Modelling Temperature Dependency of Electricity Demand - The Key to Short Term Forecasting of Energy Demand by the Half Hour

Luminita Baloi, Melbourne

Abstract As a consequence of the record hot temperatures experienced in the 1996/97 summer, the electricity maximum demand increased dramatically. Therefore, a model for the short-term forecast was developed. The paper presents an applied study in forecasting the energy requirements in order to decide the adjustments and optimise the demand profile of the hedging cover. Based on the electricity industry experience, it is known that temperature is one of the major factors that influences the energy consumption. Thus, the forecasting model is a non-linear regression as a function of maximum and minimum temperatures and other explanatory variables. An illustrative example of the application of the forecasting model is presented and it is shown that the model performs very well with highly statistical parameters. The procedure is used to predict three day ahead the energy consumption.

1. Introduction

After the electricity reform process in 1993 and 1994, the State Electricity Commission of Victoria (SECV) was separated into a transmission system owner, a market operator, five generators and five distributors. Victoria followed the UK model of privatisation. A "pool" in which the suppliers, generators and large users buy and sell electricity was formed. The pool sets the market price for electricity each half-hourly, daily. Because of the variation of the electricity demand, the volatility of the prices is also large. Therefore an accurate model for the daily energy consumption and half-hourly load is required to be developed.

2. Defining Requirements of Model

A three day ahead energy forecast is a requirement for the distributors in order to manage the customer demand.

Electricity retailers purchase electricity from the pool, with short-term pricing estimates based on the price scheduled previous day.

Because of the large volatility in the half hourly prices, having a very precise load forecast becomes an extremely important factor.

3. Determinants for Modelling

The data used in this model contain the daily energy consumption (MWh) during the period 29-Nov-96 to 26-Mar-97. The actual figures are shown in **Fig.1**.

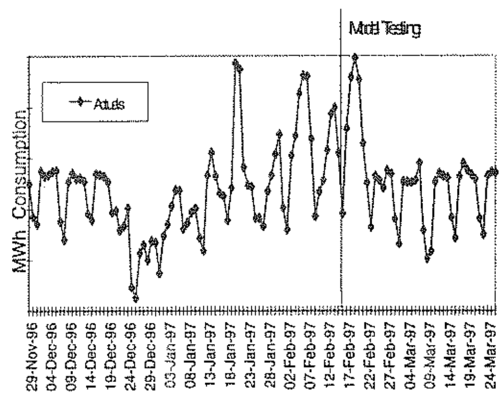

Fig.1 Initial Data

The data was divided in two data sets: the first set contains the actual figures between 29-Nov 1996 and 16-Feb 1997 and was used in the model development, while the remaining data was used for the model testing.

The Bureau of Meteorology provides only three day ahead minimum and maximum weather temperatures in Celsius Degrees. Because the forecast for rainfall or wind is not available the model cannot take in consideration these factors.

In the model, maximum temperatures (X_t) and minimum temperatures (Z_t) are the variables used. Regressing the consumption (Y_t) on the maximum and minimum temperature, the model explained only 27.4% (R-squared adjusted). Analysing the points where the model does not fit well, it has been seen that other specific attributes influenced the consumption.

Other explanatory variables are the variate types of the day, ie. weekday, weekend or holiday. There is also a difference between the type of holidays. Therefore, the following dummy variables were used: D_1 - for Christmas Holidays, D_2 - for January Holidays, D_3 - for State Holidays, D_4 - for working days, D_5 - for Mondays, D_6 - for Tuesdays, D_7 - for Saturdays, D_8 - for Sundays.

4. Analysis

All the data was analysed using SAS - Version 6.12 as the main software. The new features implemented in SAS assist to automatically fit 20 models to a time series and select the best.

4.1 Model Development

Non-linear regression is the method used to create the two models. <u>**Model 1**</u> was constructed using all data in the period 29-Nov-96 and 16-Jan-97 and the remaining data for the forecast. <u>**Model 2**</u> consists of two separate models: <u>**Model 2a**</u> which estimates weekdays and <u>**Model 2b**</u> which estimates weekend days. Using the traditional approaches for generating the Short-term forecast and using SAS as the computer package, the best model selected (from all 20 models that were automatically fitted) was <u>**Model 3**</u> (Log Winters Method - Additive).

<u>**Model 1**</u> estimates the energy consumption (Y_t) by the following equation:

$$(1)\quad Y_t = A_1 + A_2{*}X_t + A_3{*}X_t^2 + A_4{*}D_1 + A_5{*}D_2 + A_6{*}D_3 + A_7{*}D_4 + A_8{*}D_5 + A_9{*}D_6 + A_{10}{*}D_7 + A_{11}{*}D_8$$

where A_1 - A_{11} are the estimating coefficients.

Equation (1) is used to calculate the fitted values for the energy consumption and the forecast for the period between 17-Feb-97 and 26-Mar-97. The plotted values of the actual and estimated figures can be seen in **Fig.2**.

