Quantile regression model for electricity peak demand forecasting: Approximation by local triangular distribution to avoid blackouts*

Niematallah Elamin† and Mototsugu Fukushige‡

Abstract: Electricity peak demand forecasting is a key exercise undertaken to avoid power blackouts and system failure. In this paper, the next day's peak load demand is estimated and forecasted. The challenge is to generate a peak demand forecast that avoids the risk of a power blackout. We approximate the upper bound for the electricity demand utilizing estimated quantiles by quantile regression and triangular distribution. The upper bounds constructed are compared with the actual electricity demand. The proposed method successfully constructs the upper bound to avoid underprediction; i.e., it avoids the risk of power blackouts.

Keywords: Electricity peak demand, Quantile regression, Triangular distribution, Blackout.

---

* This research was financially supported by the Challenging Exploratory Research in the Grants-in-Aid for Scientific Research by the Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number JP15K12461.

† Doctoral student、Graduate School of Economics、Osaka University,
1-7, Machikaneyama-cho, Toyonaka,Osaka 560-0043, Japan.
Email address: niema.ehaimer@gmail.com.
‡ Professor, Graduate School of Economics, Osaka University.
1-7, Machikaneyama-cho, Toyonaka,Osaka 560-0043, Japan.
Email address: mfuku@econ.osakaa-u.ac.jp.

1. Introduction

Short-term load forecasting—i.e., the prediction of the load demand over an interval of up to one day—is a crucial process in managing the production, transmission, and distribution of electricity in an efficient way. As the basis of power system planning, electricity demand forecasts are needed to ensure system and supply reliability. Because electricity cannot be stored, flexibility is critical; i.e., load generation is required that has the ability to change levels rapidly in order to balance production and consumption and, therefore, to avoid system failure.

As mentioned in Almeshaiei and Sultan (2011), because of the increasing effects of changing environmental and human behavior, electric power consumption is growing rapidly and becoming more random. This increases the necessity for accuracy in forecasts. Accurate forecasts can avoid wasting energy and can prevent system failure; the former is the result of overprediction, whereas the latter arises from underprediction of demand. Furthermore, Haida and Muto (1994) stated that operating costs are increased by forecasting errors, whether positive or negative.

Peak demand forecasting is a fundamental task for ensuring the availability of sufficient supply. From an operational point of view, the key question is whether there will be any problems in meeting peak demand, as failure to meet peak demand can result in power blackouts (McSharry et al., 2005). Thus, the consequences of underprediction of demand go beyond the additional costs incurred by having insufficient capacity and the potentially serious problems of meeting load demand. Power blackouts are a critical threat, disturbing the smooth running of the economy and weakening business reliability by restricting business operations in all sectors. In order to avoid power blackouts, a model that is capable of precluding underprediction is required.

In this paper, we utilize quantile regression point forecasts to construct estimates of the upper limits for electricity demand. We attempt to forecast the 1.00 quantile point, which represent the upper limit of demand, by introducing a local triangular distribution. In the remainder of the paper, we will use “QRM” as the abbreviation for “Quantile Regression Model” and abbreviate the fitted or estimated quantile points; for example, the 0.99 quantile point is written as $\widehat{y^{(0.99)}}$.

In Section 2, we discuss the methods that are used in the electricity forecasting literature. In Section 3, we introduce the quantile regression method and propose an approximating method for the 1.00 quantile point using local triangular distribution. In Section 4, we evaluate the forecasting performances of the proposed models, using the electricity demand for the Tokyo Electric Power Company Holdings, Inc. (TEPCO), which provides the electricity for the Kanto area of Japan. In Section 5, we compare and discuss the implications of the results and provide concluding remarks.

2. Literature review

Over time, different forecasting techniques have been developed to model electricity loads, most of which focus on point forecasts. According to Hong and Fan (2016), over the past 100 or more years, both research efforts and industry practices in this area have focused primarily on point load forecasting. Weron (2007) and Taylor and McSharry (2007) provide an overview of common methods used in the literature. Many different techniques, such as multiple linear regression, the Box–Jenkins approach, and Artificial Neutral Networks (ANN) have been applied to short-term load forecasting in the past. A comprehensive review of the literature is provided by Hong (2010). Previous studies that have focused on point forecasts have attempted either to compare several forecasting techniques or to prove the superiority of their proposed technique. Engle, Brown, and Stern (1988) compare a variety of adaptive structural forecasting methods for electricity sales. Dash, Liew, and Rahman (1995) compare fuzzy neural networks for the generation of daily average and peak load forecasts. Sadownik and Barbosa (1999) compare dynamic nonlinear models and econometric models. Amjady (2001) propose an approach that incorporates ARIMA with the knowledge of experienced human operators.

Soares and Medeiros (2008) describe a forecasting model for hourly electricity loads that constructs a different model for each hour of the day, each model being based on a decomposition of the daily series of each hour into a deterministic component and a stochastic component. Ohtsuka, Oga, and Kakamu (2010) propose a spatial autoregressive (SAR) ARMA model that can simultaneously be used to capture spatial heterogeneity and spatial correlation. Whereas the abovementioned studies focus on point forecasts, Xiong, Bao, and Hu (2014) investigate interval forecasting of electricity demand. García-Ascanio and Maté (2010) apply interval time series to produce forecasts of power demand.

