

V2G 의사결정을 지원하는 전력사용량 데이터의 시계열 예측[†]

이정현¹ · 오상진² · 윤여창³ · 안영호⁴ · 김재성⁵ · 조완섭⁶ · 이성덕⁷

¹⁵⁶충북대학교 빅데이터협동과정 · ²⁷충북대학교 수학교육통계학부 · ³우석대학교 · ⁴레티그리드

접수 2018년 12월 31일, 수정 2019년 1월 21일, 게재확정 2019년 1월 22일

요약

최근 전력 수요가 급증하는 상황에서 많은 에너지 관련 연구들이 대체에너지 개발, 전력사용의 절감과 효율, 전력 사용의 시간차 분산 그리고 에너지 저장기술 등을 다루고 있다. 본 연구에서는 한 생산시설에서 수집된 전력사용량 시계열 데이터를 분석하면서, Peak 전력사용량과 그 시간대를 예측에 대한 중요성을 인지하였다. 이는 상시전력과 대체전력간의 전환, 에너지 거래 등에 있어서 필수적인 예측분석이기 때문에 더욱 실효성을 높일 수 있는 결과 도출을 위해 24시간의 단기 예측을 분석 목표로 설정하였다. 본 논문에서 수행한 시계열 분석방법으로 SARIMA모형과 계절성지수평활법이 사용되었고 이러한 분석모델이 향후 V2G 또는 ESS 관련된 현업에서 활용과 확장이 될 수 있도록 기대한다.

주요용어: 계절성 ARIMA, 계절성 지수평활법, 전력 Peak 부하, ESS, V2G.

1. 머리말

최근 제기되는 원전이슈와 함께 현재의 국내 전력수급은 위기에 맞닿아 있고 향후에도 한반도 정치적 문제와 국내 전기차 활성화 정책에 따라 추가적인 전력 수요가 급증할 것이라 예상된다. 이에 따라 에너지 절감과 효율을 높이기 위한 많은 연구가 진행되었고, Yoon (2017)의 연구에서는 건물에너지 사용량과 상관관계를 가지는 5가지 인자 (건물용도, 재실율, 준공연도, 방향, 단열형태)를 분류하여 에너지 소비와 영향이 큰 인자를 규명하였다. Kim (2014)의 연구에서는 다양한 에너지 절감요소 (자연적 환경요소와 신/재생에너지 기술)의 조합과 요소별 에너지 절감 성능평가 (자연채광, 자연환기, 단열성능, 일사열 취득계수, 기밀성능, 지열에너지 시스템)를 통해서 건축물에너지 효율등급의 단계적인 향상 방안도 제시했다. 그 외 에너지 절감 및 효율을 높이기 위한 연구분석에서 다양한 외부 환경요인과 내부 환경요인이 함께 고려되었다. 또한 공공 건물이나 사무용 빌딩과는 달리 대형 생산설비가 운용되는 시설이나 고속철도 변전소 등의 경우, 외부 환경요인이 큰 비중으로 고려되기 보다는 주요한 내부 환경요인 또는

[†] 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 대학ICT연구센터지원사업의 연구결과로 수행되었음 (IITP-2018-0-01396).

¹ (28644) 충북 청주시 서원구 충대로 1, 충북대학교 빅데이터협동과정, 박사.

² (28644) 충북 청주시 서원구 충대로 1, 충북대학교 수학교육통계학부, 학사.

³ (55338) 전북 완주군 삼례읍 삼례로 443, 우석대학교, 박사.

⁴ (05621) 서울특별시 송파구 백제고분로41길 8 송파빌딩, 레티그리드, 대표.

⁵ (28644) 충북 청주시 서원구 충대로 1, 충북대학교 빅데이터협동과정, 박사.

⁶ (28644) 충북 청주시 서원구 충대로 1, 충북대학교 빅데이터협동과정, 교수.

⁷ 교신저자: (28644) 충북 청주시 서원구 충대로 1 충북대학교 수학교육통계학부, 교수.

E-mail: sdlee@chungbuk.ac.kr

내부에 존재하는 실질적인 에너지 수요 속성이 대부분의 전력 사용량을 차지하기도 한다. Ministry of Trade, Industry and Energy (2017)의 에너지 총조사 보고서에 따르면, 에너지 수요부문 (산업부문, 수송부문, 상업 및 공공부문, 가정, 대형건물, 신재생에너지)별로 에너지 효율을 평가할 수 있는 지표도 개발되었다.

한편 최근 공공기관을 비롯한 여러 시설에서 ESS (energy storage system)을 확대 적용하고자 하는 연구와 시도가 많았다. 한 예로, Kim (2014)의 연구에서는 고속철도 변전소 피크부하를 저감시키기 위한 ESS의 적정용량과 경제성을 분석하고 평가하였다. ESS를 큰 용량으로 도입하면 Peak 전력에 대응할 수 있는 범위가 확장될 수 있지만, 그만큼 구축비용이 크기 때문에 향후에는 적절한 용량의 ESS에 대해서 V2G 기술이 접목되어 상시로 전력수급이 이루어지면서 ESS가 활용될 수 있는 대역을 확대할 수 있다.

Son (2018)의 연구에서는 배전망 단위의 소규모 전기에너지 Pool을 통해 에너지 거래가 가능하도록 LDERP (local distributed energy resource provider) 중심의 거래절차를 구성하고, 시장에 참여하는 프로슈머를 ‘판매자’로 설정하고 전력을 소비하는 측을 ‘구매자’로 설정하여 에너지 거래 모델의 운영 구조 및 절차를 연구하였다. 따라서 이러한 전력 거래가 이루어지기 위해서는 수요자의 전력 사용량에 대한 모니터링과 분석기법을 통해서 현재 시점에서 향후 24시간에 대한 전력 사용량 예측이 사전에 이루어지는 것이 중요하고, Peak 전력사용량과 그 시간대의 예측 정확도가 높아야 한다. 또한 이러한 예측의 결과가 다양한 관련 현업에서 활용가능한 차세대 시스템에 충분히 반영될 수 있는 수준이 되어야 한다.

Choi (2015)의 연구에 따르면 에너지 기술과 ICT기술의 융합은 기존에 정제되어 있던 에너지 분야에서의 기술 혁신을 이루어내기 위한 원동력이 된다. 에너지 분야에 있어서 정보통신기술의 사용은 서로 다른 건물 및 장비간의 에너지 거래 및 공유를 가능하게 하며, 에너지 생산 및 소비 단계에서의 공급 및 수요예측을 통한 효과적인 에너지 사용을 가능하게 한다. 또한 통신기술을 통해 수집한 정보를 분석하여 기존에 예측하지 못했던 다양한 부분에서의 에너지 절감을 목표로 하는 것이 가능하다. 그러나 이러한 에너지 ICT융합의 일환에서 현업에서의 차세대 서비스에 실질적인 도움이 되기 위해서는 기존 시계열 분석이나 인공지능경망 등을 통해 단순한 예측연구 결과를 도출하는 것 보다는 비즈니스 모델이 기반된 연구가 필요하다.

