

Forecasts of electricity consumption in an industry building

Minah Kim^a · Jaehee Kim^{a,1}

^aDepartment of Statistics, Duksung Women's University

(Received December 21, 2017; Revised January 30, 2018; Accepted March 8, 2018)

Abstract

This study is on forecasting the electricity consumption of an industrial manufacturing building called GGM from January 2014 to April 2017. We fitted models using SARIMA, SARIMA + GARCH, Holt-Winters method and ARIMA with Fourier transformation. We also forecasted electricity consumption for one month ahead and compared the predicted root mean square error as well as the predicted error rate of each model. The electricity consumption of GGM fluctuates weekly and annually; therefore, SARIMA + GARCH model considering both volatility and seasonality, shows the best fit and prediction.

Keywords: ARIMA, electricity consumption, forecasting, Fourier transformation, GARCH, Holt-Winters, seasonal, SARIMA, time series

1. 서론

전기는 에디슨이 백열전구를 발명한지 8년 만인 1887년, 우리나라 경복궁에서 최초의 전등불을 밝히면서 전기가 도입 되었다. 1898년 이후 우리나라 최초의 전기 회사가 설립되면서 전기 사용이 시작되었고 그 이후로 한전(한국전력공사)과 원자력 발전소가 완공되면서 전기의 보급이 수월해졌다. 그 결과 전기 사용량이 높아지고, 점차 에너지 수입의존도가 높아지게 되었다. 현재 전기의 과도한 사용을 줄이기 위해서 전기 사용량 증가에 대해 비례 이상으로 세율을 증가시키는 전기 누진세 제도가 도입되었다. 그 결과 누진세 부과 등으로 전기 사용량이 높은 경향을 보이면서 전기 사용량을 줄이는 것이 기업들과 개인 가구들의 관심사가 되었다. 그래서 한국전력공사 등 여러 기관에서는 전기 사용량에 대한 분석을 심도있게 진행하고 있다.

시계열 분석은 데이터가 시간에 따라 변동성과 추세가 존재하는 주택가격, 주식, 기후변화, 수요조사 등 여러 가지 분야에서 많이 쓰이는 분석이다. 시계열 데이터는 시간 데이터 현상에 대한 분석과 동시에 미래에 대한 예측분석이 가능하여 의사결정에도 도움을 줄 수 있다. Choi 등 (2009)은 10년 동안 국내 회

This research was supported by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry & Energy (MOTIE) of the Republic of Korea (No. 20161210200610). It was also supported by a program of the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (2016R1D1A1B03933527).

¹Corresponding author: Department of Statistics, Duksung Women's University, 419 Samyang-ro 144 Gil 33, Dobong-Gu, Seoul 01369, Korea. E-mail: jaehee@duksung.ac.kr

사의 가스 수요량에 대한 월별 데이터를 사용하여 온도효과, 확산계절효과 등을 고려한 가스 수요량 예측 모형을 구축하였다. 특히 온도 변수의 경우 회귀모형을 사용하여 가격탄력성을 계산하여 가스 수요량의 탄력성을 고려하였다.

Chujai 등 (2013)은 2006년부터 2010년까지 기간 동안 1분 단위로 샘플링이 된 한 가구의 전력 소비 데이터에 대해서 auto-regressive intergrated moving average (ARIMA) 모형으로 예측을 하였고, Abdel-Aal와 Al-Garni (1997)는 사우디 아라비아 동부 주의 1987년 8월부터 6년간의 월간 전기 소비를 모델링, 예측하고 회귀분석과 GARCH 네트워크 머신러닝 모델과 비교하였다. 그러나 autoregressive moving average (ARMA) 모형은 계절성이 있는 데이터에서 적합을 하지 못하는 단점이 있다. 이러한 모형의 단점을 보완하여 계절성(seasonal ARIMA; SARIMA), 변동성(generalized autoregressive conditional heteroskedasticity; GARCH)을 포함한 모형이 발표되었다. 자세한 모형에 대한 설명은 Cryer와 Chan (2008)에서 참고하면 된다. Kim과 Kim (2013)은 원, 달러 데이터로 ARIMA 모형과 변동성을 포함한 모형 중에서도 ARIMA + IGARCH 모형을 이용하여 한 달의 환율을 예측하였다. 또한 Tan 등 (2010)은 2002년 스페인의 전자상거래 데이터, 2006년 PJM 전자상거래 데이터를 이용하여 ARIMA, ARIMA + GARCH, 그리고 소파동 변환을 이용한 ARIMA + GARCH 등의 모형을 비교하였다.

홀트와 윈터스는 예측 방정식과 3개의 평활 방정식(smoothing equation)으로 구성되어 있는 식을 확장시켜 홀트-윈터스(Holt-Winters) 방법을 제안하였다. Billah 등 (2006)에 따르면 홀트-윈터스 방법은 간단한 모형임에도 불구하고 예측력이 비교적으로 괜찮은 평가를 받고 있는 방법이다. 홀트-윈터스 방법은 과거 데이터와 현재 데이터와 가까울수록 과거 데이터의 가중치를 크게 주어 합산한 후 다음 데이터를 예측하는 지수평활법의 여러 기법 중 하나이며, 다른 지수평활법의 종류로 단순 지수평활법, 이중지수 평활법, 삼중지수 평활법 등이 있다. 홀트-윈터스 방법은 단순 지수평활법과 이중 지수평활법의 단점을 보완한 모형으로 계절성과 추세를 동시에 고려하여 추정할 수 있다. 추세와 계절성 둘 다 가법으로 적용한 모형은 가법 홀트-윈터스 방법, 추세는 가법, 계절성은 승법으로 적용한 승법 홀트-윈터스 방법으로 분류된다. Gardner (1985)는 추세와 계절성에 따라 지수평활법을 자세하게 나타내었다. Bianchi 등 (1998)은 텔레 마케팅 센터에서 예산과 계획을 목적으로 전화가 걸려오는 수를 예측하기 위해서 ARIMA 모형과 홀트-윈터스 방법을 이용한 모형을 분석하였으며, 제공된 평균 제공 오차를 통해 예측 정확도를 측정하였다.

Lye 등 (2009)에서는 계절성과 변동성을 수용하는 SARIMA와 GARCH 모형이 있지만, 계절성이 포함된 모형에서 예측의 한계를 보인다고 언급하였다. 이를 극복하기 위해 푸리에 변수 변환 방법을 사용하였다. 푸리에 방법은 시간 도메인 데이터를 주파수 도메인 데이터로 변환할 수 있는 방법이다. 즉, 파형이 있는 진폭 데이터에서 일정한 시간동안 주기가 반복되는 정도를 측정하여 진폭 데이터를 막대그래프로 나타낸 것을 말한다. 푸리에 변환은 사인, 코사인 함수와 같이 주기성을 보이는 데이터에 좋으며, 기계 데이터와 센서 데이터 등에 많이 사용되는 변환이다. Gonzalez-Romera 등 (2008)은 1975년 1월부터 2002년 12월까지 28년간 스페인 월간 전기 수요량 데이터로 하이브리드 접근법인 신경망으로 추세를 예측하고, 푸리에 변환으로 주기를 예측하였다. Kimball (1974)은 날짜에 따라 함수량 등 재배학 변수들로 이루어진 변수들을 평활화 기법을 사용하여 분석하였으며, 푸리에 변환의 사용도 포함되어 있다. Fumi 등 (2013)은 푸리에 변환을 사용하여 패션 회사의 수요 예측을 분석하였다.

본 논문의 구성은 다음과 같다. 2장에서는 전기 사용량 데이터에 대해 설명과 탐색을 한다. 3장에서는 계절주기를 반영한 SARIMA 모형, SARIMA 모형에 변동성 GARCH를 포함한 모형, 지수평활법과 모수를 포함한 가법 홀트-윈터스 모형과 푸리에 변환을 적용한 ARIMA 모형을 적합한 후 서로 비교한다. 마지막으로 4장에서는 간단한 결론을 내린다.

