Journal of the Korean Institute of Industrial Engineers
Vol. 28, No. 1, pp. 63-75, March 2002.

Winters' Multiplicative Seasonal Model에 의한 월 최대 전력부하의 단기예측

양문희[1†] · 임상규[2]

[1]단국대학교 산업공학과/[2]경상대학교 산업시스템공학부

Short-Term Forecasting of Monthly Maximum Electric Power Loads Using a Winters' Multiplicative Seasonal Model

Moonhee Yang[1] · Sanggyu Lim[2]

[1]Departmemt of Industrial Engineering, Dankook University, Cheonan, 330-714
[2]Departmemt of Industrial Systems Engineering, Gyeongsang National University, Jinjoo, 660-701

To improve the efficiency of the electric power generation, monthly maximum electric power consumptions for a next one year should be forecasted in advance and used as the fundamental input to the yearly electric power-generating master plan, which has a greatly influence upon relevant sub-plans successively. In this paper, we analyze the past 22-year hourly maximum electric load data available from KEPCO(Korea Electric Power Corporation) and select necessary data from the raw data for our model in order to reflect more recent trends and seasonal components, which hopefully result in a better forecasting model in terms of forecasted errors. After analyzing the selected data, we recommend to KEPCO the Winters' multiplicative model with decomposition and exponential smoothing technique among many candidate forecasting models and provide forecasts for the electric power consumptions and their 95% confidence intervals up to December of 1999. It turns out that the relative errors of our forecasts over the twelve actual load data are ranged between 0.1% and 6.6% and that the average relative error is only 3.3%. These results indicate that our model, which was accepted as the first statistical forecasting model for monthly maximum power consumption, is very suitable to KEPCO.

Keywords: electricity, electric power forecasting, monthly maximum power consumption

1. Introduction

KEPCO (Korea Electric Power Corporation) has developed several forecasting models including ARIMA models and Jang's Model (KEPCO, 1999). The forecasting models for electrical consumptions currently used by KEPCO may be classified into three models depending on forecast periods; a week-term, a short-term, and finally a long-term forecasting model. In a week-term model, we forecast hourly average power consumptions one week ahead of time, which results in a daily maximum and a weekly maximum power consumption. In a short-term model, we forecast twelve monthly maximum power consumptions a year ahead of time. In a long-term model. we forecast the maximum yearly peak load for the following five or ten years usually based upon tentative economic scenarios.

For convenience we define the hourly load data, which KEPCO has recorded for forecasting and

The research was conducted by the research fund of Dankook University in 2000.
† Corresponding author : Professor Moonhee Yang, Department of Industrial Engineering, Dankook University, San 29 Anseo-Dong Cheonan, Choongnam, 330-714 Korea; Fax +82-041-550-3570; e-mail myfriend@anseo.dankook.ac.kr
Received August 2001, accepted December 2001. This paper was with the authors 2 weeks for 2 revisions.

management since 1970, not as the maximum load for an hour but as the average load for an hour. Strictly speaking, daily maximum load which we define in this paper is not an actual maximum load during the day but just the maximum value among 24 average hourly loads, which is clearly a little less than an actual maximum load during the day. In the similar way, we define a monthly maximum load is the maximum value among 28 to 31 daily maximum loads and a yearly maximum load is the maximum value among 12 monthly maximum loads.

To improve the efficiency of the electric power generation, KEPCO has made the short-term forecasts of monthly maximum electric power consumptions for next one year in advance using a simple model in which they are just the sum of the previous ones and the increments multiplied by the corresponding forecasted coefficients considering the various economic factors and weather factors. Sometimes a short-term ARIMA model was taken into consideration. KEPCO used the short-term forecasts as the fundamental input to the yearly power generating master plan, which greatly affect many relevant sub-plans in regular sequence. However, the more statistical, reliable short-term forecasts for electric power consumptions have been recently one of the main attentions in KEPCO.

In this paper, we focus on developing a short-term forecasting model and estimating the monthly maximum loads forecasted by our model from the KEPCO's past hourly loads (from January of 1977 to December of 1998). From the past hourly loads, the 22-year monthly maximum load data are computed and plotted in <Figure 1> by taking x axis and y axis as time in month and the magnitude of a load in megawatt respectively.

In Section 2 and 3, the necessary data will be selected from the raw data using a statistical argument and with the selected data, a Winters' Multiplicative seasonal model using the exponential smoothing technique will be suggested and fitted. In Section 4, the optimal weights of smoothing parameters will be obtained. In Section 5, the forecasts for the electric power consumptions and their 95% confidence intervals up to December of 1999 will be provided. In Section 6, our model will be evaluated by several statistics including F, t and Theil's U statistics as well as a graph of actual load data along with the predicted power consumptions. For reference, we would like to mention that this paper is slightly modified from the part of our past research report (Yang, 1999).

2. Determination of a Forecasting Model from Raw Load Data : Winters' Multiplicative Model

There have been emerged so many forecasting models in the world that it would not be easy even for an forecasting expert to select an adequate forecasting model without considering the validation and verification of all the existing models and without checking the difference between actual data and forecasted estimates. Hence we would not like to mention how to choose an appropriate forecasting model but choose an

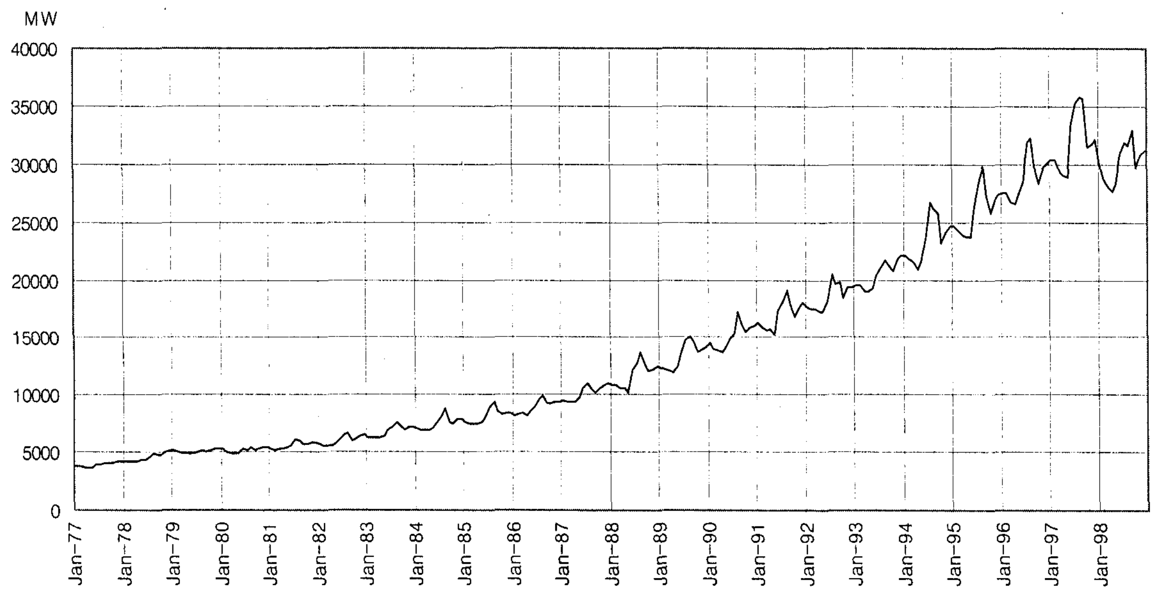

Figure 1. The curve for actual monthly maximum load data (1977-1998).

appropriate model merely by analyzing a plotted actual curve and our past experience as experts.

