Technol Econ Smart Grids Sustain Energy (2017) 2: 1
DOI 10.1007/s40866-016-0018-x

ORIGINAL PAPER

Comparative Analysis of Intelligently Tuned Support Vector Regression Models for Short Term Load Forecasting in Smart Grid Framework

Sreenu Sreekumar[1] · Jatin Verma[1] · Sujil A[1] · Rajesh Kumar[1]

Received: 21 April 2015 / Accepted: 10 December 2016 / Published online: 28 December 2016

Abstract A large amount of work has been taken place, if we talk about forecasting in the fields of power system. Various reforms in the existing techniques have proved to be helpful in providing guidance to researchers for developing efficient algorithms exhibiting greater accuracy. This paper presents three forecasting models viz. three-day-trained Support Vector Regression model and parameter optimized Support Vector Regression using Genetic Algorithm (SVRGA) and that using Particle Swarm Optimization (SVRPSO). Unlike existing models, these models accomplish accurate forecasting by optimizing the regularized structural risk function. The models make use of previous three days hourly load data for predicting next day hourly load. This paper performs a comparative study between GA and PSO on the grounds of optimization of the hyper-parameters of SVR model.

Keywords Genetic Algorithm · Hyper parameter optimization · Particle Swarm Optimization Support Vector Regression · Short Term Load Forecasting

Introduction

With the invention of fire, humans have learnt to distinguish themselves from other creatures by being able to invent things to shape the circumstances in their own favour. In the same way, when humans realised the need for electricity to be transmitted to distances, they discovered the way to transmit AC power in 1886 [1]. For few decades, with the adequacy of the conventional resources, limited demand and monopolistic nature of electricity supplying industry, the power system technology retained its fundamental framework. However, with the increase in demand in the past few decades, circumstances have again beckoned for another reformation on account of rapid depletion of conventional resources which produce electricity. For tackling this adversity, humans have come up with various new concepts of deregulation, restructuring and the most promising concept of Smart Grid [2]. Primarily, all these concepts are meant for the reduction of conventional fuel consumption along with the reduction of price. There has been a major paradigm shift towards reliable, automated and optimum computational approaches for achieving highest possible system efficiency. Since the notion is to reduce fuel consumption with optimum demand fulfilment, researchers have tried to find out approaches which could actually find out the exact amount of fuel needed to fulfil the consumers' future demand. This approach is formally known as Electrical Load Forecasting.

Electric load forecasting is the practice used to forecast upcoming electric load using known historical load as well as historical, current and forecasted weather information. Load forecasting is normally carried out to help the plant operators in creating strategic decisions associated with unit commitment, security assessments, maintenance scheduling, fuel purchase, financial investments, plant expansion, economic load dispatch and various issues that affect the efficiency and reliability of the plant [3, 4]. Most of the available and modified models for forecasting purposes have already been tested for load forecasting with appreciable

✉ Sreenu Sreekumar
sreekumarsreenu@gmail.com

[1] Malaviya National Institute Of Technology, Jaipur, Rajasthan, India

success [5]. On a broad sense, load forecasting can be divided into four major categories [6]:

Long Term Electric Load Forecasting

As it is clear from Fig. 1, long term load forecasting mainly refers to obtaining forecasts ranging from one to ten years or more. Utilities perform this type of forecasting for long-term planning regarding an expansion of the plant, large investments, maintenance scheduling, security and various economic issues, etc. In [7], Carpinteiro et al. have performed long-term peak-load and mean-load forecasting with an objective of comparing two prominent techniques viz. Hierarchical Hybrid Neural Model (HHNM) and Multi-layer Perceptron (MLP). The hierarchical topology leads to efficient prediction. However, there is a scope of increasing the accuracy by tuning the governing parameters and by the application of pre-processing techniques on the data.

Medium Term Load Forecasting

Medium term load forecasting comes under the category of obtaining forecasts ranging from one week to twelve months. Utilities perform this type of forecasting mainly to achieve two objectives- first being the determination of the amount of fuel purchase and second being the scheduling of maintenance [8]. Basically, monthly peak loads can be used to get an idea of the future demand. Growth rates may also be considered in case considerable variation has taken place during previous months. Similar to the work of Yalcinoz and Eminoglu in [8], various architectures of neural networks have been implemented with considerable accuracies but with the requirement of a large amount of data.

Short Term Load Forecasting

The short term time leads may vary from one hour or day to one week. Similar to medium term forecasting, short term forecasting also helps the utility to decide the maintenance time. It helps them to schedule their outage time and security analysis. The cost of electricity can be optimized by the implementation of economic load scheduling [9].

Very Short Term Load Forecasting

Very short term time leads may vary from few seconds to an hour or so. Very Short Term Load Forecasting holds the responsibility to reconcile the demand and generated load values in real time with time leads of around 15 minutes. Basically, very short term forecasts help to match the generation and demand values, regulate the vulnerable system frequency, and moreover, dispatch the load economically [10]. H. Y. Yang et al. [11] have designed a Fuzzy Neural System for dealing with load forecasting having 15 minutes time lead. A chaotic dynamic reconstruction technique has been employed for obtaining the value of correlation dimension for estimating the proposed model's order. In addition, the effect of any residual computational error during the estimation of the correlational dimension itself has been overcome by a dimension switching detector developed by the authors.