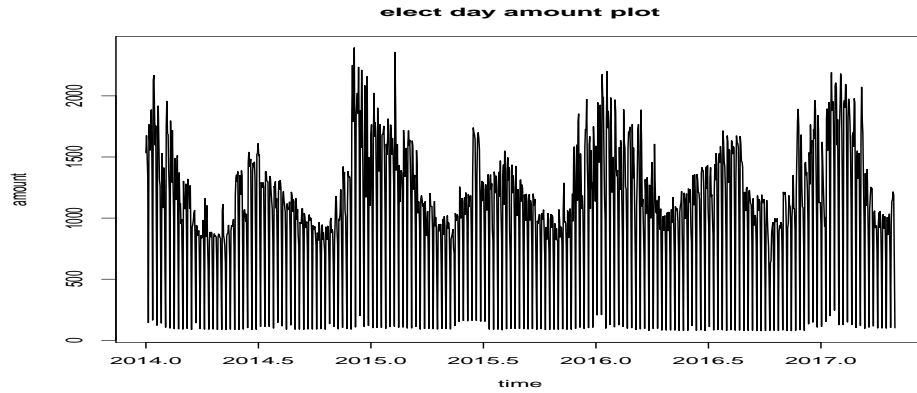


Figure 2.1. Electricity consumption data in GGM, mining industry building.

Table 2.1. Augmented Dickey-Fuller (ADF) test

	ADF statistics	Lag	<i>p</i> -value
Raw data	-4.954	10	<0.01
Log transformation	-4.727	10	<0.01
First differenced	-12.895	10	<0.01
Log transformation and first differenced	-13.132	10	<0.01

2. 전기 사용 데이터

본 논문에서 사용한 데이터는 한국전력공사(KEPCO)에서 제공받았으며, 광공업용 제조업을 하는 ‘GGM’ 건물의 2014년 1월부터 2017년 4월까지 일별 전기 사용 데이터이다. 건물은 경기도 부천시 도당동 소재에 있으며, 주생산품이 금형이다. 데이터를 살펴보면, 보통 휴일에는 공장 가동을 멈추기 때문에 데이터에 대한 주의가 필요하다. 주말과 휴일, 8월 초에 공장가동 중단(휴가, 시설정비 등), 그리고 토요일은 부분 공장가동 등의 이유로 전력이 평소보다 크게 감소하기 때문에 모형 추정에 영향을 줄 수 있다. 공장이 정상적으로 가동하는 평일 전기사용량의 예측을 극대화하기 위해서 데이터 클리닝을 다음과 같이 진행하였다. 주로 전 주와 다음 주 같은 요일의 전기 사용량의 평균을 구하여 대체하였고, 데이터 특성에 따라 일부는 전 날과 다음 날의 전기 사용량의 평균으로 대체하였다. 그리고 평일인데도 불구하고 전기 사용량이 크게 감소한 경우(400kWh 이하)와 2016년 12월 중순부터 1월 중순까지 한 달 데이터가 측정되지 않은 경우에도 이러한 방법으로 클리닝에 포함되었다.

Figure 2.1은 데이터의 움직임을 알아보기 위한 시계열 그림으로 일주일 주기, 크게 보면 1년 주기로 2가지 주기가 겹쳐진 것으로 보인다. 일주일 주기를 살펴보면 토, 일요일에는 평일보다 전기 사용량이 현저히 줄기 때문에 그래프의 진폭이 크고 파장은 짧은 편이다. 그리고 1년 주기를 보면 매년 전기 사용량이 비슷한 편이며, 전력수요가 높은 여름과 겨울이 다른 계절에 비해서 사용량이 높다. 데이터의 로그변환, 차분 등 데이터의 변환이 필요한지 알아보기 위해 Dickey-Fuller (augmented Dickey-Fuller; ADF) 검정을 사용하여 단위근 검정을 하였다. 정상성 검정에 대한 가설은 다음과 같다.

H_0 : 시계열 데이터가 정상성을 만족하지 않는다. 에 대해 H_1 : 정상성을 만족한다.

Table 2.1의 ADF 검정 결과 모두 유의수준이 0.01보다 작아서 귀무가설을 기각하게되어 시계열 데이터가 정상성을 가지고 있다고 볼 수 있다. 그리고 ADF 검정결과, Figure 2.2의 자기상관함수 그래프

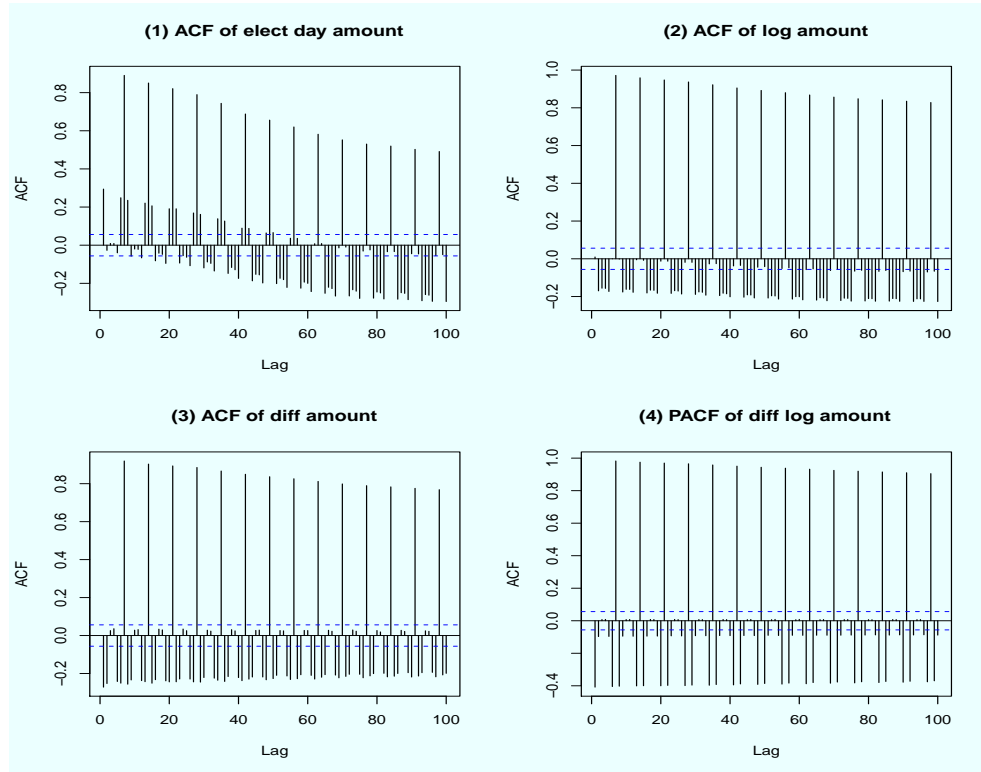


Figure 2.2. ACFs of transformed data. ACF = autocorrelation function.

와 Figure 2.3에서 부분자기상관함수 그래프를 종합해볼 때, 로그 변환된 데이터가 적절해보이므로 이후 분석에서 로그 변환값을 사용하기로 한다. $\{Y_t\}$, $t = 1, 2, \dots, n$ 는 로그 변환된 데이터를 나타낸다.

3. 전기 사용 데이터에 대한 시계열 분석

3.1. SARIMA 모형

정상성을 만족하는 시계열 데이터 $\{Z_1, Z_2, \dots, Z_n\}$ 가 자기회귀과정(autoregressive; AR)과 이동평균 과정(moving average; MA)을 혼합한 형태인 $ARMA(p, q)$ 과정인 경우 다음의 모형

$$Z_t = \mu + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

으로 표현한다. 여기서 $\{a_t\}$ 는 백색잡음과정(white noise process)으로 서로 독립이고 평균 0, 분산 σ_a^2 인 확률 변수이다. $ARMA(p, q)$ 과정에서 AR 차수 p 의 결정에는 부분자기상관함수(partial autocorrelation function; PACF)를 MA 차수 q 의 결정에는 자기상관함수(autocorrelation function; ACF)를 이용한다.

