

온도와 부하의 비선형성을 이용한 단기부하예측에서의 TAR(Threshold Autoregressive) 모델

論 文

50A-9-1

TAR(Threshold Autoregressive) Model for Short-Term Load Forecasting Using Nonlinearity of Temperature and Load

李景勳* · 李允昊** · 金鎮吾***
(Kyung-Hun Lee · Yunho Lee · Jin-O Kim)

Abstract - This paper proposes TAR(Threshold Autoregressive) model for short-term load forecasting including temperature variable. In the scatter diagram of daily peak load versus daily high or low temperature, we can find out that the load-temperature relationship has a negative slope in the lower regime and a positive slope in the upper regime due to the heating and cooling load, respectively. TAR model is adequate for analyzing these phenomena since TAR model is a piecewise linear autoregressive model. In this paper, we estimated and forecasted one day-ahead daily peak load by applying TAR model using this load-temperature characteristic in these regimes. The results are compared with those of linear and quadratic regression models.

Key Words : Short-term load forecasting, TAR(Threshold Autoregressive) model

1. 서 론

전력부하 예측의 정확도 향상은 발전 비용의 감소와 전력 공급의 신뢰도 향상의 결과로 나타나기 때문에 장기 및 단기 기간의 전력부하의 예측은 전력계통의 경제적 운용의 측면에서 매우 중요하다. 특히 그 결과는 전력계통의 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[2,3]이 있다. 이 중에서 시계열 방법, 특히 자기회귀(AR : Autoregression)[4], 통합혼합모형(ARIMA : Autoregressive Integrated Moving Average) 등은 예측 모형을 설정하는데 있어서 자기 변수만을 고려하므로 예측 모형으로는 너무 단순하고 또한 선형모델이므로 좋은 예측 결과를 기대할 수 없다. 이러한 시계열 방법과 비교하여 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 예측모델의 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].

본 논문은 온도를 변수로 갖는 단기부하예측을 위하여 TAR(Threshold Autoregressive) 모델을 도입하였다. 겨울철의 난방부하, 여름철의 냉방부하의 영향으로 온도와 부하의 관계는 임계값을 기준으로 부(負)의 상관관계와 정(正)의 상관관계를 나타낸다[6]. 이러한 부하와 온도의 상관관계를 이용하여 일별 최고 및 최저온도와 최대부하의 관계를 임계값을 기준으로 구분하는 TAR 모델을 도입하여 회귀분석하였으며 이를 선형 모델, 2차 함수 모델과 비교하여 TAR 모델 도입의 필요성을 검증하였다. 또한 실제 예측 단계에서는 예측결과에 각 영역별 부하성장률을 반영하여 예측력을 높일 수 있었다.

2. 부하와 온도의 관계

전력부하의 변동요인으로는 크게 요일의 패턴(평일, 주말, 특수일)과 기상의 변화 등이 있다. 주말, 공휴일, 공휴일 전, 후일, 연휴기간 등의 특수일은 평상일과 매우 다른 전력부하 패턴을 나타내며, 특히 온도는 습도나 일조시간 등의 다른 기상요인에 비해 전력부하의 변동에 더 많은 영향을 미친다. 따라서 부하예측에 있어서 더 좋은 예측 결과를 기대하기 위해서는 부하와 온도의 관계를 규명하는 것이 급선무이다. 그림 1은 출발시점이 서로 다른 1년 동안의 일별 최고 온도와 일별 최대부하의 관계이며, 각 데이터는 특수성을 띠는 주말과 특수일을 제외하였다. 그림에서 임계값을 기준으로 그 이상의 영역은 여름철 부하로 그 특성이 온도에 대해 정의의 상관관계로 나타나는 반면, 그 이하인 영역은 각각 다르다. 즉, 출발시점을 1월부터 잡은 그림 1의 (a)의 경우 하위영역은 뚜렷하게 부의 두 가지 선형관계가 나타나고, 출발

