Applied Mechanics and Materials

TTP TRANS TECH PUBLICATIONS

Applied Mechanics and Materials Vol 781 (2015) pp 245-249
© (2015) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMM.781.245

Submitted: 2014-09-30
Revised: 2015-03-26
Accepted: 2015-04-06

Time Series Pattern Learning and Forecasting for Long-Term Peak Electricity by Spectral Mixture Gaussian Kernel

Tuchsanai Ploysuwan[1,a*], Prasit Teekaput[2,b], Pramukpong Atsawathawichok[3,c*]

[1]Department of Electrical Engineering Siam University, Thailand
[2,3]Department of Electrical Engineering Chulalongkorn University, Thailand

[a]tuchsanai@gmail.com, [b]Prasit.T@chula.ac.th, [c]g.pramukpong@gmail.com

Keywords: Gaussian Process, Peak load, Kernel Function, long-term load forecasting

Abstract. This paper presents the mathematical model for forecasting of future long-term peak electricity load from January 2014 to December 2024 with totally 132 months from the past knowledge data of training 156 months. The new kernel method is proposed by the combination of summed weight spectral mixture Gaussian in the frequency domain and squared exponential in the time domain, which are used as components in the answer of Gaussian Process (GP). Finally, the results show the prediction error mean absolute percentage error (MAPE) by 2.3283%.

Introduction

The demand of long-term peak load forecasting is a tool to estimate the energy load or peak load in manufacture and supply for consumers in the future more than 1 year. So it's very important in the planning of electricity, the purchase - sale of electricity in advance, and for setting the electricity rate in the future. The data of peak load and energy load are saved and collected in the annual report by "Electricity Generating Authority of Thailand" (EGAT) [1] continuously every year. In addition, the peak load is also related with multiple external variables such as climate, trend of using energy in the past, population, gross domestic product (GDP), oil prices, and the electricity rate per unit. The technique of machine learning that uses external variables as an input to predict are neural network (NN) [2],[3], support vector regression (SVR) [4] etc. But in fact, it is difficult to know the long-term of external variables as an exactly input that reduced accurate of prediction.

Recently, it has introduced the GP with mix kernel function [5] to solve the above problem for searching and learning the pattern from the attributes of peak load without external variables which can work efficiently and the convergence rate is better than NN and SVR although there is a small amount of training data [6]. So this research presents the new kernel function in the frequency domain by using weighted spectral mixture Gaussian for the demand of long-term peak load forecasting to increase the precision and more complexity.

Mathematic Model

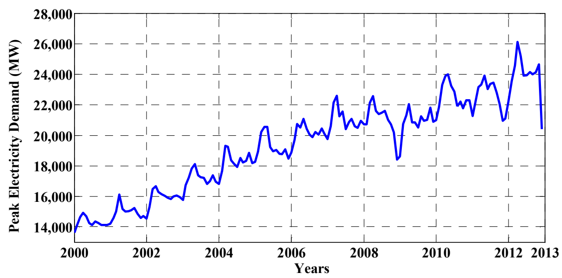

Figure 1 The monthly of peak electricity demand **[1]**

The data that used in the experiments of this paper is the monthly peak load of the EGAT between Jan. 2000 and Dec. 2013 which have total 168 months. Then the data divided into two sets, the first set contains the amount of 156 months that called the training data between Jan. 2000 to Dec. 2012 and the second set is next 12 month Jan. 2013 to Dec. 2013 respectively that called validation period.

The data set of total 168 months peak load is shown in Figure 1. Define time x_i by the index i. i is an integer and its meaning described in the above example. Then define the monthly peak load can modeled by function $f(x_i)$ and the sampling of peak load y_i is measured by EGAT of each x_i

$$y_i = f(x_i) + \varepsilon_i \tag{1}$$

$\varepsilon_i \sim \mathcal{N}(0,\sigma_\varepsilon^2)$ denotes Gaussian error. The point of this paper is to forecast the demand of peak load in long term: $f(x_j)$ by $169 \le j \le 300$ which ranges from January 2014 to December 2024 by using the trends of training period to forecast $f(x_j)$

$$f(x_j) \sim \mathcal{GP}(m(x_j)\ ,k(x_j,x_j)) \tag{2}$$

Gaussian Process

GP defined distribution over function that is $f(x_j)$ caused by the random process of GP contain with mean: $m(x_j)$ and kernel function: $k(x_j,x_j)$. Define training data vector $\mathbf{y}=\{y_i\}_{i=1}^N$ size $N\times 1$ by $N=156$ and the objective want to forecast peak load in long term $f_*=f(x_*)$ with $x_*=x_j$ by $169 \le j \le 300$. Suppose $\mathbf{y}$ and f_* have relation are joint Gaussian

$$\begin{bmatrix}\mathbf{y}\\ f_*\end{bmatrix} \sim \mathcal{N}\left(\mathbf{0},\begin{bmatrix}\mathbf{K}(X,X)+\sigma_\varepsilon^2 I_N & \mathbf{k}(X,x_*)\\ \mathbf{k}(X,x_*)^T & k(x_*,x_*)\end{bmatrix}\right) \tag{3}$$

