

The Forecasting Power Energy Demand by Applying Time Dependent Sensitivity between Temperature and Power Consumption

Jinho Kim · Chang-Yong Lee[†]

Dept. of Industrial and Systems Engineering, Kongju National University

시간대별 기온과 전력 사용량의 민감도를 적용한 전력 에너지 수요 예측

김진호 · 이창용[†]

공주대학교 산업시스템공학과

In this study, we proposed a model for forecasting power energy demand by investigating how outside temperature at a given time affected power consumption and. To this end, we analyzed the time series of power consumption in terms of the power spectrum and found the periodicities of one day and one week. With these periodicities, we investigated two time series of temperature and power consumption, and found, for a given hour, an approximate linear relation between temperature and power consumption. We adopted an exponential smoothing model to examine the effect of the linearity in forecasting the power demand. In particular, we adjusted the exponential smoothing model by using the variation of power consumption due to temperature change. In this way, the proposed model became a mixture of a time series model and a regression model. We demonstrated that the adjusted model outperformed the exponential smoothing model alone in terms of the mean relative percentage error and the root mean square error in the range of 3%~8% and 4kWh~27kWh, respectively. The results of this study can be used to the energy management system in terms of the effective control of the cross usage of the electric energy together with the outside temperature.

Keywords : Time Series, Power Energy Consumption, Temperature, Exponential Smoothing, Forecast

1. 서 론

전력 수요의 특징 분석 및 예측은 전력 시장의 가격 결정과 전력 계통의 효율적 운용을 위해서 매우 중요하다. 예를 들어 국내 전력 에너지 소비의 특징 중 하나는 냉난방의 전력의존도가 높아 여름은 물론 겨울에도 전력

피크(peak)가 발생하는 것이다[8]. 이것은 겨울철에도 전력을 사용한 난방 수요가 확대되고 있음을 의미하는 것으로 지속적인 난방 수요의 확대는 전력 소비의 변동성을 확대시킬 수 있는 잠재 요인이 되며, 전력 시장의 경제성과 안정성을 위해 전력수요 예측을 위한 모형 및 알고리즘 개발의 필요성을 뒷받침한다[7, 20]. 계절적 요인 외에도 외부 기온에 따라 단기적으로 전력 사용량이 영향을 받을 수 있기 때문에 온도와 전력 사용량 사이의 상관관계를 파악하는 것이 단기 전력수요 예측을 위해 중요하다.

Received 21 January 2019; Finally Revised 21 January 2019;
Accepted 20 February 2019

[†] Corresponding Author : cleec@kongju.ac.kr

예측을 위한 모형은 회귀 모형(regression model)과 시계열 모형(time series model)으로 크게 구분할 수 있다. 연속형 데이터에 주로 사용하는 회귀 모형은 독립변수 값이 종속변수에 미치는 영향 혹은 두 변수들 간의 관계를 적합하기 위한 수리적 모형이다[2, 14]. 온도를 고려한 전력수요 예측의 경우, 회귀 모형은 온도를 독립변수로 사용하여 종속변수인 전력 수요량과의 관계를 분석하는데 적용될 수 있다.

시계열 모형은 시간을 독립변수로 하여 과거에 측정된 데이터를 사용하여 미래를 예측하는 것으로 선형 모형과 비선형 모형으로 크게 구분할 수 있다. 선형 모형으로는 자기 회귀(auto-regression), 이동 평균(moving average), 자기 회귀와 이동 평균의 혼합 모형인 ARIMA[3], 그리고 지수평활(exponential smoothing)[21] 등이 있다. 비선형 모형으로는 지지도 벡터 회귀(SVR, support vector regression) [11]과 순환 신경망(RNN, recurrent neural networks)[17-19]의 Elman 신경망[4]과 LSTM(long short term memory) [6] 등이 있다.

전력 수요와 관련된 기존의 연구를 예측 방법 측면에서 보면, 뉴로-퍼지 모델(neuro-fuzzy model)의 구조 학습을 이용하여 전력 수요를 예측하는 모의실험을 수행한 연구[15, 22]와 다중 회귀 분석을 이용한 지역 전력 수요 예측 알고리즘[12], 그리고 SVR을 이용하여 월별 전력수요 예측 모형에 대한 연구[11] 등이 있다. 전력 수요에 대한 실용적 연구를 보면, 전력 사용량 데이터를 기반으로 소규모 사업체의 에너지 사용 추이 및 소비 예측을 분석한 연구[23]가 있으며, 시도별 16개 지역의 전력 사용량을 ARIMA 모형을 이용하여 중장기(12개월)적으로 예측하는 방법을 제안한 연구[1]가 있다. 또한 최근 각광을 받고 있는 빅데이터에 대한 연구가 전력 산업에서 활용되고 있는 현황 및 전망에 대한 연구도 있다[13].

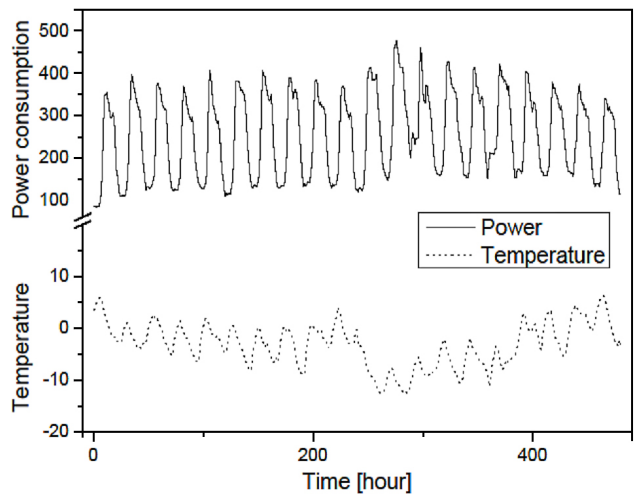
시간은 전력 수요의 단기간 예측에 중요한 인자이다. 시계열 모형은 과거와 현재에 사용된 전력 사용량을 바탕으로 미래의 전력 수요를 예측하며, 회귀 모형도 시간과 전력 사용량 간의 선형 혹은 비선형 관계가 있는 경우 최소 제곱법 등을 사용하여 전력수요 예측에 적용할 수 있다.

