

Nonlinear impact of temperature change on electricity demand: estimation and prediction using partial linear model

Jiwon Park^a · Byeongseon Seo^{b,1}

^aKorea Institute of Industrial Economics and Trade;

^bDepartment of Food and Resource Economics, Korea University

(Received June 26, 2019; Revised August 25, 2019; Accepted September 2, 2019)

Abstract

The influence of temperature on electricity demand is increasing due to extreme weather and climate change, and the climate impacts involves nonlinearity, asymmetry and complexity. Considering changes in government energy policy and the development of the fourth industrial revolution, it is important to assess the climate effect more accurately for stable management of electricity supply and demand. This study aims to analyze the effect of temperature change on electricity demand using the partial linear model. The main results obtained using the time-unit high frequency data for meteorological variables and electricity consumption are as follows. Estimation results show that the relationship between temperature change and electricity demand involves complexity, nonlinearity and asymmetry, which reflects the nonlinear effect of extreme weather. The prediction accuracy of in-sample and out-of-sample electricity forecasting using the partial linear model evidences better predictive accuracy than the conventional model based on the heating and cooling degree days. Diebold-Mariano test confirms significance of the predictive accuracy of the partial linear model.

Keywords: electricity, nonlinear impact, partial linear model, temperature

1. 서론

최근 빈번하게 발생하는 이상기온으로 인하여 전력 소비 변동성이 증가하고 있으며 전력수요에 대한 예측에 있어서 기후변화와 기온의 영향이 커지고 있다. 이상기온에 의한 기온의 반응은 이전과 달리 복잡한 비선형성을 나타내고 있다. 이상기온 영역에서는 전력수요는 기온에 대하여 급격하게 증가하는 반응으로 불특성을 나타내고 있으며 임계기온을 중심으로 중간영역에서는 전력수요는 평활한 반응을 보인다. 정부 에너지 정책과 4차 산업혁명의 전개에 따라 기온 효과를 보다 정확하게 추정하고 예측하는 것은 안정적 전력수급 관리를 위하여 중요한 과제이다. 본 연구는 기온변화가 전력수요에 미치는 비선형적·비대칭적 영향에 대하여 부분선형모형(partial linear model)을 이용하여 분석하고자 한다.

기상청 (2017)의 이상기후 보고서에 따르면 지구온난화의 영향으로 전 지구 평균기온이 계속 상승하고 있으며 우리나라 역시 꾸준히 평균기온이 상승하고 있다. 2016년 전 지구 평균기온은 20세기 평균기온

¹Corresponding author: Department of Food and Resource Economics, Korea University, 145 Anam-Ro, Sungbuk-Gu, Seoul 02841, Korea. E-mail: seomatteo@korea.ac.kr



Figure 1.1. Economic growth and electricity demand.

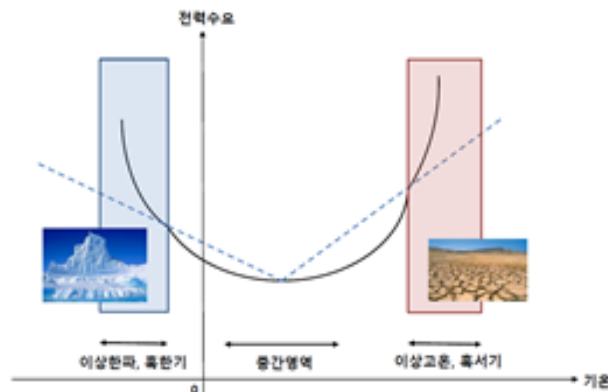


Figure 1.2. Nonlinear relationship between temperature and electricity demand.

보다 0.94°C 만큼 더 높은 14.94°C 를 기록하며 1880년 이후 전 지구 평균기온이 가장 높은 해가 될 것으로 예상했다. 한국의 2016년 평균기온은 평년보다 1.1°C 더 높은 13.6°C 로 기록되었다. 이상기온과 예상치 못한 폭염으로 전력수요가 급등하여 2016년 8월 8일 최대전력수요는 8,370만 KW로 역대 최대치를 기록하면서 전력 예비율은 7% 수준으로 하락하였다. Figure 1.1은 경제성장률과 전력수요증가율에 대한 시간 변화를 보여준다. 2000년대 초반 이전에는 전력수요증가율은 경제성장률과 동행하는 모습을 보이나, 그 이후 2008년까지 그리고 국제금융위기 이후 동조성이 약해지고 있다. 이러한 현상은 기후변화와 지구온난화로 인한 이상기후에 의한 전력수요에 미치는 영향이 더 커지고 있음을 시사한다.

기온과 전력수요의 관계는 기본적인 특성으로 임계기온을 중심으로 기온의 변화에 따라 전력 소비가 좌우 비대칭적으로 증가하는 비대칭성을 보인다. 비대칭성과 함께 최근 이상기온에 의한 기온의 효과는 복잡한 비선형성을 나타내고 있다. 특히, 폭염과 한파와 같은 이상기온 영역에서는 전력수요는 급격하게 증가하여 블록성을 나타낸다. 그리고 임계기온을 중심으로 중간영역에서의 전력수요는 평활한 반응을 보인다. 따라서 기온과 전력수요의 관계는 Figure 1.2와 같이 블록성과 평활성을 특징으로 하는 비대칭 U자 형태에 가깝다. 이는 기존의 연구에서 가정하는 이차함수 형태, CDD-HDD 모형과 같은 비대칭 V자 형태로는 설명이 어렵다. 이상기온 영역의 블록성은 지구온난화로 인해 여름철 아열대성의 이상고온이 지속화되거나 겨울철에 삼한사온 현상과는 다르게 이상한파가 계속되어 전력수요가 급격하

게 증가하는 현상을 반영한다. 그리고 중간영역에서는 단일시스템과 건축 구조 등 기술적 발전으로 기온에 대한 의존성이 감소하여 전력수요는 평활한 형태의 반응을 나타낸다.

기온의 제곱항을 사용하거나 냉난방도일에 기초한 기존의 전력수요모형은 복잡한 비선형성과 비대칭성을 설명하기 어렵다. Figure 1.2와 같이 이상기온 영역에서 전력수요는 불록성을 나타내고 중간영역에서는 평활성을 나타낸다면 점선으로 나타낸 냉난방도일에 기초한 모형은 선형적 증가를 가정하므로 이상기온과 기술적 발전에 의한 기온의 효과를 반영하지 못한다. 마찬가지로 기온의 제곱항을 이용한 모형은 대칭적 U자 형태를 가정하므로 기온 효과에서 나타나는 비대칭성을 설명하지 못한다. 본 연구에서는 기온변화와 전력수요의 비선형적 관계를 설명하기 위하여 부분선형모형을 사용한다. 전력수요와 관련한 다른 변수들에 대하여는 기존 모형과 같은 선형적 관계를 가정한다. 따라서 ‘차원의 저주’에 의한 문제를 해결하면서 예측력 향상을 위하여 모수 절약적 설정(parsimonious specification)이 중요하므로 기온 효과에 비선형성을 가정한 준모수적 부분선형모형을 사용하고자 한다.

부분선형모형을 추정하기 위하여 Robinson의 double residual (DR) 준모수 추정과 스플라인을 이용한 추정을 적용한다. Engle 등 (1986)은 기온과 미국의 가정용 전력수요의 관계에 대하여 스플라인 추정을 적용하였다. 국내연구로 Yu와 Hwang (1997)는 Robinson’s double residual 준모수 추정을 사용하였다. 최근 빈번하게 발생하는 이상기온과 기후변화의 영향을 측정하기 위해서는 최근 자료와 고주기 자료를 사용해야 한다. 본 연구는 시간 단위 고주기 자료를 사용한 점에서 이들 연구와 차별성을 갖는다.

Moral-Carcedo와 Vicens-Otero (2005), Bessec와 Fouquau (2008), Shin과 Cho (2014)은 smooth transition regression, threshold regression과 switching regression 방법으로 기온의 비선형적 효과를 설명하였다. 이들 연구에서는 기온과 전력수요의 비선형적 관계를 분석하기 위하여 모수적 방법을 적용하였으며 준모수적 접근을 사용한 본 연구와 보완성을 갖는다. 정부 에너지 정책은 신재생 3020 이행, 친환경 에너지 세계개편, 에너지 소비 산업구조 개편, 4차 산업혁명 대비 신에너지 생태계 조성으로 요약된다. 화석연료에 대한 의존성을 줄이면서 신재생에너지 방향으로 이행하는 과정에서 기온변화와 전력수요의 관계는 더욱 중요하다. 화석연료와 원전에 의존하던 시기에는 비교적 안정적인 전력공급이 가능했으나 새로운 에너지 정책에서 이상기후와 기온 효과의 중요성은 더욱 커질 것이다. 따라서 시간 단위 고주기 자료를 사용하여 기온 효과를 측정하는 점에서 본 연구는 전력수요에 대한 예측과 에너지 정책을 수립하는데 유용한 결과로 활용될 것으로 기대한다.

