

An Application of Cycle Regression Analysis Algorithm to Predict Electric Utility Demand

Mayur Mehta, Southwest Texas State University, San Marcos, TX. 78666
Walter Johnston, Southwest Texas State University, San Marcos, TX. 78666

Introduction:

Numerous methods have been set forth for developing forecasting models for predicting electric utility demand. Forecasting electric utility demand has been a particularly difficult problem due to a variety of reasons, including seasonal variations, changes in the consumption pattern, and variation in the population size.

This paper is an attempt to demonstrate the application and power of the cycle regression analysis algorithm in predicting electric utility demand. The method is illustrated using historical data for a major Central Texas utility company. It should be noted that this is only an exploratory analysis in the use of cycle regression analysis algorithm to predict electric utility demand. We are still in the process of validating as well as comparing forecasts prepared by this method with those resulting from other methods designed to analyze cyclical time-series data, such as the Box-Jenkins methodology and cubic/quadratic regression models.

Cycle Regression Analysis Algorithm:

Cycle regression analysis is a family of algorithms for extracting cycles from time series data exhibiting cyclical phenomena. In the past, numerous methods have been used to extract cycles from multi-cyclical data. Among these are spectral analysis, non-linear regression analysis, periodogram method, and fourier analysis [1,2,3,7; 8]. A handful of these included the use of sinusoidal functions to represent and identify cyclical components in time-series data [4, 7, 10; 11]. Three major weaknesses may be observed in these works. The first major weakness pertains to the lack of ease of use of the forecasting technique(s) proposed in the works cited. Most of this research utilize forecasting models that require an in-depth knowledge of mathematics, making it almost impossible for practitioners to use the methods effectively. Secondly, none of the techniques cited enable you to estimate all parameters of cyclical component simultaneously without a priori estimates. Finally, none of the research cited demonstrate the use of their forecasting models in the field of business.

Cycle Regression Analysis Algorithm (CRAA), developed by Simmons [10], overcomes the first two major weaknesses discussed above. CRAA is an easy-to-use, systematic procedure that may be used by practitioners with only a rudimentary knowledge of mathematics. Moreover, CRAA is designed to simultaneously estimate all parameters of cyclical component directly, without a-priori knowledge of cycle parameters. Applications-oriented research by Simmons et. al. [10, 11], have shown cycle regression

analysis algorithm to be powerful and, in most cases, superior to many of the above techniques in providing simultaneous estimates of all parameters of a sinusoidal model. They, however, lack business application of this powerful algorithm. This paper attempts to overcome the third weakness by illustrating the use of cycle regression analysis algorithm in predicting electric utility demand.

Cycle regression analysis employs the Marquardt's Compromise method and some heuristic procedures to develop a non-linear regression model consisting of a linear trend and a set of K sine functions. The method fits a set of time series data to the following model.

$$Y = b_0 + b_1 t + \sum_{i=1}^k b_{3i-1} \sin b_{3i} (t + b_{3i+1}) + \varepsilon \quad (1).$$

The term $(b_0 + b_1 t)$ represents the trend line. The parameters b_{3i-1} , b_i , and b_{3i+1} determine the amplitude, period, and phase, respectively, of the i^{th} cycle. ε is a random variable normally distributed with mean zero and standard deviation, σ .

Marquardt's compromise method [6] is used to estimate the unknown parameters of any non-linear regression model. It thus provides the basis for the cycle regression analysis algorithm to simultaneously estimate the amplitude, period and phase of each cycle in the model (1) directly, without a-priori knowledge of cycle parameters. The general procedure followed in cycle regression analysis algorithm is outlined below:

1. Starting Procedure to estimate trend and the first cycle.

- 1.1. Initial estimates of the slope and intercept are obtained by performing simple linear regression on the data.

$$Y = b_0 + b_1 t \quad (2)$$

- 1.2. Obtain the initial estimate of the period of the first cycle by evaluating the autocorrelations of the residuals from step 1.1.

- 1.3. Use zero as the initial estimate for its phase and amplitude.

- 1.4. Marquardt's Compromise is then used to obtain the final estimates for the trend and the first cycle by fitting the model,

$$Y = b_0 + b_1 t + b_2 \sin b_3 (t + b_4) \quad (3).$$

- 1.5. Perform a partial F-test to determine the significance of the cycle. If this cycle is found to be statistically significant then let $K = 2$ and proceed to step 2. Otherwise the model obtained in step 1.1 is the final model.