차분(difference)이란 현 시점의 데이터에서 전 시점의 데이터를 빼주는 것을 말한다. 평균이 일정하지 않은 비정상 시계열은 차분을 통하여 정상 시계열로 바꿀 수 있다. d 번 차분한 $\nabla^d Z_t \equiv W_t$ 가 $ARMA(p, q)$ 모형을 따를 때 Z_t 가 $ARIMA(p, d, q)$ 모형을 따른다고 한다. 즉 W_t 가 통계적 모형

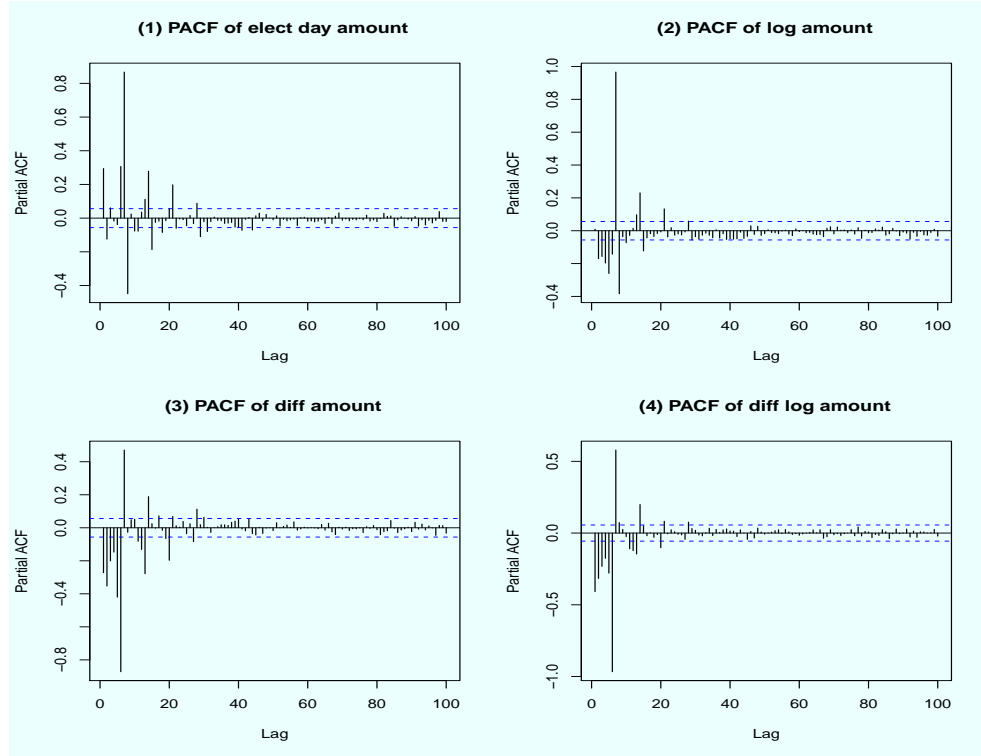


Figure 2.3. PACFs of transformed data. PACF = partial autocorrelation function.

Table 3.1. Comparison of ARIMA parts in SARIMA models

Model	$\hat{\sigma}^2$	Likelihood	AIC	Ljung-Box	RMSE	ER
SARIMA(0, 0, 1) \times (0, 1, 6) ₇	0.02305	549.57	-1085.13	0.00003	0.1514	0.0177
SARIMA(0, 0, 2) \times (0, 1, 6) ₇	0.02161	588.53	-1161.06	0.17250	0.1466	0.0171
SARIMA(0, 0, 3) \times (0, 1, 6) ₇	0.02115	601.57	-1185.14	0.50030	0.1450	0.0170
SARIMA(1, 0, 0) \times (0, 1, 6) ₇	0.02051	619.96	-1225.93	0.00783	0.1428	0.0168
SARIMA(1, 0, 1) \times (0, 1, 6) ₇	0.01979	641.41	-1266.81	0.10140	0.1403	0.0165
SARIMA(1, 0, 2) \times (0, 1, 6) ₇	0.01942	654.34	-1290.67	0.75540	0.1390	0.0164
SARIMA(2, 0, 0) \times (0, 1, 6) ₇	0.02013	631.15	-1246.31	0.53770	0.1415	0.0166

ARIMA = auto-regressive intergrated moving average; SARIMA = seasonal ARIMA; AIC = Akaike information criteria; RMSE = root mean square error; ER = error rate.

ARMA(p, q)

$$W_t = \phi_1 W_{t-1} + \cdots + \phi_p W_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}$$

를 만족한다.

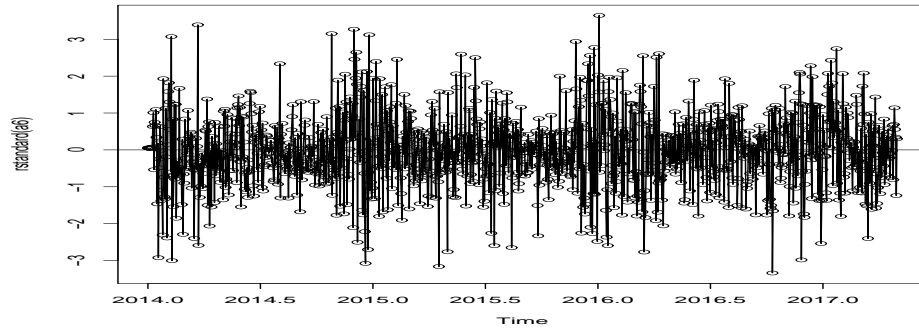
ARIMA(p, d, q) 모델을 따르며 주기 s 를 갖고 D 번 주기적 차분한 Z_t 가 ARMA(P, Q) 모델을 따를 때 Z_t 는 SARIMA(p, d, q) \times (P, D, Q) _{s} 모델을 따른다고 한다. 여기서 $\nabla_s^1 Z_t \equiv Z_t - Z_{t-s}$ 이다.

전기 데이터에 대해 SARIMA 모델을 적합한 결과는 Table 3.1에서 보여준다. 모형에 대한 비교로

Table 3.2. Comparison of seasonal parts in SARIMA models

Model	$\hat{\sigma}^2$	Likelihood	AIC	Ljung-Box	RMSE	ER
SARIMA(1, 0, 2)×(0, 1, 4) ₇	0.02037	629.91	-1245.81	0.7657	0.1423	0.0169
SARIMA(1, 0, 2)×(0, 1, 5) ₇	0.01972	647.83	-1279.66	0.7447	0.1400	0.0166
SARIMA(1, 0, 2)×(0, 1, 6) ₇	0.01942	654.34	-1290.67	0.7554	0.1390	0.0164
SARIMA(1, 0, 2)×(0, 1, 7) ₇	0.01936	654.79	-1289.58	0.7594	0.1388	0.0163

SARIMA = seasonal auto-regressive intergrated moving average; AIC = Akaike information criteria; RMSE = root mean square error; ER = error rate.

**Figure 3.1.** Standardized residual plot of SARIMA(1, 0, 2)×(0, 1, 6)₇. SARIMA = seasonal auto-regressive intergrated moving average.

우도함수 기반 오차 분산 추정량 $\hat{\sigma}^2$, Akaike information criteria (AIC) 통계량, 제곱근 평균제곱오차(root mean square error; RMSE)와 오차율(error rate; ER)이 작고 로그우도가 크며 Ljung-Box 검정으로 기각하지 않는 모형으로 추천할 수 있다. 모형의 과적합을 고려하여 p 와 q 가 3 이상인 모형은 고려대상에서 제외되었다.

Figure 2.2에서 각 lag 값에 대한 자기상관함수 ACF 모양은 서서히 내려가는 모양을 보이고 일정한 패턴이 있음을 확인할 수 있다. 또한 PACF 그래프는 lag = 7, 14, 21로 lag가 7씩 증가할 때 마다 PACF 값이 기하급수적으로 줄어들고 있으므로 주기 7을 결정하고 계절성 부분에서 AR 부분은 적합하지 않기 때문에 모형에서 추정되지 않았다. ARIMA 모형에 주기를 포함한 모형으로 SARIMA(p, d, q) × (0, 1, Q)를 고려하고자 한다. Table 3.1에서 비교한 모형 중 SARIMA(1, 0, 2) × (0, 1, 6)₇을 선택하고자 한다. Table 3.2은 ARIMA(1, 0, 2) 선택 후 계절성 부분의 탐색으로 모수결약의 원칙을 적용하여 SARIMA(1, 0, 2) × (0, 1, 7)₇ 모형보다 SARIMA(1, 0, 2) × (0, 1, 6)₇을 선택하고자하며 추정된 모형은 다음과 같다.

$$\begin{aligned}
 & Y_t - 0.94Y_{t-1} - Y_{t-12} + 0.94Y_{t-13} \\
 &= \epsilon_t - 0.75\epsilon_{t-12} - 0.08\epsilon_{t-24} + 0.003\epsilon_{t-36} + 0.02\epsilon_{t-48} - 0.09\epsilon_{t-60} - 0.11\epsilon_{t-72} \\
 &\quad - 0.49(\epsilon_{t-1} - 0.75\epsilon_{t-13} - 0.08\epsilon_{t-25} + 0.003\epsilon_{t-37} + 0.02\epsilon_{t-49} - 0.09\epsilon_{t-61} - 0.11\epsilon_{t-73}) \\
 &\quad - 0.19(\epsilon_{t-2} - 0.75\epsilon_{t-14} - 0.08\epsilon_{t-26} + 0.003\epsilon_{t-38} + 0.02\epsilon_{t-50} - 0.09\epsilon_{t-62} - 0.11\epsilon_{t-74}), \quad (3.1)
 \end{aligned}$$

여기서 $\{\epsilon_t\}$ 는 백색잡음과정으로 서로 독립이고 평균 0, 분산 σ_ϵ^2 인 확률 변수이다.

