

# Adaptive ARIMA Model based on Lazy Learning Algorithm for Short Period Electric Load Forecasting

Chengze LI, Tomohiro MURATA.

**Abstract**—The short term electric load forecasting which is generally from one hour to one week is one of the intelligent electric grid (smart grid), for control of stable load supply hour-to-hour or day-to-day. The difficulty of short time forecasting is that the trend of time series usually change, and the non-adaptive auto-regressive integrated moving average (ARIMA) could not fit accurately. To solve that problem, conventional adaptive ARIMA with constant forgetting factor that gives a larger weight to more recent train data for dealing with non-stationary change of stochastic disturbance. The forgetting factor governs the recursive least squares (RLS) algorithm. However, constant forgetting factor usually result in over-fitting that increases forecasting error. A new adaptive ARIMA is proposed in this paper to improve the accuracy with lazy learning algorithm to reduce over-fitting error.

**Index Terms**—Intelligent electric grid, RLS algorithm, forgetting factor, ARIMA, lazy learning algorithm.

## I. INTRODUCTION

THE Electric power quality is the vital consideration of steady power supply in industry. The running well power quality could be defined as stable electric load with normal range that is determined by technical requirement.<sup>[1]</sup>

The intelligent electric grid could improve the efficiency and reduce rates of power supply anomalies and fluctuations. For smart grid, the system is established with a variety of infrastructure, electric equipment and intelligence system<sup>[2]</sup>. The intelligence layer in electric utility could establish a more responsive grid that balances the power supply and demand.

In industry, technologies of intelligent grid includes analyzing data that is gathered from electric devices and integrating the intelligent operation systems. The purpose of intelligent system is to decrease the possibility rate of power outages, which could improve economic benefits, according to industry standard that excessive amplitude of power fluctuations could result in damage of electrical equipment.

Chengze LI, Waseda University Graduate School of Information Production and System (IPS), Japan. (e-mail: chengzeli@fuji.waseda.jp).

Tomohiro MURATA, Professor of Waseda University Graduate School of Information Production and System (IPS), Japan. (e-mail: t-murata@waseda.jp).

The smart monitoring device could improve the robustness and reliability of distributive power grid.<sup>[3]</sup>

Prediction of electric load is one part of the construction of intelligent grid composited with intelligent electric devices and related intelligent software. It is possible to adjust the electric load in stable way by analyzing forecasting results, which is vital part of production of power generation and supply in smart grid.<sup>[4]</sup>

There are several related methods of short term electric load forecasting are commonly used in engineering, such as neural network, regression. The neural network is a non-parametric model and has advanced supervise learning ability but needs to be trained before applied<sup>[5]</sup>. Regression is a conventional method that it is able to determine the relative influence of one or more predictor variables to the criterion value. However, it is easy for regression to result in over-fitting or under-fitting.

In this paper, ARIMA model that is usually used in nonlinear time series is selected to be improved to get result that is more accurate. However, conventional non-adaptive ARIMA has a drawback of its static model that could not refresh fitness when the trend of time series have changed. That drawback results in decrease of the forecasting accuracy.

## II. FORMULATION OF PROBLEM

### A. Difficulties of short time period forecasting

Difficulties of short time period electricity load forecasting usually include two parts. First difficulty is non-stationary change of stochastic disturbance in short-term electricity demand. Second difficulty is non-stationary change of trend in short-term electricity demand. Forecasting method for short-term electricity demand should be robust and adaptive against non-stationary change.<sup>[6]</sup>

The non-adaptive ARIMA (p,d,q) is shown as following:

$$Y_t = (\varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p}) \nabla^d + e_t + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q} \quad (1)$$