Quantile regression has not received much attention from the load forecasting community over the past 30 years (Hong and Fan, 2016). In addition, Juban et al. (2016) state that quantile regression has rarely been applied in the area of probabilistic energy forecasting. Probabilistic load forecasting provides additional information on the variability and uncertainty of future load values and can be in the form of quantiles, intervals, or density functions (Hong and Fan, 2016).

In contrast to previous studies, in this study, we employ quantile regression to construct point forecasts of peak demand, utilizing estimated quantile points and an approximation method using triangular distribution. Previous studies of electricity forecasting that have applied quantile regression differ from ours. For example, Gibbons and Faruqui (2014) develop a new estimation procedure, optimal forecast quantile regression (OFQR), to forecast annual peak load demand. Their estimation procedure establishes a loss function framework that uses only annual peak days to estimate the optimal quantile for the model, whereas all days are used to estimate the coefficients of the regression itself. Hong, Maciejowska, Nowotarski, and Weron (2014) propose a methodology for computing interval forecasts of electricity demand, which applies a quantile regression averaging (QRA) technique to a set of independent expert point forecasts. QRA is a forecast combination approach used to compute prediction intervals. It involves applying quantile regression to the point forecasts of a small number of individual forecasting models or experts. It assigns weights to individual forecasting methods and combines them to yield forecasts of chosen quantiles. Liu, Nowotarski, Hong, and Weron (2015) generate probabilistic load forecasts by performing QRA on a set of sister point forecasts. Sister forecasts are predictions generated from the same family of models. Thus, their proposed methodology involves two steps: first, generating a set of sister load forecasts; and, second, applying QRA to the sister forecasts.

Model performance is usually assessed by forecasting accuracy measures. In short-term load forecasting, root mean squared error (RMSE) and mean absolute percentage error (MAPE) are generally used to present load forecasting error (Munoz et al., 2010). Numerous performance assessment criteria are proposed and applied in the literature. In the current study, models are assessed based on their ability to preclude underprediction.

3. Forecasting method

3.1. Quantile regression

Koenker and Bassett (1978) introduced QRM, which models conditional quantiles as functions of predictors. The linear regression model specifies the change in the conditional mean of the dependent variable associated with a change in the covariates, whereas the QRM specifies changes in the conditional quantile (Hao and Naiman, 2007). According to Bremnes (2004), one of the main advantages of quantile regression is that the shape of the distribution does not have to be specified and that any information about these distributions can easily be included in the models. That is, quantile regression relaxes linear regression assumptions, and thus it produces flexible, nonsensitive estimates, properties that are not found in the linear regression models. Quantile regression is computationally simpler than constructing linear regression prediction intervals, as quantile regression is simply a point prediction made using the estimated coefficients and the values of the explanatory variables, without the need for calculating standard errors of the prediction.

Following Koenker and Bassett (1978), the QRM can be expressed as:

$$y_i = \beta_0^{(P)} + \beta_1^{(p)} x_i + \varepsilon_i^{(p)} \ ,$$

where p is the pth quantile estimated, and $0 < p < 1$ indicates the proportion of the population having scores below the quantile at p. The QRM minimizes a sum that gives asymmetric penalties $(1- p)| y_i - \hat{y}_i |$ for overprediction and $p\ | y_i - \hat{y}_i |$ for underprediction. The pth quantile regression estimators $\beta_0^{(P)}$ and $\beta_1^{(p)}$ are chosen to minimize:

$$\sum_{i=1}^{n} d_p(y_i, \hat{y}_i) = p \sum_{y_i \geq \beta_0^{(P)} + \beta_1^{(p)} x_i} | y_i - \beta_0^{(P)} - \beta_1^{(p)} x_i |$$

$$+ (1-p) \sum_{y_i \leq \beta_0^{(P)} + \beta_1^{(p)} x_i} | y_i - \beta_0^{(P)} - \beta_1^{(p)} x_i |, \qquad (1)$$

where d_p is the average weighted distance between y_i and $\hat{y}_i$. As a result, if $p = 0.95$, for example, there is a higher penalty for underprediction $(0.95|y_i - \hat{y}_i|)$, and a much lower penalty $(0.05|y_i - \hat{y}_i|)$ for overprediction. Therefore, the QRM will minimize the positive residuals that are caused by underprediction, accordingly minimizing power blackouts. Because any quantile can be used, it is possible to model any predetermined position of the distribution. Thus, we choose positions that are tailored to the needs of our specific inquiries; i.e., adoption of an approximation method involving triangular distribution. Therefore, we choose to estimate quantiles in pairs, 0.99 and 0.98, or 0.99 and 0.97, for example, and to construct the 1.00 quantile on this basis.