본 논문은 다음과 같이 구성된다. 제2절에서는 기온변화와 전력수요의 관계를 분석하기 위한 부분선형 모형과 분석방법에 대하여 살펴본다. 제3절에서는 전력 소비와 기상변수에 대한 시간 단위 고주기 자료에서 주요 특성을 얻는다. 제4절에서는 부분선형모형을 추정하고 이를 이용한 전력수요를 예측하여 평가한 결과를 보인다. 제5절에서는 주요 결과를 요약하고 시사점을 찾는다.

2. 분석 모형

2.1. 부분선형모형

본 연구에서 전력수요모형으로 적용하는 부분선형모형은 다음과 같다.

$$y_t = m(z_t) + X_t' \beta + \epsilon_t, \quad t = 1, 2, \dots, n, \quad (2.1)$$

여기서 y_t 는 로그 변환된 전력수요, z_t 는 기온, 그리고 X_t 는 전력수요와 관련된 기타 설명변수이다. 그리고 $m(z_t)$ 는 미지의 함수이며 오차항은 $E(\epsilon_t | X_t, z_t) = 0$ 을 만족한다.

식 (2.1)에서 전력수요는 설명변수 X_t 와 선형 관계를 갖고 기온변수 z_t 와는 비선형적 관계를 갖는다. 따라서 부분선형모형은 기온 효과의 비선형성을 설명하면서 비교적 많은 설명변수를 모형에 포함하고

있으므로 비모수적 모형을 적용할 경우 발생하는 차원의 저주를 피할 수 있다.

부분선형모형에 대한 추정, 분포이론 및 응용연구는 Robinson (1988), Speckman (1988), Eilers와 Marx (1996), Fan과 Li (1999), Yatchew (2003), Ruppert 등 (2003)에서 발전되어 왔다. 부분선형 모형의 추정은 미지의 함수 추정에 적용하는 방식에 따라, 커널을 이용한 Robinson's double residual 준모수 추정과 스플라인을 이용한 추정으로 이루어진다.

2.1.1. Robinson's double residual 추정

- 1단계: 우선 식 (2.1)을 z_t 에 대한 조건부 기대를 취하면 다음 식을 도출할 수 있다.

$$E(y_t|z_t) = m(z_t) + E(X_t|z_t)' \beta, \quad t = 1, 2, \dots, n. \quad (2.2)$$

식 (2.1)과 식 (2.2)를 결합하여 비모수적 부분을 나타내는 함수 $m(z_t)$ 을 제거한다.

$$y_t - E(y_t|z_t) = [X_t - E(X_t|z_t)]' \beta + \epsilon_t, \quad t = 1, 2, \dots, n, \quad (2.3)$$

여기서 $E(y_t|z_t)$ 와 $E(X_t|z_t)$ 는 커널을 이용한 비모수적 추정을 통해 $E_N(y_t|z_t)$ 와 $E_N(X_t|z_t)$ 을 얻고 식 (2.3)에 대입하면 다음 식을 얻는다.

$$y_t - E_N(y_t|z_t) \simeq [X_t - E_N(X_t|z_t)]' \beta + \epsilon_t, \quad t = 1, 2, \dots, n, \quad (2.4)$$

여기서 $E_N(y_t|z_t) = \sum_{s,s \neq t}^n K_h(z_s - z_t) y_t / \sum_{s,s \neq t}^n K_h(z_s - z_t)$, $E_N(X_t|z_t) = \sum_{s,s \neq t}^n K_h(z_s - z_t) X_t / \sum_{s,s \neq t}^n K_h(z_s - z_t)$. 그리고 K_h 는 커널이며 h 는 구간넓이(bandwidth) 모수이다. 적절한 구간 넓이 모수를 선택하기 위하여 교차검증(cross validation; CV) 방법을 적용한다.

- 2단계: 식 (2.4)를 최소제곱법으로 추정하면 다음과 같은 추정량을 얻는다. 여기서 $R_{1t} = X_t - E_N(X_t|z_t)$, $R_{2t} = y_t - E_N(y_t|z_t)$ 이다.

$$\hat{\beta} = \left[\sum_{t=1}^n R_{1t} R_{1t}' \right]^{-1} \sum_{t=1}^n R_{1t} R_{2t}. \quad (2.5)$$

- 3단계: 기온(z_t)과 전력수요(y_t)의 관계를 나타내는 $\hat{m}_N(z_t)$ 의 추정은 다음과 같다.

$$y_t^* = m(z_t) + \epsilon_t^*, \quad (2.6)$$

여기서 $y_t^* = y_t - X_t' \hat{\beta}$ 이다. 제2단계에서 얻은 $\hat{\beta}$ 을 이용하여 y_t^* 을 구하고, 커널 추정 방법으로 다음과 같은 $\hat{m}_N(z_t)$ 을 얻는다.

$$\hat{m}_N(z_t) = \frac{\sum_{s,s \neq t}^n K_h(z_s - z_t) y_t^*}{\sum_{s,s \neq t}^n K_h(z_s - z_t)}, \quad (2.7)$$

여기서 K_h 는 커널이며 h 는 구간넓이 모수이다.

2.1.2. 스플라인 추정 스플라인 추정은 식 (2.1)의 미지의 함수 $m(z_t)$ 을 정의된 구간에서 M 개의 부분 구간(knots)으로 나누고, 각 부분 구간에서의 기저 함수를 연결하여 부드러운 곡선으로 근사화한다. 따라서 스플라인 추정은 다음과 같이 이루어진다.

$$y_t = X_t' \beta + m(z_t) + \epsilon_t, \quad t = 1, 2, \dots, n, \quad j = 1, 2, \dots, M, \quad (2.8)$$

여기서 $m(z_t) = \rho_0 z_t + \sum_{j=1}^M \rho_j |z_t - \kappa_j|_+, |z_t - \kappa_j|_+ = \max(z_t - \kappa_j, 0)$.

식 (2.8)에서 스플라인의 차수는 1로 가정하였다. 실제 추정에서 다른 차수에 대하여 추정한 결과 우도 함수(restricted log likelihood)가 작아지는 결과를 얻었다. 식 (2.8)의 모수를 추정하기 위해, 벌점화 최소제곱법을 이용하여 아래의 식을 최소로 하는 모수 $\hat{\theta} = (\hat{\beta}', \hat{\rho}')$ 을 추정한다.

$$\min_{\theta} \sum_{t=1}^n \left(y_t - X_t' \beta - Z_t' \rho \right)^2 + \lambda \theta' D \theta, \quad (2.9)$$

여기서 $Z_t = (z_t, |z_t - \kappa_1|_+, |z_t - \kappa_2|_+, \dots, |z_t - \kappa_M|_+)', \beta = (\beta_1, \beta_2, \dots, \beta_k)', \rho = (\rho_0, \rho_1, \rho_2, \dots, \rho_M)', D = \begin{pmatrix} 0_k & 0 \\ 0 & I_{M+1} \end{pmatrix}$.

평활 모수 λ 는 제약에 부과되는 계수이며 스플라인 곡선의 평활도를 결정하고 추정량의 분산과 편의(bias)를 조절하는 역할을 한다. 만약 λ 가 0에 가까울수록 추정된 회귀함수는 평활성이 낮은 함수 형태로 추정되는 반면, λ 가 커질수록 평활성이 증가하여 분산은 작아지고 편의는 커진다. 따라서 평활 모수 λ 에 따라 모수에 대한 추정이 달라지기 때문에 적절한 평활 모수를 선택하기 위하여 교차검증 방법을 적용한다. 교차검증, 일반화 교차검증(generalized CV; GCV), Mallows의 C_p , Akaike information criterion (AIC), Bayesian information criterion (BIC) 등을 적용할 수 있다. 평활 모수 λ 의 선택과 함께 매듭의 개수(M)와 각 매듭의 위치를 선택하여야 한다. 본 연구에서는 Ruppert 등 (2003)의 방식에 따라서 매듭의 수를 결정하고 매듭의 위치는 $\kappa_j = (j + 1)/(M + 2)$ 번째 분위수로 선택하였다.

