Improvement of performance of short term electricity demand model with meteorological parameters

Kamal Chapagain[1,2,*], Tomonori Soto[2,], Somasak Kittipiyakul[1]
[1]School of Information Communication Technology, SIIT, Thammasat University, PathumThani-12000, Thailand
[2]Graduate School of Environmental Science, Hokkaido University, Sapporo, Japan
Corresponding Email: *kamal.chapagain02@gmail.com

***Abstract*— The accuracy of short term electricity demand forecasting is essential for operation and trading activities on energy market. This paper considers a parsimonious forecasting model to explain the importance of sophisticated weather parameters for hourly electricity demand forecasting. Temperature is the major factor that directly influence electricity demand, but what about the affect of other weather factors such as relative humidity, wind speed etc. on short term electricity demand forecasting, is the prime research question and this paper analyzed it quantitatively. We demonstrate three different multiple linear models including auto-regressive moving average ARMA (2,6) models with and without some exogenous weather variables to compare with performance for Hokkaido Prefecture, Japan. Since, Bayesian approach is used to estimate the weight of each variables with Gibbs sampling, it generates the weight of coefficients in terms of distribution as our interest. The performance of each models for complete one year out sample prediction shows that the average improvement of hourly forecast by 1 to 2 % can be achieve by including such weather factors.**

I. INTRODUCTION

ACCURATE short term load forecasting is important and crucial because electricity distribution grid requires proper balancing between electricity supply from production Company and consumption at every instant. This electricity load if not balance, it affects the system with unnecessary blackouts or load shedding and huge loss of revenue for utility companies. So, proper forecasting helps to maintain secure power system, avoids blackout risk and provides adequate electricity supply. Short term electricity demand forecasting is important for all stakeholder of electricity- such as market operator, electricity generators, electricity retailers and ultimately for general people. For market operator, forecasting is crucial for scheduling and dispatch of generators capacity. For electricity generators, strategic choice involved in bidding and re-bidding of capacity depends on demand forecast [1]. For, electricity retailer, demand forecasting affects the decisions about the balance between hedging spot acquisition of electricity, and finally these actions help for general people consistent energy supplies without black out and possibly minimum cost. Therefore modeling and forecasting of electricity load for short horizon is alarming issue today. Various authors construct their own model and estimation process for better performance. Among various approaches for predicting future data, we can generalize into two types of estimates. i) point estimate- single valued forecast, and ii) probabilistic estimate- where each parameters are treated as random variable and several possible values for the future demand is predicted. The main advantage of probabilistic forecast is that it contains additional information in terms of uncertainty. This paper employs two stage for estimation. In first stage, the point estimated values obtained from OLS is considered as prior information for Bayesian and in next stage, these parameters are used as random variable for Markov Chain Monte Carlo (MCMC) and obtained the final values in terms of distributions for probabilistic forecast. Finally, forecasting the next day demand is in terms of mean, median and 60 percentile values. More specifically, based on literature- forecasting models are classified in three categories-

- Single equation time series model: such kind of model are based on strong seasonal patterns and provides the information based on historical demand pattern [2, 3] .
- Semi/non-parametric method: emphasize non-linearity of demand [4, 5] .
- Multiple equation time series model: these models got high attention during latest papers followed by Peirson and Henely [6], Ramanathan et al. [7], and papers [8, 9, 10, 11, 12] where each period is treated as a separate forecasting problem with its own equation.

Many authors built univariate time series model without any exogenous variables. For example, Taylor [2] employed a double seasonal exponential smoothing for half hour data to predict with mean absolute percentage error (MAPE) of 1.25 to 2 %. Only historical demand may not sufficiently addressed the cause of effect on demand because temperature variation is also an important factor that directly influenced electricity demand. Harvey and Koopman [13] used a time series model with temperature to get performance up to 3 %. Satish et al. [14] studied the effect of temperature with 20 training patterns and compared the results with univariate model having no other exogenous variables except demand data. Results showed that the performance was improved up to 55 % by including temperature variable. Satish et al. [14] also mentioned that the implementation of additional sophisticated weather variables such as precipitation, wind speed, humidity and cloud cover should yield even better results. However, their individual affects were not discussed which is our concern. Friedrich L. et al. [11] investigate the results for Abu Dhabi city electricity load using multiple weather variable for 24 hour to 48 hour prediction horizon and got very promising result of 1.5 % MAPE for 24-hour and 48-hour horizon.