Importance of Short Term Load Forecasting

Great emphasis has been shown towards Short Term Load Forecasting even at the time when Neural Networks were not so prevalent in Electrical Load Forecasting [9]. As the electricity market started undergoing deregulation and restructuring, the reduction of electricity price became a prime objective for various market players. This opened a window creating a conducive environment for the players to bring resilient innovation to achieve cent percent efficiency. Similar reforms took place in the fields of electrical load forecasting aimed to implement economic load dispatch and to prevent overloading and equipment failure. The determination of consumer demand and setting up of electricity prices depends upon the load forecasts for the power supplying competitors. Thus, the accuracy of forecasts with short term time leads becomes a deciding factor for various types of bids offered by the suppliers

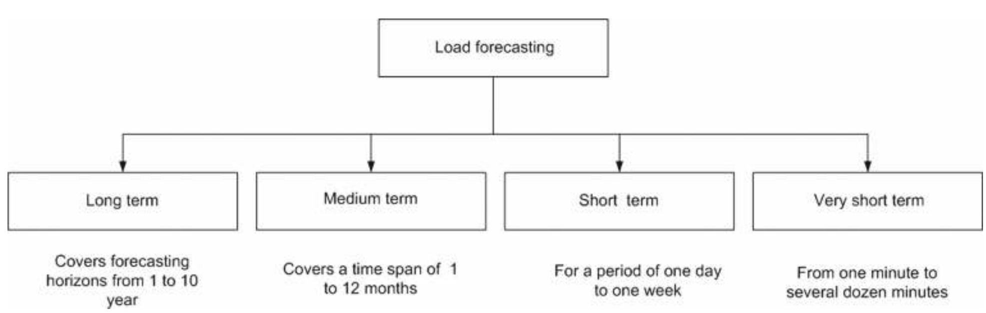

Fig. 1 Classification of Load Forecasting

[12]. In the electricity market, the energy transactions are dealt on the basis of accurate short-term load forecasts. Unit commitment and economic load dispatch take place only after properly observing the current forecast values [13].

Gross and Galiana [9] have performed a survey describing the literature based on the features which short-term load forecasting may depend on. The work was focused on a statistical approach, namely ARMA (Autoregressive Moving Average) model. In [13, 14], it can be seen that various modifications in the conventional ARMA model have resulted in better accuracy. However, these conventional statistical approaches were later replaced by linear regression techniques incorporating eclectically chosen features affecting load, which is evident from the work of Papalexopoulos and Hesterberg [15]. The work, shown by Park et al. [3], can be considered as a landmark which was followed by many researchers for implementing Neural Networks in the fields of Electrical Load Forecasting. This lead to the widening of the scope of performance improvement by integration various feature extraction techniques, such as the application of wavelet transform in the works of Reis et al. [12].

Despite the bright future of the Artificial Neural Network (ANN) techniques, it was soon found out that the performance of this technique was constricted by the phenomenon of overfitting. Also, a large amount of historical data was required for its training. This was the reason for the need of algorithms like Bayesian Neural Networks [16], Support Vector Regression (SVR), etc. which had the ability to overcome the overfitting of the model. According to Hong [17], neural networks fail to provide the most accurate forecasts because they try to minimize the empirical risk. On the other hand, algorithms under Support Vector Machines (SVMs) minimize structural risk. Unlike neural networks, SVMs provide a unique and globally optimal solution. Hong [17] has applied Immune Algorithm to optimize the governing parameters of Support Vector Regression (SVR). Another work [18] employs firefly algorithm based memetic algorithm to do the optimization task. SVR models have also been used in the field of short-term Wind Power Forecasting [19]. This paper is based on the comparative analysis of two optimization techniques, viz. Genetic Algorithm (GA) and Particle Swarm optimization (PSO), used to optimize the hyper-parameters of the SVR model [20–22].

This paper is further organized as follows: "Support Vector Regression Model" defines support vector regression model. Section "Hyper Parameter Optimization using Inteligent Optimization Techniques" explains parameter optimization using Genetic Algorithm and Particle Swarm Optimization. Section "Data Selection and Methodology" explains data selection and methodology. Section "Results and Discussion" discusses the forecasted results of SVR, SVRGA and SVRPSO. Finally, "Conclusion" derives some conclusions about the optimization techniques.

Support Vector Regression Model

Support Vector Machines have been a ground-breaking innovation that have brought a necessary reformation in the fields of supervised machine learning. These were first developed for the identification of pattern only for classification of data into specified classes. For this, a set of labelled examples has to be given to the algorithm for training purposes. Training leads to the determination of parameters of the model which is then capable of labelling new unseen examples on the basis of the features or the pattern related to the new input data. According to the concept of SVM, the algorithm tries to fit a hyperplane between different classes or categories of the data such that the distance of the nearest data from the hyperplane is maximised from all sides. Later this concept was extended to non-linear regression, which enabled the model to perform prediction as well using different types of kernels [17]. This extended version of SVM is called Support Vector Regression (SVR). Another thing that distinguishes this algorithm is that it uses selective data for training. Data points, that come under a pre-defined error tube called ϵ -tube, do not constitute the cost function used for determining the model. Similar concept applies to SVR algorithm [18, 19].

