Genetic Algorithm and Particle Swarm Optimization for Parameter Optimization of Least-Square Support Vector Regression Model in Electricity Load Demand Forecasting

Irhamah$^{1,*}$, K H Gusti$^{2}$, H Kuswanto$^{3}$, N. A. Firdausanti$^{4}$

1,2,3,4Department of Statistics, Institut Teknologi Sepuluh Nopember Surabaya
Jl. Raya ITS, Kampus ITS Sukolilo, Surabaya, Indonesia 60111

Email: irhamah@statistika.its.ac.id$^{1,*}$

Abstract. Accuracy is very important in time series forecasting where the model obtained depends on historical, linear, or nonlinear data patterns. This study aims to forecast short term electricity load in East Java, Indonesia, whereas electricity load is one of major and vital need in our daily life. The linear approach is carried out using the ARIMA method, while the nonlinear approach used in this study consist of the SVR and LSSVR models. The parameter selection greatly affects the results of accuracy, so an optimization method is needed. The usual grid search optimization method does not guarantee an optimum solution and more importantly, it is not efficient in some practical applications. Therefore metaheuristic optimization methods are needed, including GA and PSO which can find the entire possible space for the search for solutions. PSO is easier to apply but is prone to have premature convergence due to trapped in the minimum locale. To overcome this, modified PSO is developed to seek the optimal position according to the specified criteria. The result of this research is that the accuracy of the nonlinear approach is much better than the linear approach. The addition of an optimization method to the SVR provides a significant change in accuracy compared to LSSVR. Meanwhile, MPSO is the best optimization method because it produces a lowest RMSE value.

1. Introduction

Accuracy is the most important thing in forecasting which is a fundament of decision making process. There are two kinds of time series forecasting methods, which are forecasting model based on statistics model and forecasting model based on artificial intelligence, such as SVR. The forecasting result of ARIMA method tends to be less good because it tends to be flat in the range of its average value [1]. Meanwhile, the SVR method produces more accurate forecasting results than other artificial intelligence-based methods because it can overcome overfitting [2]. The solution of the SVR method is computationally complex, so the LSSVR method is proposed which only requires solving one set of linear equations without eliminating the principle of SVR [3]. The results of the accuracy of the SVR and LSSVR methods depend on the selection of parameter values that can be done by trial and error but are not effective because there are too many combinations of numbers and experiments [4].

The traditional optimization method is grid search which works by traversing all points in the parameter range and is guided by a performance matrix [5]. However, grid search optimization does...
not guarantee a global optimal solution and is impractical in some applications [6]. Alternative optimization methods that can be used are metaheuristic methods that mimic social behavior or strategies that exist in nature to find effective and comprehensive solutions. Metaheuristic methods have a superior ability to avoid local optima due to the ability to search the entire search space extensively thereby avoiding stagnation of local solutions [7]. Genetic Algorithm is one of the metaheuristic methods which can be implemented in a wider variety of problems and solution space [8]. Another metaheuristic method is the PSO method which is easier to apply because it does not require many parameters to be specified [9]. However, PSO can be trapped at a local minimum or premature convergence. Therefore, MPSO adds a parameter of PSO that will seek to find the optimal particle position according to predetermined criteria and show better accuracy than PSO [10].

The case study in this study is a short-term electricity load that has the characteristics of a complicated series pattern due to an overlapping repetitive pattern generated by linear and nonlinear processes [11]. In this study, a comparison of the SVR and LSSVR methods along with the grid search, GA, PSO, and MPSO optimization methods were carried out so that short-term power load forecasting values were obtained. The best model was determined based on the minimum RMSE value.

2. Literature Review

2.1. Genetic Algorithm. The GA method is an optimization technique based on the principles of genetics and natural selection. The population is formed from many individuals who develop according to specific selection rules by maximizing fitness. This algorithm is done by looping Darwin's concept. The main stages of the GA procedure are (1) Defining variables as chromosomes, (2) Forming an initial population consisting of several chromosomes, (3) Evaluating the ability of chromosomes based on fitness, (4) Selecting chromosomes with maximum fitness as parents, (5) Combining genetic information in parents by crossing over, (6) Carrying out a process of mutation to introduce new genes randomly, (7) The process is repeated from stage 2 until it reaches convergence [12].