본 논문에서는 이러한 예측을 위해서 시계열 분석을 수행하였고, 분석결과를 통해 V2G 환경에서 전력 수요자 측면의 전력 수급계획에 대한 의사결정을 효과적으로 지원하고자 한다. 본 연구의 구성은 2장에서 ESS와 V2G에 대한 기술을 알아보고 ESS를 활용한 전력데이터 분석 사례를 소개한다. 이러한 ESS와 V2G의 조합은 에너지 저장 Pool의 개념으로 다양한 연구에서 활용되고 있다. 3장에서는 본 연구에서 사용할 계절성 ARIMA모형 (SARIMA)과 계절지수평활법에 대한 모형이론을 확인하고, 실험데이터에 대한 설명을 기술하였다. 4장에서는 데이터를 활용한 예측분석을 위해서 다양한 연구방법 중에서 통계학적 예측방법을 통한 분석의 과정을 밝혔다. 5장에서는 총 87일간의 시간별 데이터 (1시간 수집주기, 총 관찰치 수: 1463개) 중 86일간의 데이터를 활용하여 시계열 모델을 수립하고, SARIMA와 계절지수평활법에 따른 예측기법을 통해서 87일째 되는 날의 24시간에 대한 실제값과 예측된 값을 비교하여 6장에서 결론을 도출하였다.

2. 선행 연구

2.1. ESS

신재생에너지와 스마트그리드가 연구되면서 ESS (energy storage system)이 많은 각광을 받았다. ESS는 전력에너지를 미리 저장했다가 필요한 시간대에 사용할 수 있는 에너지 저장 시스템이다. 국내

에서는 2011년 지식경제부에서 발표한 ‘에너지 저장 기술 개발 및 산업화 전략 K-ESS 2020’에서 세계 시장 진출을 위해서 6.4조원 규모의 연구 개발 및 설비 투자를 추진하였다.

Kim (2014)의 연구에서는 고속철도 변전소의 Peak 전력사용량 분석 및 예측을 통해 고가의 ESS의 최적 용량을 산정하고 그 경제성을 분석하였다. 고속철도 운행에서 발생하는 전력 사용량 15분 데이터를 분석하여 정격 15MW의 대용량 부하 환경에서 Peak 전력에 대응하기 위해 ESS의 용량 (6MW의 정격출력과 3.5MWh의 전지용량, 2011 ~ 2012년 기준, 특정 변전소 기준)을 산정하였고, 구축비용 (약53억원)에 대한 회수기간 (8.7년)을 평가하였다. Byeon (2016)의 연구에서 ESS의 효율적인 가동을 위해서 과거데이터 기반으로 Peak 전력을 저감시킬 수 있는 최적운전기법이 적용된 프로그램을 개발하였다.

2.2. V2G

Ansari (2015)의 연구에서와 같이 V2G (vehicle to grid)는 전기자동차에 기반한 서비스중 하나로서 전기요금에 싸고 공급량이 충분한 시간대에 전기자동차를 충전하고 피크시간대에 다시 전력망으로 되팔게 함으로써 전체적인 발전요구량을 낮춘다. 또한 Kim (2017)의 연구에서 V2G는 전기차 (EV), 플러그인 하이브리드 자동차 (PHEV) 등 충전식 차량을 충전했다가 주행 후 남은 전기를 다시 전력망으로 송전하게 해주는 기술이며, 전기차가 움직이는 에너지저장장치 (ESS)의 역할을 제안하였다. 따라서 이러한 V2G 기술은 풍력/태양열 발전소처럼 생산능력이 일정치 않거나, 전력 수요가 큰 변전소 등에 우선 적용될 수 있다. 이와 관련하여 Hwang (2014)의 연구에서는 전력생산이 일정하지 않고 기상에 따라 유동적인 풍력/태양열 발전소에 V2G를 적용하여 부하에 공급되는 전력의 안정성을 높이기 위해 V2G 운영기법이 제안되었다. 이러한 배경에서 Kim (2016)의 연구처럼 피크 시간대에 V2G 전기자동차의 배터리에 충전했던 전력을 계통에 공급하는 방법은 최대수요를 낮추어 높은 단가의 발전소 건설비용과 연료비용을 회피하는 효과를 가져올 수 있다. 현재 관련 연구기술의 방향으로서는 전기자동차가 전력계통에 직접 송전하는 것은 안전관리 측면의 이슈가 존재하기 때문에, ESS에 여러 대의 전기자동차가 연결된 형태의 V2G 기술이 고려될 수 있다.

2.3. 전력데이터를 활용한 분석 연구

전력 사용량 데이터를 활용하여 다양한 예측 관련된 연구가 진행되었다. Park (2017)의 연구에서는 전력 수요 패턴인 요일, 휴일, 날씨, 사회적 이벤트 등에 대해 선행연구를 진행하면서 국내 전력거래소에서 수집된 5년간의 시간 단위 데이터를 활용하여 4년간 데이터를 training data로 사용하고 나머지 1년간 데이터를 validate data로 사용하여 분석 결과를 비교 및 평가하였다. 데이터 셋에서의 각 시간대를 독립변수로 간주하였고, 24개의 독립변수를 가지고 여러 가지 군집분석을 수행하였다. 군집분석에 사용된 기법은 의사결정나무와 랜덤포레스트, 나이브베이지안, SVM이고, 각 분석기법에 대해 예측 정확도를 산출했다.

Moon (2017)의 연구에서는 대학 캠퍼스 중 일부 건물을 대상으로 총 4년간의 일간 전력 사용량 데이터를 통해서 의사결정나무를 이용하여 데이터를 분류하였고, 테스트 셋에서 유사 시계열 형태의 데이터 셋을 가지고 ARIMA 기반의 예측모델을 구성하였다. 학사 기간의 특성에 따라 전력 사용량의 여러 패턴을 보였고, 해당 기간이나 주말에 따라서 각기 다른 9개의 ARIMA 모형을 제시하였다.

Ryu (2017)의 연구에서는 미국 연구소 내 지원시설을 대상으로 분석을 수행하였으며 총 91일간의 일간 전력 데이터 중 86일 간의 데이터를 통해 시계열 모형을 수립하고 선정된 모형의 적절성을 검토한 뒤, 나머지 5일간의 실제 데이터와 새로운 5일간의 예측값을 비교하여 성능을 평가하였다.

