

# Evaluation Performance Of Hybrid Localized Multi Kernel SVR in Electrical Load Data Using 4 Different Optimization

Rezzy Eko Caraka, Sakhinah Abu Bakar

*Abstract*— The main problem using SVR is to find optimal parameter ( $\sigma$ ) by using kernel function such as radial basis, polynomial, gaussian and so on. Moreover, we also have to find optimal hyperplane parameter ( $C$  and  $\epsilon$ ). In the heart of statistics methods and data mining, the motivation of Researcher doing this is to minimize time, money and energy in the analysis at the same time the results will be more accurate. The development of such a massive technology and the availability of data is very much making progress and improvement of methods based on data mining. In this paper, we proposed 3 different optimization such as LIBSVM, MOSEK, QUADPROG, SMO applied to Localized Kernel Learning (LKL).

*Keywords*— SVR, Kernel, Optimization,

## I. INTRODUCTION

Electricity is one of the essential human needs in everyday life because almost all equipment used is a supporter of human activities using electricity as power. Today the demand for electricity is increasing significantly day. This increase is influenced by several factors such as population growth, increased use of electrical equipment, technological developments, increased welfare or public purchasing power, and various other factors. Therefore, electric power service providers must have availability of supply and reliability of electrical energy in their distribution system to meet the electricity demand for their customers. Various studies have been conducted on electricity load forecasting. Electricity load forecasting has been a field of research for more than 50 years on operating systems and electrical planning (Srivastava et al. 2016). Forecasting is done to get the accuracy of electricity load forecasting for the short term, medium term, and long-term. In general, electric load forecasting methods can be divided into two parts, namely: statistical methods of classical mathematics and approaches based on artificial intelligence.

The authors are with the Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia

In statistical methods, researchers using methods linear regression method, multiple regression, Autoregressive Moving Average Model (ARMA), Kalman Filter Technique, and Box-Jenkins Model to analysis load forecasting. The method has a weakness regarding ability that is not linear. With the Artificial Intelligence approach, such as Neural Network, Support Vector Machine, the researchers are concerned to get higher and efficient electricity load forecasting accuracy (Kavousi-Fard & Kavousi-Fard 2013)

Although different in model structure and forecasting, the two categories of methods rely on historical data of electrical loads and other exogenous factors. Nowadays it has been widely used a combination of statistical methods and artificial intelligence methods for short-term electrical load forecasting. Short-term electrical load forecasting is the prediction of load demand in the time intervals for one day to several weeks ahead (Caraka 2017) which forecast has a fundamental importance in terms of plant scheduling, fuel purchase plans, security capacity, short-term maintenance as well as for the use of short-term storage. Therefore, forecasting electrical loads with high accuracy is very important for companies providing electricity services to meet the needs of electricity demand customers (Caraka, Bakar, et al. 2017). Excessive predictions or forecasting of electrical loads will result in unnecessary use of power plants, increased operating costs. As well as forecasting of electrical loads that do not meet the electrical load requirements may cause system reliability due to insufficient power resources to meet security or cost requirements become higher due to the purchase of expensive electricity to meet the demand for electricity (Hong 2010)

## II. OPTIMIZATION

### 2.1 LIBSVM

LIBSVM is an algorithm used in SVM techniques. Its application is on the classification, regression and distribution estimates. LIBSVM is also known as a library of Support Vector Machines (SVMs). (Chang & Lin 2013) actively developing this package since the year 2000. The goal is to help users to easily apply SVM to their applications. As we know that the linear conditions are very difficult to apply to dynamic data therefore need non-linear techniques to perform the analysis. Suppose there is data point  $(x_i, d_i)$  following the equation (Qi 2013)

$$d_i(\mathbf{w}^T x_i + b) \geq +1 \text{ with the provision of } i = 1, 2, \dots, N$$

The researcher did the set of non negative slack variables,  $\{\xi_i\}_{i=1}^N$  to hyperplane, then:

$d_i(w^T x_i + b) \geq +1 = \xi_i$ , with the provision of  $i = 1, 2, \dots, N$

**2.2 MOSEK**

In the field of optimization, MOSEK is a software package used to solve conic and convex. MOSEK can also be used to solve primal and dual simplex for linear problems (LO), Quadratic (QCQO), Conic Quadratic (CQO, SOCO), Semidefinit (SDO), Separable Convex (SCOPT), and General Convex. MOSEK also capable of solving smooth convex nonlinear optimization of the form. Moreover. MOSEK also capable of solving smooth convex nonlinear optimization of the form

$$\begin{aligned} & \text{minimize } f(x) + c^T x \\ & \text{subject to } g(x) + Ax - x^c = 0 \\ & \quad l^c \leq x^c \leq u^c, \\ & \quad l^x \leq x \leq u^x, \end{aligned}$$