2.2. 모수적 모형

전력수요모형으로 분석에서 사용된 모수적 모형은 이차함수 모형과 냉난방도일 모형이다. 기온의 제곱항을 사용하는 이차함수모형은 $m(z_t) = \mu + \gamma_1 z_t + \gamma_2 z_t^2$ 와 같다. 냉난방도일을 사용하는 모수적 모형의 경우에는 $m(z_t) = \mu + \gamma_1 z_{1t} + \gamma_2 z_{2t}$ 가 된다. 여기서 z_{1t} 와 z_{2t} 는 각각 HDD_t(난방도일(heating degree days; HDD)), CDD_t(냉방도일(cooling degree days; CDD))을 의미한다. 기온과 함께 전력수요와 관련된 사회경제적 요인과 습도, 시차 변수를 추가하여 다음과 같은 모수적 모형으로 한다.

$$\text{Quadratic model : } y_t = \mu + \gamma_1 z_t + \gamma_2 z_t^2 + X_t' \beta + \epsilon_t, \quad (2.10)$$

$$\text{CDD/HDD model : } y_t = \mu + \gamma_1 z_{1t} + \gamma_2 z_{2t} + X_t' \beta + \epsilon_t, \quad (2.11)$$

여기서 $z_{1t} = \max(z_t - 18, 0)$, $z_{2t} = \max(18 - z_t, 0)$, $\epsilon_t \sim^{iid}(0, \sigma_\epsilon^2)$ 이다. 전력은 일반적으로 내구재 또는 생산을 위해 가동되는 공장기기를 통해 효율이 창출되는 파생수요이다. 전력수요는 직접 소비하는 상품이 아니라 생산과 소비과정에서 중간재로 사용된다. 따라서 전력수요는 지속성과 주기성을 갖고 이는 추세, 주기성, 계절성, 계절상관의 형태로 나타난다. 즉, 전력수요는 경제 규모의 증가와 인구증가에 따라 추세와 함께 주기성을 갖는다. 그리고 우리나라와 같이 사계절이 존재하는 경우에는 계절성이 뚜렷하게 나타난다. 추세, 주기, 계절성에 의한 전력수요의 변화는 사회경제적 요인에 의하여 비롯되는 특성으로 지정하였다. 전력수요의 추세효과를 반영하기 위하여 연도변수를 사용한다. 그리고 월별효과와 시간별 효과를 나타내기 위해 월별 더미, 시간 더미를 이용하였으며, 일하는 날과 일을 하지 않는 날의 효과를 비교하기 위해 요일 더미와 공휴일 더미를 사용한다. 공휴일 더미의 경우, 일요일, 설 연휴, 추석 연휴를 제외한 공휴일로 이뤄졌다. 그리고 전력수요 데이터의 경우, 계절 상관(serial correlation)이 강하게 작용하기 때문에 시간 의존성을 확인하고자 시차 변수의 유의성에 따라 설명변수로 설정하여 분석하였다.

기상변수는 매시간 자료를 실시간으로 관측할 수 있으나, 경제적 변수는 대부분 시차를 두고 공표되며 월별, 분기별, 연도별 주기를 갖기 때문에 고주기인 기상변수와 주기가 일치하지 않는다. 본 연구에서

사용한 전력사용량과 기온변수의 주기는 시간 단위이다. 경제적 요인으로서 생산과 소득변수를 수요에 측에 고려해야 하지만 시간 단위 주기를 갖는 경제적 변수를 찾을 수 없으므로 추세효과를 나타내는 연도변수로써 경제적 요인을 반영하고자 하였다. 마지막으로 시간 단위 평균기온(°C)과 습도(%)를 기상 요인으로 사용하였다. 그리고 시간 단위 평균기온을 이용하여 난방도일과 냉방도일을 산출하였다.

$$\text{HDD}_t = \max(T_0 - z_t, 0), \quad \text{CDD}_t = \max(z_t - T_0, 0),$$

여기서 z_t 는 시간별 평균기온이며, T_0 는 기준온도로 우리나라의 경우 18°C로 정의된다. 난방도일이란 기온이 기준온도보다 낮을 때의 기준온도와 기온의 차이를 의미하며, 반대로 냉방도일은 기온이 기준온도보다 높을 경우 기온과 기준온도의 차이를 나타낸다.

2.3. 모형의 적합성 및 예측력 평가

전력수요에 대한 부분선형모형에 대하여 추정이 이루어지면 이차함수 모형과 CDD/HDD 모형을 모수적 모형으로 고려하여 추정 결과의 적합성과 예측 결과에 대한 정확성을 비교 평가한다. Härdle과 Mammen (1993)은 모수적 모형에 의한 추정과 비모수적 추정을 비교하여 모수적 모형에 대한 적합성을 검정하는 통계량을 제시하였다.

$$T_n = nh^{\frac{1}{2}} \int [\hat{\mu}_h(z) - \Lambda_{h,n} \hat{\mu}_\theta(z)]^2 \pi(z) dz,$$

여기서 $\hat{\mu}_h(z)$ 는 비모수적 커널 추정이며 $\hat{\mu}_\theta(z)$ 는 모수적 모형의 추정이다. 그리고 h 는 구간넓이 모수, $\Lambda_{h,n} \hat{\mu}_\theta(z) = \sum_{t=1}^n K_h(z - z_t) \mu(z) / \sum_{t=1}^n K_h(z - z_t)$, $K_h(\cdot)$ 는 커널 함수, $\pi(z)$ 는 가중 함수이다. Härdle과 Mammen (1993) 검정을 위한 임계값과 p -값은 wild 부트스트래핑 방식으로 얻는다. Härdle과 Mammen (1993) 검정에서 구간넓이 모수에 대한 가정은 대부분 통용되는 방법으로 하면 성립한다. 따라서 mean integrated squared error (MISE)를 최소화하는 plug-in 선택과 교차검증이 이 가정을 만족하는데 부트스트래핑 검정에 시간이 소요되므로 plug-in 선택을 적용하였다.

다음은 부분선형모형을 추정한 결과를 사용하여 전력수요에 대한 예측을 얻고 이를 모수적 모형에 의한 예측과 비교하여 정확성을 평가한다. 부분선형모형과 모수적 방법에 의하여 표본내 예측(in-sample forecast)과 표본외 예측(out-of-sample forecast)을 얻고 예측력을 비교하기 위하여 root mean square error (RMSE)와 mean absolute error (MAE), mean absolute percentage error (MAPE)를 구하여 예측력을 비교한다.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n e_t^2}, \quad \text{MAE} = \frac{1}{n} \sum_{t=1}^n |e_t|, \quad \text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right|,$$

여기서 e_t 는 예측오차, y_t 는 실제 관측치, \hat{y}_t 는 예측치이다. 그리고 예측력 평가를 위하여 Diebold-Mariano (DM) (1995) 검정을 한다. 두 개의 예측을 비교하여 손실의 차이는 $d_t = L(e_t^A) - L(e_t^B)$ 와 같다. 여기서 $e_t^i = y_t - \hat{y}_t^i$ 는 각각의 모형($i = A, B$)을 통해 산출된 예측오차이다. DM 검정에 손실함수 $L(e_t)$ 는 예측오차의 제곱(squared-error loss, e_t^2) 또는 절댓값(absolute error loss, $|e_t|$)으로 하였다. DM 검정의 귀무가설과 검정통계량은 다음과 같다.

$$H_0 = E(d_t) = 0 \quad \text{against} \quad H_1 = E(d_t) \neq 0,$$

$$\text{DM} = \frac{\bar{d}}{\text{s.e.}(\bar{d})}, \quad (2.12)$$

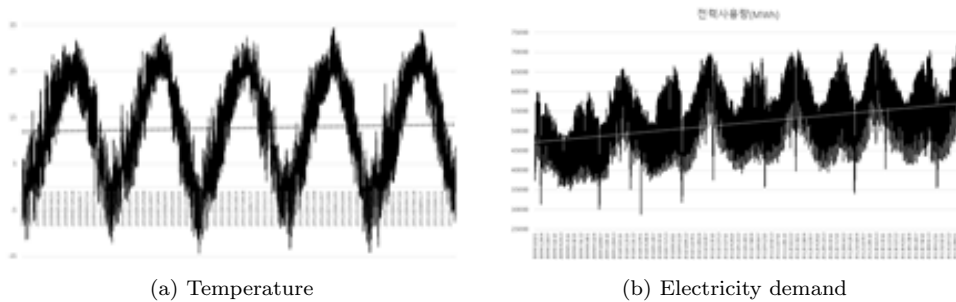


Figure 3.1. Plot of temperature and electricity demand.

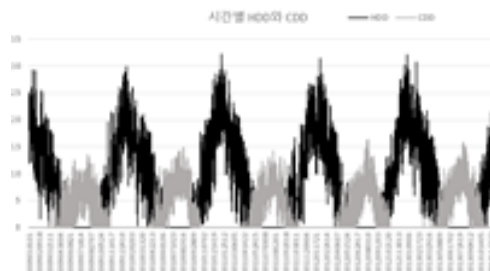


Figure 3.2. Plot of CDD/HDD. CDD = cooling degree days; HDD = heating degree days.

여기서 $\bar{d} = (1/n) \sum_{t=1}^n d_t$ 이다. 예측오차는 계열상관과 이분산성을 나타낼 수 있으므로 표준편차 $s.e(\bar{d})$ 는 Newey-West 방식을 통해 얻는다. 만약 두 개의 예측에서 얻은 손실함수를 비교하여 차이가 없다면 d_t 는 평균적으로 0이 되고, DM 검정통계량은 점근적으로 표준정규분포를 따른다.

