INTEGRATION OF EXTRAPOLATION TECHNIQUE AND TIME SERIES STATISTICAL MODEL FOR LONG TERM FORECAST

Sharon H. Coelho,
Student M.E (I.T)
Vivekanand Institute of Technology,
Chembur, Mumbai.

Prof. Gresha Bhatia,
Deputy Head Of Department,
Vivekanand Institute of Technology,
Chembur, Mumbai.

ABSTRACT

Extrapolation is a Numerical Analysis technique primarily used for estimating future variables and trend on the basis of past given data. In this paper, we propose a technique which integrates Time Series Statistical technique with Extrapolation technique for accurate load forecasting. Least Square Method of Time Series Statistical Model is used for generating the pattern and establishing a relation between the past data considered as variables which are then extrapolated resulting in relatively efficient long term prediction. As we will see, this integrated model involves some aspect of artificial intelligence to smoothen the pattern generated using Time Series.

General Terms:

Integration of Extrapolation Technique with Time Series model for efficient long term load forecasting.

Keywords:

Long Term Load Forecasting, Extrapolation, Smoothing Technique, Time Series.

1. INTRODUCTION

Load forecasting is used by power companies to anticipate the amount of power needed to supply the demand of consumers. For decision makers in the electricity sector, the decision process is complex with several different levels that have to be taken into consideration. The most challenging aspect of advanced planning and forecasting is predicting the requirement over a long period of Time ranging from 1 to 5 years and even more.

Although load forecasting for long term depends of several factor such as population, economical growth etc which can be considered as growth rate of an area or a region, it can still be predicted efficiently by efficient study of past consumption patterns of consumers of that region.

In spite of the numerous literatures on Load Forecasting published since 1960s, the research work in this area is still a challenge because of its high complexity. How to estimate the future load with the historical data has remained a difficulty up to now, But with the recent development of new mathematical, data mining and artificial intelligence tools, it is potentially possible to improve the forecasting result.

There are various approaches applied to load forecasting (Kyriakides and Polycarpou, 2007 [1]; Taylor and McSharry [2]). These range from regression-based approaches over time-series approaches towards artificial neural networks and expert systems. Regression models are quite common in load forecasting (Kyriakides and Polycarpou, 2007 [1]) and used to model the relationship between the load and external factors, for instance weather and calendar information or customer types (Feinberg and Genethliou, 2005 [3]).

Hor et al. [4] developed a multiple regression model and analyzed the impact of weather variables on the load demand for England and Wales (Hor et al., 2005 [4]). They used data from 1989 to 1995 for model training and from 1996 to 2003 for testing the accuracy. Their aim was to provide an accurate model for a long-term prediction of the monthly demand.

Time-series approaches (Box and Jenkins, 1970 [5]) are among the oldest methods applied in load forecasting. They can be distinguished on several levels. First of all, there are uni-variate (Taylor et al., 2006; Taylor and McSharry, [2]) and multivariate methods. The former are usually used for very short-term load forecasts whereas the latter are applied for all time- horizons. The time series is usually assumed to be linear. It should be noted, though, that the assumption of linearity usually does not comprise the influence of the temperature. In this case, there appears to be a unanimous agreement that the non-linearity of this relationship has to be preserved in the model.

2. EXTRAPOLATION TECHNIQUE

Extrapolation is the process of estimating, beyond the original observation range, the value of a variable on the basis of its relationship with another variable.[6] This technique is very useful when the future value has to be predicted on the basis of know past values. Extrapolation may also apply to human experience to project, extend, or expand known experience into an area not known or previously experienced so as to arrive at a (usually conjectural) knowledge of the unknown [7]. It can be seen as a combination of Artificial Intelligence and Time Series Statistical Models. The main requirement of using extrapolation over a set of variables is to have or built a certain relationship amongst those variables. It is not possible to extrapolate when the given set of variables are not interlinked.

Given a set of values ranging from x_1 to x_n having some sort of a relation between them, which can be represented by a mathematical derivation or a formula, extrapolation technique gives the values beyond x_n like x_{n+1}, x_{n+2} and so on. Let $X=\{x_1, x_2, x_3, x_4, x_5\}$ be the given set of known variables

The values of "a" and "b "are computed as follows:

$$a = \frac{\sum Y}{n} - b\left(\frac{\sum t}{n}\right)$$

$$b = \frac{n\sum tY - (\sum Y)(\sum t)}{n\sum t^2 - (\sum t)^2}$$

Although extrapolation is a powerful technique for prediction of unknown variables, it is subject to greater uncertainty and a higher risk of producing meaningless results if the relationship between the given set of variables is not properly defined. The success rate of extrapolation thus depends on continuity, smoothness and periodicity of the given data.