Where m is the number of constraint, n is the number of decision variables.  $x \in \mathbb{R}^n$  is a vector of decision variables,  $x^c \in \mathbb{R}^m$  is a vector of constraints or slack variables,  $c \in \mathbb{R}^n$  is the vector of decision variables,  $A \in \mathbb{R}^{m \times n}$  is the constraint matrix,  $l^c \in \mathbb{R}^m$  is the lower limit on the activity for the constraints,  $u^c \in \mathbb{R}^m$  is the upper limit on the activity for the constraints,  $l^x \in \mathbb{R}^n$  is the lower limit on the activity for the variables,  $u^x \in \mathbb{R}^n$  is the upper limit on the activity for the variables,  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  is a nonlinear function,  $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$  is a nonlinear vector function. In other words, it means that he ith constraint has the form

$$l_i^c \leq g_i(x) + \sum_{j=1}^n a_{ij} x_j \leq u_i^c$$

when the  $x_i^c$  variable has been eliminated

**2.3 QUADPROG**

QUADPROG is one of special program in mathematics also statistics can be formulated as

$$f(x) = \frac{1}{2} x^T Q x + c^T x$$

Assume x exists in space  $\mathbb{R}^n$ , Q is symmetric matrix nxn and c is any nx1 vector

1. Minimize against x
2. With the limits  $Ax \preceq b$  and  $Ex = d$

If Q is a mat 0 (zero), then the problem will turn into linear programming. From the optimization theory, the requirement for x to be Global minimizer is x must satisfy the condition of Karush-Kuhn-Tucker (KKT) (Yasin et al. 2016) and (Caraka, Yasin, et al. 2017). Convex quadratic programming is a particular case of convex optimization. This method consists of two stages, first intended to convert into a convex quadratic function. Where to change the function of this goal produces a new formula that has tractable continuous relaxation, which can be placed on branch-and-bound logarithms.

$$\begin{aligned} & \sum_{t=1}^T \sum_{p=1}^P q_{tp} x_{tp} + \sum_{t=1}^{T-1} \sum_{t'=t+1}^T \sum_{p=1}^P \sum_{p'=1}^P c_{tp t' p'} X_{tp t' p'} \\ & \sum_{p=1}^P X_{tp t' p'} = X_{tp} \quad t = 1, \dots, T ; t' = 1, \dots, T ; p = 1, \dots, P \\ & X_{tp tp} = X_{tp} \quad t = 1, \dots, T \quad p = 1, \dots, P \end{aligned}$$

X is a symmetric matrix (T x P) and can be considered as part of the relaxation of semidefinite programming. Any optimal solution of the dual problem will result in a convex

formulation of the objective function. We can assume that  $(\alpha^*, u^*)$  is the optimal solution for dual. We can produce a 0-1 quadratic formulation as follows

$$\begin{aligned} F^{\alpha^* u^*}(x) = & F(x) + \sum_{t=1}^T \sum_{t'=1}^T \sum_{p=1}^P \alpha^*_{tp t' p'} x_{tp} \left( \sum_{p=1}^P x_{t' p'} - 1 \right) \\ & + \sum_{t=1}^T \sum_{p=1}^P u^*_{tp} (x_{tp}^2 - x_{tp}) \end{aligned}$$

With limit

$$\sum_{p=1}^P x_{tp} = 1 \quad t = 1, \dots, T \quad x \in \{0, 1\}^{T \times P}$$

**2.4 Sequential Minimal Optimization (SMO)**

One method that can be used to optimize hyperplane, that is to solve quadratic programming with defined constraints is Sequential Minimal Optimization (SMO). Sequential Minimal Optimization (SMO) is an algorithm to solve Quadratic Programming (QP) problems that arise during training in the Support Vector Machine. SMO is a simple algorithm that can solve QP problems in SVM quickly. According to Karatzoglou et al. (2006), the SMO algorithm solves the QP problem in SVM without using numerical QP optimization steps. Instead, SMO chooses to resolve the smallest possible optimization problem by involving two  $\alpha_i$  elements due to the necessity to satisfy a linear divider equation. Each step, SMO selects two  $\alpha_i$  to be optimized together and finds the optimal values for the  $\alpha_i$  values analytically, thereby avoiding numerical QP optimization, and updating SVM to provide the latest optimum values.

Stages of hyperplane optimization in SVM include: Complete Two Lagrange Multipliers and Counting b. Stages of optimization of SVM hyperplane can be described, as follows:

1. Complete Two Lagrange Multipliers

QP problem in SVM training is stated, as follows:

$$\begin{aligned} & \max_{\alpha} W(\alpha) = \\ & \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l y_i y_j K(\mathbf{x}_i \mathbf{x}_j) \alpha_i \alpha_j \\ & 0 \leq \alpha_i \leq C, \quad \forall i \\ & \sum_{i=1}^l \alpha_i y_i = 0 \end{aligned}$$

The QP problem is solved with the SMO algorithm. A point on the equation is the optimum point when the Karush-Kuhn-Tucker (KKT) condition is satisfied and  $y_i y_j K(\mathbf{x}_i \mathbf{x}_j)$  is positive semi-definite. QP problems are met if all i:

$$\begin{aligned} \alpha_i = 0 & \rightarrow y_i f(x_i) \geq 1 \\ 0 < \alpha_i < C & \rightarrow y_i f(x_i) = 1 \\ \alpha_i = C & \rightarrow y_i f(x_i) \leq 1 \end{aligned}$$

With  $f(x_i)$  is the predicted value of  $x_i$  data. The SMO algorithm will be lit up until the condition is met. For QP the condition of HCFs is a sufficient condition as well as a necessary condition.