3. 시계열 모형 이론과 실험데이터

3.1. Seasonal ARIMA 모형의 분석 방법 이론

본 연구에서는 시계열 자료가 계절성 (seasonality)이나 주기적 특성 (periodicity)이 있을 때 활용할 수 있는 시계열 분석 (time series analysis) 기법중 하나인 seasonal autoregressive integrated moving average (SARIMA)을 활용하였다. Lee (2017)의 연구에 따라, ARIMA 모형은 과거의 관찰된 값들로 설명되는 자기회귀 모형 (autoregressive model)과 시계열을 오차항으로 설명하는 모형으로 이동평균 모형 (moving average model)을 혼합하고 차분 과정을 거친 모형이다. 이러한 ARIMA 모형에 주기적인 특성이나 계절성을 반영한 모형이 SARIMA 모형 이다. SARIMA 모형은 ARIMA 모형처럼 예측하고자 하는 시점으로부터 가까운 과거 자료들을 이용할 뿐만 아니라, 자료의 주기적인 특성, 계절성을 반영했기 때문에 이전 주기의 자료를 추가적으로 활용한다. SARIMA(p,d,q)×(P,D,Q)_s 모형의 일반식은 식(3.1)과 같다.

$$\begin{aligned}
 \phi_p(B)\Phi_P(B)W_t &= \theta_q(B)\Theta_Q(B)e_t, \\
 \phi_p(B) &= 1 - \phi_1 B - \dots - \phi_p B^p, \\
 \Phi_P(B) &= 1 - \Phi_1 B^s - \dots - \Phi_P B^{Ps}, \\
 \theta_q(B) &= 1 - \theta_1 B - \dots - \theta_q B^q, \\
 \Theta_Q(B) &= 1 - \Theta_1 B^s - \dots - \Theta_Q B^{Qs}, \\
 W_t &= \nabla^d \nabla_s^D Z_t,
 \end{aligned} \tag{3.1}$$

여기서 Z_t 는 원시계열 자료, W_t 는 차분된 시계열 자료, t 는 시간을 나타내는 연산자, e_t 는 $N(0, \sigma^2)$ 을 따르는 오차항 (백색잡음), B 는 후향연산자, p 는 AR항의 차수, q 는 MA항의 차수, d 는 차분의 차수, P 는 계절 AR항의 차수, Q 는 계절 MA항의 차수, D 는 계절차분의 차수, s 는 계절시차를 나타낸다.

3.2. 계절지수평활법에 의한 시계열 분석

평활법은 관측된 과거 및 현재의 시계열 자료에 대해서 이동평균 (moving average), 가중이동평균 (weighted moving average) 등의 방법을 사용하여 불규칙변동을 평활하여 시계열 값을 예측하는 분석방법이다. 이러한 평활법에는 이동평균평활법과 가중평균평활법인 지수평활법 등이 있으며, 시계열의 유형에 따라 다른 평활법이 적용될 수 있고 계절변동성에 대해서는 계절지수평활법을 사용할 수 있다. 또한 계절변동은 가법적 계절변동과 승법적 계절변동의 두 형태가 있기 때문에 계절변동의 형태에 따라 서로 다른 예측방법이 적용되어야 한다. 가법적 계절변동의 특성은 시계열의 계절적 진폭이 시간의 흐름에 따라 일정하지만 승법적 계절변동은 시계열의 진폭이 점차적으로 증가 혹은 감소하게 된다. 본 논문에서의 데이터는 시간의 흐름에 따라 계절적 진폭이 일정하기 때문에 가법적 계절지수평활법 (additive seasonal exponential smoothing method)을 활용하였다. 가법적 계절지수평활법에서의 예측식은 시계열 패턴의 세 가지 성분인 수평성, 추세성, 계절성을 평활한 세 성분의 합으로 나타낸다. 즉 현재 시점이 n 인 경우에 1 시점 후의 예측값은 식 (3.2)와 같이 나타낼 수 있다.

$$\begin{aligned}
F_{n+l} &= a_n + b_n l + S_{n+l-L}, \quad l = 1, 2, \dots, L \\
a_n &= \alpha(Z_n - S_{n-L}) + (1 - \alpha)(a_{n-1} + b_{n-1}) : \text{수평성분}, \\
b_n &= \beta(a_n - a_{n-1}) + (1 - \beta)b_{n-1} : \text{추세성분}, \\
S_n &= \gamma(Z_n - a_n) + (1 - \gamma)S_{n-L} : \text{계절성분},
\end{aligned} \tag{3.2}$$

여기서 $l = 1, 2, \dots, L$ 은 계절성의 길이, S_n =계절인자, α, β, γ 는 평활상수를 나타낸다.

3.3. 실험 데이터

본 연구에서 사용된 데이터는 도시 외곽에 위치한 A사의 생산시설에서 사용된 전력사용량 데이터이다. 해당 시설 수배전반 전력설비에 설치된 다수의 IoT센서에서는 전력량 및 전력 품질에 관한 다양한 정보가 수집된다.

IoT센서는 기존 전력량계가 제공하는 전력량 이외에 전력 품질 데이터에 대해서 5개의 값 (INITIAL, END, MIN, MAX, AVG)이 세부속성으로 분류된다. 또한 사용자가 선택할 수 있는 수집주기 (5분, 15분, 1시간, 1일)로 구분된 데이터가 임시 저장된다. 본 연구에서 사용할 데이터의 수집기간은 2015년 5월 2일 ~ 2015년 7월 27일까지의 총 87일간의 데이터이고, 시설 내부에 IoT센서로 수집되는 각 전력 수요 수요설비의 전력 사용량 Max 값을 합산한 값과 1시간 수집주기 데이터를 사용하였다. 데이터 전 처리에서 결측치는 전체자료에 대한 4분위수 기준으로 다음과 같이 처리하였다. 5월 28일 15시 결측 데이터를 그 이전 값인 14시 데이터 기준 (305.64kw, 4분위수의 최소값 근접)으로 최소값 (305.6kw)을 적용했고, 같은 날 16시 결측 데이터를 그 이후 값인 17시 데이터 기준(1,256.05kw, 4분위수의 최대값)으로 최대값 (1256.05kw)을 적용하였다. 분석과정에 사용된 도구는 R (3.5.1)과 RStudio (1.1.456)이다.

분석을 위해서 87일 중 86일간의 시간별 데이터를 통해 모형 선정을 수행하고, 이후 1일간의 24시간에 대해 수요예측을 하되, 그 예측결과를 2015년 7월 27일 실제 데이터와 비교하였다.