3.2. Approximation by local triangular distribution

To avoid electricity blackouts, we should construct the estimates of the upper bound for electricity demand. Of course, the QRM can provide 0.99 or higher quantiles when the sample size is relatively large. However, this method cannot provide the 1.00 quantile because of the property of its objective function (1).

In this paper, we introduce the assumption that the upper tail distribution of the error terms can be approximated by triangular distribution and then attempt to construct the 1.00 quantile. As Figure 1 indicates, when we obtain a pair of 0.99 and 0.98 quantile points, or 0.99 and 0.97, and assume the local triangular distribution, we can estimate the 1.00 quantile point.

Insert Figure 1 about here

For example, using the deviation between the 0.99 and 0.98 quantile points and assuming the triangular distribution, we can calculate the 1.00 quantile point as follows:

$$\mathrm{y}^{\widehat{(1.00)}}{}_{99\&98} = \mathrm{y}^{\widehat{(0.99)}} + \left(1 + \sqrt{2}\right) * \left(\mathrm{y}^{\widehat{(0.99)}} - \mathrm{y}^{\widehat{(0.98)}}\right).$$

This relation is easy to obtain by applying the rule for determining the area of the right triangle. When we utilize the deviation between the 0.99 and 0.97 quantile points, the 1.00 quantile point is given by:

$$\mathrm{y}^{\widehat{(1.00)}}{}_{99\&97} = \mathrm{y}^{\widehat{(0.99)}} + \left(1 + \sqrt{3}\right) * {\left(\mathrm{y}^{\widehat{(0.99)}} - \mathrm{y}^{\widehat{(0.97)}}\right)}\Big/{2},$$

and when the deviation between the 0.99 and 0.95 quantile points is used, the 1.00 quantile point is given by:

$$\mathrm{y}^{\widehat{(1.00)}}{}_{99\&95} = \mathrm{y}^{\widehat{(0.99)}} + \left(1 + \sqrt{5}\right) * {\left(\mathrm{y}^{\widehat{(0.99)}} - \mathrm{y}^{\widehat{(0.95)}}\right)}\Big/{4}.$$

From the deviation between the 0.95 and the 0.90 quantile points, the 1.00 quantile point is given by:

$$\mathrm{y}^{\widehat{(1.00)}}{}_{95\&95} = \widehat{\mathrm{y}^{(95)}} + \left(1 + \sqrt{2}\right) * \left(\mathrm{y}^{\widehat{(0.95)}} - \mathrm{y}^{\widehat{(0.90)}}\right).$$

We consider that the 1.00 quantile point is an estimate of the upper limit of the electricity peak demand.

4. Empirical comparison

To check the usefulness of the proposed method, we will apply the method to the actual electricity demand data. In additional, we alter the estimation period and the in-sample and out-of-sample forecasting periods to check the robustness of the proposed method's practical performance.

4.1. Data

Peak demand is the highest load observed during a (short) unit of time. In this paper, we obtain hourly electricity demand for TEPCO and adopt the highest hourly electricity demand in 10,000 kilowatts between 0:00 and 23:59 hours as the daily peak demand. On June 10, 2015, we downloaded the hourly demand from January 1, 2008 to December 31, 2015 from the website of TEPCO (http://www.tepco.co.jp/forecast/html/download-j.html). However, on March 11, 2011, the East Japan Great Earthquake and tsunami occurred. In its aftermath, some nuclear power plants (including the Fukushima first and second power plants) were shut down, and planned electricity shutdowns took place. The aftermath of the event itself, and the consequent self-restraint demand adjustments that TEPCO requested of its customers, affected electricity demand. The demand data points affected are not suitable for constructing a forecasting model, so we only use data from after January 1, 2012 to construct our daily peak demand data series. Thus, the models are estimated using daily peak demand data from 2012 through 2014. Model performance is evaluated for the in-sample period (2012–2014) and an out-of-sample period (2015).

Insert Figure 2 about here

Figure 2 shows strong seasonal fluctuations, with daily peak demand being higher in summer than in winter. The highest yearly peak is observed during August, whereas the lowest peak appears during May. In addition, weekends show a lower peak demand than weekdays.

National holidays have a load pattern that differs from working days. In Japan, some of the public holiday dates change every year, based on the Happy Monday System, which refers to a set of modifications to Japanese laws in 1998 and 2001 to move a number of public holidays to Mondays, creating three-day weekends for those with five-day work weeks. Table 1 shows the national holidays of Japan.

Insert Table 1 about here

The data do not exhibit either an upward or downward trend but show a high degree of multiple seasonality (based on the day of the week and month of the year effects).

4.2. Model selection

In this section, models are estimated using daily peak electricity demand from January 1, 2012 to December 31, 2014 (the in-sample period). Static forecasts are constructed by substituting realized values of the previous period's peak demand into the explanatory variable in the estimated regression, and are then generated for the period January 1, 2015 to December 31, 2015 (the out-of-sample period).

