

## 한국 최대 전력량 예측을 위한 통계모형

윤상후<sup>1,a</sup>, 이영생<sup>a</sup>, 박정수<sup>a</sup>

<sup>a</sup>전남대학교 통계학과

### 요약

한국의 경제규모가 꾸준히 커감에 따라 가정, 건물, 공장 등에서 필요로 하는 전력량이 지속적으로 증가하고 있다. 전력공급의 안정화를 위해서는 최대전력량보다 전력공급능력이 높아야 한다. 월별 최대전력량을 잘 설명할 수 있는 통계모형을 찾기 위해 Winters 모형, 분해 시계열모형, ARMA 모형, 설명변수를 통해 추세성분과 계절성분을 교정한 모형을 살펴보았다. 모형의 예측력 비교 기준으로 모형적합으로부터 구한 RMSE와 MAPE가 사용되었다. 여름철 최대전력량을 예측하기 위해 평균기온과 열대야 일수를 설명변수로 갖는 시계열 모형이 가장 우수하였다. 아울러 외부요인을 갖는 극단분포 모형을 이용한 분석을 시도하였다.

**주요용어:** Winters 모형, 분해시계열 모형, ARMA 모형, 설명변수가 있는 시계열 모형, 일반화 극단분포.

### 1. 서론

국내 모든 산업분야의 발달로 인하여 전력에너지 사용량이 폭발적으로 증가되어 국가전체의 전력공급 안정성이 크게 위협받고 있는 실정이다 (전국전력노동조합, 2003). 국가적인 에너지정책과 전력수급의 안정화를 위해 수요예측의 불확실성, 전원 입지난의 가중문제, 장기적인 전원관리를 위한 설비확충 등을 해결할 필요성이 있다. 아울러 안정적인 전력공급이 유지될 수 있도록 발전소의 돌발사고 등에 대비해 비축하고 있는 예비 전력의 수준인 전력공급 예비율을 높이는 것이 중요하다. 이 때문에 최대전력량에 대해 경험적 기법이나 다양한 통계적 기법을 통해 단기 및 장기수요를 예측하는 시스템 개발 필요성이 높아지고 있다 (한국전력공사, 1998, 2000).

본 논문의 목적은 한국 최대전력량 자료에 적합한 통계적 모형을 찾고 앞으로 3년간 최대전력량이 얼마나 될지 예측하는데 있다. 전력예측에 관련된 국내의 연구는 남봉우 등 (2008), 송경빈 (2007) 등이 다중 회귀분석법과 퍼지 최소자승 선형회귀모형을 이용하였고, 국외의 연구는 Azadeh (2007), Nowicka-Zagrajek과 Weron (2002), Samer 등 (2001), Tsekouras 등 (2007) 등이 신경망구조와 ARMA 모형 등을 고려하였다. 본 연구에서 이용한 모형으로는 Winters 모형, ARMA 모형, 설명변수가 포함된 시계열 모형이고 분석은 SAS 9.1과 R을 이용하였다. 추가적으로 최대전력량과 같은 극값이 지니는 특성을 활용한 외부요인을 갖는 극단분포 모형을 적용하였다.

### 2. 분석 자료

본 연구에 사용된 자료는 에너지경제연구원 (2008)과 한국전력공사 (2008)의 통계에서 수집하였다. 수집된 자료는 1991년 1월부터 2007년 12월까지 월별 최대전력량이다. 그림 1은 시간에 따른 월

본 연구는 한국학술진흥재단의 2005년 지역대학우수과학자 연구비 지원에 의해 수행되었음 (KRF-2005-205-C00072).

<sup>1</sup> 교신저자: (500-757) 광주시 북구 용봉동 300, 전남대학교 통계학과, 박사과정. E-mail: statstar96@gmail.com

