

온도와 부하의 관계를 이용한 단기부하예측

이경훈\* · 이윤호\* · 김진오\* · 이효상\*\*  
 \*한양대학교 전기공학과 \*\*한국전력거래소

Short-Term Load Forecasting using Relationship of Temperature and Load

Kyung-Hun Lee\* · Yunho Lee\* · Jin-O Kim\* · Hyo-Sang Lee\*\*  
 \*Dept. of EE, Hanyang University · \*\*Korea Power Exchange(KPX)

**Abstract** - This paper presents a model for short-term load forecasting using relationship of temperature and load. We made one-day ahead load forecasting model using hourly normalized load and 11 dummy variables that were classified by day characteristics such as day of the week, holiday, and special holiday.

1. 서 론

전력 부하의 예측의 정확도 향상은 발전 비용의 감소와 전력 공급의 신뢰도 향상의 결과로 나타나기 때문에 장기 및 단기간의 전력 부하의 예측은 전력 계통의 경제적 운용의 측면에서 매우 중요하다. 특히 그 결과는 전력 계통의 제어 및 전력조류계산의 입력 자료로 사용되기 때문에 지금까지 다양한 방식에 의한 예측 기법이 제안되었다. 이러한 종래의 방법으로는 크게 시계열 방법(1)과 인공지능 기법의 신경회로망(3,4) 등이 있다. 시계열 방법, 특히 자기회귀(AR : autoregression)(2), 통합혼합모형(ARIMA : autoregressive integrated moving average) 등은 예측 모형을 설정하는데 있어서 자기 변수만을 고려하므로 예측 모형으로서 단순하다. 또한, 부하 예측 문제에 내재하는 비선형성, 유연한 입력 선택, 모델링의 복잡성 등과 같은 문제를 적절히 해결할 수 있는 방법으로 신경회로망이 알려져 있으나 기본적으로 몇 가지 단점이 있다. 즉, 학습을 통해 유추된 예측 모델의 규칙을 표현할 수 없고, 최적의 성능을 얻기 위한 다층 퍼셉트론(MLP : multilayer perceptron)의 은닉층의 수와 각 은닉층의 노드 수를 결정하는 효과적이며 체계적인 방법이 존재하지 않는다(5).

따라서 본 논문에서는 기존의 전력 부하 예측 방식처럼 부하를 유형별(평일, 월요일, 토요일, 일요일, 특수일)로 분류하는 대신 시간에 대해 연속성을 갖는 부하 특성과 주기성을 고려한 더미변수(Dummy Variable)를 도입하여 최대 및 최소부하를 추정하고 예측일에 따라 normalized 부하를 계산한 후 최종 24시간에 대해 부하를 예측하는 모델을 제안한다.

2. 부하의 특성에 따른 더미변수의 도입

기상이나 특수일과 같은 부하의 변동 요인이 없다면 하루의 부하 곡선은 보통 화요일부터 금요일까지 그대로 유지된다. 그림 1은 2000년 12월 4일 월요일부터 12월 10일 일요일까지 한 주간의 부하 특성 곡선을 나타내는데 토요일, 일요일, 월요일은 평일에 비해 전 시간대에 걸쳐 혹은 시간대별로 낮은 수요치를 보이지만 전주의 동일 요일의 부하 곡선과는 유사한 경향을 보인다. 따라서 부하 예측에 있어서 한 주간의 요일을 보통 화요일부터 금요일까지의 평일, 토요일, 일요일, 월요일, 그리고 특수일로 나누는 것이 일반적이었다.

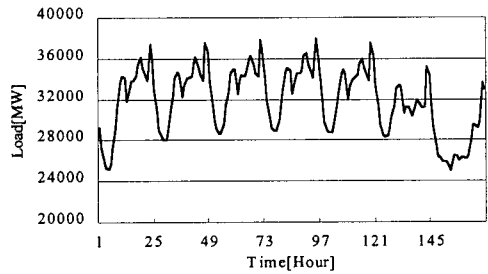


그림 1. 일반적인 한 주간의 부하특성곡선

그러나 전력 부하를 시간에 대해 연속성을 갖는 시계열 자료로 간주하여 더미변수를 도입하면 부하패턴의 효과를 하나의 예측 모델에 표현하는 것이 가능하다. 따라서 본 논문에서는 요일을 유형별로 분류하는 대신 더미변수를 도입하여 예측 모델 수립의 자동화를 꾀하고자 한다.

