

Electric Load Forecasting Using Genetic Based Algorithm, Optimal Filter Estimator and Least Error Squares Technique: Comparative Study

Khaled M. EL-Naggar, and Khaled A. AL-Rumaih

Abstract—This paper presents performance comparison of three estimation techniques used for peak load forecasting in power systems. The three optimum estimation techniques are, genetic algorithms (GA), least error squares (LS) and, least absolute value filtering (LAVF). The problem is formulated as an estimation problem. Different forecasting models are considered. Actual recorded data is used to perform the study. The performance of the above three optimal estimation techniques is examined. Advantages of each algorithms are reported and discussed.

Keywords—Forecasting, Least error squares, Least absolute Value, Genetic algorithms.

I. INTRODUCTION

ONE of the primary tasks of an electric utility is to accurately predict load requirements at all times. Results obtained from load forecasting process are used in different areas such as planning and operation. For example, long-term load forecasting, one to ten years ahead monthly and yearly values, is applied to expansion planning, inter-tie tariff setting and long-term capital investment return problems. While short-term load forecast results, one day to one month ahead hourly and daily values, are needed in unit commitment, maintenance and economic dispatch problems. Thus there is a need for accurate load forecasting techniques in general. First, the model that describes the load growth pattern should be selected, then the parameters of the model should be estimated using historical data. The estimated parameters are then used to predict future load values. Finally the resultant error from the forecasting process should be evaluated.

Many classic approaches have been proposed and applied to long-term load forecasting to estimate model parameters, including static and dynamic state estimation techniques [1],[2],[3],[4]. Methods based on artificial intelligence such as artificial Neural networks and expert systems have been also proposed and shown promising and encouraging results[5], [6]. While the LS technique, has been the most famous static estimation technique and in use for a long time as the preferred technique for optimum estimation in general, some limitation and disadvantages are associated with this approach. For example, in case of the data set is contaminated with bad measurements, the estimates may be inaccurate unless a large number of data points are used.

Khaled M. EL-Naggar is with the College of Technological Studies, Kuwait (e-mail: knaggar@ieec.org).

Khaled A. AL-Rumaih is with the College of Technological Studies, Kuwait.

An other powerful approach to state estimation, is the dynamic approach. Kalman filtering and the least absolute value filtering algorithms are examples of such dynamic approach. Unlike static approaches, where the whole set of data is used to obtain the optimal solution, dynamic filters are recursive algorithms. In recursive filters, the estimates are updated using each new measurements. Dynamic filters are well suited to on-line digital processing as data are processed recursively. They been used extensively in estimation problems for dynamic systems [3]. Dynamic filters have the advantage of their ability in handling measurements that change with time.

Genetic Algorithms have recently received much attention as robust stochastic search algorithms for various problems. This class of methods is based on the mechanism of natural selection and natural genetics, which combines the notion of survival of the fittest, random and yet structured, search and parallel evaluation of the points in the search space. GAs have been successfully applied in various areas such as, load flow problems, fault detection, stability analysis, economic dispatch, power system control [7],[8],[9],[10].

This paper presents performance comparison of three optimal estimation techniques for long-term peak load forecasting in electric power systems. Different forecasting models are considered. The state space representation for each model is presented. The three algorithms, LS, LAVF and GA are then, used to estimate the parameters coefficient of each model using actual recorded data for the Egyptian unified network. Forecasting results are obtained and evaluated.

II. MATHEMATICAL FORMULATION

Load models are developed to mathematically represent the relationship between load and influential variables such as time, weather etc. The coefficients of the model formulated are identified and used to predict the future loads by extrapolating the relationship to desired lead time. The final accuracy of the forecast process depends on the model selected and the accuracy of the estimated parameters. Reviewers of load forecasting models have found that techniques almost in use today, can be categorized as being of multiple regression, general exponential smoothing and statistical methods [2].