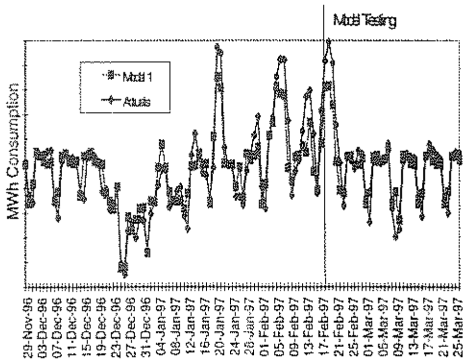

Fig.2 Model 1

<u>**Model 2**</u> uses two components for the estimation of the energy consumption (Y_t): <u>**Model 2a**</u> for the working days (equation **(2)**) and <u>**Model 2b**</u> for the weekend days (equation (3)).

$$(2)\quad Y_t = B_1 + B_2{*}X_t + B_3{*}X_t^2 + B_4{*}Z_t^2 + B_5{*}X_t{*}Z_t + B_6{*}X_{t-1} + B_7{*}D_1 + B_8{*}D_2 + B_9{*}D_3 + B_{10}{*}D_4$$

$$(3)\quad Y_t = C_1 + C_2{*}X_t + C_3{*}X_t^2 + C_4{*}X_{t-1} + C_5{*}D_7 + C_6{*}D_1$$

where B_1 - B_{10} and C_1 - C_6 are the estimating coefficients for (2) and (3).

Equations (2) and (3) are used to calculate the fitted values for the energy consumption and the forecast for the period between 17-Feb-97 and 26-Mar-97. The plotted values of the actual and estimated figures can be seen in **Fig.3**.

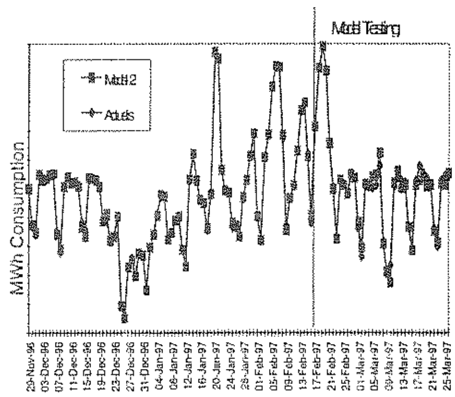

Fig.3 Model 2

<u>**Model 3**</u> - Additive Log Winters Method was fitted automatically using SAS software. This model was the best of the 20 models fitted automatically by the software, although its

results are disappointing. The estimated parameters are shown in **Table 1.**

PARAMETERS	VALUE
Level Smoothing Weight	0.999
Trend Smoothing Weight	0.001
Seasonal Smoothing Weight	0.001
Smoothed Level	9.59436
Smoothed Trend	0.0007163
Smoothed Seasonal Factor 1	0.03858
Smoothed Seasonal Factor 2	-0.05343
Smoothed Seasonal Factor 3	-0.11004
Smoothed Seasonal Factor 4	0.01821
Smoothed Seasonal Factor 5	0.04814
Smoothed Seasonal Factor 6	0.02331
Smoothed Seasonal Factor 7	0.03522

Table 1 Estimated parameters for Model 3

Table 1 is used to calculate the fitted values for the energy consumption and the forecast for the period between 17-Feb-97 and 26-Mar-97.
The plotted values of the actual and estimated figures can be seen in **Fig.4**.

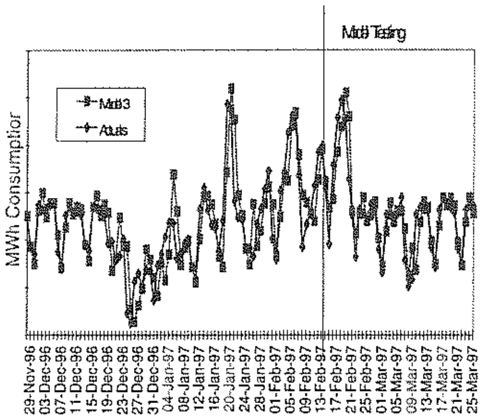

Fig.4 Model 3

4.2 Model Testing

In order to compare the performances of the three models mentioned above, the following common statistics were calculated: R-squared (**R-sq**) and R-squared adjusted (**R-sq adj**), J_p - known also as Final Prediction Error (FPE), Hocking's Criterion (S_p), Hannan and Quinn Criterion (**HQ**), Generalized Cross-Validation Criterion (**GCV**), Mean Square Error (**MSE**), Mean Absolute Deviation (**MAD**), Mean Absolute Percent Error (**MAPE**), Mean Error - Bias (**M**), Mean Percent Error - Percent Bias (**%M**). These tests were performed for all the three models, for:

- only data used in model construction - results in **Table 2**
- only the test data - results in **Table 3**
- all data between 29-Nov-96 and 26-Mar-97 - results in **Table 4.**