더미변수(Dummy Variable)는 특성의 존재 유무에 따라 1 또는 0의 값을 할당받음으로써 종속변수의 질적 변화를 나타내기 위한 변수로 쓰인다(6). 본 논문에서는 표 1과 같은 입력 더미변수의 구성을 통하여 특정일의 더미변수만을 1의 값을 취하여 요일별 부하패턴이 미치는 영향도(影響度)를 종속변수, 즉 예측부하에 반영하고자 한다. 공휴일의 경우, 해당 공휴일이 주말이면 부하 패턴은 평상시의 주말과 크게 다르지 않은 반면, 평상일이면 대체로 작게 나타난다. 따라서 공휴일의 경우는 표 1과 같이 주말과 평일을 구분하였다.

표 1. 더미변수의 구성

평 일	공 휴 일	명 절
토요일(D <sub>1</sub> )	D <sub>7</sub>	전 날(D <sub>9</sub> )
일요일(D <sub>2</sub> )		
월요일(D <sub>3</sub> )	D <sub>8</sub>	명 절(D <sub>10</sub> )
수요일(D <sub>4</sub> )		
목요일(D <sub>5</sub> )		
금요일(D <sub>6</sub> )		
		다음날(D <sub>11</sub> )

표 1의 공휴일과 명절(신정, 설날, 추석)의 더미변수는 다른 더미변수보다 우선권을 갖는다. 예를 들어 토요일이면서 공휴일인 경우, 더미변수 D<sub>7</sub>에만 1의 값이 할당된다. 즉 화요일을 제외한 모든 날은 11개의 더미변수 중 하나의 더미변수에 대하여 1의 값을 갖고 모든 더미변수가 0인 경우는 화요일에 해당된다.

3. 예측 모델 및 예측 절차

본 논문에서 일간 최대부하 및 최소부하와 시간별 normalized 부하의 관계를 이용하여 일전예측(one-day ahead forecasting) 모델을 수립하는 예측 절차를 그림 2에 나타내었다. 단계 1에서는 더미변수를 포함하는 일간 최대부하와 최소부하의 예측 모델을 수립하고 예측일의 최대부하와 최소부하를 추정하고 단계 2에서는 표 1의 더미변수의 구성을 참조하여 예측일에 해당하는 시간별 normalized 부하를 계산한 후 단계 3에서 실제 부하 예측에 적용한다. 이러한 부하 예측 절차를 자세히 설명하면 다음과 같다.

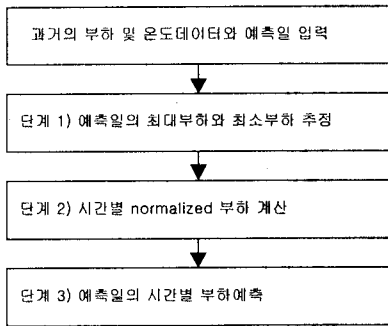


그림 2. 부하예측 절차

### 3.1 단계 1 - 최대부하와 최소부하의 추정

예측하고자 하는  $d$ 일을 기준으로 본 논문에서 제시하는 과거 최대 및 최소부하 데이터와 최고 및 최저온도 데이터, 일간 총부하, 그리고 더미변수를 입력으로 하는 예측 모델은 다음 식 (1)과 같다.