Moreover, the coefficients of function are:

$$\Psi = (\varphi_1, \varphi_2, \dots, \varphi_p) \quad (2)$$

$$\Theta = (\theta_1, \theta_2, \dots, \theta_q) \quad (3)$$

Difference operator is:

$$\nabla^d Y_t = Y_t - Y_{t-d} \quad (4)$$

### B. CONVENTIONAL ADAPTIVE ARIMA MODEL

That above non-adaptive ARIMA (p,d,q) model need to be updated according to non-stationary change of time series data including sharp demand change in short period. To solve those problems, considering the adaptive ARIMA model to fit

better. In addition, conventional adaptive ARIMA is based on RLS-forgetting factor algorithm, as one of the Kalman filters family, which is used to increase the weight of more recent train data. Because the performance of RLS algorithm in stability and tracking is based on forgetting factor, as mention above, constant forgetting factor has an unwanted performance of overfitting. Conventional Adaptive ARIMA of RLS-forgetting factor update parameters<sup>[7]</sup> in this following way. The error between forecasted data and actual data  $Y_t$ :

$$L(t) = \sum_{t=p+1}^n \sum_{t=q+1}^n Y_t - [(\varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p}) \nabla^d + e_t + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q}] \quad (5)$$

The weighted least squares error function  $C$  which is desire to minimize:

$$C(\Psi) = \lambda^{n-1} L(t)^2 \quad (6)$$

$$C(\Theta) = \lambda^{n-1} L(t)^2 \quad (7)$$

$\lambda$  is constant forgetting factor,  $\lambda \in (0,1)$ .

Adaptation of forecasting function's coefficient  $\Psi$  and  $\Theta$  to minimize  $C(\Psi)$  and  $C(\Theta)$ , and the  $\Psi^*$  and  $\Theta^*$  are calculated as following:

$$\Psi^* = \operatorname{argmin} C(\Psi) \quad (8)$$

$$\Theta^* = \operatorname{argmin} C(\Theta) \quad (9)$$

However, the drawbacks of the conventional Adaptive ARIMA model is that the RLS-Forgetting factor algorithm by using constant  $\lambda$  is proposed but constant  $\lambda$  need to be decided by experiment and is not enough for Adaptation of ARIMA against overfitting problem<sup>[8]</sup>. For example, when  $\lambda$  is close to 1, which results in low misadjustment & good stability and tracking capabilities are reduced in the same time. When  $\lambda$  is smaller, it could improves tracking capabilities, but increases misadjustment. Therefore, we need variable weight parameter, and it is proposed by using recursive least square algorithm, and KNN-weight algorithm.

### III. PROPOSED METHOD

#### A. Observation time window

Lazy learning algorithm provides useful training algorithms based on local time series data. In this paper, KNN that is one of the lazy learning algorithms is used to select the train data for local learning could create observation time window.

The k-nearest neighbors (KNN) is a kind of lazy learning algorithm, which is usually used to applied into local learning. Generally, KNN is based on the Euclidean distance to calculate similarity between center points and other samples. However, this paper proposes a method to select the train data to establish the observation time window by using KNN's idea.

The observation time window could choose  $k$  points nearest to the most recent data point  $x_s$ , and  $x_t$  is the  $t$ -th sample point nearest to  $x_s$ . In the observation time window, measuring the Euclidean distance between  $x_t$  and  $x_s$ , and that is related to Gaussian RBF for giving larger weight to more recent data. The observation time window is used for solving the problem of overfitting and underfitting.

#### B. Update of parameters

The RLS algorithm is used for the update of ARIMA model's parameters to make adaptive prediction. It is noteworthy that parameter update could avoid overfitting caused by the not good global learning.<sup>[9]</sup>

Additionally, the idea of lazing learning algorithm, which is do local learning, could solve the problem of overfitting with a novel selection method of train data and determination of adaptive weights that has better performance than constant forgetting factor.<sup>[10]</sup>

Proposed adaptation method of weight and  $(p, d, q)$  parameters based on lazy learning algorithm is delivered to improve the prediction performance. Every time a point is predicted, the parameter is updated once. The proposed parameters update of adaptive ARIMA  $(p, d, q)$  is given as follows:

$$Y_t = (\varphi_1^* Y_{t-1} + \varphi_2^* Y_{t-2} + \dots + \varphi_p^* Y_{t-p}) \nabla^d + e_t + \theta_1^* e_{t-1} + \dots + \theta_q^* e_{t-q} \quad (10)$$

