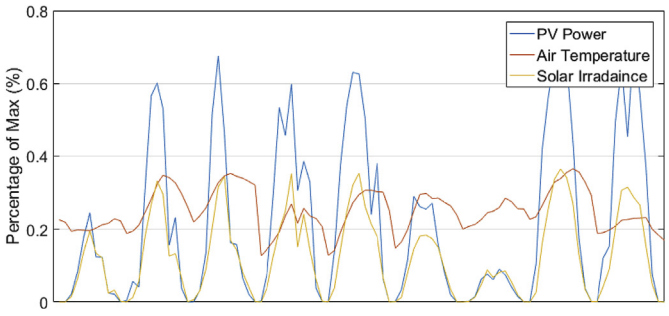


Fig. 11. Representation of the percentage of max.

Table 2
Optimized parameters.

Parameters	Range of values
Kernel scale	$\infty > Kernel\ scale > 0$
Box constraint	$\infty > Box\ constraint > 0$
Solver	• Sequential Minimal Optimization (SMO) • L1 Soft Margin optimization • Iterative Single Data Algorithm
Singular - supervisory	• One vs All • One vs One
Kernel functions	• Linear:$K(x_j, x_i) = x'_j x_i$ • Polynomial:$K(x_j, x_i) = (1 + x'_j x_i)^n$ • Gaussian:$K(x_j, x_i) = e^{\|x_j - x_i\|_2^2}$

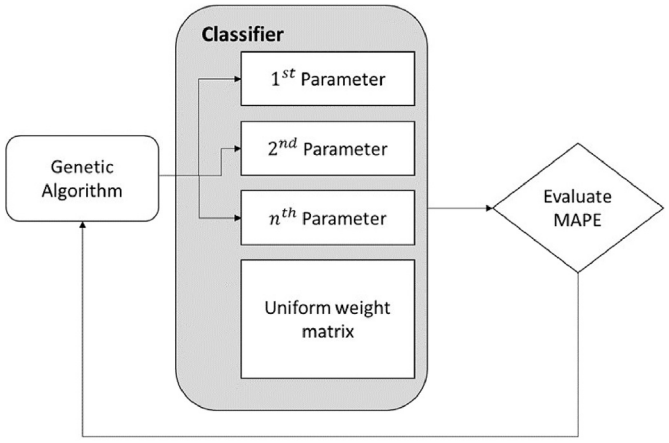


Fig. 12. Architectural optimization methodology.

square error (RMSE) and mean absolute percentage error (MAPE). These performance metrics are selected to be evaluated because of the wide use in the existing literature and are also considered as the primary measures to establish the strength of the classifiers.

5.2.1. Mean absolute percentage error

The MAPE approach is common in the industry and involves considering the mean of all accuracies on a per point basis to evaluate the total mean accuracy. While mathematically simple, this approach has its limitations. The MAPE approach fails in its evaluation when the actual value is equal to zero. Despite the significance of this limitation, it is often eliminated due to the filtering of the results that tend to approach zero. The effect of this limitation is proportional to the percentage of actual values of the data set being zero. The mathematical presentation is as shown in Equation (16)

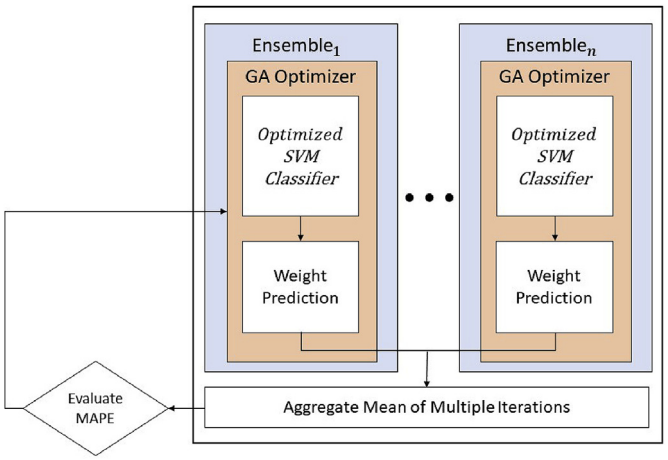


Fig. 13. Supervisory SVM classifier.