$$L_{peak(valley)}^d = a_0 \cdot peak(valley) + \sum_{i=1}^{N_i} b_i \cdot peak(valley) L_{peak(valley)}^{d-i} + \sum_{j=1}^{N_j} c_j \cdot peak(valley) T_{high}^{d-j} + \sum_{k=1}^{N_k} d_k \cdot peak(valley) T_{low}^{d-k} (1) + \sum_{l=1}^{N_l} f_l \cdot peak(valley) L_{total}^{d-l} + \sum_{m=1}^{N_m} g_m D_m + e_i$$

여기서,  $L_{peak(valley)}^d$  : 예측일( $d$ 일)의 일간 최대 및 최소부하의 추정치[MW]

$T_{high}^d, T_{low}^d$  :  $d$ 일의 일간 최고 및 최저온도[°F]

$L_{total}^d$  :  $d$ 일의 총부하[MW]

$D$  : 더미변수

$N_i, N_j, N_k, N_l$  : 일간 시차(lag)의 개수

$N_m$  : 더미변수의 개수(11개)

$a_0, b_i, c_j, d_k, f_l, g_m$  : 상수항 및 각 변수에 대한 회귀계수

$e_i \sim i.i.d.(0, \sigma^2)$

식 (1)의 오른쪽 식에서 첫째 항은 상수항이고 예측하고자 하는  $d$ 일 기준으로 둘째 항은 과거  $N_i$ 일 동안의 최대 및 최소부하량이며 셋째 항과 넷째 항은 각각  $N_j, N_k$ 일 동안의 최고 및 최저온도를 뜻한다. 여기에서 최고 및 최저온도는 지역별 수요구성비(서울:0.51, 부산:0.22, 대구:0.13, 대전:0.09, 광주:0.05)에 따라 가중평균한 온도를 입력데이터에 맞게 화씨온도로 바꾼 값이다. 또한 다섯째 항은  $N_l$ 일 동안의 총부하를 여섯째 항은 더미변수를 나타내며, 표 1에서 밝힌 바와 같이 본 논문에서 설정한 더미변수의 개수  $N_m$ 는 11이다. 마지막 항은 평균 0, 분산  $\sigma^2$ 인 독립적 동일분포를 갖는 확률변수로서 백색잡음(white noise)을 나타낸다. 이러

한  $(1+N_i+N_j+N_k+N_l+N_m)$ 개의 입력 변수에 대한 상수항 및 회귀계수의 값은 잔차제곱합(Residual Sum of Squares, RSS)을 최소로 하는 계수를 구하는 기법인 최소제곱법(Ordinary Least Squares, OLS)을 이용하여 구하였다. 이러한 회귀계수들의 값을 토대로 예측 모델의 모형을 설정하여  $d$ 일의 최대 및 최소부하를 추정한다.

### 3.2 단계 2 - normalized 부하 계산

더미변수( $D$ )에 따른 예측일( $d$ )의  $h$ 시간에 대한 normalized 부하의 계산은 다음 식 (2)를 이용한다.

$$L_{nor}^d(h, D) = \frac{1}{n} \sum_{n=1}^N \frac{L^{d-n}(h, D) - L_{valley}^{d-n}}{L_{peak}^{d-n} - L_{valley}^{d-n}} (2)$$

$L_{nor}^d(h, D)$ 는 0과 1 사이의 값으로서 해당 예측일의 더미변수를 기준으로 최근 2일의  $h$ 시간의 평균치를 normalized 부하로 정의하였다. 즉,  $N_n$ 은 2이다.

### 3.3 단계 3 - 시간별 부하 예측

단계 1과 2에서 구한 최대 및 최소부하와 시간별 normalized 부하를 입력으로 하는 다음 식 (3)과 같은 예측 모델을 이용하여  $d$ 일의 최종 24시간 부하를 예측한다.

$$L_{nor}^d(h) = \frac{d_{nor}}{d_{nor}}(h, D) (L_{peak}^d - L_{valley}^d) + L_{valley}^d (3)$$

### 3.4 예측 성능 평가

제한한 예측 모델의 효용성을 검증하기 위하여 예측 부하의 실제 부하에 대한 평균 오차는 식 (4)와 같이 시간별 및 요일별 평균 오차 절대값의 합을 평균하여 구한 절대 평균 백분율 오차(Mean Absolute Percentage Error, MAPE(%))로 정의하였다[4].