2.2. Particle Swarm Optimization. The PSO method was developed based on the behaviour of birds or fish in search of food sources. The direction of movement towards food is influenced by instinct, which in the algorithm is identical to that of exploration to obtain optimum results in different regions [13]. In the PSO algorithm, a set of m moving particles is influenced by (1) Inertia (2) Best local position (3) Best global position. In general, the motion of particles is formulated as follows:

$$v_{i}^{k+1} = v_{i}^{k} + c_{1}r_{1}(p_{best_{i}}-x_{i}^{k}) + c_{2}r_{2}(g_{best}-x_{i}^{k})$$  \hspace{1cm} (1)

$$x_{i}^{k+1} = x_{i}^{k} + v_{i}^{k+1}$$  \hspace{1cm} (2)

The greater the value of $\omega$, the greater the ability of PSO to find possible solutions in the search space, and vice versa. In general, if $\omega$ is one, then in the next iteration there is a partial lack of search capability. The position and velocity of the particles will be updated until convergence is reached [14].

2.3. Modified Particle Swarm Optimization. In complex problems, the PSO algorithm risks premature convergence which results in particles trapped in the local optimum solution. The MPSO algorithm adds a parameter of mean particle distance and operator mutation to guide population selection and evaluate convergence. The formulation of the average particle distance is written in equation (3) and the operator mutation capable of controlling the convergence speed of the PSO algorithm is formulated in equation (15).

$$D(t) = \frac{1}{mL} \sum_{\text{all}} \sum_{i=1}^{m} \sqrt{\sum_{j=1}^{n} (x_{ij} - x_{ij})}$$  \hspace{1cm} (3)
\[ \sigma^2 = \sum \left( \frac{f - \bar{f}}{f} \right)^2 \]

\[ f = \begin{cases} \max |f - \bar{f}|, & \max |f - \bar{f}| > 1 \\ 1, & \min |f - \bar{f}| \leq 1 \end{cases} \]  

(4)

With the increasing number of iterations, the values \( D(t) \) and \( \sigma^2 \) will get smaller. When \( D(t) < \alpha \) and \( \sigma^2 < \beta \) (\( \alpha, \beta \) are the specified threshold values), the algorithm is considered to be in the final search stage and the population will experience premature convergence. The MPSO algorithm maintains the optimal location of the current group of particles and recovers the particle solution space so that it can get out of the optimal locale [10].

3. Methodology

3.1. Data Sources

The data used in this study is secondary data obtained from PT PLN Distribution East Java regarding the demand for electricity loads in the districts and cities of East Java. Electricity load is recorded per 30 minutes from January 2016 to December 2016. The data is divided into two groups: training, and testing. The training data used to build the model is from January 2016 to November 2016 and testing data in 2016.

3.2. Data Structure

The data used in this study are short-term electricity loads recorded per 30 minutes at a 20kV substation (distribution section) in East Java. The research data is arranged in a structure which is presented in Table 1.

| Time  | Day  | Electricity Loads | Description |
|-------|------|-------------------|-------------|
| ...   | ...  | ...               | ...         |
| 16031 | 334  | \( Z_{1,16031} \) | Training    |
| 16032 | 334  | \( Z_{1,16032} \) | Training    |
| ...   | ...  | ...               | ...         |
| 17520 | 365  | \( Z_{1,17520} \) | Testing     |

3.3. Steps of Analysis

The analysis steps in this research are as follows:

- Obtain an electric load forecasting model based on the ARIMA, SVR, and LSSVR models
- Get the best optimization results on the SVR and LSSVR models with the GA, PSO, and MPSO optimization methods. The selection of the best model is done by evaluation using RMSE and SMAPE

4. Result and Discussion

4.1. Characteristics of Electricity Load

The characteristics of the demand for electricity in the East Java region are obtained through descriptive statistical analysis by exploring the information contained in the data. Figure 1(a) shows that the data is non-stationary in mean. The consumption pattern of public electricity loads is presented in Figure 1(b). It can be seen that the lowest average consumption of electricity load occurs at 07:00 in line with the commencement of community activities to work outside the home. The average
consumption of electric loads reaches the highest peak at 18:30 as the end of industrial activities and work outside the home. After that, the average electricity load gradually drops along with the entry of public rest hours and agencies that do not operate at night. Based on Figure 1 (b) it is known that during the day, the consumption of electric loads has high variance. This is related to the consumption pattern which is dominated by the industrial sector.

The characteristics of the seasonal pattern from the data are shown in Figure 2 (a) which shows the amount of electricity consumption from January 1 to March 31, 2016. This illustrates the weekly seasonal pattern where on Saturdays and Sundays the consumption of electricity loads is lower than during the active day because it was contributed by industrial consumption which was operating on that day. Meanwhile, Figure 2 (b) shows the amount of electricity consumption from 1 January to 7 January 2016, which illustrates a pattern of low electricity consumption at night to early morning and then an increase in the morning to evening. This phenomenon shows the presumption of a seasonal daily pattern in the data. Based on the information above, it is estimated that there are daily variations (per 48 hours) and weekly variations (per 336 hours).