4. 분석과정

Lee (2017)의 연구에 따라 Box-Jenkis 모형 중 seasonal ARIMA모형은 시계열자료의 예측분석에 널리 사용되어 왔고, 이 모형을 통해 시계열 자료를 다루고자 할 때는 모형의 식별과 추정 그리고 검정 단계를 거친다.

4.1. 모형의 식별

모형식별 (model identification) 과정에서는 자료의 산점도를 통해서 시계열의 경향이나 계절성, 이상치, 일정하지 않은 분산, 비정상성 등 비정상성 (non-stationary)을 갖고 있는지 판단한다.

본 연구에서 사용된 데이터에 대해서 추세성과 계절성, 랜덤 경향을 알아보기 위해서 분해법을 적용해 보면 Figure 4.1에서와 같이 추세는 증가한 후 감소하고 다시 증가한 후 유지되는 것으로 보인다.

또한 Figure 4.2는 전체 데이터에 대해 요일별로 시간대별 평균값을 나타내는데, 정오와 저녁 시간대에 전력소모량이 크고 밤과 새벽에는 낮은 사용량을 가지는 계절성의 확인도 필요함을 판단하였다.

이러한 특성들을 통해서 본 연구에서는 ARIMA모형에 계절성을 반영한 SARIMA (seasonal ARIMA) 모형을 사용하였다. 그리고 Figure 4.3에서와 같이 데이터의 ACF와 PACF를 확인해보면 자기상관이 존재한다는 것을 알 수 있고, 이는 비정상시계열자료이기 때문에 정상화 과정이 필요하다. 정상성을 만

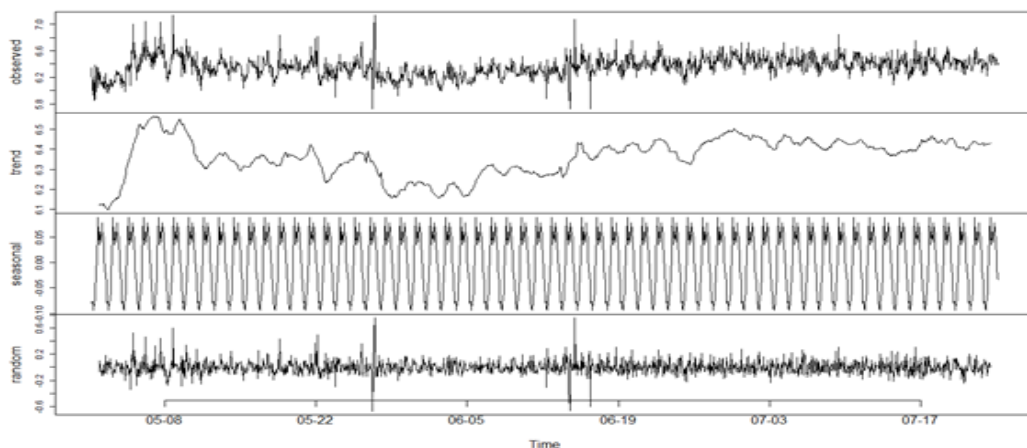


Figure 4.1 Decomposition of time series data

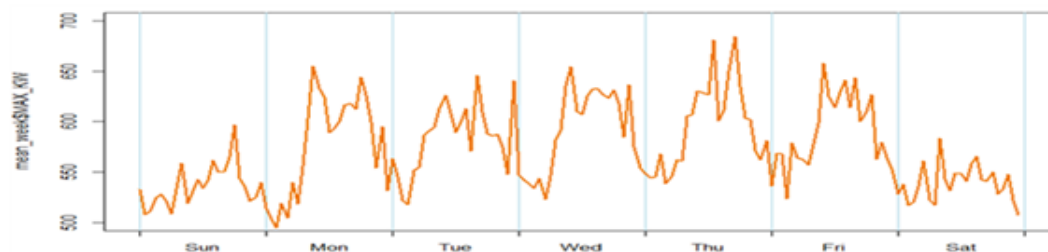


Figure 4.2 Average by day of the week on observation data

측시하기 위해 1차 차분, 계절차분 (24차 차분), 로그 변환을 해주었는데, 1차 차분의 필요는 뚜렷하지 않기 때문에 R에서 지원하는 `ndiffs()` 함수를 사용하여 1차 차분이 필요함을 확인하였다. 이 자료는 24 주기로 계절성을 가지기 때문에 계절차분을 적용, 분산의 안정화를 위해서 로그 변환을 실시하였다.

로그변환과 1차 차분, 계절차분을 처리한 결과, Figure 4.4와 같이 0을 중심으로 데이터가 산포되어 있고, 분산이 비교적 안정화되었다. 이를 통해 정상성 만족을 알 수 있다. 또한 여기에서 변동성이 변하는 성질이 보이는데 본 논문에서는 평활법과 SARIMA 방법을 이용한 모델을 다루었지만, 향후 연구에서는 ARCH나 GARCH 모델을 적용한 분석도 고려해볼 수 있다.

로그 변환, 1차 차분, 계절차분을 수행한 자료에 대해 ACF와 PACF를 확인해보면 Figure 4.5에서와 같이 ACF에서 시차 1, 24에서 최하를 나타내고 절단되는 형태를 보이고, PACF에서 시차 24, 48에서 점차 감소하는 형태를 보이고 있다. 이를 통해서 계절성 MA 차수를 1 또는 2로 식별할 수 있다. 비계절성 파트를 식별하기 위해서 ACF를 보면 1, 24 근처에서 하나 최하를 찍고 절단되는 형태를 보이고 PACF는 1, 24, 48 근처에서 지수적으로 감소하는 형태를 보이므로 비계절성 MA 차수로 1 또는 2를 식별하였다. 따라서 $SARIMA(0,1,1) \times (0,1,1)_{24}$ 이 잠재적 모델로 선정될 수 있음을 판단하였다.