$$MAPE_{\%} = \frac{100}{N}\sum_{i=1}^{N}\frac{|A_i - F_i|}{A_i} \tag{16}$$

where A_i represents the actual PV power, $(A_i \neq 0)$, F_i represents the forecasted PV power and N indicates the number of values. Correspondingly accuracy of the proposed approach depends on the MAPE error according to the following relationship. Correspondingly the accuracy is given by the following Equation (17).

$$Accuracy = 100 - MAPE_{\%} \tag{17}$$

5.2.2. Root mean square error

RMSE is another common approach used in the evaluation of the accuracy of the classifier. RMSE provides a means of assessing mean error with regards to PV power. This indication is important since methods such as MAPE does capture the accuracy in terms of the output power value which is the main objective of the optimization and therefore can present good accuracy with high losses in the system. Correspondingly the RMSE of the system is given by the following Equation (18).

$$RMSE_{\%} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(A_t - F_t)^2} \tag{18}$$

6. Results and discussion

6.1. Numerical results

The performance of the proposed method is verified by comparing the outputs with the original SVM and the proposed GASVM. The results obtained from the performance evaluation allow a performance measure to be established by comparing SVM without optimization and the proposed method which includes optimization. The final SVM architecture obtained from the supervisory SVM classifier in the GASVM is shown in Table 3. The Gaussian kernel is the most accurate and preferred over the polynomial-based kernels, which ranged from linear to quadratic kernel functions. Also, the Gaussian kernel is generally more accurate in terms of the data used. It again outperformed the other kernels. The solver used was the iterative single data algorithm (ISDA). The selected kernel scale obtains a finer gram matrix, which essentially controls the scale of the space where the data resides. This scale factor allows the space to be expanded or shrunk to construct a hyperplane with freedom between points. The box constraint controls the maximum penalty incurred when misclassified points exist past their respective boundaries.

Table 3
Optimized SVM architecture.

Parameters	Values
Kernel	Gaussian
Solver	Iterative single data algorithm
Kernel scale	21
Box constraint	155

GASVM are based on a predefined envelope of modifications within the internal classifier of the supervisory SVM classifier has been tabulated in Table 3. The GA systematically proceeds through the parameters and attempts to change individual clusters instead of changing all at once. Therefore, by this process, the target parameter is determined. The decrease in error is evident, as shown in the results. Illustrations of the response when the GA determines a target area clearly stand by this observation. The results also show that, at this early point in the process, the accuracy is relatively poor and then rapidly improves until it converges to the local minima. Table 4 indicated the performance comparison of the SVM and proposed GASVM model.

The predictions from GASVM after optimization through Step 1 are shown in Fig. 14. These results show that the evaluation of RMSE and MAPE differ with a performance at different regions of PV Power. The RMSE is worst in the higher region of PV Power and therefore more greatly affects the net power prediction. Whereas, MAPE is worst in the lower regions and causes the accuracy to drop more significantly in those regions. In addition to this, Fig. 15 shows that most of the peaks of error are due to the region of PV Power that can be approximated to 0 Watts. Therefore, to illuminate these spikes, a filter was used at the 20 Watts mark that had an insignificant effect on the net PV power whilst drastically increase the accuracy of the results.

Fig. 16 shows the comparison of the SVM and GASVM results which clearly indicate that the GASVM outputs are the closest to the actual outputs of the PV system indicating the accuracy of the proposed forecasting method which predicts the output with a forecasting horizon on an hour. Furthermore, the results of the architecture of each classifier are outlined individually for each algorithm. These results are based on the optimized classification for a given uniform weight, which is optimized through GA.

6.2. Weight matrix results

The SVM weight matrix was generated through GA that optimized it to fit the testing and validation data. Prior to these results, the weight matrix was uniform. The GA Optimized weight matrix results can be seen in Fig. 17. Each weight matrix point was tested to remove points that had no bearing on the accuracy of the classifier. These results highlight the rigidity of the GASVM weight matrix. In addition to this, the results from GASVM at this point is quite

Table 4
Performance after architectural optimization.