Setting $\tau_i^* = x_i - x_*$ and from (3) the conditional probability of $P(f_* \mid x_*,\mathbf{y})$ can find an answer of $m(x_*)$ and $k(x_*,x_*)$ by

$$\begin{aligned} m(x_*) &= \mathbf{k}(X,x_*)^T\left(\mathbf{K}(X,X)+\sigma_\varepsilon^2 I_N\right)^{-1}\mathbf{y} \\ k(x_*,x_*) &= \mathbf{k}(x_*,x_*)-\mathbf{k}(X,x_*)^T\left(\mathbf{K}(X,X)+\sigma_\varepsilon^2 I_N\right)^{-1}\mathbf{k}(X,x_*) \end{aligned} \tag{4}$$

From the definition of GP we can use minimum mean square error to predict $f(x_*)$ from $m(x_*)$ in (4) and we know that the accuracy of prediction will depend on the chosen kernel function $k(\tau_j)$. Furthermore, Bochner's theorem [7] has shown that the kernel function $k(\tau)$ has a spectral density $S(f)$ and it's also important parameters to model kernel in this paper.

Proposed Kernel Function

Considering the nature of monthly peak load in Figure 1 shows that the data consists of two forms: 1) long-term smooth and increasing trend 2) Irregular nonlinear change with seasonal trend in each year. In order to combine the features of the kernel function $k(\tau)$ to qualify the two forms that is mentioned. So let

$$k(\tau \mid \theta) = k_{SE}(\tau \mid \theta_{SE}) + k_{SM}(\tau \mid \theta_{SM}) \tag{5}$$

$k_{SE}(\tau \mid \theta_{SE})$ is squared exponential: SE, $k_{SM}(\tau \mid \theta_{SM})$ is spectral mixture Gaussian kernel and θ is called hyper-parameter which contains $\theta = \{\theta_{SE},\theta_{SM}\}$.

Squared Exponential Kernel

$$\begin{aligned} k_{SE}(\tau \mid \theta_{SE}) &= \sigma^2 \exp\left(-\alpha \|\tau\|^2\right) \\ \theta_{SE} &= \{\sigma^2,\alpha\} \end{aligned} \tag{6}$$

The characteristic of squared exponential kernel is a smooth function of time and change by σ^2 controlling an amplitude and α controlling the sensitivity to change the time.

Spectral Mixture Gaussian Kernel

Highlight of GP depends on a selection of an appropriate kernel function for each problem. Sometimes, a simple kernel function can't solve the problem if it's too complicated which is designed to use with GP [6]. So in order to analyze the structure of complicated data, we create a new kernel that it can adapt itself to the structure of the data [8]. Therefore, analysis and design kernel function in terms of spectral density is motivated in this research. By the structural design in frequency domain is good enough then it will be more accurate in time domain [8]. Defining, the spectral density is the sum of weighted Gaussian mixture model of Q in the frequency domain.

$$S(f)=\sum_{q=1}^{Q} w_q \times \psi^{(q)}(f) \quad (7)$$

By w_q is weighting co-efficient that is equivalent with $\sum_{q=1}^{Q} w_q = 1$

$$\phi^{(q)}(f;\ \mu_q,\upsilon_q)=\frac{1}{\sqrt{2\pi\upsilon_q}}\exp\left(-\frac{1}{2\upsilon_q}\left(f-\mu_q\right)^2\right) \quad (8)$$

$\phi^{(q)}(f;\ \mu_q,\upsilon_q)$ is a "q" members of spectral Gaussian consist mean: μ_q, variance: υ_q and $\psi^{(q)}(f)$ is a "q" members of average of spectrum which have symmetry at $f=0$. So we can compute Spectral mixture Gaussian kernel by substituting equation (7) into Bochner's theorem [7] and use table of integrals. We will get the final equation

$$k_{SM}(\tau\,|\,\theta_{SM})=\sum_{q=1}^{Q} w_q \exp\left(-2\pi^2\tau^2 v_q\right)\cos\left(2\pi\tau\ \mu_q\right)$$
$$\theta_{SM}=\left\{w_1,\ldots w_q,\left\{\left\{\mu_q,v_q\right\},\ldots,\left\{\mu_q,v_q\right\}\right\}_{q=1}^{Q}\right\} \quad (9)$$