그러나 시간뿐만 아니라 외부 온도와 같은 요인을 추가로 고려한 전력 수요의 예측을 위해서는 시간을 고려한 모형에 더하여 외부 온도의 영향을 포함하는 예측 모형을 설정해야한다. 이를 위해서는 온도와 전력 사용량 간의 관계를 분석한 결과를 시간을 고려한 모형에 추가한 모형을 설정해야 한다. 따라서 새로운 모형은 시간을 고려한 시계열 모형과 온도와 전력 사용량 간의 관계를 위한 회귀 모형을 결합한 모형이 될 가능성이 높다.

본 연구에서는 외부 온도가 전력 사용량에 미치는 영향을 분석하고, 분석 결과를 시간뿐만 아니라 외부 온도를 고려한 전력수요 예측 모형에 대한 연구를 수행하였다. 이를 위하여 전력 사용량 시계열 데이터 외에 외부 온도 시계열 데이터를 추가적으로 고려하여 온도가 전력 사용량에 미치는 영향을 분석하고, 그 결과를 전력수요 예측에 적용하였다. 또한 설정한 모형의 적합성 및 예측의 정확성 등을 분석하고, 그 결과를 향후 전력 에너지 관리 시스템 개발에 적용할 수 있는 기반을 제공하고자 한다.

본 연구에서는 201*년 A 테크노파크에서 사용된 전력 시계열 데이터와 기상청에서 제공하는 A 테크노파크 지역의 온도 시계열 데이터 등 2종류의 시계열 데이터를 사용하였다. 전력 사용량은 15분 간격으로 측정되었으나 기상청의 온도 데이터는 1시간 단위로 측정되었기 때문에 시간 단위를 맞추기 위해 1시간 동안 4번 측정된 전력 사용량의 평균값을 전력 사용량 시계열 데이터로 사용하였다. 따라서 본 연구에서 사용한 시계열 데이터는 시간 i 에서 온도 $\tau(i)$ 와 평균 전력 사용량 $p(i)$ 로 구성된 $\{\tau(i), p(i) | i = 1, 2, \dots, N\}$ 로 나타낼 수 있다. <Figure 1>은 1월의 전력 사용량과 온도 시계열 데이터를 나타낸 것이다.

본 논문은 다음과 같이 구성되어 있다. 제 1장 서론에 이어 제 2장에서는 온도에 따른 단기간 전력수요 예측 모형을 위하여 회귀 모형과 시계열 모형을 결합한 혼합 모형을 제안하였다. 제 3장에서는 제안한 예측 모형을 실제 데이터에 적용하여 그 유의성을 검증하였고, 마지막으로 제 4장에서 본 연구의 요약과 결론을 맺었다.



<Figure 1> The Time Series of Power and Temperature in January

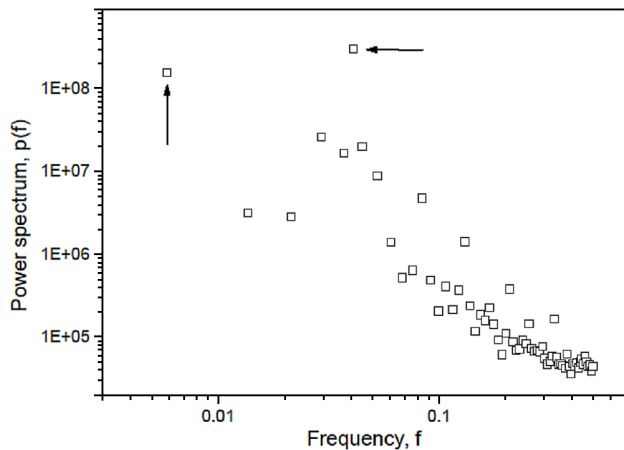
2. 온도 변화에 따른 단기간 전력 사용량 예측을 위한 모형

2.1 파워 스펙트럼(Power Spectrum) 분석

<Figure 1>에서 볼 수 있듯이 전력 사용량 시계열 데이터 $\{p(i)\}$ 에 주기가 있는 것으로 판단됨으로 주기성을 보다 체계적으로 분석하기 위하여 전력 사용량 시계열 데이터에 파워 스펙트럼을 적용하였다. 파워 스펙트럼은 N 개의 시계열 데이터 $p(1), p(2), \dots, p(N)$ 을 주파수 측면에서 분석하여 시계열 데이터에 내재한 주기성을 규명하기 위한 방법이다. 주파수 $f_n = \frac{n}{N}$, $n = -\frac{N}{2}, n = \frac{N}{2}$ 에서 전력 사용량 시계열 데이터 $\{p(i)\}$ 에 대한 파워 스펙트럼 $F(f_n)$ 은 아래와 같이 계산할 수 있다.

$$F(f_n) = |H_n|^2 + |H_{-n}|^2, H_n = \sum_{t=1}^N p(t)e^{2\pi n t/N} \quad (1)$$

위의 식 (1)은 보통 빠른 푸리에 변환(FFT, fast Fourier transform)[5]을 사용하여 계산하며, <Figure 2>는 전력 사용량 시계열 데이터에 대하여 파워 스펙트럼을 실행한 결과를 나타낸 것이다.