	SVM	GASVM
RMSE (W)	680.85	11.226
MAPE (%)	100.47	1.7052

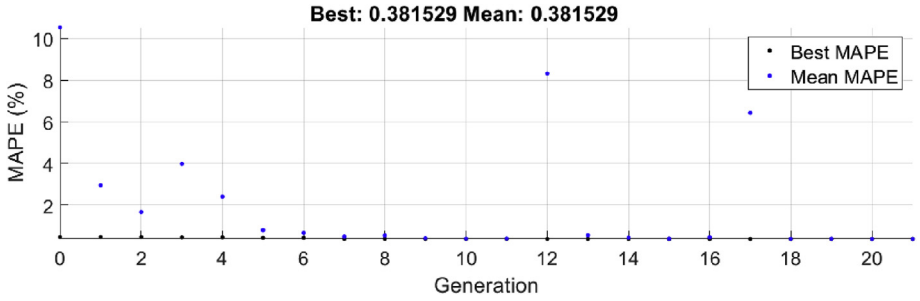


Fig. 14. GASVM genetic algorithm results.

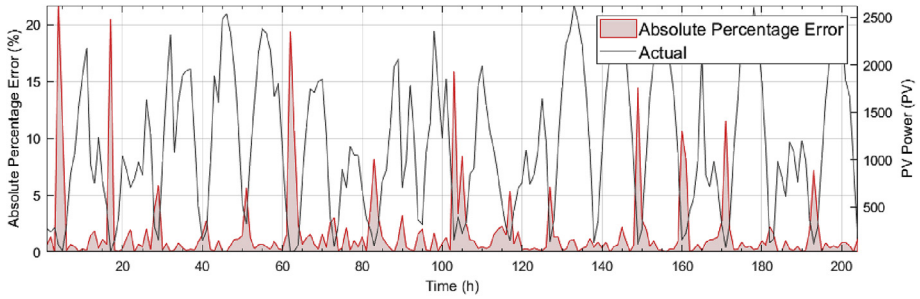


Fig. 15. GASVM prediction error (%) results.

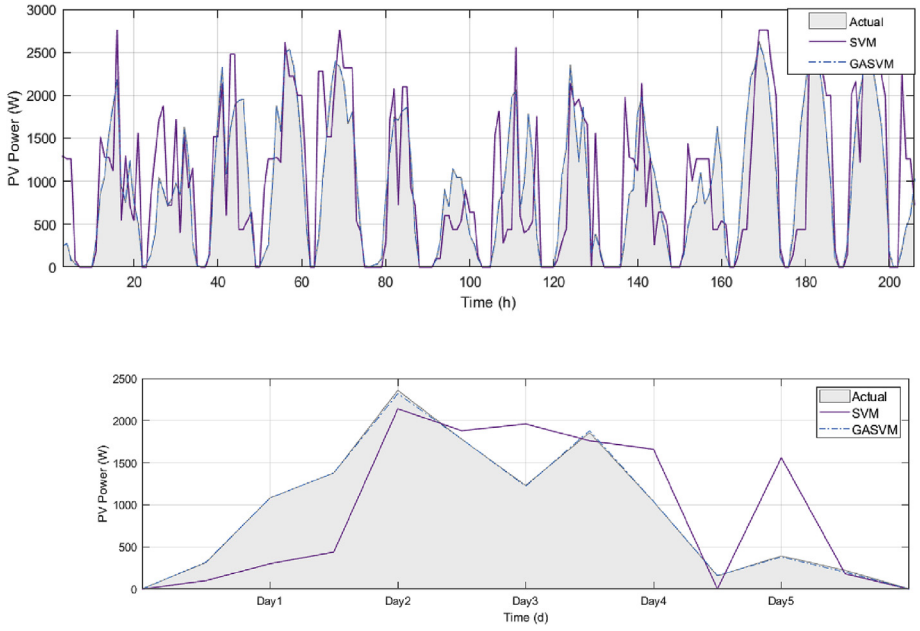
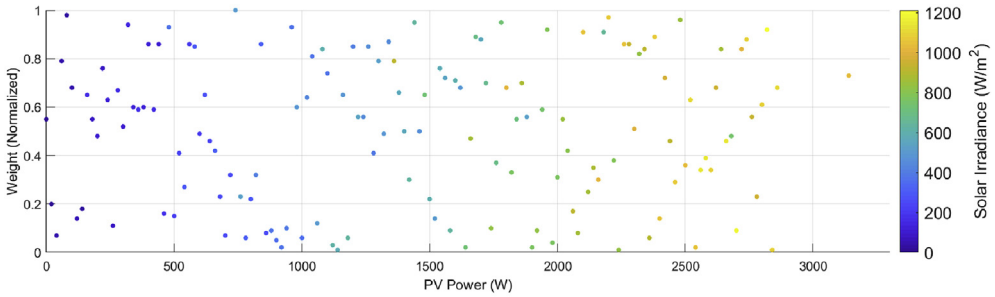