The notion behind SVR is nothing but to map the original data x into a higher-dimensional space [23].

Consider a set of data

$$G = (x_i, d_i) \tag{1}$$

where, x_i, d_i and N are input vector, actual values and number of data pattern respectievely. The SVR function is given by

$$Y = f(x) = W\psi(x) + b \tag{2}$$

Where, $\psi(x)$ is the feature which is obtained by nonlinear mapping of input space. The coefficients w and b are calculated by minimizing the regularized risk function which is given by,

$$R(C) = (C/N)\sum_{i=1}^{N} L_\epsilon(d_i, y_i) + \frac{\|W^2\|}{2} \tag{3}$$

where, C and ϵ are prescribed parameters, and

$$L_\epsilon(d, y) = \begin{bmatrix} 0 & if|d-y| <= 0 \\ |d-y| - \epsilon & otherwise \end{bmatrix} \tag{4}$$

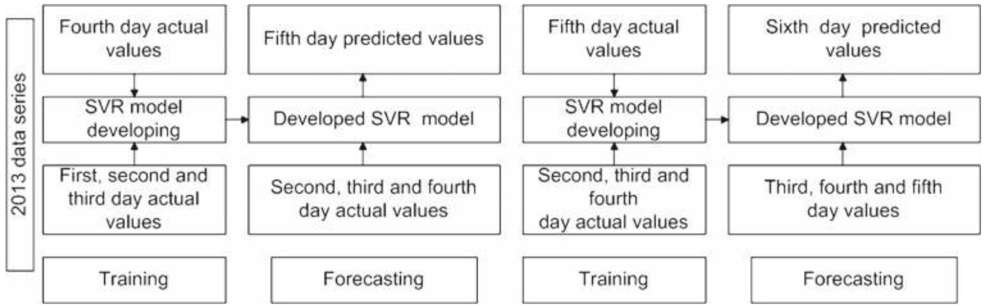

Fig. 2 Methodology for forecasting in SVR

This function creates a tube which has error less than ϵ. $L_\epsilon(d, y)$ is called the ϵ insensitive loss function. If the forecasted values are within the ϵ tube, the loss function become zero. The flatness of the function is measured by the second term $\frac{\|W^2\|}{2}$ [23]. C is the trade off between empirical risk and model flatness, user can define both of these parameters - C and ϵ. There are two slack variables, μ and μ^*, which represent the distance from the actual values to the corresponding boundary values of the ϵ-tube [24]. The ϵ-tube and support vectors for the training data from November 14, 2013 is shown in Fig. 5.

The Eq. 2 can be written as

$$min\ R(W, \mu, \mu^*) = \|W^2\| + C(\sum_{i=1}^{N}(\mu_i + \mu^*{}_i) \qquad (5)$$

With constraints,

$$W\psi(x_i) + b - d_i \leq \epsilon + \mu^*{}_i \qquad (6)$$

$$d_i - W\psi(x_i) - b \leq \epsilon + \mu_i \qquad (7)$$

$$\mu_i, \mu^*{}_i \geq 0 \qquad (8)$$

where, $i = 1, 2, 3........N$.

Primal lagrangian equation is used for solving this constrained optimization problem, which is in the form of,

$L(W, b, \mu_i, \mu_i^*, \alpha_i, \alpha_i^*, \beta_i, \beta_i^*) = \|W^2\| + C(\sum_{i=1}^{N}(\mu_i + \mu^*{}_i) - \sum_i^N \beta_i(W\psi(x_i) + b - d_i + \epsilon + \mu^*{}_i) - \sum_i^N \beta^*{}_i(d_i - W\psi(x_i) - b + \epsilon + \mu^*{}_i) - \sum_i^N(\alpha_i\mu_i + \alpha^*\mu^*{}_i)$

This equation is maximized with respect to nonnegative lagrangian multipliers $\alpha_i, \alpha_i^*, \beta_i$ and β_i^*, minimized with respect to the primal variables W, b, μ_i and μ_i^*, which leads to the equations

$$\frac{\partial L}{\partial W} = W - \sum_i^N(\beta_i - \beta_i^*)\psi(x_i) = 0 \qquad (9)$$

$$\frac{\partial L}{\partial b} = \sum_i^N(\beta_i - \beta_i^*) = 0 \qquad (10)$$

$$\frac{\partial L}{\partial \mu_i} = C - \beta_i - \alpha_i = 0 \qquad (11)$$

$$\frac{\partial L}{\partial \mu_i^*} = C - \beta_i^* - \alpha_i^* = 0 \qquad (12)$$

