

Electricity Demand Forecasting for Daily Peak Load with Seasonality and Temperature Effects

Sang-Wook Jung^a · Sahn Kim^{a,1}

^aDepartment of Applied Statistics, Chung-Ang University

(Received October 6, 2014; Revised October 7, 2014; Accepted October 7, 2014)

Abstract

Accurate electricity demand forecasting for daily peak load is essential for management and planning at electrical facilities.

In this paper, we first, introduce the several time series models that forecast daily peak load and compare the forecasting performance of the models based on Mean Absolute Percentage Error(MAPE).

The results show that the Reg-AR-GARCH model outperforms other competing models that consider Cooling Degree Day(CDD) and Heating Degree Day(HDD) as well as seasonal components.

Keywords: Time Series models, daily peak load, CDD, HDD, seasonality, MAPE.

1. 서론

현재 산업의 고도화와 다양한 수요에 발맞추어 국내의 전기소비량이 지속적으로 증가하고 있다. 이것은 우리사회의 전기에 대한 의존도 또한 높아지고 있음을 의미한다. 이러한 현실에서 예기치 못한 정전사태는 사회 구성원들로 하여금 많은 물질적 피해와 큰 정신적 고통을 야기한다. 특히 전기사용량이 전력공급량을 초과할 때 대한민국 전체에 전기가 나가는 대 정전 현상을 일컫는 블랙아웃(Blackout)이 일어나기라도 한다면 악몽과 같은 일이 될 것이다. 대표적 경고 사례가 바로 2011년 9월 15일 전국적으로 발생했던 정전사고이다. 블랙아웃 직전까지 갔던 전력 부족 상황을 막기 위해, 한국 전력거래소에서 선제적으로 전국 각 지역별로 돌아가면서 전기를 차단하는 방법으로 더 큰 사고를 막은 사건이다.

이렇듯 전기사용량이 전력공급량을 초과할 때 나타나는 블랙아웃현상이 실제로 일어날 수 있는 현상이다. 따라서 보다 효율적이고 합리적인 전력수급계획을 수립 가능케 하며, 더 큰 피해를 사전에 예방할 수 있다는 측면에서 전력수요량의 정확한 예측에 대한 중요성이 대두되고 있으며, 또한 정확도 높은 예측을 위한 여러 가지 방법들이 연구 되고 있다.

전력수요 예측은 현재 전 세계적인 빅 이슈이며 이제까지의 대표적인 연구로서 Amjady (2001)는 전력수요와 관계있는 독립변수를 고려한 ARIMA 모형을 이용해 예측력을 높였다. Taylor (2003)는 이중계

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(2014-027173).

¹Corresponding author: Department of Applied Statistics, Chung-Ang University, 221 Heukseok-Dong, Dongjak-Gu, Seoul 156-756, Korea. E-mail: sahm@cau.ac.kr

질 지수평활모형을 사용하여 단기 전력수요를 예측하였고, Huang과 Shih (2003)은 ARIMA 모형을 이용하여 단기 전력수요를 예측하였다. Sigauke와 Chikobvu (2011)는 분산모형을 이용하여 일별최대전력수요량에 대해 예측하였다. Fan과 Hyndman (2012)은 준모수(semi-parametric) 방법을 근거로 단기 전력수요를 예측하였다. Lee 등 (2013)은 독립변수를 고려한 계절형 ARIMA모형을 통해 단기 전력수요를 예측하였다.

우리나라의 경우 외국에 비해 사계절이 뚜렷해 여름과 겨울 사이의 기온차이가 매우 크며, 이 때문에 봄·가을과 비교하여 여름·겨울의 전력수요가 높아 1년 동안 전력수요량의 변동성이 크다는 특징을 가진다. 본 논문에서는 이러한 특징을 고려하기 위해 회귀 모형, ARIMA 모형, 그리고 GARCH 모형에 각각 일평균온도를 사용하였을 때와 냉·난방지수를 사용하였을 때로 구분하여 1년간 장기수요예측을 통해 모형의 성능을 비교하고자 한다.

2장에서는 이 논문에서 언급하고자 하는 모형들에 대한 설명을, 3장에서는 분석 데이터에 대한 설명할 것이다. 이를 바탕으로 4장에서는 비교 모형별 예측력에 대해 비교해보고, 5장에서는 기존의 일별 최대 전력 수요 예측 방법보다 우수한 방법에 대한 것과 향후 연구 방향에 대한 결론으로 끝맺음 할 것이다.