식 (3.1)에 대한 모형진단을 하기 위해서 표준화 잔차 그래프와 잔차의 ACF, Ljung-Box 검정의 결과를 검토하고자 한다. 표준화 잔차 그래프로 Figure 3.1를 보면 약 10개의 점들이 -3과 3을 벗어난다.

Table 3.3. Comparison of SARIMA(1, 0, 2) × (0, 1, 6)₇ + GARCH models

Model	Information criterion		Ljung-Box test		Error	
	AIC	BIC	Q(10) for a_t^2	p-value	RMSE	ER
SARIMA(1, 0, 2) × (0, 1, 6) ₇ + GARCH(1, 0)	-1.1326	-1.1200	9.4849	0.4868	0.1390	0.0164
SARIMA(1, 0, 2) × (0, 1, 6) ₇ + GARCH(0, 1)	-1.1129	-1.1003	21.8332	0.0160	0.1403	0.0165
SARIMA(1, 0, 2) × (0, 1, 6) ₇ + GARCH(2, 0)	-1.1567	-1.1399	10.6513	0.3853	0.1390	0.0164
SARIMA(1, 0, 2) × (0, 1, 6) ₇ + GARCH(1, 1)	-1.1826	-1.1659	11.8795	0.2932	0.1390	0.0164
SARIMA(1, 0, 2) × (0, 1, 6) ₇ + GARCH(1, 2)	-1.1818	-1.1608	11.8361	0.2962	0.1390	0.0164
SARIMA(1, 0, 2) × (0, 1, 6) ₇ + GARCH(2, 1)	-1.1807	-1.1597	11.8305	0.2966	0.1390	0.0164
SARIMA(1, 0, 2) × (0, 1, 6) ₇ + GARCH(2, 2)	-1.1802	-1.1550	11.8361	0.2962	0.1390	0.0164

SARIMA = SARIMA = seasonal auto-regressive intergrated moving average; GARCH = generalized autoregressive conditional heteroskedasticity; AIC = Akaike information criteria; BIC = Bayesian information criteria; RMSE = root mean square error; ER = error rate.

여기서 $(-3, 3)$ 을 벗어나는 경우는 극단값으로 판단할 수 있다. 그리고 연말과 연초에는 잔차가 커지는 편이고 6-8월에는 데이터의 잔차가 적은 편으로 잔차의 일정한 패턴이 보인다고 할 수 있다. 잔차에 대한 ACF에서는 자기상관성이 유의하지 않으나 잔차 제곱에 대한 ACF와 잔차의 절대값에 대한 ACF는 lag = 150까지 자기상관성을 나타내므로 (잔차에 대한 ACF 그림은 생략) 모형에 변동성을 포함한 조절이 필요해 보인다. 잔차에 대한 Ljung-Box 검정에 대한 가설은 다음과 같다.

H_0 : 시계열 데이터의 상관관계가 없다. 에 대해 H_1 : 상관관계가 있다.

Table 3.2에서 SARIMA(1, 0, 2) × (0, 1, 6)₇ 모형에 대한 Ljung-Box 검정 결과 p -값 = 0.7754 > α = 0.05으로 유의수준 5%에서 귀무가설을 기각하지못해 잔차에 대해 상관성은 없다고 할 수 있다.

3.2. SARIMA + GARCH 모형

Engle (1982)는 조건부 분산이 시간에 따라 변동하는 autoregressive conditional heteroskedasticity (ARCH) 모형을 발견하였다. 그 이후 Bollerslev (1986)는 ARCH 모형의 일반화로 GARCH 모형을 정의하였다. 그 이후로도 integrated GARCH (IGARCH), exponential GARCH (EGARCH), threshold GARCH (TGARCH) 등 여러 가지의 변형된 ARCH 모형들이 등장하였으며, 모형에 대한 간단한 설명과 데이터를 적용한 결과는 Ali (2013)의 논문에서 설명되어 있다. ARCH 효과에 대한 검정을 하기 위해서는 여러 방법이 있다. Engle (1982)는 라그랑즈 승수 검정, Broock 등 (1996)은 BDS 검정을 언급하였다. 여기서는 McLeod와 Li (1983)가 제안한 McLeod-Li 검정을 적용하고자 한다.

변동성을 갖는 확률변수 $\{a_t\}$ 에 대해 GARCH(p, q) 모형은 다음과 같다:

$$a_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i a_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2,$$

여기서 $\omega > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$ 을 만족하며 $\{\epsilon_t\}$ 는 평균 0, 분산 1을 갖는 서로 독립이고 동일한 분산을 갖는 확률 변수이다.

ARCH 효과 존재 여부에 대한 검정으로 McLeod-Li 검정에 대한 가설은 다음과 같다.

H_0 : ARCH의 효과가 없다. 에 대해 H_1 : ARCH의 효과가 있다.

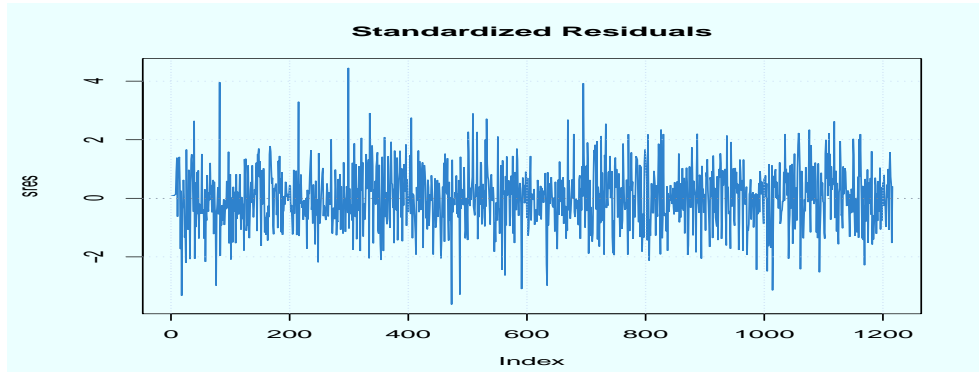


Figure 3.2. Standardized residual plot of $SARIMA(1, 0, 2) \times (0, 1, 6)_7 + GARCH(1, 1)$. SARIMA = seasonal autoregressive integrated moving average.

전기 사용 데이터에 대해 $SARIMA(1, 0, 2) \times (0, 1, 6)_7$ 모형 적합후 잔차에 대해 변동성 모형을 추가하고자 한다. McLeod-Li 검정 결과 그림(그림 생략)을 보면 30 이하 lag에서 모두 p -값이 유의수준 0.05 보다 매우 작아 귀무가설을 기각하게 된다. 초기 4개 lag에서 p -값은 다음과 같다.

$$\text{lag1} : 4.56 \times 10^{-11}, \quad \text{lag2} : 1.11 \times 10^{-11}, \quad \text{lag3} : 6.16 \times 10^{-13}, \quad \text{lag4} : 5.63 \times 10^{-14}.$$

그러므로 ARCH 모형의 효과가 있으므로 오차에 대한 변동성 모형을 적용해야 함을 알 수 있다. 변동성에 대해 적절한 GARCH 모형을 찾고자 한다.

