

온도 효과를 고려한 다항 회귀분석법을 이용한 특수일 최대 전력 수요 예측 알고리즘

*위영민, *문국현, *이재희, *주성관, **송경빈
*고려대학교, **송실대학교

Load Forecasting for the Holidays using a Polynomial Regression Incorporating Temperature Effect

*Young-Min Wi, *Guk-Hyun Moon, *Jae-Hee Lee, *Sung-Kwan Joo, **Kyung-Bin Song
*Korea University, **Soongsil University

Abstract - 본 논문은 특수일 전력 수요 예측을 위해 온도 효과를 고려한 데이터 추출법을 이용하여 특수일 전력 수용 예측 오차율을 감소시키는 방법을 제시한다. 제안된 기법의 타당성을 확인하기 위해 논문에서는 통계학에서 사용되는 결정계수를 이용한다. 결정계수를 이용하여 온도효과의 고려 여부가 오차율에 미치는 영향을 분석하였다. 또한 제안된 기법은 1996년 특수일 오차율을 기존 논문의 결과와 비교 분석하여 기존 방식 대비 특수일 전력 수요예측 관련 우수성을 보였으며, 최근 데이터인 2006년 특수일 전력 수요 예측을 통하여 검증하였다.

1. 서 론

전력 수요 예측은 발전량 계획관련 수요 예측 뿐 아니라, 전력 시스템의 안정성(전력 산업 설비의 유지, 보수, 투자), 시간별 급전, 전력시장 운영 등에 중요한 지표로 사용된다. 하지만 전력 수요는 그 특성이 비선형적이기 때문에 예측하는 것에 어려움이 있다. 전력 수요가 비선형적인 특성을 갖는 원인은 전력 수요가 사회·경제적 요소, 시계절적 요소, 기후적 요소, 전력시장 요소, 기타 비선형적 요소에 의해 강한 영향을 받기 때문이다. 이러한 모든 요소를 감안하여 전력 수요를 예측하기 위한 시스템을 계획하기는 현실적으로 어려운 일이다. 지금까지 전력산업 관련 연구 분야에서 전력 수요 예측이라는 자체가 가지는 비선형적인 점을 제거하여 보다 좋은 정확도를 가진 예측 시스템을 개발하려는 많은 연구 노력과 성과가 있어 왔으며, 현재도 국·내외에 연구 기관에서 계속 진행 중이다.

전력 수요 예측 기법은 매우 다양하게 발전되어 왔다. 초창기 시계열법, 회귀분석법 등이 연구 되었으며, 국·내외로 전문가 지식 기반의 인공지능기법과 퍼지 방법을 통한 예측 방법이 연구되어 지고 있다. 또한 최근에는 기존 방법들을 결합한 하이브리드 방식 알고리즘에 대한 연구가 활발히 진행 중이다[1]-[2]. 현재까지 진행된 연구 내용에서 알 수 있듯 평일 대상의 전력 수요 예측은 대부분 높은 정확성을 보인다. 하지만 특수일에 대한 정확도는 평일에 비해 많이 떨어지는 것을 확인 할 수 있다. 국내 연구진에 의해 몇몇 특수일 전력 수요 예측 관한 논문이 국·내외 논문으로 발간되었다. 하지만 평일 수준의 정확성에는 아직 미치지 못하는 것이 현실이다.

본 논문에서는 특수일 수요 예측의 정확성을 평일 수준으로 끌어올리길 위한 알고리즘을 제시하고 제시된 방법에 따른 시험 결과를 예시하였다.

2. 본 론

2.1 특수일 전력 수요 예측

특수일 전력 수요 예측의 어려움은 평일 대비 특수일

이 가진 과거 데이터의 회소성이 근본적 원인이다. 정확한 예측을 하기 위해서는 많은 데이터가 있어야 한다. 그런 면에서 특수일은 일 년에 한 번 밖에 없기 때문에 과거 데이터가 부족한 것이 현실이다. 또한 특수일 중에는 설날, 석가탄신일, 추석과 같은 음력을 따르는 날들이 있기 때문에 과거 데이터의 구성이 매우 힘들다. 특수일의 이런 특성들은 특수일 전력 수요 패턴을 평일과 다르게 한다. 이러한 단점들의 극복 방법으로 본 논문에서는 온도 효과를 고려한 데이터 선정을 통한 다항회귀분석법을 제시한다. 온도 효과란 전력 수요 예측을 위한 시물레이션 데이터 선정 시 기준을 의미한다. 과거의 데이터 중에 예측하려는 특수일과 온도 관점에서 가장 밀접한 데이터를 선정하는 근거를 제시하며, 이러한 과정을 통해 얻은 데이터는 전력 수요 예측의 정확성을 높이는 결과를 만든다. 또한 다항회귀를 이용하여 비선형적인 수요 예측의 특성을 반영할 것이다.