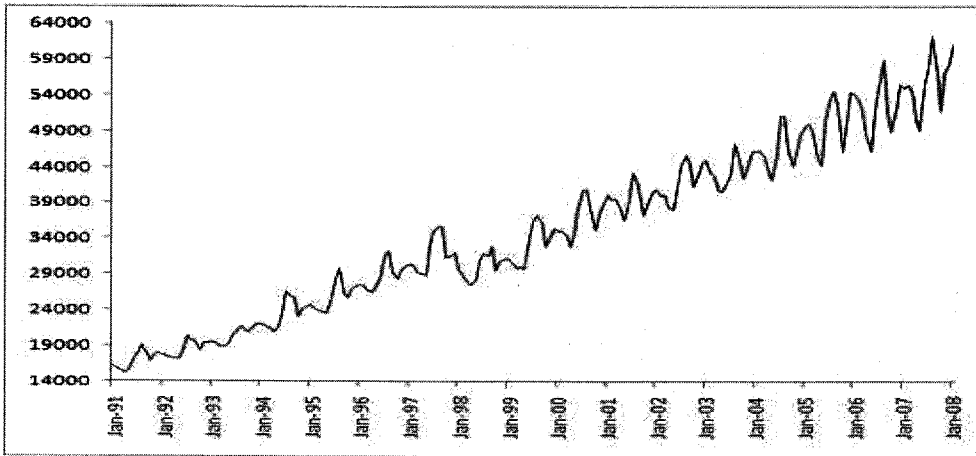


그림 1: 시간에 따른 최대전력량의 시계열 도표 (단위: MW)

별 최대전력량의 시계열도표이다. 그림 1을 보면 최대전력량이 시간의 흐름에 따라 뚜렷한 추세성과 계절성을 띄고 있다. 91년 월별최대전력량은 약 15,000MW에서 최근 약 62,000MW정도로 18년간 약 4.2배 성장하였다. 계절변동성은 90년대 중반까지 크지 않지만 점점 증가하여 최근 그 변동성이 커지고 있다. 추세성의 경우 IMF 시기를 제외하고 시간이 흐름에 따른 추세성을 일정해 보인다. 따라서 앞으로의 예측은 변동성이 유지되면서 증가추세가 계속 될 것이다. 본 논문에서는 최근 패턴과 비슷한 IMF 시기 이후의 자료인 2000년 1월부터 2006년 12월까지 자료를 사용하여 모형에 적합시켰다. 또한 모형의 예측력을 비교하기 위해 2007년 1월부터 2007년 12월까지의 자료를 사용하였다. 모형이 좋은지 나쁜지를 판정하기 위한 척도로 아래와 같은 제공근 평균제곱오차 (RMSE: Root Mean Squared Error)와 절대평균 백분위 예측오차 (MAPE: Mean Absolute percentage Prediction Error)를 사용하였다:

$$RMSE = \left( \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \right)^{1/2}, \quad (2.1)$$

$$MAPE = \frac{100}{m} \sum_{i=1}^m \left| \frac{y_i - \hat{y}_i}{y_i} \right|. \quad (2.2)$$

### 3. 월별 최대전력량에 적합한 통계모형

추세성과 계절성을 고려한 통계모형으로 Winters모형, 분해시계열 모형, ARMA 모형, 설명변수를 포함한 추세성분과 계절성분이 고려된 모형이 고려되었다. 각 모형에서 종속변수는 월별 최대전력량이고 독립변수는 시간이다.

#### 3.1. Winters 계절지수 모형

Winters 모형 (Winters, 1960)은 가법 계절요인(additive seasonal factor)과 승법 계절요인(multiplicative seasonal factor)으로 계절적 변동에 대처하고 시간추세를 고려한 후 과거의 관측값들에 지수적으로 감소하는 가중값을 주어 예측값을 구한다. 모형들의  $n+1$  시점에서의 종속변수의 모형은 추세성분  $T_{n+1}$ 과 계절주기  $s$ 를 가지는 계절성분  $S_{n+1}$  그리고 불규칙 성분인 오차항  $I_{n+1}$ 으로 구성된다. 여기서

각 성분들을 평활법에 의해 추정한 후 이를 이용하여 예측값을 구하면 다음과 같다:

$$\begin{aligned} \hat{T}_{n+1} &= w_1(Z_{n+1} - \hat{S}_{n+1-s}) + (1 - w_1)(\hat{T}_n + \hat{\beta}_{1,n}), \\ \hat{\beta}_{1,n+1} &= w_2(\hat{T}_{n+1} - \hat{T}_n) + (1 - w_2)\hat{\beta}_{1,n}, \\ \hat{S}_{n+1} &= w_3(Z_{n+1} - \hat{T}_{n+1}) + (1 - w_3)\hat{S}_{n+1-s}. \end{aligned} \tag{3.1}$$

평활상수들은 1-시차후 예측오차합(SSE)를 최소로 하는 가중치로 결정되며 가법모형은  $(w_1, w_2, w_3) = (0.2, 0.0, 0.4)$  그리고 승법모형은  $(w_1, w_2, w_3) = (0.3, 0.0, 0.5)$ 가 선택되었다. Winters 지수평활 모형들의 잔차는 가법모형과 승법모형 모두 2003년 7월, 2004년 7월, 2005년 10월, 2006년 9월 등의 순으로 큰 차이를 보이고 있어 여름철 예측력이 좋지 못하다(구체적 결과는 생략).