4.2. ARIMA Estimation
The stationary in mean can be identified by ACF and PACF. The data is not stationary in mean because the ACF plot decays very slowly, so it is necessary to do differencing on 1st lag. However, ACF plot after differencing still shows repeating patterns at multiples of 48, so it is necessary to do differencing at 48th lag. After the second differencing, the data is still not stationary. The high significance of the PACF lag at the multiple of 336 indicates that the data needs to be differencing at 336th lag. Therefore, the possible models are as follows:

\[ ARIMA(28,31,47), [1, [11,12,31]])(0,1,1)^{48}(0,1,2)^{336} \]
\[ ARIMA(28,47), [1, [11,12]])(0,1,1)^{48}(0,1,2)^{336} \]
The parameter estimation is shown in Table 2. Both models have significant parameters so that forecasting was done and compared with testing data. This procedure resulted in RMSE values of 618.1697 and 646.465 with sMAPE values of 13.76 and 14.48 for model 1 and model 2, respectively. The best model is the 1st model that can be written as follows.

\[
\left(1 - \phi_1 B^{28} - \phi_2 B^{31} - \phi_3 B^{47} - \phi_4 B^{48}\right) \left(1 - B\right) \left(1 - B^{11}\right) \left(1 - B^{336}\right) \left(1 - B^{672}\right) z_t
\]

\[
= \left(1 - \theta_1 B^{11} - \theta_2 B^{12} - \theta_3 B^{13}\right) \left(1 - \Theta B^{24}\right) \left(1 - \Theta B^{108}\right) \left(1 - \Theta B^{602}\right) a_t
\]

| Table 2. Parameter Estimation of ARIMA Models |
|---------------------------------------------|
| Model 1 | Est | p-value | Model 2 | Est | p-value |
|         |     |         |         |     |         |
| \( \phi_{11} \) | -0.0519 | <.0001 | -0.0525 | <.0001 |
| \( \phi_{12} \) | -0.0292 | 0.0002 | -0.0274 | 0.0006 |
| \( \phi_{13} \) | -0.2297 | 0.0012 | -        | -        |
| \( \theta_{11} \) | 0.7072 | <.0001 | 0.7057 | <.0001 |
| \( \theta_{12} \) | 0.7862 | <.0001 | 0.78644 | <.0001 |
| \( \phi_{18} \) | -0.0311 | 0.0002 | -0.03102 | 0.0002 |
| \( \phi_{19} \) | -0.0199 | 0.0103 | -0.0192 | 0.0167 |
| \( \phi_{21} \) | -0.2403 | 0.0006 | -        | -        |
| \( \phi_{28} \) | 0.0783 | <.0001 | 0.08006 | <.0001 |

4.3. Parameter Optimization using Grid-search

Grid-Search optimization is done by describing the ARIMA model and then a significant lag is obtained used as an exogenous variable. The significant lag focuses on the AR model so that there are 111th data lag. Two input parameters, which are the cost parameter \( (C) \) and the RBF kernel function parameter \( (\sigma^2) \) for the LSSVR method and three parameters for the SVR method, which are cost \( (C) \), epsilon \( (\varepsilon) \) and RBF kernel function parameters \( (\sigma^2) \) were optimized using the Grid-Search. The parameter ranges for LSSVR used are \{50,200\}, \{1, 25.75, 50.5, 75.25, 100\} for \( \sigma^2 \) and \( C \), while the SVR is \{\((1 e - 2.1), (1 e -4.10), (1 e -2.50)\)\} for \( \varepsilon \), \( C \), and \( \sigma^2 \). Furthermore, the model obtained was used to model the next 1488 and a half hours. The accuracy criteria used are the minimum RMSE and the optimal parameters obtained for \( \sigma^2 \) and \( c \) are 50 and 25.75, and the SVR method results in optimal parameters for \( \varepsilon \), \( C \), and \( \sigma^2 \) are 0.01, 10, and 0.01, respectively. The RMSE values obtained for LSSVR and SVR were 51.114845812 and 58.11766134.

