J. Soc. Korea Ind. Syst. Eng Vol. 39, No. 1 : 25-30, March 2016
http://dx.doi.org/10.11627/jkise.2016.39.1.025
ISSN : 2005-0461(print)
ISSN : 2287-7975(online)

Short-term Forecasting of Power Demand based on AREA

S. H. Kwon* · H. S. Oh**†

*Department of Statistics, Hannam University
**Department of Industrial & Management Engineering, Hannam University

AREA 활용 전력수요 단기 예측

권세혁* · 오현승**†

*한남대학교 경상대학 비즈니스통계학과
**한남대학교 공과대학 산업경영공학과

It is critical to forecast the maximum daily and monthly demand for power with as little error as possible for our industry and national economy. In general, long-term forecasting of power demand has been studied from both the consumer's perspective and an econometrics model in the form of a generalized linear model with predictors. Time series techniques are used for short-term forecasting with no predictors as predictors must be predicted prior to forecasting response variables and containing estimation errors during this process is inevitable. In previous researches, seasonal exponential smoothing method, SARMA (Seasonal Auto Regressive Moving Average) with consideration to weekly pattern Neuron-Fuzzy model, SVR (Support Vector Regression) model with predictors explored through machine learning, and K-means clustering technique in the various approaches have been applied to short-term power supply forecasting. In this paper, SARMA and intervention model are fitted to forecast the maximum power load daily, weekly, and monthly by using the empirical data from 2011 through 2013. $ARMA(2, 1, 2)(1, 1, 1)_7$ and $ARMA(0, 1, 1)(1, 1, 0)_{12}$ are fitted respectively to the daily and monthly power demand, but the weekly power demand is not fitted by AREA because of unit root series. In our fitted intervention model, the factors of long holidays, summer and winter are significant in the form of indicator function. The SARMA with MAPE (Mean Absolute Percentage Error) of 2.45% and intervention model with MAPE of 2.44% are more efficient than the present seasonal exponential smoothing with MAPE of about 4%. Although the dynamic repression model with the predictors of humidity, temperature, and seasonal dummies was applied to foretaste the daily power demand, it lead to a high MAPE of 3.5% even though it has estimation error of predictors.

Keywords : SARMA, Intervention Model, Forecasting, MAPE, Power Demand, Seasonality

1. 서 론

전력수요에 대한 예측기간은 활용 목적별로 차이가 있다. 전력공급시설을 건설하는데 적어도 2년, 최대 10년까지 소요되므로 국가의 산업발전 및 경제성장 수준에 맞춘 중장기 수요 예측이 필요하고, 전력 예비율은 매일, 매주, 매월, 1년까지의 단기적 예측이 필요하다. 본 연구의 연구대상인 최대부하(전력수요 중 한 시간 평균 최댓값)의 예측절차는 국가법령에 의해 정하고 있고 월간, 주간, 일간 예측을 하도록 규정하고 있다[9].

전력 예비율은 전력 공급이 가장 높은 피크타임의 수요를 채우고 남는 상태를 비율로 나타낸 것으로 (총 공급

Received 30 December 2015; Finally Revised 4 February 2016; Accepted 5 February 2016
† Corresponding Author : hsoh@hnu.kr

전력량-최대 전력수요)/(최대 전력수요)으로 계산된다. 안정된 전력공급으로 판단하는 전력예비율의 기준 값은 10%로 그 값 이하로 내려가면 비상상황으로 간주한다[6].

전력수요에 대한 과소 예측은 일정 지역의 전기공급이 동시에 중단되는 블랙아웃을 가져오게 하여 국가는 산업 동력의 상실로 인한 막대한 경제적 손실은 물론 사회 구성원이 겪는 정신적 피해로 인한 혼란까지 감수해야 한다. 과다 예측은 전력 가격 상승을 유발하고 소비되고 남은 전력을 관리하는 비용이 증가하게 된다. 그러므로 정확한 전력수요 예측은 중장기적으로는 설비투자와 수급안정 결정, 단기적으로는 전력비용 책정에 영향을 미쳐 국가경제에서 매우 중요한 역할을 하게 된다.