Fig. 16. Comparison of SVM and GASVM output results.

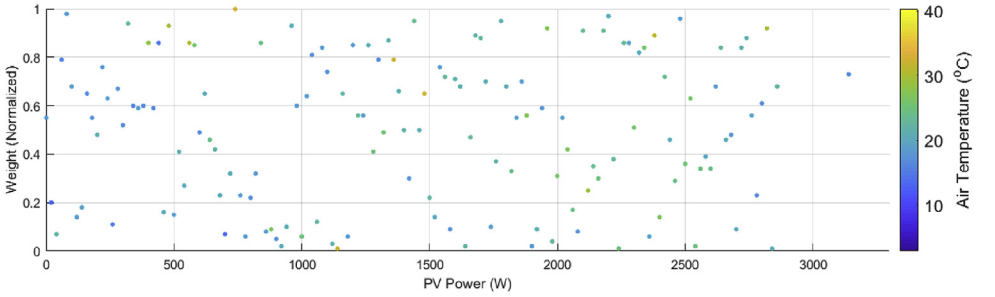
accurate, and no clear relationship between the weights and their properties was found.

The mean weighting factor was calculated from the results and plotted against the PV power. The constraint of the training data has mostly unique PV output throughout the data set. Thus, the PV power should be divided into 20 W segments, where the mean weight of the segments can be calculated. Fig. 17 also shows that the mean weighting factor becomes increasingly dispersed with the increase of PV power due to the reduction in data density for the middle and higher bands, contrary to the low PV power band. Furthermore, the trend line in the figure shows that a higher mean weighting factor is given to the higher PV power data points. Fig. 18 show a different trend due to the significant reduction in the effectual weighting factors.

At either end of the trend line of the weighting matrix, relatively higher mean weighting factors exist for the PV data. The reason is believed to be unrelated to each other. The first low values of PV data are a result of the limitations imposed by using MAPE as the fitness value in the GA. MAPE is sensitive to low values of accuracy. As the actual value approaches 0 W the MAPE value approaches



(a) Solar irradiance



(b) Air temperature

Fig. 17. Weight distribution results GASVM.

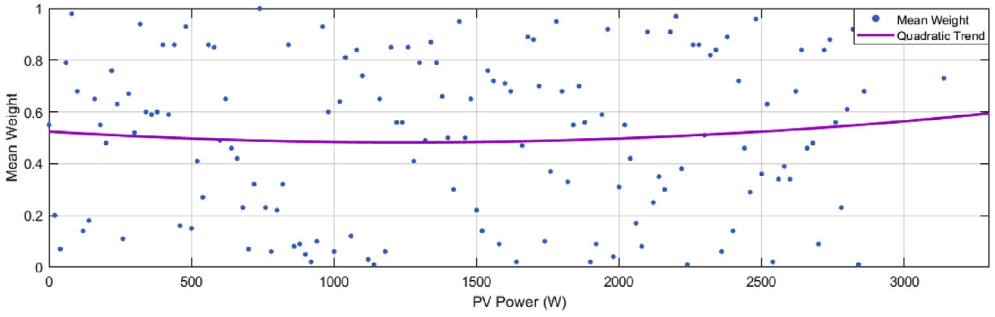


Fig. 18. Mean weight results GASVM support vector machines.