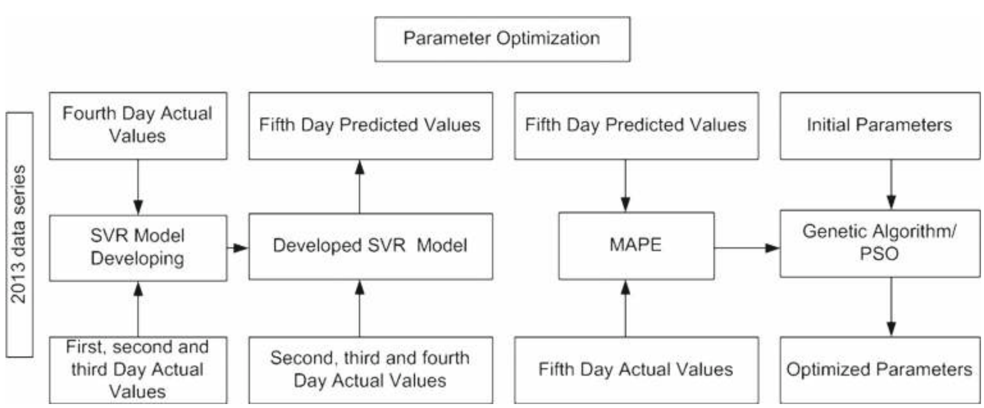

Fig. 3 Methodology for paremeter optimization in SVRGA and SVRPSO

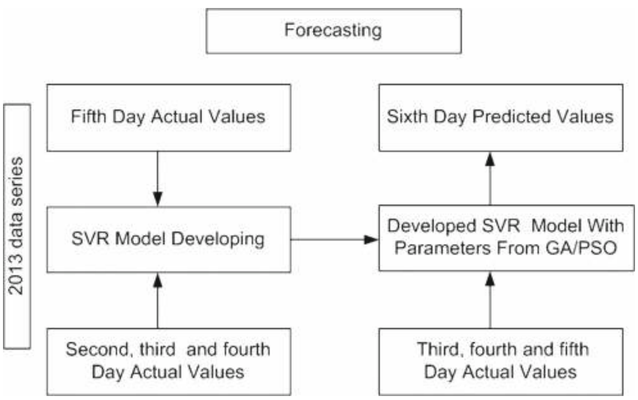

Fig. 4 Methodology for forecasting in SVRGA and SVRPSO

The application of Karush-Kuhn-Tucker conditions, the substitution of Eqs. 6 to 9 into Eq. 5 and the kernel $K(x_i, x_j) = \psi(x_i * x_j)$ gives the dual Lagrangian equation

$$D(\beta_i, \beta_i^*) = \sum_i^N d_i(\beta_i - \beta_i^*) - \sum_i^N \epsilon(\beta_i + \beta_i^*) - (1/2)\sum_i^N\sum_j^N(\beta_i - \beta_i^*)(\beta_j - \beta_j^*)K(x_i, x_j) \tag{13}$$

Subject to the constraints,

$$\sum_i^N(\beta_i - \beta_i^*) = 0 \tag{14}$$

$$0 \leq \beta_i \leq 0 \tag{15}$$

$$0 \leq \beta_i^* \leq 0 \tag{16}$$

$$\beta_i * \beta_i^* = 0 \tag{17}$$

There are different types of kernel functions like Gaussian kernel, polynomial kernel, laplacian kernel, exponential kernel, Cauchy kernel, generalised T- student kernel etc., which are presently used for non-linear mapping. From these kernels, Gaussian kernel gives best performance for load forecasting [25]. Gaussian function is created by composing the exponential function with a concave quadratic function. Thus, the Gaussian functions are those functions whose logarithm is a concave quadratic function.

$$K(x_i, x_j) = Ae^{-(x_i - x_j)^2/\sigma^2} \tag{18}$$

Where, A is the amplitude (for load forecasting, we have taken it as 1), x_i and x_j are two input vectors, σ is the standard deviation or Gaussian RMS width. The optimization of lagrangian multipliers β_i and β_i^* can be calculated by quadratic programming. The maximization quadratic function for the above equation is given by

$$Max(\beta) = -0.5\beta^T H\beta + f^T\beta \tag{19}$$

which is subject to the same constraints. Here, H is the hessian matrix given by,

$$H = \begin{bmatrix} h & -h \\ -h & h \end{bmatrix} \tag{20}$$

$$h(i, j) = K({x_i}^T, x_j) + 1 \tag{21}$$

$$f = \left[\epsilon - y_1 \;\; \epsilon - y_2 \;\; .... \;\; \epsilon - y_n \;\; \epsilon + y_1 \;\; .... \;\; \epsilon + y_n\right] \tag{22}$$

Where $y_1, y_2 ....y_n$ are the training stage output values.