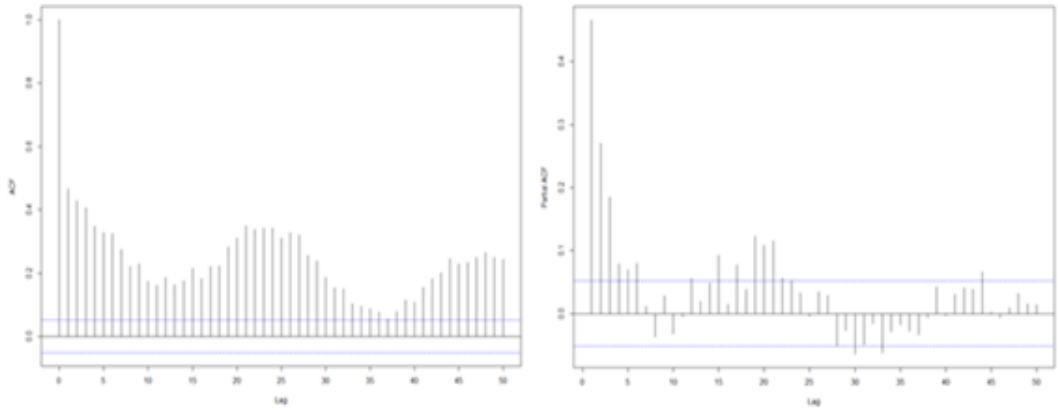


Figure 4.3 ACF and PACF on observation data



Figure 4.4 Apply log, diff and seasonal diff on observation data

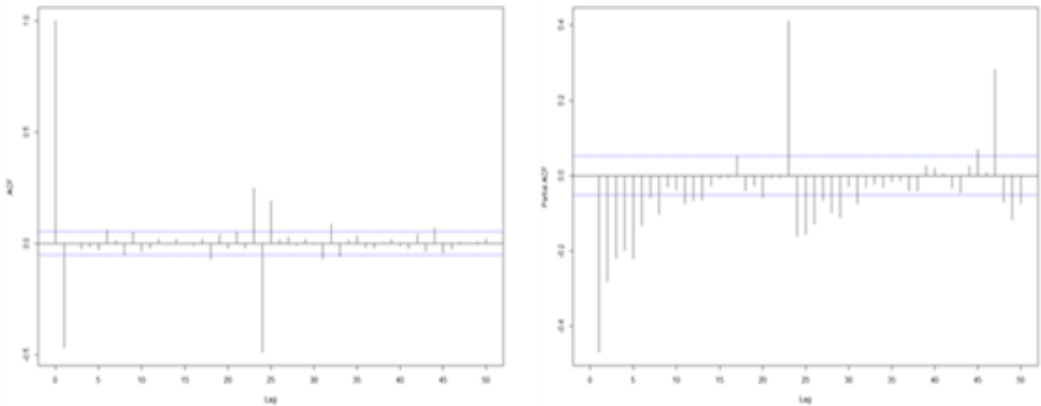


Figure 4.5 ACF and PACF after applying log, diff and seasonal diff

4.2. 모형의 추정

SARIMA(p, d, q) \times (P, D, Q) $_s$ 모형을 추정하는데 있어서 p, d, q, P, D, Q 의 값을 찾는 것이 이 부분에서의 핵심이다. 우선 $d = 1, D = 1$ 이라는 것을 정상성 검정을 통해 알 수 있다. 다음으로 p, q, P, Q 의 값을 추정하여야 하는데 p, q, P, Q 를 추정하기 위해서 AIC (Akaike's information criterion) 통계량과 BIC (Bayesian information criterion) 통계량 등을 이용한다. AIC와 BIC는 모형의 적합도와 간명성을 동시에 고려하는 통계량이다. 모형의 AIC, BIC 값이 작을수록 더 적합한 모델이고 식(4.1)으로부터 도출된다.

$$AIC = -2\ln(\hat{L}) + 2k, \quad BIC = -2\ln(\hat{L}) + \ln(n)k, \quad (4.1)$$

여기서 \hat{L} 은 maximum likelihood, k 는 모수의 개수, n 은 관측수이다.

모델이 아무리 높은 예측력을 가지고 있다 하더라도 모델이 간결하지 않으면 실용성이 떨어진다. 모델을 추정함에 있어서 모수 축약의 원칙을 따라 각 계수들을 2까지만 설정 했다.

Table 4.1 AIC, BIC, RMSE on SARIMA model

	AIC	BIC	RMSE
SARIMA(0,1,1)×(0, 1, 1) ₂₄	-1890.03	-1874.268	0.1197656
SARIMA(0,1,1)×(0, 1, 2) ₂₄	-1890.67	-1869.648	0.1197733
SARIMA(0,1,2)×(0, 1, 1) ₂₄	-1888.56	-1867.538	0.1197525
SARIMA(0,1,2)×(0, 1, 2) ₂₄	-1889.03	-1862.757	0.1197604
SARIMA(0,1,1)×(0, 0, 1) ₂₄	-1799.5	-1783.683	0.1291253
SARIMA(0,1,1)×(0, 0, 2) ₂₄	-1811.86	-1790.777	0.1284579
SARIMA(0,1,2)×(0, 0, 1) ₂₄	-1803.4	-1782.315	0.1288492
SARIMA(0,1,2)×(0, 0, 2) ₂₄	-1817.95	-1791.587	0.1280833

최적의 모형을 찾기 위해 AIC와 BIC 그리고 RMSE를 Table 4.1와 같이 도출하였다. 또한 계절차분의 필요성에 대한 추가 근거로 계절차분을 적용한 모델과 적용하지 않은 모델의 AIC, BIC, RMSE도 비교해보았다. AIC는 SARIMA(0,1,1)×(0, 1, 2)₂₄ 모형이 가장 낮은 값을 보였지만 다른 모형들과의 AIC 차이가 크지 않다. 반면에 BIC를 확인하면 SARIMA(0,1,1)×(0, 1, 1)₂₄모형이 -1874.268로 가장 작은 값을 보였다. 다른 모델들과의 BIC 값 차이가 AIC보다 다소 큰 경향을 보였다. 이러한 근거로 본 연구에서는 BIC를 기준으로 최적의 모델을 선정하였다. 최적의 모델로 추정된 SARIMA(0,1,1)×(0, 1, 1)₂₄ 모델의 계수와 식은 Table 4.2와 식(4.2)과 같다.

Table 4.2 Coefficients on SARIMA(0,1,1)×(0, 1, 1)₂₄ model

	Estimate	Std. error	Z value	Pr(> z)
ma1	-0.873326	0.013930	-62.674	< 2.2e-16***
sma1	-0.971069	0.015197	-63.900	< 2.2e-16***

$$\nabla^1 \nabla_{24}^1 Z_t = (1 - \theta_1 B)(1 - \Theta_1 B^{24})e_t, \quad e_t \sim WN(0, \sigma^2). \quad (4.2)$$

4.3. 모형의 검정

추정된 모형이 통계적으로 유의한지 판단하기 위해서 모형 검정이 필요하고, 이는 잔차검정을 통해서 확인한다. 잔차의 자기상관이 존재한다면 추정의 단계로 돌아가 모형을 다시 추정하게 되고, 자기상관이 나타나지 않는다면 해당 모형이 적합하다고 할 수 있다. Figure 4.6에서 확인되는 바와 같이 잔차는 0을 기준으로 분포하고 있고 특정한 패턴은 보이지 않으므로 i.i.d. 가정을 충족하는 것으로 보인다. 잔차의 ACF도 1 이상의 시차에서 모두 유의하게 작은 값을 가지므로 독립성을 만족한다고 할 수 있다. 또한 백색잡음 검정을 위한 포트맨토 검정의 유의확률이 모든 시차에서 0.05보다 크므로 귀무가설 $H_0 : \rho(k) = 0, k \geq 1$ 을 기각할 수 없기 때문에 백색잡음이 자기상관성이 없다고 할 수 있다. 따라서 SARIMA(0,1,1)×(0, 1, 1)₂₄ 모형은 통계적으로 유의하다고 판단되어 최종모형으로 선정하였다.