To develop a mathematical model several factors that directly or indirectly influence on electricity demand have to take in account. For example- weather, calendar, and historical demand data. Impact of weather variables on electric power demand in England, Australia, Jordan and many more regions are found in literature, but mostly focused on effects of temperature. Failure of power supply in 1995 was due to the excessive demand of electricity in Malaysia because of excessive temperature increment. Such increment of demand is also possible if the temperature lowers by extreme. Countries having cold regions have the peak demand during winter and summer are usually lower compared to the peak demands of winter. This indicates that human activities during winter season is higher. Various weather variables can be considered for demand forecasting: temperature and humidity are the most commonly used, but wind, radiation and cloud cover are often excluded. The affect of meteorological factors such as temperature, humidity, solar radiation, precipitation, and wind speed varies according to the season and hence varies electricity demand significantly. However, most of the paper exclude other factors and include only temperature for their analysis.

So far we didn't get the quantitative comparison such as what is the improvement of performance if we include the weather variables. The weather variables we discussed here are wind speed, humidity, cloud coverage, precipitation. Therefore, this is our interest to analyze the effect of these variables quantitatively.

II. LITERATURE REVIEW

In literature, several models are discussed and pay attention having better performance. Electricity demand in Japan has strong time correlation with lagged dependent variable such as Ohtsuka Y. et al. [3], therefore there are several papers that relate with ARMA time series structure. Ohtsuka Y. [3], proposed univariate ARMA with Bayesian estimation procedures and got good performance. Moreover, ARIMA process can perform good forecasting for non-stationary mean by a differencing process [15] and was the most popular model for forecasting. [10] presented a model for hourly electricity load forecasting state space model based on stochastically time-varying process which were designed to account for changes in customer behavior and in utility production efficiency. Multivariate sophisticated weather variables such as, temperature, precipitation, wind speed, cloud cover, humidity are employed for modeling electricity consumption load in [8,10]. This paper investigates the individual effects of these weather variables. However, Talylor et al. [17] uses univariate model that includes only historical observations of electricity data, but predict well. But, for the robust and reliable performance, many papers follow classical approach for short-term forecasting normally employ regression methods with dummy variables for example Ramanathan et al. [7], Cancelo et al. [16], Cottet et al. [8], Soares et al. [9]. Therefore, this paper follow similar strategy for the construction of multivariate mathematical model based on including and excluding some weather parameter with types of day, months, interaction of temperature with months as a categorical variables.

In literature, many authors develop univariate time series model without any exogenous variables with competitive forecasting performance. For example, Taylor [2] employed a double seasonal exponential smoothing for half hour data to predict very good result with mean absolute percentage error (MAPE) of 1.25 to 2 %. However, only historical demand data set may not sufficiently address the cause of effect on demand because temperature variation is also an important factor that directly influence electricity demand. After 2003, climate change significantly affect the variation on demand such as modification of annual daily load curve, shifting of the peak demand occurrence from evening to morning in Jordan [18]. In Europe, extremely high temperatures during summer of 2003 was created significantly greater electricity demand. Therefore, it is worth to specifically examine the influence of each meteorological parameters for electricity demand.

Some multivariate weather parameters such as, temperature, precipitation, wind speed, cloud cover, humidity are employed for modeling electricity consumption load [8][10]. They also mentioned that the use of additional weather variables such as precipitation, wind speed, humidity and cloud coverage should yield even better results. That means the performance is improved and consistent nature with 20 training patterns and compare the results with univariate model having no other due to such meteorological variables. Friedrich L. et al. [11] investigate the results for Abu Dhabi city electricity load using multiple weather variable for 24 hour to 48 hour prediction horizon and got very promising result of 1.5 % MAPE for 24-hour and 48-hour horizon. Apadula et al. [19] analyze weather, and calendar variables effects on monthly electricity demand using MLR model for Italy. Including good meteorological variable estimates highly improve monthly demand forecast with MAPE around 1.3 %. However, they haven't analyzed the performance including and excluding individual meteorological parameters.

Another important factor found in literature is day types. Dordonant et al. [10], Chapagain and Kittipiyakul [12], forecast the electricity load considering a normal day, and make adjustments with other dummy variable for treating as weekend or other special days which is also taken into account during modeling. But, our intention here in this paper is to analyse the improvement of performance when we include such weather variables. So far we are not getting any quantitative comparison among the weather variables, such as what is the improvement of performance if we include meteorological variables for example- wind speed, humidity, cloud coverage, precipitation. Therefore this is our interest to analyze it quantitatively.

III. DESCRIPTION OF DATA

March 25, 1878, Institute of Technology in Toranomon, Tokyo was the first day of electricity in Japan. By 1951, nine Electricity utilities - Hokkaido, Tohoku, Tokyo, Chubu, Hokuriku, Kansai, Chungoku, Shikoku and Kyushu Electric power Companies with the responsibility of supplying electricity to each region were established. According to the populations and industrialization, demand is highest in Tokyo region, whereas lowest in the Hokkaido region. The electricity produced in one region is not necessarily for only that same region, they are connected by overhead and undersea transmission line for sharing the power to other utilities companies, if they face shortages of energy. However, the Okinawa electric company supplies, established as the tenth utilities company at Okinawa Prefecture and is not connected to other companies.