In regard to the dependent variable, logarithms of the series are modeled in order to reduce the effect of heteroskedasticity that may be present because of the characteristics of the data set. First, daily peak demand is expressed as a function of the previous day's peak demand, and sets of fixed effects for weekends are estimated using the linear regression model utilizing ordinary least squares (OLS). Additional information on the past daily average load, as well as additional dummy variables to account for the month-of-year effect, were introduced into the model and then removed because of their lack of contribution to improving the model's diagnostics and performance, as measured by the AIC, the SBC, and the standard error (S.E.) of the regression. Then, we chose a model with dummy variables for the day of the week for Saturday (SAT), Monday (MON), and Tuesdays after Happy Mondays (HTuesday), and national holiday dummy variables (Holiday1), which had the best diagnostics. The following models are estimated by OLS:

$$\mathrm{Log}(\mathrm{Peak}_t) = \beta_0 + \beta_1 \log(\mathrm{Peak}_{t-1}) + \beta_2 \mathrm{SAT} + \beta_3 \mathrm{MON} + \beta_4 \mathrm{HTuesday} + \beta_5 \mathrm{Holiday1}$$

Insert Table 2 about here

Here, the Holiday1 dummy variable takes a value of one only for the national holidays listed in Table 1, and a value of zero for other days. For the dummy variable for holidays, we have two other candidates: Holiday2, which takes a value of one for holidays listed in Table 1 and January 2 and 3, which are not national holidays but are traditional holidays in Japan; and Holiday3, which takes a value of one during the period August 12–16, in addition to the Holiday2 dates, because most Japanese companies set these days as summer holidays. In Table 2, the OLS estimation results are reported only for the case that includes Holiday1. When Holiday2 or Holiday3 is adopted as the independent variable, its estimated coefficient is statistically significant, but its t-value is less than that for Holiday1. However, in the results for the 0.99 quantile of the QRM in Table 2, the estimated coefficients for Holiday1 and also for Holiday2 and Holiday3 are statistically insignificant at the 5% level. Therefore, in the rest of the paper, we estimate the following model using the QRM:

$$\log(\mathrm{Peak}_t) = \beta_0 + \beta_1 \log(\mathrm{Peak}_{t-1}) + \beta_2 \mathrm{SAT} + \beta_3 \mathrm{MON} + \beta_4 \mathrm{HTuesday}.$$

In Table 3, the estimation results for the 0.98, 0.97, 0.95, and 0.90 quantile models are reported.

Insert Table 3 about here

Before proceeding to the evaluation of the proposed method, we need to check the performance of the simple forecasting by the QRM. In Table 4, we count the underestimated cases and calculate their percentages.

Insert Table 4 about here

Because the forecasts by the QRM are simply quantile points below one, underestimated cases occurred around the quantiles rates. In other words, we cannot avoid shutdowns owing to shortages in meeting demand. To avoid shutdowns, we must construct the 1.00 quantile point.

4.3. Forecasting performance

In Table 5, we report the numbers and percentages for the underestimated cases for construction of the 1.00 quantile point, approximated by triangular distribution for the in-sample and out-of-sample periods.

Insert Table 5 about here

Except for the 1.00 quantiles constructed from the 0.95 and 0.90 pairs, underestimated cases are less than 0.1% for the in-sample period and are zero for the out-of-sample period in all approximations. From these results, we conclude that we have successfully estimated the upper limits of the daily peak demand. However, if this method is applied for actual peak demand forecasting, we should check the efficiencies of the proposed method. In Table 6, we calculate the average rates of overestimation for the in-sample and out-of-sample periods.

Insert Table 6 about here

As for the demonstration rates of the forecasting, the 1.00 quantiles constructed from the 0.95 & 0.90 pairs result in a 13.59% overestimation compared with the actual demand, but their performance in terms of underestimation cases is relatively poor. (Table 6) The 1.00 quantiles constructed from the 0.99 & 0.98, 0.99 & 0.97, and 0.99 & 0.95 pairs perform similarly, but the 1.00 quantile estimated from the 0.99 & 0.98 pair results in an overestimate of 18.34% compared with the actual demand, and its rate is smallest compared to other pairs. At this stage, we can conclude that constructing the 1.00 quantile from the 0.99 and 0.98 quantiles is sufficient to estimate the upper limit for the daily peak demand. Of course, this 18.34% overestimation should be considered from other points of view, including its economic and financial consequences. In addition, we should consider another candidate for the independent variables to improve the forecasting models.

4.4. Robustness

To check the robustness of the proposed method, we calculate the underestimated cases with several combinations of estimation periods and forecasting periods. Because of the data availability, which is from April 1, 2012 to March 31, 2015, we set estimation periods of one-year lengths (2012, 2013, 2014) and two-year lengths, and we set the forecasting periods one or two years ahead and one or two years back. In Table 7, we report the numbers and percentages of underestimated cases in each year.