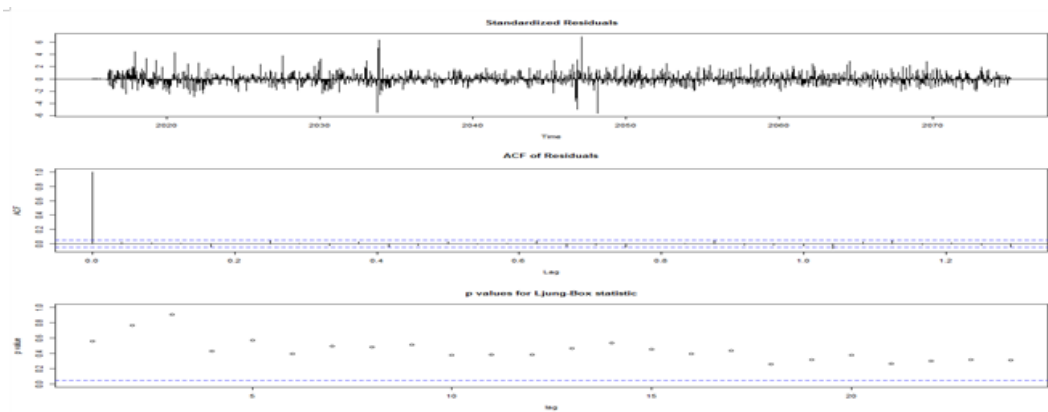


Fig. 4.6 Check of Residuals on $\text{SARIMA}(1,1,2) \times (0,1,1)_{24}$ model

Figure 4.6 Check of residuals on $\text{SARIMA}(1,1,2) \times (0,1,1)_{24}$ model

5. 예측 결과

5.1. SARIMA 모형의 예측

추정된 $\text{SARIMA}(0,1,1) \times (0,1,1)_{24}$ 모형의 미래 24시간에 대한 시계열 데이터를 Table 5.1에서와 같이 실제값과 각 시간대별 예측값, 그리고 신뢰구간에 대한 값을 시간대별로 도출하였다. 이 데이터를 그래프로 나타낸 시계열도는 Figure 5.1과 같고, 파란색 선으로 나타낸 부분이 예측값, 회색 음영 부분이 예측값의 신뢰구간을 나타낸다.

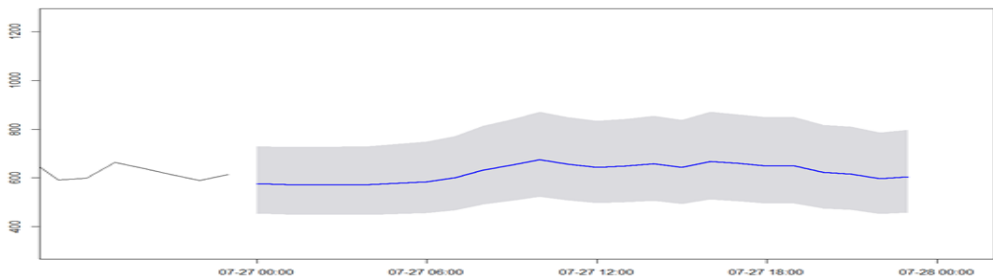


Figure 5.1 Time series plot on following 24 points forecasting using $\text{SARIMA}(0,1,1) \times (0,1,1)_{24}$ model

5.2. 계절지수평활법에 따른 예측

R에서 `HoltWinters()` 함수를 사용하면 모델의 SSE (Sum of square error)를 최소로 해주는 평활상수 α , β , γ 가 지정되고 해당 결과 값은 Table 5.2와 같다.

Table 5.1 Following 24 points forecasting on SARIMA(0,1,1)×(0,1,1)₂₄ model

	True value	Point Forecast	Lo 95	Hi 95
07-27 00:00	502.90	576.2563	454.6596	730.3736
07-27 01:00	534.88	572.7201	451.0146	727.2678
07-27 02:00	490.16	570.9086	448.7440	726.3308
07-27 03:00	448.17	571.0786	448.0415	727.9030
07-27 04:00	601.05	572.0229	447.9528	730.4569
07-27 05:00	600.53	577.8440	451.6811	739.2465
07-27 06:00	610.33	584.5117	456.0610	749.1410
07-27 07:00	626.62	601.0126	468.0879	771.6845
07-27 08:00	572.33	632.6650	491.8555	813.7860
07-27 09:00	593.77	652.5446	506.4064	840.8552
07-27 10:00	750.44	675.6271	523.3917	872.1423
07-27 11:00	744.31	656.3648	507.5760	848.7689
07-27 12:00	591.71	643.5006	496.7593	833.5889
07-27 13:00	606.80	649.3525	500.4075	842.6304
07-27 14:00	582.02	657.9032	506.1237	855.1992
07-27 15:00	566.64	643.0086	673.8188	837.2707
07-27 16:00	734.75	668.2961	512.3666	871.6800
07-27 17:00	611.60	659.2444	504.5730	861.3284
07-27 18:00	632.71	648.2539	495.3282	848.3933
07-27 19:00	574.57	648.1441	494.4181	849.6672
07-27 20:00	600.20	622.1547	473.8059	816.9517
07-27 21:00	622.30	616.2071	468.5030	810.4775
07-27 22:00	623.30	596.6911	452.9218	786.0966
07-27 23:00	565.25	603.7610	457.5418	496.7081

위의 표에서는 가법적 계절지수평활법과 승법적 계절지수평활법을 비교할 수 있는데 MAE값에서는 승법적 계절지수평활법이 더 낮지만 RMSE, MAPE, SSE에서는 가법적 계절지수가 더 적은 값으로 도출되었다. 따라서 가법적 계절지수평활법이 더 올바른 모델로 평가할 근거가 되어 가법적 계절지수평활법을 최종모형으로 선정하였다. 이 때의 가법적 계절지수평활법은 식 (5.1)와 같다.

$$F_{n+l} = a_n + b_n l + S_{n+l-L}, \quad (5.1)$$

여기서 $a_n = 0.1073573(Z_n - S_{n-L}) + (1 - 0.1073573)(a_{n-1} + b_{n-1})$, $b_n = b_{n-1}$, $S_n = 0.08545172(Z_n - a_n) + (1 - 0.08545172)S_{n-L}$, $l = 1, 2, \dots, 24$ 이다.