전력연구원[6]은 선진국의 중장기 전력수요 예측방법을 최종소비자 방법과 계량경제학적 방법으로 나누어 기술하고 있다. 전자는 최종 소비자 들을 주거, 상업, 산업, 교통 등으로 분류하고 기술 변화에 따른 효율성의 증가와 사회경제적 요인의 변화를 고려하여 각 부문의 미래 에너지 소비를 예측한 후 전력수요가 차지하는 비중을 계산하여 전력수요를 예측하는 방법이다. 후자는 전력수요에 영향을 미치는 요인들을 외생변수로 하여 모형을 적합 시켜 전력수요를 예측한다. 이 방법들은 외생변수들의 미래 예측 값이 있어야 전력수요를 예측할 수 있어 계산된 오차보다 불확실성이 더 많이 내포되어 있어 단기예측에는 활용 빈도도 낮다.

단기예측에서는 예측변인을 고려할 필요가 없는 시계열 분석방법인 지수평활법이나 ARMA(Auto Regressive Moving Average) 모형을 활용하여 자신의 과거 값들의 패턴을 인식하여 미래 값을 예측하는 것이 필드 활용성도 높고 예측오차를 포함한 예측변인을 활용한 모형보다 적합하다. 단기 전력수요는 기상 요인들과 상관관계가 높으므로 이를 이용한 ARMA에 의한 전력수요 예측이나[3], 시간대별 전력수요에 대한 기온 민감도를 계산하여 ARMA를 적용한 예측방법이 적용되었다[13]. 전력수요가 주마다 상이한 패턴이 나타나므로 과거 전력수요를 입력데이터로 사용하여 주간 주기성과 패턴을 반영한 계절성 ARMA 모형과 회귀분석을 통해 계산된 기온과 전력수요사이의 변동 량을 예측에 반영한 지수평활 모형을 제안하였다[4, 10, 12].

최근에는 데이터마이닝 기법이나 빅데이터 기법을 이용한 전력수요 예측연구가 진행되고 있다. 박영진[11] 등은 뉴로-퍼지 모델을 이용하여 전력수요를 시간, 일간, 주간 단위로 예측하였고, 이형도[8] 등은 전력수요에 영향을 미치는 방대한 예측변인들 중 기계학습법에 의해 최소한의 예측변인을 탐색・도출하여 SVR(Support Vector Regression) 방법으로 전력수요를 예측하였다. 최상률[2] 등은 의사결정나무기법으로 일간, 특수요일별 전력수요량의 패턴을 분석하여 단기 전력수요량을 예측하였다. 김철홍[5] 등은 매일의 전력 최대부하를 K-평균 비계층적 군집분석에 의해 4개 계절로 군집화 하고 K-최단 이웃거리(Nearest Neighbor) 방법으로 요일을 군집화 하여 일간 최대 부하 값을 예측하였다.

기존 연구에서 제안된 다양한 예측모형들이 있음에도 활용의 간편성과 다른 예측변인 없이 이전 최대 전력부하 값으로 예측 가능한 부하 값을 현장에서는 주로 사용한다. 하여, 본 연구에서는 지수평활법보다 예측력을 높일 수 있는 SARMA 모형을 적합하되 기존 SARMA 연구들에서 반영하지 않았던 시계열의 지시함수 형태 변화인 여름휴가, 겨울, 긴 휴일기간을 예측변인으로 한 개입모형을 적합 시켰다. 예측을 위하여 예측변인으로 사용한 기상변인에 대한 예측 값이 필요한 기존 SARMA연구와는 달리 본 연구에서 제안된 개입변인(예측변인)들은 측정오차가 없으므로 제시된 MAPE가 최종오차이다.

본 연구의 실증분석은 2011년~2013년 3개년 일별 최대 전력부하 데이터를 이용하여 일별, 주별, 월별 예측모형을 적합 시켜 비교하였다. 제 2장에서는 본 연구에서 사용된 연구방법론을 기술하였고, 실증분석에 활용되는 데이터 기술과 최종 적합 된 전력수요 예측모형은 제 3장에 정리하였다. 제 4장에는 연구결과에 대한 논의와 향후 연구방향을 기술하였다.