<Figure 1> shows some characteristics of a trend and seasonal variations. The trend intuitively looks like the linear or quadratic or exponential upward movement of the data over the period of time except the 1998 load due to IMF environment. The regular seasonal pattern in the time series appears from 1981 and keeps forming a series of higher peaks continuously from that year, i.e., in which a pronounced peak during summer and a blunt hill during winter completes itself within an shifted calendar year and then is repeated on a yearly basis. Finally, we can imagine the error component, the erratic movements in the load data that have no definable pattern.

It may be clearly recommended from <Figure 1> that one of seasonal forecasting models such as the classical decomposition model or Winters' seasonal model could be a strong candidate for forecasting model. Or one might suggest at first glance that a multiple linear regression model could be a useful candidate if some independent variables including the time series of population data, GDP data, weather data, the number of cooling and heating devices distributed in Korea could be available. However, it is almost impossible to collect monthly GDP's (note that quarterly GDP's are recently available in Korea) as well as other economic variables. Moreover, it is not easy problem to interpret weather data such as temperature and humidity which are measured every three hours at different locations throughout this country, into a series of interpretable monthly weather data. It follows that one of existing seasonal models could be an appropriate forecasting model to this series of load data. In addition, since the seasonal and trend components are not constant but increasing their magnitudes continuously over time, we could determine that a multiplicative seasonal model would be much better than an additive seasonal model in which seasonal variations are assumed to be constant.

At this point in time, the first practical problem we confronted with is whether a cyclic component should be quantitatively included in our forecasting model or not, since one may intuitively argue that in <Figure 1> there seems to exist a 22-year cycle period forming from 1977 to 1998 due to IMF surroundings. However, this argument might results in a wrong way because we can not quite assure if the next successive economic shocks like the recent IMF shock would come periodically in every 22 years. That is, no one dare to say that the period of cycle must be 22 years. It is clear that there is no way to detect a cycle period with our load data. According to the book (Gaynor and Kirkpatrick, 1994), if it takes 2 to 10 years to complete a cycle and three or four complete cycles of data are needed to establish a clear pattern, one could conceivably need 30 to 40 years of data to obtain a good model. If their comment is true and if our cycle period is 22 years, we need at least 66 to 88 years of load data. Note that the cyclic component is one of the hardest to model because of somewhat irregular nature of a cycle and the required amount of data it takes to establish a cyclical pattern. In most cases, when it is impossible to model cycle quantitatively, the cyclical component can be considered as part of the irregular fluctuations in the trend or the trend is modeled a quadratic.

In this paper we assume that the cyclical component is negligible enough to be the part of the irregular fluctuations in the trend. Also we assume that the past load data is the product of trend, seasonal, and random error components, and that the seasonal factor for any one season is the same for each year and that if trend or seasonal variation is missing, then its value is assumed to be 1. That is, our forecasting model is the Winters' multiplicative model with exponential smoothing technique. For convenience we summarize the Winters' multiplicative model in the book (Gaynor and Kirkpatrick, 1994) as follows. Note that their procedure is slightly different from the original paper (Winters, 1960).

$$Y_t = f(Tr_t, Sn_t, \varepsilon_t) = Tr_t \cdot Sn_t \cdot \varepsilon_t \quad (1)$$

$$Tr_t = a(0) + b(0)t \quad (2)$$

where

Y_t : actual monthly maximum load at time period t

Tr_t : trend component at time t

Sn_t : seasonal component at time t

ε_t : random component at time t

To update the level, $a(t)$, the slope, $b(t)$, and the seasonal component, $Sn_{t+L}(t)$, of the series respectively, we use

$$a(t) = \alpha\left[\frac{Y_t}{Sn_t(t-L)}\right] + (1-\alpha)\{a(t-1) + b(t-1)\} \quad (3)$$

$$b(t) = \beta\{a(t) - a(t-1)\} + (1-\beta)b(t-1) \quad (4)$$

$$Sn_{t+L}(t) = \gamma\left[\frac{Y_t}{a(t)}\right] + (1-\gamma)Sn_t(t-L) \quad (5)$$

where

$a(t)$: the smoothed estimate for the level at time period t

$b(t)$: the smoothed estimate for the slope at time period t

$Sn_{t-L}(t)$: the new smoothed estimate for the seasonal

component at time period t
L : the number of seasons (=12 months in this paper)
α : the weighting factor (constant) for the level
β : the weighting factor (constant) for the slope
γ : the weighting factor (constant) for the seasonal components

After the level, slope, and seasonal estimates have been smoothed, a τ-step-ahead forecast, $\{\hat{Y}_{t+\tau}(t)\}$ and an approximate $100(1-\alpha)\%$ confidence interval, $[(\widehat{YL}_{t+\tau}, \widehat{YU}_{t+\tau})]$, for τ-step-ahead forecast are obtained respectively by the following equations.

$$\hat{Y}_{t+\tau}(t) = \{a(t) + b(t)\tau\} Sn_{t+\tau}(t+\tau-L) \quad (6)$$

$$\widehat{YL}_{t+\tau} = \hat{Y}_{t+\tau}(t) - 1.25 Z_{\frac{\alpha}{2}} MAE_n, \quad \widehat{YU}_{t+\tau} = \hat{Y}_{t+\tau}(t) + 1.25 Z_{\frac{\alpha}{2}} MAE_n \quad (7)$$

where
$\hat{Y}_{t+\tau}(t)$: the forecast for the next time period $(t+\tau)$, estimated at time period t
$Z_{\frac{\alpha}{2}}$: the confidence coefficient for $\frac{\alpha}{2}$ found by using the normal curve tables
$MAE_n = \frac{1}{n}\sum_{t=1}^{n}|e(t)|$: the mean absolute error of the deseasonalized load data
n : the number of the deseasonalized load data given
$e(t) = \left|\frac{Y_t}{Sn_t(t-L)} - \{a(t-1) + b(t-1)\}\right|$: residual error at time period t

3. Initial Estimates for Winters' Multiplicative Model

The initial estimates for the Winters' multiplicative model are $a(0)$, $b(0)$, and the twelve constant seasonal factors. The estimates, $a(0)$ and $b(0)$, may be obtained by fitting Equation (2) to a series of deseasonalized load data which can be computed from the multiplicative decomposition analysis. Note that the initial estimates for the level, slope, and seasonal factors for the Winters' multiplicative model are usually the estimates obtained from the multiplicative decomposition analysis. There are, however, formulas for obtaining the initial estimates that eliminate the need for the decomposition analysis (Holt, 1957; Montgomery and Johnson, 1976; Thompoulos, 1980).