3. 분석 자료

본 연구는 부분선형모형을 이용하여 기온과 전력수요가 나타내는 보다 복잡한 형태의 비선형·비대칭적 관계를 추정하고, 예측력을 측정하고 평가하고자 한다. 분석 자료의 표본기간은 2009년 1월 1일부터 2013년 12월 31일까지이며 시간 단위 데이터로서 총 표본의 크기는 43,824이다. 기온과 습도 등 기상변수 역시 시간 단위 데이터이다. 이들 자료는 전력거래소와 통계청 DB에서 구하였다. 전력수요에 대한 표본 내 예측의 예측기간은 2009년 1월 1일부터 2013년 12월 31일까지이고, 표본 외 예측에는 2009년부터 2012년까지 자료를 이용하여 추정하고 2013년 1월 1일부터 2013년 12월 31일까지에 대하여 예측이 이루어졌다. Figure 3.1은 시간 단위 평균기온(°C)과 시간 단위 전력소비량의 변화를 보여준다. 전력소비량의 단위는 MWh이다. 평균기온의 추세를 보면, 뚜렷한 식별이 어렵지만 평균기온이 상승하는 것으로 보인다. 기온변수와 마찬가지로 전력수요는 뚜렷한 계절성과 주기를 나타내고 시간의 흐름에 따라 증가하고 있음을 알 수 있다. Figure 3.2는 난방도일(HDD)과 냉방도일(CDD)의 시간 변화를 보여준다. 규칙적 변화가 뚜렷하게 나타나고 있으며 계절성과 주기를 반영한다.

Figure 3.3(a)는 주중과 주말의 시간별 평균전력사용량이고, Figure 3.3(b)는 공휴일과 비공휴일의 시간별 평균전력사용량이다. 근무일인 주중과 비공휴일, 그리고 비근무일인 주말과 공휴일의 시간별 평균전력사용량을 비교할 수 있다. 일반적으로 근무일에는 비근무일에 비해 더 많은 전력을 사용하는 것으로 보인다. 그리고 근무일인 주중과 비공휴일에는 오전 3시에 가장 적은 전력사용을 나타내고 오전 11시와 오후 2시에 가장 많은 전력을 사용하는 것으로 나타나며 오후 1시에 감소하고 다시 증가하는 패턴을 볼

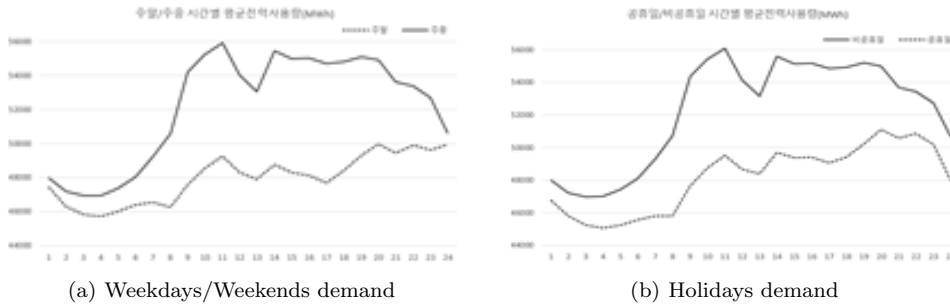


Figure 3.3. Patterns of electricity demand.

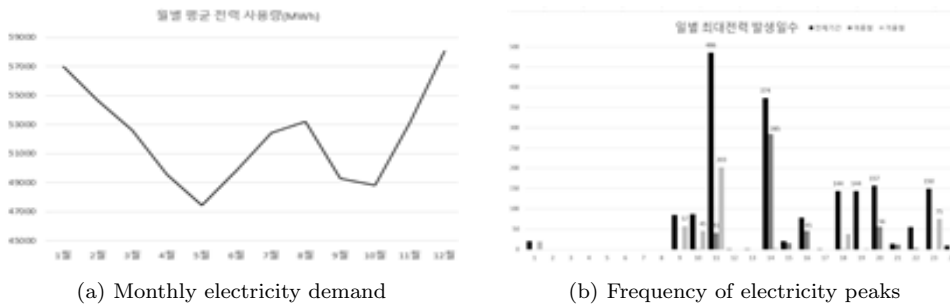


Figure 3.4. Characteristics of electricity demand.

수 있다. 비근무일인 주말과 공휴일의 경우, 오전 3-4시에 가장 적은 전력사용을 보이는 것은 비슷하지만, 일중 최대전력을 사용하는 시간대는 오후 8시에서 10시로 저녁 시간대로 나타난다.

Figure 3.4(a)의 월별 평균전력사용량을 보면 봄에서 여름으로 그리고 가을에서 겨울로 계절이 바뀌면서 전력사용량이 증가한다. 반대로 여름에서 가을로 그리고 겨울에서 봄으로 계절이 바뀔 때는 사용량이 감소한다. 전반적으로 봄인 5월에 가장 적은 전력을 사용하고, 겨울인 12월에 가장 많은 전력을 소비하기 때문에 하절기와 동절기에 가장 많은 전력을 사용하고 있음을 알 수 있다. Figure 3.4(b)에서 전체기간의 일별 시간대별 최대전력수요(표본수: 1,826)가 발생하는 시간대 분포를 살펴보면 오전 11시에 26.62%로 발생빈도가 가장 높게 나타났으며 오후 2시(20.48%), 오후 8시(8.6%) 순으로 나타난다. 또한 여름철(6-8월)의 일별 시간대별 최대전력수요(460개) 발생빈도의 경우, 오후 2시(61.96%)로 압도적으로 높은 발생빈도를 차지했다. 이어서 오후 8시(12.17%), 오후 4시(9.78%), 오전 11시(8.91%) 순으로 나타났다. 그리고 겨울철(12, 1, 2월) 일별 최대전력수요(451개)의 경우, 오전 11시(45.01%)에 몰려 있으며, 오후 11시(16.63%), 오전 9시(12.64%), 오전 10시(9.98%) 순으로 나타났다.

Table 3.1은 변수명과 변수에 대한 설명, 그리고 기초통계량을 정리하였다. 전체기간(2009-2013년)의 시간 단위 평균전력사용량은 52,174MWh으로 나타났고, 전력사용량의 최소값과 최대값은 각각 28,744MWh, 72,255MWh로 변동성이 크게 나타났다. 분석 기간의 평균기온은 12.7℃ 이고, 최소값과 최대값은 각각 영하 14.3℃, 영상 34.4℃로 기록되었다. 그리고 기준온도인 18℃를 기준으로 시간 단위 평균 난방도일(HDD)은 약 7.45이며 평균 냉방도일(CDD)은 2.15이다. 일반적으로 난방기기 사용이 많은 동절기의 전력사용량이 하절기의 전력사용량보다 많은데 이는 평균 난방도일이 냉방도일보다 높은 기상 특성을 반영하는 결과이다. 습도의 평균은 66.7%이며, 최소값과 최대값은 각각 14.1%, 95.7%로 큰 변동성을 보였다.

Table 3.1. Descriptive statistics

	Variable	Average	St. dev.	Min	Max
ElecD	전력사용량(MWh)	52,174	7,176.30	28,744.42	72,255.11
log(ElecD)	로그변환 전력	10.8530	0.1395	10.2662	11.188
Temp	기온(°C)	12.6987	10.4571	-14.3	34.4
HDD	난방도일	7.4499	8.0980	0	32.3
CDD	냉방도일	2.1486	3.4292	0	16.4
Humid	습도(%)	66.7269	17.6832	14.1	95.7

HDD = heating degree days; CDD = cooling degree days.

Table 4.1. Unit root testing

	ADF test		PP test	
	Intercept only	With trend	Intercept only	With trend
ElecD	-13.6679***	-16.1433***	-46.9943***	-48.3545***
log(ElecD)	-13.7142***	-16.2553***	-45.8620***	-46.9219***
Temp	-4.5566***	-4.5413***	-7.6238***	-7.6224***
Humid	-16.7026***	-16.7639***	-22.7931***	-22.7602***
HDD	-5.4750***	-5.4658***	-8.5298***	-8.5263***
CDD	-4.6943***	-4.6876***	-18.2622***	-18.3183***

***: 1%, **: 5%, *: 10% significant. ADF = augmented Dickey-Fuller; PP = Phillips-Perron; HDD = heating degree days; CDD = cooling degree days.

4. 주요 결과

4.1. 단위근 검정

Table 4.1은 주요 변수에 대한 단위근 검정결과이다. 단위근을 갖는 비정상적 시계열은 가성적 회귀(spurious regression)를 발생하므로 분석에 앞서 이를 검정하였다. 단위근 검정으로 augmented Dickey-Fuller (ADF) 검정과 Phillips-Perron (PP) 검정을 하였다. 단위근 검정에는 절편을 포함하였고 절편과 함께 추세를 포함한 모든 경우, 전력 소비 (로그변환 자료 포함), 기온, 습도, 난방도일, 냉방도일 주요 변수들이 단위근을 갖지 않고 정상성을 만족함을 확인하였다. ADF 검정에서 AR 적정 차수는 BIC로 선택하였다.