* 準會員 : 漢陽大 大學院 電氣工學科 碩士課程

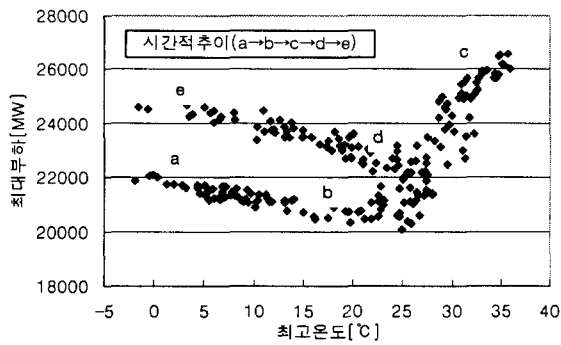
** 準會員 : 漢陽大 大學院 電氣工學科 研究員

*** 正會員 : 漢陽大 電氣工學科 副教授 · 工博

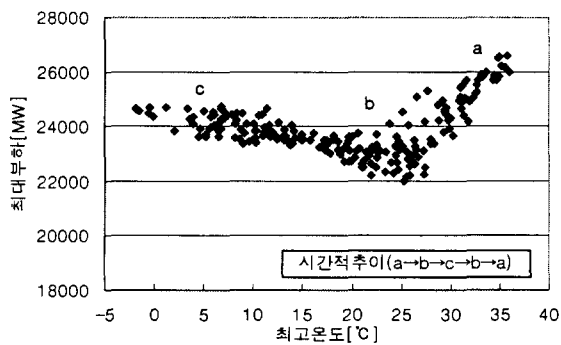
接受日字 : 2000年 11月 28日

最終完了 : 2001年 6月 7日

시점을 7월부터 잡은 (b)의 경우는 하나의 부의 선형관계가 나타난다. 이를 자세히 살펴보기 위해 시간적인 추이에 따라 부하와 온도의 특성관계를 보면 (a)의 경우 겨울철인 1월부터 시작하므로 a 영역에서 출발하여 시간이 흐를수록 온도는 올라가지만 겨울철 난방부하의 감소로 부하는 상대적으로 떨어진다. 이러한 추세는 어느 시점(b 영역)에 이르면, 즉 더위를 느끼는 여름철이 되면 온도의 상승에 따른 냉방부하의 영향으로 부하도 함께 상승하는 형태로 바뀌게 된다. 이러한 현상이 c 영역까지 계속되다가 다시 온도가 떨어지는 가을철이 되면 그와 반대로 냉방부하의 감소로 인해 부하는 다시 떨어지며(d 영역), d 영역이 지나면서 서서히 증가하는 겨울철 난방부하의 영향으로 온도의 하강에 대해 부하는 상승하게 된다. 여기서 주목할 점은 겨울철과 봄에 해당하는 a와 e의 영역에서 비슷한 온도에 대해 부하패턴이 다르게 나타난다는 것이다. 그 이유는 각 영역이 시간적으로 거의 1년의 차이가 나기 때문에 연간 대략 10%에 해당하는 부하의 성장률이 반영된 결과이다.



(a) 1994. 1. 1~1994. 12. 31



(b) 1994. 7. 1~1995. 6. 31

그림 1 최대부하와 최고온도의 관계
Fig. 1 Scatter diagram of daily peak load vs. high temp.

그림 1의 (a)와는 다르게 출발시점을 7월부터 잡은 (b)의 경우는 (a)의 c 영역부터 출발하는 것과 같으므로 (a)에서와 동일한 이유로 인해 시간적인 추이는 일단 a 영역에서

b 영역으로, b 영역에서 다시 c 영역으로 이동된다. c 영역은 겨울철에 해당되는데 해가 바뀌면서 점차 온도는 다시 상승하게 되고 (a)에서와 마찬가지로 난방부하의 감소로 인해 부하는 여름철이 시작되는 b 영역까지 하강하며 그 후에는 a 영역까지 온도의 상승에 대해 부하도 함께 증가한다. 겨울철과 봄에 있어서 시간적인 차이로 인해 두 영역으로 나뉘었던 (a)의 경우와는 달리 (b)의 경우는 시간적으로 별 차이가 없기 때문에 한가지 영역으로 나타난다. 따라서 본 논문은 모델을 단순화하고 예측력을 높이기 위해 그림 1의 (b)와 같이 1년을 7월부터 다음 해 6월까지로 설정하였으며, TAR 모델 도입의 필요성을 설명하고자 한다.