2. 연구방법

본 연구에서는 최대 부하전력에 대한 예측모형 도출을 위하여 이전 최대 부하전력 시계열 패턴을 도형화 하는 ARMA 기법과 시계열 패턴의 이상점 변화를 결정적(deterministic) 예측변인으로 반영할 수 있는 개입모형을 고려하였다. 개입 ARMA 모형 적합은 시계열 데이터에 최적 ARMA 모형 적합 시키고, 그 모형에 개입변인을 적용하여 최종 예측모형을 도출하게 된다.

t를 시간, t기 관측 값을 Y_t , 관측된 데이터 총 개수를 T라 하면 $\{Y_t, t=1, 2, \cdots, T\}$은 시계열 데이터 표현이다. Z_t를 Y_t의 표준화 관측값, d를 차분차수 d, $(p,\ q)$를 각각 AR, MA의 차수라 하면, $ARMA(p, d, q)$ 모형은 다음과 같다. B는 후진연산자(Back shift operator)로 $B(Z_t) = Z_{t-1}$이다[1].

$$\phi(B)(1-B)^d Z_t = \theta_q(B)e_t \tag{1}$$

i) $\{Y_t\}$가 백색잡음인지 검정하여 백색잡음이 기각되어야 ARMA 모형 적합이 가능하다.

ii) 정상적(stationary) 시계열 데이터인지 판단하여 비정상적인 시계열의 경우에는 정상적 시계열 데이터로 변환한다. 분산이 일정하지 않는 비정상성은 분산 안정

화 방법(로그 변환을 가장 많이 사용)을 적용하고 랜덤 워크 모형(AR(1)이고 $Y_{(t-1)}$의 계수가 1)은 1차 차분 $\{\nabla Y_t\}$에 모형을 적용한다. 분산 안정화 방법 적용이 불가능한 이분산시계열모형 GARCH(Generalized Auto Regressive Conditional GARTH) 모형을 적용한다.

iii) ACF(Auto Correlation Function)과 PACF(Partial Auto Correlation Function)을 이용하여 차수 (p, d, q)을 결정한다. 이 과정을 모형 식별이라고 한다. 일반적으로 p, q 차수는 2 이하를 권장한다.

iv) 결정된 차수에 의해 모수를 추정하고 유의성을 검정한다. 순서 (iii)에서 결정한 (p, d, q)의 회귀계수가 모두 유의해야 한다. 모수 추정은 최대우도추정법과 조건부 최소제곱추정법을 주로 사용한다.

v) 추정 적합모형에 의해 절차를 구하고 절차가 백색잡음인지를 검정한다. 모형 적합이 완전하면 잔차는 시계열 데이터의 어떠한 패턴도 남아 있으면 아니 되므로 백색잡음이어야 한다. 만약 잔차가 백색잡음을 따르지 않으면 모형 적합이 불완전함을 의미한다.

vi) 본 연구에서는 예측모형의 적합성 비교를 위하여 AIC (Akin Information Criterion)와 MAPE(평균 절대 퍼센트 편차)를 활용하였다. 예측모형의 최대우도함수 값을 L, 예측변수 수를 k, t기 예측 값을 $\widehat{Y_t}$라 하면 AIC와 MAPE는 아래와 같이 정의된다. 각 통계량 값이 작을수록 모형의 적합 도는 높다.

$$AIC = 2k - 2\ln(L) \tag{2}$$

$$MAPE = \frac{1}{T}\sum_{t=1}^{T}|(Y_t - \widehat{Y_t})/Y_t| \times 100(\%) \tag{3}$$