### 3.2. 분해시계열 모형과 ARMA 모형

최대전력량 자료는 추세성과 계절성을 지니고 있어 정상성 자료로 볼 수 없다. 시점  $t$ 에서 종속변수의 모형  $Z_t$ 은 추세성분  $T_t$ 과 계절주기  $s$ 를 가지는 계절성분  $S_t$ , 그리고 불규칙 성분인 오차항  $I_t$ 으로 구성되어 있으므로, 추세성분  $T_t$ 과 계절성분  $S_t$ 을 제거한다면 불규칙 성분이 정상 시계열모형이 된다. 위의 요인들의 결합방식에 따라 분해시계열 모형과 ARMA (Auto-Regressive Moving Average) 모형으로 구분할 수 있다.

$$\text{분해시계열 모형 : } Z_t = T_t + S_t + I_t$$

$$\text{ARMA 모형 : } Z_t = T_t \times S_t \times I_t$$

시간을 독립변수로 가지는 단순 회귀분석을 통해 추세성분  $T_t$ 를 제거하고 지시함수를 이용한 자기회귀오차 모형을 통해 계절성분  $S_t$ 를 제거하면 불규칙성분  $I_t$ 가 정규 시계열패턴을 갖는 시계열 자료가 얻어지므로 평균수준이 0인 정상 ARMA( $p, q$ ) 모형에 적합시킬 수 있다. 모수  $p$ 와  $q$ 는 AIC (Akaike Information Criterion) 값을 최소로 하는 값을 선택하였다. 정상 시계열  $Z_t$ 는 다음 수식으로 표현된다 (조신섭과 손영숙, 2002).

$$\phi(B)Z_t = \theta(B)\epsilon_t, \tag{3.2}$$

$$\phi(B) = 1 - \phi_1(B) - \phi_2(B)^2 - \dots - \phi_p(B)^p \quad : \text{AR 특성함수,}$$

$$\theta(B) = 1 - \theta_1(B) - \dots - \theta_q(B)^q \quad : \text{MA 특성함수.}$$

여기서  $t$ 는 시간을 나타내는 지표이며,  $\epsilon_t$ 는 서로 독립인  $N(0, \sigma^2)$ 을 따르는 오차항으로 백색 잡음이다. 분해시계열 모형에서 잔차는 2003년 7월, 2007년 7월, 2006년 7월 순으로 큰 차이를 보이고, ARMA 모형에서 잔차는 2003년 7월 2007년 10월, 2007년 7월 순으로 큰 차이를 보이고 있어 Winters 모형과 마찬가지로 여름철 예측력이 좋지 못하다(구체적 결과는 생략).

### 3.3. 설명변수를 포함한 시계열 모형

Winters 모형과 ARMA 모형은 단변량 시계열자료의 분석이다. 따라서 분석의 대상이 되는 시계열과 밀접한 관계가 있는 변수를 이용할 수 없다는 단점이 있다. 일반적으로 전력량은 경제성장지표인 GDP, 인구수, 전력가격, 날씨 등에 영향을 받는다 (Samer 등, 2001). 실제 단변량 시계열 예측결과를 보면 여름철인 7월, 8월, 9월의 최대전력량 예측력이 좋지 못하다. 여름철 전력소비는 기온에 많은 영향을 받으므로, 월별 평균기온과 열대야 일 수를 고려하였다. 그 외에도 매년 전력소비량의 증가율을

표 1: 연간 전력소비량과 성장률 및 예측값 (단위: MW)

Year	총에너지소비량(GWH)	연간 성장률	정규화 변수
2000	266,398	1.000	1.513
2001	285,192	1.071	1.414
2002	306,473	1.075	1.315
2003	322,429	1.052	1.250
2004	342,148	1.061	1.178
2005	364,639	1.066	1.106
2006	381,180	1.045	1.058
2007	403,124	1.058	1.000
2008	(427,718)	(1.061)	(0.943)
2009	(453,812)	(1.061)	(0.888)
2010	(481,498)	(1.061)	(0.837)