본론 2.2와 2.3에서는 온도효과를 고려한 데이터 선택과 다항회귀분석법에 대한 알고리즘을 제시할 것이다. 본론 2.4에서는 제안된 알고리즘을 이용하여 2006년 특수일 전력 수요 예측을 통하여 제안된 알고리즘을 검증하였다.

2.2 온도효과를 이용한 전력 수요 데이터 선택

전력 수요는 기온과 밀접한 관계를 가진다. 그림 1은 최고 기온과 최대 전력 수요와의 관계를 한국의 2006년 데이터를 이용하여 나타낸다.

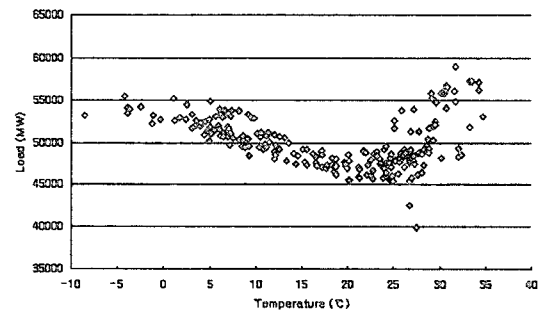


그림 1 2006년 최고 기온과 최대 수요

특히 그림 1에서 높은 온도 영역에서는 작은 온도차에서 큰 전력 수요의 변화를 확인할 수 있다. 이러한 점은 여름철의 전력 수요 예측은 다른 계절에 비해 온도의 영향을 더 받는다는 것을 확인시켜 준다. 위와 같은 전력 수요와 기온의 관계를 전력 수용 예측 기법에 적용하기 위해서는 예측을 위해 필요한 과거 데이터의 선정 기준이 있어야 한다. 전력 수요 예측 위한 데이터 선정은

다음과 같은 절차를 따른다.

- Step 1) 예측하려는 특수일을 선정한다.
 Step 2) 예측하려는 특수일과 같은 타입의 요일을 가지는 과거의 특수일을 결정한다. 요일의 타입은 평일(화요일, 수요일, 목요일, 금요일), 토요일, 일요일, 월요일 로 4가지로 구분한다[3]. 특수일 데이터에는 최고 기온과 최대 수요 데이터를 포함한다.
 Step 3) Step2에서 결정된 특수일 전의 평일 데이터 집합을 만든다. 평일 데이터에는 최고 기온과 최대 수요 데이터 포함한다.
 Step 4) 특수일의 최고 기온과 평일 최고 기온의 차를 이용하여 평일 데이터의 집합의 원소를 재구성한다. 기온 차에 의해 오름차순으로 정리 후, 계절에 따른 기준 지표를 이용하여 데이터를 선별한다. 재구성의 기준으로는 온도 차를 이용한 지표를 사용하게 되는데 이 지표는 특수일이 속한 계절에 따라 달라진다. 그림 1에서 확인했듯이 온도가 높은 날의 단위 온도 당 전력 수요 변화량이 온도가 낮은 날의 전력 수요 변화량 보다 크기 때문이다.
 Step 5) 재구성된 평일 데이터에서 같은 성향을 가지는 평일 데이터 클러스터(cluster)를 찾는다. 클러스터(cluster)에 속하는 평일 데이터를 찾기 위해 본 논문에서는 유클리드 거리(Euclidean-distance)를 사용한다. 유클리드 거리(Euclidean-distance)값 합을 최소로 하는 평일 데이터들을 선택한다.
 Step 6) Step2에서 Step5을 반복하여 예측하려는 특수일과 같은 과거의 특수일 데이터 쌍(특수일과 평일)을 만든다.

2.3 다항회귀분석

회귀분석이란 어떤 변수(독립변수)를 지정해서 그 지정변수에 대응해서 변화하는 다른 변수(종속변수)의 값을 추정하는 방법이다. 통계학에서는 비선형적인 요소가 강한 데이터에서는 단순회귀분석이 보다 다항회귀분석을 주로 이용한다. 일반적인 2차 다항회귀 모델은 다음의 수식(5)과 같다.

$$y = \alpha + \beta x + \gamma x^2 \quad (1)$$

여기서 α, β, γ 는 2차 회귀분석 모델 계수이다.