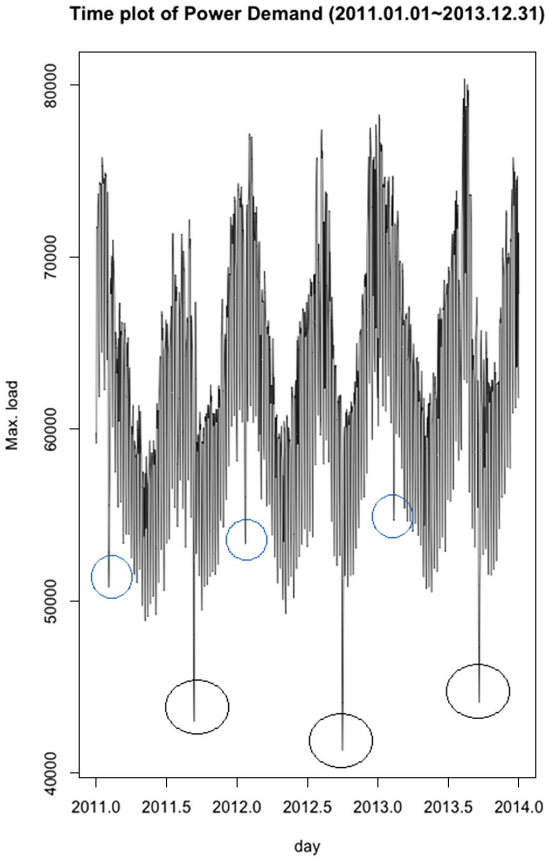
본 연구의 시계열 데이터, 일별 전력 최대부하 데이터는 주기 7의 계절성, 월별 전력 사용량 데이터는 주기 12의 계절성을 가지고 있을 가능성이 높으므로 계절성 ARMA 모형을 고려해야 한다. 계절성 주기가 s인 $ARMA(p, d, q)(P, D, Q)_s$모형은 다음과 같고 모형 적합 절차는 위와 동일하다.

$$\begin{aligned}&\phi(B)\Phi(B^s)(1-B^s)^D(1-B)^d Z_t \\ &= \theta_q(B)\Theta(B^s)e_t\end{aligned} \tag{4}$$

예측에 사용되는 데이터 기간이 길어 시계열 패턴의 변화가 생길 가능성이 높은 경우 개입(intervention) SARMA 모형을 고려하게 된다. 개입변수는 일정 시점에 한 번 영향을 미치는 지시변수 형태와 일정 시점 이후 지속적으로 영향을 미치는 계단함수 형태가 있다. 본 연구에서는 일별 전력수요량 데이터에서 계절 영향, 추석과 구정의 긴 연휴에 지시변수 형태의 개입이 발생할 수 있어 적합시켜 SARMA에 개입변인을 고려하여 모형 적합도를 높였다.

3. 실증분석

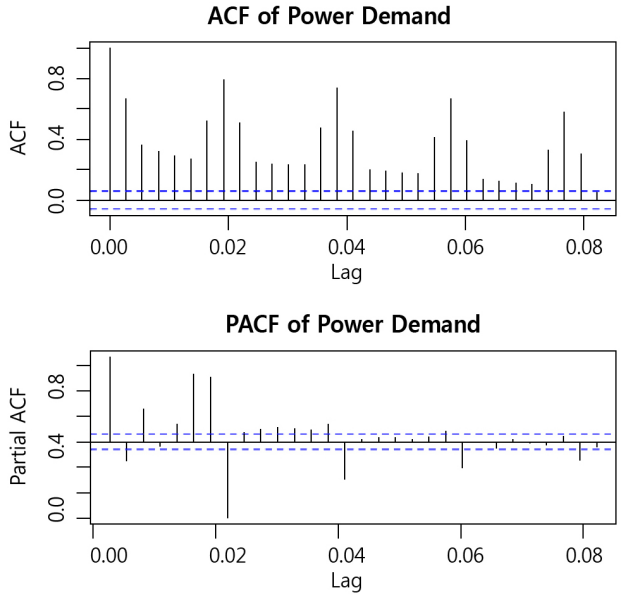
분석 대상인 일별 전력수요량은 백색잡음이 아닌 시계열 데이터이고(유의확률 < 0.001), 단일근 검정 결과 안정적이므로(유의확률 0.01) ARMA 모형 적합이 가능하다. <Figure 1>은 2011년 1월 1일부터 2013년 12월 31일까지 일별 전력 최대부하에 대한 시간도표이다.