2. 전력수요 예측을 위한 시계열 모형의 소개

최근에 전력수요 예측을 위한 방법으로 다양한 시계열 모형이 사용되고 있다. 본 논문에서는 이제까지 제시된 주요 모형을 고개하여 모형의 성능분석 및 비교를 위한 기초 작업을 제시하고자 한다.

2.1. SARIMA 모형

SARIMA(Seasonal ARIMA) 모형은 Box 등 (1994)에 의해 제안된 모형으로 시계열 $\{Y_t | t = 1, 2, \dots, T\}$ 가 평균이 μ 인 $ARIMA(p, d, q) \times (P, D, Q)_s$ 를 따를 때의 모형은 식 (2.1)과 같이 나타난다.

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D(Y_t - \mu) = \theta_q(B)\Theta_Q(B^s)\varepsilon_t. \quad (2.1)$$

B 는 후진연산자(backshift operator)를 의미하고, p, d, q, P, D, Q 는 정수들이며, s 는 계절주기를 의미한다. $\phi_p(B), \Phi_P(B^s), \theta_q(B), \Theta_Q(B^s)$ 는 각각 B 에 대한 p, P, q, Q 차 다항식이며, 각각 $\phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p$, $\Phi_P(B) = 1 - \Phi_1 B^s - \dots - \Phi_P B^{Ps}$, $\theta_q(B) = 1 - \theta_1 B - \dots - \theta_q B^q$, 그리고 $\Theta_Q(B^s) = 1 - \Theta_1 B^s - \dots - \Theta_Q B^{Qs}$ 를 의미한다. d 는 차분 차수를 의미하며, D 는 계절차분 차수를 의미한다. 또한 Y_t 는 t 시점($t = 1, 2, \dots, T$)에서의 관측값을 의미하며, ε_t 는 평균이 0이고 분산이 고정된 백색잡음과정(White noise process)이다.

본 논문에서는 SARIMA 모형이 벤치마킹 모형으로 지정하였으며 많은 연구 논문에서도 사용된다 (Sigauke와 Chikobvu, 2011).

2.2. Reg-ARIMA 모형

Reg-ARIMA 모형은 기존 ARIMA모형에 회귀모형의 독립변수 효과를 고려한 모형으로 Regression과 ARIMA를 합친 단어이다. k 개의 독립변수를 고려한 시계열 $\{Y_t | t = 1, 2, \dots, T\}$ 가 $ARIMA(p, d, q) \times (P, D, Q)_s$ 를 따를 때의 모형은 식 (2.2)와 같이 나타난다.

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D \left(Y_t - \sum_{i=1}^k \beta_i x_{ti} \right) = \theta_q(B)\Theta_Q(B^s)\varepsilon_t. \quad (2.2)$$

각 요소들의 의미는 2.1절에서와 동일하며, β_i 들은 각각 독립변수 x_{ti} 의 계수를 의미하며 실수값을 갖는다. 본 논문에서는 독립변수로 온도에 해당하는 변수, 공휴일, 설·추석 연휴를 고려하였다.

2.3. SGARCH(Seasonal GARCH) 모형

앞서 언급한 ARIMA 모형의 경우 오차항의 분산이 시간에 관계없이 일정한 값을 갖는 등분산성 가정이 필요하다. 이에 처음으로 이분산성을 고려한 ARCH(Autoregressive conditionally heteroscedasticity) 모형이 Engle (1982)에 의해 제안되었고, Bollerslev (1986)이 현재 분산이 과거 분산의 영향을 받는 형태로 확장시켜 보다 일반화된 형태인 GARCH(Generalized autoregressive conditionally heteroscedasticity) 모형을 제안하였다. 이에 GARCH(r, m) 모형은 식 (2.3)와 같이 나타난다.

$$\begin{aligned} Y_t &= \mu_t + \varepsilon_t, \\ \varepsilon_t &= \sqrt{h_t}e_t, \\ h_t &= \alpha_0 + \sum_{i=1}^m \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^r \beta_i h_{t-i}. \end{aligned} \quad (2.3)$$

이 때, $e_t \sim WN(0, \sigma^2)$ 이며, 각 모수들은 $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$, $\alpha_i + \beta_i < 1$ 의 조건을 갖는다. 또한, m 은 ARCH 부분의 차수이고, r 은 GARCH 부분의 차수를 뜻한다. 만약 r 이 0일 경우, 즉, GARCH($0, m$)모형은 ARCH(m)과 같다.