3. TAR(Threshold Autoregressive) 모델

TAR(Threshold Autoregressive) 모델은 임계값을 기준으로 데이터를 구분하여 각각을 다른 회귀 모형으로 취급한다. 일반화된 다변수 TAR 모델의 형태는 다음 식 (1)과 같다.

$$Y_t = \begin{cases} a_0 + (a_1 Y_{t-1} + a_2 Y_{t-2} + \dots + a_i Y_{t-i}) \\ + (b_1 X_{t-1} + b_2 X_{t-2} + \dots + b_j X_{t-j}) + e_t, \\ \text{if } X_{t-d} < A \\ \\ c_0 + (c_1 Y_{t-1} + c_2 Y_{t-2} + \dots + c_k Y_{t-k}) \\ + (d_1 X_{t-1} + d_2 X_{t-2} + \dots + d_l X_{t-l}) + e_t, \\ \text{if } X_{t-d} \geq A \end{cases} \quad (1)$$

여기서, A : 임계값(threshold value)
d : 지연 파라미터(delay parameter)
X_{t-d} : 임계기준 변수(threshold variable)
e_t ~ iid(0, σ²)

임계값을 기준으로 두 변수간의 비선형 관계를 규명하기 위하여, 다변수 시계열을 다루고 상태 변수를 포함하는 일반화된 TAR 모델을 적용한다. TAR 모델의 각 파라미터를 추정할 때 지연 파라미터 d, 임계값 A를 어떻게 결정할 것인가가 문제가 되는데 이러한 값들을 계산하기 위해서 일반적으로 TONG[7,8]의 방법을 따른다. 일반적으로 종속변수에 대해 임계값을 기준으로 그 이하인 영역을 하위영역(Lower Regime)이라 하고, 그 이상인 영역을 상위영역(Upper Regime)이라고 한다.

다음 식 (2)는 본 논문에서 제시하는 일별 최대부하를 종속변수, 일별 최고 및 최저온도를 독립변수로 갖는 TAR 모델이다.

$$L_{peak} = \begin{cases} a_{0,L} + a_{1,L}T + e_t, & \text{if } T < A \\ a_{0,U} + a_{1,U}T + e_t, & \text{if } T \geq A \end{cases} \quad (2)$$

여기서, L_{peak} : 일별 최대부하[MW]
T : 일별 최고 및 최저온도[°F]

$a_{0,L}, a_{1,L}, a_{0,U}, a_{1,U}$: 하위영역(Lower Regime), 상위영역(Upper Regime)의 회귀계수

TAR 모델을 2차 함수 모델과 비교할 때 식 (2)의 T 가 만약 섭씨온도이면 영하의 온도와 영상의 온도가 서로 구분되지 않으므로 본 논문에서는 온도 T 를 화씨온도로 정하였다. 또한 본 논문은 부하와 온도의 상관관계를 밝히는 것에 초점을 두고 부하와 온도에 대해 규칙성을 띠지 않는 주말이나 특수일의 데이터를 제외하였으므로, 과거의 시계열 Y_{t-i} 또는 Y_{t-k} 까지 모델에 반영한다면 변수간 간격이 일정하지 않게 되어 신뢰성 있는 모델을 기대할 수 없다. 따라서 일별 최고 및 최저온도만을 독립변수로 설정하였다.

3.1 임계값(Threshold Value) 선정

TAR 모델을 수립하기 위해서는 먼저 오차를 가장 작게 하는 임계값 A 를 찾아야 한다. 임계값을 선정하는 절차를 그림 2에 나타내었다. 먼저 n 개의 임의의 임계값을 선택하여 예측치와 실제치의 오차가 가장 작도록 하는 함수를 구하는 기법인 최소자승법(OLS : Ordinary Least Squares)을 이용하여 모든 n 에 대하여 하위영역과 상위영역 각각의 계수를 구한다. 세 번째와 네 번째 단계에서는 모든 n 에 대하여 $RSS1$ 과 $RSS2$, $TRSS$ 와 $Pooled SE$ 를 계산한다. 여기서 $RSS1$ 과 $RSS2$ 는 각각 하위영역과 상위영역의 잔차제곱합(RSS : Residual Sum of Squares), $TRSS$ 는 $RSS1$ 과 $RSS2$ 의 합이며 $Pooled SE$ (Pooled Standard Error)는 다음 식 (3)과 같다[9].

$$Pooled SE = (TRSS / 자유도)^{0.5} \quad (3)$$

마지막 단계에서는 최소 $Pooled SE$ 값을 갖는 임계값 A_i 를 선정한다.