infinity. In addition, the low values of PV power should be filtered out because it represents an insignificant amount compared with the maximum power of the PV panel. On the other end of the trend line, as shown in Fig. 18, the mean weight is caused by consistently predicting lower values than the actual values. The low density of data for higher values of PV increases the dispersal of the points in that region, which also holds a greater tendency of increasing MAPE. Therefore, ideally, a high volume of data would be desirable in calculating MAPE accurately. Furthermore, MAPE converges to a static mean as the volume of data is increased.

The improvement of these results, as opposed to that seen prior to the weight matrix optimization, highlights the need to optimize for further accuracy beyond the architectural design.

The raw data obtained from the self-energized data logger system is used to understand the relationship between the output power and weather parameters in order to classify the PV power. These two parameters are used because they are closely connected to PV power. The PV power is more effectively classified using more data from the environmental factors associated with the PV power which is then used to train the classifier. The assumption is that the other environmental factors, such as humidity, wind direction, and wind speed can limit the ability of the classifiers. They are disregarded in this experiment due to lack of proper data related to these environmental factors highlighting the restriction of the scope of this study.

The methodology of this study divides the process of constructing a classifier in key steps. The hypothesis was that the optimization of the architecture and weight matrix are independent of each other and can be combined to form the most optimal configuration. This assumption was developed to obtain a manageable problem in terms of time and processing power. This simplification allows feasible solutions to the complex problem. However, the impact of the premise statements should be evaluated to ensure that it is not destructive to the result. However, the measurement is desirable to justify the assumption.

The GA has been used as the primary optimization technique throughout this proposed method. The optimization of the architecture was conducted by utilizing the discrete nature of GA to change the discrete property within the classifier. This discrete property changes enabled the architecture to be modified drastically between modes of operation. However, the changes are not well-suited to an optimization technique that responds better in a

continuous space.

GA method of optimization uses random initial values values and therefore the results are not exactly reproducible with regards to the training of the GASVM model. During the iterative training process, GA illustrated the tendency in finding local minima that reduced the optimization effectiveness until it finds its way out of the local minima.

MAPE is the primary performance metric used in the present study. This performance metric, however, had limitations due to the actual data used in the calculation of MAPE. The data increased the MAPE value, introducing bias into the metric results. This bias was imposed through low values of actual data, which exponentially increased the error or the inclusion of low density of points that reduced the convergence and increased the MAPE for those areas of the dataset.

RMSE was at times a better metric in terms of low actual values, where MAPE struggled to provide reliable results. MAPE was a better metric for high RMSE values given that the percentage of the RMSE and the actual data was low. Thus, the inclusion of both factors is beneficial to allow the distinction between the qualities of the performance metrics.

Another method is adjusting the training data to enable more accurate evaluations of the performance metrics. This adjustment would be in the form of removing low values of PV power. Its benefit to RMSE is that the lower PV values add the insignificant amount to the total RMSE value. The removal also eliminates bias from the MAPE value by reducing the increase that is directly related to low values of the actual PV power. Thus, the accuracy of MAPE results significantly increases. But, the expected result would be a significant reduction in MAPE for each classifier.

A limitation, which is related to the platform that was used in training the classifiers, was identified. This limitation was imposed through lack of flexibility within the toolboxes that was used, creating an irregular response to accuracy and processing time, contrary to the nature of the classification models. To overcome this drawback, the package that initially suits the nature of the classification problem should be identified or code the entire classifier from scratch. The lack of control reduced the ability to solve the problem and manipulate the classifier to produce more accurate results.

The outcomes of this study strongly highlight the improvements in the classification models by the use of the proposed GASVM optimization techniques. An appropriate model has also been highlighted and should be considered in conjunction with the optimization technique. Table 5 clearly highlights that the GASVM model proposed in this paper outperforms the conventional SVM model by the margin of about 669.624 W in the RMSE value and 98.7648% of the MAPE error.