<Figure 1> Time Plot of Power Demand

시간도표를 보면 큰 동그라미 부분은 극단적으로 낮은 전력수요를 보이는데 이는 추석연휴, 작은 동그라미 부분은 구정연휴로 긴 휴일 기간에 지시변수 형태의 개입모형을 고려할 필요가 있다. 그리고 시계열 데이터가 직선적으로 증가하는 경향을 보이고 있어 차분차수 $d=1$를 고려하는 것이 적절해 보인다. 일별 전력수요는 특성 상 요일별 계절성을 가지고 있어 계절주기 $s=7$을 고려할 필요가 있다.

모형인식을 위한 자기상관함수(ACF)과 부분자기상관함수(PACF)는 <Figure 2>에 제시하였다. ACF는 일정한 주기(s = 7)를 두고 피크를 이루고 피크 구간 내에서는 직선적으로 감소하고 있다. ACF의 직선적 감소는 시계열 데이터에 직선적 증가 경향이 있음을 진단하고 피크의 주기는 계절성 주기가 된다. PACF도 주기 7에서 피크가 반복되고 있어 ACF가 진단한 결과와 동일하게 주기 7인 계절성이 있음을 보여주고 있다. ACF와 PACF 모두 일정 시점 이후 값이 0이 되는 형태를 지니고 있지 않아 (p, q) 차수 모두 1 이상 고려하여 모형을 적합 시키는 것이 적절하다.



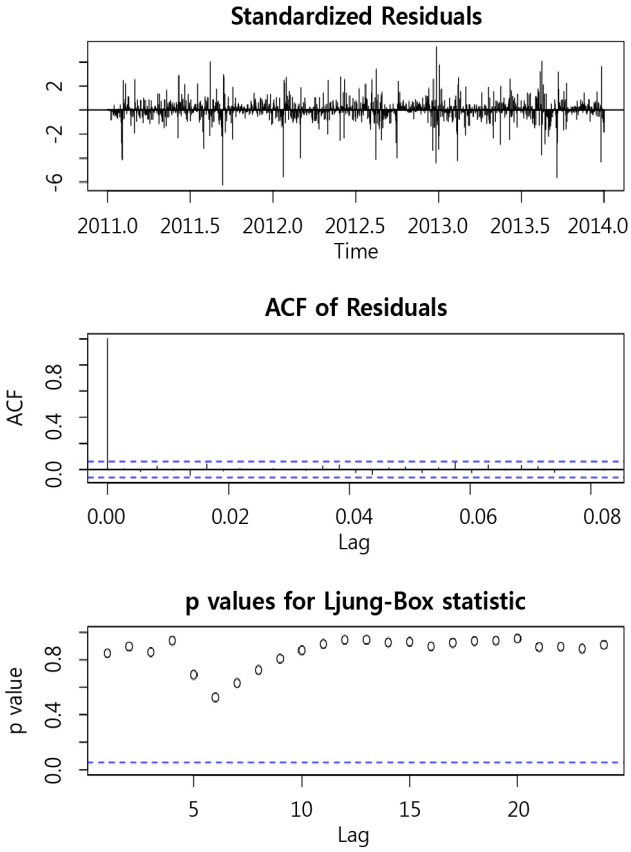
<Figure 2> ACF & PACF of Daily Power Demand

PACF, ACF 분석 결과 주기 7의 계절성, 선형적 증가를 조정하기 위한 1차 차분 후에 최대 차수(p = 2, q = 2)로 하여 추정회귀계수가 유의하고 절차가 백색잡음을 따르며 AIC 값이 가장 낮은 모형을 최종적합모형으로 도출하였다. <Table 1>은 최종적합모형 $ARMA(2,1,2)(1,1,1)_7$의 추정 결과이며(AIC = 19,996), 잔차에 대한 백색잡음 검정 결과(최하단 그래프-잔차 자기상관 유의성 검정 결과 유의확률이 유의수준 위에 있어 백색잡음)은 <Figure 3>에 제시하였다.