The regression hyperplane's optimal desired weight vector is given by

$$W = \sum_i^N(\beta_i - \beta_i^*)\psi(x) \tag{23}$$

Table 1 MAE and MAPE of SVR and SVRGA models for different days

Days	MAE			MAPE		
	SVR	SVRGA	SVRPSO	SVR	SVRGA	SVRPSO
January 9	851.79	851.79	851.79	2.64	2.64	2.64
February 14	771.75	552.93	535.05	2.36	1.70	1.65
March 27	697.26	690.11	671.34	2.37	2.36	2.27
April 21	662.70	627.67	618.94	2.64	2.45	2.39
May 3	670.38	651.77	651.77	2.60	2.56	2.56
June 21	594.12	565.25	548.93	1.74	1.66	1.61
July 16	1182.24	1182.24	1182.24	2.84	2.84	2.84
August 13	741.08	697.27	634.53	2.24	2.15	1.91
September 18	472.37	404.95	356.08	1.69	1.44	1.27
October 31	592.35	592.35	592.35	2.21	2.21	2.21
November 14	765.94	679.45	660.31	2.4	2.12	2.06
December 4	726.43	647.54	618.42	2.31	2.06	1.98

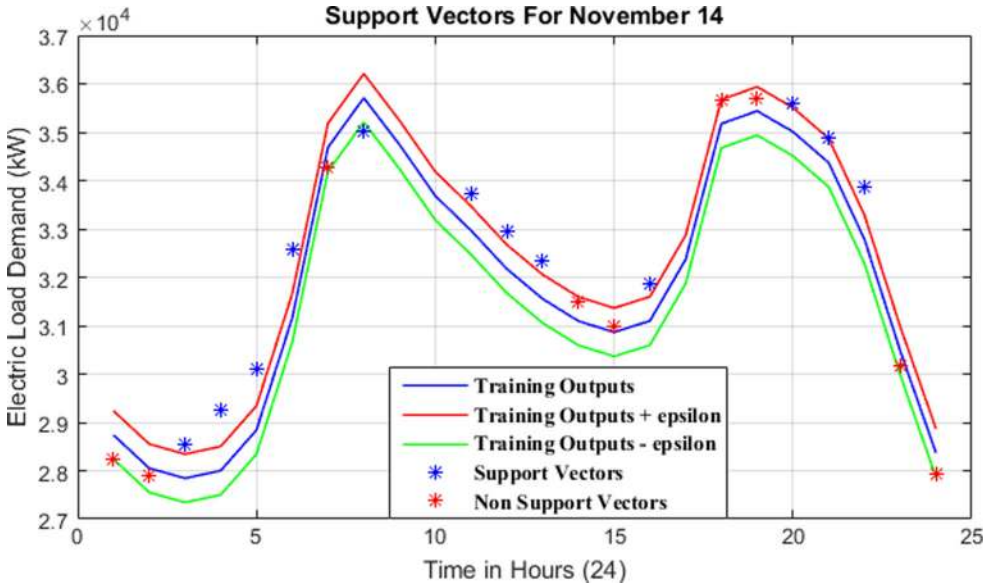

Fig. 5 Support Vectors And Epsilon Tube For November 14

If β_i lies is in between 0 and C, the co-efficient b is given by

$$b = y_i - W\psi(x) - \epsilon \tag{24}$$

If β_i^* lies is in between 0 and C, the co-efficient b is given by

$$b = y_i - W\psi(x) + \epsilon \tag{25}$$

The predicted value Q is given by

$$Q = \sum_{i}^{N} (\beta_i - \beta_i^*) K(x_i, x_j) + b \tag{26}$$

Hyper Parameter Optimization using Inteligent Optimization Techniques

The choice of the parameters (σ, C and ϵ) of an SVR model is crucial for prediction accuracy. In simple SVR model, the parameters σ, C and ϵ are chosen as 10000, 0.78 and 500 respectively based on previous knowledge. There are number of prevailing practical methodologies towards selection of parameters such as user-defined, based on previous knowledge and experiences, cross-validation, and asymptotical optimization. Nevertheless, manual tactics for the selection of these hyper-parameters are unreliable, and we need an efficient optimization algorithm to do this task. Therefore, genetic algorithm (GA) and Particle Swarm Optimization (PSO) are used in the proposed SVR model for parameter optimization. Thousands iterations are used for each set of parameter [22].