여기서 오차항에서 계절적인 요소를 고려한 SGARCH(Seasonal GARCH)모형은 식 (2.5)와 같이 나타난다.

$$\begin{aligned} Y_t &= \mu_t + \varepsilon_t, \\ \varepsilon_t &= \sqrt{h_t}e_t, \\ h_t &= \alpha_0 + \sum_{i=1}^m \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^r \beta_i h_{t-i} + \sum_{i=1}^M \alpha_{is} \varepsilon_{t-i}^2 + \sum_{i=1}^R \beta_{is} h_{t-i}. \end{aligned} \quad (2.4)$$

이 때, s 는 계절주기를 의미하고, M, R 은 각각 GARCH 부분과 ARCH 부분의 계절차수를 의미한다.

2.4. Reg-AR-SGARCH 모형

AR-GARCH 모형은 기존 GARCH 모형에서 평균모형 μ_t 가 AR(p) 모형일 때를 의미하며 AR(p) - GARCH(r, m)은 식 (2.5)와 같이 나타난다.

$$\begin{aligned} \phi_p(B)(Y_t - \mu) &= \varepsilon_t, \\ \varepsilon_t &= \sqrt{h_t}e_t, \\ h_t &= \alpha_0 + \sum_{i=1}^m \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^r \beta_i h_{t-r}. \end{aligned} \quad (2.5)$$

이 때, $\phi_p(B)$ 는 후진연산자 B에 대한 다항식으로 $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ 를 의미하고, 분산모형은 GARCH 모형에서와 같이 $e_t \sim WN(0, \sigma^2)$ 이며, 각 모수들은 $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$, $\alpha_i + \beta_i < 1$ 의 조건을 갖는다.

이에 회귀모형의 형태를 포함하여 k 개의 독립변수 효과를 고려한 Reg(k) - AR(p) - SGARCH(r, m) \times (R, M) $_s$ 모형은 식 (2.6)과 같다.

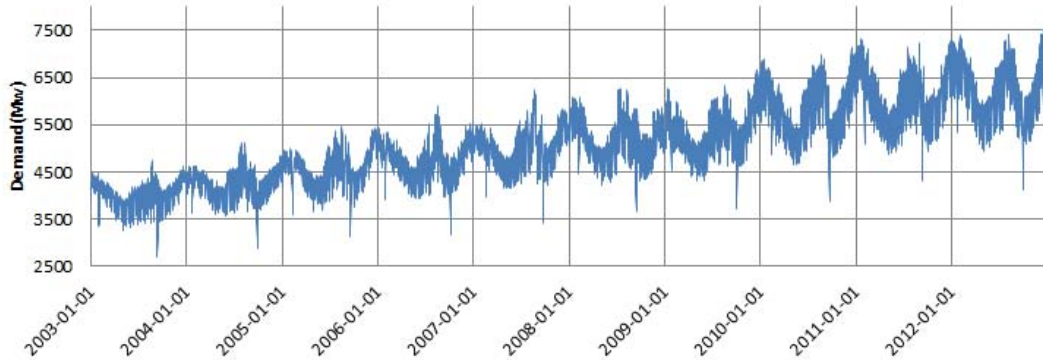


Figure 3.1. Original time series data

$$\begin{aligned} \phi_p(B) \left(Y_t - \sum_{i=1}^k \gamma_i x_{it} \right) &= \varepsilon_t, \\ \varepsilon_t &= \sqrt{h_t} e_t, \\ h_t &= \alpha_0 + \sum_{i=1}^m \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^r \beta_i h_{t-i} + \sum_{i=1}^M \alpha_{is} \varepsilon_{t-i}^2 + \sum_{i=1}^R \beta_{is} h_{t-i}. \end{aligned} \quad (2.6)$$

3. 자료 소개 및 분석

본 논문에서 이용한 데이터는 2003년 1월 1일부터 2012년 12월 31일까지 한국 전력거래소에서 수집된 일별 최대 전력수요 데이터이다. 이 중 2003년 1월 1일부터 2011년 12월 31일까지 9년 동안의 데이터를 트레이닝 데이터(training data)로써 모형을 적합하는데 사용하고 2012년 1월 1일부터 2012년 12월 31일까지 1년 치의 데이터를 테스트 데이터(test data)로써 모형의 성능을 평가하는데 사용하였다.