Regression analysis or trend analysis is the study of the behavior of a time series or process in the past and its mathematical modeling so that the future behavior can be

extrapolated from it. A time variant event such as power system load can be broken down into four components, basic, trends, seasonal variations, cyclic variations and random variations. The last three variations have a long-term zero mean. Regression curves used in power system load forecasting are; linear, polynomial, exponential and power. The term linear is used to mean either straight line relation between two variables or a model in which the parameters appear linearly. In general a multi-variable regression model can be related to $n+1$ independent variables (repressors) and can be written as:

$$P(t) = a_0 + \sum_{i=1}^n a_i t^i + r(t) \quad (1)$$

where $P(t)$ is the peak load demand at time t , a_0 , a_i are the regression coefficients relating the load $P(t)$ to the time t . $r(t)$ is the residual load at year (t) . Although the relationship between P and t may be non-linear for $i=2,3,\dots$ the model is still said to be linear since t and t^2 can be transformed into Y_1 and Y_2, \dots where $Y_1=t^2$, $Y_2=t^3, \dots$

Another type of regression technique involves nonlinear regression models. Nonlinear regression models are not linear in terms of the parameters and can not be made so by any transformation.

For many years, generation planners have used regression techniques as an aid in predicting annual peak system demands. Peak demands are known to be influenced by weather conditions, number and type of consumers and general economic conditions. However, a simple relationship in which demand increases exponentially with time is generally found to yield an adequate forecast for system peak demand. Forecasts are frequently obtained from the following simple relationship:

$$P(t) = e^{a+bt} \quad (2)$$

Simply this equation can be transferred into a linear form by having the (\ln) of both sides.

In order to identify the most adequate model for forecasting application among all available linear and nonlinear regression models, different types of graphs must be examined. A visual inspection of a graph of a given observation against time can often reveal both obvious and less apparent characteristics of the data. After estimation process, the resultant residuals are subjected to whiteness test. The objective of such test is to insure that the selected model adequately describes the given data series [4].

A. Whiteness Test

The objective of the whiteness test is to ensure that a selected model adequately describes a given set of data. The whiteness test can be achieved by the following two steps:

- 1- examination of the estimated residual graph(exploratory analysis); and
- 2- calculation of the residual autocorrelation function(RACF) at different time lags (confirmatory analysis)

The RACF can be calculated as:

$$RACF_k = \frac{\sum_{t=k+1}^m \omega_t \omega_{t-k}}{\sum_{t=1}^m \omega_t^2} \quad (3)$$

where, $RACF_k$ is the RACF at time lag k ,

ω_t is the estimated residual at time t

The RACF value ranged from -1 to +1. If a given value (rather than the first one) is significantly different from zero, it will fall outside a confidence interval level [4].

Before beginning the forecasting process, one must select a forecasting method, construct a model and finally test the constructed model. As mentioned before, the regression technique is the most widely used because of its simplicity and ease of use. Therefore, this technique is considered for modeling. It is very important to emphasize that the primary objective of this paper is to present the application of the three techniques to load forecasting problem and evaluate the results obtained. The objective is not to present different models and comparison between them. However, the error analysis will show that the selected model is appropriate.

To identify the most appropriate regression model, the data set given in reference 3, must be examined. A perusal of figure 1 reveals that the data set is nonstationary, since the demand is always increasing with time. Furthermore, the pattern of increasing indicates that the best load growth model for such data is the multiple linear regression representation given by equation 1 using $I=1$ or 2 or 3. Therefore, in this paper three model are considered, i.e $i=1$ or 2 or 3. Given the peak load (P) at each year T , an equation just like equation 1 can be written for each load. If the data set consists of m years and the corresponding peak load, then there will be (m) equations in (n) unknowns. This system of equation is an overdetermined system ($m>n$). Then for m years, a discrete system of equations in state space form can be written as:

$$Z(t) = H(t) X + r(t) \quad (4)$$

where : $Z(t)$ is the load demand vector
 X is the parameter vector to be estimated
 $r(t)$ is the error vector associated with $P(T)$
 $H(t)$ is a row vector that relates $P(T)$ to X

In this study the three models used are;

Model 1 ($i=1$)

$$H(t) = [1 \quad T], \quad T=1,2,\dots,m \quad \text{and} \quad X = [A \quad B]^T \text{Transpose}$$

Model 2 ($i=2$)

$$H(t) = [1 \quad T \quad T^2], \quad T=1,2,\dots,m \quad \text{and} \quad X = [C \quad D \quad E]^T \text{Transpose}$$

Model 3 ($i=3$)

$$H(t) = [1 \quad T \quad T^2 \quad T^3], \quad T=1,2,\dots,m \quad \text{and} \quad X = [F \quad G \quad I \quad J]^T \text{Transpose}$$

Now, the problem is to find an estimate for the parameter vector X for any model, that minimize the error vector $r(t)$.