3. TIME SERIES TECHNIQUE

Time Series is a fundamental statistical model for forecasting and prediction. It uses historical data to predict the future values. It can also be defined as a collection of data recorded over a period of time—weekly, monthly, quarterly, or yearly. An analysis of history—a time series—can be used by management to make current decisions and plans based on long-term forecasting.

Using Time Series, we could draw the relation between various consumptions of a consumer over the past data. Using Least Square Method, Consumption data (variables) are plotted against time to form a pattern of usage. These patterns are usually Linear and are assumed to continue into the future.

Least Square Method

The long term trend of many consumers often follows a pattern which if plotted gives a straight line which can be represented in the form:

$$Y' = a + b\,t$$

where:

- Y' is the projected value for time t.

- a is the Y-intercept.

- b is the slope of the line.

- t is any value of time that is selected.

having linear relationship between them represented by a graph as shown in fig 1. Extrapolation is used to predict the values x_6, x_7, x_8 using given set of data.

4. PROPOSED TECHNIQUE INTEGRATING EXTRAPOLATION WITH TIME SERIES

For the higher success rate of extrapolation, it is extremely important to establish a relationship between the given set of variables. To achieve this, we propose a new model involving integration of Time Series and extrapolation.

Our main aim in this study, is to predict the future energy requirement based on past consumption data. The variables available to us are the actual energy consumed by a particular set of consumers over a period of last 5 years. Using this data we predict the energy consumption for the next 5 years.

To establish a relation amongst these given variables, we use time series statistical technique called as Least Square Technique which plots the graph of consumption variables against time. This gives us a relation between the given variables.

Based on the relationship, we can use different extrapolation techniques. If the graph formed is Linear, we use Linear Extrapolation Technique. If the graph is a curve, we could use polynomial Extrapolation Technique. In our study on Load Forecasting, most of the consumers, if not all, have a linear pattern consumption over time as the growth rate increases with time, the need and demand for energy and power increases as well with time and hence it's no surprise that in Load Forecasting, we use Linear Extrapolation.

Proposed Extrapolation Formula:

$$y(x) = y\,(x-1) + \frac{1}{2}\sum_{n=x-1}^{2} y\,(n) - y(n-1)$$

Where,

Y(x) is the value to be forecasted in time t.

Y(x-1) is the value at time t-1.

n represents t as time.

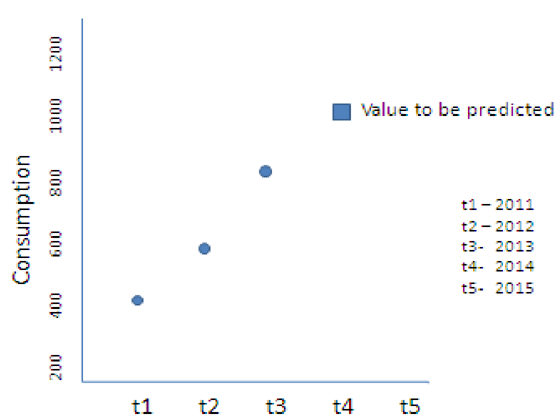

Fig. 1

In the above Fig 1, we have the values of consumption for time t_1, t_2 and t_3. Based on this data we have to predict future long term values t_4, t_5 … representing 2014, 2015…

Calculations:

$x = t_4$

$y\ (t_4) = y(t_3)\ + \frac{1}{2}\sum_{n=3}^{2} y\ (n) - y(n-1)$

$\sum_{n=3}^{2} y\ (n) - y(n-1)\ =\ (y_3 - y_2) + (y_2 - y_1)$

$= (830 - 610) + (610 - 400)$

$= 220 + 210$

= 430 KW

$y\ (t4) = 830\ + \frac{1}{2} * 430$

$= 830 + 215$

=1045 KW

Consumption at t_4(2014) is 1045 KW

Similarly, the consumption at t_5, t_6 and so on could be computed.

5. RESULTS AND ANALYSIS

Predicting a data over a longer period is always more challenging as many factors that could affect the result had to be taken into consideration. Keeping this in mind, if the factors were to change for a time t_4, the effect should be made in t_5. But if the factors do not change too much during t_5 then t_6 should also not reflect too much change in behavior.

This theory is implemented in our proposed Extrapolation over Time Series technique where the latest time has the most effect on the next time interval.

As discussed earlier, the formula for computing consumption is given by,

$$y(x) = y\,(x-1) + \frac{1}{2}\sum_{n=x-1}^{2} y\,(n) - y(n-1)$$

Where,

y(x) is the value to be forecasted in time t.

y(x-1) is the value at time t-1.

Here y(x-1) is a significant factor in computing the consumption at y(x).

The other consumptions prior to y(x-1) are also considered by applying the summation of their differences.

This technique has yielded better results for long term forecast as any change in the parameters during the present year has reflected in the forecast for the next year.