Table 3.3은 SARIMA + GARCH 모형에 대한 비교 결과 $SARIMA(1, 0, 2) \times (0, 1, 6)_7 + GARCH(1, 1)$ 모형이 선택되고 적합된 모형은 식 (3.2)와 같다. 그리고 Ljung-Box 검정 결과 시계열 데이터에 대해 상관성이 없다는 귀무가설을 유의수준 5% 에서 기각하지 않는다(p -값 = 0.2932).

$$\begin{aligned} Y_t &- 0.943Y_{t-1} - Y_{t-12} + 0.943Y_{t-13} \\ &= \epsilon_t - 0.751\epsilon_{t-12} - 0.076\epsilon_{t-24} + 0.003\epsilon_{t-36} + 0.024\epsilon_{t-48} - 0.086\epsilon_{t-60} - 0.109\epsilon_{t-72} \\ &\quad - 0.489(\epsilon_{t-1} - 0.751\epsilon_{t-13} - 0.076\epsilon_{t-25} + 0.003\epsilon_{t-37} + 0.024\epsilon_{t-49} - 0.086\epsilon_{t-61} - 0.109\epsilon_{t-73}) \\ &\quad - 0.185(\epsilon_{t-2} - 0.751\epsilon_{t-14} - 0.076\epsilon_{t-26} + 0.003\epsilon_{t-38} + 0.024\epsilon_{t-50} - 0.086\epsilon_{t-62} - 0.109\epsilon_{t-74}), \\ \epsilon_t &= \sigma_t a_t, \\ \sigma_t^2 &= 0.00091 + 0.0846\epsilon_{t-1}^2 + 0.8687\sigma_{t-1}^2, \end{aligned} \quad (3.2)$$

여기서 a_t 는 서로 독립이며 평균 0, 분산 1인 확률변수이다.

선택한 모형 (3.2)에 대한 모형진단으로 Figure 3.2는 표준화 잔차그림이고 Figure 3.3은 표준화 잔차에 대한 Q-Q 그림으로 GARCH 모형을 포함하지 않은 모형에 비해 꼬리부분의 두꺼운 정도가 약해진 편이다. Figure 3.4는 식 (3.2) 모형에서 표준화 잔차의 ACF를 lag = 30까지 보여준다.

Figure 3.5은 2014년도부터 2017년 4월까지 $SARIMA(1, 0, 2) \times (0, 1, 6)_7 + GARCH(1, 1)$ 모형으로 추정값과 실제 데이터에 대한 그림이며, 좀 더 자세한 부분을 보여주기위해 Figure 3.6은 2014년도 1년간 그림이다. GARCH가 포함되지 않은 모형에 비해, 값이 1500kWh 이상 높거나 1000kWh 정도로 낮을 때, GARCH를 포함한 모형이 좀 더 적합이 잘 되어보인다. 이는 GARCH(1, 1) 모형이 SARIMA 모형에 추가되면서 변동성을 설명해주기 때문인 것으로 여겨진다.

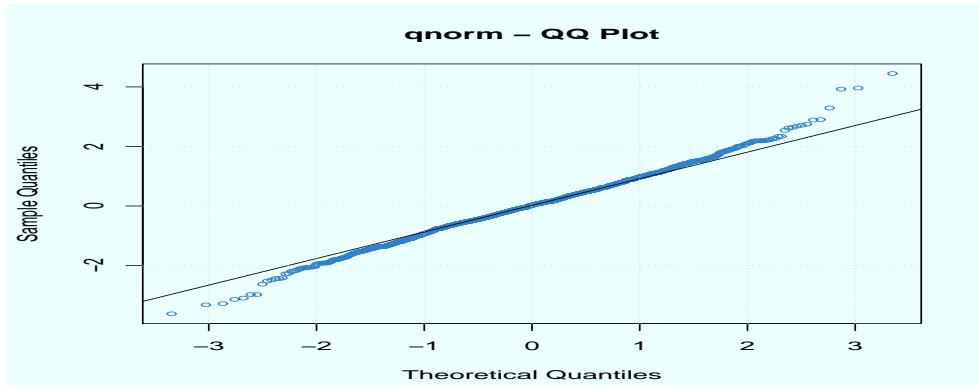


Figure 3.3. Q-Q plot of SARIMA(1,0,2)×(0,1,6)₇ + GARCH(1,1). SARIMA = seasonal auto-regressive intergrated moving average; GARCH = generalized autoregressive conditional heteroskedasticity.

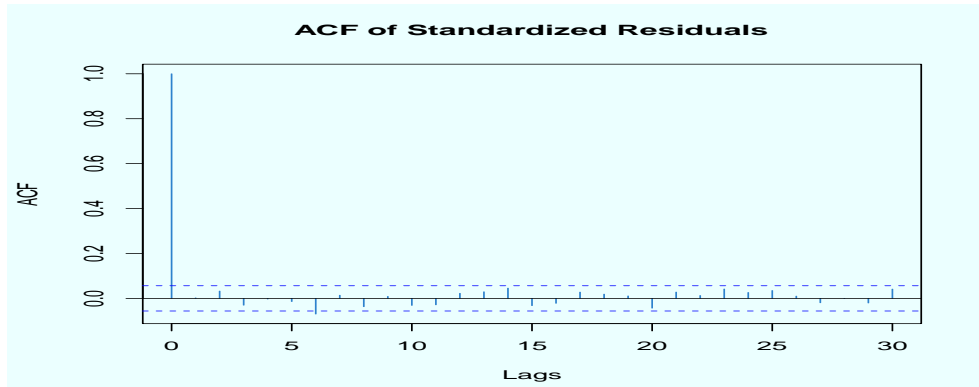


Figure 3.4. ACF of standardized residuals in SARIMA(1,0,2)×(0,1,6)₇ + GARCH(1,1). ACF = autocorrelation function; SARIMA = seasonal auto-regressive intergrated moving average; GARCH = generalized autoregressive conditional heteroskedasticity.

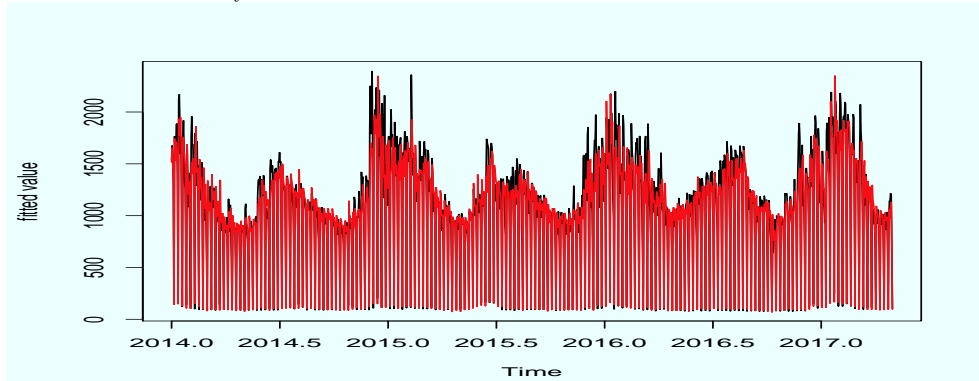


Figure 3.5. Fitted values of SARIMA(1,0,2)×(0,1,6)₇ + GARCH(1,1). SARIMA = seasonal auto-regressive intergrated moving average; GARCH = generalized autoregressive conditional heteroskedasticity.

3.3. 홀트-윈터스 모형

승법과 가법 홀트-윈터스 방법 중 데이터가 일정하게 증가하는 추세를 나타낼 때 가법을 사용하는 것이

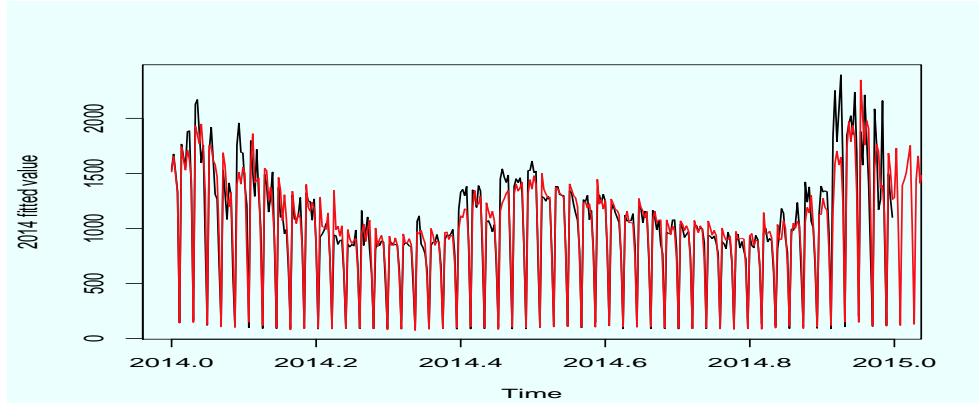


Figure 3.6. Fitted values of SARIMA(1,0,2)×(0,1,6)₇ + GARCH(1,1) in 2014. SARIMA = seasonal autoregressive intergrated moving average; GARCH = generalized autoregressive conditional heteroskedasticity.