<Figure 2> The Power Spectrum of the Time Series of Power Consumption. Note that Axes of Abscissa and Ordinate are in the Logarithmic Scale

파워 스펙트럼 분석을 통해 시계열 데이터에 내재되어 있는 두 가지 주기성을 발견할 수 있다. 파워 스펙트럼은 $f \approx 0.041$ (<Figure 2>의 ← 화살표)에서 최대임을 알 수 있는데, 이것은 $\frac{1}{f} \approx 24$ 의 주기를 가지고 있음을 의미한다. 전력 사용량 시계열 데이터는 매 시간 측정된 것임으로 $\frac{1}{f} \approx 24$ 은 1일에 해당된다. 따라서 전력 사용량

은 하루의 주기를 가짐을 알 수 있다. 다음으로 세기가 큰 파워 스펙트럼의 주파수는 $f \approx 0.006$ (<Figure 2>의 ↑ 화살표)이며, 주기는 $\frac{1}{f} \approx 167$ 시간으로 약 7일에 해당한다. 따라서 전력 사용량은 하루의 주기성이 가장 뚜렷하며, 또한 한 주 주기로 변화함을 알 수 있다.

본 연구의 목적이 비교적 짧은 시간(1~2시간) 내의 전력 수요 예측이고 전력 사용량은 하루의 주기성이 가장 강함으로, 하루 단위로 전력 수요를 예측해도 모형의 매개변수 값은 변할지라도 모형의 전반적인 형태는 변하지 않음을 예상할 수 있다. 따라서 본 연구에서는 시계열 데이터를 하루 단위로 나누어 분석과 예측을 수행하고자 한다.

2.2 시간대별 온도와 전력 사용량의 관련성 분석

외부 온도를 고려하지 않고 시간에 따른 전력수요 예측을 위해서는 일반적인 시계열 방법을 사용할 수 있다. 그러나 전력 사용량 외에 외부 온도 시계열 데이터를 추가적으로 고려하여 전력 수요를 예측하기 위해서는 온도와 전력 사용량 사이의 관계를 먼저 분석해야 한다. 두 데이터 사이의 관계를 수치적으로 표현하는데 가장 널리 사용하는 방법으로 상관계수(correlation coefficient)[10]를 들 수 있다. 온도와 전력 사용량 시계열 데이터 $\{\tau(i), p(i) | i = 1, 2, \dots, N\}$ 사이의 상호상관계수(correlation coefficient) ρ 는

$$\rho = \frac{cov(\tau, p)}{\sqrt{var(\tau)var(p)}} \quad (2)$$

로 주어진다. 여기서 $cov(\cdot)$ 와 $var(\cdot)$ 는 각각 공분산과 분산을 나타낸다.

1년 동안 측정된 온도와 전력 사용량 시계열 데이터 $\{\tau(i), p(i) | i = 1, 2, \dots, N\}$ 를 사용하여 상관계수 ρ 를 추정했으며, 그 결과 $\hat{\rho} = -0.10$ 을 얻었다. $\hat{\rho} < 0$ 인 것은 온도가 올라가면 전력 사용량은 감소(혹은 온도가 내려가면 전력 사용량은 증가)함을 의미하는데, 그 크기가 0.1로 시간에 따른 온도와 전력 사용량 사이에는 상관관계가 거의 존재하지 않음을 의미한다. 이 결과는 온도를 고려한 전력 수요를 시계열적으로 예측하기 힘들음을 의미한다.

위의 결과를 좀 더 세밀하게 분석하기 위하여 시계열 데이터를 월별로 나누어 온도와 전력 사용량 간의 상관관계를 분석하였다. <Table 1>은 온도와 전력 사용량의 월별 시계열 데이터에 대하여 상관계수를 계산한 결과이다. <Table 1>을 통해 볼 때, 온도와 전력 사용량 시계열 데이터는 모든 월에 대하여 음의 상관관계를 가지며, 이것은 기온이 내려감에 따라 전력 사용량은 증가함을 의미

한다. 그러나 상관계수의 절대 크기는 0.5을 넘지 않음으로 온도에 따른 전력 사용량 사이의 관계를 시계열적으로 파악하기에는 한계가 있다.

<Table 1> Monthly Correlation Coefficients

Month	1	2	3	4	5	6
ρ	-0.27	-0.18	-0.23	-0.34	-0.25	-0.27
Month	7	8	9	10	11	12
ρ	-0.15	-0.07	-0.27	-0.24	-0.28	-0.17

전력 사용량 시계열 데이터가 하루의 주기를 가짐에 착안하여 하루 동안 동일한 시간 i 에서 온도 $\{\tau(i)\}$ 에 따른 전력 사용량 $\{p(i)\}$ 의 관계를 조사하였고, 그 결과 중 일부를 <Figure 3>에 나타내었다. <Figure 3>에서 볼 수 있듯이 온도에 따른 주중과 주말의 전력 사용량은 큰 차이가 있는데, 온도 변화에 따른 주말의 전력 사용량의 변화는 주중에 비하여 적음을 알 수 있다. 이것은 주말에는 고정 설비만 가동되고 일반 사무실에 필요한 냉난방 가동은 중지됨을 의미한다. 특히 온도에 따른 전력 사용량은 단순한 관계가 아니며 동일한 온도에서 소모한 전력에는 큰 차이가 있음을 알 수 있다. 이러한 특성은 날짜에 무관하게 유지되었고, 이 결과는 시계열적으로 전력 사용량을 온도의 함수로 표현하기 어려움을 의미한다.