7. Conclusion

From the analysis made on the results of the proposed GASVM model used in short-term forecasting of residential PV systems to predict an hour ahead values, the accuracy of the individual classifiers is greatly improved. The optimization through GA was the key step in increasing accuracy. The performance indexes of GASVM model indicates that it outperforms the conventional SVM model by a difference of about 669.624 W in the RMSE value and 98.7648% of the MAPE error. However, increasing the accuracy to an acceptable level is insufficient, which points toward greater efforts with optimization. The limiting factors, such as software and the lack of training data for optimization, significantly reduced the accuracy to less than its perceived actual value. The future implications of this study include identifying prediction models for PV power prediction that focus on the comparative study of classification models and optimization techniques. The method of predicting PV power also contributes to the prediction of bypassing factors, such as efficiency of the PV panel and influence of other environmental factors. The environmental factors that are considered for alternative modelling techniques contribute to the reliability and accuracy of the system. The proposed model greatly simplifies the problem for future evaluations of different locations without requiring a detailed study on the relationships of the PV panel, local environment, and climatic pattern.

Table 5
Statistical performance values.

	SVM	GASVM	Δ
RMSE (W)	680.85	11.226	669.624
MAPE (%)	100.47	1.7052	98.7648

References

[1] N.Y. Mansouri, R.J. Crookes, T. Korakianitis, A projection of energy consumption and carbon dioxide emissions in the electricity sector for Saudi Arabia: the case for carbon capture and storage and solar photovoltaics, Energy Policy 63 (2013) 681–695.

[2] T. Jacob, J. Wahr, W.T. Pfeffer, S. Swenson, Recent contributions of glaciers and ice caps to sea level rise, Nature 482 (7386) (2012) 514.

[3] P. Wolfram, T. Wiedmann, M. Diesendorf, Carbon footprint scenarios for renewable electricity in Australia, J. Clean. Prod. 124 (2016) 236–245.

[4] A.S. Werulkar, P.S. Kulkarni, A case study of residential solar photovoltaic system with utility backup in nagpur, India, Renew. Sustain. Energy Rev. 52 (2015) 1809–1822.

[5] L.J. Walston Jr., K.E. Rollins, K.E. LaGory, K.P. Smith, S.A. Meyers, A preliminary assessment of avian mortality at utility-scale solar energy facilities in the United States, Renew. Energy 92 (2016) 405–414.

[6] A.S. Mundada, Y. Nilsiam, J.M. Pearce, A review of technical requirements for plug-and-play solar photovoltaic microinverter systems in the United States, Sol. Energy 135 (2016) 455–470.

[7] A. Raziei, K. Hallinan, R. Brecha, Clean energy utility for multifamily housing in a deregulated energy market, Energy Build. 127 (2016) 806–817.

[8] M.Q. Raza, M. Nadarajah, C. Ekanayake, On recent advances in pv output power forecast, Sol. Energy 136 (2016) 125–144.

[9] S. Sperati, S. Alessandrini, L. Delle Monache, An application of the ecmwf ensemble prediction system for short-term solar power forecasting, Sol. Energy 133 (2016) 437–450.

[10] A. Hammer, D. Heinemann, C. Hoyer, R. Kuhlemann, E. Lorenz, R. Müller, H.G. Beyer, Solar energy assessment using remote sensing technologies, Remote Sens. Environ. 86 (3) (2003) 423–432.

[11] P. Lauret, C. Voyant, T. Soubdhan, M. David, P. Poggi, A benchmarking of machine learning techniques for solar radiation forecasting in an insular context, Sol. Energy 112 (2015) 446–457.

[12] L. Olatomiwa, S. Mekhilef, S. Shamshirband, K. Mohammadi, D. Petković, C. Sudheer, A support vector machine–firefly algorithm-based model for global solar radiation prediction, Sol. Energy 115 (2015) 632–644.

[13] A. Lemmens, C. Croux, Bagging and boosting classification trees to predict churn, J. Market. Res. 43 (2) (2006) 276–286.