한편, 전력수요에 가장 큰 영향을 미치는 기온에 대한 효과를 고려하는 변수로 일평균기온과 냉방지수(Cooling Degree Day; CDD), 그리고 난방지수(Heating Degree Day; HDD)를 사용하였다. 일평균 기온은 기상청으로부터 수집된 전국 5개 지역(서울, 강릉, 대전, 광주, 부산)의 일별 평균 기온의 평균치를 나타냈고, 냉·난방지수는 식 (3.1)과 같이 나타낸다.

$$\text{CDD} = \begin{cases} T_t - 24, & \text{if } T_t \geq 24 \\ 0, & \text{if } T_t < 24 \end{cases}, \quad \text{HDD} = \begin{cases} 18 - T_t, & \text{if } T_t \leq 18 \\ 0, & \text{if } T_t > 18 \end{cases}. \quad (3.1)$$

이 때, T_t 는 일평균기온을 의미한다. 냉방지수와 난방지수의 기준 온도는 실내에서 사람이 쾌적하게 지낼 수 있는 기온으로 민족, 연령, 사회·경제적 여건, 목적 등에 따라 다르다. 우리나라의 경우 냉·난방지수의 기준온도를 각각 24°C와 18°C로 사용한다.

일별 최대 전력 수요량에 대한 원 데이터의 시계열 그래프를 보면 Figure 3.1과 같다. Figure 3.1로부터 데이터가 1년을 주기로 같은 패턴을 나타내는 것을 알 수 있다. 1년을 주기로 나타나는 패턴을 분석해보면 전력소비가 큰 냉·난방용품의 사용량이 많은 여름과 겨울의 전력수요량이 많게 나타나며, 그렇지 않은 봄, 가을의 전력수요량은 비교적 작게 나타난다. 또 눈에 띄는 특이한 현상으로는 설 연휴와 추석 연휴에 있어서 정규적인 패턴과 다르게 전력수요량이 눈에 띄게 적은 현상을 꼽을 수 있다. 음력달력으

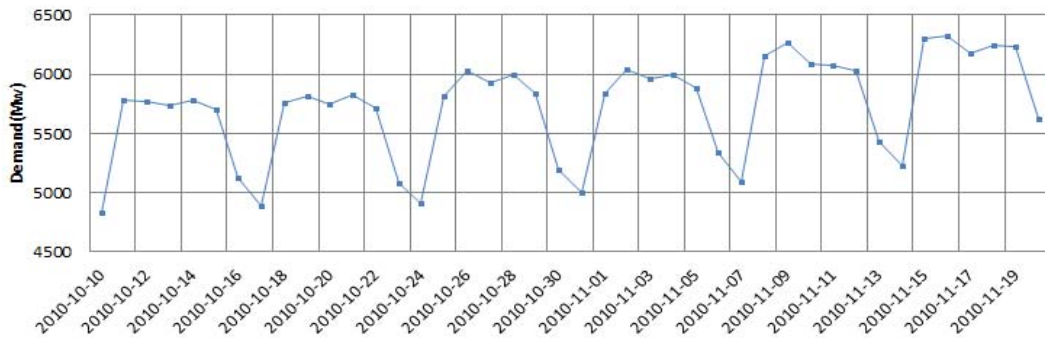


Figure 3.2. Time plot for weeks without holidays and special days

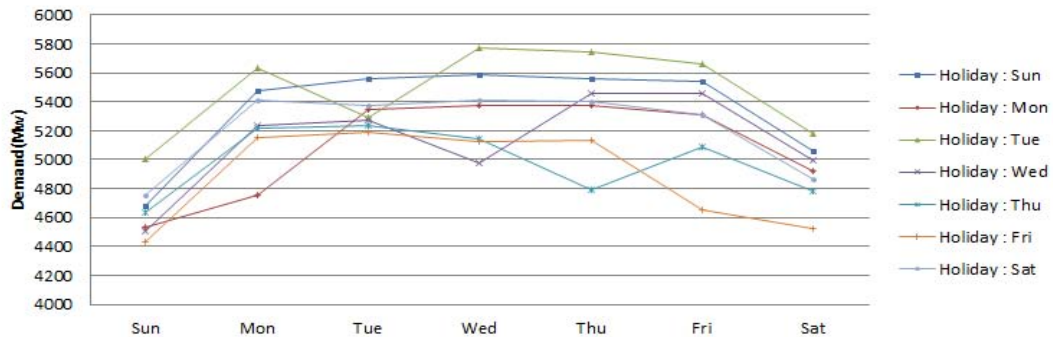


Figure 3.3. Daily impact with holidays

로 발생하는 설 연휴와 추석 연휴의 경우 그 일정이 매년 다르게 나타나며, 그 효과 또한 다른 공휴일들과 전혀 상이하므로 연휴변수만을 개별적인 더미변수로 설정하여 연휴효과를 모형에 반영하도록 한다.