In this paper, we have played with the hourly electricity demand data from Jan 1, 2013 through December 31, 2015 for Hokkaido Prefecture downloaded from official web-page of Hokkaido Electric Power Company (HEPCO) which is freely available. These 3 years of hourly electricity demand data consist 26280 samples without any missing data. Metrological dataset also available free but in limited access on public webpage of Japan Meteorological Agency (JMA). Some missing data for meteorological data are filled up with simple interpolation. Since, JMA has many stations within Hokkaido Prefecture, we are limited to the data of Sapporo station, a highly populated and commercial city of Hokkaido region.

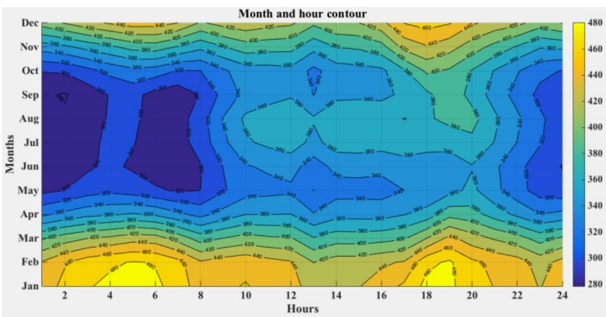

Fig.1. Trend of demand pattern: Historical observation (×10 MW)

Figure 1 demonstrates how the annual average electricity demand for each months varies hour-wise. Yellow color in contour map indicates the maximum demands upto 4800MW during morning (approx. 4 to 6 AM) and evening (approx 6 to 7 PM) at winter season, specially in January. Since, Hokkaido Prefecture suffers from very cold climate during winter season with temperature -20 °C, and use of heating appliances for house, office are excessive for warming purpose. The next reason of high demand at the morning time is due to its price which is relatively cheaper during off hours and automatic appliances set to be get on during off period. The lowest demand on the similar time period are found in summer season, specially May to September. This is exactly the opposite affect of winter and exhibits seasonal variation.

More specific explanation of electricity demand throughout 24 hours of the day is illustrate in figure 2, based on types of day. During weekdays, there is high demand of electricity due to industrial and official works where as during weekends only household electricity demand cause high variations of demand during weekdays, weekends and public holidays. Therefore, day type should be treated carefully during modeling. Various peaks of demand can be observed at morning, day and evening time period. However, day type doesn't have any effect for demand pattern, and looks similar pattern.

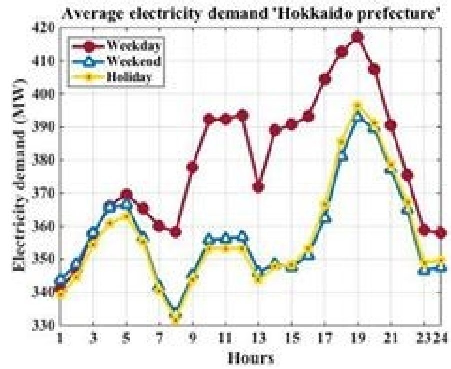

Fig.2. Day type variation of electricity demand

IV. FORECASTING HORIZON

In forecast literature, we find the range of models and methods for generating forecast values. There is no consensus in the literature as to what the thresholds should actually be, but based on their overall time horizon in literature that we found, four kind of forecasting horizons are popularly used-

- Very short-term forecasts: minutes to one hour ahead for observation of instant fluctuation of demand
- Short-term forecasts: hours to one week ahead ensuring for system stability
- Medium-term forecasts: one week to six months ahead for maintenance scheduling
- Long-term forecasts: six months to ten years ahead for capital planning

The generalized block diagram of forecasting methods in figure 3, can be consider a simple forecasting system that we follow in this paper. Based on our model formulation, historical demand, calendar and weather variables are consider as the input and parameters are estimated using two different estimation approach- first using ordinary least square (OLS) and refined it with Bayesian approach.

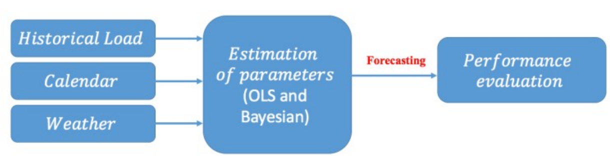

Fig.3. Generalized block for electricity demand forecasting

Most of the electricity demand forecasting methods are dedicated to short-term (hours to a few days ahead) forecasting but not as much for the long-term or intermediate terms forecasting [20]. This may be the due to the research challenge because of high variation of intraday uncertainty of demand profile for short term prediction compared to other horizon. In this paper we compare the forecasting results based on an hour ahead prediction between three models named as model A, B, and C. These models A, B, and C are discussed and developed in next two sub sections.