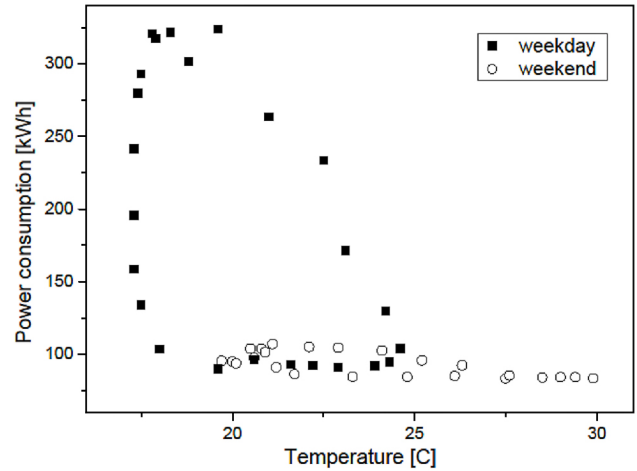
<Table 1>과 <Figure 3>을 통해 시간에 따른 온도와 전력 사용량 사이에 함수 관계를 유추하기 힘들음을 알 수 있다. 이것은 온도와 전력 사용량의 관계가 시간에 따라 다를 가능성을 내포하고 있다. 예를 들어, 오전 시간 중의 온도와 전력량 사이의 관계와 오후 시간 중의 관계가 다를 수 있다. 이것은 전력 사용량 및 온도 시계열 데이터를 시간대별로 구분하여 각 시간대에 온도에 따른 전력 사용량의 변화를 분석할 필요를 의미한다.

이를 위하여 온도와 전력 사용량 시계열 데이터 $\{\tau(i), p(i) | i=1, 2, \dots, N\}$ 를 고정된 시간 $t(t=1, 2, \dots, 24)$ 에 대한 시계열 데이터

$$\{\tau_t(j), p_t(j) | j=1, 2, \dots, m\}, \quad m = N/24 \quad (3)$$

로 나누었다. 위의 식 (3)에서 아래 첨자 t 는 하루 중 고정된 시간을 나타내며, $\tau_t(j)$ 와 $p_t(j)$ 는 고정된 시간 t 에 대한 온도와 전력 사용량 시계열 데이터이다. 예를 들어 $\{\tau_9(j), p_9(j)\}$ 는 $\{\tau(i), p(i) | i=1, 2, \dots, N\}$ 에서 9시에 측정된 온도와 전력 사용량만으로 구성된 시계열 데이터를 의미한다.

고정된 시간 $t=1, 2, \dots, 24$ 에 대하여 1년간 측정된 온도 $\tau_t(j)$ 와 전력량 $p_t(j)$ 의 관계를 조사하였으며, 대표적으로 $t=15$ 의 결과를 <Figure 4>에 나타내었다. 모든 시



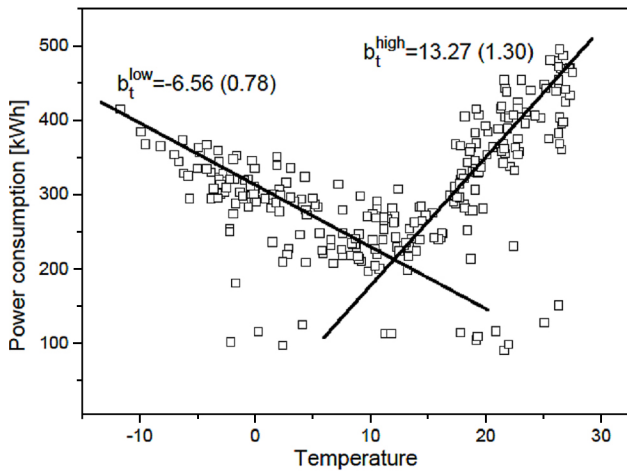
<Figure 3> Plots of the Power Consumption Versus Temperature in One Weekday and One Weekend of June

간대에 대한 온도와 전력 사용량 사이의 관계를 분석한 결과 두 가지 특징을 파악할 수 있다. 첫째, 온도와 전력 사용량의 관계는 측정 시간대 t 에 의존하였다. 업무 시간 ($t=9, 10, \dots, 18$)의 경우에는 <Figure 4>에서 볼 수 있듯이 대략적인 선형적 관계를 유지하였으나, 업무 시간 외에는 온도에 따른 전력 사용량의 경향을 정량화하기 힘들었다. 둘째, 업무 시간의 경우, 임계 온도(critical temperature)를 기준으로 온도에 따른 전력 수요량은 선형적으로 증가 혹은 감소하였다. <Figure 4>을 통해 볼 수 있듯이 $t=15$ 의 경우에는 임계 온도 약 10도($\tau_c=10$)를 기준으로 임계 온도보다 낮은 경우에는 온도가 오름에 따라 대체적으로 전력 사용량이 줄어들고, 임계 온도보다 높은 경우에는 전력 사용량이 증가한다. 이러한 특성은 증가 및 감소의 정도와 임계온도 값의 차이가 있지만, 모든 업무 시간대에 나타났다.

업무 시간대의 온도와 전력 사용량 사이의 선형적 특징을 바탕으로 업무 시간($9 \leq t \leq 18$) 각각에 대하여 온도에 따른 전력 소모량의 관계를 선형 회귀식으로 표현할 수 있다. 즉, 시간 t 에서 전력사용량 $p_t(j)$ 와 온도 $\tau_t(j)$ 사이에는 아래와 같은 선형 회귀식

$$p_t(j) = a_t + b_t \tau_t(j) + e(j), \quad e(j) \sim N(0, \sigma^2) \quad (4)$$

을 설정할 수 있다. 여기서 a_t 는 절편, b_t 는 기울기를 나타내며, $e(j)$ 는 잔차(residual)로 정규 분포를 따른다고 가정한다. 온도에 따른 전력 사용량의 민감도는 식 (4)의 기울기로 표현되며, 기울기 추정치 \hat{b}_t 는 측정 시간 t 뿐만 아니라 온도 $\tau_t(j)$ 와 임계온도 τ_c 의 관계에 따라 달라진다. <Figure 4>에서 \hat{b}_t^{low} 와 \hat{b}_t^{high} 는 각각 $\tau_t < \tau_c$ 와 $\tau_t \geq \tau_c$ 인 경우의 기울기 추정치를 나타낸다.