Figure 3.2는 2010년 10월 10일 부터 2010년 11월 20일 까지 공휴일을 미 포함한 6주 동안의 시계열그림으로 전력수요량은 1주일을 주기로 계절성의 특징을 갖는 것을 알 수 있다. 주중 5일 동안의 전력수요량은 비슷한 형태를 띄며, 주말은 주중보다 낮은 소비형태를 띤다.

Figure 3.3은 연휴를 제외한 공휴일이 전력수요량에 미치는 영향을 각 요일별로 구분하여 나타내고 있다. 공휴일이 토요일, 일요일인 경우에는 그 효과가 크지 않았지만, 주중에 공휴일이 나타나는 경우 전력수요량이 비공휴일에 비하여 큰 폭으로 줄어든 것으로 보아 주중 공휴일의 영향이 있음을 알 수 있다. 따라서 공휴일이 주말인 경우를 제외한, 공휴일이 주중에 위치하는 날에 대해서 더미변수를 생성하여 그 효과를 모형에 반영하였다.

Figure 3.1로부터 전력수요량은 시간이 흐름에 따라서 변동 폭이 증가하는 특징을 가지고 있으며, Figure 3.2로부터 시간이 흐름에 따라 일주일 단위의 주기 형태는 그대로이되 수요량이 변화하는 비정상적 특징을 가진다. 따라서 로그변환 후 7차 차분을 하여 정상성 조건을 만족하는 데이터를 모형설정에 이용하도록 하며 그 형태는 Figure 3.4와 같다.

3.1. 모형의 적합 결과

회귀모형과 ARIMA 모형, 그리고 AR-GARCH 모형을 각각 일평균온도를 고려했을 경우와 냉·난방지

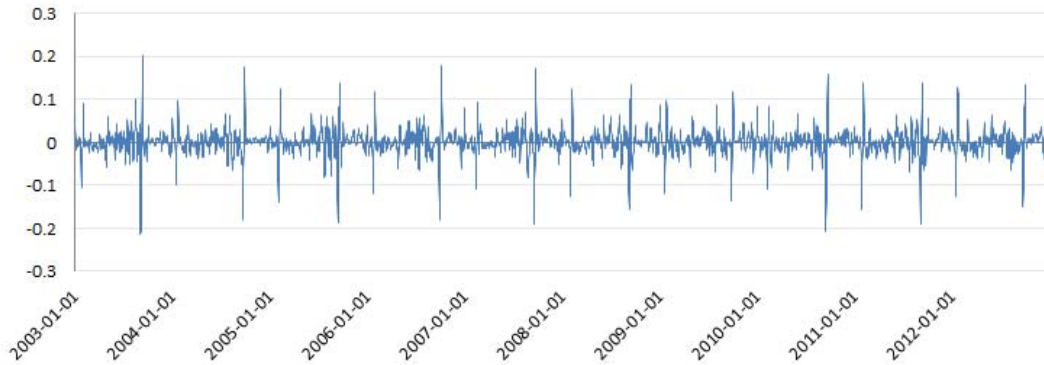


Figure 3.4. Log-different(7) data

Table 3.1. Estimates of Regression

모수	추정값	표준오차	p-value
t	0.7015	0.0048	< .0001
β_0	3705.6170	20.4429	< .0001
β_{holiday}	-376.2010	32.1354	< .0001
β_{temp}	-15.7929	0.9651	< .0001
$\beta_{\text{specialday}}$	-1126.4700	36.6292	< .0001
β_{Sun}	-284.6110	17.0453	< .0001
β_{Mon}	410.6279	17.0460	< .0001
β_{Tue}	434.1098	17.0455	< .0001
β_{Wed}	427.7745	17.0354	< .0001
β_{Thu}	448.6093	17.0360	< .0001
β_{Fri}	417.3200	17.0361	< .0001
β_{Jan}	406.8842	21.8460	< .0001
β_{Feb}	290.6286	20.7836	< .0001
β_{Jul}	470.6570	19.6231	< .0001
β_{Aug}	583.5090	20.2178	< .0001
β_{Sep}	177.8439	18.8407	< .0001
β_{Dec}	323.3178	20.3704	< .0001

수를 고려한 경우로 나누어 6가지 모형을 비교하였다.