5.3. 가법적 계절지수평활법 모델과 SARIMA(0,1,1)×(0,1,1)₂₄모델 비교

Prediction sum-of-squares (PRESS)는 적합치와 관측치 사이 편차의 측도이다. 이를 사용하여 모형의 예측 능력을 평가할 수 있고, 일반적으로 PRESS 값이 작을수록 모형의 예측 능력이 더 좋다. PRESS는 식 (5.2)와 같이 나타낸다.

$$PRESS = \sum_{i=1}^n (y_i - f_i)^2. \quad (5.2)$$

이 경우 n 은 예측한 데이터의 개수, y_i 는 실제 값, f_i 는 예측 값으로 표현하고, 가법적 계절지수평활법과 SARIMA(0,1,1)×(0,1,1)₂₄에 대한 PRESS 결과는 Table 5.3과 같다.

Table 5.2 Forecasting on following 24 points using seasonal exponential smoothing method

	t	Observation	Additional seasonal exponential smoothing method	Multiplicative seasonal exponential smoothing method
	t	Observation	Additional	Multiplicative
05-04 00:00	1	561.82	470.14	470.49
05-04 01:00	2	566.05	444.62	444.09
05-04 02:00	3	470.65	460.07	460.15
\vdots	\vdots	\vdots	\vdots	\vdots
07-24 21:00	1462	522.12	612.30	612.16
07-24 22:00	1463	478.71	589.41	589.06
07-24 23:00	1464	530.72	609.15	608.45
	h		F_{t+h}	F_{t+h}
07-27 00:00	1		586.7636	587.7002
07-27 01:00	2		581.2675	582.8412
07-27 02:00	3		586.3590	587.3935
\vdots	\vdots		\vdots	\vdots
07-27 21:00	22		613.7761	613.8064
07-27 22:00	23		590.5821	590.3285
07-27 23:00	24		610.5787	610.0697
	α		0.1073573	0.106532
	β		0	0
	γ		0.08545172	0.1031945
MAE (Mean absolute error)			54.244467	54.17757
RMSE (Root mean squared error)			79.0664	79.17757
MAPE (Mean absolute percentage error)			9.090714	9.135394
SSE (Sum of square)			8852117	8877028

Table 5.3 Result of PRESS between additive seasonal exponential smoothing method and SARIMA(0,1,1) $\times(0,1,1)_{24}$ model

	Additive seasonal exponential smoothing method	SARIMA(0,1,1)
PRESS	93773.87	82684.91

Table. 5.3에서 확인되는 바와 같이, SARIMA(0,1,1) $\times(0,1,1)_{24}$ 모형의 PRESS 값이 가법적 계절지수평활법의 PRESS 값보다 더 적은 82684.91이 도출 되었다. 따라서 SARIMA(0,1,1) $\times(0,1,1)_{24}$ 모형이 더 좋은 예측력을 가지고 있다고 할 수 있다. 최종적으로 예측구간 24시점에 대한 예측 그래프는 Figure 5.2과 같다.

SARIMA 모델과 가법적 계절지수평활법을 이용한 모델의 두 그래프를 보면, 시각적으로는 큰 차이를 보이지 않는다. 그럼에도 SARIMA 최종모델로 선택한 것은 그 PRESS 값이 가법적 계절지수평활법의 값보다 더 작고 값의 차이가 적지않다. 또한 평활법은 단순한 예측기법인 것에 비하여 SARIMA는 이론적으로 충실한 모형이며 잘 정립된 모형으로서, 식별, 모수추정, 모형검진을 거친 통계적 모형으로 이론적인 강점이 있다. 그리고 새로운 데이터 입력에 대한 추가적인 분석에서도 SARIMA가 사용된다면, 모형 요소의 차수만 변경하여 해당 자료의 특성을 반영할 수 있는 일반화의 강점도 있다.

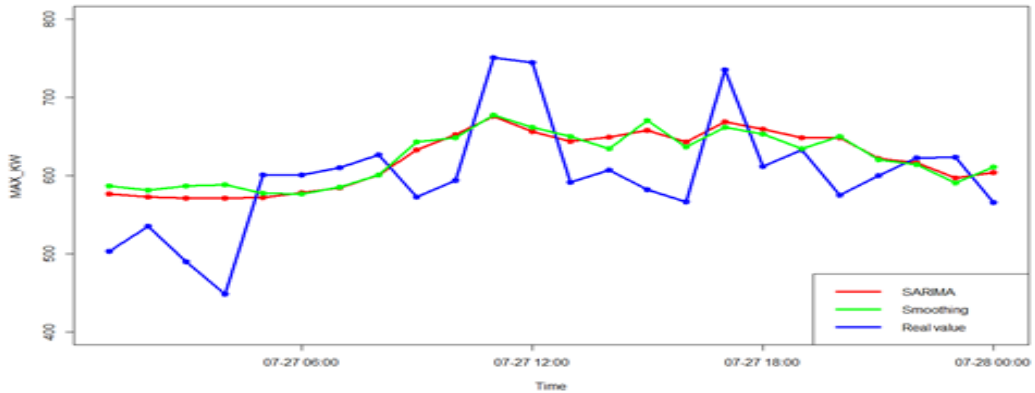


Figure 5.2 Comparison 24 points of forecasting with observation data using SARIMA and seasonal exponential smoothing method

따라서 현업 등의 환경에서 향후 계속적으로 축적되는 데이터에 대해서 예측분석이 필요할 경우, SARIMA(0,1,1)×(0,1,1)₂₄모델을 통해 향후 24시간에 대한 전력사용 및 Peak전력 예측에 활용할 수 있다.