<Figure 4 > Plots of Power Consumption Versus Temperature at $t = 15$

<Table 2> Slope Estimates for All Business Hours. The Estimate Errors are in Parentheses

t	τ_c	\hat{b}_t^{low}	\hat{b}_t^{high}
9	19	-8.74(0.53)	17.88(2.04)
10	18	-8.24(0.54)	17.43(2.08)
11	18	-7.56(0.55)	18.42(2.27)
12	17	-6.64(0.55)	17.78(2.25)
13	13	-7.11(0.70)	14.66(1.46)
14	11	-6.87(0.82)	13.73(1.23)
15	11	-6.56(0.78)	13.27(1.30)
16	11	-6.29(0.71)	13.19(1.38)
17	11	-7.16(0.72)	12.48(1.31)
18	10	-7.00(0.66)	8.06(0.97)

모든 업무 시간($9 \leq t \leq 18$)에 대하여 기울기를 추정 한 결과를 <Table 2>에 나타내었다. <Table 2>에서 볼 수 있듯이 임계온도 τ_c 는 측정 시간 t 에 따라 다른데, 전반적으로 시간이 흐름에 따라 감소하는 경향을 보인다. 추정된 기울기 역시 시간 t 에 의존하며, 온도 τ 가 임계온도 보다 높을 때 기울기의 절대값이 낮을 때에 비하여 높음을 알 수 있다. 만약 온도의 변화에 따른 전력 사용량 증감의 주된 원인이 냉난방이라면, 이러한 특징은 냉방을 위한 전력이 난방보다 더 많이 필요함을 의미한다.

2.3 온도를 고려한 전력수요 예측 모형

위에서 분석한 온도에 따른 전력 사용량 민감도는 온도 변화에 따른 전력 사용량의 증가 혹은 감소 정도를 파악하거나, 기존의 전력수요 예측 방법을 보완하기 위한 방법으로 사용될 수 있다. 전력수요 예측을 위한 기존

시계열 방법[2, 3, 21]들은 과거 전력 사용량만 사용하였으나, 온도를 추가로 고려함으로 온도가 전력 사용량에 미치는 영향을 정량적으로 파악할 수 있을 뿐만 아니라 보다 정확한 예측이 가능할 수 있다.

본 연구에서는 전력 수요 예측을 위한 시계열 모형으로 선형 모형 중에서 지수평활 모형을 사용하고, 온도에 따른 전력 사용량 민감도를 추가로 적용하여 민감도가 수요 예측에 미치는 영향을 분석할 수 있는 방법과 모형을 제안한다. 지수평활 모형을 선택한 이유는 다음과 같다. 비선형 모형인 SVR과 RNN은 전력량 시계열 데이터와 모형의 오차가 최소가 되도록 모형의 매개변수(parameter)들을 추정하기 때문에 SVR 혹은 RNN에 온도에 따른 전력 사용량 효과가 이미 함축적으로 내재되었다고 간주할 수 있다.

선형 모형 중에서 ARIMA 모형 등은 시계열 데이터 자체를 적합하지 않는다는 점에서는 지수평활 모형과 유사한 특징을 가지고 있지만 모형에 포함된 매개변수가 상대적으로 많기 때문에 추정된 매개변수 값에 따라 성능의 차이가 지수평활 모형에 비하여 클 수 있다. 반면 지수평활 모형의 매개변수는 평활계수(smoothing coefficient)가 유일하기 때문에 매개변수의 추정으로 인한 복잡성을 피하고 온도에 따른 전력 사용량의 변화가 예측에 미치는 효과를 보다 명확하게 분석할 수 있는 장점이 있다.

지수평활 모형은 예측할 시점과 가까운 시점의 데이터에 더 큰 가중치를 부여하고 멀어질수록 가중치를 지수적으로 줄여나가는 방법으로, 관측치 $\{p(1), p(2), \dots\}$ 에 대하여 예측치 $\{\hat{p}(2), \hat{p}(3), \dots\}$ 은 $t = 2, 3, \dots$ 일 때

$$\hat{p}(t+1) = \alpha p(t) + (1-\alpha)\hat{p}(t) \quad (5)$$

로 주어진다[16]. 여기서 $\hat{p}(2) = p(1)$ 이며, α 는 평활계수로 $0 < \alpha < 1$ 의 값을 가진다. α 는 예측 시간과 가까운 시간에서의 관측치에 대한 가중치를 나타내는 것으로 α 가 클수록 예측시간과 가까운 관측치에 대한 중요도가 높아진다. 편의상 식 (5)를 Model (I)이라 한다.

식 (4)를 사용하여 추정된 회귀모형의 기울기 \hat{b}_t 는 전력 수요의 시계열적 예측에 적용할 수 있다. 시간이 t 에서 $(t+1)$ 로 변화할 때 온도 변화는 $(\tau_{t+1} - \tau_t)$ 임으로 이에 따른 전력 사용량의 변화 $\Delta p(t)$ 는 식 (4)를 이용하면

$$\Delta p(t) \approx \hat{b}_t (\tau_{t+1} - \tau_t),$$

$$\hat{b}_t = \begin{cases} \hat{b}_t^{low}, & \text{if } \tau_t < \tau_c \\ \hat{b}_t^{high}, & \text{if otherwise} \end{cases} \quad (6)$$

와 같이 주어진다. 시간에 따른 회귀모형의 기울기 \hat{b}_t 는 <Table 2>에 주어졌으므로 <Table 2>와 식 (6)을 사용하면 시간이 t 에서 $(t+1)$ 로 변할 때 온도 변화에 따른 전력 사용량의 변화 $\Delta p(t)$ 를 구할 수 있다.