회귀모형은 온도변수 외에 시간, 공휴일, 연휴, 요일 변수, 여름(7, 8월)과 겨울(12 ~ 2월)에 해당하는 월별 변수를 독립변수로 포함하였다. 이 때 일평균온도를 고려했을 때의 모수 추정치는 Table 3.1, 냉·난방지수를 고려했을 경우의 모수추정치는 Table 3.4와 같다.

ARIAM모형에서 온도변수 위의 독립변수로는 공휴일, 연휴를 고려하였다. 모형 선정은 AIC(Akaike's Information Criterion)가 최소인 모형으로 선택하였다. AIC를 고려한 최적 모형은 $ARIMA(2, 0, 1) \times (1, 0, 2)_7 \times (1, 0, 1)_{364}$ 였다. 잔차의 독립성을 보여주는 포트만트 검정(Portmanteau test) 결과 0.05보다 커 잔차들 사이에 독립성을 만족하였고, 추정된 모든 모수의 p -value 는 0.05보다 작아 유의하게 나타났다. 이 때 일평균온도를 고려했을 때의 모수 추정치는 Table 3.2, 냉·난방지수를 고려했을 경우의 모수추정치는 Table 3.5와 같다.

AR-GARCH모형에서 독립변수로는 $AR(2)(2)_7(1)_{364} - ARCH(3)(1)_7$ 모형과 마찬가지로 냉·난방지

Table 3.2. Estimates of ARIMA(2, 0, 1)(1, 0, 2)₇(1, 0, 1)₃₆₄

모수	추정값	표준오차	p-value
θ_1	-0.7609	0.0145	< .0001
Θ_7	1.7112	0.0363	< .0001
Θ_{14}	-0.7189	0.0342	< .0001
Θ_{364}	0.7418	0.0441	< .0001
ϕ_2	0.6833	0.0162	< .0001
Φ_7	0.8968	0.0278	< .0001
Φ_{364}	0.9172	0.0341	< .0001
β_{temp}	-0.0006	0.0001	< .0001
$\beta_{specialday}$	-0.0641	0.0021	< .0001
$\beta_{holiday}$	-0.0300	0.0011	< .0001

Table 3.3. Estimates of AR(2)(2)₇(1)₃₆₄ - ARCH(3)(1)₇

모수	추정값	표준오차	p-value
ϕ_2	-0.0012	0.0001	< .0001
Φ_7	-0.0928	0.0006	< .0001
Φ_{14}	-0.0329	0.0010	< .0001
Φ_{364}	-0.5077	0.0090	< .0001
α_0	0.2291	0.0099	< .0001
α_3	0.0777	0.0078	< .0001
α_7	-0.1249	0.0097	< .0001
β_{temp}	0.0001	2.66E ⁻⁶	< .0001
$\beta_{specialday}$	0.4106	0.0259	< .0001
$\beta_{holiday}$	0.3528	0.0246	< .0001

Table 3.4. Estimates of Regression with CDD and HDD

모수	추정값	표준오차	p-value
t	0.6908	0.0041	< .0001
β_0	3345.1840	13.5793	< .0001
$\beta_{holiday}$	-380.2370	27.2695	< .0001
β_{CDD}	141.3261	5.7575	< .0001
β_{HDD}	32.1629	0.9832	< .0001
$\beta_{specialday}$	-1127.3300	31.0861	< .0001
β_{Sun}	-279.3480	14.4663	< .0001
β_{Mon}	416.6521	14.4672	< .0001
β_{Tue}	435.8567	14.4664	< .0001
β_{Wed}	428.9985	14.4573	< .0001
β_{Thu}	445.4219	14.4578	< .0001
β_{Fri}	415.0950	14.4580	< .0001
β_{Jan}	202.9466	19.5081	< .0001
β_{Feb}	137.4234	18.1606	< .0001
β_{Jul}	297.4281	16.4102	< .0001
β_{Aug}	269.7504	18.8947	< .0001
β_{Sep}	182.8168	15.4657	< .0001
β_{Dec}	164.0145	17.8875	< .0001