표 2: 2007년 1년간 각 모형별 RMSE (단위: MW)

모형		1분기	2분기	3분기	4분기	전체
Winters 모형	가법모형	665.63	812.67	1686.36	1203.88	1161.53
	승법모형	864.87	1025.83	1569.58	1291.55	1217.79
설명변수 없음	분해시계열모형	621.98	1420.96	2023.92	1617.74	1510.92
	ARMA모형	737.86	998.31	1495.60	1558.72	1245.74
설명변수 포함	분해시계열모형	671.05	973.09	369.41	1168.46	851.32
	ARMA모형	617.33	981.80	348.77	1132.60	829.07

고려하여 년 간 총전력 소비량을 표준화하였다 (Abdel-Aal, 2008). 최종적으로 선택된 설명변수는 평균기온, 표준화된 총전력 소비량, 열대야 일수이다. 날씨 자료는 기상청 (2008)의 통계를 이용하였다. 추세성분은 시간외에 설명변수로 연간 전력소비 증가률(표 1)을, 여름철 계절성분(7월, 8월, 9월)은 설명변수로 월평균 기온과 열대야 일수를 고려하였다.

$$\text{추세성분: 정규화된 } Z_t = T_t + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2),$$

$$\text{계절성분(7월, 8월, 9월): } S_t = I_{t, \text{month}} + A_t + D_t + v_t$$

$$v_t = \epsilon_t - \theta_1 v_{t-1} - \dots - \theta_m v_{t-m},$$

$I$ : 지시함수,  $A$ : 월평균 기온,  $D$ : 열대야 일수.

여기서 표준화된 총전력 소비량은 정규화 변수(normalization factor)와 연간 총 에너지 소비량의 곱이다. 전력수요 예측기술의 수준을 나타내는 평균예측오차는 약 1.3%이며, 최대오차는 약 3%를 나타낸다.

설명변수가 없는 모형과 비교하면 년 간 전력소비량의 증가율을 고려하여 전반적인 예측력이 향상되었고, 여름철 전력소비량에 큰 영향을 미치는 날씨요인인 열대야 일수와 평균온도를 고려하여 여름철(7월, 8월, 9월) 예측력도 향상되었다. 분해시계열 모형이 ARMA 모형에 비해 예측력이 조금 좋았다.

#### 3.4. 모형 비교와 예측

위에서 고려한 모형들을 RMSE 기준으로 비교하였다. 표 2는 2007년 1년간 RMSE를 세가지 모형에 대해 요약한 결과이다. 전반적으로 설명변수를 포함한 모형이 다른 모형들에 비해 좋다. 세부적으로 확인하면, Winters의 가법모형은 2007년 2분기만 다른 모형들에 비해 예측력을 높였지만 3분기의 예측력이 좋지 못했다. 설명변수를 포함한 ARMA 모형은 설명변수가 없는 ARMA 모형에 비해 전반

표 3: 설명변수가 포함된 ARMA 모형의 연간 전력 예측값과 95% 신뢰구간 (단위: MW)

Date	예측값	95%하한	95%상한	Date	예측값	95%하한	95%상한
Jan-09	61466	58577	64367	Jan-10	65122	62061	68195
Feb-09	60971	58105	63848	Feb-10	64597	61561	67646
Mar-09	60133	57307	62971	Mar-10	63710	60715	66717
Apr-09	57230	54540	59931	Apr-10	60634	57784	63496
May-09	56426	53774	59090	May-10	59783	56973	62604
Jun-09	61292	58411	64185	Jun-10	64937	61885	68002
Jul-09	65994	62892	69109	Jul-10	69939	66652	73240
Aug-09	68118	64917	71333	Aug-10	72199	68806	75607
Sep-09	63825	60825	66837	Sep-10	67666	64486	70860
Oct-09	58896	56128	61676	Oct-10	62399	59467	65345
Nov-09	61959	59047	64883	Nov-10	65644	62559	68742
Dec-09	63989	60981	67009	Dec-10	67795	64608	70995

적인 예측력이 향상되었고 특히 3분기 예측력이 좋았다. Winters 모형들과 설명변수를 포함한 ARMA 모형을 비교하면 2분기에 Winters 모형의 예측력이 높는데 반해 3, 4분기에는 ARMA 모형이 높았다.