Therefore, spectral mixture Gaussian kernel has number of hyper-parameters when bring equations (6) and (9) into equation (5) we found that the $k(\tau\,|\,\theta)$ has number of hyper-parameters is $3Q+2$

$$\theta=\{\theta_{SE},\theta_{SM}\}=\left\{\sigma^2,\alpha,w_1,\ldots w_q,\left\{\left\{\mu_q,v_q\right\},\ldots,\left\{\mu_q,v_q\right\}\right\}_{q=1}^{Q}\right\} \quad (10)$$

Simulation Result

The simulation result of the demand of peak load forecasting between January 2014 to December 2024 is divided into 3 sections via training period, validation period and forecasting period.

Training Period

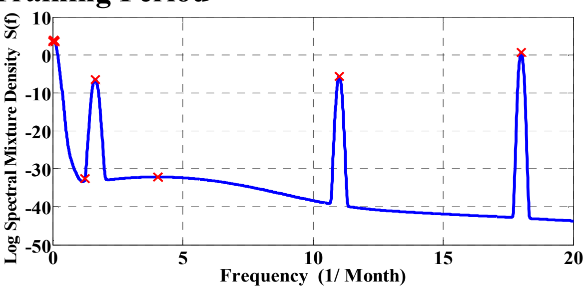

Figure 2 Spectral density $10\log_{10}|S(f)|$ in dB

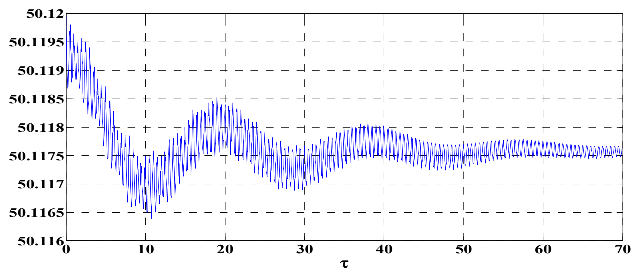

Figure 3 Kernel function $k(\tau\,|\,\theta)=k_{SE}(\tau\,|\,\theta_{SE})+k_{SM}(\tau\,|\,\theta_{SM})$

Using training data $\{y_i\}_{i=1}^{156}$ after calculating the hyper-parameters in [6] We found that the $k_{SE}(\tau\,|\,\theta_{SE})$ has θ_{SE} is $\{\sigma^2 = 0.7297, \alpha=12.3970\}$ and Spectral density of mixture Gaussian kernel by using Q=10 have shown in Fig. 2 and inverse Fourier transform of spectral density which $k_{SM}(\tau\,|\,\theta_{SM})$ is sum of kernel in (5) have shown in Fig. 3

When considering the effect of the kernel $k_{SE}(\tau\,|\,\theta_{SE})$ and $k_{SM}(\tau\,|\,\theta_{SM})$ to structure of peak load function: $f(x_i)$. We found that the peak load function consists of one from $f_{SE}(x_i)$ that obtains from $k_{SE}(\tau\,|\,\theta_{SE})$ and $f_{SM}(x_i)$ that obtain from $k_{SM}(\tau\,|\,\theta_{SM})$ in equation (4) have shown in Fig. 4. Moreover, the sum of peak load function $f_{SE}(x_i)+f_{SM}(x_i)$ are according to the Fig. 5 (blue line) which close to the peak load from data of EGAT (red line).

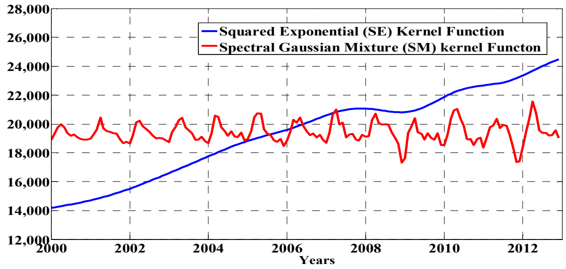

Figure 4 Analysis of each kernel function.