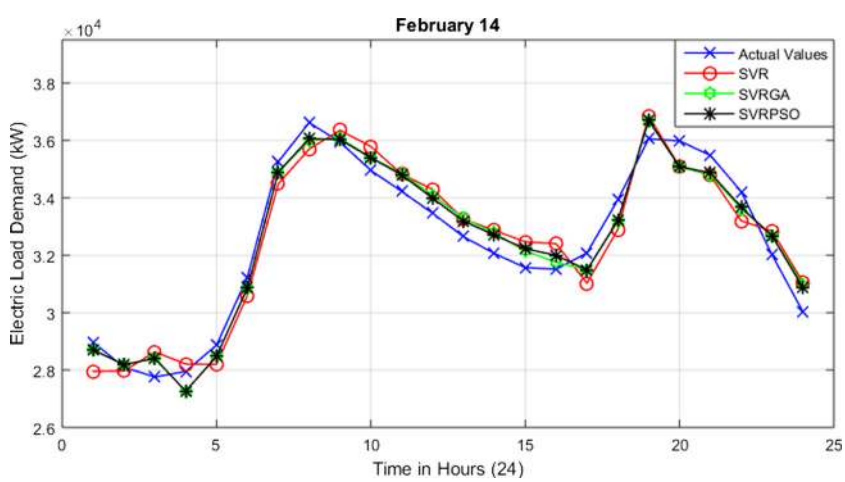

Fig. 6 Performance Plot Of February 14

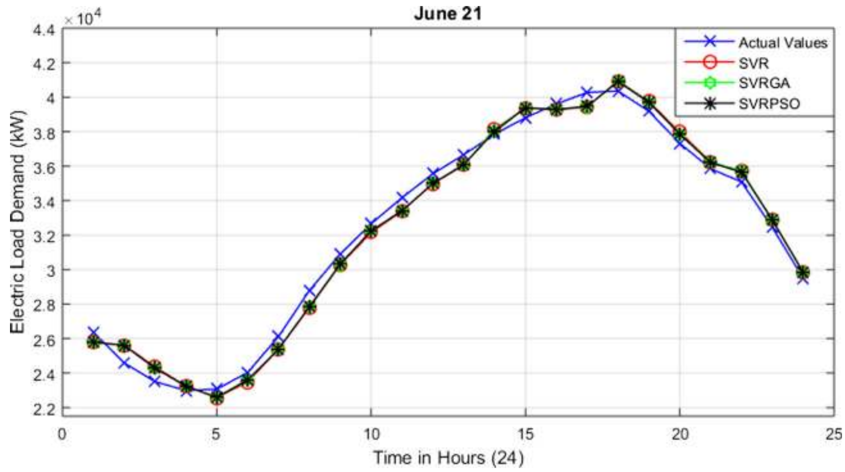

Fig. 7 Performance Plots Of June 21

Genetic Algorithm

A genetic algorithm embodies mainly three stages (operations), which are the mutation, the crossover and fitness selection. First of all, a certain population is initialized. Then, one-by-one each of the aforementioned operations are performed on the population. This leads to the final selection of a population consisting of elements that are considered to give the best result in terms of the value of the objective function. If the algorithm is likely to get stuck in a local minima, then mutation comes handy by changing certain characteristics of the element so as to take it out of the local minima zone. This change in characteristics in mutation is random. During the iteration, certain elements of population are allowed to crossover to generate elements with better fitness values. Also, this helps in converging towards the result faster [26]. This is a nature-inspired algorithm.

Particle Swarm Optimization

This optimization technique uses various randomnly initialized particles in the search space. Then, a rigorous search is run to look for an optimum solution in the nearby region. Thus, we have a certain number of candidate solutions, here named particles, that seem promising. As stated earlier, these so-called particles are allowed to enquire in their surroundings for better solution. Their movement is governed by a mathematical formula through which the direction and speed of the particles can be controlled. The particles' position can be dependent on the initialization of the population. Nevertheless, during the process, the swarm of particles tends to move towards better solutions that could fit the constraints well [27].

The fitness function for the parameter optimization in both SVRGA and SVRPSO is MAPE, which is given by

$$MAPE = \sum_{i}^{N} \frac{\|(Y_i - ((\beta_i - \beta_i^*)K(x_i, x_j) + b))\|}{Y_i * 24} * 100 \tag{27}$$

Where Y_i is the actual value. For calculating β, kernel and b, we need optimized parameters $C, \epsilon and \sigma$ which have lower an upper limit as 3000 to 10000, 0 to 1000 and 0 to 1 respectively [22].

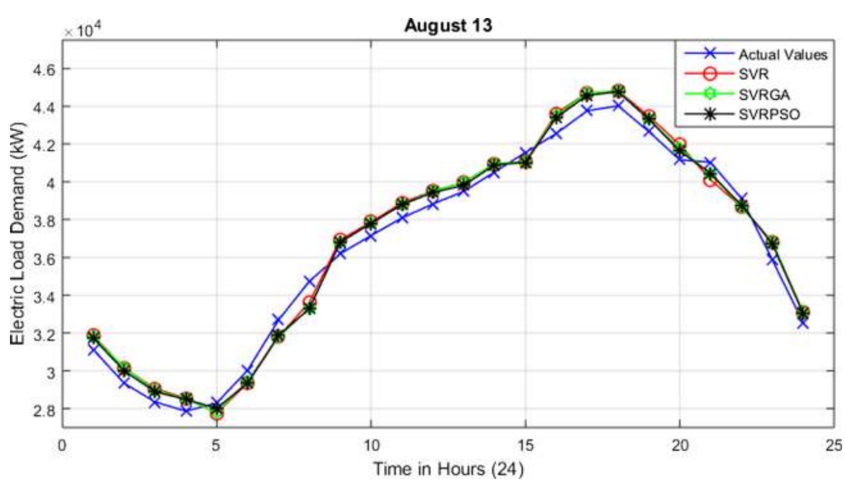

Fig. 8 Performance Plots Of August 13

Data Selection and Methodology

The load consumption data for training and testing of the proposed models has been taken from PJM Mid-Atlantic Region [28]. In simple SVR and SVRGA model for the prediction of load consumption on a particular day in 2013, previous 72 hours load data is used. For the sake of fair comparison between the effects of optimization techniques, temperature data has not been considered as an extra feature. The methodologies for SVR and SVRGA are shown in Figs. 2, 3 and 4.