To complete the two Lagrange multipliers, SMO first calculates the delimiter on the multipliers. If the target  $y_i$  is not the same as the  $y_j$  target, then the following limits apply to  $\alpha_j$ :

$$L = \max\{0, \alpha_j - \alpha_i\}, H = \min\{C, C + \alpha_j - \alpha_i\}$$

If the target  $y_i$  is the same as the  $y_j$  target, then the following limits apply to  $\alpha_j$ :

$$L = \max\{0, \alpha_j + \alpha_i - C\}, H = \min\{C, \alpha_j + \alpha_i\}$$

We will look for the value of  $\alpha_j$  to maximize the objective function  $W(\alpha)$ . If the value falls outside the boundary  $L$  and  $H$ , the value of  $\alpha_j$  will be cut to fit within that range. Optimal  $\alpha_j$  can be formulated, as follows:

$$\alpha_j^{\text{new}} = \alpha_j + \frac{y_j(E_i - E_j)}{\eta}$$

with  $E_i = f(x_i) - y_i$  or error training to-i

$$E_j = f(x_j) - y_j \text{ or error training to-j}$$

$$\eta = K(\mathbf{x}_i, \mathbf{x}_j) + K(\mathbf{x}_j, \mathbf{x}_i) - 2K(\mathbf{x}_i, \mathbf{x}_j)$$

Under normal conditions,  $\eta$  will be worth less than 0. In the next step, value  $\alpha_j$  clip to within range  $[L, H]$  :

$$\alpha_j^{\text{new,clipped}} = \begin{cases} H, & \text{if } \alpha_j^{\text{new}} \geq H \\ \alpha_j^{\text{new}}, & \text{if } L < \alpha_j^{\text{new}} < H \\ L, & \text{if } \alpha_j^{\text{new}} \leq L \end{cases}$$

After finding the solution for  $\alpha_j$ , searchable value of  $\alpha_i$ , as follows:

$$\alpha_i^{\text{new}} = \alpha_i + s(\alpha_j - \alpha_j^{\text{new,clipped}})$$

with  $s = y_i y_j$

According to Platt (1998), beyond normal conditions,  $\eta$  will not be negative.  $\eta$  worth of zero can occur if more than one training instance has the same input as vector  $x$ . On some occasions, SMO will keep working even when  $\eta$  is not negative, in this case, the objective function  $\Psi$  which must be evaluated at each end of the line segment.

$$f_i = y_i(E_i + b) - \alpha_i K(\mathbf{x}_i, \mathbf{x}_i) - \alpha_j K(\mathbf{x}_i, \mathbf{x}_j)$$

$$f_j = y_j(E_j + b) - \alpha_i K(\mathbf{x}_i, \mathbf{x}_j) - \alpha_j K(\mathbf{x}_j, \mathbf{x}_j)$$

$$L_i = \alpha_i + s(\alpha_j - L)$$

$$H_i = \alpha_i + s(\alpha_j - H)$$

$$\Psi_L = L_i f_i + L f_j + \frac{1}{2} L_i^2 K(\mathbf{x}_i, \mathbf{x}_i) + \frac{1}{2} L^2 K(\mathbf{x}_j, \mathbf{x}_j) + s L L_i K(\mathbf{x}_i, \mathbf{x}_j)$$

$$\Psi_H = H_i f_i + H f_j + \frac{1}{2} H_i^2 K(\mathbf{x}_i, \mathbf{x}_i) + \frac{1}{2} H^2 K(\mathbf{x}_j, \mathbf{x}_j) + s H H_i K(\mathbf{x}_i, \mathbf{x}_j)$$

## 2. Calculate b

The value of  $b$  is always recalculated at each end of the stage so that the conditions of KKT are met for both optimization problems. The following  $b_i$  will be valid if the new  $\alpha_i$  is not on the limit:

$$b_i = E_i + y_i(\alpha_i^{\text{new}} - \alpha_i)K(\mathbf{x}_i, \mathbf{x}_i) + y_j(\alpha_j^{\text{new,clipped}} - \alpha_j)K(\mathbf{x}_i, \mathbf{x}_j) + b$$

$b_j$  will be valid if new  $\alpha_j$  not on the limit:

$$b_j = E_j + y_j(\alpha_j^{\text{new}} - \alpha_j)K(\mathbf{x}_j, \mathbf{x}_j) + y_i(\alpha_i^{\text{new,clipped}} - \alpha_i)K(\mathbf{x}_j, \mathbf{x}_i) + b$$

When both  $b_i$  and  $b_j$  are valid, they are the same. If the two new Lagrange multipliers are at the limit, and if  $L$  is not equal to  $H$  then the interval between  $b_i$  and  $b_j$ , all is a threshold consistent with the terms of the KKT. SMO selects the middle value between  $b_i$  and  $b_j$