For convenience, we summarize the procedure of the multiplicative decomposition method (Gaynor and Kirkpatrick, 1994) as follows. For the actual time series, we compute the quotient corresponding to $Sn_t \cdot \varepsilon_t$ and the normalized seasonal estimate ($\widehat{Sn}_j$) of the j-th month using equations (8) and (9).

$$Sn_t \cdot \varepsilon_t = \frac{Y_t}{CMA_t} \quad (8)$$

$$\widehat{Sn}_j = \frac{L\overline{Sn_j}}{\sum_{j=1}^{L}\overline{Sn_j}} \quad (9)$$

where
$CMA_t = \frac{1}{2}(MA_{t-1} + MA_t)$: a centered moving average of length L at time period t
$MA_t = \frac{1}{L}\sum_{i=t-L+7}^{t+L-6} Y_i$: the moving average at time period t
$\overline{Sn_j}$: the average seasonal estimate of the j-th month for $j = 1, 2, \cdots, L$

Next, we compute the deseasonalized load data (d_t) corresponding to month j using equation (10). Note that we omit the index j to avoid a complex expression.

$$d_t = \frac{Y_t}{\widehat{Sn}_j} \quad (10)$$

Using equations (8), (9) and (10), the deseasonalized load data are plotted in <Figure 2>.

In order to fit the trend from the deseasonalized load data, we derive a simple linear and an exponential regression equations minimizing the sum of squares of error respectively. It turns out that the coefficient of determination (R^2) of the simple linear regression is as low as 92% not so enough to explain the actual load data well. On the other hand, the R^2 of the exponential regression is very high as 99%.

At this point, pinpointing at the overall trend and the temporal downward movement of load data due to the IMF shock at the end part of the curve as shown in <Figure 2>, we suggested to the most of experts in KEPCO that both the future load demand and seasonal variations would grow linearly or exponentially for the time being and that some recent data should be used in order to reflect more strongly the recent trend and seasonal variations into our forecasting model. We also suggested that the forecasts for the load triggered by the IMF shock could be statistically more reliable and could hopefully reflect the IMF shock by the some recent data if we used an exponential smoothing

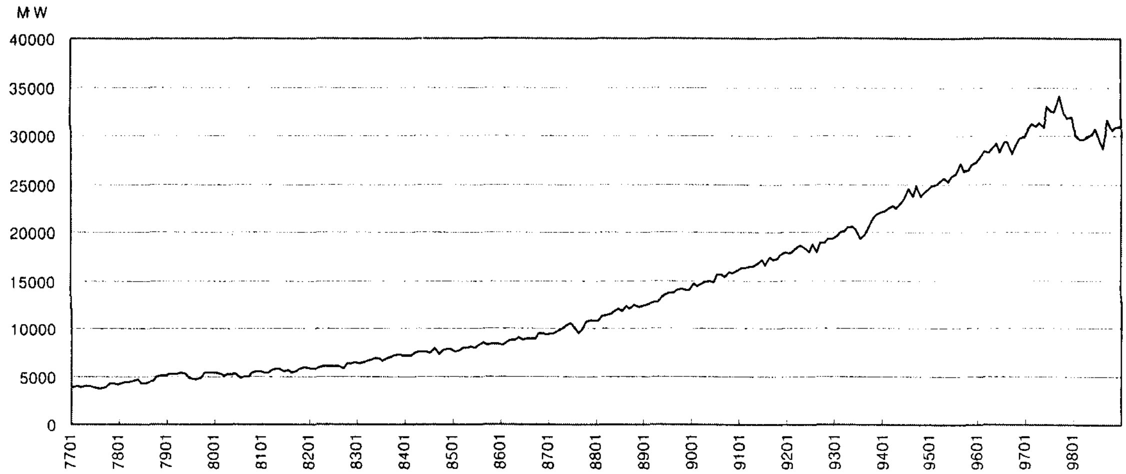

Figure 2. Deseasonalized electric load data by the multiplicative decomposition analysis (1977-1998).

technique and if we selected the proper values of smoothing parameters in the Winters' multiplicative model. Thus the second practical problem we confronted with at this time was how to select some recent data. That is, how to split the raw data into two groups, an old data group and a recent data group so that the recent seasonal variation could be reflected into those monthly maximum loads.

In fact, there has been few theoretically best method about selecting a suitable input data for better forecasts. The method to be used here is to find roughly a splitting year such that two different linear regression equations based on a splitting year have the minimum total sum of squares of error. The reason that we use two linear regressions is that the load exponential curve could be roughly explained with them. Using the SPSS package, 42 regression analyses are made and summarized in <Table 1>.

As shown in Table 1, the first column represents a splitting year, for example, the splitting year, 1988 means that the load data from 1977 to 1987 belong to the first group and the load data from 1988 to 1999 belong to the second group, i.e., to the recent data group. The successive columns of the first group represent the estimate of intercept, slope, the sum of squares of errors (SSE_1), and the coefficient of determination (R_1^2) from the 1977-1987 linear regression equation respectively. In this case, the fitted linear regression equation can be expressed as $d_{1t} = 3{,}343.82 + 47.86t$ where d_{1t} denotes the deseaonalized load data of the first group. The representation above can be applied to the second group, and $d_{2t} = -12{,}957.49 + 173.68t$, except the last column which represents the sum of SSE_1 and SSE_2. Since the minimum is $1.868E+8$, the splitting year becomes 1988 and the 132 load data from 1988 to 1998 as shown in Appendix will be used for modelling. By applying equations (8) and (9) to the recent load data, the seasonal variations were reestimated and summarized in <Table 2>.

The nonparametric Kruskal-Wallis test for seasonality can be used in this case and for convenience, the statistic is given in equation (11).

$$H = \frac{12}{(n-11)(n-12)} \sum_{k=1}^{12} \frac{R_k^2}{n_k} - 3(n-11) \quad (11)$$

where

R_k : the sum of the rankings in season k

n_k : the number of rankings in season k

The computed Kruskal-Wallis statistic($=108.2$), which compared with the five percent critical value for the number of rankings ($=10$), $\chi^2_{10,0.05} = 18.31$, indicates that there is seasonality in the load data. Hence, using these seasonal estimates, the deseasonalized load data are recalculated as shown in Appendix and plotted in <Figure 3>.