Insert Table 7 about here

As for the 1.00 quantiles constructed from the 0.95 & 0.90 pairs, in most setups, the underestimated cases are over 1.0%. As discussed in subsection 4.3, these 1.00 quantiles are not practical. As for the 1.00 quantiles constructed from the 0.99 & 0.98 pairs, their underestimated cases for the out-of-sample period are zero, except for performance in 2014, for almost all of the setups. Thus, the proposed method successfully constructs the upper limits of the peak demand, and its robustness in construction of the upper limits is confirmed. However, for 2014 in particular, for the out-of-sample period, the performances of the 1.00 quantiles are relatively poor in all estimation period setups. This indicates that there may have been some days with abnormally high demand in 2014. The reason that such days exist is beyond the scope of this research.

5. Conclusion

Assuming that, in practice, the most crucial issue is to prevent system failure and to eliminate power blackouts, this article aimed to generate a model that is capable of precluding underprediction. 1.00 quantiles constructed by the quantiles estimated by the QRM were compared with the actual demand to investigate the proposed method based on its ability to avoid power blackouts (i.e., to avoid underprediction of demand). From the empirical comparison, we can conclude that we successfully constructed the upper limits of the daily peak demand. Additional robustness checks validated these results. From the performance in relation to the overestimation rates, at this stage, we can state that constructing the 1.00 quantile from the 0.99 and 0.98 quantiles is sufficient to estimate the upper limit for the daily peak demand. Two problems remain to be solved: first, whether the proposed method is useful from an economic or financial perspective; and second, whether we should consider further candidates for the independent variables to improve the forecasting models.

References

Almeshaiei, E., Soltan, H. 2011. A methodology for electric power load forecasting. *Alexandria Engineering Journal* **50**: 137–144. doi: 10.1016/j.aej.2011.01.015.

Amjady, N. 2001. Short-term hourly load forecasting using time-series modeling with peak load estimation capability. *IEEE Transactions on Power Systems* **16(3)**: 498–503. doi: 10.1109/59.962429.

Bremnes, J. B. 2004. Probabilistic wind power forecasts using local quantile regression. *Wind Energy* **7**: 47–54. doi: 10.1002/we.107.

Dash, P. K., Liew, A. C., Rahman, S. 1995. Comparison of fuzzy neural networks for the generation of daily average and peak load profiles. *International Journal of System Science* **26**: 2091–2106. doi: 10.1080/00207729508929156.

Engle, R. F., Brown, S. J., Stern, G. 1988. A comparison of adaptive structural forecasting methods for electricity sales. *Journal of Forecasting* **7**: 149–172. doi: 10.1002/for.3980070302.

García-Ascanio, C., Maté, C. 2010. Electric power demand forecasting using interval time series: A comparison between VAR and iMLP. *Energy Policy* **38(2)**: 715–725. doi: 10.1016/j.enpol.2009.10.007.

Gibbons, C., Faruqui, A. 2014. Quantile regression for peak demand forecasting. Available at http://ssrn.com/abstract=2485657.

Haida, T., Muto, S. 1994. Regression-based peak load forecasting using a transformation technique. *IEEE Transaction on Power Systems* **9 (4)**: 1788–1794. doi: 10.1109/59.331433.

Hao, L., Naiman, Q. D. 2007. *Quantile Regression*, SAGE publications: Thousand Oaks.

Hong, T. 2010. Short-term electric load forecasting. PhD dissertation, North Carolina State University.

Hong, T., Fan, S. 2016. Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting* **32**: 914–938. doi: 10.1016/j.ijforecast.2015.11.011.

Hong, T., Maciejowska, K., Nowotarski, N., Weron, R. 2014. Probabilistic load forecasting via quantile regression averaging of independent expert forecasts. Hugo Steinhaus Center, Wrocław University of Technology http://www.im.pwr.wroc.pl/~hugo/RePEc/wuu/wpaper/HSC_14_10.pdf

Juban, R., Ohlsson, H., Maasoumya, M., Poirier, L., Kolter, J. Z. 2016. A multiple quantile regression approach to the wind, solar, and price tracks of GEFCom2014. *International Journal of Forecasting* **32**: 1094–1102. doi: 10.1016/j.ijforecast.2015.12.002.

Koenker, R., Bassett, G. 1978. Regression quantiles. *Econometrica* **46(1)**: 33–50.

Liu, B., Nowotarski, J., Hong, T., Weron, R. 2016. Probabilistic load forecasting via quantile regression averaging on sister forecasts. *IEEE Transactions on Smart Grid* (in press).

McSharry, P. E., Bouwman S., Bloemhof, G. 2005. Probabilistic forecast of the magnitude and timing of peak electricity demand. *IEEE Transaction on Power Systems* **20(2)**: 1166–1172. doi: 10.1109/TPWRS.2005.846071.

Munoz, A., Sanchez-Ubeda, E. F., Cruz, A., Marin, J. 2010. Short-term forecasting in power systems: a guided tour. *Energy Systems* **2**: 129–160. doi: 10.1007/978-3-642-12686-4_5.

Ohtsuka, Y., Oga, T., Kakamu, K. 2010. Forecasting electricity demand in Japan: A Bayesian spatial autoregressive ARMA Approach. *Computational Statistics & Data Analysis* **54(11)**: 2721–2735.