A. Statistical Model

In this paper, we develop multiple linear regression (MLR) model with autoregressive AR(2) inspired by the seminal paper of Ramanathan et al. [7] who developed multiple regression model with separate equations for each hour of the day. Since, sampling of the data is observed every hour, there are N=24 samples in a day. We estimate the demand for the first hour of the day with one equation and the second for the second half hour of the day from next equation and so on. Therefore, we need24 individual equations for the complete prediction of demand in one day.

$$y_T = \mu_{h(t)}(\boldsymbol{x}_t) + \epsilon_t \tag{1}$$

$$\epsilon_t = N(0, \sigma^2) \tag{2}$$

for $t = 1, \ldots, T,$ where $\mu_{h(t)}$is the regression mean function at hour $h(t)$ and ϵ_t is the random residual demand with zero mean as eq. 2. The mean function $\mu_{h(t)}$varies over intraday periods and we need to estimate the N mean functions, $\mu_1, \mu_2, \ldots, \mu_N$, separately. As an example, the mean function μ_h for a particular hour $h = h(t)$ is

$$\mu_h(\boldsymbol{x}_t) = \beta_h^T \boldsymbol{x}_t \tag{3}$$

where $\boldsymbol{x}_t = [co - variats]$ and selection of variable plays the important role to improve performance and selection of covariates determines the types of models. So, we develop three different model based on selection of variables to develop Multiple linear regression models.

- **Model A:** includes all meteorological parameters such as temperature, humidity, rainfall, snowfall, radiation, and wind speed
- **Model B:** model B is created by removing daily and hourly variables of humidity, rainfall, snowfall, solar radiation, and wind speed.
- **Model C:** Model C is created by removing some low weight covariate from model B.

Covariates for model A, B, and C can be written in column vector form and we formulate it in prototype formulation sub-section.

B. Prototype Formulation

General model for hourly demand is developed as,

$$Demand_{h,d} = Deterministic_{h,d} + Meteorology_{h,d} + HistoricalData_{h,d} + v_{h,d} \tag{4}$$

where h indicates the hour of the day, and d indicates the daily observations and $v_{h,d}$ contains the correlated error term with some order of lag data. We use some special technique to select the appropriate order of q

$$v_{d,h} = \textstyle\sum_{i=1}^{q} \rho_i \, \epsilon_{h,d-i} + \epsilon_{h,d} \tag{5}$$

and, $\epsilon_{h,d} = N(0, \sigma^2)$.

Deterministic variable refers predictable variables such as days of week, months, and years. From historical demand analysis, some sort of similar pattern can be noticed for daily and weekly demand profile as shown in figure 1 or 2 for example- daily pattern shows the higher demand during the business week (Monday to Friday) than that of weekend (Saturday and Sunday) or public holidays. Various kind of events have influence on demand such as public holiday, day before holiday, day after holiday, long holidays, bridge holidays. For consistent variables such as days and months, we use dummies on some reference. For example- For the case of all days of week, we can consider Saturday as reference dummy so that other days can be compare with respect to that day. Procedure holds same for months dummies and seasons dummies. And the model $Deterministic_{h,d} = Det_{h,d}$ is modeled with 25 variables with consistant term β_0.

$$Det_{h,d} = \beta_0 + \textstyle\sum_{i=1}^{6} \beta_i \, DayDum_{h,d} + \sum_{i=7}^{17} \beta_i \, MonthDum_{h,d} + \sum_{i=18}^{20} \beta_i \, SeasonDum_{h,d} + \beta_{21} Wkday_{h,d} + \beta_{22} WkdayBfHoliday_{h,d} + \beta_{24} HolidayBridge + \beta_{23} WkdayAfHoliday_{h,d} \tag{6}$$

Generally, the coefficient on the weekday dummies are expected to be greater than the coefficient on weekend dummy and also based on reference day. For example, our model is referenced on Saturday as a base level. That means Sunday and Monday should have higher deviation on their weight because of huge demand deviation. Similar effects should be on months dummies that means there should be some deviation on weights between winter and summer months.

Meteorological effects is the another factor that determines the demand of electricity. Although, in paper [2] Taylor implemented a weather-based regression approach and found that it did not show any significant effect on very short-term load forecasting, but worked well for hours ahead prediction. Temperature parameter in various form such as $temp, temp^2 temp^3, AveTemp$, minimum and maximum temperature, mid-night moving average daily and hourly temperature deviation, temperature interaction with months are the factors we consider here. Some pre-processing of temperature is done finding some correlation between temperature and demand, which implies that $17.1 \deg celcius$ as the reference point where there is no effect of temperature for demand. Therefore, effective temperature is:

$$Temp = Temp_{eff} = Abs(Temp_{actual} - 17.1) \qquad (7)$$

Now, we include some other meteorological variables such as relative humidity, wind speed, precipitation(rain or snowfall), solar radiation are also accounted for formulations. Therefore in $Meteorological_{h,d}$ term, there are 34 (temp related 22 and other meteorological related 12) variables altogether.