2. Estimate the parameters for the K Cycles as follows:

- 2.1. For the initial estimates of the trend parameters, use the estimates obtained in the previous fit.

- 2.2. For the initial estimate of the period, b_{3i} , use the corresponding estimates from the

previous fit for all except the K^{th} cycle. For the K^{th} cycle, evaluate the autocorrelations of residuals of the model obtained in the previous iteration.

- 2.3. For the initial estimate of the phase, b_{3i+1} , use the corresponding estimates from the previous fit for all but the K^{th} cycle. Use 0.1 for the K^{th} cycle.
- 2.4. For the initial estimate of the amplitude, b_{3i-1} , use the sum of the amplitudes obtained in the previous fit divided by K .
- 2.5. Use the Marquardt's Compromise to fit the model in (4) using the initial estimates obtained from steps 2.1 - 2.4 to obtain the final estimates for the trend and K cycles.

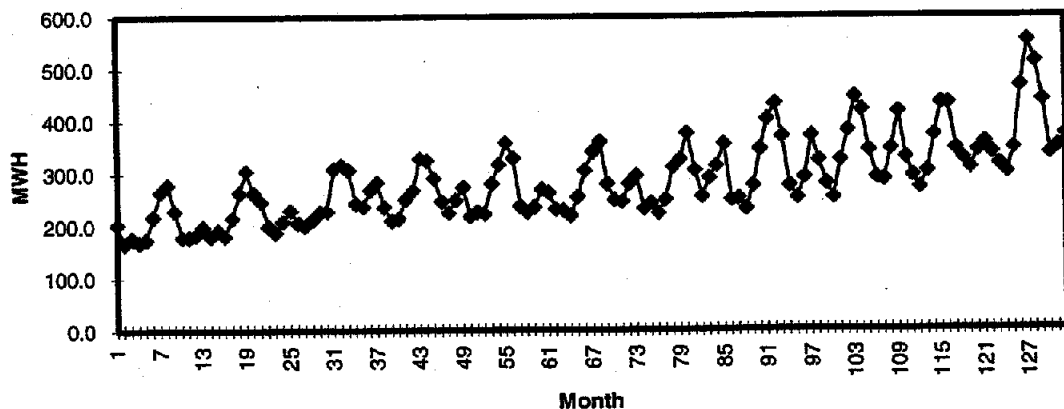
$$Y = b_0 + b_1 t + \sum_{i=1}^k b_{3i-1} \sin b_{3i} (t + b_{3i+1}) \quad (4).$$

- 2.6. Perform a partial F-test to determine the significance of the K^{th} cycle. If it is found to be significant, increase K by 1 and repeat Step 2; otherwise, use the model obtained in step 2.5 above.

Predicting Electric Utility Demand

Cycle regression analysis algorithm was used to predict the average monthly electric utility demand for a large central Texas utility company. The demand is the total MWHs sold to the retail and wholesale customers of the company. Historical data on the average monthly MWH consumption for the period January 1980 through December 1990 were used to prepare a non-linear forecasting model using the cycle regression analysis algorithm. Figure 1 shows a time-series plot of these 132 observations.

Figure 1. Average Monthly Electric Utility Demand (MWH)