Sadownik, R. Barbosa, E. P. 1999. Short-term forecasting of industrial electricity consumption in Brazil. *Journal of Forecasting* **18**: 215–224. doi: 10.1002/(SICI)1099-131X(199905)18:3<215::AID-FOR719>3.0.CO;2-B.

Soares, L. J., Medeiros, M. C. 2008. Modeling and forecasting short-term electricity load: A comparison of methods with an application to Brazilian Data. *International Journal of Forecasting* **24**: 630–644. doi: 10.1016/j.ijforecast.2008.08.003.

Taylor, J. W., McSharry, P. E. 2007. Short-term load forecasting methods: An evaluation based on European data. *IEEE Transaction on Power Systems* **20(2)**: 1166–1172. doi: 10.1109/TPWRS.2007.907583.

Weron, R. 2007. *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. Wiley: Chichester.

Xiong T., Bao, Y., Hu, Z. 2014. Interval forecasting of electricity demand: A novel bivariate EMD-based support vector regression modeling framework. *International Journal of Electrical Power and Energy Systems.* doi: 10.1016/j.ijepes.2014.06.01. Accepted.

Table 1. Japan's national holidays

Holiday dates	
January 1	Third Monday of July
Second Monday of January	August 11 [1)]
February 11	Third Monday of September
March 20 or 21	September 22 or 23
April 29	Second Monday of October
May 3	November 3
May 4	November 23
May 5	December 23

Note: 1) This holiday commenced in 2016.

Table 2. Estimation results for the model selection

	OLS		Quantile Regression (0.99 quantile)							
			Model 1		Model 2		Model 3		Model 4	
	Coefficients	t-value	Coefficients	t-value	Coefficients	t-value	Coefficients	t-value	Coefficients	t-value
Constant	0.471	4.83	0.747	1.53	0.740	1.32	0.528	0.97	0.747	1.45
Log ($Peak_{t-1}$)	0.942	79.67	0.923	15.48	0.924	13.53	0.950	14.37	0.923	14.74
Sat	–0.067	–15.93	–0.070	–5.93	–0.070	–5.14	–0.065	–4.58	–0.070	–5.36
Mon	0.108	24.29	0.121	6.85	0.121	6.08	0.125	6.37	0.121	7.22
HTuesday	0.937	6.76	0.052	2.14	0.052	2.15	0.055	2.12	0.052	2.29
Holiday1	–0.032	–4.20	–0.031	–1.03	-	-	-	-	-	-
Holiday2	-	-	-	-	0.001	0.005	-	-	-	-
Holiday3	-	-	-	-	-	-	0.012	0.57	-	-
R^2	0.8602		0.4809		0.4766		0.4790		0.4763	
S.E. of regression	0.04827		-		-		-		-	

Note: The R^2 for the OLS is the adjusted R^2, whereas for the quantile regressions, it is a pseudo R^2.

Table 3. Results of other quantile regressions

	0.98 quantile		0.97 quantile		0.95 quantile		0.90 quantile	
	Coefficients	t-value	Coefficients	t-value	Coefficients	t-value	Coefficients	t-value
Constant	1.006	1.59	0.332	0.58	–0.069	–0.22	–0.095	–0.51
Log ($Peak_{t-1}$)	0.890	11.68	0.970	13.97	1.016	26.94	1.016	45.06
Sat	–0.068	–4.03	–0.047	–2.33	–0.060	–3.55	–0.066	–7.28
Mon	0.104	5.73	0.119	7.46	0.119	8.93	0.128	14.52
HTuesday	0.064	2.48	0.089	2.33	0.085	2.74	0.069	2.80
R^2	0.5111		0.5340		0.5684		0.6166	

Note: In each year, data from April 1 to March 31 are used for estimations or forecasting. The R^2 for the quantile regressions is a pseudo R^2.

Table 4. Underestimated forecasts with 0.95, 0.97, 0.98, and 0.99 quantile regressions

Quantiles	In-sample		Out-of-sample	
	#	%	#	%
0.99	14	1.278	1	0.273
0.98	25	2.283	5	1.369
0.97	35	3.196	9	2.465
0.95	52	4.748	25	6.849

Note: In each year, data from April 1 to March 31 are used for estimation or forecasting. The total number of observations for the in-sample period is 1,095.

Table 5. Underestimation by forecasting with triangular approximation

Year	In or out of sample	0.99 & 0.98		0.99 & 0.97		0.99 & 0.95		0.95 & 0.90	
		#	%	#	%	#	%	#	%
2012	In-sample	0	0.00	0	0.00	0	0.00	2	0.54
2013	In-sample	0	0.00	0	0.00	0	0.00	6	1.64
2014	In-sample	1	0.27	2	0.54	1	027	3	0.82
From 2013 to 2014		1	0.09	2	0.18	1	0.09	11	1.00
2015	Out-of-sample	0	0.00	0	0.00	0	0.00	1	0.27

Note: # and % indicate the number and percentage of underestimations, respectively.