## III. APPLICATION OF LOCALIZED MULTIPLE KERNEL

Kernel machines learn a decision function in terms of kernel values between training instances,  $\{\mathbf{x}_i\}_{i=1}^N$  dan test instance  $\mathbf{x}$  as follows:

$$f(\mathbf{x}) = \sum_{i=1}^N \alpha_i k(\mathbf{x}_i, \mathbf{x}) + b$$

Where researchers can use various types of kernels such as linear, polynomial, radial base, Gaussian, and others. In fact, researchers can use multi-Kernel by converting the equations into

$$k_\eta(\mathbf{x}_i, \mathbf{x}_j) = f_\eta(\{k_m(\mathbf{x}_i^m, \mathbf{x}_j^m)\}_{m=1}^P)$$

According to this, we can use combination function of  $f_\eta$  with linear or nonlinear kernel. Feature representations with kernel function is  $\{k_m(\cdot, \cdot)\}_{m=1}^P$ .  $P$  is a representative feature of the data. Where  $\mathbf{x}_i = \{\mathbf{x}_i^m\}_{m=1}^P$ ,  $\mathbf{x}_i^m \in \mathbb{R}^{D_m}$ . Also,  $D_m$  is a dimension that describes the feature

Based on the above equation, we can get an example for a combination of more than two kernels

$$k_\eta(\mathbf{x}_i, \mathbf{x}_j) = \sum_{m=1}^P k_m(\mathbf{x}_i^m, \mathbf{x}_j^m)$$

$$k_\eta(\mathbf{x}_i, \mathbf{x}_j) = \prod_{m=1}^P k_m(\mathbf{x}_i^m, \mathbf{x}_j^m)$$

The equation is also mentioned as pairwise kernel so it is also obtained:

$$\begin{aligned} k(\{\mathbf{x}_i^a, \mathbf{x}_j^a\}, \{\mathbf{x}_i^b, \mathbf{x}_j^b\}) \\ = k(\mathbf{x}_i^a, \mathbf{x}_i^b), k(\mathbf{x}_j^a, \mathbf{x}_j^b) \\ + k(\mathbf{x}_i^a, \mathbf{x}_j^b), k(\mathbf{x}_j^a, \mathbf{x}_i^b) \end{aligned}$$

or can be written:

$$k_\eta(\{\mathbf{x}_i^a, \mathbf{x}_j^a\}, \{\mathbf{x}_i^b, \mathbf{x}_j^b\}) = \sum_{m=1}^P k_m(\{\mathbf{x}_i^a, \mathbf{x}_j^a\}, \{\mathbf{x}_i^b, \mathbf{x}_j^b\})$$

$$k_\eta(\{\mathbf{x}_i^a, \mathbf{x}_j^a\}, \{\mathbf{x}_i^b, \mathbf{x}_j^b\}) = k_\eta(\mathbf{x}_i^a, \mathbf{x}_i^b), k_\eta(\mathbf{x}_j^a, \mathbf{x}_j^b) + k_\eta(\mathbf{x}_i^a, \mathbf{x}_j^b), k_\eta(\mathbf{x}_j^a, \mathbf{x}_i^b)$$

Besides using  $f_\eta(\cdot)$  we can also use function for parameterization by using parameter  $\Theta$ . The simplest way to parameterize is to use weights in the kernel function  $\{\eta_m\}_{m=1}^P$ . Kernel matrix from  $\omega(\mathbf{K})$  can be explained as to complete canonical SVM optimization using soft margin and  $\ell_1$  - norm is a slack variables. So that :

$$\text{maximize } \omega(\mathbf{K}) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j k(\mathbf{x}_i, \mathbf{x}_j)$$

$$\text{with respect to } \boldsymbol{\alpha} \in \mathbb{R}_+^N$$

$$\text{subject to } \sum_{i=1}^N \alpha_i y_i = 0$$

$$C \geq \alpha_i \geq 0 \quad \forall_i$$

The combinations of the kernel matrix can be exposed as  $K_L = \{\mathbf{K}; \mathbf{K} = \sum_{m=1}^P \eta_m \mathbf{K}_m, \mathbf{K} \succeq 0, \text{tr}(\mathbf{K}) \leq c\}$

The resulting optimization problem that minimizes the implausibility of the combined

kernel matrix (the objective function value of the corresponding soft margin SVM

$$\text{Minimize } \omega(\mathbf{K}_\eta^{\text{tra}})$$

$$\text{With respect to } K_\eta \in K_L$$

$$\text{Subject to } \text{tr}(\mathbf{K}_\eta) = c$$

Where  $K_\eta^{tra}$  is a kernel matrix that is calculated only if the training data on the problem of researchers follow semidefinite programming (SDP) with the following equation:

Minimize  $t$

$$\text{With respect to } \eta \in \mathbb{R}^p, t \in \mathbb{R}, \lambda \in \mathbb{R}, v \in \mathbb{R}_+^N, \delta \in \mathbb{R}_+^N$$

$$\text{Subject to } \text{tr}(\mathbf{K}_\eta) = c$$

The main problem since the width of  $\epsilon - margin$  holds the ability to affect the complexity and the generalization of the regression function indirectly. Also, it is important to seek another optimal value for different applications.