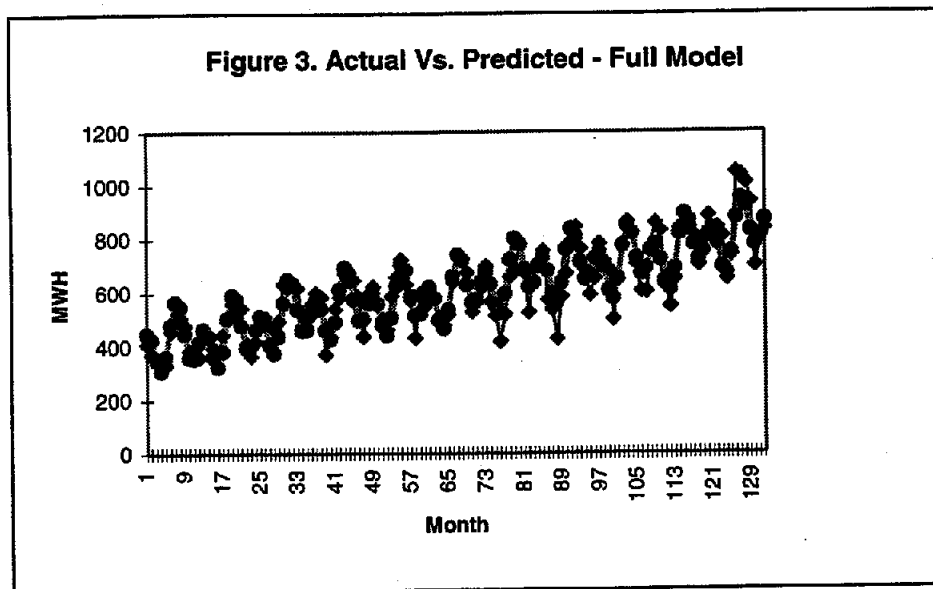
In developing the model, two different procedures were used for the purpose of determining the accuracy of the fit and, therefore, the accuracy of the forecasts. In the first procedure, all of the 132 historical observations on the average monthly demand were included in the analysis.

This procedure resulted in a non-linear forecasting model consisting of trend and three cyclical

components. The lengths of the three cycles were estimated to be 6-, 12-, and 44-months. The overall R^2 was observed to be 88.01%. Figure 2 presents various details of this model while Figure 3 shows a time series plot of the original data and forecasts prepared by this model.

Figure 2: CRAA Forecasting Model to Predict Average Monthly Utility Demand (MWH) based on 132 Observations

After Extracting	Partial F	R-Square
Trend	-----	0.6322
1st Cycle	36.30	0.8030
2 nd Cycle	24.03	0.8750
3 rd Cycle	2.76	0.8801
Model Obtained: $Y = 403.68 + 3.14t + 91.2 \sin 1.05(t + 0.21) + 59.49 \sin 0.52(t + 7.36) + 15.84 \sin 0.14(t + 16.77)$		
Overall R^2	0.8801	
Cycle Length	6-month, 12-month, 44-month	



To better judge the predictive accuracy of the cycle regression algorithm, a holdout sample methodology was also utilized. In this case, historical data on average monthly demand for the first 100 months (January 1980 to April 1987) were used to develop the forecasting model. Data for May 1987 to December 1990 (the last 32 months) were held back to test the predictive accuracy of the forecasting model.

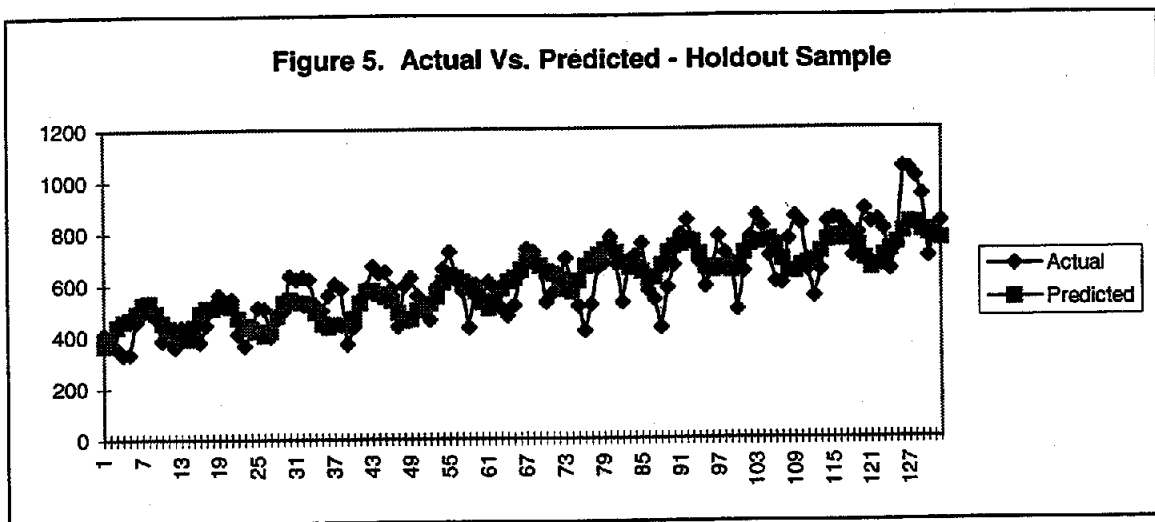
Using this methodology, the cycle regression analysis algorithm was able to successfully identify four

significant cycles, in addition to the trend component. Cycles of length 4-, 6-, 12-, and 78-months were identified by the cycle regression analysis algorithm. The overall R^2 was computed to be 87%. This is a fairly good indication of a strong fit between the data and forecasting model. The detailed results of this method are shown in Figure 4.