Table 6. Ratios of overestimation to actual electricity demand

Year	In or out of sample	0.99 & 0.98 %	0.99 & 0.97 %	0.99 & 0.95 %	0.95 & 0.90 %
2012	In-sample	18.45	18.05	18.07	13.89
2013	In-sample	18.41	18.47	18.54	13.91
2014	In-sample	18.39	18.98	19.07	13.73
From 2013 to 2014		18.41	18.50	18.56	13.84
2015	Out-of-sample	18.34	19.83	19.99	13.59

Note: The symbol % indicates the percentage of overestimations compared with actual electricity demand.

Table 7. Robustness check using other estimations and forecasting periods

Year	In or out of sample	0.99 & 0.98		0.99 & 0.97		0.99 & 0.95		0.95 & 0.90	
		#	%	#	%	#	%	#	%
Estimation period: From Jan 1, 2013 to Dec 31, 2014 (two years)									
2012	Out-of-sample	0	0.00	0	0.00	0	0.00	2	0.54
2013	In-sample	1	0.27	0	0.00	0	0.00	7	1.91
2014	In-sample	3	0.82	3	0.82	2	0.54	2	0.54
2015	Out-of-sample	0	0.00	0	0.00	0	0.00	2	0.54
Estimation period: From Jan 1, 2014 to Dec 31, 2014 (one year)									
2012	Out-of-sample	0	0.00	0	0.00	1	0.27	2	0.54
2013	Out-of-sample	0	0.00	0	0.00	0	0.00	5	1.36
2014	In-sample	1	0.27	1	0.27	3	0.82	3	0.82
2015	Out-of-sample	0	0.00	0	0.00	0	0.00	1	0.27
Estimation period: From Jan 1, 2012 to Dec 31, 2013 (two years)									
2012	In-sample	1	0.27	1	0.27	2	0.54	5	1.36
2013	In-sample	3	0.82	2	0.54	2	0.54	10	2.73
2014	Out-of-sample	5	1.36	3	0.82	2	0.54	7	1.91
2015	Out-of-sample	0	0.00	0	0.00	0	0.00	8	2.19
Estimation period: From Jan 1, 2011 to Dec 31, 2013 (one year)									
2012	Out-of-sample	1	0.27	2	0.54	1	0.27	5	1.36
2013	In-sample	3	0.82	4	1.09	1	0.27	7	1.91
2014	Out-of-sample	3	0.82	3	0.82	4	1.09	3	0.82
2015	Out-of-sample	1	0.27	1	0.27	1	0.27	3	0.82
Estimation period: From Jan 1, 2012 to Dec 31, 2012 (one year)									
2012	In-sample	1	0.27	2	0.54	2	0.54	4	1.09
2013	Out-of-sample	2	0.54	2	0.54	2	0.54	7	1.91
2014	Out-of-sample	2	0.54	3	0.82	3	0.82	5	1.36
2015	Out-of-sample	0	0.00	0	0.00	0	0.00	3	0.82

Note: # and % indicate the number and percentage of underestimations, respectively.