$\Psi^*$  is such that  $\min(w(z_t)L(t)^2)$

$\Theta^*$  is such that  $\min(w(z_t)L(t)^2)$

Where

$$w(z_t) = \exp\{-\frac{\varepsilon}{2} \|z_t - z_s\|^2\} \quad (11)$$

$\varepsilon$  is smoothing factor, and equals 0.5;

$z_t, z_s$  are normalized  $x_t$  and  $x_s$ :

$$z_t = \frac{x_t - \min(X)}{\max(X) - \min(X)} \quad (12)$$

$$z_s = \frac{x_s - \min(X)}{\max(X) - \min(X)} \quad (13)$$

Where

$$X = (x_1, \dots, x_n)$$

$$z_t \subseteq [0,1]$$

$$z_s \subseteq [0,1]$$

And error function is:

$$L(t) = \sum_{t=p+1}^n \sum_{t=q+1}^n Y_t - [(\varphi_1^* Y_{t-1} + \varphi_2^* Y_{t-2} + \dots + \varphi_p^* Y_{t-p}) \nabla^d + e_t + \theta_1^* e_{t-1} + \dots + \theta_q^* e_{t-q}] \quad (14)$$

$\Psi^*$  and  $\Theta^*$  are given by solving the equation of partial derivatives, and get coefficient  $\Psi^*$  and  $\Theta^*$ :

$$\begin{aligned} \frac{\partial C(\Psi)}{\partial \Psi} &= \sum_{t=p+1}^n 2w(z_t)L(t) \frac{\partial L(t)}{\partial \Psi} \\ &= - \sum_{t=p+1}^n 2w(z_t)L(t) \nabla^d Y(t-\tau) = 0 \end{aligned} \quad (15)$$

$$\tau = 1, 2, 3, \dots, p$$

$$\begin{aligned} \frac{\partial C(\Theta)}{\partial \Theta} &= \sum_{t=q+1}^n 2w(z_t)L(t) \frac{\partial L(t)}{\partial \Theta} \\ &= - \sum_{t=q+1}^n 2w(z_t)L(t)e(t-\tau) = 0 \end{aligned} \quad (16)$$

$$\tau = 1, 2, 3, \dots, q$$

Replacing  $\Psi$  with  $\Psi^*$ , and getting forecasted time series  $Y_t$ :

$$Y_t = (\varphi_1^* Y_{t-1} + \varphi_2^* Y_{t-2} + \dots + \varphi_p^* Y_{t-p}) \nabla^d + e_t + \theta_1^* e_{t-1} + \dots + \theta_q^* e_{t-q} \quad (17)$$

#### IV. EXPERIMENT

The experiments and associated results are based on above methodology. The proposed method is tested with a case study of a city's electric load of south China, selecting two days data from summer period (case data-1) and spring period (case data-2). Additionally, first 15 hours are train data and left 8 hours are test data.

The measurement of prediction accuracy are based on Mean Absolute Percentage Error (MAPE) and Root-Mean-Square Error (RMSE) and they are defined as follows:

$$MAPE = \frac{\sum_{i=1}^N \left| \frac{x_i - \hat{x}_i}{x_i} \right|}{N} \times 100\% \quad (18)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (x_i - \hat{x}_i)^2}{N}} \quad (19)$$

The vital part of proposed method is to determine the size of observation time window. In experimental test, the size  $k$  which represents the number of points nearest to  $x_s$  of observation time window is to set up from 1 to 15 for finding out the best size  $k$  of train data  $x_t$  to minimize forecasting error. According to the fifteen results of best size  $k$ , and the results are shown in figure 5-1 and figure 5-2, the error rate of MAPE and RMSE is lowest when  $k$  equals 9. Due to that experimental test is based on a case study of summer period data, so the best  $k$  of summer period is 9.