4.2. 기온과 전력수요의 관계

기온과 전력 소비의 기본적인 특성인 비선형성과 비대칭성은 임계기온을 기준으로 기온변화에 따라 전력수요가 좌우 기울기가 다르게 증가하는 것을 의미한다. 최근 지구온난화로 인한 이상기온은 전력수요의 변동성을 증가시키고 전력수요 예측에는 기상변수의 영향이 커지고 있다. 이상기온을 고려하여 전력수요를 분석하기 위해서는 기온과 전력 소비가 나타내는 특성을 파악하는 것이 필요하다. 기후변화에 의한 이상고온 영역에서는 전력 소비가 U자 형태로 급격하게 증가한다. 그리고 중간영역에서는 기술적 발전으로 기온변화에 대한 전력 소비의 반응은 평활한 형태이다. 분석 결과, 실제 데이터에서 나타나는 특성은 불특성과 평활성을 갖는 비대칭적 U자형 반응형태로 나타났다.

Figure 4.1은 시간별 기온과 전력수요 두 변수만을 이용하여 얻은 반응함수로 실제 관측된 기온에 따른 전력 소비의 산점도와 함께 보여주고 있다. 첫 번째 행은 모수적인 방법을 적용한 것으로 좌측은 이차함수 모형이고 우측은 지정된 기준온도 18°C를 기준으로 산출한 냉난방도일 모형이다. 이차함수모형을

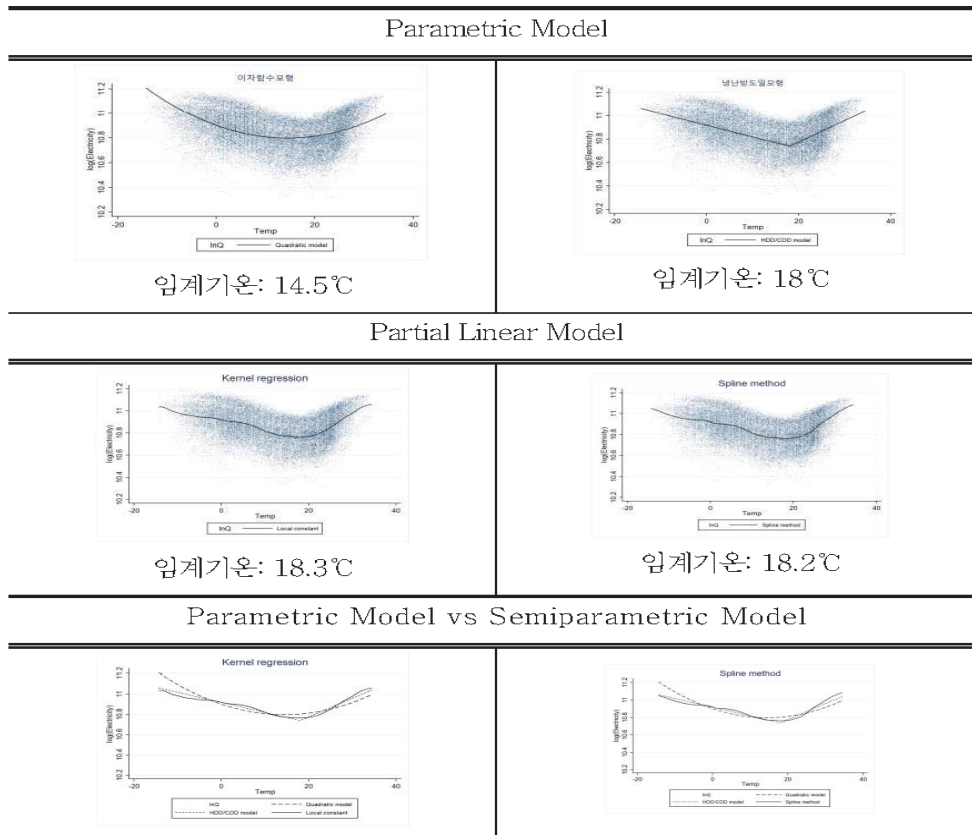


Figure 4.1. Temperature and electricity demand.

적용하여 저점을 추정하면 14.5°C이다. 이를 기준으로 좌우대칭인 U자 형태로 추정되었다. 우측의 냉난방도일 모형은 18°C를 기준으로 좌우로는 선형인 V자 형태의 반응함수가 도출되었다. 그리고 두 번째 행은 부분선형모형의 추정으로 좌측은 Robinson의 double residual 준모수 추정을 하였고 우측은 스플라인 추정으로 도출된 반응함수이다. Robinson의 준모수 추정과 스플라인 추정의 임계기온은 각각 18.3°C와 18.2°C로 추정되었다. 마지막 행은 모수적 모형과 부분선형모형을 비교한 결과이다. 부분선형 모형의 추정을 보면 기온과 전력수요의 관계는 모수적인 방법에서 가정하고 있는 U자형 또는 V자형과 같은 단순한 형태가 아닌 보다 복잡한 비선형적 비대칭적 관계를 갖는다.

계절별로 기온과 전력수요 두 변수의 관계를 추정하여 Figure 4.2를 얻었다. 여름철의 경우, 기온이 올라갈수록 전력수요가 가파르게 증가한다. 특히, 이상고온 영역에서 모수적 모형의 기본가정인 선형의 V자나 대칭적인 U자 형태와는 다른 특징을 보인다. 겨울철에는 전력 소비는 여름철과 비교하여 비교적 완만하게 증가하는 형태를 보이지만 이상 한파 영역에서는 전력수요가 급격히 증가하는 형태를 나타낸다. 봄과 가을에는 기온에 따른 전력수요의 반응은 비교적 완만한 형태이다. 이는 냉난방도일 모형에 의한 V자형과 달리 중간영역에서의 평활한 형태를 의미한다.

Figure 4.3은 여름철과 겨울철 기온과 최대전력수요의 관계를 연도별로 살펴보았다. 전반적으로 모든 계절에서 2009년에서 2013년 순으로 기온에 따른 전력수요 그래프가 위로 이동하고 있으며 특히 여름

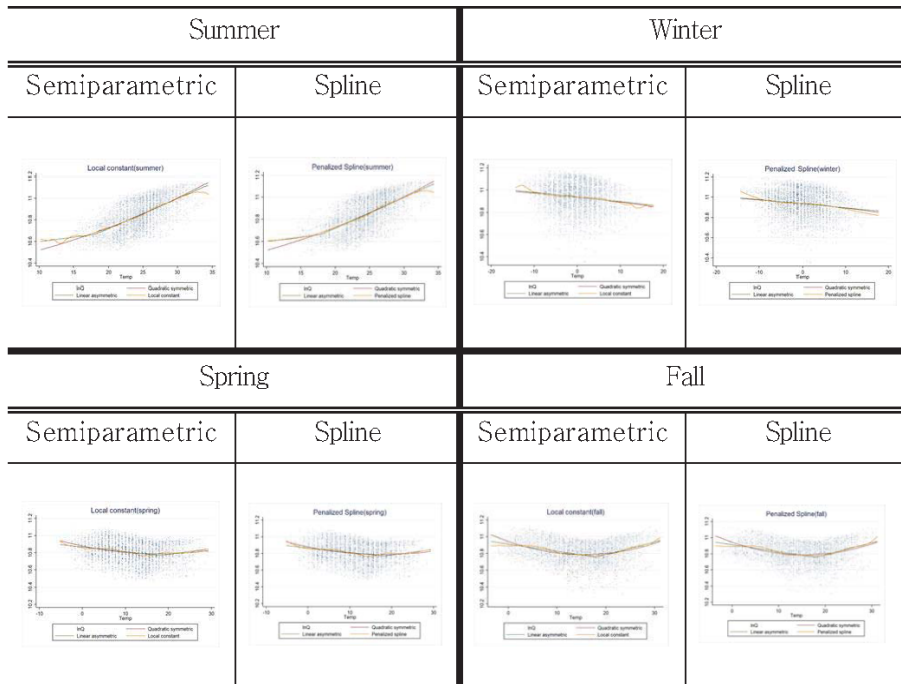


Figure 4.2. Temperature and seasonal electricity demand.