4.4. Parameter Optimization using Genetic Algorithm

Optimization with GA is done by making the parameters from the Grid-Search as a good initial value. The procedure was carried out by generating 15 chromosomes for 5 generations with a cross over the chance of 0.8, the chance for mutation of 0.1, the number of chromosomes that survive in the next generation, namely 1, the solution space for \( c = \{49,250\} \) and \( \sigma^2 = \{25,100\} \). The selection of parents uses the RWS method, the combination of genes uses a single point crossover process, the mutation process uses a uniform mutation that replaces one of the genes with a uniformly distributed random number (0.1), the fitness value used is RMSE. Based on the optimum fitness, the optimal parameters are 197.2273 and 43.88546 for \( \sigma^2 \) and \( C \) with an RMSE value of 43.25871.

In the SVR method, the procedure performed for GA optimization is the same as for the LSSVR method. The size of genes in chromosomes in the SVR method is under many optimized parameters, namely cost \( (C) \), epsilon \( (\varepsilon) \), and RBF kernel function parameters \( (\sigma^2) \) so that there are 3 genes. The fitness value used in the SVR is -RMSE. Based on the optimum fitness, the optimal parameter is 0.00839; 11.29062; 0.001119 for \( \varepsilon \), \( C \), and \( \sigma^2 \) with an RMSE value of 42.901.

4.5. Parameter Optimization using Particle Swarm Optimization
Optimization with PSO is done by making the parameters from the Grid-Search as the good initial value. The procedure is carried out by generating 15 particles with an inertia value of 0.9, the coefficient of each particle and the overall particle of 2, the maximum speed limit of 2% of the upper limit of each particle position, the fitness used is RMSE and the convergence criterion by limited by 5 iterations. Based on the optimum fitness, the optimal parameters are 196,1141 and 40,597 for $\sigma^2$ and $C$ with an RMSE value of 41.87447.

In the SVR method, the procedure performed for PSO optimization is the same as for the LSSVR method. The particle size in the SVR method corresponds to many optimized parameters, which are cost ($C$), epsilon ($\epsilon$), and the RBF kernel function parameter ($\sigma^2$). The fitness value used in the SVR is RMSE. Based on the optimum fitness, the optimal parameter is 0.005567992; 9,481282; 0.002875503 for $\epsilon$, $C$, and $\sigma^2$, with an RMSE value of 43.6688.

4.6. Parameter Optimization using Modified Particle Swarm Optimization
Optimization using MPSO is done by making the parameters of the Grid-Search as the good initial value. The procedure is carried out by determining the same parameter values as PSO, plus the MPSO parameters, which are alpha and beta as the algorithm guides to retrieve the optimal position and diagonal length of the search space. The values assigned to each alpha, beta, and diagonal length are 0.5, 0.05, and 100. Based on the optimum fitness, the optimal parameters are 188.0168 and 47.67838 for $\sigma^2$ and $C$ with RMSE values of 45.403822.

In the SVR method, the procedure carried out for MPSO optimization is the same as for the LSSVR method, including the values of alfa, beta, and diagonal length. The particle size in the SVR method corresponds to many optimized parameters, which are cost ($C$), epsilon ($\epsilon$), and the RBF kernel function parameter ($\sigma^2$) so that there are 3 sizes. The fitness value used in the SVR is RMSE. Based on the optimum fitness, the optimal parameter is 0.0001051; 8,3459835649; 0.0044634277 for $\epsilon$, $C$, and $\sigma^2$, with an RMSE value of 45.403822.

4.7. The Comparison of the Optimization Algorithms
The performance comparison of each optimization algorithm were measured based on the RMSE values shown as in Table 3.

| Metod          | RMSE  | Metod          | RMSE  | rank | Metode  | RMSE  | rank |
|---------------|-------|---------------|-------|------|---------|-------|------|
| DSARIMA       | 618,1697 | LSSVR        | 42,02479 | 4       | SVR     | 48,7939 | 4     |
|               |       | LSSVR-GA     | 41,86809 | 2       | SVR-GA  | 42,4174 | 3     |
|               |       | LSSVr-PSO    | 41,87447 | 3       | SVR-PSO | 41,4365 | 2     |
|               |       | LSSVR-MPSO   | 41,85998 | 1       | SVR-MPSO | 40,1367 | 1     |

It can be seen that the LSSVR method the best optimization method is the GA algorithm. In contrast to [16] in this study, MPSO did not improve the accuracy of PSO, this was possible because of an error in the selection of the guiding parameters. Overall, the accuracy of the SVR method is better than the LSSVR when combined with the optimization method, this can be seen from the high difference in RMSE values, while LSSVR has a not too significant difference in accuracy between optimization methods. When compared to DSARIMA, artificial intelligence-based methods have better accuracy because of the complex patterns in the short-term time series cannot be accommodated by the linear model.