Tests	Model 1	Model2	Model3
R-sq	0.8336	0.9988	0.6446
R-sq adj	0.8099	0.9986	0.6307
Jp	57,143,634	404,767	102,354,595
Sp	9,202	65	16,226
HQ	64,304,615	450,246	106,082,954
GCV	58,050,676	409,955	102,498,734
MSE	555,563	4,036	1,186,943
MAD	542.5	27.6	784.4
MAPE	3.71	0.21	5.59
M	42.7	-1.0	-39.7
%M	-0.00099	-0.00005	-0.00545

Table 2

Tests	Model 1	Model2	Model3
R-sq	0.8135	0.9842	0.7446
R-sq adj	0.7469	0.9793	0.7227
Jp	34,951,303	2,793,526	32,705,245
Sp	26,969	2,123	23,461
HQ	40,229,539	3,177,615	34,233,481
GCV	37,551,846	2,959,539	32,910,365
MSE	536,533	45,360	734,713
MAD	515.8	141.0	615.6
MAPE	3.49	1.00	4.12
M	140.2	5.3	-46.6
%M	0.00505	0.00040	-0.00546

Table 3

Tests	Model 1	Model2	Model3
R-sq	0.8297	0.9946	0.6772
R-sq adj	0.8139	0.9942	0.6688
Jp	76,839,477	2,384,534	129,285,42

			7
Sp	5,610	174	9,373
HQ	84,491,275	2,597,399	133,035,072
GCV	77,395,318	2,398,487	129,369,047
MSE	549,435	17,344	1,041,310
MAD	533.9	64.1	730.1
MAPE	3.64	0.46	5.12
M	74.1	1.0	-41.9
%M	0.00096	0.00009	-0.00545

Table 4

Another important criterion is given by the ability to duplicate turning points or rapid changes in the data. Therefore, the percentages for false signals, missed signals and good forecast were calculated. This test was performed for all three models, for:

- only data used in model construction - results in **Table 5**
- only the test data - results in **Table 6**
- all data between 29-Nov-96 and 26-Mar-97 - results in **Table 7.**

	Model 1	Model2	Model3
False Signals	56.82	0.00	52.78
Missed Signals	44.12	6.00	50.00
Good Forecast	48.72	97.00	54.00

Table 5

	Model 1	Model2	Model3
False Signals	55.00	38.10	50.00
Missed Signals	43.75	18.75	37.50
Good Forecast	52.63	71.05	57.89

Table 6

	Model 1	Model2	Model3
False Signals	56.25	15.09	51.79
Missed Signals	44.00	10.00	46.00
Good Forecast	50.00	88.79	55.17

Table 7

5. Interpretation of Results

As it can be seen from **Table 2, Table 3** and **Table4**, **Model 2** has the best results for all the tests, with all the statistic numbers very low comparing with the other models. This means that **Model 2** has a high accuracy. The model with constantly bad results is **Model 3**.

From **Table 5,** it can be observed that **Model 2** generated:

- false signals 0.0% of the time (that is, 0.0 percent of the time when the model predicted a change in direction, it was incorrect)
- missed turning points 6.0% of the time (that is, 6.0 percent of the time when the turning points occurred they have not been predicted)
- the percentage of correct forecast is 97.0 (that is, 97.0 percent of the time when the turning points occurred they have been predicted)

It can be concluded from **Table 5**, **Table 6** and **Table7**, **Model 2** performed very accurately. Therefore, analysing all the test results and the forecast, **Model 2** is the best model for the short-term forecast.

6. Conclusions and Recommendations

The short-term forecast of the daily load plays a crucial role in the day-to-day operations of a utility - the most important being the load management. Therefore a precise and practical forecasting model is required.

The traditional models fail to predict short term demand accurately. Using the results given by **Model 2**, the three day ahead energy consumption can be forecast very accurately. As can be seen from the graphs for **Model 2**, the prediction for the very hot days was very good. At the same time, the method is both very practical and very simple.

Furthermore, in the same way, a model for every season can be created. The model for winter was verified in practice to forecast the last 11 days of July. The forecasts were afterwards compare with the actuals. The results are again very good. The variation between the sum of the forecasted energy consumption and the sum of actual consumption for the above period was insignificant (0.2%).

Having a precise daily load forecast, a model for calculating half-hourly load, as a percentage from the total daily load, can be developed.

In developing a much improved model, we must never lose sight of the basic uncertainties. Unpredictable errors can appear due to the inaccuracy of the weather forecast. At the same time, the impossibility of developing 100% accurate forecasts and the simple fact that we cannot foretell the future must be recognised.

7. References

Bails, D.G., and Peppers, L.C., Business Fluctuations -Forecasting Techniques and Applications, 1993

Roberts, L., Liss, P.S., and Saunders, P.A.H. Power Generation and Environment, 1990

Hanke, J.E., and Reitsch, A.G., Business Forecasting, 1992

Makridakis, S., Willwright, S., and McGee, V., Forecasting - Methods and Applications, 1983

Box, G.E., and Jenkins, G.M., Time Series Analysis - Forecasting and Control, 1976