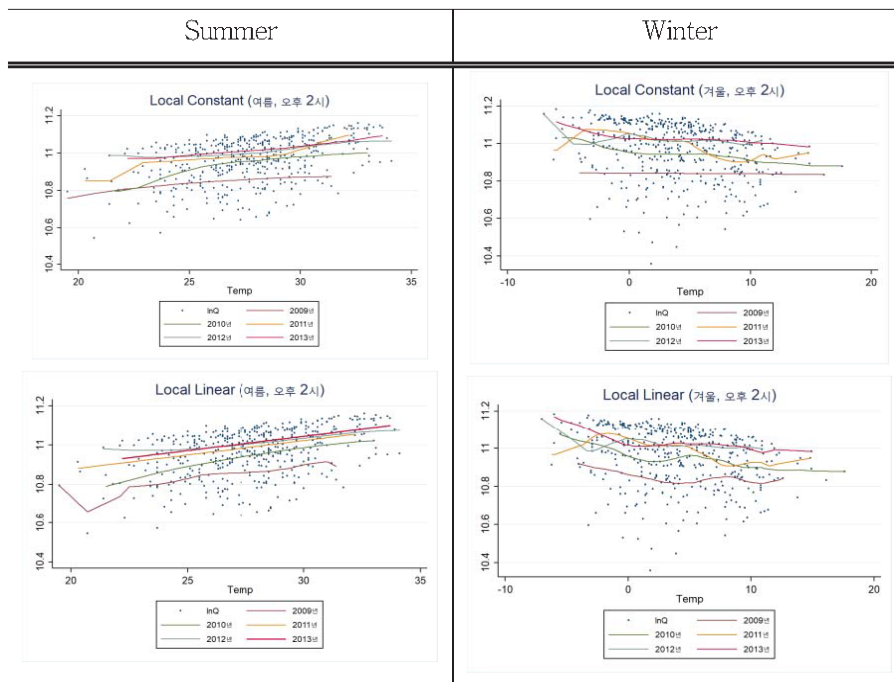


Figure 4.3. Temperature and seasonal electricity peak demand.

철에는 기온기가 급해지는 것을 더욱 명확히 볼 수 있다. 이는 겨울과 여름철의 경우, 냉난방 기구의 사용이 집중되면서 전력 소비가 증가하고 경제 성장과 함께 전력 소비가 증가한 것으로 보인다. 그리고 연도별 전력수요 그래프가 오른쪽으로 이동하는 것을 볼 수 있다. 이는 평균기온이 높아지고 있음을 의미한다.

4.3. 추정 결과

Table 4.2는 기상변수로 기온과 습도를 사용하고 사회경제적 요인으로 연도, 월별 더미, 시간 더미, 공휴일 더미 변수를 사용하여 부분선형모형과 냉난방도일 모형, 이차함수 모형을 추정한 결과이다. 전력수요의 시간 의존성을 설명하기 위하여 전기의 전력 소비 Y_{t-1} , Y_{t-2} , Y_{t-3} , Y_{t-6} , Y_{t-9} , Y_{t-12} 와 Y_{t-24} 을 모형에 포함하였다. Table 4.2에서 준모수 추정은 Robinson's double residual 추정을 의미하며 스플라인 추정은 스플라인 방식으로 부분선형모형을 추정한 결과이다. 추정된 계수는 대부분 통계적 유의성을 갖는다. 그리고 부분선형모형에 대한 준모수 추정과 스플라인 추정 결과는 크게 다르지 않았다. 모형선택에는 계열 상관 등 사후 검증을 하였고 Akaike 정보함수로 적정 모형을 선택하였다.

이차함수 모형에서 기온과 기온 제곱항의 추정치는 각각 -0.00101 , 0.00004 로 임계기온을 기준으로 블록한 U자형 함수 형태이다. 냉난방도일 모형에서 CDD, HDD의 계수는 각각 0.0028 , 0.00079 로 냉방도일(CDD)에 대한 계수가 난방도일(HDD)의 계수보다 더 크므로 기준온도인 18°C 를 기준으로 좌우 비대칭적인 V자형 형태이다. 그리고 기상변수로 기온과 함께 고려한 습도의 계수는 정(+),의 부호를 갖고 유의하게 추정되었다.

전기 종속변수의 계수들은 통계적 유의성을 나타내므로 전력수요에서 시간 의존성이 존재하고 지속성이 있음을 의미한다. 한 시간 전의 전력수요 계수가 가장 크고 시차가 증가할수록 점차 계수의 크기가 작아지는 것을 확인할 수 있다. 연도변수와 연도 제곱항의 계수 추정치는 0.007 , -0.001 로 추정되었으며 이는 전력소비량은 시간 흐름에 대하여 증가하고 있으며 증가 폭은 체감하고 있음을 의미한다. 연도변수는 2009년을 기준연도로 2009년 $t = 0$, 2010년 $t = 1$ 순으로 추세 함수로 정의하였다.

전력수요의 계절성을 반영하기 위하여 월별 더미 변수를 사용하였으며 기준 월은 12월로 하였다. 추정 계수의 값이 모두 부(-)의 부호로 나타나서 전력 소비는 12월에 가장 많고 5월에 가장 적은 전력을 사용하는 것으로 나타났다. 전력수요의 일간 주기성을 측정하고자 시간 더미 변수를 사용하였는데 기준 시간은 24시로 하였다. 추정 계수를 보면 오전 1시와 정오에 부(-)의 부호를 갖고 정오에 가장 작은 값을 갖는다. 그리고 오전 9시와 오후 2시는 정(+),의 부호 중 가장 큰 값을 갖는 것으로 추정되었다. 따라서 정오에 가장 적은 전력을 사용하고 있으며 오전 9시와 오후 2시에 가장 많은 전력을 사용한다.

요일 더미 변수와 공휴일 더미 변수의 추정 계수를 살펴보면, 근무일인 주중과 비공휴일에 비근무일인 주말과 공휴일과 비교하여 전력 소비가 더 크다. 특히 설날과 추석 연휴의 경우 전력사용량은 약 3% 적은 것으로 나타났다.

기온의 비선형적 효과를 반영한 부분선형모형과 모수적 모형에 대하여 적합성을 분석하기 위하여 Härdle과 Mammen 검정을 하였다. Table 4.3에서 기온과 전력수요의 관계를 이차함수로 하는 모수적 모형에 대하여 Härdle과 Mammen 검정 통계량은 46.50 이며 5% 임계값보다 크다. 여기서 5% 임계값과 p -값은 부트스트래핑으로 구하였다. 따라서 모수적 모형으로 고려한 이차함수 모형에 의한 추정은 기온과 전력수요의 복잡한 비선형성과 비대칭성을 반영하는 부분선형모형과 비교하여 적합하지 않음을 의미한다. 그리고 부분선형모형에 대한 스플라인 추정에서 우도비 검정 통계량은 유의하며 이는 모수적 모형으로 설명되지 않는 비선형성이 유의함을 의미한다. 모수적 모형으로 선형함수를 가정하면 검정 통계량은 더욱 유의하므로 부분선형모형의 적합성을 나타낸다.