Results and Discussion

The load forecasted by these models were used to calculate the error. The primary criteria used to measure the performance of the proposed model are taken to be mean absolute percentage error (MAPE) and mean average error (MAE). Table 1 contains MAE and MAPE of different models. Error is calculated by taking the difference between actual value and forecasted value for each data point Fig. 5.

$$E = Y - Q \tag{28}$$

Where Y=actual output, Q= predicted value. Mean average error is calculated by taking the mean of absolute error.

$$MAE = E_A/N \tag{29}$$

$$E_A = E_1 + E_2 + E_3 .....E_{24} \tag{30}$$

Where E_1, E_2...E_{24} are absolute values of individual errors. Percentage error is calculated by dividing absolute value of error by corresponding actual value and then multiplied it by 100. Mean average percentage error is the mean of percentage error

$$MAPE = (1/N) \sum |PE| \tag{31}$$

The mean absolute percentage error in SVR is 2.33 % and that in SVRGA and SVRPSO is 2.18 % and 2.11 % respectively. it is clear that there is 6.8 % reduction in MAPE from SVR model to SVRGA model and 10.42 % reduction in MAPE from SVR model to SVRPSO model. Table 1 shows MAE and MAPE of different days. Figures 6, 7, 8 and 9 shows the days February 14, June 21, August 13 and September 18 respectively.

One of the main factors affecting load consumption profile of mid atlantic region is temperature. We have considered average monthly temperature data for discussing the predicted load consumption pattern from US climatic data center [29]. The months January, February, March, April, October, November and December have temperature ranging from -4 to 20 degree celsius. This range corresponds to cool weather, thus electric load demand increases in the form of heating load, for e.g. heaters, geysers, etc, and therefore overall load consumption is higher as is evident from figures of corresponding months. Rest of the months have moderate weather. July and August are the hottest months with temperature reaching upto 31 degree Celsius, which again increases electric load demand in the form of cooling load, for e.g. coolers, ACs, etc. and hence their peak loads are higher than other months; the same can be seen in the corresponding figures. SVR, SVRGA and SVRPSO have minumum mean absolute percentage error on September 18, because in this day the load pattern exactly follows the previous three days. January, July, March and October have same results for all the models, because in these months for a wide range parameters (σ, C and ϵ) it shows almost similar error. July is the hottest month; in each day temperature variation is higher, which results highest MAPE (2.84 %). January, April, June and July have MAPE in between 2.5 % to 2.9 %, rest days have MAPE less than 2.5 %. In August 13, the unexpected error in forecasting at 1 am was caused by the failure of one part of the system.

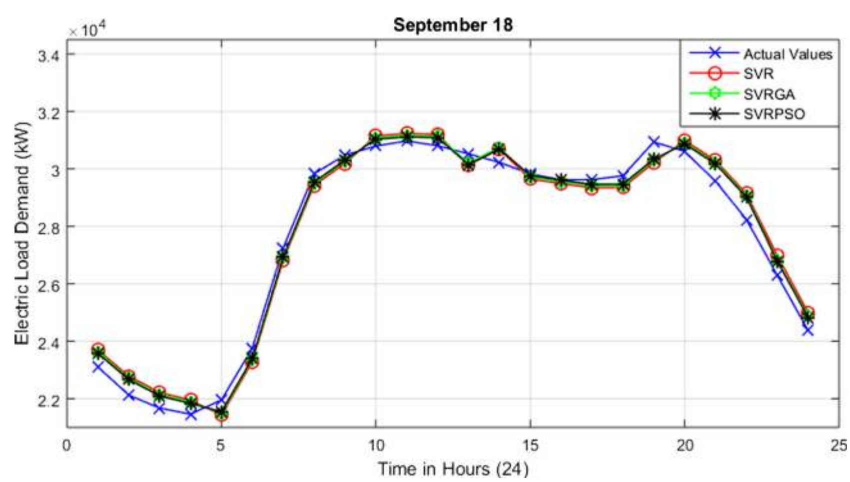

Fig. 9 Performance Plots Of September 18

Conclusion

In this paper, three models based on support vector regression are implemented for short term load forecasting. Support vector regression has proved to be very efficient for forecasting problems which reduces the requirement of huge amount of training data which is essential for the existing models like ANN. The proposed three day trained support vector regression with fixed parameter model, GA and PSO optimized support vector regression model give 97.67 %, 97.82 % and 97.89 % accuracies respectively. Results show that these methods are very effective in real time forecasting, because these models use only 72 hour historical data for 24 hour ahead prediction. However, GA and PSO based optimization are time consuming. Thus, an efficient algorithm having less processing time is required for hyper parameter optimization in SVR. The addition of more factors affecting load demand and fast method for hyper parameter optimization will be considered in the future work.