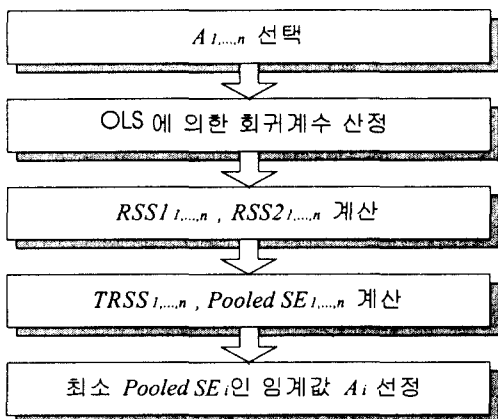


그림 2 임계값 선정 절차
Fig. 2 Procedure of the selection of optimal threshold value

3.2 검정 통계량

(1) 결정계수

관측된 데이터를 이용하여 회귀모형이 추정되면 이 추정된 회귀선이 표본자료들을 얼마나 잘 설명하는가를 검토해야 한다. 이때 사용되는 척도가 결정계수(Coefficient of Determination, R^2)이다. 즉 결정계수란 추정된 회귀선이 관측점들에 얼마나 적합한지를 나타내는 값으로서 결정계수 R^2 는 1에 가까울수록 추정된 회귀선은 관측점들에 적합하다.

(2) DW(Durbin-Watson)값

최소자승법을 사용하여 회귀계수를 추정할 때 잔차는 서로 독립적으로 random하게 발생해야 한다는 잔차항에 대한 독립성 가정을 하게 되는데 이때 잔차항이 서로 상관되어질 때 자기상관(Autocorrelation)이 존재한다고 말하고 그 존재여부는 DW값을 사용하여 검정할 수 있으며 그 값이 2에 가까울수록 독립적이라 할 수 있다.

(3) t-값

회귀계수에 대한 추정결과는 일반적으로 추정된 회귀계수, 각 회귀계수의 표준오차, 각 회귀계수들에 대한 검정 통계량 t-값을 의미한다. 결정계수 R^2 는 회귀모형식에서 독립변수 전체가 종속변수를 얼마나 설명할 수 있는가를 의미하는데 반하여 t-값은 각각의 독립변수가 종속변수를 얼마나 잘 설명할 수 있는가를 검증하는 통계량이다. 이는 회귀계수를 각각의 표준오차로 나누면 회귀계수 각각의 t-값을 얻을 수 있는데 이 값이 2 이상이 되어야 유의값을 갖고 있다고 분석된다.

4. 사례 연구

TAR 모델의 성능을 검증하기 위하여 사례 연구는 두 가지 방법으로 행하여 졌다. 먼저 동일한 관찰치에 대해 임계값을 선정하여 TAR 모델을 수립한 후, 이를 선형, 2차 함수 모델과 비교한다. 이를 1절에 나타내었다. 2절에서는 수립된 모델을 기반으로 연간 부하성장률을 반영하여 데이터를 보정한 후 실제 예측에 의한 결과물 각 모델별로 비교한다.

4.1 각 모델의 결과 비교

1994년 7월부터 1995년 6월까지의 부하와 온도 데이터 중 특수성을 띠는 주말과 특수일의 데이터를 제외한 일별 최대 부하 데이터와 등가화된 최고 및 최저온도 데이터의 관계를 TAR 모델을 이용하여 회귀분석 하였다. 온도의 등가화는 지역별 수요구성비(서울:0.51, 부산:0.22, 대구:0.13, 대전:0.09, 광주:0.05)를 가중평균 하여 산출하였다[10]. 그림 2의 절차에 따라 임계값을 선정하여 표 1에 나타내었다.

일별 최고온도에 대해서 일별 최대부하는 26℃에서 최소 $Pooled SE$, 445.029를 갖으며 따라서 26℃에서 임계값을 갖는다. 마찬가지로 일별 최저온도에 대해서는 14℃에서 임계값을 갖는다.