III. GENETIC ALGORITHMS (GA)

Genetic algorithms are a numerical optimization technique. More specifically, they are parameter search procedures based upon the mechanics of natural genetics. They combine a Darwinian survival-of-the-fittest strategy with a random, yet structured information exchange among a population of artificial "chromosomes". This technique has gained popularity in recent years as a robust optimization tool for a variety of problems in engineering, science, economics, finance, etc. GA accommodate all the facets of soft computing, namely uncertainty, imprecision, non-linearity, and robustness. Some of the attractive features can be summarized here in the next paragraph.

Learning: GA are the best known and widely used global search techniques with an ability to explore and exploit a given operating space using available performance (or learning) measures. **Generic Code Structure:** GA operate on an encoded parameter string and not directly on the parameters. This enables the user to treat any aspect of the problem as an optimizable variable. **Optimality of the Solutions:** In many problems, there is no guarantee of smoothness and unimodality. Traditional search techniques often fail miserably on such search spaces. GA are known to be capable of finding near optimal solutions in complex search spaces. **Advanced Operators:** This includes techniques such as niching (for discovering multiple solutions), combinations of Neural, Fuzzy, and chaos theory, and multiple-objective optimization.

The GAs approach presented in this work is employed to find the optimum values of the state vector X that minimizes the absolute summation of the forecasting error $r(t)$. In order to emphasize the "best" string and speed up convergence of the iteration procedure, fitness is normalized into range between 0 and 1. The fitness function (ff) adopted is [6]:

$$ff = \frac{1}{1 + k \sum_{t=1}^m |r(t)|} \quad (5)$$

where k is a scaling constant ($k=0.0001$ in this study)

Like other stochastic methods, the GA has a number of parameters that must be selected. Size of population, probability of crossover and probability of mutation. Extensive runs showed that the following values are appropriate for this study: population size=50, probability of crossover = 0.8 and probability of mutation=0.04.

IV. LEAST ABSOLUTE VALUE FILTER (LAVF)

The complete derivation of the proposed filter equations is beyond the scope of this paper and is given in reference [12]. The dynamic filter works on the discrete state space model described by the measurement equation given as equation 4 and the state transition equation in the following form.

$$X(t+1) = \Phi(t)X(t) + w(t) \quad (6)$$

As mentioned before the measurement error vector $r(t)$ is assumed to be white sequence with known covariance. The covariance matrix for $w(k)$ is given as:

$$E\{w(k)w(j)^T\} = \begin{cases} 0 & ; j \neq k \\ Q(k) & ; j = k \end{cases} \quad (7)$$

The estimate is updated using the filter gain matrix $K(k)$ at step (k) as

$$\hat{X}(t) = \bar{X}(t) + K(t)[Z(t) - H(t)\bar{X}(t)] \quad (8)$$

The process is repeated until the last measurement is used. It is assumed that the covariances and the transition matrices are known.

It is very important to mention here that the difference between the proposed least absolute value filter (LAVF) and Kalman filter (KF) method lies in the gain equation, due to the difference in the nature of the objective function used in deriving the filter equation. In KF, the function is the weighted least square error, but in LAVF, the function is the weighted least absolute error [3].

V. LEAST ERROR SQUARES (LS)

The well known LS technique is used to minimize the sum of square of the residuals $r(t)$ of equation 4. The LS solution of an overdetermined system of equation such the one described by equation 4 is given by [5],[11]:

$$X = [H^T H]^{-1} H^T Z(T) = H^+ Z(T) \quad (9)$$

Where H^+ is the left pseudo inverse matrix of H .

VI. PRACTICAL APPLICATION AND RESULTS

The proposed method is used to forecast the peak load demand of the unified Egyptian Network. Actual recorded data is used to perform the study. The data given in reference 3 is plotted in figure 1. This figure represents the peak load demand of the national unified power system of Egypt during the period from 1977 to 1993 [3]. The three forecasting models discussed earlier are used to represent the load growth. The data set is divided into two parts. The first twelve years, up to the year 1988, are used to establish an overdetermined system of equations. This system of equations is solved using the three estimation techniques to find the optimal parameters for different models parameters. The next part of the data set, from 1989 to 1993, is used to evaluate the estimation process. This is simply by using the parameters obtained from the estimation process to forecast the peak load during the period 1989-1993 and compare those values with the actual data given in the reference.