5. Conclusions and Suggestions

5.1. Conclusions
Forecasting using the LSSVR method is carried out based on the results of the ARIMA model first. Short-term electricity load data for the East Java area is known to have a Double Seasonal model

Table 3. The Performance of Each Model

| Metoden | RMSE  | Metoden | RMSE  | rank | Metoden | RMSE  | rank |
|---------|-------|---------|-------|------|---------|-------|------|
| DSARIMA | 618,1697 | LSSVR   | 42,02479 | 4    | SVR     | 48,7939 | 4    |
|         |       | LSSVR-GA | 41,86809 | 2    | SVR-GA  | 42,4174 | 3    |
|         |       | LSSVR-PSO | 41,87447 | 3    | SVR-PSO | 41,4365 | 2    |
|         |       | LSSVR-MPSO | 41,85998 | 1    | SVR-MPSO | 40,1367 | 1    |
ARIMA ([28,31,47], 1, [11,12,32]) (0,1,1) 48 (0,1,2) that has 111 significant lags. Forecasting using the LSSVR method produces a much better accuracy of forecast results than ARIMA.

Overall, the optimization method that has the best accuracy is MPSO, GA, then PSO, respectively. This is because the MPSO optimization results are ranked first in each model. As for the results of the accuracy in the comparison of the SVR and LSSVR methods, it is concluded that the SVR method has the best accuracy when combined with the optimization method, while in the LSSVR method the addition of the optimization method does not give a significant change in the accuracy results.

Acknowledgement
The authors greatly appreciate Ministry of Research, Technology, and Higher Education, Republic of Indonesia and Institut Teknologi Sepuluh Nopember Surabaya Indonesia for financial support under “Penelitian Tesis Magister”.

References
[1] D. Wiyanti and R. Pulungan, "Peramalan Deret Waktu Menggunakan Model Fungsi Basis Radial (RBF) dan Auto Regressive Integrated Moving Average (ARIMA)," Jurnal MIPA, vol. 35, no. 2, pp. 175-182, 2012.
[2] M. Amin and M. Hoque, "Comparison of ARIMA and SVM for Short-Term Load Forecasting," pp. 205-210, 2019.
[3] J. Suykens and J. Vandewalle, "Least Square Support Vector Machine Classifiers," Neural Processing Letters, vol. 9, pp. 293-300, 1999.
[4] S. Prangga, "Optimasi Parameter pada Support Vector Machine Menggunakan Metode Taguchi untuk Data High-Dimensional," Institut Teknologi Sepuluh Nopember, Surabaya, 2017.
[5] L. Septiningrum, H. Yasin and Sugito, "Prediksi Indeks Harga Saham Gabungan Menggunakan Support Vector Regression (SVR) dengan Algoritma Grid Search," vol. 4, no. 2, pp. 315-321, 2015.
[6] P. Cortez, "Modern Optimization with R," Gurmaraes.
[7] A. Pangestu, B. Arie and E. Jovi, "Optimizing Neural Network Classifier for Diabetes Data Using Metaheuristic Algorithms," ITSMART : Jurnal Ilmiah Teknologi dan Informasi, vol. 6, no. 2, pp. 85-91, 2017.
[8] S. Sivandam and S. Deepa, Introduction to Genetic Algorithm, Berlin, Heidelberg: Springer, 2008.
[9] W. Hu, L. Yan, K. Liu and H. Wang, "PSO-SVR : A Hybrid Short-Term Traffic Flow Forecasting Method," International Conference on Parallel and Distributed Systems, pp. 553-561, 2015.
[10] D. Niu and S. Dai, "A Short-Term Load Forecasting Model with A Modified Particle Swarm Optimization Algorithm and Least Square Support Vector Machine Based on the Denoising Method of Empirical Model Decomposition and Grey Relational Analysis," Energies, vol. 10, no. 408, pp. 1-20, 2017.
[11] F. Fahmi and H. Sofyan, "2017," International Conference on Electrical Engineering and Informatics (ICELTICs) , pp. 97-102, 2017.
[12] Z. Ismail and Irhamah, "Solving the Vehicle Routing Problem with Stochastic Demands via Hybrid Genetic Algorithm-Tabu Search," Journal of Mathematics and Statistics, vol. 4, no. 3, pp. 161-167, 2008.
[13] Ghalia and B. Mounir, "Particle Swarm Optimization with An Improved Exploration-Exploitation Balance," Midwest Symposium on Circuit and Systems, pp. 759-762, 2008.
[14] Q. Bai, "Analysis of Particle Swarm Optimization Algorithm," Computer and Information Science, vol. 3, no. 1, pp. 180-184, 2010.