식 (6)를 Model (I)에 추가하면 전력 사용량 민감도를 추가로 고려한 모형이 되며, 이것은 지수평활 모형에 온도 변화에 따른 전력 사용량을 보정한 모형으로 다음과 같이 주어진다.

$$\hat{p}^{adj}(t+1) = \hat{p}(t+1) + \Delta p(t), t = 1, 2, \dots \quad (7)$$

여기서 $\hat{p}(t+1)$ 과 $\Delta p(t)$ 는 각각 식 (5)와 식 (6)에서 구할 수 있다. 편의상 식 (7)을 Model (II)라 한다.

식 (7)은 시계열 모형인 지수평활 모형과 회귀 분석으로 추정된 온도 변화에 따른 전력 사용량을 결합한 것으로 시계열 모형과 회귀 모형의 혼합 모형이다. 또한 식 (7)의 혼합 모형을 사용하면 전력수요 예측 $\hat{p}^{adj}(t+1)$ 에서 온도에 의한 요인인 $\Delta p(t)$ 가 차지하는 비율을 계산할 수 있다.

3. 실험결과

위에서 제안한 두 모형[Model (I)과 (II)]을 단기간 전력 수요 예측에 적용하였다. 제안한 모형의 성능 평가를 위한 예측 정확도 척도로 일반적으로 많이 사용되는 평균 상대 오차 백분율(MRPE, mean relative percentage error)과 평균 제곱근 오차(RMSE, root mean square error)를 사용하였다. 전력 사용량 시계열 데이터 $\{p(t)|t=1, 2, \dots, n\}$ 를 Model (I)을 사용하여 $t=t_0$ 부터 $t=n-1$ 까지 예측한 결과를 $\{\hat{p}(t+1)|t=t_0, 2, \dots, n-1\}$ 라 하면, MRPE와 RMSE는 각각

$$MRPE = \frac{1}{n-t_0} \sum_{t=t_0}^{n-1} \left| \frac{p(t+1) - \hat{p}(t+1)}{p(t+1)} \right| \times 100, \quad (8)$$

$$RMSE = \sqrt{\frac{1}{n-t_0} \sum_{t=t_0}^{n-1} \{p(t+1) - \hat{p}(t+1)\}^2}$$

로 주어진다. 위의 식 (8)에서 $\hat{p}(t+1)$ 대신 $\hat{p}^{adj}(t+1)$ 를 사용하면 Model (II)에 대한 예측 정확도가 된다.

임의로 선택한 10일치 온도와 전력 사용량 시계열 데이터에 대하여 Model (I)과 (II)를 사용하여 하루($t_0=1, n=24$) 동안 전력수요를 예측하였다. 두 모형의 예측 성능을 비교하기 위하여 MRPE와 RMSE를 구하였고, 온도에 따른 전력 사용량이 수요 예측에 미치는 영향을 파악하기 위해서

$$r \equiv \Delta p(t) / \hat{p}^{adj}(t+1) \times 100 \quad (9)$$

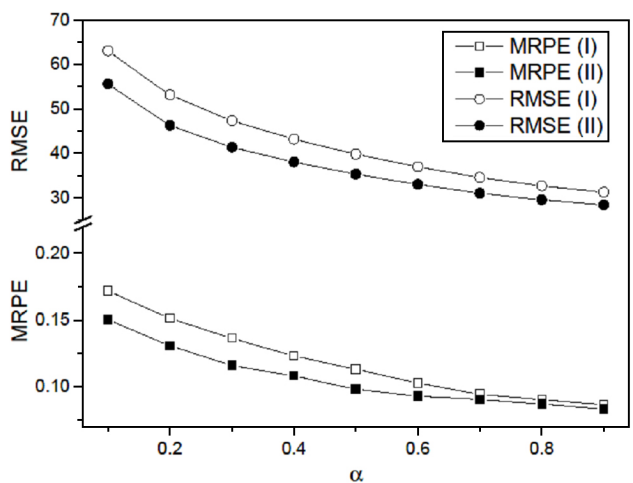
을 계산하였으며, 그 결과를 <Table 3>에 나타내었다.

<Table 3>을 통해 크게 3가지를 알 수 있다. 첫째, 온도 변화에 따른 전력 사용량이 수요 예측에 미치는 영향은

<Table 3> Results of Experiments by Using the Two Models. We used $\alpha = 0.8$, and SD Stands for the Standard Deviation

Sample	r	MRPE		RMSE	
		Model (I)	Model (II)	Model (I)	Model (II)
1	0.81	4.45	3.86	15.92	13.58
2	2.05	3.74	2.85	13.69	10.07
3	0.04	3.78	3.85	15.90	17.19
4	1.23	3.65	2.85	12.43	9.63
5	1.19	4.33	3.59	16.03	14.07
6	2.52	5.18	4.33	11.38	10.71
7	0.16	5.41	5.19	15.93	15.51
8	0.90	4.40	4.09	11.49	10.54
9	1.46	3.16	7.68	4.22	9.38
10	4.77	7.53	7.29	27.39	27.03
mean	1.51	4.56	4.56	14.44	13.77
SD	1.37	1.25	1.69	5.81	5.38

표본에 따라 0.04%에서 약 5%까지 다양한데, 평균적으로 약 1.5%를 차지하고 있다. 이것은 외부 온도의 변화가 전체 전력 사용량에 큰 영향을 미치지 않음을 의미한다. 둘째, 전력수요 예측 오차 중 MRPE는 약 3%~8%(평균 5%)이며, RMSE는 약 4~27kWh(평균 14kWh)로, 예측 오차가 전력 사용량의 약 5% 정도를 차지하고 있음을 의미한다. 셋째, 온도를 고려하지 않은 Model (I)과 온도를 고려한 Model (II)에 대한 예측 정확도를 비교한 결과, 표본 9를 제외한 모든 표본에 대하여 Model (II)의 예측율이 높음을 알 수 있다. 이것은 온도 변화에 따른 전력 사용량을 추가로 고려한 혼합 모형의 예측 성능이 온도를 고려하지 않은 모형보다 나음을 의미한다.