권장되므로 다음과 같이 가법 홀트-윈터스 방법을 적합하고자 한다.

$$\begin{aligned}
 L_t &= \alpha(Y_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + b_{t-1}) \\
 b_t &= \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1} \\
 S_t &= \gamma(Y_t - L_t) + (1 - \gamma)S_{t-s} \\
 F_{t+m} &= L_t + b_tm + S_{t-s+m},
 \end{aligned} \tag{3.3}$$

여기서 α 는 평균수준, β 는 추세, γ 는 계절성의 모수이며, F_{t+m} 는 현재 t 시점에서 m 시점을 더 앞서 예측한 값이다. Y_t 는 t 시점에서의 관측값, L_t 는 t 시점에서의 평균수준, b_t 는 t 시점에서의 추세, S_t 는 t 시점에서의 계절성이며, s 는 계절성의 주기이다. 전기 데이터에 적합한 결과 추정된 모형은 식 (3.4)이다.

$$\begin{aligned}
 L_t &= L_{t-1} + b_{t-1} \\
 b_t &= b_{t-1} \\
 S_t &= 0.73077(Y_t - L_t) + 0.26923S_{t-s} \\
 F_{t+m} &= L_t + b_tm + S_{t-s+m}.
 \end{aligned} \tag{3.4}$$

평균수준과 추세는 과거 데이터의 영향을 받지 않지만 계절성에는 영향을 준다. 모형 (3.4)로 전기 데이터를 적합한 결과는 Figure 3.7에서 보여준다. Figure 3.7에서 보면 적합값과 실제값이 거의 같기 때문에 두 선이 겹쳐져 검정색 선으로 보인다. 그러나 전기사용량이 봉우리에서 높은 값(피크)이 되기 직전의 적합값이 갑자기 높아지는 경향을 보인다. 그리고 전체적으로 추세가 점차적으로 증가한다. 직전 관측값이 증가하면 현재 추정값이 그 값에 영향을 받기 때문에 평일의 값의 영향으로 주말의 값이 높아진다. 또한 주말의 값이 높아지는 것의 영향으로 다음 주 평일 값도 증가하는 패턴의 반복을 보이는 것이라 짐작된다.

3.4. 푸리에 변환을 사용한 ARIMA 모형

주기가있는 특정한 점을 적합하기 위해서는 Walker (1991)가 제안하는 푸리에 변환(Fourier transformation)식을 이용하여 이산형 분해를 하고자 한다. 데이터 $\{Z_t\}$ 를 푸리에 표현으로 나타내면 식

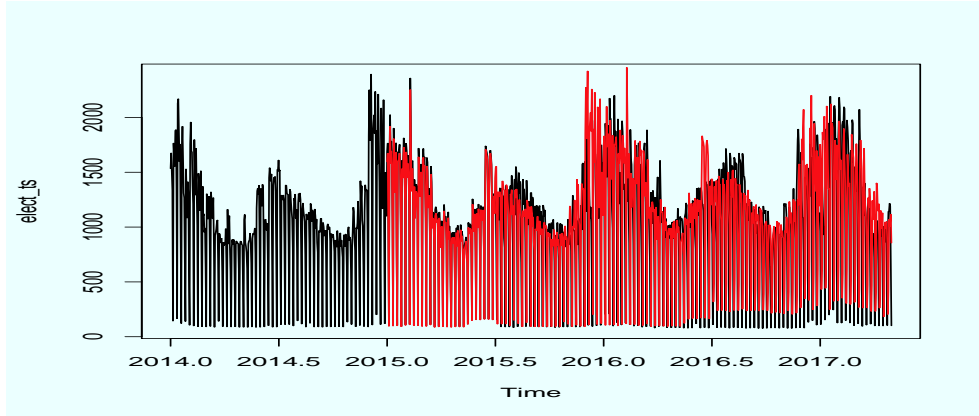


Figure 3.7. Fitted values of additive Holt-Winters model.

(3.5)와 같다.

$$Z_t = a_0 + \sum_{k=1}^K \left(a_k \cos \frac{2\pi k}{n} t + b_k \sin \frac{2\pi k}{n} t \right) + \epsilon_t, \quad (3.5)$$

여기서 a_0 는 계절성 주기의 평균, K 는 푸리에 변환 예측을 위한 각각의 주기로 고려되는 주파수의 주기이다. ϵ_t 는 평균이 0이고 분산이 σ^2 인 서로 독립인 오차항이다.

$$a_k = \frac{2}{n} \sum_{t=1}^n \left(Z_t \cos \frac{2\pi k}{n} t \right), \quad (3.6)$$

$$b_k = \frac{2}{n} \sum_{t=1}^n \left(Z_t \sin \frac{2\pi k}{n} t \right). \quad (3.7)$$

모형에 대한 자세한 설명은 Asante-Darko 등 (2016)을 참고한다. 이 절에서는 SARIMA + Fourier 모형으로 계절적 주기를 가진 시계열에 대한 ARIMA 모형을 적합하고 나머지 주기에 대해서는 푸리에 변환 적합활용을 통해 모형의 적합도를 높이하고자 한다. 실제 데이터 분석 연구로 Huang 등 (2014)은 뉴질랜드 관광산업수요 예측을 하기위해 SARIMA + Fourier 방법으로 적합하였다.

전기 사용 데이터에 적용할 ARIMA 모형은 R의 함수 `auto.arima()`에서 선택된 ARIMA(2,0,2)를 선택한다. 각각 7일과 1년으로 2개의 주기가 중복된 모형이므로 ARIMA(2,0,2)에 대해 푸리에 변환을 한 모형은 식 (3.8)과 같다.

$$Y_t + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} = \epsilon_t - \beta_1 \epsilon_{t-1} - \beta_2 \epsilon_{t-2} + \sum_{k=1}^{K_1} \left(a_k \cos \frac{2\pi k}{n_1} t + b_k \sin \frac{2\pi k}{n_1} t \right) + \sum_{k=1}^{K_2} \left(p_k \cos \frac{2\pi k}{n_2} t + q_k \sin \frac{2\pi k}{n_2} t \right), \quad (3.8)$$

여기서 K_1 과 K_2 의 수는 각 주기의 절반 값인 $7/2$, $365/2$ 보다 작은 수로 사용해야 한다. n_1, n_2 는 주기를 나타내며 n_1 은 7일, n_2 는 365일로 모형설정을 하였다. K 값에 따라서 모수의 개수, 모형의 진폭과 안정성이 결정되므로 적당한 값을 찾아야 한다. K 값이 클수록 모형의 과적합 위험이 있으므로 K 값의 선택에 신중해야 한다.

Table 3.4. Comparison of ARIMA + Fourier models

Model	RMSE	ER
$K_1 = 1, K_2 = 1$	0.450213	0.058997
$K_1 = 2, K_2 = 1$	0.300636	0.038675
$K_1 = 3, K_2 = 1$	0.148641	0.018020
$K_1 = 3, K_2 = 10$	0.144062	0.017511
$K_1 = 3, K_2 = 30$	0.141258	0.017053
$K_1 = 3, K_2 = 50$	0.137843	0.016720
$K_1 = 3, K_2 = 100$	0.130972	0.015983

ARIMA = auto-regressive intergrated moving average; RMSE = root mean square error; ER = error rate.

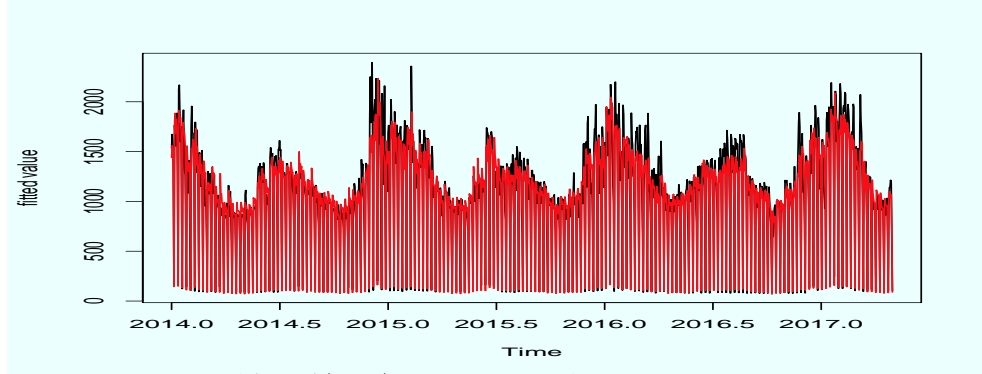
**Figure 3.8.** Fitted values of ARIMA(2,0,2) with Fourier transformation.