Table 4.2. Estimation of electricity demand models

	Parametric model		Partial linear model	
	Quadratic model	CDD/HDD model	Spline	Semiparametric
temp	-0.00101*** (0.000)		-0.00067 (0.0006)	
temp ²	0.00004*** (0.000)			
CDD		0.00128*** (0.000)		
HDD		0.00079*** (0.000)		
Humid	0.00002*** (0.000)	0.00003*** (0.000)	0.00003*** (0.000)	0.00003*** (0.000)
Y _{t-1}	1.29561*** (0.046)	1.28723*** (0.046)	1.28451*** (0.005)	1.28406*** (0.046)
Y _{t-2}	-0.46766*** (0.065)	-0.46526*** (0.065)	-0.46436*** (0.007)	-0.46390*** (0.065)
Y _{t-3}	0.07043*** (0.027)	0.07034*** (0.027)	0.06954*** (0.005)	0.06939* (0.027)
Y _{t-6}	-0.01807*** (0.004)	-0.01603*** (0.004)	-0.01517*** (0.003)	-0.01526*** (0.004)
Y _{t-9}	-0.06996*** (0.005)	-0.06959*** (0.005)	-0.06832*** (0.003)	-0.06824*** (0.005)
Y _{t-12}	0.06106*** (0.003)	0.06282*** (0.003)	0.06272*** (0.002)	0.06285*** (0.003)
Y _{t-24}	0.04653*** (0.003)	0.04370*** (0.003)	0.04299*** (0.002)	0.04298*** (0.003)
t	0.00679*** (0.000)	0.00720*** (0.000)	0.00729*** (0.000)	0.00729*** (0.000)
t ²	-0.00091*** (0.000)	-0.00097*** (0.000)	-0.00099*** (0.000)	-0.00099*** (0.000)
Jan	-0.00331*** (0.000)	-0.00267*** (0.000)	-0.00277*** (0.000)	-0.00274*** (0.000)
Feb	-0.00309*** (0.000)	-0.00337*** (0.000)	-0.00333*** (0.000)	-0.00333*** (0.000)
Mar	-0.00407*** (0.000)	-0.00462*** (0.000)	-0.00425*** (0.000)	-0.00425*** (0.000)
Apr	-0.00703*** (0.001)	-0.00610*** (0.001)	-0.00566*** (0.001)	-0.00566*** (0.001)
May	-0.00999*** (0.001)	-0.00715*** (0.001)	-0.00741*** (0.001)	-0.00743*** (0.001)
Jun	-0.00792*** (0.001)	-0.00556*** (0.001)	-0.00577*** (0.001)	-0.00582*** (0.001)
Jul	-0.00603*** (0.001)	-0.00514*** (0.001)	-0.00522*** (0.001)	-0.00527*** (0.001)
Aug	-0.00558*** (0.001)	-0.00499*** (0.001)	-0.00533*** (0.001)	-0.00537*** (0.001)
Sep	-0.00708*** (0.001)	-0.00434*** (0.001)	-0.00465*** (0.001)	-0.00467*** (0.001)
Oct	-0.00687*** (0.001)	-0.00464*** (0.001)	-0.00453*** (0.001)	-0.00454* (0.001)
Nov	-0.00244*** (0.000)	-0.00227*** (0.000)	-0.00187*** (0.000)	-0.00187*** (0.000)
H1	-0.00323*** (0.001)	-0.00368*** (0.001)	-0.00383*** (0.001)	-0.00383*** (0.001)
H2	0.02923*** (0.001)	0.02831*** (0.001)	0.02804*** (0.001)	0.02802*** (0.001)
H3	0.02342*** (0.001)	0.02253*** (0.001)	0.02224*** (0.001)	0.02223*** (0.001)
H4	0.02741*** (0.000)	0.02653*** (0.000)	0.02618*** (0.001)	0.02617*** (0.001)
H5	0.03371*** (0.001)	0.03292*** (0.000)	0.03258*** (0.001)	0.03257*** (0.001)
H6	0.03435*** (0.001)	0.03379*** (0.001)	0.03354*** (0.001)	0.03351*** (0.001)
H7	0.04089*** (0.000)	0.04063*** (0.000)	0.04047*** (0.000)	0.04042*** (0.000)
H8	0.03980*** (0.000)	0.03977*** (0.000)	0.03971*** (0.000)	0.03965*** (0.000)
H9	0.07842*** (0.000)	0.07870*** (0.000)	0.07877*** (0.000)	0.07870*** (0.000)
H10	0.01785*** (0.002)	0.01859*** (0.002)	0.01892*** (0.001)	0.01889*** (0.002)
H11	0.03449*** (0.001)	0.03518*** (0.001)	0.03556*** (0.001)	0.03550*** (0.001)
H12	-0.00752*** (0.001)	-0.00687*** (0.001)	-0.00645*** (0.001)	-0.00649*** (0.001)
H13	0.02814*** (0.001)	0.02844*** (0.001)	0.02873*** (0.001)	0.02867*** (0.001)
H14	0.07746*** (0.001)	0.07778*** (0.001)	0.07805*** (0.001)	0.07802*** (0.001)
H15	0.01059*** (0.003)	0.01116*** (0.003)	0.01147*** (0.001)	0.01147*** (0.003)
H16	0.04479*** (0.001)	0.04518*** (0.001)	0.04537*** (0.001)	0.04533*** (0.001)
H17	0.03361*** (0.001)	0.03403*** (0.001)	0.03421*** (0.001)	0.03418*** (0.001)
H18	0.04746*** (0.001)	0.04793*** (0.001)	0.04807*** (0.001)	0.04803*** (0.001)
H19	0.04697*** (0.002)	0.04757*** (0.002)	0.04775*** (0.001)	0.04773*** (0.002)

(Continued)

Continued

	Parametric model		Partial linear model	
	Quadratic model	CDD/HDD model	Spline	Semiparametric
H20	0.03772*** (0.001)	0.03829*** (0.001)	0.03844*** (0.001)	0.03843*** (0.001)
H21	0.01510*** (0.000)	0.01552*** (0.000)	0.01569*** (0.001)	0.01567*** (0.000)
H22	0.03708*** (0.001)	0.03729*** (0.001)	0.03742*** (0.001)	0.03738*** (0.001)
H23	0.01982*** (0.001)	0.01999*** (0.001)	0.02007*** (0.001)	0.02004*** (0.001)
Mon	0.02539*** (0.001)	0.02568*** (0.001)	0.02579*** (0.000)	0.02579*** (0.001)
Tue	0.01980*** (0.001)	0.02019*** (0.001)	0.02033*** (0.000)	0.02033*** (0.001)
Wed	0.01840*** (0.001)	0.01895*** (0.001)	0.01911*** (0.000)	0.01909*** (0.001)
Thu	0.01869*** (0.001)	0.01930*** (0.001)	0.01949*** (0.000)	0.01947*** (0.001)
Fri	0.01783*** (0.001)	0.01844*** (0.001)	0.01862*** (0.000)	0.01860*** (0.001)
Sat	0.00657*** (0.000)	0.00691*** (0.000)	0.00696*** (0.000)	0.00694*** (0.000)
DH	-0.00830*** (0.001)	-0.00844*** (0.001)	-0.00855*** (0.001)	-0.00854*** (0.001)
DNYDays	-0.02715*** (0.002)	-0.02801*** (0.002)	-0.02854*** (0.001)	-0.02853*** (0.002)
DTGDays	-0.02816*** (0.002)	-0.02925*** (0.002)	-0.02965*** (0.001)	-0.02971*** (0.002)
상수항	0.84387*** (0.036)	0.88110*** (0.037)	0.90920*** (0.020)	
R-squared	0.9867	0.9868		
Adj R-squared	0.9866	0.9867		

***: 1%, **: 5%, *: 10% significant. Standard errors are in the parentheses. DH = 공휴일; DNYDays = 설날 연휴; DTGDays = 추석 연휴. CDD = cooling degree days; HDD = heating degree days; .

Table 4.3. Härdle and Mammen and LR testing

Parametric model	Testing	Statistics	5% C.V.	p-value
Quadratic Model	Härdle and Mammen	46.49856	1.95996	0.00000
	LR (Spline)	431.22	5.991465	0.00000
Linear Model	Härdle and Mammen	290.81595	1.95996	0.00000
	LR (Spline)	1214.61	5.991465	0.00000

4.4. 예측력 분석

Table 4.4는 부분선형모형과 모수적 모형으로 추정하고 전력수요에 대하여 예측력을 평가한 결과이다. 모수적 모형으로 이차함수 모형과 냉난방도일 모형을 사용하였고 부분선형모형에 대하여 Robinson의 double residual 추정과 스플라인 추정을 적용하였다. 전력수요에 대한 예측 정확성은 RMSE, MAE, MAPE로 측정하였으며 이차함수 모형에 의한 예측을 기준으로 상대적 예측력을 평가하였다.

표본내 예측을 보면 RMSE를 기준으로 부분선형모형을 이용한 예측은 이차함수 모형에 의한 예측과 비교하여 스플라인 추정과 준모수 추정으로 각각 6.0%, 4.5% 향상되었다. MAE를 기준으로 부분선형 모형을 이용한 예측은 이차함수 모형에 의한 예측과 비교하여 스플라인 추정과 준모수 추정으로 각각 6.4%, 4.5% 예측력 향상을 보인다. 냉난방도일 모형에 의한 예측은 이차함수 모형과 비교하여 RMSE, MAE 기준으로 각각 3.6%, 3.7% 예측 정확성이 향상되었다.

표본외 예측은 추정기간을 2009-2012년으로 하고 2013년을 예측기간으로 하여 얻은 결과이다. Table 4.4를 살펴보면, 표본 외 예측에서 부분선형모형에 대한 스플라인 추정과 준모수 추정으로 얻은 예측은 이차함수 모형에 의한 예측과 비교하여 RMSE 기준으로 각각 6.1%, 5.1% 정확성이 향상됨을 보인다. MAE, MAPE가 모두 부분선형모형에 의한 예측의 정확성이 비슷한 크기로 향상되는 결과를 얻었다.

Table 4.5는 DM 검정으로 예측력을 비교 평가한 결과를 보인다. DM 검정은 기준이 되는 모형과 비교

Table 4.4. Prediction accuracy.

		Partial linear model		Parametric model	
		Semiparametric	Spline	Quadratic model	CDD/HDD model
In-sample forecasting	RMSE	0.12071	0.11887	0.12641	0.12182
	MAE	0.09839	0.09644	0.10300	0.09919
	MAPE	0.00908	0.00890	0.00951	0.00916
In-sample forecasting (relative ratio)	RMSE	0.95486	0.94034	1.00000	0.96367
	MAE	0.95524	0.93631	1.00000	0.96301
	MAPE	0.95520	0.93637	1.00000	0.96308
Out-of-sample forecasting	RMSE	0.12514	0.12280	0.13191	0.12661
	MAE	0.10360	0.10220	0.10910	0.10490
	MAPE	0.00946	0.00933	0.00996	0.00958
Out-of-sample forecasting (relative ratio)	RMSE	0.94868	0.93095	1.00000	0.95983
	MAE	0.94959	0.93676	1.00000	0.96150
	MAPE	0.95038	0.93753	1.00000	0.96213

CDD = cooling degree days; HDD = heating degree days; RMSE = root mean square error; MAE = mean absolute error; MAPE = mean absolute percentage error.