Figure 4: CRAA Forecasting Model - Average Monthly Utility Demand (MWH) based on Holdout Sample

After Extracting	Partial F	R-Square
Trend	-----	0.4569
1 st Cycle	36.30	0.6079
2 nd Cycle	24.03	0.8239
3 rd Cycle	2.76	0.8519
4 th Cycle	2.10	0.8700
Model Obtained: $Y = 410.55 + 2.95t$ $+ 64.31 \sin 0.52 (t + 8.12)$ $+ 9.55 \sin 0.92 (t + 6.22)$ $+ 29.28 \sin 0.08 (t + 18.31)$ $+ 14.98 \sin 1.47 (t + 14.84)$		
Overall R^2	0.8700	
Cycle Length	4.27-month, 6.8 -month, 12.08-month, 78.5-month	
R^2 for Holdout sample	0.7683	

The model in Figure 4 was then used to prepare forecasts for May 1987 through December 1990. 32 months that were held back . A comparison of forecasts with the actual values observed during this time frame resulted in an R^2 value of 76.83%. Although not as high as was expected, an R^2 value of 76.83% is still respectable. It is indicative of both the success with which cycle regression analysis algorithm is able to identify cycles contained within multi-cyclical data as well as the accuracy with which it generates the predictions. Figure 5 displays a time-series plot of the actual data versus forecasts generated by this model.



Closing Remarks

Based on this simple exploratory analysis of using the cycle regression analysis algorithm in predicting the average monthly electric utility demand for a large Central Texas utility company, it can be concluded that the cycle regression analysis algorithm is a powerful business forecasting tool. Not only was this algorithm able to predict the average monthly demand to within 10% of the actual values, it was also highly successful in tracking the movement of the data. Obviously, the real power of cycle regression analysis algorithm can only be determined by comparing forecasts generated by cycle regression with those obtained by other methods designed to analyze multi-cyclical data. We are currently in the process of doing this and hope to have the comparative results by the time of the conference.

REFERENCES

1. Bloomfield, P. (1976) Fourier Analysis of Time Series: An Introduction., New York: John Wiley.
2. Bolviken, E. (1982), "New Tests of Significance in Periodogram Analysis," Scandinavian Journal of Statistics, 10, 1-9.
3. Campbell, M.J., and A. M. Walker (1977), "A Survey of Statistical Work on the Mackenzie River Series of Annual Lynx Trappings for the Years 1821 - 1934 and a New Analysis," Journal of the Royal Statistical Society, Ser. A, 140, 411-431.
4. Damsleth, E., and E. Spjotvoll (1982), "Estimation of Trigonometric Components in Time Series," Journal of the American Statistical Association, 77, 381-387.

5. Draper, N. R., and H. Smith (1966), Applied Regression Analysis, New York: John Wiley.
6. Marquardt, D. W. (1963), "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," Journal of the Society for Industrial and Applied Mathematics, 11, 431-441.
7. Morris, M. J. (1977), "Forecasting Sunspot Cycle," Journal of the Royal Statistical Society, Ser. A, 140, 437-448.
8. Ottestad, P. (1979), "The Sunspot Series and Biosperic Series Regarded as Results Due to a Common Cause," Scientific Reports of the Agricultural University of Norway, 58, 1-20.
9. Simmons, L. F., Mehta, M. R., and L. P. Simmons (1983), "Cycle Regression Analysis Made Objective - The T-Value Method," Cycles, 34, 99-102.
10. Simmons, L. F., and D. R. Williams (1982), "A Cycle Regression Algorithm for Extracting Cycles from Time-Series Data," Journal of Computers and Operations Research, 9, 243-254.
11. Simmons, L. F., and D. R. Williams (1983), "The Use of Cycle Regression Analysis to Predict Civil Violence," Journal of Interdisciplinary Cycle Research, 14, 203-215.