[14] C. Paoli, C. Voyant, M. Muselli, M.-L. Nivet, Forecasting of preprocessed daily solar radiation time series using neural networks, Sol. Energy 84 (12) (2010) 2146–2160.

[15] J. Wang, R. Ran, Z. Song, J. Sun, Short-term photovoltaic power generation forecasting based on environmental factors and ga-svm, J. Electr. Eng. Technol. 12 (2017) 64–71.

[16] I. Rish, An empirical study of the naive bayes classifier, in: IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence, vol. 3, IBM, 2001, pp. 41–46.

[17] M. Diagne, M. David, P. Lauret, J. Boland, N. Schmutz, Review of solar irradiance forecasting methods and a proposition for small-scale insular grids, Renew. Sustain. Energy Rev. 27 (2013) 65–76.

[18] F.H. Gandoman, F. Raeisi, A. Ahmadi, A literature review on estimating of pv-array hourly power under cloudy weather conditions, Renew. Sustain. Energy Rev. 63 (2016) 579–592.

[19] G. Reikard, Predicting solar radiation at high resolutions: a comparison of time series forecasts, Sol. Energy 83 (3) (2009) 342–349.

[20] M. Moharrampour, S. Sohrabi, J. Vakili, Comparison of Support Vector Machines (SVM) and Autoregressive integrated moving average (ARIMA) in daily flow forecasting, J. River Eng. 1 (2013).

[21] S. Leva, A. Dolara, F. Grimaccia, M. Mussetta, E. Ogliari, Analysis and validation of 24 hours ahead neural network forecasting of photovoltaic output power, Math. Comput. Simulat. 131 (2017) 88–100.

[22] J. Cao, X. Lin, Application of the diagonal recurrent wavelet neural network to

solar irradiation forecast assisted with fuzzy technique, Eng. Appl. Artif. Intell. 21 (8) (2008) 1255–1263.
[23] F. Almonacid, P. Pérez-Higueras, E.F. Fernández, L. Hontoria, A methodology based on dynamic artificial neural network for short-term forecasting of the power output of a pv generator, Energy Convers. Manag. 85 (2014) 389–398.
[24] K.M. Rajpoot, N.M. Rajpoot, Wavelets and support vector machines for texture classification, in: Multitopic Conference, 2004. Proceedings of INMIC 2004. 8th International, IEEE, 2004, pp. 328–333.
[25] E.W. Law, A.A. Prasad, M. Kay, R.A. Taylor, Direct normal irradiance forecasting and its application to concentrated solar thermal output forecasting–a review, Sol. Energy 108 (2014) 287–307.
[26] M. Ghofrani, M. Ghayekhloo, R. Azimi, A novel soft computing framework for solar radiation forecasting, Appl. Soft Comput. 48 (2016) 207–216.
[27] P. Chen, L. Yuan, Y. He, S. Luo, An improved svm classifier based on double chains quantum genetic algorithm and its application in analogue circuit diagnosis, Neurocomputing 211 (2016) 202–211.
[28] Y. Li, Y. He, Y. Su, L. Shu, Forecasting the daily power output of a grid-connected photovoltaic system based on multivariate adaptive regression splines, Appl. Energy 180 (2016) 392–401.
[29] J.C. Mojumder, H.C. Ong, W.T. Chong, S. Shamshirband, et al., Application of support vector machine for prediction of electrical and thermal performance in pv/t system, Energy Build. 111 (2016) 267–277.
[30] R.-E. Fan, P.-H. Chen, C.-J. Lin, Working set selection using second order information for training support vector machines, J. Mach. Learn. Res. 6 (Dec) (2005) 1889–1918.
[31] E.P. Ijjina, C.K. Mohan, Human action recognition using action bank features and convolutional neural networks, in: Asian Conference on Computer Vision, Springer, 2014, pp. 328–339.
[32] Z. Qiongbing, D. Lixin, A new crossover mechanism for genetic algorithms with variable-length chromosomes for path optimization problems, Expert Syst. Appl. 60 (2016) 183–189.