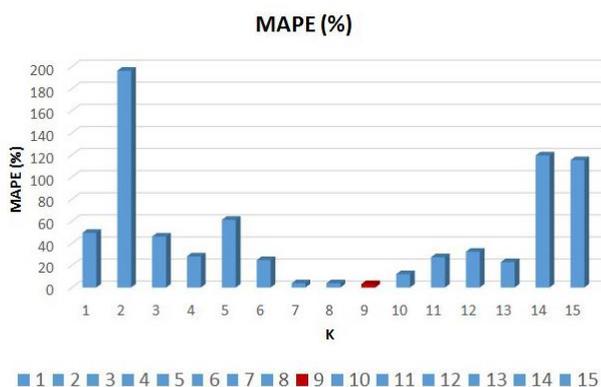


Fig 1. MAPE of proposed adaptive ARIMA model with different size of observation window.

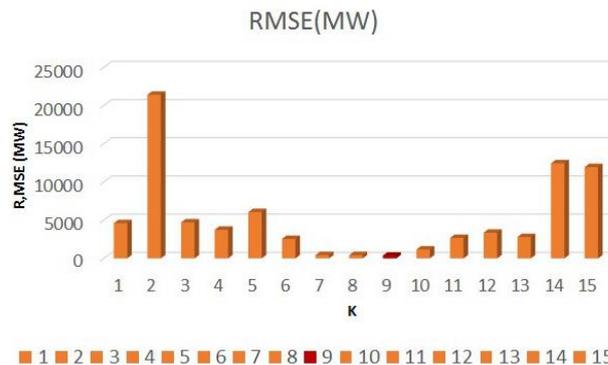


Fig 2. RMSE test of proposed adaptive ARIMA model with different size of observation window.

#### A. Forecasting accuracy evaluation

This part is the evaluation by comparing proposed adaptive ARIMA, conventional adaptive ARIMA with constant forgetting factor and non-adaptive ARIMA, and the training is based on summer period data (case data-1).

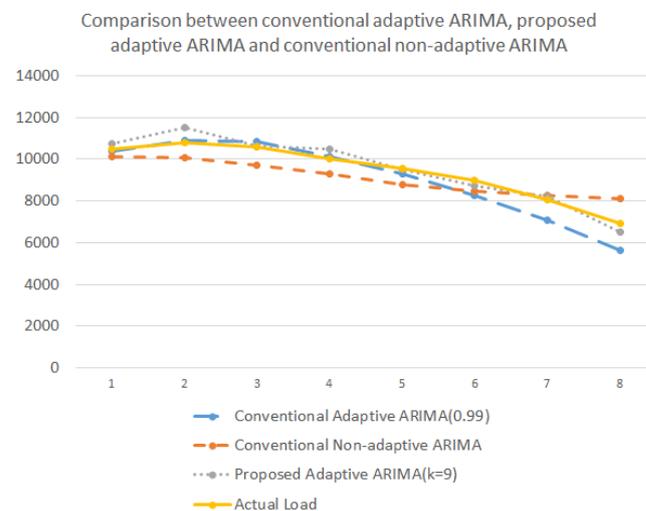


Figure 3. Comparison between conventional adaptive ARIMA, proposed adaptive ARIMA and conventional non-adaptive ARIMA.

The MAPE of proposed adaptive ARIMA model is 3.117601044 % that is lower than the conventional adaptive ARIMA model. That means the proposed adaptive ARIMA model is better than conventional RLS-forgetting factor method, and improves 84.821% in MAPE and 75.073% in RMSE error evaluation.

Table I MAPE & RMSE OF THREE METHODS UNDER SUMMER PERIOD

	Conventional Adaptive ARIMA(0.99)	Proposed Adaptive ARIMA(K=9)	Conventional Non-adaptive ARIMA
MAPE (%)	5.76195944754	3.117601044	7.423626457
RMSE (MW)	630.7553231	360.280489	731.9617385

According to MAPE and RMSE evaluation of accuracy, the proposed method has the best performance compared with conventional non-adaptive ARIMA and conventional adaptive ARIMA.