Similarly for for $HistoricalData_{h,d}$, we have to analyze its pattern of historical demand. Since, auto-regressive (AR) component captures the pattern of load in hour $h = i$ for any given day is a good indication that load will be higher in hour $h = i$ on the following day(s). More specifically, demand is strongly linearly correlated with day ahead same time load and previous hour load and its error term from previous 6 days. Therefore we have included 7 parameters for this term.

In this way, model A is developed substituting all covariates of these terms in equation 4 which consist altogether 74 variables including 6 correlated variables. This model A looks like more variables, but still demand continuously subjected to vary due to random kind of disturbances. For example, unknown working hours of large steel mills, shutdown of industrial activities, days with extreme weather or sudden change in weather are the promising factors that affect on demand. Although, we are trying to address extreme weather or sudden weather change by inserting hourly and daily deviation terms, but unscheduled holidays (eg- 28 Dec 2014 to 3rd Jan 2015) is still a limitation in our study.

As our interest is to analyze the effect of other meteorological factors on the demand forecasting, now we develop next mode called model B. Where we only exclude some meteorological variables from model A such as rain, snow, wind, radiation, humidity, cloud and their interactions (12 variables). Therefore, model B will consist 56 exogenous variables excluding 6 correlated coefficients, for prediction of demand.

In a mean time we remove some more variables from model B, having very low weight on their coefficient. Removing the variables having low weight, number of covariates now reduced to 40 excluding 6 correlated coefficients, named model C and is expressed as,

The auto-regressive moving average (ARMA) component represent the appropriate model of cyclicality for off-peak and peak hours. The $ARMA(p,q)$ representation can be specified as-

$$ARMA_{h,d} = \sum_{i=1}^{p} \beta_i \, Demand_{h,d-1} + \sum_{i=1}^{q} \beta_i \, \epsilon_{h,d-1} + \epsilon_{h,d} \qquad (8)$$

We conduct Ljung-Box Q-test after analysing the pattern of residuals. Patterns shows that residuals are not random but exhibits some correlations with previous hour error and previous day error. We perform BIC test here and implement our model as $ARMA(2,6)$.

V. Parameters Estimation

Bayesian estimation is used to estimate the parameter $\hat{\beta}$ and $\hat{\sigma}^2$ of general model equation ?. Bayesian principle, estimates the posterior distribution for $\hat{\beta}$ and $\hat{\sigma}^2$ with the product of prior belief and likelihood function. Prior belief is the initial assumption of $\hat{\beta}$ with some value of variation and likelihood function is the observational data with some mean and variance. Since, we consider the prior belief is normally distributed so that our posterior is also normally distributed. So,

$$P(\hat{\beta}, \hat{\sigma}^2 | Y) = N(Mean *, Var *)$$

Where,

$$Mean *= (\sum o^{-1} + \frac{1}{\sigma^2} X'X)^{-1} (\sum_{0}^{-1} \beta_0 + \frac{1}{\sigma^2} X'Y)$$

and,

$$Var *= (\sum o^{-1} + \frac{1}{\sigma^2} X'X)^{-1}$$

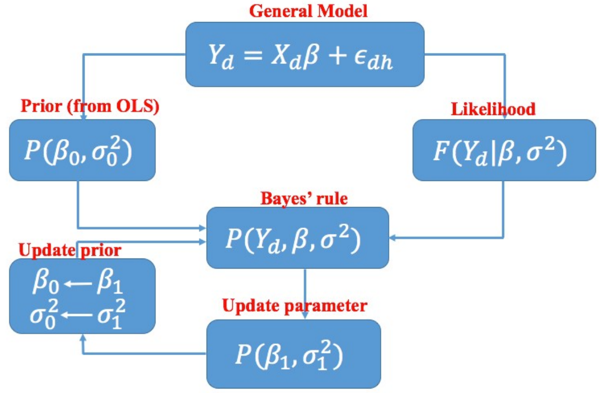

Fig. 4. General flow diagram for estimation of parameters

From this $Mean *$ and $Var *$ we draw the number of samples using Gibbs sampling. These samples are stored after the burn in stage of Markov chain and discards other samples. In our algorithm only 1000 samples are saved out of 5000 samples. So, the series of samples are stored as

$$\hat{\beta} = Mean * + [\bar{\beta} * (Var *)^{\frac{1}{2}}]$$

and $\hat{\sigma}^2$ we sample scalar value from inverse Gamma distribution with degree of freedom $\frac{T}{2}$ and scale parameter $\frac{\theta}{2}$

$$\hat{\sigma}^2 = \Gamma^{-1}(\frac{T}{2}, \frac{\theta}{2})$$

Here, classical regression technique is used for OLS estimation and hence obtained β values are used as the prior for posterior estimation form Bayesian technique.