<Table 1> Results of Estimation

Variable	AR1	AR2	SAR1	MA1	MA2	SMA1
Estimate	1.606	-0.705	0.193	1.762	-0.796	0.974
Standard Error	0.065	0.046	0.035	0.069	0.065	0.011
t-statistic	24.8	15.2	5.56	25.3	12.3	92.4

차기 최대 전력부하에는 현재 전력부하(AR1)와 일주일 전 전력부하(SAR), 현재 전력부하에 의해 설명되지 못한 부분(MA1), 동일하게 일주일 전 영향인 SMA1이 양의 영향을 미치고 있다. 이 중 가장 영향을 많이 미치는 변인이 일주일 전 최대전력부하이고 그 다음이 MA1이다. 일별 특성을 갖는 차기 최대 부하전력에 일주일 전 최대 전력부하의 영향이 가장 큰 것은 당연하지만 다음 영향력 높은 변인이 현재 최대 부하전력에 의해 설명되지 못하는 변동의 영향이 크다. 그러므로 예측력을 높이기 위하여 새로운 예측변인을 고려할 필요가 있어 개입모형을 적합하였다.



<Figure 3> Residual Plot of $ARMA(2,1,2)(1,1,1)_7$

<Table 1>의 과거 최대 부하전력이 설명하지 못하는 변인, SMA1의 영향력을 줄이기 위한 새로운 예측변인을 찾기 위하여 시간도표를 재분석하였다. <Figure 1>의 동그라미로 볼 수 있듯이 긴 연휴(long weekend)에는 전력수요가 낮아진다. 그리고 여름휴가(summer, 7월 15일~8월 15일), 겨울(winter) 등도 전력수요량에 영향을 미칠 것으로 판단하여 이를 예측변인으로 사용하는 개입 ARMA 모형을 적합하여 <Table 2>에 정리하였다(AIC = 19,989).

<Table 2> Intervention Model

Variable	Estimate	SUE	t Value	p-value
AR1	1.597	0.068	23.61	< .0001
AR2	-0.696	0.048	14.49	< .0001
SAR1	0.195	0.035	5.64	< .0001
MA1	1.760	0.073	24.07	< .0001
MA2	-0.791	0.068	11.6	< .0001
SMA1	0.974	0.011	92.43	< .0001
Summer Holidays	-35.96	18.82	-1.91	0.056
Long Holidays	112.4	57.04	1.97	0.049
Winter	1710.7	833.54	2.05	0.040

개입변인의 삽입으로 적합모형의 예측력은 높아졌고(AIC가 낮아짐) SMA1의 영향력도 전기 최대 부하전력보다 낮아져 기존 최대 부하전력 시계열 데이터의 설명력이 높아졌음을 알 수 있다. 긴 공휴일은 산업용 전력수요가 적어 전력수요가 낮고 여름휴가와 겨울에는 냉난방기 사용으로 인하여 전력수요가 높아졌음을 알 수 있다. 개입변인 중 겨울변인이 최대 부하전력에 개입하는 영향력이 가장 큼을 알 수 있다.

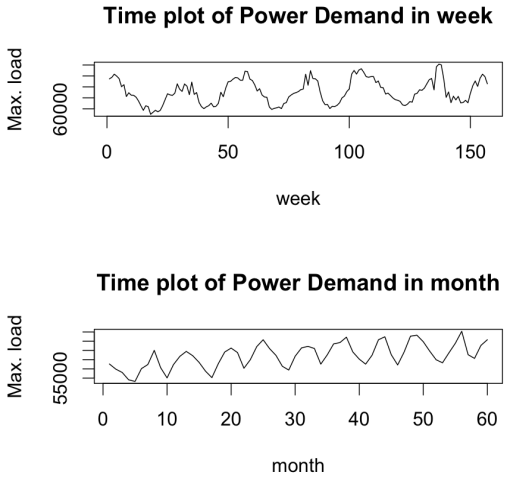
주별, 월별 전력 최대부하 값은 각 기간의 최댓값으로 데이터를 사용하였다. <Figure 4>는 주별, 월별 전력수요량의 시간도표이다. 주별 전력수요량은 단일근 모형으로 ARMA 모형 추정이 불가능하였다. 월별 전력수요량에 모형 인식을 위하여 필요한 ACF, PACF 그림은 <Figure 5>에 정리하였다.