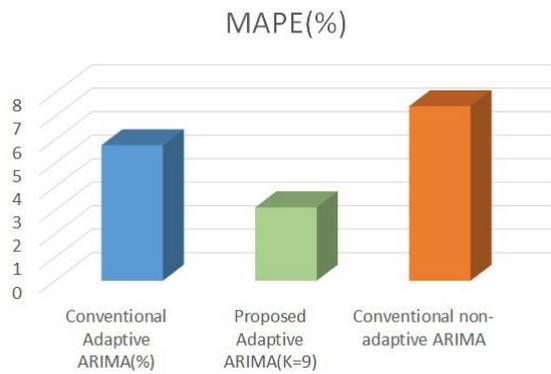


Figure 4. MAPE of conventional adaptive ARIMA, proposed adaptive ARIMA and conventional non-adaptive ARIMA in summer period.

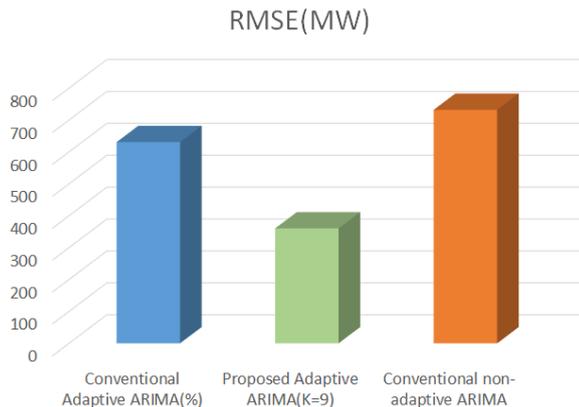


Figure 5. RMSE of conventional adaptive ARIMA, proposed adaptive ARIMA and conventional non-adaptive ARIMA in summer period.

### B. Stability evaluation of proposed method

For the evaluation of proposed method's forecasting accuracy stability or robustness, training proposed adaptive ARIMA model with spring period data (case data-2).

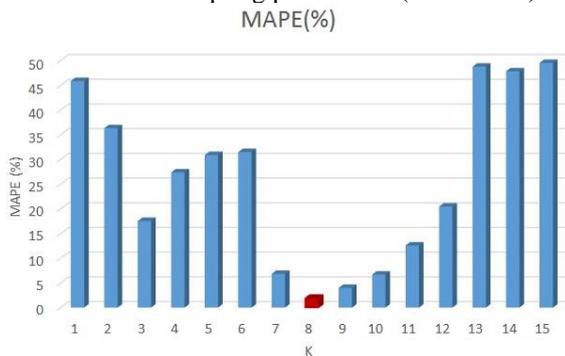


Figure 6. MAPE of conventional adaptive ARIMA, proposed adaptive ARIMA and conventional non-adaptive ARIMA in spring period

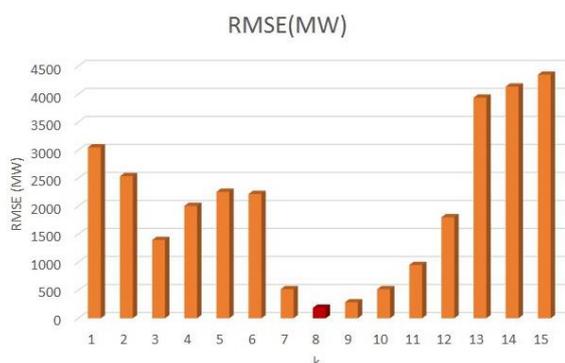


Figure 7. RMSE of conventional adaptive ARIMA, proposed adaptive ARIMA and conventional non-adaptive ARIMA in spring period

Table II. MAPE & RMSE OF PROPOSED METHOD (k=8)

	Proposed Adaptive ARIMA(K=8)
MAPE (%)	1.88668354394
RMSE (MW)	187.134295961

The best k is 8 according to the figures above, and establishing model with that k. The train performance is robust both in summer time and spring time. In the same quarter, the patterns of power fluctuations are similar, so we could set up the observation window and use it during each quarter that means there are different values of k during different quarter.

### C. Evaluation of weight measure effect on accuracy

For the evaluation by comparing proposed adaptive ARIMA with variable weight based on different type RBF, since the RBF includes Multiquadric, Inverse quadric, Inverse multiquadric and Gaussian. It is necessary to find which RBF is the best measurement of weight with the summer time train data.