As shown in <Figure 3>, the trend model could be a linear, or a quadratic or an exponential model. However, since the initial estimates of the Winters' multiplicative seasonal method include $a(0)$, the level at time period 0 and $b(0)$, the slope at time period 0, a quadratic or an exponential model can not be applied without modifying the smoothing equations, (3), (4) and (5). Note that the Winters' multiplicative seasonal method heavily depends upon the exponential smoothing method in order to absorb a quadratic or an exponential trend model.

As summarized in <Table 3> (the summary statistics for the linear regression model), it is clear that all the statistics except the Durbin-Watson statistic are good; The observed value of t-statistics and $F(=4518.6)$

Table 1. Summary for estimates of slopes and intercepts, SSE_1 and SSE_2, R^2 of two linear regression equations given a splitting year

Splitting Year	First Group				Second Group				
	Intercept	Slope	SSE_1 ($\times 10^3$)	R_1^2	Intercept	Slope	SSE_1 ($\times 10^3$)	R_2^2	SSE ($\times 10^3$)
78	3,789.66	21.99	276	0.20	-1,811.27	117.79	1,457,764	0.93	1,458,040
79	3,619.78	49.59	807	0.78	-2,627.24	122.27	1,250,276	0.93	1,251,082
80	3,666.84	46.42	1,911	0.81	-3,945.69	129.35	945,237	0.95	947,147
81	3,831.99	34.92	3,328	0.77	-5,073.38	135.27	747,129	0.95	750,457
82	3,867.38	32.98	3,701	0.84	-6,259.09	141.37	579,257	0.96	582,958
83	3,873.77	32.62	4,001	0.89	-7,449.13	147.37'	451,127	0.97	455,128
84	3,782.95	36.02	5,220	0.92	-8,745.84	153.77	337,511	0.97	342,731
85	3,687.40	39.16	7,125	0.94	-10,066.07	160.16	250,002	0.98	257,128
86	3,615.77	41.23	8,436	0.95	-11,227.94	165.68	200,802	0.98	209,239
87	3,512.04	43.91	11,814	0.96	-12,187.93	170.16	176,573	0.98	188,387
88	3,343.82	47.86	22,784	0.95	-12,957.49	173.68	164,043	0.97	*186,827
89	3,067.02	53.81	53,163	0.93	-13,860.92	177.74	153,892	0.97	207,056
90	2,756.94	59.94	98,210	0.92	-14,648.26	181.22	148,923	0.96	247,134
91	2,406.17	66.38	164,831	0.91	-15,478.42	184.82	145,161	0.95	309,992
92	2,045.31	72.55	246,084	0.91	-16,005.03	187.07	143,650	0.92	389,734
93	1,706.26	77.98	328,909	0.92	-14,739.22	181.78	138,955	0.88	467,865
94	1,349.70	83.35	435,498	0.92	-9,662.02	160.88	115,785	0.80	551,283
95	876.68	90.05	640,794	0.91	-2,471.29	131.73	97,608	0.62	738,402
96	380.82	96.71	891,235	0.91	14,431.13	64.34	65,884	0.20	957,119
97	-155.91	103.55	1,217,914	0.91	50,486.80	-77.07	30,790	0.18	1,248,704
98	-738.88	110.62	1,648,807	0.91	2,315.56	107.64	5,406	0.23	1,654,213

Table 2. The normalized seasonal estimates ($\widehat{Sn}_j$)

Month \ Year	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	$\overline{Sn_j}$	$\widehat{Sn_j}$
Jan.	-	0.9817	1.0121	1.0051	0.9907	1.0015	1.0082	0.9969	0.9981	1.0012	0.9571	0.9953	0.9952
Feb.	-	0.9609	0.9624	0.9743	0.9766	0.9925	0.9742	0.9671	0.9865	0.9891	0.9325	0.9716	0.9716
Mar.	-	0.9332	0.9430	0.9485	0.9662	0.9619	0.9472	0.9474	0.9527	0.9422	0.9178	0.9460	0.9460
Apr.	-	0.9096	0.9268	0.9190	0.9472	0.9530	0.9071	0.9336	0.9400	0.9232	0.9094	0.9269	0.9268
May	-	0.9433	0.9378	0.9393	0.9399	0.9599	0.9310	0.9235	0.9618	0.9122	0.9370	0.9386	0.9385
Jun.	-	1.0204	0.9828	1.0265	0.9804	1.0031	1.0093	1.0034	0.9902	1.0453	1.0240	1.0085	1.0085
Jul.	1.0735	1.0980	1.1062	1.0558	1.0980	1.0197	1.1303	1.0834	1.1012	1.1037	-	1.0870	1.0869
Aug.	1.1357	1.1022	1.1175	1.1150	1.0509	1.0452	1.0981	1.1330	1.1064	1.1264	-	1.1030	1.1030
Sep.	1.0343	1.0574	1.0263	1.0323	1.0458	1.0177	1.0748	1.0305	0.9982	1.1244	-	1.0442	1.0441
Oct.	0.9881	0.9854	0.9870	0.9665	0.9677	0.9841	0.9547	0.9582	0.9588	0.9957	-	0.9746	0.9746
Nov.	0.9884	0.9928	0.9969	1.0044	1.0072	1.0225	0.9863	0.9928	0.9966	1.0061	-	0.9994	0.9994
Dec.	0.9960	0.9878	1.0017	1.0186	0.9998	1.0261	1.0011	1.0002	1.0006	1.0228	-	1.0055	1.0054
											sum	0.999953	1.0000

* The entries in <Table 2> indicate $Sn_t \cdot \varepsilon_t$ and $\overline{Sn_j}$ represents the average of the real numbers in the same row j.

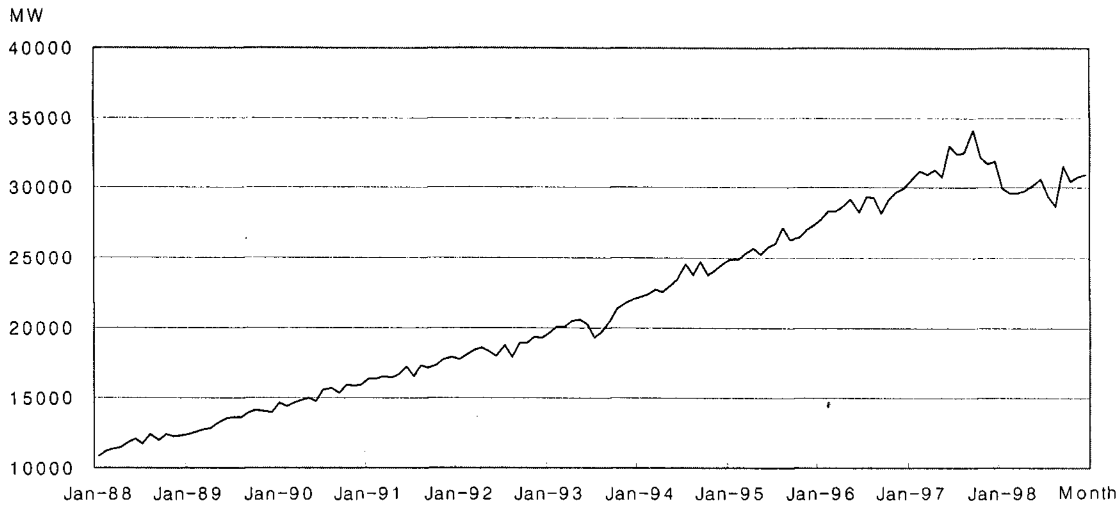

Figure 3. Deseasonalized electric load data by the multiplicative decomposition analysis (1988-1998).