2차 회귀분석 모델의 계수는 최소자승법에 의해서 구해진다. 최소자승법은 잔차 자승의 합이 최소가 될 때의 계수 α, β, γ 를 찾아내는 기법이며, 수식은 다음과 같다.

$$\text{minimize } \sum_{i=1}^n (y_i - \alpha - \beta x_i - \gamma x_i^2)^2 \quad (2)$$

본 논문에서 다항회귀분석을 위한 입력 데이터(X_i, Y_i)는 다음과 같은 방식으로 만들어진다. 입력 데이터 X_i 는 특수일 전 평일 4일의 값을 토대로 만들어진다. 특수일 직전 평일 4일의 최대값이 각각 m_1, m_2, m_3, m_4 라 하면 이 중 가장 큰 값(M)으로 정규화하여 이들의 평균으로 정한다[1].

$$x_i = \frac{m_1 + m_2 + m_3 + m_4}{4M} \quad (3)$$

Y_i 는 특수일 수요량을 m_5 라 하면 다음과 같이 표현된다[1].

$$y_i = \frac{m_5}{M} \quad (4)$$

2.4 특수일 전력 수요 예측 사례연구

온도 영향성을 고려한 다항회귀분석 알고리즘을 적용하여 사례연구를 수행하였다. 수요예측의 에러는 다음의 식(0)으로 구한다.

$$\text{Error}(\%) = \frac{|P^{\text{Forecast}} - P^{\text{Actual}}|}{P^{\text{Actual}}} \times 100 \quad (5)$$

여기서 P^{Forecast} 는 예측 수요량이며, P^{Actual} 는 실제 수요량이다.

2006년 석가탄신일 전력 수요 예측을 제안된 방법으로 테스트를 진행하였다. 표 1에서는 예측에 사용된 데이터들의 온도 편차를 보여준다. 그림 1에서 확인했듯이 25℃를 전·후하여 온도에 대한 최대 전력 수요 변화율의 부호가 달라지기 때문에 2006년 석가탄신일 경우 변화율의 부호가 바뀌는 근처에 데이터가 존재함으로써 온도 효과의 비중이 커진다. 표 2는 표 1에서 사용된 데이터를 이용한 전력 수요 예측 오차를 테스트 결과를 나타낸다. 표 1과 2를 통해 전력 수요 예측에 사용되는 평일 및 특수일 데이터 간의 온도차가 작을수록 예측의 정확성을 높일 수 있음을 확인할 수 있다.

표 1. 2006년 석가탄신일 전력 수요 예측 데이터

석가탄신일	평일과 특수일 데이터의			특수일 실제온도
	최대온도 편차		제안된 방식	
	온도효과 비교	비교		
과거	1	16	0.7	14.2
데이 터	2	8.7	1.1	24.8
	3	4.8	4.3	19.4
2006년 데이터	11.6		7.5	23.6

표 2. 2006년 석가탄신일 전력 수요 예측 오차율

특수일	온도효과		제안된 방법
	비교	다항회귀	
석가탄신일	2.6051	1.8337	

3. 결 론

특수일은 평일과는 다른 수요패턴으로 인해 수요 예측의 정확성이 크게 떨어진다. 이러한 현상을 개선하기 위해 본 논문에서는 온도 효과를 고려한 다항회귀분석 모델을 제안하였다. 다항회귀모델의 계수를 찾기 위해 최소자승법을 사용하였다. 2006년 석가탄신일을 대상으로 사례연구를 진행한 결과 온도효과를 고려한 다항회귀분석법의 특수일 전력 수요 예측 정확성이 향상됨을 입증하였다.

향후 온도 효과를 고려한 데이터 선택법을 이용하여 특수일 뿐만 아니라 평일의 전력 수요예측에도 적용에 대한 연구를 진행할 예정이다. 또한 퍼지 다항회귀에도 온도 효과를 고려한 데이터 선택법을 적용하여 특수일 전력 수요 예측의 정확성을 높이는 연구를 지속할 것이다.

[참 고 문 헌]

- [1] S. Fan and L. Chen, "Short-Term Load Forecasting Based on an Adaptive Hybrid Method," IEEE Trans. on Power Systems, vol.21, no.1, pp.392-401, 2006.
- [2] K.B. Song, S.K. Ha, J.W.Park,D.J.Kweon,K.H. Kim, "Hybrid Load Forecasting Method with Analysis of Temperature Sensitivities," IEEE Trans. on Power Systems, vol.21, no.2, pp. 869-876, 2006.
- [3] K.B. Song, Y.S. Baek, D.H. Hong, and G.S. Jang, "Short-Term Load Forecasting for the Holidays Using Fuzzy Linear Regression Method," IEEE Trans. on Power Systems, vol.20, no.1, pp.96-101, 2005.