$$MAPE[\%] = \frac{1}{N} \sum_{n=1}^N \frac{|Actual(n) - Forecast(n)|}{Actual(n)} \times 100 (4)$$

여기서,  $N$  : 예측 날짜의 개수

Actual : 실제부하[MW]

Forecast : 예측부하[MW]

## 4. 사례 연구

예측일의 최대 및 최소부하를 추정하는 식 (1)의 입력값은 예측일을 기준으로 과거 300일의 실제 우리나라의 부하 및 온도데이터를 이용하였다. 즉, 예측일의 수만큼 예측 모델식 수립하였다. 또한 예측구간은 1994년 1월 1일부터 12월 31일까지 정하였으며, 최대 및 최소부하의 예측 모델의 수립에 사용되는 일간 최대 및 최소부하와 총부하의 시차 개수  $N_i$ 와  $N_j$ 는 7, 최고 및 최저온도에 대한 시차 개수  $N_k$ 와  $N_l$ 는 1로 정하였다.

표 2는 각 시간별 예측 결과를 다시 분기별로 정리한 것으로 여름철인 3/4분기의 오차가 가장 크고 가을 겨울철인 4/4분기의 오차가 가장 작게 나타났다. 표 3은 예측 결과를 월별로 분석한 것으로 최소부하에 비하여 최대부하의 예측 결과는 2배 이상 나쁘게 나타났다. 예측일의 24시간에 대한 부하를 구하는 식 (3)의 결과는 최대부하 및 최소부하의 추정치에 좌우되는데 최대부하 추정치의 오차가 너무 크게 나타나서 전체 24시간에 대한 예측 결과에 나쁜 영향을 미친 것으로 사료된다. 향

후 입력변수의 가감이나 새로운 입력변수의 도입을 통하여 좀 더 정확한 최대부하의 추정치를 얻는다면 더 좋은 예측 결과를 기대할 수 있을 것으로 판단된다.

표 2. 시간별 예측 결과[%](1994.1.1~12.31)

시간	1/4	2/4	3/4	4/4	1년
1	1.22	1.11	1.79	0.97	1.28
2	1.02	1.08	1.57	0.95	1.15
3	0.86	1.04	1.33	0.95	1.04
4	0.94	0.98	1.25	0.93	1.03
5	0.95	0.98	1.21	0.94	1.02
6	1.16	1.17	1.32	1.01	1.17
7	1.56	1.15	1.49	1.05	1.31
8	1.85	1.58	1.62	1.31	1.59
9	2.31	1.99	2.32	1.72	2.08
10	2.55	2.18	2.77	1.96	2.36
11	2.52	2.35	3.04	1.97	2.47
12	2.47	2.49	3.18	1.96	2.52
13	2.26	2.62	3.53	1.76	2.54
14	2.42	2.76	3.33	1.70	2.55
15	2.65	2.84	3.40	1.75	2.66
16	2.73	2.72	3.39	1.79	2.66
17	2.75	2.71	3.33	1.80	2.65
18	2.71	2.69	3.05	2.04	2.62
19	2.49	2.64	2.74	1.80	2.42
20	2.47	2.23	2.60	1.60	2.22
21	2.42	2.28	2.78	1.44	2.23
22	2.42	2.31	2.98	1.31	2.25
23	2.41	2.08	2.88	1.28	2.16
24	2.42	2.02	2.92	1.33	2.17
평균	2.06	2.00	2.49	1.47	2.01

표 3. 월별 예측 결과[%](1994.1.1~12.31)

월	최대부하	최소부하	24시간
1	1.77	1.05	1.71
2	3.03	1.13	3.16
3	1.49	0.69	1.44
4	1.41	0.84	1.30
5	2.21	1.21	1.92
6	4.03	0.97	2.78
7	3.25	1.08	2.58
8	3.00	1.38	2.48
9	2.75	1.36	2.41
10	1.75	1.23	1.58
11	1.09	0.77	1.08
12	2.38	0.90	1.75
평균	2.34	1.05	2.01