<Figure 5> Plots of MRPE and RMSE Versus α for Two Models

또한 지수평활 모형의 매개변수인 α 에 따른 성능을 비교하기 위하여 $0.1 \leq \alpha \leq 0.9$ 에 대하여 두 모형의 예측 오차를 계산하였으며, 그 결과를 <Figure 5>에 나타내었다. <Figure 5>를 통해 볼 때 모든 α 값에 대하여 Model (II)의 예측 오차가 더 작음을 알 수 있다. 이것은 온도를 고려한 모형이 매개변수 값에 무관하게 상대적으로 더 향상된 성능을 나타냄을 의미한다.

논문의 제 2.3절에 언급했듯이 외부 온도에 의한 전력 사용량을 분석하기 위해서는 온도의 변화에 따른 전력량 효과를 포함할 가능성이 있는 모형을 배제해야 한다. 비선형 모형이 이에 해당되는데, 예를 들어 SVR 모형은 시계열 데이터를 최적으로 적합시키는 매개변수를 사용할 수 있기 때문에, 모형에 이미 온도의 변화에 따른 전력량 효과를 함축적으로 포함하고 있다. 이를 구체적으로 살펴보기 위하여 SVR 모형과 SVR에 전력 사용량 민감도를 추가한 모형인 SVR II 등 두 모형을 사용하여 예측 오차를 분석하였다. SVR 모형은 정규 커널(radial basis function)을 사용하였고, 매개변수는 비용 함수 $C=8.0$ 그리고 커널 매개변수 $\gamma=5.0$ 을 사용하였다. 또한 <Table 3>의 경우와 동일한 데이터를 사용하였으며, 그 결과를 <Table 4>에 나타냈다.

<Table 4>에서 볼 수 있듯이 모든 표본에 대하여 SVR 모형이 SVR II 모형에 비하여 예측 오차가 작음을 알 수 있다. 이러한 결과는 제 2.3절에서 언급했듯이 SVR 모형 자체가 이미 온도에 따른 전력 사용량 효과를 포함하기 때문임을 입증하며, 따라서 SVR에 온도 영향이 포함된 SVR II는 과적합(over-fitted)된 모형으로 간주할 수 있다.

또한 <Table 3>과 <Table 4>를 비교해 보면 SVR 모형이 지수평활 모형에 온도 효과를 추가한 Model (II)보다 전력수요 예측에 보다 우수한 성능을 나타냄을 알 수 있다.

<Table 4> Results of Experiments by Using SVR Together with SVR and $\Delta p(t)$, Denoted as SVR II. SD Stands for the Standard Deviation

Sample	MRPE		RMSE	
	SVR	SVR II	SVR	SVR II
1	0.85	1.36	2.99	6.05
2	1.06	2.10	4.26	8.22
3	0.53	0.70	2.37	3.24
4	0.54	1.50	2.02	5.85
5	1.23	1.74	5.07	7.96
6	1.39	2.83	4.01	8.22
7	1.5	1.72	4.37	5.06
8	1.45	1.80	4.33	5.44
9	0.62	5.36	1.00	7.47
10	1.84	4.91	6.65	16.3
mean	1.10	2.40	3.71	7.39
SD	0.46	1.54	1.64	3.54

그러나 SVR 모형은 온도가 전력 사용량에 미치는 효과를 파악하기는 힘들기 때문에, 온도가 전력수요 예측에 미치는 효과 등을 분석하는 모형으로는 적절하지 않음을 알 수 있다. 따라서 SVR과 Model (II)는 상호보완적 모형이라 할 수 있다.

4. 요약 및 결론

본 논문에서는 외부 기온이 전력 사용량에 미치는 영향을 분석한 결과를 바탕으로 전력수요 예측을 위한 모형과 그 방법에 대한 연구를 수행하였다. 이를 위하여 A 테크노 파크에서 201*년에 측정된 전력 사용량 시계열 데이터와 기상청에서 매 시간 제공하는 외부 기온 시계열 데이터를 사용하였다.

온도 변화를 고려한 전력 수요 예측 모형을 설정하기 위하여 온도와 전력 사용량 시계열 데이터 사이의 상관관계를 분석하였다. 분석 결과 두 시계열 데이터 사이에는 미미한 음의 상관관계가 있었고, 월별 상관관계 분석에서도 거의 유사한 결과를 얻었다. 이러한 결과는 외부 온도가 올라감에 따라 전력 사용량은 감소하나 유의한 정도는 아님을 의미한다. 따라서 온도 변화를 고려한 전력 수요 예측은 시계열적 방법 외에 추가적인 방법이 필요함을 알 수 있었다.

주어진 시간대에 대하여 온도와 전력 사용량 사이의 관계를 분석한 결과, 온도와 전력 사용량 사이에는 대략적으로 선형적인 관계가 있으며 임계 온도가 존재함을 알 수 있었다. 임계 온도보다 낮은 온도에서는 온도가 감소함에 따라 전력 사용량은 증가하였고, 높은 온도에서는 전력 사용량은 감소하였다. 이러한 결과를 회귀 분석을 통하여 정량적으로 전력 수요 예측에 적용하였다.

전력수요 예측을 위한 모형은 온도에 따른 효과를 분석하기 위해 지수평활 모형을 사용하였으며, 온도와 전력 사용량의 관계를 지수평활 모형에 추가적으로 고려하여 새로운 모형을 제안하였다. 두 모형의 성능을 비교하기 위하여 MRPE와 RMSE 등 예측 오차 척도를 사용하였고, 실험을 통하여 온도에 따른 전력 사용량의 의존도를 포함한 모형이 보다 효율적임을 입증하였다.