Table 3. Summary statistics for the linear regression model

	estimate	Standard Error	t-Statistic	p-value
$a(0)$	9,968.744	196.659	50.691	0.000
$b(0)$	173.684	2.566	67.689	0.000
Durbin-Watson	0.320			
R^2	0.972			
R^2_{adj}	0.972			
$\sqrt{MSE}$	1123.3			

indicate that we do not accept each null hypothesis. However, the Durbin-Watson statistic of the linear model, $d=0.320$, which compared with the five percent critical values for $n=130$, $d_L=1.6540$, $d_U=1.6940$, $4-d_U=2.3060$, $4-d_L=2.3460$ indicates that the residuals are positively autocorrelated.

The presence of autocorrelation in the errors has several affects on the ordinary least squares regression procedure. According to the book (Montgomery and Peck, 1982), when the errors are positively autocorrelated, the residual mean square(MSE) may be seriously underestimated and consequently the standard errors of the regression coefficients may be too small even if the ordinary least squares regression coefficients are still unbiased. It follows that the confidence intervals and tests of hypothesis based on t and F distributions are no longer appropriate. Generally, underestimated MSE gives the analysis a false impression of accuracy.

There are many approaches for dealing with the problem of the autocorrelation. In general, various transformation techniques are used to estimate the model parameters accurately. The well-known Cochrane-Orcutt iterative procedure (1949) happened not to work in our recent deseasonalized load data. Note that the Cochrane-Orcutt procedure is not always successful due to the estimate of $\hat{\rho}$, which is used in the Cochrane-Orcutt procedure, is biased downward. However, our procedure based on the Chatterjee and Price procedure (1977) works well on them. Our procedure is to estimate $a(0)$, $b(0)$, and ρ simultaneously by minimizing

$$S(a(0), b(0), \rho) = \sum_{t=2}^{n}[Tr_t - \rho Tr_{t-1} - a(0)(1-\rho) - b(0)\{t-\rho(t-1)\}]^2 \qquad (12)$$

as long as $d_U < d < 4-d_U$. Note that the original variables are transformed as follows :

$$t' = t - \hat{\rho}(t-1) \text{ for } t=2, 3, \cdots, n \qquad (13)$$

$$Tr_t' = Tr_t - \hat{\rho}\, Tr_{t-1} \text{ for } t=2, 3, \cdots, n \qquad (14)$$

In our case, it turns out that $\widehat{a(0)}=9,959.77$, $\widehat{b(0)}=173.2411$, $\hat{\rho}=0.625$, and $S(\widehat{a(0)}, \widehat{b(0)}, \hat{\rho})=5.655E+7$. Since the estimate of $a(0)$ can be expressed in equation (15)

$$\widehat{a(0)} = \frac{\widehat{a'(0)}}{1-\hat{\rho}} \qquad (15)$$

the ordinary least squares fit to the transformed variables yields with $F_0=1803.6$

$$Tr_t' = \widehat{a'(0)} + \widehat{b(0)}\, t' = 3,734.891 + 173.2411\, t' \qquad (16)$$

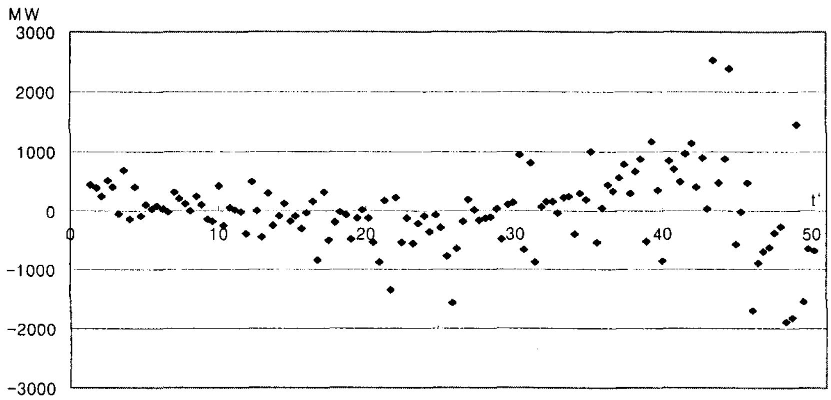

Figure 4. Plot of residuals by the transformed linear regression, equation(14).

Table 4. Summary statistics for the transformed linear regression model

	estimate	Standard Error	t-Statistic	p-value
$a(0)$	3734.891	119.916	31.146(0.000)	0.000
$b(0)$	173.243	4.079	42.469(0.000)	0.000
Durbin-Watson	1.697			
R^2	0.933			
R^2_{adj}	0.933			
$\sqrt{MSE}$	662.1			

The Durbin-Watson statistic for the transformed model, 1.697, indicates that the residuals are not autocorrelated. The plots of the residuals are given in <Figure 4>. The summary statistics for the transformed variables are summarized in <Table 4>.

Now using $\widehat{a(0)}=9{,}959.77$, $\widehat{b(0)}=173.2411$, and the normalized seasonal estimates given in <Table 2>, we summarize the results of the multiplicative decomposition procedure in Appendix. The columns show the 132 time periods corresponding to year and month, actual monthly maximum power consumption, MA_t, CMA_t, the quotient $Sn_t \cdot \varepsilon_t$, $\widehat{Sn_j}$, d_t, $\widehat{a(0)}+\widehat{b(0)}t$ and residuals in order.

Next, these initial estimates are smoothed using the equations in (3), (4), and (5). The Winters' multiplicative procedure can be then continued by updating the estimates and computing one-step-ahead forecasts for each time period. after determining the model parameters.

4. Optimal weights of Smoothing Parameters

After the initial estimates are found, the procedure for obtaining the optimal values of α, β, and γ for the smoothing process is simply a matter of finding the combination of weights that yields the smallest sum of squares of errors (SSE) as follows.

$$SSE=\sum_{t=1}^{n}\{Y_t-\widehat{Y}_t(t-1)\}^2 \qquad (17)$$

where $\widehat{Y}_t(t-1)$ is the forecast for the next time period t, estimated at time period $(t-1)$.