실제 부하에 대해 분석한 결과 11월의 오차가 가장 작고, 6월에 가장 크게 나타났다. 이 결과를 그림 3과 4에 나타내었다. 6, 7, 8월의 부하 예측 결과가 더 나쁘게 나타나는 이유는 여름철의 부하는 과거의 부하 및 온도 외에 불쾌지수나 습도 등 다른 기상요인 등에 의해 영향을 받기 때문인 것으로 판단된다.

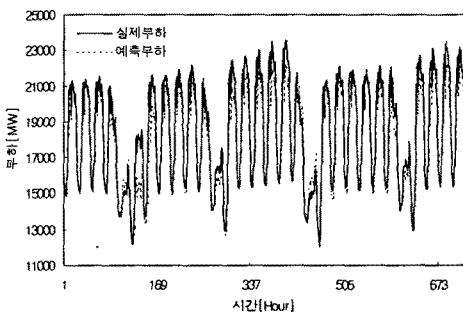


그림 3. 월별 예측 결과(6월)

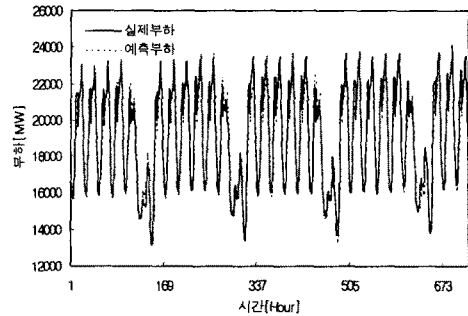


그림 4. 월별 예측 결과(11월)

#### 4. 결 론

본 논문에서는 부하 데이터를 유형별로 분류하지 않고 요일 및 특수일에 관계없이 모든 날에 대해 24시간 동안의 시간별 전력 부하를 예측하였다. 부하 데이터의 주기의 규칙성과 연속성을 이용한 더미변수(Dummy Variable)를 도입하여 최대 및 최소부하의 추정치를 구하고, 시간별 normalized 부하를 정의하였다. 예측일의 24시간에 대한 최종 예측 단계에서는 최대 및 최소부하의 추정치와 normalized 부하의 관계를 이용한 예측 모델식이 사용되었다.

향후 사용자의 보정이나 입력 변수의 가감과 같은 별도의 방법을 예측 모델에 적용하여 예측 모형을 수립한다면 더 좋은 예측 결과를 얻을 수 있을 것으로 사료된다.

#### (참 고 문 헌)

- [1] M. T. Hagan and S. M. Behr, "The Time Series Approach to Short-Term Load Forecasting", *IEEE Trans. on Power Systems*, Vol.2, No.3, pp.785~791, 1987.
- [2] A. D. Papalexopoulos and T. C. Hesterberg, "A Regression-Based Approach to Short-Term System Load Forecasting", *IEEE Trans. on Power Systems*, Vol.5, No.4, pp.1535~1547, 1990.
- [3] C. N. Lu and H. T. Wu "Neural Network Based Short Term Load Forecasting", *IEEE Trans. on Power Systems*, Vol.8, No.1, pp.336~342, 1993.
- [4] M. Daneshdoost, M. Lotfalian and G. Bumroongit, "Neural Network with Fuzzy Set-Based Classification for Short-Term Load Forecasting", *IEEE Trans. on Power Systems*, Vol.13, No.4, pp.1386~1391, 1998.
- [5] 박영진, 심현정, 왕보현, "뉴로-퍼지 모델을 이용한 단기 전력 수요 예측시스템", *전기학회 논문지*, 49권 3호, pp.107~117, 2000.
- [6] Joseph F. Hair, Rolph E. Anderson, and Ronald L. Tatham, William C. Black, *Multivariate Data Analysis*, Prentice Hall, Inc., New Jersey, 1998.