표 1 임계값 선정 결과

Table 1 Selection of optimal threshold value

임계값 (°C)		RSS1	RSS2	TRSS	Pooled SE
최고 온도	24	15508566	26253024	41761590	450.251
	25	17484629	23803861	41288490	447.694
	26	23601950	17196505	40798455	445.029
	27	29517008	13293705	42810713	455.872
	28	37928256	7224277	45152534	468.174
최저 온도	12	14561747	35104972	49666719	491.020
	13	16667494	33498856	50166351	493.484
	14	23404378	25681720	49086097	488.141
	15	25399780	23885176	49284957	489.129
	16	30872486	20334256	51206742	498.574

이러한 최고 및 최저온도에 대한 임계값에 따라 3장의 식 (2)와 같은 TAR 모델을 수립하였으며, 본 논문에서 제시하는 TAR 모델과의 비교를 위한 선형, 2차 함수 모델의 형태는 다음 식 (4), (5)와 같이 주어진다.

$$L_{peak} = a_0 + a_1 T + e_t, \quad \forall T \quad (4)$$

$$L_{peak} = a_0 + a_1 T + a_2 T^2 + e_t, \quad \forall T \quad (5)$$

여기에서 온도 T 는 식 (2)에서와 마찬가지로 화씨온도를 의미한다.

독립변수가 각각 최고온도, 최저온도일 때를 구분하여 각 모델의 결과인 검정 통계량과 회귀계수를 표 2와 3에 정리하였다. 괄호는 각 회귀계수의 표준오차를 의미한다. 3장에서 언급한 바와 같이 각 독립변수의 종속변수에 대한 설명력을 알아보기 위한 회귀계수 각각의 t -값을 구하려면 회귀계수를 표준오차로 나누면 된다. 각 모델별로 t -값을 살펴보면 회귀계수 a_0 에 대해서는 최고 및 최저온도 모두 TAR 모델의 하위영역에서 가장 큰 값을 갖고 a_1 에 대해서는 2차 함수 모델에서 가장 큰 값을 갖는다.

표 2 각 모델의 결과 비교(최고온도기준)

Table 2 Comparison of coefficients and test statistics for each model as a function of high temp.

	선형	2차 함수	TAR	
			하위	상위
a_0	23269 (246.43)	32216 (573.23)	25710 (140.30)	8239.6 (1276.2)
a_1	9.446 (3.615)	-289.67 (18.524)	-36.702 (2.360)	188.81 (14.605)
a_2	-	2.313 (0.142)	-	-
R^2	0.0318	0.5755	0.6156	0.7524
DW	0.164	0.494	1.85	1.96
SE	923.763	613.11	395.353	559.163
Pooled SE	923.763	613.11	445.029	

표 3 각 모델의 결과 비교(최저온도기준)

Table 3 Comparison of coefficients and test statistics for each model as a function of low temp.

	선형	2차 함수	TAR	
			하위	상위
a_0	23233 (188.20)	28071 (311.40)	25267 (128.26)	14775 (642.24)
a_1	13.426 (3.618)	-209.62 (13.409)	-41.777 (3.182)	139.95 (9.3356)
a_2	-	2.255 (0.133)	-	-
R^2	0.0621	0.6058	0.5609	0.7599
DW	0.168	0.468	1.74	2.28
SE	909.194	590.87	416.372	601.427
Pooled SE	909.194	590.87	488.141	

잔차들의 자기상관 정도를 검정하는 통계량인 DW값은 선형, 2차 함수 모델의 경우 잔차들의 독립성을 가늠할 수 있는 2라는 수치와는 동떨어진 영점대의 값을 갖는다. 그러나 TAR 모델은 어떠한 경우든 상, 하위 영역 모두 2에 가까운 값을 갖으며, 이는 다른 모델과 비교하여 잔차들의 독립성이 더 우수함을 입증한다. 각 모델의 SE(Standard Error)는 최고온도, 최저온도 모두 TAR 모델의 경우가 가장 작게 나타나고, TAR 모델 중에서는 상위영역, 즉 여름철에 SE가 상대적으로 크게 나타난다. 이는 일별 최고온도가 26°C보다 큰 여름철에는 온도 뿐 아니라 그 밖의 다른 기상 요인(강수량, 습도, 불쾌지수 등)에 의해 전력부하의 변동이 크게 나타난 것으로 판단된다.

4.2 각 모델의 예측 결과

본 논문에서 제시하는 TAR 모델은 일전예측(one day-ahead forecasting) 모델로서 TAR 모델의 효용성을 검증하기 위해 예측 기간을 1년으로 설정하여 사례연구를 행하였다. 부하의 예측은 앞 절에서 수립한 선형, 2차 함수, TAR 모델을 이용하고, 부하성장률을 반영하여 데이터를 보정함으로써 이루어졌다. 예측기간은 1995년 7월부터 1996년 6월까지이며 마찬가지로 주말과 특수일을 제외하였다.