References

1. Rajkumar N (2011) Direct current and alternating current systems, Electrical Engineering Vol. I Encyclopedia of Life Support Systems (EOLSS)
2. Gungor VC, Sahin D, Kocak T, Ergut S, Buccella C, Cecati C, Hancke GP (2011) Smart grid technologies: communication technologies and standards. IEEE Trans Ind Inf 7(4):529–539
3. Park DC, El-Sharkawi MA, Marks RJ, Atlas LE, Damborg MJ (1991) Electric load forecasting using an artificial neural network. IEEE Trans Power Syst 6(2):442–449
4. Metaxiotis K, Kagiannas A, Askounis D, Psarras J (2003) Artificial intelligence in short term electric load forecasting: a state-of-the-art survey for the researcher. Energy Convers Manage 44(9):1525–1534
5. Chen B-J, Chang M-W, Lin C-J (2004) Load forecasting using support vector machines: A study on EUNITE competition 2001. IEEE Trans Power Syst 19(4):1821–1830
6. Feinberg EA, Genethliou D (2005) Load forecasting. In: Applied Mathematics for Restructured Electric Power Systems. Springer US, pp 269–285
7. Carpinteiro OAS, Leme RC, Zambroni de Souza AC, Pinheiro CAM, Moreira EM (2007) Long-term load forecasting via a hierarchical neural model with time integrators. Electr Power Syst Res 77(3):371–378
8. Yalcinoz T, Eminoglu U (2005) Short term and medium term power distribution load forecasting by neural networks. Energy Convers Manage 46(9):1393–1405
9. Gross G, Galiana FD (1987) Short-term load forecasting. Proc IEEE 75(12):1558–1573
10. Liu K, Subbarayan S, Shoults RR, Manry MT, Kwan C, Lewis FL, Naccarino J (1996) Comparison of very short-term load forecasting techniques. IEEE Trans Power Syst 11(2):877–882
11. Yang HY, Ye H, Wang G, Khan J, Tongfu H (2006) Fuzzy neural very-short-term load forecasting based on chaotic dynamics reconstruction. Chaos, Solitons Fractals 29(2):462–469
12. Reis AJR, Alves da Silva AP (2005) Feature extraction via multiresolution analysis for short-term load forecasting. IEEE Trans Power Syst 20(1):189–198
13. Huang S-J, Shih K-R (2003) Short-term load forecasting via ARMA model identification including non-Gaussian process considerations. IEEE Trans Power Syst 18(2):673–679
14. Amjady N (1990) Short-term hourly load forecasting using time-series modeling with peak load estimation capability 5(4):1535–1547
15. Papalexopoulos AD, Hesterberg TC (2001) A regression-based approach to short-term system load forecasting. IEEE Trans Power Syst 16(1):498–505
16. Lauret P, Fock E, Randrianarivony RN, Manicom-Ramsamy J-F (2008) Bayesian neural network approach to short time load forecasting. Energy Convers Manage 49(5):1156–1166
17. Hong W-C (2009) Electric load forecasting by support vector model. Appl Math Modell 33:2444–2454
18. Hu Z, Bao Y, Xiong T Electricity Load Forecasting Using Support Vector Regression with Memetic Algorithms. Hindawi Publishing CorporationThe Scientific World Journal Volume 2013, Article ID 292575
19. Zeng J, Qiao W (2011) Support vector machine-based short-term wind power forecasting. In: Power Systems Conference and Exposition (PSCE), 2011 IEEE/PES. IEEE, pp 1–8
20. Pan F, Cheng H-z, YANG J-f, Zhang C, PAN Z-d (2004) Power system short-term load forecasting based on support vector machines. Power System Technology 21
21. Mohandes MP (2002) Support vector machines for short term electrical load forecasting. Int J Energy Res 26(4):335–345
22. Sreekumar S, A S, Verma J, Kumar R One Day Ahead Prediction of Hourly Electrical Load using Genetically Tuned Support Vector Regression for Smart Grid Frame Work, IEEE Conference on Recent advances in engineering and computational Sciences (RAECS 2015, 21-22 December), 978-1-4673-8253-3/15/$31.00 2015 IEEE
23. Elattar EE, Goulermas J, Wu QH (2010) Electric load forecasting based on locally weighted support vector regression. In: IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol 40, pp 438–447
24. Okasha MK (2014) Using Support Vector Machines in Financial Time Series Forecasting. In: International Journal of Statistics and Applications, vol 4, pp 28–39
25. Kim K-j (2003) Using support vector machines financial time series forecasting. In: Neurocomputing, vol 55, pp 307–319
26. Davis L (1991) Handbook of genetic algorithms
27. Poli R, Kennedy J, Blackwell T (2007) Particle swarm optimization. Swarm Intell 1(1):33–57
28. http://www.pjm.com/markets-and-operations/energy/real-time/loadhryr.aspx
29. http://www.usclimatedata.com/climate/philadelphia/pennsylvania/united-states/uspa1276/2015/1