#### IV. APPLICATION OF GUI MATLAB HYBRID LMK-SVR

In this simulation we have the electrical load data which is divided on Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, Sunday, Ramadhan, and National Holiday. Moreover, this data cover P3BS areas UPT Banda Aceh, UPT Medan, UPT Pematang Siantar, UPT Padang, UPT Palembang, UPT Tanjung Karang, UPT North Sumatra and UPT South Sumatra. Based on figure 1 it can be seen is a two-dimensional plot of electrical loads of the Northern Sumatera (SBU) system from 2012 to 2015. From the plot it can be seen that there is a seasonal pattern, which is a pattern that shows repeated changes periodically in the annual series. This plot provides information that increasing the electrical load each year indicates that the electricity consumption level of customers increases for each year.

Based on the analysis we can classify the electrical data in northern Sumatra as follows:

##### 1. For business days:

A. Around 04.00 am there was an increase in electrical load with a peak load at 06.30 am, and this is because customers start to use electricity for prayer, cooking breakfast and preparing for work. After 6:30 am, the electricity load gradually decreased until 08.00 am. It happens because at that time the sun has risen and the lights have been extinguished.

b. After 08.00 am, the load of electricity starts to increase, and this happens because the community began to move, office activities and industry began to operate. During the day between 12.00 and 13.00 noon, the electrical load has decreased slightly because at this time it is the worker's break for prayer and lunch.

c. In the afternoon between 1800 hours the load has increased dramatically, this happens because people who are at home start to turn on electrical appliances such as lights for lighting,

television for entertainment, and other electrical equipment used simultaneously. The peak load occurs at this time, around 7:30 pm. Starting at 21.30 people start a lot that reduces its activity, turn off the TV, reduce lights and sleep.

##### 2. Saturday

On Saturday, the hourly electrical charge chart is the same as the electricity load chart on weekdays, but the consumed power load is lower than the electrical load on weekdays. This happens because of a company that does not work on Saturday.

##### 2. Monday, National Holiday

Expenses on Sundays differ from 08.00 am to 17.00 pm where the electricity load is very low compared to weekdays and Saturdays. The pattern of the daytime ordinary holiday loads and special holidays for Northern Sumatra from 2012 s.d. 2015 on figure 1 provides information that special holidays such as Ramadhan, New Year, etc. At the same hour, the electrical load is lower than the usual holiday day (Saturday and Sunday). This is due to the particular holiday industry activities, and the company does not exist.