TAR 모델의 효용성을 검증하기 위하여 예측 부하의 실제 부하에 대한 평균오차는 식 (6)과 같이 일별 평균오차 절대값의 합을 평균하여 구한 절대평균 백분율오차(MAPE : Mean Absolute Percentage Error, [%])로 정의하였다[3].

$$MAPE[\%] = \frac{1}{N} \sum_{i=1}^N \frac{|Actual(i) - Forecast(i)|}{Actual(i)} \times 100 \quad (6)$$

여기서, N : 예측 날짜의 개수

$Actual$: 실제부하[MW]

$Forecast$: 예측부하[MW]

그림 3은 1988년부터 2000년 5월까지 12년 이상 일주일용

한 단위로 실제 우리나라 전력부하의 총합을 나타낸 것으로 IMF 시기를 제외한 매년 대략 10[%] 정도씩 증가해왔다. 이러한 추세는 인구의 증가와 경제성장률의 상승에 기인한다.

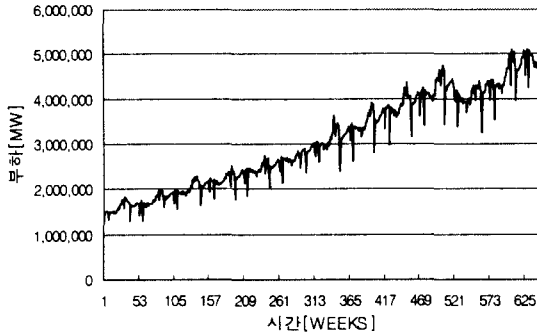


그림 3 부하의 증가 추세
Fig. 3 Weekly load curve for the past 12 years

그림에 나타나듯이 전력부하는 연간 꾸준히 성장해 왔으므로 1년 동안의 부하데이터를 통해 얻은 예측모델을 그대로 이용하여 다음 1년을 예측한다면 온도 외의 자기변수인 과거의 부하데이터가 독립변수에 포함되어 있지 않기 때문에 좋은 결과를 기대할 수 없다. 즉, 부하성장률을 예측모델에 반영해야 한다.

본 논문에서는 부하성장률을 고려하여 예측부하를 보정하기 위해 다음 식 (7), (8)을 이용하였다.

$$k_{L(U)} = \frac{1}{N_y} \sum_y \frac{L_{L(U),y}}{L_{L(U),y-1}} \quad (7)$$

$$L'_{peak} = \begin{cases} k_L L_{peak}, & \text{if } T < A \\ k_U L_{peak}, & \text{if } T \geq A \end{cases} \quad (8)$$

여기서, $k_{L(U)}$: 하위 및 상위영역의 평균 부하성장률
 $L_{L(U),y}$: y년 하위 및 상위영역의 일평균 부하량 [MW]
 L'_{peak} : 부하성장률에 따라 보정된 일별 최대부하 [MW]

표 4 하위 및 상위영역의 부하성장률
Table 4 Load growth rate of each regime

	최고온도	최저온도
하위영역	1.1193	1.1209
상위영역	1.1161	1.114

식 (7)의 평균 부하성장률을 구하기 위한 기간은 1988년부터 1994년까지로 정하였으며, 하위 및 상위영역의 평균 부하성장률의 결과는 표 4와 같다. 식 (2)의 TAR 모델을 통해 얻은 최고 및 최저온도에 대한 각 영역의 예측부하

L_{peak} 에 표 4의 평균 부하성장률을 식 (8)과 같이 적용하여 보정된 TAR 모델의 부하예측 결과를 얻을 수 있었으며, 또한 상/하위 영역을 따로 구분하지 않는 선형 및 2차 함수 모델의 경우는 영역을 따로 구분하지 않고 y년의 일평균 부하량 L_y 만을 식 (7)에 적용하여 구한 평균 부하성장률 k (1.1186)를 이용하여 보정된 부하예측 결과를 얻을 수 있었다. 사례연구 4.1의 예측 모형에서 구한 각 모델별 예측부하의 보정을 위해 식 (7)과 (8)을 이용하여 1995년 7월부터 1996년 6월까지 주말과 특수일을 제외한 219일의 예측 결과를 표 5와 6에 정리하였다. 여기에서 각 모델의 MAPE를 비교하기 위해서 Pooled MAPE를 도입하였고, 그 값은 Pooled SE를 구하는 방법과 마찬가지로 상, 하위 영역 각각의 MAPE를 각각의 예측치 개수에 따라 평균한 값이다. 표 5와 6에서 MAPE는 각 모델별로 비교할 때 TAR 모델에서 가장 작게 나타나고, TAR 모델에 있어서는 하위 영역에 비해 최고온도와 최저온도 모두 상위영역에서 그 값이 거의 3배정도 크게 나타난다.