VI. FORECASTING RESULTS

We have used three years of data set upto 2015 through 2012, where complete 2 years (730 days) of moving windows is used as training dataset to predict the out sample demand for the year of 2015. The multiple equations modeled here is estimated for each hour with separate equations having its own covariants. Therefore, each hour of the day, they have a different weight of parameter value.

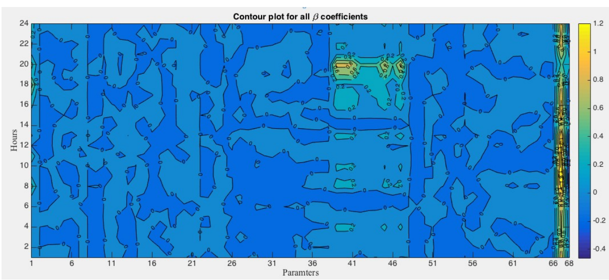

Fig. 5. Weight of all the variables as contour plot

Contour plot of figure 5 shows averaged values for whole year β coefficients for separate hours. Color intensity indicates the weight of parameter. Positive value corresponds positive effect on demand prediction where as negative values helps to reduce the forecasted value. For example, variables 38-47 (months dummies) cause increment of demand specially at evening time 16 to 22 hour, where as 67-68 (previous hour demand) have a very high positive influence though out all 24 hours. Some other parameters such as 35 through 25 includes day type dummies with both positive and negative effects on demand based on hours of the days. To discuss day type dummies in detail, we have plot the coefficients values in figure 6 below. For- example during morning time, demand has negative

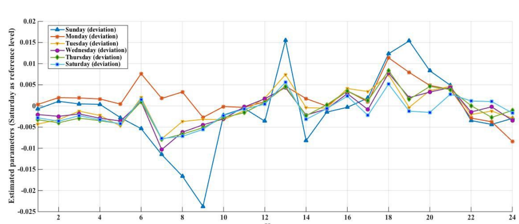

Fig. 6. Hourly analysis of coefficient for week dummies

In our models, we represent dummy variables Sunday to Friday assuming Saturday as a base level. The largest coefficient values are seen to occur on Monday because of previous day's demand, which are substantially lower than Monday demand, is being used to predict Monday forecast ($AR(2) effect: demand_{h,d-1}, demand_{h,d-2}$). The smallest coefficient specially during morning time (exactly same time of sharply increment of demand during weekdays) significantly decreases to generate lower weekends loads. Coefficients for different weekdays are found almost similar patterns indicating similar effects though out 24 hours, but during morning and night hours, coefficients are negative indicating decrease of demand than that of day hours specially evening 18 to 20 hour.

In figure 7, forecasted electricity demand for first week of January, 2015 is plotted and compared with actual demand. Since, we have implemented mean, median and 60 percentile forecast from the distribution of predicted data, which is the beauty of Bayesian estimation. For future prediction, such information are quite helpful to express demand prediction in terms of uncertainty. This first week consists various types of day types such as scheduled public holiday, unscheduled holidays, weekends and weekdays. Overall MAPE for this week is 0.82%, but still it is over estimated on Jan 1, due to non holiday effect of 31 Dec, under estimated on Jan 5, and 7 due to significant rise of peak on that day compared with the peaks of previous day. We have forecast the electricity demand for complete year 2015, with Bayesian approach with MAPE 0.69% and 0.15% variation. We have used, MAPE, and Root Mean Squared Error (RMSE) are widely used for performance analysis purpose.

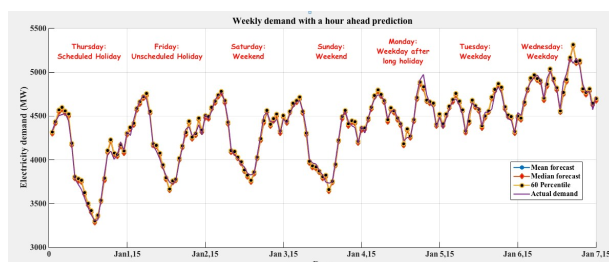

Fig. 7. Weekly demand variation for the first week of Jan 2015

VII. PERFORMANCE ANALYSIS

The performance is based on the capability of model that can predict the future value with better accuracy. Various methods are found in literature for this purpose in which MAPE, and RMSE are widely used in forecasting papers. MAPE is the most common measure of forecast error and considered as best when there are no extremes to the data because with zeros and near zero values, it provide infinite high value and vice versa.

$$MAPE = \frac{1}{N}\sum_{i=1}^{N} \left| \frac{(Y_i - \hat{Y}_i)}{Y_i} \right| \times 100 \tag{9}$$

RMSE depends on the scale of depend variable because it accounts the relative measurement with actual value. The smaller the error, better the forecasting ability of that model.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Y_i - \hat{Y}_i)^2} \tag{10}$$

The performance of model A is compared with other models (model B and model C) in figure 8 and 9 on the basis of each months of the year 2015. Figure 8 compares the performance between model A (that include all weather parameters) and model B (that exclude other weather parameter except temperature). On each month comparision, model A shows minimum average MAPE except some summer months. There is high deviation (approx. 0.2%) on average MAPE during the winter months December and January for both models A and B, but model A can improve it by more than 3% on same period. In overall, model A predict with the performance more than 2% for 6 months.