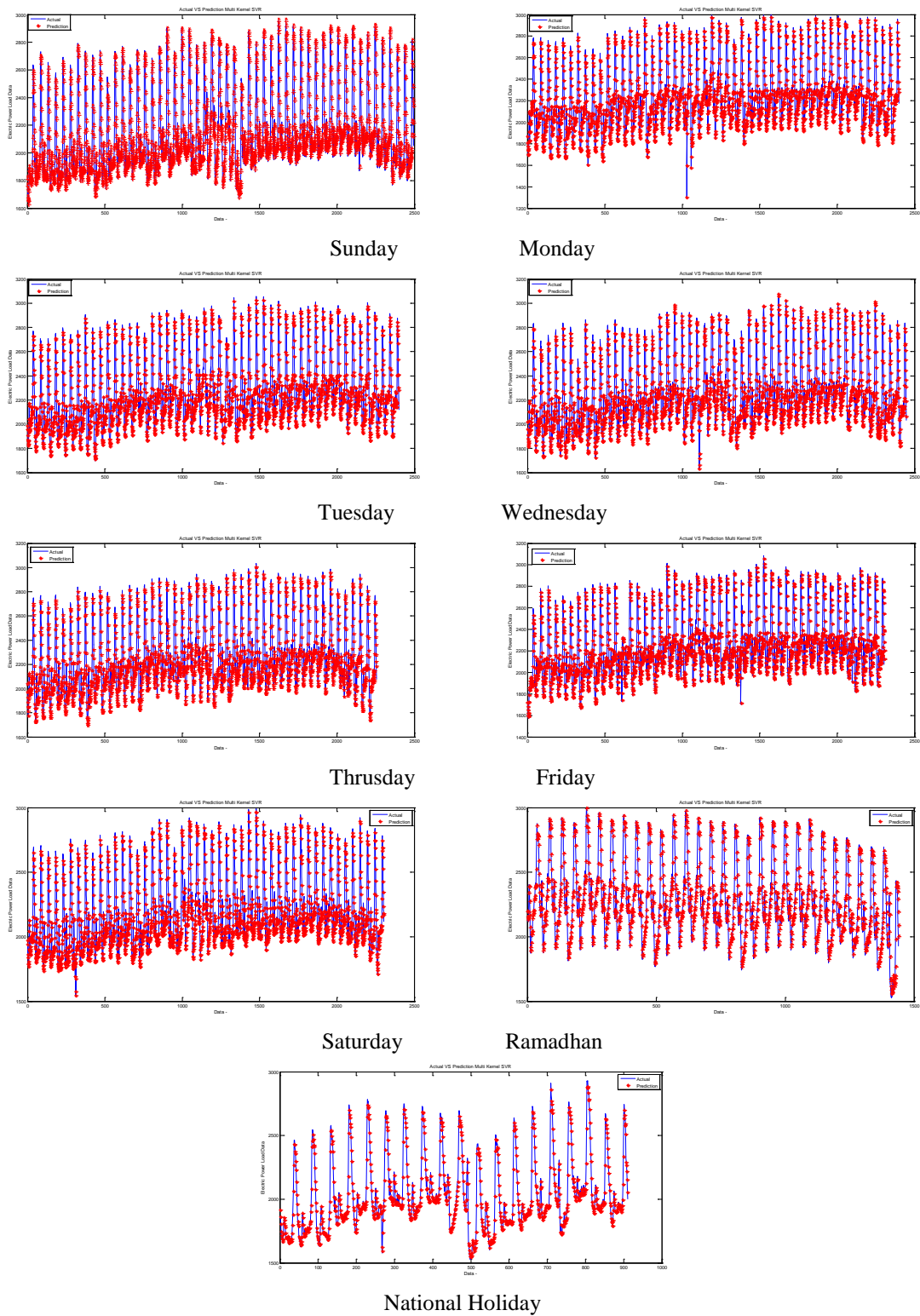


Figure 1. Simulation Based on LMKL with Different Optimization

### 3. Holy Ramadhan

In the month of Holy Ramadhan electrical load has increased compared with regular months. For example, 03:00 to 04:30 This is because the people, especially the Muslims, have been waking up to having food and perform dawn and dawn prayers. Around 05:30 to 7:30 the electrical load in the fasting month is lower than the regular month due to the early morning activities such as preparing breakfast using electrical appliances to switch to the dawn activities.

Besides, The significance of this research is that we do a combination of several types of optimization by looking at the MAPE and RMSE values as the accuracy of the model.

$$RMSE = \sqrt{(f - o)^2}$$

With f=forecast (expected values or unknown results) and o=observed values (known results)

$$RMSE_{f_0} = \left[ \sum_{i=1}^N \frac{(z_{fi} - z_{oi})^2}{N} \right]^{\frac{1}{2}}$$

Where :  $\sum$ =summation ;  $(z_{fi} - z_{oi})$  sup > 2 =differences, squared and ; N= Sample size

After the simulation results, we found that the MAPE value is of little value that can be justified that the model is good to use. For more details can be seen in table 1

**Table 1. Optimization using LMKSVR**

Data	N	Regularization	MAPE	RMSE
Monday	2400	C=100 ; Tube width(e)= 0.5 Optimition= SMO, Locality =Linear	0.0249	88.6135
		C=25 ; Tube width(e)= 0.5 Optimition= MOSEK, Locality =Quadratic	0.0249	89.1185
		C=50 ; Tube width(e)= 0.5 Optimition=SMO, Locality =Quadratic	0.0250	88.6291
		C=100 ; Tube width(e)= 0.5 Optimition= SMO, Locality =Linear	0.0233	84.1870
Tuesday	2400	C=25 ; Tube width(e)= 0.5 Optimition= MOSEK, Locality =Quadratic	0.0233	84.4312
		C=50 ; Tube width(e)= 0.5 Optimition=SMO, Locality =Quadratic	0.0240	82.5484
		C=100 ; Tube width(e)= 0.5 Optimition= SMO, Locality =Linear	0.0236	85.0443
		C=25 ; Tube width(e)= 0.5 Optimition= MOSEK, Locality =Quadratic	0.0244	82.3021
Wednesday	2400	C=50 ; Tube width(e)= 0.5 Optimition=SMO, Locality =Quadratic	0.0236	84.9959
		C=100 ; Tube width(e)= 0.5 Optimition= SMO, Locality =Linear	0.0239	84.9062
		C=25 ; Tube width(e)= 0.5 Optimition= MOSEK, Locality =Quadratic	0.0241	84.1622
		C=50 ; Tube width(e)= 0.5 Optimition=SMO, Locality =Quadratic	0.0239	0.0239
Thrusday	2256	C=100 ; Tube width(e)= 0.5 Optimition= LIBSVM, Locality =Linear	0.0240	86.1961
		C=25 ; Tube width(e)= 0.5 Optimition= MOSEK, Locality =Quadratic	0.0247	84.3274
		C=50 ; Tube width(e)= 0.5 Optimition=SMO, Locality =Quadratic	0.0241	85.6633
		C=100 ; Tube width(e)= 0.5 Optimition= LIBSVM, Locality =Linear	0.0232	85.4363
Friday	2304	C=25 ; Tube width(e)= 0.5 Optimition= MOSEK, Locality =Quadratic	0.0232	85.3757
		C=50 ; Tube width(e)= 0.5 Optimition=SMO, Locality =Quadratic	0.0232	85.4044
		C=100 ; Tube width(e)= 0.5 Optimition= SMO, Locality =Linear	0.0220	85.1964
		C=50 ; Tube width(e)= 0.5 Optimition= Quadrpog, Locality =Quadratic	0.0220	85.0730
Saturday	2304	C=25 ; Tube width(e)= 0.5 Optimition= LIBSVM, Locality =Linear	0.0219	84.9332
		C=50 ; Tube width(e)= 0.5 Optimition=SMO, Locality =Quadratic		
		C=100 ; Tube width(e)= 0.5 Optimition= SMO, Locality =Linear		
		C=25 ; Tube width(e)= 0.5 Optimition= LIBSVM, Locality =Linear		
Sunday	2496			