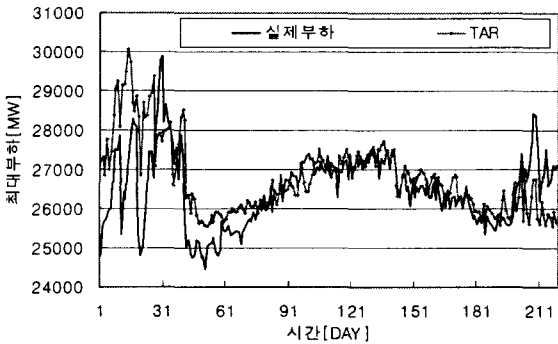
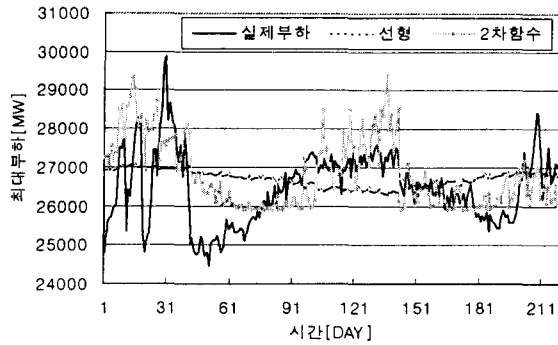
표 5 예측 결과 비교[%](최고온도기준)
Table 5 Comparison of forecasting results for each model as a function of high temp.[%]

모델	최소오차	최대오차	MAPE	Pooled MAPE	
선형	0.0261	9.8112	3.1353	3.1353	
2차 함수	0.0022	12.671	2.8425	2.8425	
TAR	하위	0.0056	7.1068	1.4771	2.4683
	상위	0.5248	14.965	5.095	

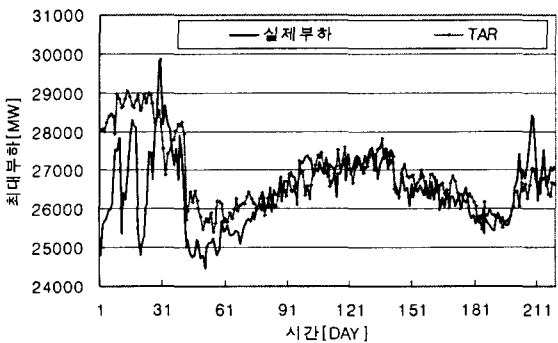
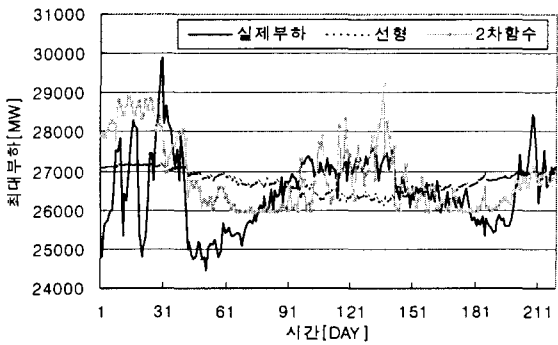
표 6 예측 결과 비교[%](최저온도기준)
Table 6 Comparison of forecasting results for each model as a function of low temp.[%]

모델	최소오차	최대오차	MAPE	Pooled MAPE	
선형	0.0327	9.7785	3.1916	3.1916	
2차 함수	0.0110	14.337	2.786	2.786	
TAR	하위	0.0113	4.9165	1.3233	2.4563
	상위	0.0830	15.019	4.3127	

그림 4는 본 논문에서 예측 기간으로 잡은 219일 간의 각 모델별 예측 결과이다. 여름철, 즉 주말과 특수일을 제외한 1995년 7월부터 9월에 해당하는 1~60여 일은 다른 시기와 비교하여 오차가 다소 나타나지만, 독립변수를 최고 및 최저온도로 설정한 (a), (b) 모두에서 TAR 모델은 다른 모델에 비해 실제부하와 가장 근접하는 예측 결과를 나타낸다. 여름철 부하 즉 상위영역은 표준편차와 마찬가지로 하위영역에 비해 예측력이 떨어지는 결과를 보이는데 그 이유는 최고 및 최저온도가 반드시 최대부하와 일치하지 않기 때문이며 상위영역에서 좀더 정확한 예측 결과를 얻기 위해서는 온도의 자기변수, 즉 과거의 온도데이터나 다른 기상요인을 독립변수에 추가하여 새로운 예측모델을 개발해야 한다고 판단된다.