Using the EXCEL of Microsoft Office 2000 Professional (Dodge and Craig, 1999), the optimal $(\alpha^*, \beta^*, \gamma^*)$ can be obtained easily as (0.6311, 0.0000, 0.0000) and the resulted SSE is 5.2063×10^7. Note that the slope and seasonal component depends completely upon the past observations since $\beta^*=\gamma^*=0$ and that in case of level, more emphasis is placed on current observations.

Since MAE is 407.96, the maximum of absolute forecasted errors is 2,357 MW, and the maximum of relative forecasted errors is 7.71%, our model is, with a few exceptions, a very good fit to the historical load data.

5. Forecast and 95% Confidence Interval by Our Model

Using the equations (6) and (7), we can obtain the forecasts and the approximate 95% confidence intervals for electric loads up to December of 2000, which correspond to $133\le t\le144$ as summarized in <Table 5>. Note that the time period for forecasting is December of 1998 corresponding to $t=132$. As shown in the last column, the relative errors of the forecasts over the actual load data are ranged between 0.1% and 6.6% and they increases gradually as normally expected. The average relative error of the forecasts is only 3.3%. These results indicate that our model is

Table 5. Forecasts for monthly maximum electric loads in 1999

Month	Actual monthly maximum load	monthly maximum load			
		Forecast	95% Confidence Interval		Error(%)
			L.B.	U.B.	
Jan	31,241	30,993	29,995	31,990	0.8
Feb	30,812	30,424	29,427	31,421	1.3
Mar	29961	29,786	28,789	30,783	0.6
Apr	30,396	29,344	28,347	30,342	3.5
May	29,850	29,877	28,880	30,874	0.1
Jun	34,006	32,279	31,282	33,276	5.1
Jul	36,577	34,978	33,981	35,975	4.4
Aug	37,293	35,686	34,689	36,683	4.3
Sep	36,338	33,962	32,965	34,959	6.5
Oct	33,061	31,870	30,872	32,867	3.6
Nov	34,034	32,853	31,855	33,850	3.5
Dec	35,562	33,226	32,229	34,223	6.6

very suitable to KEPCO. In addition, using the forecasted maximum monthly load, we suggest 35,686 MW as the 1999 yearly maximum load. As a point of interest, it turned out that the 1999 peak load was 37,293 MW and that the relative errors were 4.3%.

6. The Evaluation of Our Model

The Winters' exponential smoothing method is basically intuitive and gives the positive results obtained in the practical application of the method. Unfortunately, there are a number of theoretical weaknesses in its approach. However, once the model has been built, there must be several ways to measure its accuracy.

First, since a graph of actual values along with the predicted values of Y_t is an excellent means of showing how well the model fits the data, we present <Figure 5>, in which there are almost no difference between the actual values and the predicted values except in 1993 and 1998 and the reliability of our model could be judged with naked eyes.

Second, we showed in Section 3 that all the statistical tests for regression, which include a residual analysis, the test for autocorrelation, goodness of fit, and slopes, were met for the trend component.

Another measure of model accuracy is to compute Theil's U statistic written in equation (18). It turns out that the accuracy of our model is greatly high since

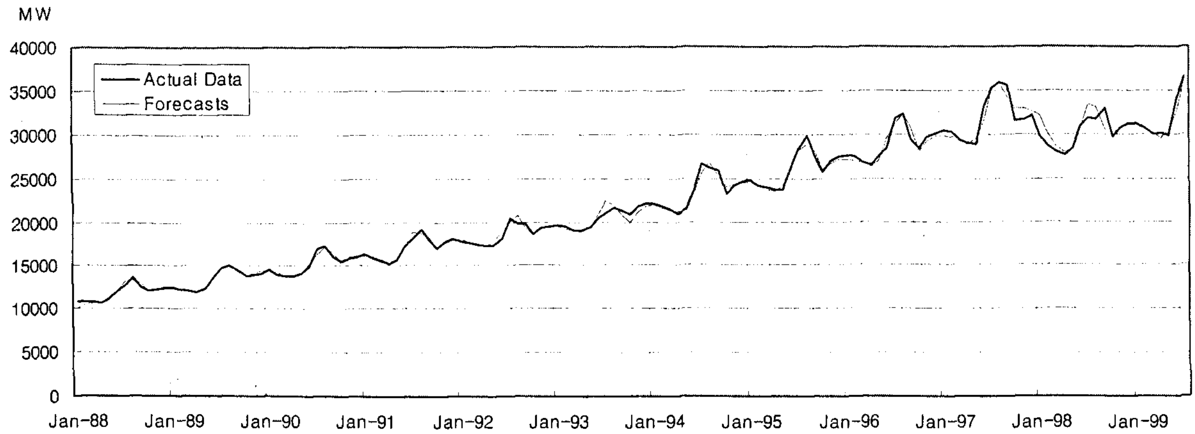

Figure 5. The graph for the actual maximum monthly load data and predicted load data by the Winters' multiplicative model.

Theil's U value is computed as 0.0135. Note that any U equal to or less than 0.55 indicates a very good fit.

$$U=\frac{\sqrt{\frac{1}{n}\sum e(t)^2}}{\sqrt{\frac{1}{n}\sum Y_t^2}+\sqrt{\frac{1}{n}\sum \hat{Y}_t(t-1)^2}} \quad (18)$$

7. Conclusion

In this paper, we applied the Winters' multiplicative model with exponential smoothing technique in order to forecast the electric power consumptions and their 95% confidence interval up to December of 1999. It turned out that the relative errors of our forecasts over the actual load data were ranged between 0.1% and 6.6% and that the average relative error was only 3.3%. Thus our model, which was accepted as the first statistical forecasting model for monthly maximum power consumption, showed the desirable results to KEPCO. In addition, using our forecasted maximum monthly load, we suggest 35,686 MW as the 1999 yearly maximum load. As a point of interest, it turned out that the 1999 peak load was as 37,293 MW and that the relative errors were only 4.3%.

A forecasting model can not be perpetual but evolving continuously with the new added data so that if there seems to be any abnormal phenomenon along the sequence of forecasting errors, it should be immediately updated to a new model or should be fundamentally changed to another different type of model. In case we build a forecasting model periodically or yearly, the effort and cost as well as forecasting experts would be required. Hence we recommended to KEPCO a long-term project on a particular expert system for building a forecasting model of electric power consumptions. We expect that the project could have more possibility of success in the future since this expert system have some internal special properties in electrical power consumptions.