A. Model Adequacy

The three methods are used to identify each model parameters. The three models, namely, linear, second and third order polynomials (models 1,2,3 respectively) are used to find an estimation for the parameters in each case via GA, LAVF and LS approaches. Estimated parameters obtained for each model using the three methods are then used to forecast the loads during the period 1989-1993. Results obtained using the three methods for forecasting confirms that the best model that represents this data set is Model 2. This conclusion was also reported in reference 3. In order to insure this, the

whiteness test described before is applied. Examinations of the estimated residual and the RACF graphs shown in figures 2 and 3, reveals that the model 2 is appropriate for forecasting the load of the given data set. The RACF is not significantly different from zero; i.e. The model have estimate and removed the pattern of the given data set, and what is left over is white noise.

Figure 4 gives sample of the results obtained for the peak load forecasting error for the three model using GA. It is very clear that the best model that describes the load growth accurately and gives less error in forecasting process, is model 2. It gives maximum error about 2.75%.

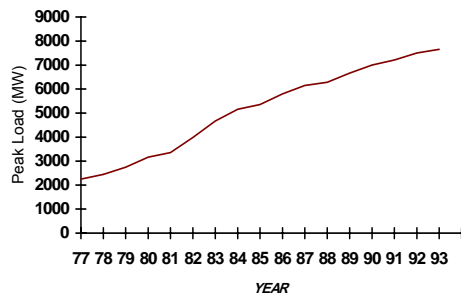


Fig. 1 Annual peak demand

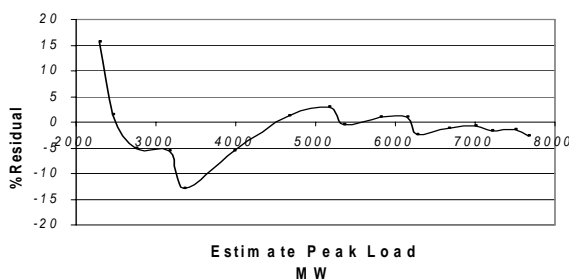


Fig. 2 Estimate residual via model2 Using GA

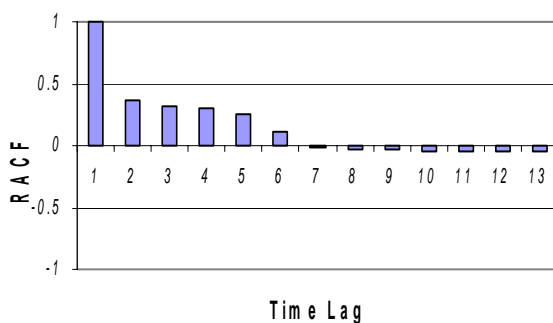


Fig. 3 RACF via model2 using GA

B. Comparative Study

In this section model 2 is used to set-up a comparison between GA solution and least error squares (LS) and least absolute filter (LAVF) solutions. The following table gives the estimation of the model parameters using the three methods after solving the overdetermine system of equations using the first twelve years data.

TABLE I
PARAMETERS ESTIMATION USING MODEL 2

	GA	LS	LAVF
C	1426.65	1651.10	1623.9
D	508.86	404.70	486.70
E	-7.644	0.0185	-6.736

TABLE II
PERCENTAGE ERROR IN FORECASTED LOADS

Year	Actual MW	GA	LAVF	LS
1989	6664	1.30	2.23	3.76
1990	7004	0.70	1.69	4.51
1991	7215	1.73	2.60	7.10
1992	7503	1.44	2.43	8.48
1993	7657	2.75	3.89	11.5

Forecasted loads using the three methods are compared to the exact recorded data in figure 5. The corresponding percentage absolute error, which is given in table 2, is also presented in figure 6. Examining table 2 and figures 5, 6 reveals that, the estimation made via GAs approach, is much closer to the exact value than the others. The maximum error in estimated loads occurs at the year 1993 and it was 2.75%, only. The maximum error associated with other methods is 3.89% for LAVF and 11.5% with the LS method.