The experimental results show that Gaussian RBF that is used in proposed method has the best performance compared with others.

Table III. MAPE & RMSE OF DIFFERENT ALGORITHMS (k = 9, using data of case-1)

	Multiquadric	Inverse quadric	Inverse multiquadric	Gaussian
MAPE (%)	36.42854145	5.58506678	32.47059133	3.117601044
RMSE (MW)	3728.53073	675.8263875	3296.194113	360.280489

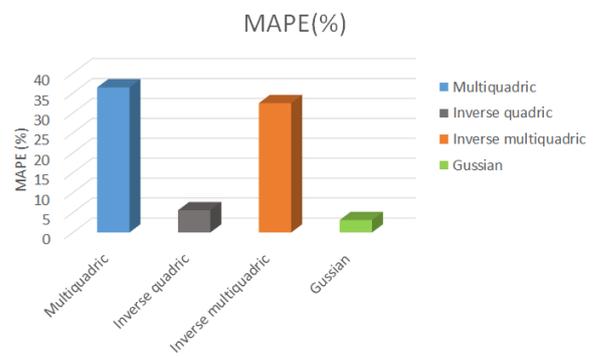


Figure 8. RMSE of Multiquadric, Inverse quadric, Inverse multiquadric and Gaussian RBF

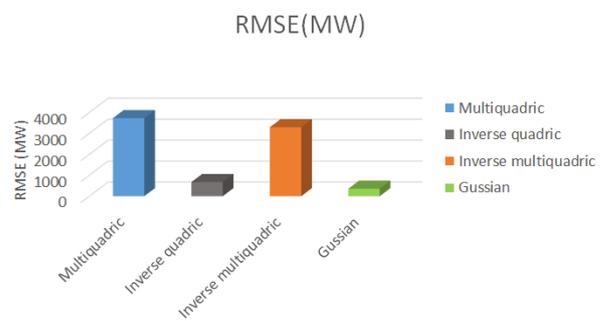


Figure 9. RMSE of Multiquadric, Inverse quadric, Inverse multiquadric and Gaussian RBF

In addition, the different weight measure is applied into spring period to test the accuracy. The experimental results show that Gaussian RBF that is used in proposed method has the best performance compared with others.

The results are showed that the Gaussian function has the best performance compared with other functions, and Gaussian function has good robustness both in spring and summer period.

Table IV. MAPE & RMSE OF DIFFERENT ALGORITHMS  
 (k = 8, using data of case-2)

	Multiquadric	Inverse quadric	Inverse multiquadric	Gaussian
MAPE (%)	27.2854783	17.77760811	13.54369115	1.88668354394
RMSE (MW)	1864.684048	1552.856603	1007.548254	187.134295961

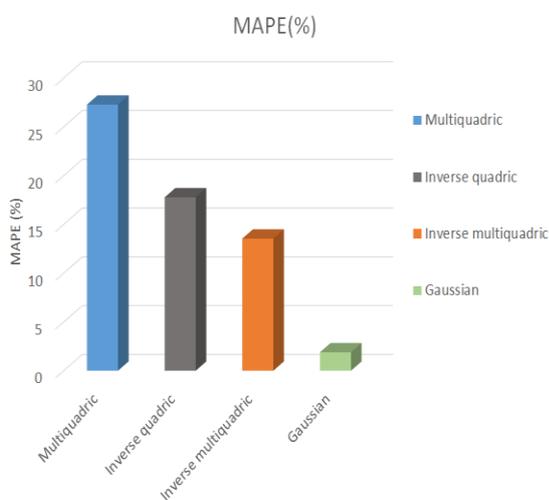


Figure 10. RMSE of Multiquadric, Inverse quadric, Inverse multiquadric and Gaussian RBF (case-2, spring period)