Table 3.4는 K_1 과 K_2 의 수에 따라 제공된 평균제곱오차와 오차율을 보여준다. K_1 이 증가할수록 모형의 제공된 평균제곱오차와 오차율이 두드러지게 줄어들기 때문에 $K_1 = 3$ 이 적당하며, K_2 도 증가할수록 모형이 좋아지지만 K_2 의 수가 과도하게 크면 과적합의 위험이 있다. $K_2 = 30$ 이후로 감소폭이 줄기 때문에 $K_2 = 30$ 인 모형이 가장 적합하다. $K_1 = 3, K_2 = 30$ 으로 푸리에 변환을 한 ARIMA(2,0,2)의 적합된 식은 식 (3.9)이며, 적합된 그래프는 Figure 3.8에서 보여준다.

$$\begin{aligned}
 & Y_t - 0.1721Y_{t-1} - 0.6850Y_{t-2} \\
 &= \epsilon_t + 0.0976\epsilon_{t-1} - 0.6674\epsilon_{t-2} + 6.7041 + 0.7332\sin \frac{2\pi}{7}t + 0.2823\cos \frac{2\pi}{7}t \\
 &\quad - 0.4166\sin \frac{4\pi}{7}t + 0.5748\cos \frac{4\pi}{7}t - 0.4666\sin \frac{6\pi}{7}t - 0.4055\cos \frac{6\pi}{7}t \\
 &\quad + 0.0392\sin \frac{2\pi}{365}t + 0.1023\cos \frac{2\pi}{365}t + \cdots - 0.0072\sin \frac{60\pi}{365}t + 0.0071\cos \frac{60\pi}{365}t. \quad (3.9)
 \end{aligned}$$

Table 3.5에서는 각 모형의 예측값, 제공된 평균제곱오차, log를 취한 시계열에 대한 적합 오차율(log fitted error rate), 예측 오차율(forecast error rate)을 보여준다. 전반적으로 보면 SARIMA + GARCH 모형을 추천하지만 SARIMA 모형과 근소한 차이를 보여 상대적으로 변동성이 작은 편이라고 볼 수 있다. Figures 3.9와 3.10에서는 각 모형으로 예측한 결과를 선으로 나타내며 실제값은 굵은 실선으로 표시된다. SARIMA와 SARIMA + GARCH 모형은 예측값이 비슷해보인다. 가법적 홀트-윈터스 방법으로는 다른 모형에 비해서 주말 예측값이 높게 예측되며 평일 예측값은 작은 편이다. 푸리에 변환을 적용한 ARIMA 모형에서는 주말은 SARIMA, SARIMA + GARCH 모형보다 예측값이 작은 경향을 보인다.

Table 3.5. Comparison of forecasted values in one month ahead in May 2017

Date	Actual values	Model with predicted values			
	Electricity consumption	SARIMA (1, 0, 2) × (0, 1, 6) ₇	SARIMA (1, 0, 2) × (0, 1, 6) ₇ + GARCH(1, 1)	Addition Holt-Winters	Fourier + ARIMA(2, 0, 2)
2017.05.01(월)	1169.635	1108.629	1102.571	241.284	1140.245
2017.05.02(화)	958.860	1176.828	1170.397	872.134	1094.476
2017.05.03(수)	1151.418	1070.486	1064.637	1094.621	1030.903
2017.05.04(목)	977.930	1165.640	1159.271	1043.150	1052.753
2017.05.05(금)	1148.687	1140.042	1133.812	1166.060	1013.400
2017.05.06(토)	768.483	737.737	733.706	1052.845	683.386
2017.05.07(일)	98.350	99.416	98.873	839.838	84.010
2017.05.08(월)	1102.920	1157.645	1151.320	299.447	1058.957
2017.05.09(화)	1217.365	1216.655	1210.006	890.660	1063.334
2017.05.10(수)	1027.470	1112.711	1106.631	1064.666	981.907
2017.05.11(목)	1051.860	1203.373	1196.797	1083.844	1029.492
2017.05.12(금)	1130.880	1177.865	1171.429	1083.575	979.479
2017.05.13(토)	755.160	762.984	758.815	1021.409	669.287
2017.05.14(일)	102.200	100.421	99.873	806.448	81.690
2017.05.15(월)	1123.080	1210.714	1204.098	314.988	1039.242
2017.05.16(화)	1304.260	1248.770	1241.947	865.342	1045.265
2017.05.17(수)	1173.670	1157.584	1151.259	1151.303	979.769
2017.05.18(목)	1123.450	1230.271	1223.548	1152.583	1042.053
2017.05.19(금)	1026.220	1200.521	1193.961	1226.594	1015.144
2017.05.20(토)	810.020	785.622	781.329	992.194	709.731
2017.05.21(일)	100.930	101.089	100.537	742.997	88.847
2017.05.22(월)	1227.310	1242.240	1235.452	326.945	1153.305
2017.05.23(화)	1368.650	1276.207	1269.234	989.033	1178.382
2017.05.24(수)	1321.940	1205.994	1199.404	1212.086	1112.648
2017.05.25(목)	1380.180	1268.023	1261.094	1085.947	1184.176
2017.05.26(금)	1288.960	1218.095	1211.439	1175.940	1146.679
2017.05.27(토)	740.270	798.982	794.617	1248.260	794.258
2017.05.28(일)	101.730	101.298	100.745	968.358	98.370
2017.05.29(월)	1225.230	1268.188	1261.259	340.070	1265.162
2017.05.30(화)	1237.690	1291.683	1284.625	1016.034	1283.569
2017.05.31(수)	1082.590	1217.664	1211.010	1271.069	1206.112
RMSE		0.13896	0.13901	1.09317	0.14126
Log fitted ER		0.01637	0.01635	0.11233	0.01705
Forecast ER		0.00886	0.00884	0.13137	0.01618

ARIMA = auto-regressive intergrated moving average; SARIMA = seasonal ARIMA; GARCH = generalized autoregressive conditional heteroskedasticity; RMSE = root mean square error; ER = error rate.

시계열 분석의 목적 중 하나는 미래에 대한 예측이다. 본 연구에서 적합해본 모형을 이용하여 모형구축에 사용된 시계열에 대한 추정값을 구해 추정오차를 구하였다. 예측으로는 각 모형에 의해 미래 한 달간의 예측값을 얻은 후 실제로 얻은 값과 비교하였다. Figure 3.9는 2017년 1월부터 4월까지의 모형 구축시 사용된 시계열 데이터(최근 4 개월간 데이터)와 5월에 대한 각 모형의 예측값을 실제값과 함께 보여준다. 2017년 5월의 전기사용량은 모형구축시 사용되지 않았으므로 예측후 실제값과 비교가능하다. 한

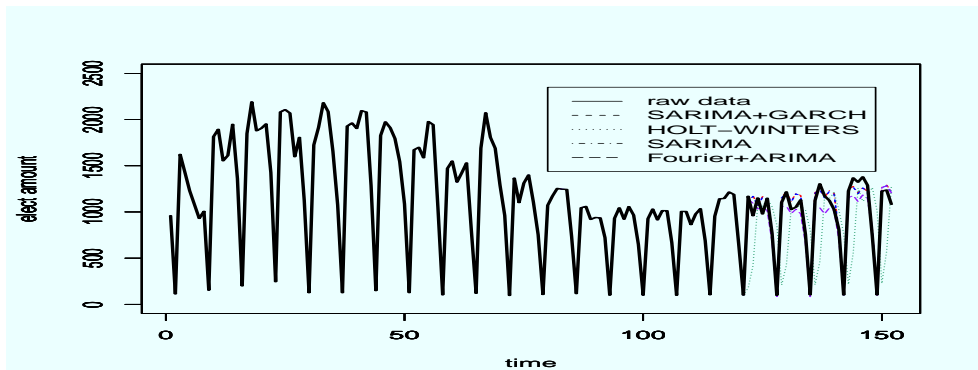


Figure 3.9. Comparison plot of forecasting for each model in 2017. SARIMA = seasonal auto-regressive intergrated moving average; GARCH = generalized autoregressive conditional heteroskedasticity.