Holy Ramadhan	1440	C=100 ; Tube width(e)= 0.5 Optimisation= LIBSVM, Locality =Linear	0.0231	76.3535
		C=50 ; Tube width(e)= 0.5 Optimisation= Quadrpog, Locality =Quadratic	0.0230	76.4327
		C=25 ; Tube width(e)= 0.5 Optimisation= LIBSVM, Locality =Linear	0.0232	75.4842
National Holiday	912	C=100 ; Tube width(e)= 0.5 Optimisation= SMO, Locality =Linear	0.0229	84.6098
		C=50 ; Tube width(e)= 0.5 Optimisation= QUADPROG, Locality =Linear	0.0228	84.8262
		C=25 ; Tube width(e)= 0.5 Optimisation= SMO, Locality =Linear	0.0236	83.0314

What is interesting about the method of support vector regression is that researchers can study in the global learning and local learning. The parameters that play a significant role in building the SVR model are the kernel. As has been explained earlier sections that the kernel that can be used very much like Gaussian, radial basis, polynomial and others. So is the type of optimization. In the simulation using electrical data, we found no significant difference between the use of optimization. On the other hand, local learning methods do not estimate a distribution from data. Instead, they focus on extracting only the local information which is directly related to the learning task LKSVR is demonstrated to provide an automatic scheme also systematic to locally and flexibly adapt the margin based on SVR. Besides, LKSVR may be tolerated of the noise. The significance of this research is we managed to create a Matlab graphical user interface (GUI) on a localized multi kernel (LMK).

Matlab (Matrix Laboratory) is a programming software that works with matrix concepts. Matlab is a high-level programming language devoted to technical computing. Matlab is developed as a programming language as well as a visualization tool that offers many capabilities to resolve cases directly related to the disciplines of mathematical science, such as engineering, physics, statistics, computing and modeling. Other Matlab capabilities that can create a GUI (Graphical User Interface) is a Window-based applications

The GUI is basically a graphical display medium instead of a text command (syntax) for users interacting, so the GUI is called Window-based applications. Creating Window-based applications with Matlab can be done using the GUI Designer (GUIDE). The GUI Designer provides a place to design the GUI as well as the components required in making the GUI