ACF 함수의 경우 주기 12에서 피크가 발생하고 있고 월별 자료이므로 계절성 주기 12를 설정하여 도출한 최종 적합모형은 $ARMA(0,1,1)(1,1,0)_{12}$이고 추정결과는 <Table 3>에 정리하였다. 적합 모형의 타당성을 보여주는 절차의 백색잡음 검정결과는 <Figure 6>에 제시하였다. 추정계수가 유의하고 절차가 백색잡음을 따르므로 최종 적합 모형은 유의하였다(AIC = 872).

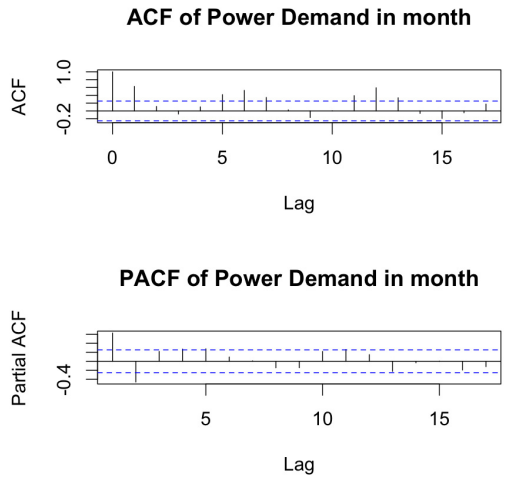
<Table 3>에서 볼 수 있듯이 월 최대 부하전력은 일 년 전 전기 월 최대 부하전력에 의해 양의 영향을 받고, 월별 최대 부하전력의 기존 값들이 설명하지 못하는 변인(MA1)에 위해 음의 영향을 받는다. MA1 영향력이 SAR1보다 크고 매우 유의하므로 예측력을 높일 수 있는 새로운 예측변인들에 연구가 향후 필요하다.

<Table 3> Results of Estimation

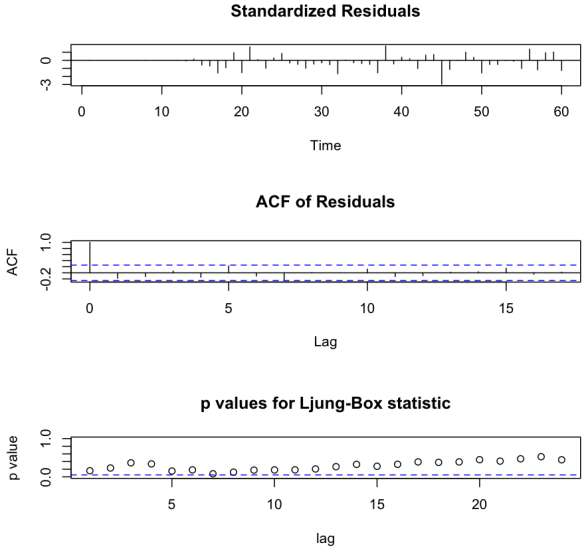
Variable	MA1	SAR1
Estimate	-0.804	0.274
Standard Error	0.074	0.164
t-statistic	-10.824	1.664



<Figure 4> Time Plot of Monthly Power Demand



<Figure 5> ACF and PACF of Monthly Power Demand



<Figure 6> Residual Plot of $ARMA(0,1,1)(1,1,0)_{12}$

4. 결 론

2011년 1월 1일부터 2013년 12월 31일, 3년 전력 최대 부하 데이터를 활용하여 일별, 주별, 월별 전력수요량을 ARMA 시계열방법으로 예측하였다. 일별 최대 부하전력은 $ARMA(2, 1, 2)(1, 1, 1)_7$, 월별 최대전력은 $ARMA(0, 1, 1)(1, 1, 0)_{12}$ 모형이 가장 적합하였으나 주별 최대 부하전력은 단일근 모형이고 1차 차분 결과 백색잡음이여서 ARMA 모형 적합을 할 수 없었다. 일별 최대 부하전력의 SARMA 모형 적합결과 데이터의 이상적 변화가 발생하는 시기 예측력이 낮아 여름휴가, 겨울, 긴 연휴가 기간을 지시함수 형태의 개입변수로 하여 개입 SARMA 모형을 적합한 결과 적합 도를 높일 수 있었다.