Table 4.5. Diebold-Mariano testing

		Reference	Partial linear model		CDD/HDD model
			Spline	Semiparametric	
In-sample forecasting	Quadratic model	Squared-error Loss	7.054***	7.125***	5.341***
		Absolute error Loss	11.500***	11.630***	7.342***
	CDD/HDD model	Squared-error Loss	5.704***	5.905***	
		Absolute error Loss	6.544***	6.643***	
Out-of-sample forecasting	Quadratic model	Squared-error Loss	5.725***	4.408***	3.671***
		Absolute error Loss	5.551***	4.371***	3.602***
	CDD/HDD model	Squared-error Loss	4.108***	1.959**	
		Absolute error Loss	3.491***	1.451	

***: 1%, **: 5%, *: 10% significant. CDD = cooling degree days; HDD = heating degree days.

하는 모형의 예측 오차를 이용하여 예측력을 평가한다. 기준모형은 이차함수 모형과 냉난방도일 모형으로 지정하였다. DM 검정에서 손실함수는 제곱 오차 손실, 절대 오차 손실로 하였다. 이차함수 모형을 기준이 되는 모형으로 하였을 때 부분선형모형에 의한 표본내 예측은 유의수준 1%에서 예측력이 유의하게 향상됨을 보였다. 다음 냉난방도일 모형을 기준이 되는 모형으로 하면 부분선형모형에 의한 표본내 예측은 유의수준 1%에서 예측 정확성이 유의하게 향상되었다. 따라서 부분선형모형에 의한 표본내 예측은 모수적 모형인 이차함수 모형, 냉난방도일 모형에 의한 예측과 비교하여 예측력 향상이 유의함을 얻었다.

표본외 예측에 대한 DM 검정에서도 부분선형모형에 의한 예측은 모수적 모형인 이차함수 모형에 의한 예측과 비교하여 통계적으로 유의한 예측력의 향상을 나타냈다. 손실함수로 고려한 제곱 오차 손실과 절대 오차 손실 모든 경우에 이차함수 모형에 의한 예측과 비교하여 부분선형모형에 의한 예측 정확성이 향상됨을 DM 검정을 통하여 확인하였다. 냉난방도일 모형과 비교하여 부분선형모형에 스플라인 추정을 적용할 경우 예측력의 향상은 유의함을 보였다. 부분선형모형을 커널 방식에 의한 준모수 추정을 하여 전력수요를 예측하면 제곱 오차 손실로는 5% 유의함을 보였으나 절대 오차 손실의 경우에는 예측력 향상은 유의하지 않았다.

5. 결론

본 연구는 전력수요를 보다 효율적으로 예측하기 위해 전력수요에 대한 기온 효과를 보다 면밀하게 파악하고 주요 특성을 반영하여 기존의 모수적인 방법보다 유연하게 분석할 수 있는 부분선형모형을 적용함으로써 기온변화와 전력수요의 비선형적 관계에 대하여 추정하고 예측력을 분석하였다. 시간 단위 고주기 자료를 사용함으로써 전력수요와 기상변수의 실시간 변동성을 충분히 반영하고자 하였다.

2009년 1월 1일부터 2013년 12월 31일까지의 시간 단위 고주기 자료를 사용하여 기온변화와 전력수요의 관계에 대하여 분석한 결과는 다음과 같이 정리할 수 있다. 첫째, 최근 기후변화와 함께 나타나는 이상기온 영역에서는 기온과 전력수요의 관계에서 블록성이 나타나고 중간영역에서는 기술적 발전에 의한 평활성이 나타나면서 비선형성과 비대칭성은 더욱 뚜렷하게 발견된다. 지금까지 전력수요의 기본모형으로 사용된 선형의 비대칭적인 V자 형태인 냉방도일 모형이나 좌우대칭인 U자 형태의 이차함수 모형과 같은 모수적 모형은 이러한 특성을 반영하지 못하므로 추정과 예측 시 오차가 발생한다. 본 연구는 이러한 특성을 반영할 수 있는 부분선형모형을 사용하여 전력수요에 대하여 추정한 결과를 얻고 모형의 적합성을 측정하였다. 부분선형모형을 추정한 결과, 고려한 변수들의 계수는 전반적으로 유의하고 기온 효과는 자료의 특성에 부합한 결과를 얻었다.

둘째, 예측력 측면을 살펴보면 표본 내 예측과 표본 외 예측 모두 부분선형모형의 예측 정확성이 향상되는 결과를 보였다. 예측 오차에서 얻은 RMSE, MAE, MAPE를 측정한 결과 모수적 방법보다 예측력이 개선되었다. Diebold-Mariano 검정을 통해 부분선형모형의 예측은 모수적 모형에 의한 예측과 비교하여 유의하게 예측력이 향상됨을 확인하였다. 따라서 부분선형모형은 전력수요에 대한 기온 효과를 반영하고 예측력 향상을 얻을 수 있다는 점에서 전력에 대한 계획수립과 에너지 정책과 관련하여 유용한 결과로 활용될 것이다.

References

- Bessec, M. and Fouquau, J. (2008). The non-linear link between electricity consumption and temperature in Europe: a threshold panel approach, *Energy Economics*, **30**, 2705–2721.
- Diebold, F. and Mariano, R. (1995). Comparing predictive accuracy, *Journal of Business and Economic Statistics*, **13**, 253–263.
- Eilers, P. H. and Marx, B. D. (1996). Flexible smoothing with B-splines and penalties, *Statistical Science*, **11**, 89–102.
- Engle, R. F., Granger, C. W. J., Rice, J., and Weiss, A. (1986). Semiparametric estimates of the relation between weather and electricity sales, *Journal of the American Statistical Association*, **81**, 310–320.
- Fan, Y. and Li, Q. (1999). Root-n-consistent estimation of partially linear time series models, *Journal of Nonparametric Statistics*, **11**, 251–269.
- Härdle, W. and Mammen, E. (1993). Comparing nonparametric versus parametric regression fits, *Annals of Statistics*, **21**, 1926–1947.
- Moral-Carcedo, J. and Vicens-Otero, J. (2005). Modelling the Non-linear Response of Spanish Electricity Demand to Temperature Variations, *Energy Economics*, **27**(3), 477–494.
- Office of Meteorology (2017). Report on Extreme Weather in 2016.
- Robinson, P. (1988). Root-n-consistent Semiparametric Regression, *Econometrica*, **56**, 931–954.
- Ruppert, D., Wand, M. P. and Carroll, R. J. (2003). *Semiparametric Regression*, Cambridge University Press.
- Shin, D. and Cho, H. (2014). Empirical study on climate sensitivity and threshold temperature of daily peak electricity demand in Korea, *Economic Research*, **32**, 175–212.
- Speckman, P. (1988). Kernel Smoothing in Partial Linear Models, *Journal of the Royal Statistical Society, Series B (Methodological)*, **50**(3), 413–436.

- Yatchew, A. (2003). *Semiparametric Regression for the Applied Econometrician*, Cambridge University.
- Yu, B. and Hwang, Y. (1997). *Research on Forecasting the Peak Electricity Demand*, Korea Energy Economics Institute, 1-56.

기온변화가 전력수요에 미치는 비선형적 영향: 부분선형모형을 이용한 추정과 예측

박지원^a · 서병선^{b,1}

^a산업연구원, ^b고려대학교 식품자원경제학과

(2019년 6월 26일 접수, 2019년 8월 25일 수정, 2019년 9월 2일 채택)

요약

최근 빈번하게 발생하는 이상기온과 기후변화로 인하여 전력수요의 변동성이 커지고 있으며 기온 영향의 증가와 함께 기온변화에 대한 전력수요의 반응은 비선형성과 비대칭성으로 나타나고 있다. 정부 에너지 정책의 변화와 4차 산업혁명의 전개에 따라 기온 효과를 보다 정확하게 추정하고 예측하는 것은 안정적 전력수급 관리를 위하여 중요한 과제이다. 본 연구는 기온변화에 대한 전력수요의 비선형적 반응에 대하여 부분선형모형을 이용하여 분석하고자 한다. 기온변화와 전력수요의 비선형·비대칭적 관계를 측정하기 위하여 Robinson의 double residual 준모수적 추정과 스플라인 추정을 적용하였다. 기상변수와 전력 소비에 대한 시간 단위 고주기 자료를 사용하여 부분선형모형으로 추정한 기온변화와 전력 소비의 관계는 기존 모수적 모형과는 다른 비선형성과 비대칭성을 갖고 있음을 확인하였다. 부분선형모형을 이용하여 얻은 전력수요에 대한 표본내·표본외 예측은 이차함수 모형과 냉난방도일 모형과 비교하여 우수한 예측력을 보였다. Diebold-Mariano 검정결과, 부분선형모형에서 얻은 예측력 향상은 통계적으로 유의하였다.

주요용어: 기온변화, 부분선형모형, 비대칭성, 비선형성, 전력수요 예측

¹교신저자: (02841) 서울시 성북구 안암로, 고려대학교 식품자원경제학과. E-mail: seomatteo@korea.ac.kr