수, 공휴일, 연휴를 고려하였고, 최소의 AIC를 갖는 모형을 선택하였다. 이 때, 최적 모형은 GARCH

Table 3.5. Estimates of ARIMA(2, 0, 1)(1, 0, 2)₇(1, 0, 1)₃₆₄ with CDD and HDD

모수	추정값	표준오차	p-value
θ_1	-0.6815	0.0156	< .0001
Θ_7	1.7132	0.0376	< .0001
Θ_{14}	-0.7214	0.0351	< .0001
Θ_{364}	0.7549	0.0415	< .0001
ϕ_2	0.5983	0.0170	< .0001
Φ_7	0.8988	0.0295	< .0001
Φ_{364}	0.9358	0.0313	< .0001
β_{CDD}	0.0101	0.0005	< .0001
β_{HDD}	0.0017	0.0001	< .0001
$\beta_{\text{specialday}}$	-0.0678	0.0019	< .0001
β_{holiday}	-0.0297	0.0010	< .0001

Table 3.6. Estimates of AR(2)(2)₇(1)₃₆₄ - ARCH(3)(1)₇ with CDD and HDD

모수	추정값	표준오차	p-value
ϕ_2	-0.4212	0.0112	< .0001
Φ_7	0.2558	0.0145	< .0001
Φ_{14}	0.0697	0.0098	< .0001
Φ_{364}	-0.1192	0.0105	< .0001
α_0	0.0001	2.13E ⁻⁰⁶	< .0001
α_3	0.3726	0.0232	< .0001
α_7	0.2980	0.0216	< .0001
β_{CDD}	0.0104	0.0002	< .0001
β_{HDD}	0.0019	6.78E ⁻⁰⁵	< .0001
$\beta_{\text{specialday}}$	0.0959	0.0005	< .0001
β_{holiday}	-0.0325	0.0009	< .0001

Table 3.7. MAPE of the models

	TEMP			HDD-CDD		
	Regression	ARIMA	AR-GARCH	Regression	ARIMA	AR-GARCH
MAPE	4.6700	3.4840	3.4610	4.1786	3.1750	3.1403

부분은 유의하지 않은 었다. 이 때 일평균온도를 고려했을 때의 모수 추정치는 Table 3.3, 냉·난방지수를 고려했을 경우의 모수추정치는 Table 3.6과 같다.

3.2. 모형의 성능비교

본 논문에서는 모형의 예측력 비교를 위해 MAPE(Mean Absolute Percentage Error)를 사용했으며, 이는 다음과 같이 정의된다.

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - F_t}{Y_t} \right| \times 100(\%), \quad (3.2)$$

이 때, n 은 예측에 사용한 데이터의 수이고, Y_t 는 t 시점에서의 실제 값, F_t 는 t 시점에서의 예측 값을 의미한다.

Table 3.7은 각 모형별로 1년(2012.01.01. ~ 2012.12.31.)에 해당하는 MAPE 값으로 온도변수에 대한

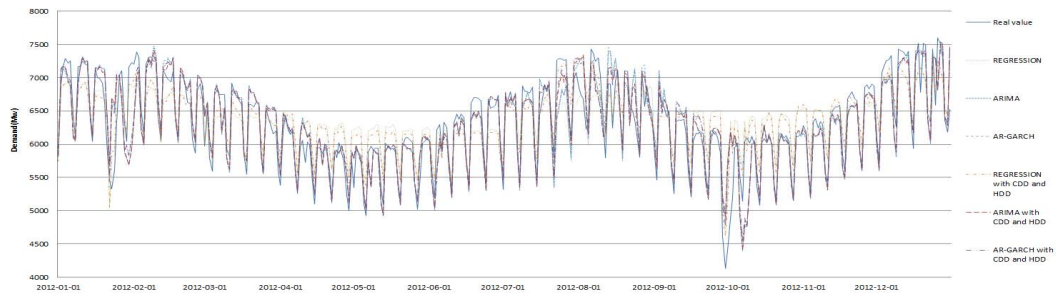


Figure 3.5. Comparison of forecasts

차이를 비교해보면 냉·난방지수를 고려한 모형이 일평균온도를 고려한 모형보다 오차율이 작다. 이는 냉·난방지수의 일별 최대전력수요량에 대한 설명력이 우수함을 나타낸다. 냉·난방지수를 포함한 AR-GARCH 모형의 MAPE가 3.14%로 가장 우수한 성능을 보였다. Figure 3.5는 실제값과 각 모형별 예측값을 나타낸 그래프이다.