Further research may be concentrated on finding the appropriate values of parameters, (α, β, γ) in Winters' multiplicative model. There is no definite principle in that we must minimize the sum of squares of errors in order to estimate the model parameters. We might define the optimal values of (α, β, γ) such that mean absolute error may be minimized or such that the sum of relative or absolute forecasted error squares may be minimized. On the other hand, since it is especially important to forecast the yearly maximum load in KEPCO, we might define the optimal values of model parameters such that the sum of error squares of August is minimized and so on.

Appendix : The electric load data from 1988 to 1998, used for modeling

Number	Year/Month	Actual Monthly Maximum Load $Y=Tr*Sn*e$	Moving Average (MA)	Centered Moving Average (CMA)	$Sn*e$	$\widehat{Sn}_j$	$d=Tr*e$	$\widehat{a}(0)+\widehat{b}(0)t$	residual
1	8801	10822				0.9952	10874	10142	731
2	8802	10879				0.9716	11198	10316	881
3	8803	10792				0.9460	11408	10490	919
4	8804	10627				0.9268	11466	10663	802
5	8805	11107				0.9385	11835	10837	997
6	8806	12136	11838			1.0085	12034	11011	1023
7	8807	12777	11967	11902	1.0735	1.0869	11755	11185	571
8	8808	13656	12082	12025	1.1357	1.1030	12381	11358	1023
9	8809	12550	12185	12134	1.0343	1.0441	12020	11532	488
10	8810	12092	12289	12237	0.9881	0.9746	12407	11706	702
11	8811	12201	12401	12345	0.9884	0.9994	12209	11879	330
12	8812	12412	12524	12462	0.9960	1.0054	12345	12053	292
13	8901	12378	12694	12609	0.9817	0.9952	12437	12227	211
14	8902	12253	12811	12752	0.9609	0.9716	12612	12400	211
15	8903	12034	12982	12896	0.9332	0.9460	12721	12574	147
16	8904	11871	13120	13051	0.9096	0.9268	12808	12748	60

Number	Year/Month	Actual Monthly Maximum Load $Y = Tr*Sn*e$	Moving Average (MA)	Centered Moving Average (CMA)	$Sn*e$	$\widehat{Sn}_j$	$d = Tr*e$	$\widehat{a(0)} + \widehat{b(0)}t$	residual
17	8905	12447	13271	13196	0.9433	0.9385	13262	12921	341
18	8906	13611	13408	13339	1.0204	1.0085	13496	13095	401
19	8907	14821	13588	13498	1.0980	1.0869	13636	13269	367
20	8908	15057	13733	13660	1.1022	1.1030	13651	13442	209
21	8909	14601	13885	13809	1.0574	1.0441	13984	13616	368
22	8910	13759	14041	13963	0.9854	0.9746	14118	13790	328
23	8911	14006	14174	14107	0.9928	0.9994	14015	13963	51
24	8912	14053	14279	14226	0.9878	1.0054	13977	14137	-160
25	9001	14540	14453	14366	1.0121	0.9952	14610	14311	299
26	9002	13997	14636	14544	0.9624	0.9716	14407	14485	-78
27	9003	13856	14752	14694	0.9430	0.9460	14648	14658	-11
28	9004	13740	14899	14825	0.9268	0.9268	14825	14832	-7
29	9005	14043	15049	14974	0.9378	0.9385	14963	15006	-43
30	9006	14872	15215	15132	0.9828	1.0085	14747	15179	-432
31	9007	16911	15358	15287	1.1062	1.0869	15558	15353	206
32	9008	17251	15515	15437	1.1175	1.1030	15640	15527	114
33	9009	15997	15661	15588	1.0263	1.0441	15321	15700	-379
34	9010	15519	15786	15723	0.9870	0.9746	15924	15874	50
35	9011	15805	15924	15855	0.9969	0.9994	15815	16048	-233
36	9012	16053	16129	16026	1.0017	1.0054	15966	16221	-255
37	9101	16255	16217	16173	1.0051	0.9952	16333	16395	-62
38	9102	15876	16373	16295	0.9743	0.9716	16341	16569	-228
39	9103	15604	16528	16451	0.9485	0.9460	16495	16742	-247
40	9104	15241	16641	16585	0.9190	0.9268	16444	16916	-472
41	9105	15704	16798	16720	0.9393	0.9385	16733	17090	-357
42	9106	17328	16963	16881	1.0265	1.0085	17182	17263	-81
43	9107	17973	17082	17023	1.0558	1.0869	16536	17437	-902
44	9108	19123	17221	17151	1.1150	1.1030	17338	17611	-273
45	9109	17858	17376	17299	1.0323	1.0441	17103	17785	-681
46	9110	16876	17544	17460	0.9665	0.9746	17316	17958	-642
47	9111	17686	17673	17609	1.0044	0.9994	17697	18132	-435
48	9112	18036	17739	17706	1.0186	1.0054	17939	18306	-367
49	9201	17676	17944	17841	0.9907	0.9952	17761	18479	-719
50	9202	17548	17994	17969	0.9766	0.9716	18062	18653	-591
51	9203	17464	18155	18074	0.9662	0.9460	18462	18827	-365
52	9204	17257	18285	18220	0.9472	0.9268	18619	19000	-381
53	9205	17250	18423	18354	0.9399	0.9385	18380	19174	-794
54	9206	18117	18536	18480	0.9804	1.0085	17964	19348	-1383
55	9207	20437	18691	18613	1.0980	1.0869	18802	19521	-719
56	9208	19726	18851	18771	1.0509	1.1030	17884	19695	-1811
57	9209	19782	18979	18915	1.0458	1.0441	18946	19869	-923
58	9210	18437	19124	19052	0.9677	0.9746	18918	20042	-1125
59	9211	19348	19296	19210	1.0072	0.9994	19360	20216	-856
60	9212	19388	19487	19391	0.9998	1.0054	19283	20390	-1106
61	9301	19538	19531	19509	1.0015	0.9952	19632	20563	-932