예측은 분석된 통계적 모형에 자료를 외삽(extrapolation)시켜서 미래를 추측하는 과정이다. 전력 소비량 증가율은 2000년 이후로 비슷한 패턴을 보이기 때문에 2000년부터 2007년까지 자료를 토대로 앞으로 3년간 연간성장률이 동일하다고 가정하여 연도별 전력 소비량 예측값과 95% 신뢰구간을 구하면 표 3와 같다. 여름철(7월, 8월, 9월) 평균기온은 1980년부터 2007년까지 서울지역의 월별 평균기온을 토대로 단순회귀분석을 이용하여 연도별 평균기온 예측치를 계산하였고, 그 식은 다음과 같다.

$$y\text{년 } 7\text{월 평균기온} = 0.026 \times \text{year} - 27.038$$

$$y\text{년 } 8\text{월 평균기온} = 0.038 \times \text{year} - 50.21$$

$$y\text{년 } 9\text{월 평균기온} = 0.059 \times \text{year} - 96.538.$$

여름철 열대야 일 수는 2000년 이후에 서울지역에 발생했던 월 평균 열대야 일 수가 미래에도 똑같이 발생한다는 가정하였다(7월: 2.34일, 8월: 4.03일, 9월: 0.07일). 최종적으로 2000년 1월부터 2007년 12월까지의 자료를 통해 앞으로 3년간 예측값과 95% 신뢰구간을 구하면 그림 2이다. 설명변수를 포함한 ARMA 모형의 경우 각각 변수들에 대한 불확실성을 고려하지 못하고 분석하였으므로 신뢰구간이 과소추정되었음을 주의해야 한다. Winters 가법모형과 설명변수를 포함한 ARMA 모형 간 앞으로 3년간 예측값을 비교하면 그림 3와 같다. 전반적으로 Winters 가법모형의 월별 최대전력량 수요예측값이 ARMA 모형에 비해 낮게 예측되고 있다.

#### 4. 외부 요인을 갖는 극단분포 모형

자연현상이나 기타 다양한 분야에서 사용되는 극값들은 모집단의 분포와 무관하게 자료의 수가 증가하면 일반화 극단분포(generalized extreme value distribution)에 수렴하며, 모양 모수(shape parameter)  $\xi$ 의 값에 따라 Gumbel, Frechet, negative Weibull 분포가 된다 (Coles, 2001). 최대전력량은 극값 자료이므로 1971년부터 2007년까지의 일일 최대전력량 중 연간 최대값들을 사용하여 일반화 극단분포에 적합시켰다. 일반화 극단분포의 누적분포함수는 다음과 같다.

$$G(x) = \exp \left\{ - \left[ 1 + \xi \left( \frac{x - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}, \quad 1 + \xi(x - \mu)/\sigma > 0 \tag{4.1}$$

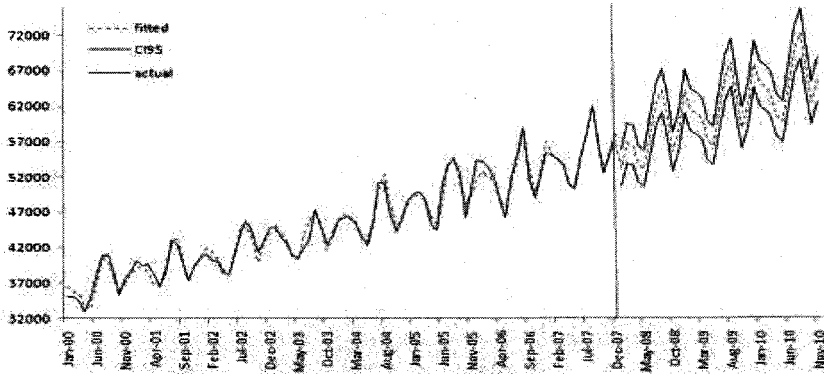


그림 2: 설명변수를 포함한 ARMA 모형의 앞으로 3년간 월별 최대전력량과 95% 신뢰구간 (단위: MW)