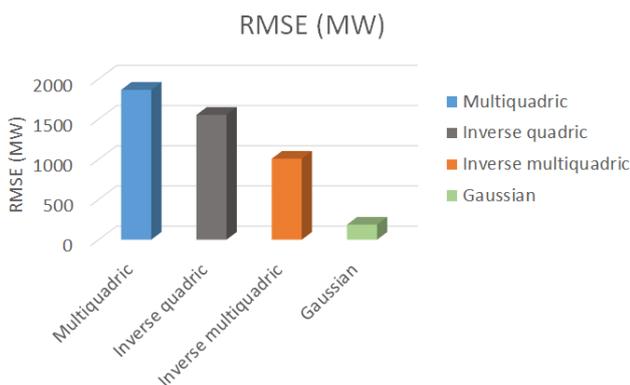


Figure 11. RMSE of Multiquadric, Inverse quadric, Inverse multiquadric and Gaussian RBF (case-2, spring period)

### V. CONCLUSION

This paper proposes a novel method of adaptive ARIMA compared with conventional adaptive ARIMA. The experimental results show the advanced performance of proposed method. Because there is little fluctuation in daily electric load during the same quarter, selecting fixed observation for each quarter is acceptable. Moreover, the parameters of ARIMA model could be updated with the forecasting of each points.

The proposed method improves 84.821% in MAPE and 75.073% in RMSE error evaluation compared with conventional adaptive ARIMA that is based on the training of summer data. In addition, the proposed method also need to be trained with other city’s data to test the robustness in the future work.

### ACKNOWLEDGMENT

Thanks for the supervisor Prof. MURATA who is my tutor, guide, and give my research valuable advice. The case study shows the train performance and advantages of my proposed method. In addition, I want to thank the help of my friends in Waseda University.

### REFERENCES

- [1] Carlo Cecati, Janusz Kolbusz, A novel RBF training algorithm for short-term electric load forecasting and comparative studies. IEEE Transaction on Industrial Electronics (Volume: 52, Issue: 10, Oct. 2015)
- [2] Mohamed E. El-hawary. The Smart Grid-State of the Art and Future Trends. Pages 239-250. 19, Nov 2013. Electric Power Components and System.
- [3] W.J.N.Turner, I.S. Walker, J.Roux. Peak load reductions: Electric load shifting with mechanical pre-cooling of residential buildings with low thermal mass. Feb, 2015. Volume 82, 15 March 2015, Pages 1057-1067, Energy.
- [4] Etimad Fadel, V.C. Gungor, Laila Nassef, Nadine Akkari, M.G. Abbasa Malik, Suleiman Almasri, Ian F. Akyildiz. Volume 71, 1 November 2015, Pages 22-23. Computer Communications.
- [5] Hao Quan, Dipti Srinivasan, Abbas Khosravi. Short-term load and wind power forecasting using neural network-based prediction intervals. (Volume: 25, Issue 2, Feb.2014). IEEE Transactions on Neural Networks and Learning Systems.
- [6] Shu-Hung Leung, Gradient-based variable forgetting factor RLS algorithm in time-varying environments. IEEE Transactions on Signal Processing (Volume: 53, Issue 8, Aug.2005). 18 July 2005, pp. 3141–3150.
- [7] Constantin Paleologu, Jacob Benesty, Silviu Ciochina. A Practical Variable Forgetting Factor Recursive Least-Squares Algorithm. Electronics and Telecommunications (ISETC), 2014 11<sup>th</sup> International Symposium on. 15 January 2015.
- [8] Constantin Paleologu, Jacob Benesty, Silviu Ciochina. A Robust Variable Forgetting Factor Recursive Least-Squares Algorithm for System Identification. IEEE Signal Processing Letters. 07 October 2008 (Volume: 15). Pages: 597-600.
- [9] Hiroki Nakayama, Shingo Ata, Ikuo Oka. Predicting Time Series of Individual trends with Resolution Adaptive ARIMA. 14 November 2013.
- [10] PatriciaSholl, R.KennethWolfe. The Kalman Filter as an adaptive forecasting procedure for use with Box-Jenkins arima models. Received 6 October 1984, Available online 18 February 2003. Computers and Industry Engineering.