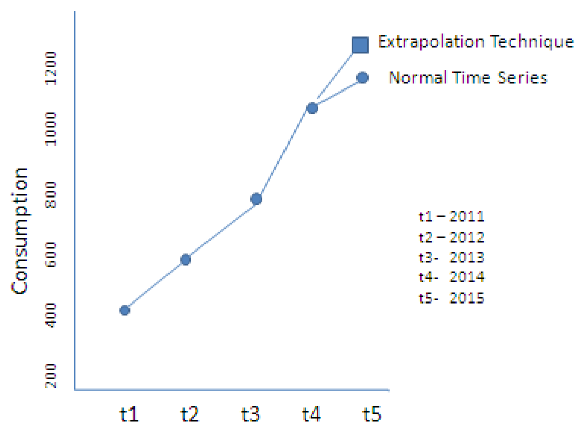

Fig. 2

In the above Figure, The consumption at time t_4 has a steep rise in comparison to t_1, t_2 and t_3. This rise is generally due to change in policy or various other parameters which directly affect the growth rate.

This change in the growth rate should be reflected in the coming year t_5 which again proves that it's the latest year that is more significant that the other years.

If a normal Time series is implemented over the given data as shown in Fig. 1, The consumption at t_5 is computed by computing the slope of the line involving all variables (t_1, t_2, t_3, t_4) and then the Y-Intercept is predicted.

The main problem with this approach is, although there is a steep rise in consumption pattern at t_4, it does not get reflected in t_5.

This same set of data when implemented using the proposed technique, yielded much better results as the steep change in t_4 can be seen in t_5 as well.

The difference does not look too significant in the given scenario as the graph for consumption is somewhat linear in trend. But if a non linear graph or a curve is considered, the difference seen is much larger.

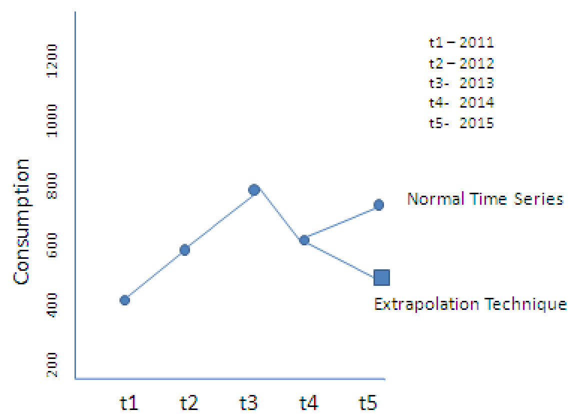

Fig. 3

In the above Figure, with non-linear consumption pattern, Extrapolation Technique gain gives proper results as against to generalized results.

6. CONCLUSION

The proposed Extrapolation over Time Series model yields in much better and efficient results for forecast over longer periods. It gives larger significance to the variable closest to the forecasted one and takes the latest growth rate in to account.

It serves better for linear as well as non linear trends. Coupled along with time series, it certainty is an efficient way for long term predictions.

7. ACKNOWLEDGEMENT

We would like to express our gratitude to all those who gave us the possibility to complete this work related to our research and helped throughout in overcoming the obstacles we encountered.

8. REFERENCES

[1] Kyriakides, E., Polycarpou, M., 2007. Short term electric load forecasting: A tutorial. In: Chen, K., Wang, L. (Eds.), Trends in Neural Computation, Studies in Computational Intelligence, vol. 35. Springer, pp. 391–418 (Chapter 16).

[2] Taylor, J.W., McSharry, P.E., in press. Short-term load forecasting methods: An evaluation based on European data. IEEE Transactions on Power Systems.

[3] Feinberg, E.A., Genethliou, D., 2005. Load forecasting. In: Chow, J.H., Wu, F.F., Momoh, J.J. (Eds.), Applied Mathematics for Restructured Electric Power Systems: Optimization, Control and Computational Intelligence, Power Electronics and Power Systems. Springer, US, pp. 269–285.

[4] Hor, C.-L., Watson, S.-J., Majitha, S., 2006. Daily load forecasting and maximum demand estimation using ARIMA and GARCH. In: International Conference on Probabilistic Methods Applied to Power Systems, 2006, PMAPS 2006, pp. 1–6.

[5] Box, G.E.P., Jenkins, G.M., 1970. Time Series Analysis. Forecasting and Control. Holden-Day, San Francisco.

[6] http://en.wikipedia.org/wiki/Extrapolation

[7] Extrapolation, entry at Merriam–Webster

[8]http://www.forecastingprinciples.com/files/LRF-Ch7b.pdf

[9] M. T. Hagan , Demuth, H.B., Beale, M.H., Neural Network Design, PWS Publishing Company, Boston, Massachusetts, 1996.

[10] J. M. Zurada, Introduction to Artificial Neural Systems, West Publishing Company, Saint Paul, Minnesota, 1992.