현재 현장에서 일별 최대부하 예측에 주로 사용되는 모형인 계절 지수평활법의 MAPE는 4%인데 비해 본 연구에서 제안된 SARMA는 2.45%, SARMA 개입모형은 2.44%로 기존 모형에 비해 예측력이 높았다. 그리고 본 연구에서 제안된 모형의 현장 활용 편이성은 지수평활법과 동일하며, 개입변인들도 측정 오차가 없는 확정적 변인이므로 현장 적용이 가능하다. 현재 MAPE가 4% 수준에서 예비전력을 10% 보유하고 있는데, 본 연구 결과를 활용하면 예비전력을 7%대 수준으로 낮출 수 있을 것으로 판단된다.

예측 가능한 전력수요량의 예측변인들, 기온, 습도, 월, 계절, 분기 등을 고려한 동적 회귀모형을 일별 예측에 적용한 결과 MAPE는 3.5%였다. 실제 예측변인에 내포된 오차까지 고려하면 적합모형으로 활용할 수 없다.

Acknowledgement

This study has been partially supported by the 2015 University Research Fund of Hannam University, Daejon, Korea.

References

[1] Box, E.P., Jenkins, G.M., Repines, G.C., and Lung, G.M., Time Series Analysis : Forecasting and Control(5th ed.), Wiley, 2015.

[2] Chaos, S.Y. and Kim, H.J., Short-term demand forecasting Using Data Mining Method, *Journal of the Korean Institute of Illuminating and Electrical Installation Engineers*, 2007, Vol. 21, No. 10, pp. 126-133.

[3] Jung, H.W. and Song, K.B., Daily Maximum Electric Load Forecasting for the Next 4 Weeks for Power System Maintenance and Operation, *The Transactions of The Korean Institute of Electrical Engineers*, 2014, Vol. 63, No. 11, pp. 1497-1502.

[4] Kim, C.H., Forecasting of Domestic Power Demand using Multiple Seasonal Exponential Smoothing Techniques, The annual report of Korea Enright Economics Institute, 2013.

[5] Kim, C.H., Koa, B.G., and Park, J.H., Short-term Electric Load Forecasting Using Data Mining Technique, *Journal of Electrical Engineering Technology*, 2012, Vol. 7, No. 6, pp. 807-813.

[6] Korea Power Exchange, A Study on the criteria of the electricity demand forecast evaluation and the confidence interval, annual report of 2011.

[7] Kwon, S.H. and Oh, H.S., Forecasting Model for Flood Risk at Bo Region, *Journal of Society of Korea Industrial and Systems Engineering*, 2014, Vol. 37, No. 1, pp. 91-95.

[8] Lee, H.R., Park, K.H., and Shin, H.J., Electricity Demand Forecasting based on Machine Learning Algorithms, *Korea Academic Association of Business Administration, Proceedings of 2011.*

[9] Ministry of Governmen Legislation, Procedure of forecasting power demand, National Law, 2011.

[10] Oh, H.S. and Moon, G.J., A comparison of technological growth models, *Journal of the Korea Society for Quality Management*, 1994, Vol. 22, No. 2, pp. 51-68.

[11] Park, Y.J. and Wang, B.H., Neuron-Fuzzy Model based Electrical Load Forecasting System : Hourly, Daily, and Weekly Forecasting, *Urge and Intelligence System Institute*, 2004, Vol, 14, No. 5, pp. 533-538.

[12] Shin, S.C., Oh, H.S., and Choi, J.H., Price Forecasting of Natural Resources with Restricted Market, *Journal of Society of Korea Industrial and Systems Engineering*, 2014, Vol. 37, No. 4, pp. 187-191.

[13] Song, K.B., Development of Short-Term Load Forecasting Algorithm Using Hourly Temperature, *The Transactions of the Korean Institute of Electrical Engineers*, 2011, Vol. 63, No. 4, pp. 451-454.

ORCID

S. H. Kwon | http://orcid.org/0000-0001-6195-9141

H. S. Oh | http://orcid.org/0000-0002-7773-3750