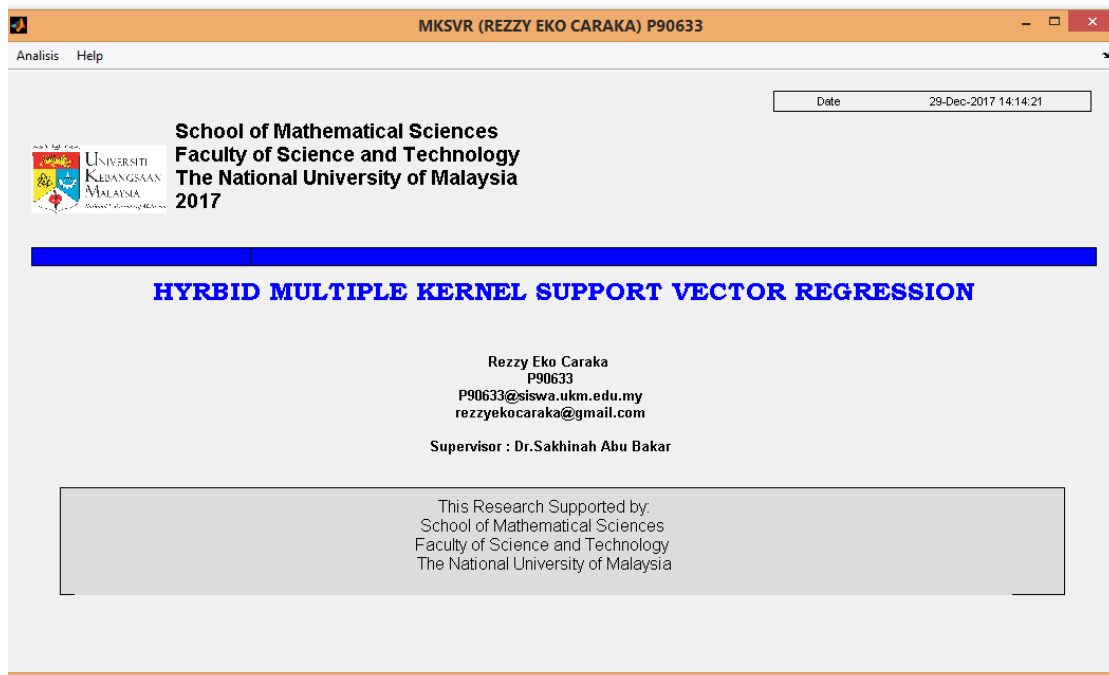


Figure 2. Main GUI MKSVR

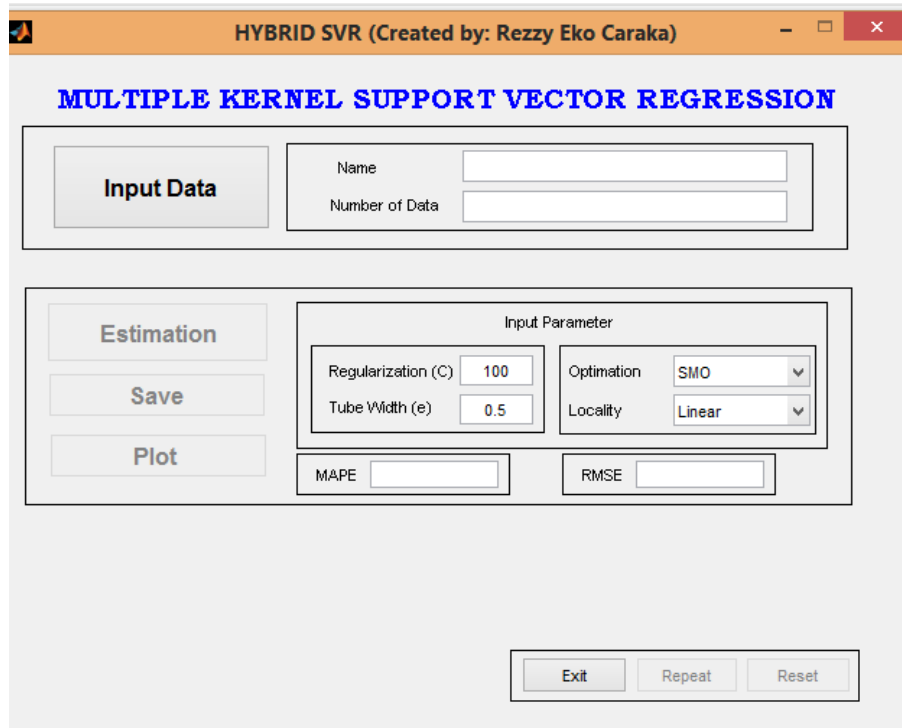


Figure 3. GUI MKSVR



## V. CONCLUSION

Based on the results achieved, it can be concluded that MKSVR model is very good in forming electrical load model. This can be seen from the value of MAPE generated all of them less than 2%. Local learning adopts a largely different way to construct parameters. This type of learning is even more task-oriented than Minimum Error Minimax Probability Machine and Maximal Conditional learning.

## REFERENCES

- Caraka, R. E. 2017. Ramadhan Short-Term Electric Load: A Hybrid Model of Cycle Spinning Wavelet and Group Method Data Handling (CSW-GMDH). *American Institute of Physics*.
- Caraka, R. E., Bakar, S. A., Pardamean, B. & Budiarto, A. 2017. HYBRID SUPPORT VECTOR REGRESSION IN ELECTRIC POWER LOAD DURING NATIONAL HOLIDAY SEASON. *International Conference on Innovative and Creative Information Technology*.
- Caraka, R. E., Yasin, H. & Basyiruddin, A. W. 2017. Peramalan Crude Palm Oil ( CPO ) Menggunakan Support Vector Regression Kernel Radial Basis. *Matematika* 7(1): 43–57. doi:<https://doi.org/10.24843/JMAT.2017.v07.i01.p81>
- Chang, C. & Lin, C. 2013. LIBSVM: A Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology (TIST)* 2: 1–39. doi:10.1145/1961189.1961199
- Hong, T. 2010. *Short Term Electric Load Forecasting. Proquest*. North Carolina State University.
- Kavousi-Fard, A. & Kavousi-Fard, F. 2013. A new hybrid correction method for short-term load forecasting based on ARIMA, SVR and CSA. *Journal of Experimental and Theoretical Artificial Intelligence*. Taylor & Francis. doi:10.1080/0952813X.2013.782351
- Qi, S. 2013. Support Vector Machine and its realization through LIBSVM: A Superficial Review 1–11. Retrieved from [www.songquestc.com/uploads/svm\\_personal\\_survey.pdf](http://www.songquestc.com/uploads/svm_personal_survey.pdf)
- Srivastava, A. K., Pandey, A. S. & Singh, D. 2016. Short-term load forecasting methods: A review. *2016 International Conference on Emerging Trends in Electrical Electronics & Sustainable Energy Systems (ICETEESES)* 130–138. doi:10.1109/ICETEESES.2016.7581373
- Yasin, H., Caraka, R. E., Tarno & Hoyyi, A. 2016. Prediction of crude oil prices using support vector regression (SVR) with grid search - Cross validation algorithm. *Global Journal of Pure and Applied Mathematics* 12(4).