Number	Year/Month	Actual Monthly Maximum Load $Y = Tr*Sn*e$	Moving Average (MA)	Centered Moving Average (CMA)	$Sn*e$	$\widehat{Sn}_j$	$d = Tr*e$	$\widehat{a(0)} + \widehat{b(0)}t$	residual
62	9302	19466	19695	19613	0.9925	0.9716	20036	20737	-701
63	9303	19007	19825	19760	0.9619	0.9460	20093	20911	-818
64	9304	18988	20023	19924	0.9530	0.9268	20487	21085	-598
65	9305	19318	20227	20125	0.9599	0.9385	20584	21258	-675
66	9306	20403	20454	20340	1.0031	1.0085	20231	21432	-1201
67	9307	20965	20668	20561	1.0197	1.0869	19288	21606	-2317
68	9308	21703	20860	20764	1.0452	1.1030	19677	21779	-2103
69	9309	21336	21069	20964	1.0177	1.0441	20434	21953	-1519
70	9310	20810	21226	21147	0.9841	0.9746	21353	22127	-774
71	9311	21800	21416	21321	1.0225	0.9994	21814	22300	-486
72	9312	22112	21684	21550	1.0261	1.0054	21993	22474	-481
73	9401	22102	22161	21923	1.0082	0.9952	22208	22648	-440
74	9402	21771	22533	22347	0.9742	0.9716	22408	22821	-413
75	9403	21520	22908	22721	0.9472	0.9460	22749	22995	-246
76	9404	20867	23103	23005	0.9071	0.9268	22514	23169	-654
77	9405	21598	23297	23200	0.9310	0.9385	23013	23342	-329
78	9406	23621	23510	23403	1.0093	1.0085	23422	23516	-94
79	9407	26696	23729	23619	1.1303	1.0869	24561	23690	871
80	9408	26167	23932	23830	1.0981	1.1030	23724	23863	-140
81	9409	25828	24132	24032	1.0748	1.0441	24737	24037	699
82	9410	23153	24371	24251	0.9547	0.9746	23757	24211	-454
83	9411	24124	24545	24458	0.9863	0.9994	24139	24384	-245
84	9412	24669	24741	24643	1.0011	1.0054	24536	24558	-22
85	9501	24731	24874	24807	0.9969	0.9952	24850	24732	118
86	9502	24204	25183	25028	0.9671	0.9716	24913	24906	7
87	9503	23923	25317	25250	0.9474	0.9460	25290	25079	210
88	9504	23736	25533	25425	0.9336	0.9268	25610	25253	357
89	9505	23687	25767	25650	0.9235	0.9385	25239	25427	-188
90	9506	25970	25995	25881	1.0034	1.0085	25751	25600	151
91	9507	28292	26233	26114	1.0834	1.0869	26029	25774	255
92	9508	29878	26509	26371	1.1330	1.1030	27088	25948	1141
93	9509	27438	26745	26627	1.0305	1.0441	26279	26121	157
94	9510	25739	26981	26863	0.9582	0.9746	26410	26295	115
95	9511	26941	27290	27135	0.9928	0.9994	26958	26469	490
96	9512	27398	27495	27393	1.0002	1.0054	27250	26642	608
97	9601	27593	27794	27644	0.9981	0.9952	27725	26816	909
98	9602	27518	27994	27894	0.9865	0.9716	28324	26990	1334
99	9603	26747	28153	28074	0.9527	0.9460	28275	27163	1111
100	9604	26567	28374	28264	0.9400	0.9268	28664	27337	1327
101	9605	27401	28602	28488	0.9618	0.9385	29196	27511	1685
102	9606	28431	28825	28713	0.9902	1.0085	28192	27684	507
103	9607	31871	29061	28943	1.1012	1.0869	29322	27858	1464
104	9608	32282	29296	29178	1.1064	1.1030	29268	28032	1236
105	9609	29350	29508	29402	0.9982	1.0441	28110	28206	-96
106	9610	28393	29716	29612	0.9588	0.9746	29133	28379	754

Number	Year/Month	Actual Monthly Maximum Load $Y = Tr*Sn*e$	Moving Average (MA)	Centered Moving Average (CMA)	$Sn*e$	$\widehat{Sn}_j$	$d = Tr*e$	$\widehat{a(0)} + \widehat{b(0)}t$	residual
107	9611	29678	29841	29778	0.9966	0.9994	29697	28553	1144
108	9612	30063	30247	30044	1.0006	1.0054	29901	28727	1174
109	9701	30425	30527	30387	1.0012	0.9952	30571	28900	1671
110	9702	30342	30825	30676	0.9891	0.9716	31230	29074	2156
111	9703	29289	31350	31087	0.9422	0.9460	30962	29248	1714
112	9704	29061	31606	31478	0.9232	0.9268	31355	29421	1934
113	9705	28909	31775	31690	0.9122	0.9385	30803	29595	1208
114	9706	33303	31945	31860	1.0453	1.0085	33023	29769	3254
115	9707	35231	31894	31920	1.1037	1.0869	32413	29942	2471
116	9708	35851	31762	31828	1.1264	1.1030	32504	30116	2388
117	9709	35655	31658	31710	1.1244	1.0441	34148	30290	3859
118	9710	31463	31537	31598	0.9957	0.9746	32283	30463	1820
119	9711	31707	31490	31514	1.0061	0.9994	31727	30637	1090
120	9712	32105	31290	31390	1.0228	1.0054	31932	30811	1121
121	9801	29815	31012	31151	0.9571	0.9952	29958	30984	-1027
122	9802	28755	30661	30836	0.9325	0.9716	29597	31158	-1561
123	9803	28038	30439	30550	0.9178	0.9460	29640	31332	-1692
124	9804	27615	30294	30366	0.9094	0.9268	29795	31506	-1711
125	9805	28347	30213	30253	0.9370	0.9385	30204	31679	-1475
126	9806	30898	30133	30173	1.0240	1.0085	30638	31853	-1215
127	9807	31899				1.0869	29348	32027	-2679
128	9808	31629				1.1030	28676	32200	-3524
129	9809	32996				1.0441	31602	32374	-772
130	9810	29718				0.9746	30493	32548	-2055
131	9811	30745				0.9994	30765	32721	-1957
132	9812	31139				1.0054	30971	32895	-1924

References

Chatterjee, S. and Price, B. (1977), *Regression Analysis by Example*, Wiley, New York.

Cochrane, D. and Orcutt, G. H. (1949), Application of least squares regression to relationships containing autocorrelated error terms, *J. of Amer. Statist. Assoc.*, **44**, 32-61.

Dodge, M. and Craig, S. (1999), *Running Microsoft Excel 2000*, Microsoft Press.

Gaynor, P. E. and Kirkpatrick, R. C. (1994), *Introduction to Time-series Modeling and Forecasting in Business and Economics*, McGraw-Hill Inc, 347-355.

Holt, C. C. (1957), *Forecasting Seasonal and Trends by Exponentially Weighted Moving Averages*, Carnegie Institute of Technology, Pittsburgh, Pa.

KEPCO (1999), *Analysis of 1998 Summer Cooling Load, Division of Electric Economics*, Reg. No. 98-0574 Dan-107.

Montgomery, D. C. and Johnson, L. A. (1976), *Forecasting and Time Series Analysis*, McGraw-Hill Book Co.

Montgomery, D. C. and Peck, E. A. (1982), *Introduction to Linear Regression Analysis*, John Wiley and Sons, New York, 348-353.

Thompoulos, N. T. (1980), *Applied Forecasting Methods*, Prentice-Hall, Inc., 167-175.

Winters, P. R. (1960), Forecasting Sales by Exponentially Weighted Moving Averages, *Management Science*, April, 324-342.

Yang, M. (1999), *Development of Short/Long-term Forecasting Models for Yearly Maximum Load and Cooling Load*, Dankook University, Cheonan.