본 연구에서는 A 테크노파크의 전력량을 사용했는데, 연구 결과 및 동기의 범용성을 위해서는 향후 가정용 전력 사용량에 대한 연구를 추가로 수행할 필요가 있다. 본 연구의 결과는 향후 전력 에너지 시계열 데이터 예측 모형과 전력 소모량 상황을 비교 검토하고, 날씨 등 다양한 정보를 포함한 예측 모형의 신뢰성 검증을 통해 보다 효율적인 에너지 관리를 위한 전력 에너지의 부하 조절과 전력과 가스의 혼용을 위한 의사결정 시스템[9]에 적용할 수 있을 것으로 판단된다.

Acknowledgement

This work was supported by the Power Generation and Electricity Delivery Core Technology Program of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) granted financial resource from the Ministry of Trade, Industry and Energy, Republic of Korea (No. 2016-2020108170).

References

- [1] Ahn, B., Choi, H., Lee, H., Regional long-term/mid-term load forecasting using SARIMA in South Korea, *Journal of the Korea Academia-Industrial Cooperation Society*, 2015, Vol. 16, No. 12, pp. 8576-8584.
- [2] Amjady, N., Short-term hourly load forecasting using time-series modeling with peak load estimation capability, *IEEE Trans. on Power Systems*, 2002, Vol. 16, No. 4, pp. 798-805.
- [3] Box, G., Jenkins, G., and Reinsel, G., *Time Series Analysis : Forecasting and Control*, 4th Edition, Wiley, Hoboken, New Jersey, 2008.
- [4] Elman, J., Finding Structure in Time, *Cognitive Science*, 1990, Vol. 14, pp. 179-211.
- [5] Heideman, M., Johnson, D., and Burrus, C., Gauss and the history of the fast Fourier transform, *IEEE ASSP Magazine*, 1984, Vol. 1, No. 4, pp. 14-21.
- [6] Hochreiter, S. and Schmidhuber, J., Long short-term memory, *Neural Computation*, 1997, Vol. 9, pp. 1735-1780.
- [7] Kang et al., BIM-based Data Mining Model for Effective Energy Management, *Journal of the Korea Academia-Industrial Cooperation Society*, 2015, Vol. 16, pp. 5591-5599.
- [8] Lee et al., A study on the estimation and prediction of electricity peak, *Korean Energy Economic Review*, 2010, Vol. 9, pp. 83-99.
- [9] Lee, C., Song, G., and Kim, J., A Study on the Prediction of Power Consumption in the Air-Conditioning System by Using the Gaussian Process, *Journal of Society of Korea Industrial and Systems Engineering*, 2016, Vol. 39, No. 1, pp. 64-72.
- [10] Lee, C., Song, G., and Kim, J., Correlation Analyses of the Temperature Time Series Data from the Heat Box for Energy Modeling in the Automobile Drying Process, *Journal of Society of Korea Industrial and Systems Engineering*, 2014, Vol. 37, pp. 27-34.
- [11] Lee, H. and Shin, H., Electricity Demand Forecasting based on Support Vector Regression, *IE Interfaces*, 2011, Vol. 24, pp. 351-361.
- [12] Nam, B., Song, K., Kim, K., and Cha, J., The spatial electric load forecasting algorithm using the multiple regression analysis method, *Journal of the Korean Institute of Illuminating and Electrical Installation Engineers*, 2008, Vol. 22, No. 2, pp. 63-70.
- [13] Oh, D., The present and future of big data in the electric industry, *Journal of the Electric World*, 2014, pp. 18-23.
- [14] Papalexopoulos, A. and Hesterberg, T., A Regression-Based Approach to Short-Term System Load Forecasting, *IEEE Trans. on Power Systems*, 1990, Vol. 4, No. 4, pp. 1535-1547.
- [15] Park, Y. and Wang, B., Neuro-fuzzy model based electrical load forecasting system : Hourly, daily, and weekly forecasting, *Journal of Korean Institute of Intelligent Systems*, 2004, Vol. 14, No. 5, pp. 533-538.
- [16] R : Holt-Winters Filtering, stat.ethz.ch. Retrieved 2018-01-05.
- [17] Saini, L. and Soni, M., Artificial neural network-based peak load forecasting using conjugate gradient methods, *IEEE Trans. on Power Systems*, 2002, Vol. 17, No. 3, pp. 907-912.
- [18] Schmidhuber, J., Deep Learning in Neural Networks : An Overview, *Neural Networks*, 2015, Vol. 61, pp. 85-117.
- [19] Senjyu et al., One-Hour-Ahead Load Forecasting Using Neural Network, *IEEE Trans. on Power Systems*, 2002, Vol. 17, No. 1, pp. 113-118.
- [20] Shin et al., A volatility analysis of Korean energy consumption, *Economic Study*, 2015, Vol. 63, pp. 71-119.
- [21] Shmueli, G. and Lichtendahl Jr., C., *Practical Time Series Forecasting with R : A Hands-On Guide*, 2nd Edition, Axelrod schnell publishers, 2018.
- [22] Sohn, K., Kim, S., and Shon, E., Fuzzy time series models with triangular fuzzy numbers as parameters, *Journal of Korean Data Analysis Society*, 2001, Vol. 3, No. 2, pp. 149-162.
- [23] Song, J., Seo, S., Yun, S., Kim, Y., and Cho, C., A study on the energy profile analysis and the forecasting method of the retail shop, in *the Proceedings of Korean Communication Society of 2016*, pp. 1117-1118.

ORCID

Jin Ho Kim | <http://orcid.org/0000-0003-4243-7318>
 Chang-Yong Lee | <http://orcid.org/0000-0003-1778-6532>