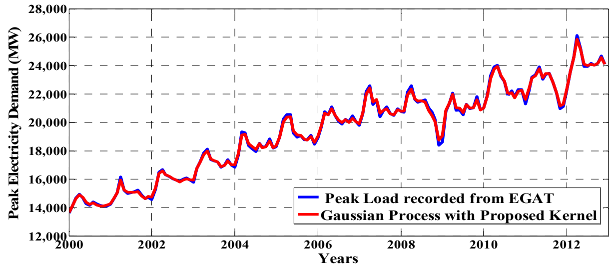

Figure 5 Peak load function by GP

Validation Period

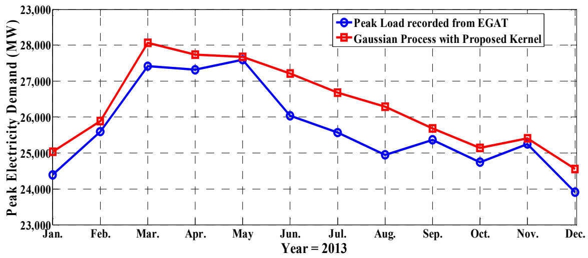

Figure 6 Peak load function in validation period

In order to test the accuracy of peak load forecasting by using spectral mixture Gaussian kernel, the forecast result during validation period using hyper-parameters: θ in the previous training period section have shown in Fig. 6. So, MAPE of proposed kernel is 2.3283% and MAPE in [5] is 2.423% respectively.

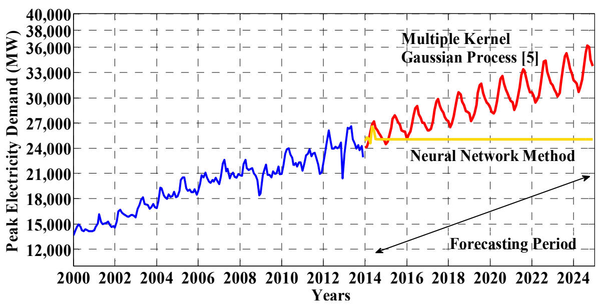

Figure 7 The comparison forecasting results on Jan. 2014 to Dec. 2024 [2], [5]

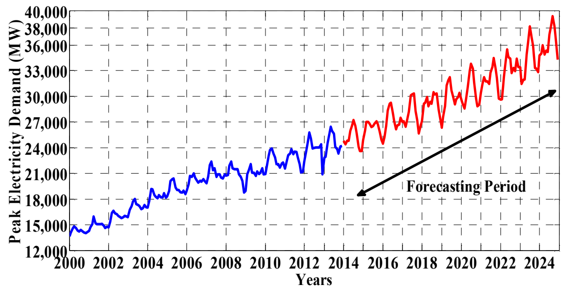

Figure 8 The forecasting result on Jan. 2014 to Dec. 2024 of proposed kernel

Forecasting Period

In order to compare the results, Fig. 7 has shown that the result of mix kernel [5] and time series using neural network (NN) matlab toolbox. The inputs which using in NN toolbox are the same as training data **y** in GP. In addition, the forecasting result of proposed kernel has shown in Fig. 8.

Conclusion

This paper presents long-term peak load forecasting by the design of new kernel. As seen above in validation period that shows MAPE of proposed kernel better than mix kernel in [5]. The trends of peak load in forecasting period, it has complex feature, non-periodic and seasonal trends are similar to the real data.

Reference

[1] "Annual Report " Electricity Generating Authority of Thailand (EGAT) 2000-2013.

[2] M. A. Mamun and K. Nagasaka, "Artificial neural networks applied to long-term electricity demand forecasting," in *Hybrid Intelligent Systems,* 2004, pp. 204-209.

[3] Z. Shi and W. Dingwei, "Medium and Long-Term Load Forecasting Based on PCA and BP Neural Network Method," in *Energy and Environment Technology*, 2009, pp. 389-391.

[4] L. Ghelardoni, A. Ghio, and D. Anguita, "Energy Load Forecasting Using Empirical Mode Decomposition and Support Vector Regression," *IEEE Trans on Smart Grid* vol. 4, pp. 549-556, 2013.

[5] P. Atsawathawichok, P. Teekaput, and T. Ploysuwan, "Long term Peak Load Forecasting in Thailand using Multiple Kernel Gaussian Process," in *IEEE ECTI-CON*, 2014, pp. 210-213.

[6] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning*. Cambridge, Mass.: MIT Press, 2006.

[7] S. Bochner, *Lectures on Fourier integrals; with an author's supplement on monotonic functions, Stieltjes integrals, and harmonic analysis*: Princeton University Press, 1959.

[8] A. G. Wilson and R. P. Adams, "Gaussian process covariance kernels for pattern discovery and extrapolation " in *ICML Conference*, 2013.