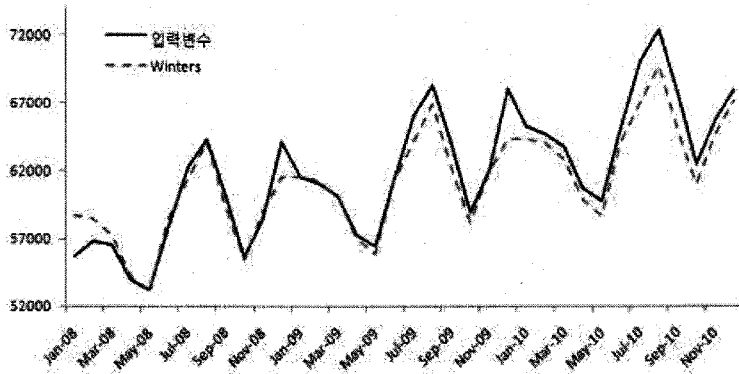


그림 3: Winters 가법모형과 설명변수를 고려한 ARMA 모형 예측값 비교 (단위: MW)

식 (4.1)에서  $\mu$ 는 위치 모수,  $\sigma$ 는 척도 모수(scale parameter),  $\xi$ 는 모양 모수이다. 식 (4.1)에서  $\xi=0$ 이면 다음과 같이 표현되고

$$G(x) = \exp \left\{ - \exp \left[ - \left( \frac{x - \mu}{\sigma} \right) \right] \right\}, \quad -\infty < x < \infty \quad (4.2)$$

이를 Gumbel 분포라고 한다. 이영생 등 (2008)에 따르면 연간 한국 최대전력량은 Gumbel 분포를 따르고, Gumbel 분포의 위치 모수  $\mu$ 는 경제성장지표인 GDP와 1년 전의 연간 최대전력량( $X_{t-1}$ )에 영향을 받으므로 다음과 같이 모형을 세우고 모수를 추정해 보았다 (단, IMF시기인 1998년은 모형적합에 적절치 않아 제외시켰다):

$$\mu_t = a + b \times X_{t-1} + c \times \text{GDP}_t.$$

Coles (2001)의 R 프로그램을 이용하여 최우추정법으로 모수를 추정한 결과  $a = 308269$ ,  $b = 1.04$ ,  $c = 0.76$ ,  $\sigma = 6337758$ 이었고, 실제자료가 모형에 잘 적합되었다(그림 4). 미래 예측을 위해서 일단 GDP를 과거 자료에 대해 선형회귀식을 적용하여 구했다. 즉, GDP 예측값은  $\text{GDP}_{2009} = 853120$ ,  $\text{GDP}_{2010} = 884134$  십억원이다. 이를 바탕으로 예측한 미래 연간 최대전력량은  $X_{2009} = 69863$ ,  $X_{2010} = 73149$  MW이다. 이들은 앞 절에서 구한 시계열 모형을 이용한 예측값보다 높게 나왔다.

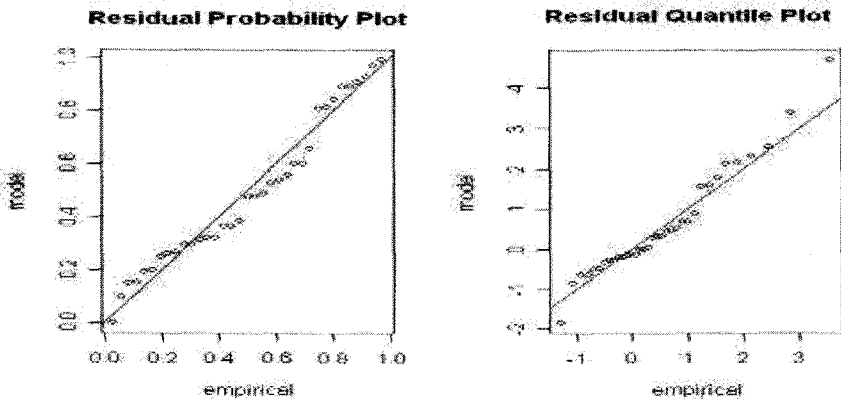


그림 4: 외부요인을 갖는 극단분포(Gumbel)의 P-P plot과 Q-Q plot.

### 5. 결론 및 토의

본 논문에서는 월별 최대전력량을 예측하기 위하여 Winters 모형, 분해시계열 모형, ARMA 모형, 설명변수가 포함된 시계열 모형을 고려하였고, 이 중 설명변수를 포함한 모형이 다른 모형에 비하여 예측력이 우수함을 보였다. 시계열자료의 분석 및 예측에 있어 설명변수와 외부요인의 추가는 통계적 모형의 설명력과 예측력을 보완할 수 있다. 또한 극값에 대해 외부요인을 갖는 극단(Gumbel)분포를 적합시켰다. 앞으로의 월별 최대전력량은 변동성이 유지되면서 증가추세가 계속 될 것이라는 가정을 하였지만, 최근 유가상승 및 세계적인 경제침체의 영향을 받을 수 있어서 향후 추이가 주목된다. 미래 연구과제는 설명변수들의 상태공간모형을 이용한 상태전이 방정식을 통해 설명변수들을 좀 더 정확히 예측한 뒤, 이를 본 논문에서 고려한 모형에 입력변수로 활용하는 것이다.