4. 결론

본 논문에서는 시계열 모형을 이용하여 일별 최대전력수요량에 대해 장기(1년)예측을 실시하였다. 시계열 자료의 분석과 예측에 있어서 전통적으로 많이 이용되는 회귀모형과 ARIMA 모형, GARCH 모형에 각각 독립변수로 일평균온도를 사용한 모형과, 냉·난방지수를 사용한 모형 총 6가지 모형을 사용하여 일별 최대전력수요량에 대한 예측을 시행하였다. 이 때, 냉·난방지수를 독립변수로 한 GARCH 모형을 통한 예측의 결과가 오차율 3.14%로 가장 우수한 성능을 보였다. 이는 사계절이 뚜렷하여 1년 간 온도의 변동 폭이 매우 크다는 우리나라의 특징을 큰 변동성에 적합한 GARCH 모형을 고려함과 동시에, 여름·겨울에 전력수요량과 밀접한 연관성을 가진 냉·난방시설의 가동이 많다는 특징을 냉·난방지수를 독립변수로 넣음으로써 그 설명력을 높였기 때문으로 판단된다.

본 논문에서는 하루 중 가장 전력수요량이 높은 시간에 측정된 일별 최대전력수요량을 사용하여 예측하였지만, 시간별 전력수요량을 이용할 경우 일별 최대전력수요량에 비해 변동성이 더 커지며, 시간별 자료의 특징인 주기성까지 고려할 수 있어 GARCH 모형을 통해 보다 뛰어난 예측력을 얻을 수 있을 것으로 생각된다. 그리고 또한, 블랙아웃 사고 등 전력사고는 주로 전력수요량을 고려하지 못해, 충분한 전력공급이 이루어지지 못할 때 발생한다는 점에서 전력수요량이 낮은 시점보다 높은 시점이 더욱 중요하게 여겨진다는 점을 고려해볼 때, 전력소모가 큰 냉·난방기구의 사용량이 높은 여름·겨울의 데이터 간의 연관성을 집중적으로 고려한 모형에 대한 연구가 향후 필요하다고 사료된다.

References

- Amjady, N. (2001). Short-term hourly load forecasting using time-series modeling with peak load estimation capability, *IEEE Transactions on Power Systems*, **16**, 498–505.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**, 307–327.
- Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (1994). Time series analysis: forecasting and control, *Princeton-Hall International*.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica: Journal of the Econometric Society*, **50**, 987–1007.

- Fan, S. and Hyndman, R. J. (2012). Short-term load forecasting based on a semi-parametric additive model, *IEEE Transactions on Power Systems*, **27**, 134–140.
- Huang, S. J. and Shih, K. R. (2003). Short-term load forecasting via ARMA model identification including non-Gaussian process considerations, *IEEE Transactions on Power Systems*, **18**, 673–679.
- Lee, J. S., Sohn, H. G. and Kim, S. (2013). Daily peak load forecasting for electricity demand by time series models, *The Korean Journal of Applied Statistics*, **26**, 349–360.
- Sigauke, C. and Chikobvu, D. (2011). Prediction of daily peak electricity demand in South Africa using volatility forecasting models, *Energy Economics*, **33**, 882–888.
- Taylor, J. W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing, *Journal of the Operational Research Society*, **54**, 799–805.

계절성과 온도를 고려한 일별 최대 전력 수요 예측 연구

정상욱^a · 김삼용^{a,1}

^a중앙대학교 응용통계학과

(2014년 10월 6일 접수, 2014년 10월 7일 수정, 2014년 10월 7일 채택)

Abstract

급증하고 있는 전력수요에 대한 신뢰성 있는 예측은 합리적인 전력수급계획 수립 및 운용에 있어서 매우 중대한 사안이다. 본 논문에서는 여러 시계열 모형의 비교를 통해 전력수요량과 밀접한 연관성이 있는 온도를 어떠한 형태로 고려할 것인지, 또한 4계절이 뚜렷하여 계절별 기온 차가 많이 나는 우리나라의 특성을 어떻게 고려할 것인지에 대하여 연구하였다. 모형 간 예측력을 비교하기 위하여 Mean Absolute Percentage Error(MAPE)를 사용하였다.

모형의 성능비교 결과는 냉·난방지수와 계절요인을 동시에 고려하면서 큰 변동성을 잘 고려해줄 수 있는 Reg-AR-GARCH 모형이 가장 우수한 예측력을 나타냈다.

주요용어: 시계열 모형, 일별 최대전력수요, 난방지수, 냉방지수, 계절성, MAPE.

이 논문은 2014년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임(2014-027173).

¹교신저자: (156-756) 서울 동작구 흑석동 221, 중앙대학교 응용통계학과. E-mail: sahm@cau.ac.kr