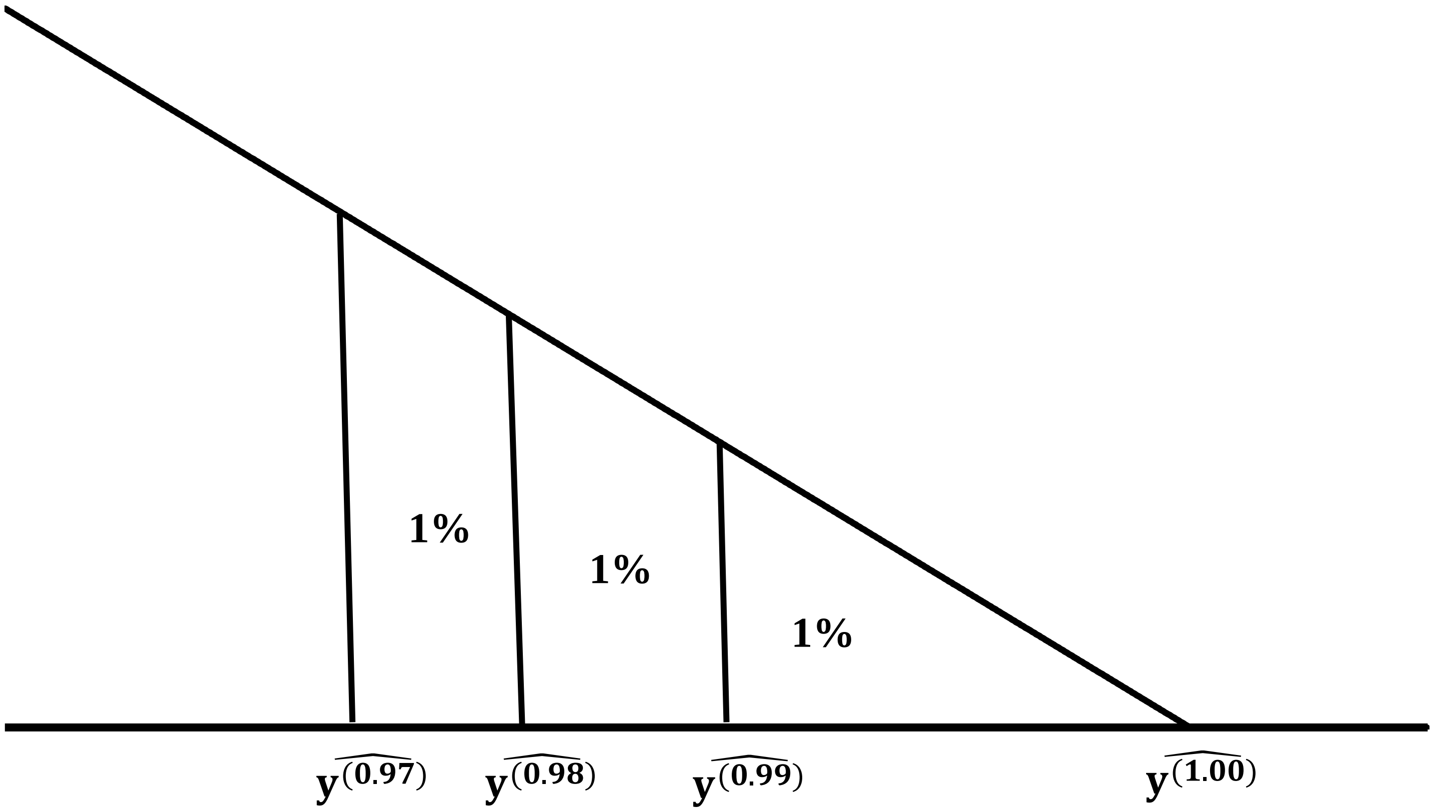

Figure 1. Quantile points and approximation with triangular distribution

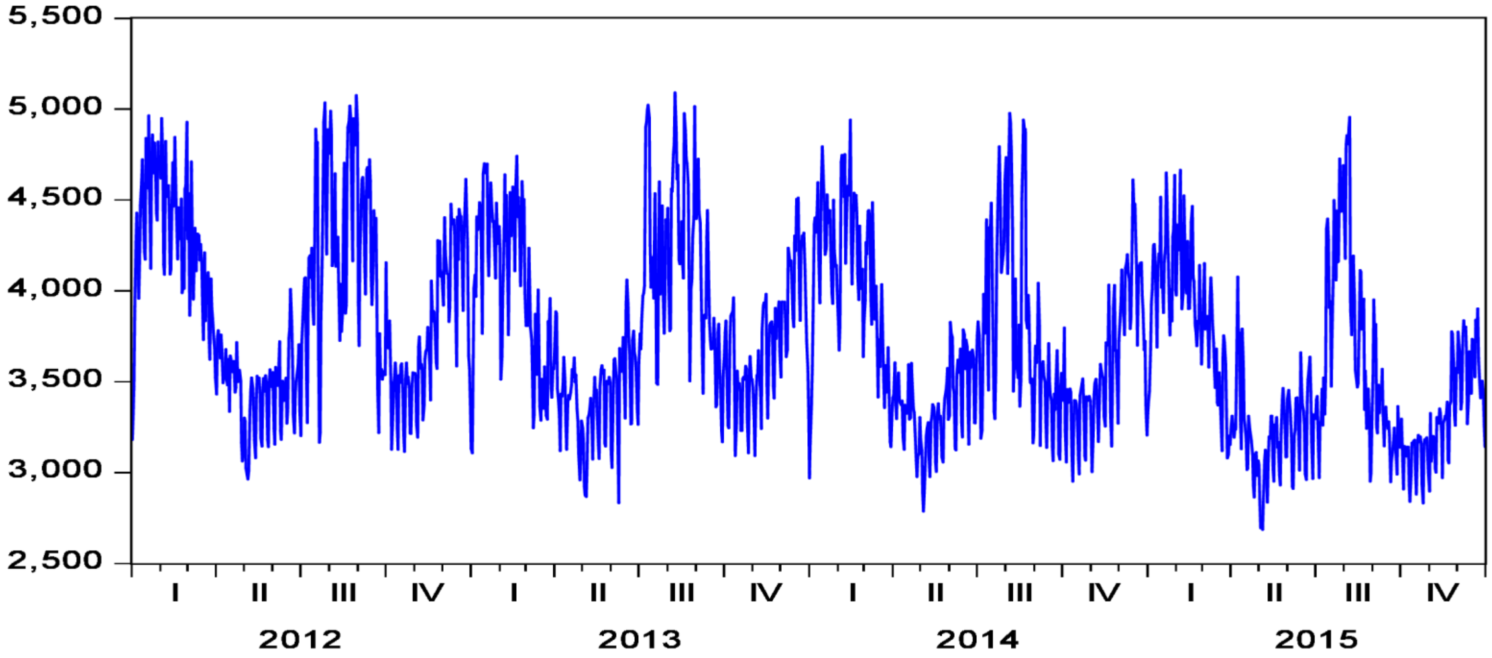

Figure 2. Daily peak demand from 2012 to 2015 (in 10,000 kilowatts)