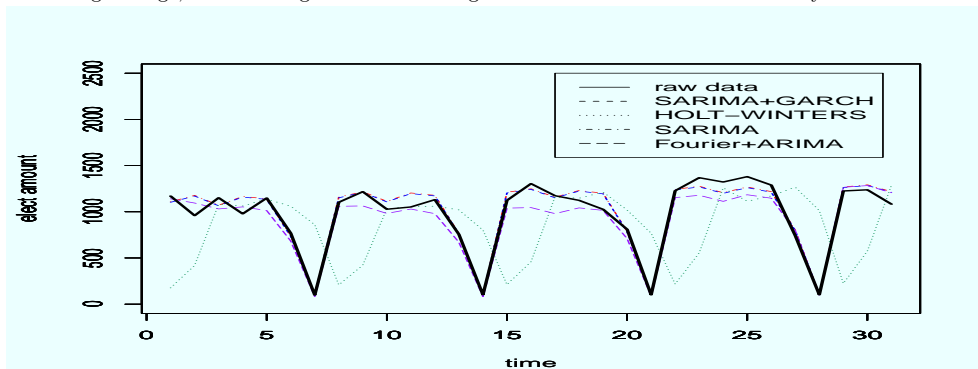


Figure 3.10. Comparison plot of forecasting for each model in May 2017. SARIMA = seasonal auto-regressive intergrated moving average; GARCH = generalized autoregressive conditional heteroskedasticity.

달간 예측부분을 좀 더 자세히 살펴보기 위해 Figure 3.10은 2017년 5월의 전기사용량 예측 그래프를 실제 사용량과 함께 나타낸 그래프이다. Figure 3.10을 보면 전반적으로 일주일 주기를 바탕으로 평일에는 사용량이 증가하고 주말에는 감소하는 예측을 보여준다.

4. 결론

본 연구에서는 공업용 건물의 전기 사용 데이터에 대한 시계열 분석으로 적절한 모형을 찾고 이를 기반으로 사용량 예측을 하였다. 로그변환 데이터를 사용하여 자기상관, 주기와 변동성을 고려한 모형으로 SARIMA 모형, SARIMA + GARCH 모형, 홀트-윈터스 모형과 푸리에 주기성을 추가로 적용한 모형을 추정하고 예측을 비교하였다. 한 달간 전기 사용량에 대한 예측 결과 SARIMA + GARCH 모형의 예측 오차율이 가장 작았으며 이는 변동성이 예측에 영향을 준 것으로 여겨진다. 가중지수평활법을 사용한 홀트-윈터스 모형은 변동성이 심한 전기 데이터에 대한 적합도가 약간 떨어지는 편이다. 주기가 고려된 푸리에 변환을 적용한 ARIMA 모형은 대체로 적합이 잘 되었지만, 평일의 예측값이 실제값에 비해 작게 예측되어 예측 오차율이 약간 높았다.

일반적으로 전기 사용량은 건물의 이용 특성에 따라 다른 패턴을 보여준다. 그러므로 건물의 전기사용적 특성과 데이터 특징을 잘 파악하고 탐색적 분석을 통해 적절한 모형을 찾아야한다. 전기사용량 분석은 전기수급 조절과 사용 예측 및 준비 등에 매우 중요한 연구이므로 지속적인 데이터 수집과 분석이 필

요하다. 또한 전기 사용량은 외부상황 즉 기온, 습도, 휴일여부 등에 영향을 받으므로 향후 연구로는 이러한 설명변수들을 포함한 모형 기반 시계열 분석을 기대한다.

References

- Abdel-Aal, R. E, and Al-Garni, A. Z. (1997). Forecasting monthly electric energy consumption in eastern Saudi Arabia using univariate time-series analysis, *Energy*, **22**, 1059–1069.
- Ali, G. (2013). EGARCH, GJR-GARCH, TGARCH, AVGARCH, NGARCH, IGARCH and APARCH models for pathogens at marine recreational sites, *Journal of Statistical and Econometric Methods*, **2**, 57–73.
- Asante-Darko, D., Adabor, E. S., and Amponsah, S. K. (2016). A Fourier series model for forecasting solid waste generation in the Kumasi metropolis of Ghana, *WIT Transactions on Ecology and the Environment*, **202**, 173–185.
- Bianchi, L., Jarrett, J., and Hanumara, R. C. (1998). Improving forecasting for telemarketing centers by ARIMA modeling with intervention, *International Journal of Forecasting*, **14**, 497–504.
- Billah, B., King, M. L., Snyder, R. D., and Koehler, A. B. (2006). Exponential smoothing model selection for forecasting, *International journal of forecasting*, **22**, 239–247.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**, 307–327.
- Broock, W. A., Scheinkman, J. A., Dechert, W. D., and LeBaron, B. (1996). A test for independence based on the correlation dimension, *Econometric Reviews*, **15**, 197–235.
- Choi, B., Kang, H., Lee, K. Y., and Han, S. T. (2009). A development of time-series model for City Gas demand forecasting, *The Korean Journal of Applied Statistics*, **22**, 1019–1032.
- Chujai, P., Kerdprasop, N., and Kerdprasop, K. (2013). Time series analysis of household electric consumption with ARIMA and ARMA models, *the International MultiConference of Engineers and Computer Scientists*, **1**, 295–300.
- Cryer, J. and Chan, K. (2008). *Time Series Analysis 2nd*, Springer, New York.
- 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.
- Fumi, A., Pepe, A., Scarabotti, L., and Schiraldi, M. M. (2013). Fourier analysis for demand forecasting in a fashion company, *International Journal of Engineering Business Management*, **5**, 30.
- Gardner, E. S. (1985). Exponential smoothing: The state of the art, *Journal of forecasting*, **4**, 1–28.
- Gonzalez-Romera, E., Jaramillo-Moran, M. A., and Carmona-Fernandez, D. (2008). Monthly electric energy demand forecasting with neural networks and Fourier series, *Energy Conversion and Management*, **49**, 3135–3142.
- Huang, Y.-F., Chen, P.-J., and Nguyen, T.-L. (2014) Forecasting with Fourier residual modified ARIMA model- an empirical case of inbound tourism demand in New Zealand, *Recent Researches in Applied Economics and Management*, **2**, 61–65.
- Kim, B. and Kim, J. (2013). Time series models for daily exchange rate data, *The Korean Journal of Applied Statistics*, **26**, 1–14.
- Kimball, B. A. (1974). Smoothing data with Fourier transformations, *Agronomy journal*, **66**, 259–262.
- Lye, K. W., Yuan, M., and Cai, T. X. (2009). A spectrum comparison method for demand forecasting, *SIMTech technical reports*, **10**, 32–35.
- McLeod, A. I. and Li, W. K. (1983). Diagnostic checking ARMA time series models using squared residual autocorrelations, *Journal of Time Series Analysis*, **4**, 269–273.
- Tan, Z., Zhang, J., Wang, J., and Xu, J. (2010). Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models, *Applied Energy*, **87**, 3606–3610.
- Walker, J. S. (1991). *Fourier Series*, Oxford University Press, New York.

광, 공업용 건물의 전기 사용량에 대한 시계열 분석

김민아^a · 김재희^{a,1}

^a덕성여자대학교 정보통계학과

(2017년 12월 21일 접수, 2018년 1월 30일 수정, 2018년 3월 8일 채택)

요약

본 연구는 2014년 1월부터 2017년 4월까지 광, 공업용 제조업을 하는 건물(GGM)의 전기 사용량에 대한 예측을 살펴보고자 한다. SARIMA, SARIMA + GARCH, Holt-Winters 방법, Fourier 변환으로 분해를 한 ARIMA 모델을 중심으로 네 가지 모형에 대한 적합을 하였다. 또한 2017년 5월 사용량에 대한 예측하고, 실제값을 고려하여 각 모형에 대해 예측 제공된 평균 제공 오차와 예측 오차율을 비교하였다. GGM 건물의 전기 사용량에 대한 변동이 심하기 때문에 여러 가지 모형 중에서도 변동성과 주기를 함께 고려한 SARIMA + GARCH 모형의 적합과 예측이 가장 뛰어난 것을 확인하였다.

주요용어: 계절성, 시계열모형, 예측모형, 전기 사용량, 푸리에 변환, 홀트-윈터스, ARIMA, GARCH, SARIMA

이 논문은 산업통상자원부(MOTIE)와 한국에너지기술평가원(KETEP)의 지원을 받아 수행된 연구과제입니다 (No. 20161210200610). 또한 2017년도 미래창조과학부 재원으로 한국연구재단 기초연구 지원을 받아 수행되었습니다 (NRF-2016R1D1A1B03933527).

¹교신저자: (01369) 서울시 도봉구 삼양로 144길 33, 정보통계학과 덕성여자대학교.

E-mail: jaehee@duksung.ac.kr