6. 결론

서론에서도 언급한 바와 같이 현재의 국내 전력수급 능력은 탈원전 이슈와 한반도 문제, 전기차 보급과 충전소 확대 등으로 결코 넉넉한 형편이 아니다. ESS를 활용하여 심야의 전력을 저렴하게 저장하고 낮 시간대의 Peak 전력시에 효과적으로 활용하고자 하는 다양한 연구들은 전반적인 전력수급의 안정화를 가져오고, 추가적인 발전원을 줄일 수 있으며, 최근까지도 발생했던 일시/순환 정전을 대비하는데 도움을 줄 것이다. 이와 더불어 비용 효율적으로 구축된 ESS의 용량을 더욱 효과적으로 활용하기 위해서 V2G기술의 접목을 통해 추가적으로 전력을 공급받는 기반이 될 수 있다. 이러한 추가적 전력공급을 위해서 Peak전력량과 그 시간대를 예측하는 것이 중요하다. 중장기적인 예측결과도 물론 쓰임이 있겠지만 본 연구에서는 24시간 기반의 단기 예측을 통해 실험을 수행하였고, 이러한 단기예측이 현업에서의 관련 차세대 시스템에서는 보다 높은 신뢰성으로 기능할 수 있을 것이다.

국내 거의 대부분의 현업에서는 부족한 주요 데이터와 주변 환경 데이터를 가지고 최적의 예측결과를 요구한다. 또한 최적의 예측결과가 기반될 수 있는 데이터를 확보하기 위한 여유시간도 많지 않다. 따라서 예측 정확도를 높일 수 있는 계속된 연구에서는 현재 주어진 데이터를 가지고도 최적의 결과를 도출할 수 있도록 여러 파라미터의 조정과 다양한 분석기법의 접목이 필요할 것이다.

향후 연구의 방향은 동일한 데이터를 가지고 인공지능 분석기법을 활용하여 본 연구에서의 시계열 예측결과와 비교하여 현업에 실질적으로 도움이 되는 연구가 되고자 한다.

References

- Ansari, M. *et al.* (2015). Coordinated bidding of ancillary services for vehicle-to-grid using fuzzy optimization. *IEEE Transaction on Smart Grid*, **6**, 261-270.
- Byeon, G. S. *et al.* (2016). Development of daily operation program of battery energy storage system for peak shaving of high-speed railway substations. *The Transactions of the Korean Institute of Electrical Engineers*, **65**, 404-410.
- Choi, J. W. and Lee, I.D. (2015). Energy ICT convergence with big data services. *Journal of the Korean Data & Information Science Society*, **26**, 1141-1154.
- Hwang, J. H. *et al.* (2014). The V2G operating method considering micro-grid. *The Transactions of the Korean Institute of Electrical Engineers*, **63**, 137-137.
- Kim, G. S., *et al.* (2014). Stepwise technique for improving building energy efficiency rating utilizing quantified simulation model. *Korea Institute of Ecological Architecture and Environment*, **14**, 65-73.
- Kim, H. C. (2017). Trends and implications of V2G business in Korea and other countries. *Jeju Research Institute, Brief of Policy and Issue*, 281.
- Kim, J. H. and Hwang, S. W. (2016). A study on the decision of appropriate subsidy levels regarding electric vehicles for V2G as load management resources. *The Transactions of the Korean Institute of Electrical Engineers*, **65**, 264-268.
- Kim, S. K. *et al.* (2014). Sizing and economic analysis of battery energy storage system for peak shaving of high-speed railway substations. *The Transactions of The Korean Institute of Electrical Engineers*, **63**, 27-33.
- Lee, S. D. (2017). *Time series analysis using SAS and R*, Free Academy.
- Ministry of Trade, Industry and Energy. (2017). *Energy consumption survey*.
- Moon, J. H. *et al.* (2017). Power consumption forecasting scheme for educational institutions based on analysis of similar time series data. *Journal of KIISE*, **44**, 954-965.
- Park, D. I. and Yoon, S. H. (2017). Clustering and classification to characterize daily electricity demand. *Journal of the Korean Data & Information Science Society*, **28**, 395-406.
- Ryu, J. A. *et al.* (2017). Forecasting daily building heating energy demand using ARIMA model. *Architectural Institute of Korea*, **37**, 608-611.
- Son, E. T. *et al.* (2018). A study on the retail electricity market based on the local pool for small energy prosumers. *The Transactions of the Korean Institute of Electrical Engineers*, **67**, 165-172.
- Yoon, J. H. *et al.* (2017). Analysis of energy consumption patterns according to usage behavior and architectural planning factors in university buildings. *Architectural Institute of Korea*, **37**, 527-530.

Analysis of time series to support decision making on V2G using energy consumption data[†]

Jeong Hyun Lee¹ · Sang Jin Oh² · Yeochang Yoon³ · Young Ho Ahn⁴ · Jae Sung Kim⁵ ·
Wan Sup Cho⁶ · Sung Duck Lee⁷

¹⁵⁶Department of Bigdata, Chungbuk National University

²⁷Department of Information Statistics, Chungbuk National University

³Woosuk University

⁴RETIGRID

Received 31 December 2018, revised 21 January 2019, accepted 22 January 2019

Abstract

The domestic energy demand is currently in crisis, and this is related with the phase-out of nuclear power plant and expansions of EV charging stations. Moreover, the additional demand will be increased because of the political issues of the Korean peninsula, too. Many studies have focus on and dealt with alternative energy development, energy efficiency and energy storage system. In this study, we analyzed time series data of the energy consumption. As the result, we recognized the importance of predicting the amount of energy peak load and its time zone when it is converted from main power to ESS, and energy trading. Therefore our goal is to get more reliable results through a short time period analysis based on 24 hours. In order to conduct this research study, Seasonal ARIMA (SARIMA) model and seasonal exponential smoothing method were used. We hope this analyzed result will be utilized positively on real environment of V2G or ESS in the future.

Keywords: ESS, Energy peak load, Seasonal ARIMA, Seasonal exponential smoothing, V2G.

[†] This research is supported by a research fund of university ICT center on Ministry of Science and ICT/IITP(IITP-2018-0-01396).

¹ Ph. D., Chungbuk National University, 1 ChungDae-ro, Seowon-gu, Chungbuk, Korea.

² Bachelor, Chungbuk National University, 1 ChungDae-ro, Seowon-gu, Chungbuk, Korea.

³ Ph. D., Woosuk University, 443 Samnye-ro, Samnye-eup, Wanju-Gun, Jeonbuk, Korea.

⁴ CEO, RETIGRID, Songpa bldg, 8 Baekjegobun-ro 41-gil, Songpa-gu, Seoul, Korea.

⁵ Ph. D., Chungbuk National University, 1 ChungDae-ro, Seowon-gu, Chungbuk, Korea.

⁶ Professor, Chungbuk National University, 1 ChungDae-ro, Seowon-gu, Chungbuk, Korea.

⁷ Corresponding author: Professor, Department of Information Statistics, Chungbuk National University, 1 ChungDae-ro, Seowon-gu, Chungbuk, Korea. E-mail: sdlee@chungbuk.ac.kr