(a) 최고온도기준



(b) 최저온도기준

그림 4 부하 예측 결과
 Fig. 4 Forecasted daily peak loads for each model as a function of (a) high temp. and (b) low temp.

5. 결 론

본 논문은 임계값을 기준으로 상위영역과 하위영역에서 각각 정(正)의 상관관계와 부(負)의 상관관계를 보이는 부하와 온도의 특성을 이용하여 일별 최대부하와 일별 최고 및 최저온도의 관계를 TAR 모델을 도입하여 회귀분석 하였으며 이를 선형 모델, 2차 함수 모델과 비교하였다. 비교는 검정 통계량과 MAPE로 이루어졌으며 두 경우 모두에서 TAR 모델은 다른 모델에 비해 우수한 결과를 나타내었고 이로써 TAR 모델 도입의 필요성을 검증하였다. 또한 실제 예측 단계에서는 평균 부하성장률만을 고려한 데이터 보정만으로도 예측력을 높일 수 있었다. 향후 자기변수 즉 과거의 부하데이터를 독립변수에 추가하기 위한 방법으로 더미변수(dummy variable)를 도입하여 주말과 특수일의 부하데이터도 입력에 추가한다면 평상일과 매우 다른 부하 패턴을 나타내는 주말과 특수일의 부하도 예측할 수 있을 것이며, 예측모델에 온도 외의 다른 기상요인을 추가하는 방법이나 혹은 부하성장률을 정확하게 반영할 수 있는 방법 등을 통하여 더 좋은 예측 결과를 기대할 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] 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.
- [3] M. Daneshdoost, M. Lotfalian and G. Bumroonggit, "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.
- [4] 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.
- [5] 박영진, 심현경, 왕보현, "뉴로-퍼지 모델을 이용한 단기 전력 수요 예측시스템", *전기학회 논문지*, 49권 3호, pp. 107-117, 2000.
- [6] E. H. Barakat and J. M. Al-Qasem, "Methodology for Weekly Load Forecasting", *IEEE Trans. on Power Systems*, Vol. 13, No. 4, pp. 1548-1555, 1998.
- [7] Tong, H., "Threshold Models in Nonlinear Time Series Analysis", *Lecture Notes in Statistics* 21, Berlin: Springer Verlag, 1983.
- [8] Tong, H., "Nonlinear Time Series Analysis: A Dynamical System Approach", New York: Oxford Science Publications, 1990.
- [9] Yunho Lee, "Nonlinear Dynamic Models: Applications of Chaos Theory and Threshold Autoregressive Model to Agricultural Prices", Thesis, Texas A&M University, 1998.
- [10] 황갑주, 김광호, 김성학, "주간수요예측 전문가 시스템 개발", *전기학회 논문지*, 48권 4호, pp. 365-370, 1999.

저 자 소 개



이 경 훈 (李 景 勳)

1975년 1월 2일 생. 2000년 한양대 전기공학과 졸업. 2000년~현재 동 대학원 전기공학과 석사과정

Tel : 02-2290-0347, Fax : 02-2297-1569

E-mail : ceilrule@ihanyang.ac.kr



김 진 오 (金 鎮 吾)

1956년 1월 17일 생. 1980년 서울대 전기공학과 졸업. 1983년 동 대학원 전기공학과 졸업(석사). 1991년 Texas A&M Univ. 전기공학과 졸업(공학박사). 현재 한양대 공과대 전자전기컴퓨터공학부 부교수

Tel : 02-2290-0347, Fax : 02-2297-1569

E-mail : jokim@hanyang.ac.kr



이 윤 호 (李 允 昊)

1954년 6월 6일 생. 1978년 연세대 경영학과 졸업. 1985년 동 대학원 경제학과 졸업(석사). 1998년 Texas A&M Univ. 농업경제학과 졸업(박사). 현재 (주)데이타네오넷 대표

E-mail : yh14265@dreamwiz.com