If model A is compare with model C (figure 9), similar improvement on average MAPE can be seen throughout a year, however some of the summer months such as July and August can be improved with both model B or C.

Figure (a)

Months	Model A performance				Model B performance				Improved (%)
	Average	StdDev	Max	Min	Average	StdDev	Max	Min	
Jan	0.696	0.197	1.301	0.396	0.719	0.209	1.324	0.410	3.224
Feb	0.641	0.190	1.024	0.360	0.650	0.191	1.160	0.387	1.426
Mar	0.676	0.147	1.062	0.443	0.696	0.152	1.102	0.459	2.871
Apr	0.690	0.126	0.932	0.451	0.715	0.123	0.940	0.474	3.642
May	0.793	0.164	1.132	0.407	0.821	0.170	1.173	0.437	3.476
Jun	0.679	0.136	0.940	0.439	0.682	0.155	1.039	0.418	0.497
Jul	0.684	0.121	0.983	0.499	0.676	0.131	0.933	0.456	−1.049
Aug	0.740	0.133	1.025	0.483	0.739	0.141	1.118	0.514	−0.198
Sep	0.715	0.135	1.101	0.410	0.731	0.127	1.009	0.451	2.184
Oct	0.724	0.153	1.080	0.441	0.724	0.162	1.050	0.407	−0.118
Nov	0.689	0.146	0.988	0.402	0.699	0.173	1.222	0.382	1.331
Dec	0.732	0.238	1.677	0.371	0.760	0.244	1.676	0.315	3.793

Fig 8. Comparision of performances for model A and B

Figure (b)

Months	Model A performance				Model C performance				Improved (%)
	Average	StdDev	Max	Min	Average	StdDev	Max	Min	
Jan	0.696	0.197	1.301	0.396	0.713	0.208	1.263	0.383	2.462
Feb	0.641	0.190	1.024	0.360	0.646	0.189	1.128	0.423	0.784
Mar	0.676	0.147	1.062	0.443	0.693	0.145	1.046	0.457	2.446
Apr	0.690	0.126	0.932	0.451	0.690	0.126	0.902	0.469	−0.077
May	0.793	0.164	1.132	0.407	0.824	0.169	1.153	0.427	3.931
Jun	0.679	0.136	0.940	0.439	0.686	0.137	1.017	0.448	1.132
Jul	0.684	0.121	0.983	0.499	0.673	0.117	0.923	0.466	−1.524
Aug	0.740	0.133	1.025	0.483	0.729	0.144	1.123	0.525	−1.541
Sep	0.715	0.135	1.101	0.410	0.733	0.140	1.038	0.414	2.477
Oct	0.724	0.153	1.080	0.441	0.725	0.153	1.036	0.423	0.069
Nov	0.689	0.146	0.988	0.402	0.690	0.174	1.143	0.319	0.121
Dec	0.732	0.238	1.677	0.371	0.757	0.246	1.701	0.317	3.333

Fig 9. Comparision of performances for model A and C

VIII. CONCLUSIONS

In this paper we developed three models based on literature about multiple equation demand forecasting model. During modeling, we pays particular attention to the weather variables that effects electricity demand and try to analyze quantitatively. We have analyzed these models based on their forecasting performance for complete one year out-sample prediction. Since, models were categorized based on all weather parameter include or not, they have a bit variation on their performances. More specifically, comparing with model B and C, model A that include all available weather parameters can improve the performance at least by 2% for 6 different months and by 1% for 8 months. Interestingly, during summer months (July and August) both model B and C looks better. One complexity for prediction during summer season is due to its high variation of demand. Sudden changes of temperature due to rainfall or wind speed also cause immediate fluctuations on demand. But, model B and C are succeed to address such a variation of demand. This indicates that optimization of exogenous variable is also necessary to improve performance. One more key point of this paper is that the overall an hour ahead forecast performance for all three models are impressive and better compared with Mandal P. et al. [21] forecasted for Okinawa Prefecture, Japan.

IX. REFERENCES

[1] J W Taylor. Short-term electricity demand forecasting using double seasonal exponential smoothing. J Oper Res Soc, 54(8):799–805, 2003.