### 참고 문헌

기상청 (2008). 지난 날씨 통계자료. [http://www.kma.go.kr/sfc/sfc\\_03\\_02.jsp](http://www.kma.go.kr/sfc/sfc_03_02.jsp)

남봉우, 송경빈, 김규호, 차준민 (2008). 다중회귀분석법을 이용한 지역전력수요예측 알고리즘, *Journal of the Korean Institute of Illuminating and Electrical Installation Engineers*, **22**, 63-67.

송경빈 (2007). 전력수요예측을 위한 다양한 퍼지 최소자승 선형회귀 모델, *Journal of the Korean Institute of Illuminating and Electrical Installation Engineers*, **21**, 61-67.

에너지경제연구원 (2008). 에너지 월별통계. <http://www.keei.re.kr>

이영생, 박정수 (2008). 극단분포를 활용한 최대전력 소비량 분석, 한국통계학회 추계학술대회, 중앙대학교.

전국전력노동조합 (2003). 전력산업구조개편 정책의 문제점과 대안, 전국전력노동조합. <http://power.inochong.org/datamenu/DataView.html?SID=5>

조신섭, 손영숙 (2002). <SAS/ETS를 이용한 시계열분석>, 을국출판사, 서울.

한국전력공사 전기연구원 (1998). A study on the improvement of forecasting method for regional demand and load characteristics, 내부 연구문서, 한국전력공사.

한국전력공사 전력연구원 (2000). 신경회로망을 이용한 배전부하 수요예측 연구, 최종 연구보고서, 한국전력공사, 대전.

한국전력공사 (2008). 전력통계. <http://www.kepco.co.kr>

Abadel-Aal, R. E. (2008). Univariate modeling and forecasting of monthly energy demand time series using abductive and neural networks, *Computers & Industrial Engineering*, **54**, 903-917.

- Azadeh, A., Ghaderi, S. F. and Sohrabkhani, S. (2007). Forecasting electrical consumption by integration of neural network, time series and ANOVA, *Applied Mathematics and Computation*, **186**, 1753–1761.
- Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*, Springer Verlag, London.
- Nowicka-Zagrajek, J. and Weron, R. (2002). Modeling electricity loads in California: ARMA models with hyperbolic noise, *Signal Processing*, **82**, 1903–1915.
- Samer, S., Elie, B. and Gerge N. (2001). Univariate modeling and forecasting of energy consumption: The case of electricity in Lebanon, *Energy*, **26**, 1–14.
- Tsekouras, G. J., Dialynas, N. D. and Hatzigiorgios, S. K. (2007). A non-linear multivariable regression model for midterm energy forecasting of power systems, *Electric Power Systems Research*, **77**, 1560–1568.
- Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages, *Management Science*, **6**, 324–342.

2008년 10월 접수; 2009년 1월 채택



# Statistical Modeling for Forecasting Maximum Electricity Demand in Korea

Sang Hoo Yoon<sup>1,a</sup>, Young Saeng Lee<sup>a</sup>, Jeong Soo Park<sup>a</sup>

<sup>a</sup>Dept. of Statistics, Chonnam National Univ.

---

## Abstract

It is necessary to forecast the amount of the maximum electricity demand for stabilizing the flow of electricity. The time series data was collected from the Korea Energy Research between January 2000 and December 2006. The data showed that they had a strong linear trend and seasonal change. Winters seasonal model, ARMA model were used to examine it. Root mean squared prediction error and mean absolute percentage prediction error were a criteria to select the best model. In addition, a nonstationary generalized extreme value distribution with explanatory variables was fitted to forecast the maximum electricity.

**Keywords:** RMSE, Winters seasonal model, ARMA model, additional explanatory variable, generalized extreme value distribution.

---

This paper was supported by The Korean Research Foundation, 2005(KRF-2005-202-C00072).

<sup>1</sup> Corresponding author: Ph. D. student, Department of statistics, Chonnam National University, Gwangju 500-757, Korea.  
E-mail: statstar@hanmail.net