[2] Yoshihiro Ohtsuka, Takashi Oga, and Kazuhiko Kakamu. Forecasting electricity demand in japan: A bayesian spatial autoregressive arma approach. Computational Statistics & Data Analysis, 54(11):2721–2735, November 2010.

[3] D. Asber, S. Lefebvre, J. Asber, M. Saad, and C. Desbiens. Non-parametric short-term load forecasting. International Journal of Electrical Power & Energy Systems, 29(8):630 – 635, 2007.

[4] Pierre-Andr´e Cornillon, Nick Hengartner, Vincent Lefieux, and Eric Matzner-Løber. Fully Nonparametric Short Term Forecasting Electricity Consumption, pages 79–93. Springer International Publishing, Cham, 2015.

[5] John Peirson and Andrew Henley. Electricity load and temperature: Issues in dynamic specification. Energy Economics, 16(4):235–243, October 1994.

[6] Ramu Ramanathan, Robert Engle, Clive W. J. Granger, Farshid Vahid-Araghi, and Casey Brace. Short-run forecasts of electricity loads and peaks. International Journal for Forecasting, (13):161–174, 1997.

[7] R. Cottet and Michael Stanley Smith. Bayesian modeling and forecasting of intraday electricity load. Journal of the American Statistical Association, 98:839–849, 2003.

[8] Lacir J. Soares and Marcelo C. Medeiros. Modeling and forecasting short-term electricity load: A comparison of methods with an application to brazilian data. International Journal of Forecasting, 24(4):630–644, October 2008.

[9] Virginie Dordonnat, Siem Jan Koopman, and Marius Ooms. Dynamic factors in periodic timevarying regressions with an application to hourly electricity load modelling. Computational Statistics & Data Analysis, 56(11):3134–3152, November 2012.

[10] Luiz Friedrich and Afshin Afshari. Short-term forecasting of the abu dhabi electricity load using multiple weather variables. Energy Procedia, 75:3014 – 3026, 2015.

[11] Kamal Chapagain and Somsak Kittipiyakul. Short-term electricity load forecasting model and Bayesian estimation for Thailand data. In 2016 Asia Conference on Power and Electrical Engineering (ACPEE 2016), volume 55, pages –, 2016.

[12] Andrew Harvey and Siem Jan Koopman. Forecasting hourly electricity demand using timevarying splines. Journal of the American Statistical Association, 88(424):1228–1236, December 1993.

[13] B. Satish, K.S. Swarup, S. Srinivas, and A.Hanumantha Rao. Effect of temperature on short term load forecasting using an integrated fANNg. Electric Power Systems Research, 72(1):95 – 101, 2004.

[14] J.W. Taylor and R. Buizza. Neural network load forecasting with weather ensemble predictions. Power Systems, IEEE Transactions on, 17(3):626–632, 2002.

[15] Jos´e Ram´on Cancelo, Antoni Espasa, and Rosmarie Grafe. Forecasting the electricity load from one day to one week ahead for the spanish system operator. International Journal of Forecasting, 24(4):588–602, October 2008.
[16] James W. Taylor. An evaluation of methods for very short-term load forecasting using minute-by minute british data. International Journal of Forecasting, 24(4):645–658, October 2008.
[17] Momani M. Awad. Factors affecting electricity demand in jordan. 5:50–58, 2013.
[18] Francesco Apadula, Alessandra Bassini, Alberto Elli, and Simone Scapin. Relationships between meteorological variables and monthly electricity demand. Applied Energy, 98:346 – 356, 2012.
[19] H.M. Al-Hamadi and S.A. Soliman. Longterm/ mid-term electric load forecasting based on short-term correlation and annual growth. Electric Power Systems Research, 74(3):353 – 361, 2005.
[20] Paras Mandal, Tomonobu Senjyu, Naomitsu Urasaki, and Toshihisa Funabashi. A neural network based several-hour-ahead electric load forecasting using similar days approach. International Journal of Electrical Power & Energy Systems, 28(6):367 – 373, 2006.

X. ACKNOWLEDGEMENTS

This research work is conducted under the Regional Climate System Laboratory, Hokkaido University, Japan. Authors are thankful to PARE exchange program and Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan that provide partial funding for this work. HEPCO, and JMA Japan are always thankful for providing necessary data used here for analysis.

XI. BIOGRAPHY

Kamal Chapagain was born in Jun 1978. He received B.E degree in Electronics and Communication Engineering from Pokhara Engineering College, Pokhara, Nepal in 2003 and M.Sc engineering from Pulchowk campus in 2013. He received excellent foreign scholarship (EFS) to conduct PhD at school of Information Communication Technology, Srindhron International Institute of Technology Thailand, and MEXT scholarship as long term exchange to Hokkaido University, Hokkaido, Japan. Currently, he is PhD candidate and his research is on short term electricity demand forecasting for Thailand.