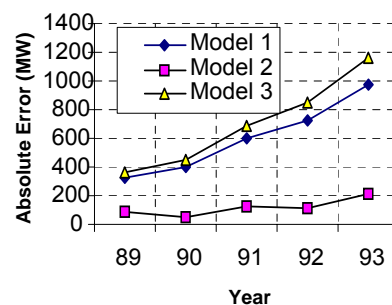


Fig. 4 Absolute error in peak load forecast using GA technique

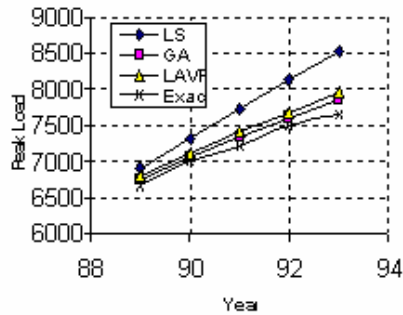


Fig.5 Peak Load Forecast: Comparison

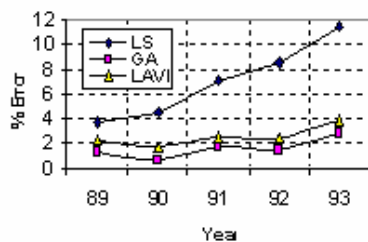


Fig. 6 Percentage Error: Comparison

C. Bad Data Effects

The performance of the three algorithms is examined when the data set is contaminated with bad data points. Two tests are performed. In the first test one random point of load data is changed deliberately. This represents about 8% of the total data. The load at the year 1979 is changed to 7242 instead of 2742. The three algorithms are tested using the new data set and the resultant parameters are used to calculate the loads during the period 1989-1993. The maximum resultant errors are found to be 40%, 14% and 6% for the LS, GA and LAVF respectively. In the second test, 2 points of bad data are introduced, the resultant errors are increased to 76%, 34% and 24% respectively. It is clear that the LS method is badly affected with bad data while the LAVF is less sensitive to the presence of these bad points. GA gives reasonable results when the data set is contaminated with less than 10% of bad points. If the data set is contaminated with more than 10% of bad points, all algorithms will give poor results.

Table 3 summarizes the performance of the three algorithms. In this table (t) represents the CPU time needed for GA calculations. Indeed this time depends on several factors such as computer type, software used and GA parameters selected. In this work, with the GA parameters mentioned before, and using PC, 1000 MHz processor, t was in the order of about 40 seconds. Since the load forecasting calculations are always performed off-line, the calculation time in such application is not important as the accuracy. Therefore, the GA approach can be considered as a powerful tool for load forecasting.

TABLE III
PERFORMANCE EVALUATION

	GA	LAVF	LS
Calculation time	t	0.7 t	0.2 t
Accuracy	high	low	lower
% error (bad data)	reasnable	low	high

VI. CONCLUSIONS

This paper presents the application of GA, LAVF and LS algorithms for long term load forecasting in power systems. The problem is formulated as an optimization problem. The solution framework was implemented and tested using actual recorded data. Three different models were used and the quadratic model was proven to be the best one that reopresents the data available. This model is then used with the actual recorded data to test the performance of the three algorithms. The forecast using the GA method has been compared with those obtained with other methods. Forecasting results using GA were found to be the best. This indicates that the GAs approach is quite promising and deserves serious attention because of its robustness and suitability for parallel implementation..

REFERENCES

- [1] Load Forecasting Bibliography: Part 1, IEEE Trans.,Power APP., Vol. PAS-99, #1,Jan.1980.
- [2] H. Willis and J.E.D. Northcote, *Comparison tests of 14 distribution load forecasting methods*, IEEE Trans., PAS, Vol-PAS 103 , #3, June 1994, pp 1190-1197.
- [3] K.M.EL-Naggar, *Long -Term Load Forecasting Using Recursive Algorithm*, Fifth International Middle East Power Conference (MEPCON97), Alexandria, Egypt, Jan. 1997, pp-320-324.
- [4] H.K. Temraz, K.M.EL-Naggar and M.M.A. Salama, *Application of Noniterative least Absolute Value Estimation for Forecasting Annual Peak Electric Power Demand*, Canadian Journal for Electric & Computer Engineering, Vol. 23, No.4, Oct. 1998, pp.141-146.
- [5] Kandil, M.S., El-Debeiky, S.M., Hasanien, N.E. , "Long-term load forecasting for fast developing utility using a knowledge-based expert system Power Systems", IEEE Transactions on PS, Vol. 17, Issue:2, May 2002, pp. 491 – 496.
- [6] Parlos, A.G.; Oufi, E.; Muthusami, J.; Patton, A.D.; Atiya,A.F.; "Development of an intelligent long-term electric load forecasting system", Proceedings of the International Conference on Intelligent Systems Applications to Power Systems (ISAP '96), 28 Jan.-2 Feb. 1996, pp.288-292.
- [7] Khaled M. EL-Naggar and Khaled AL-Roumaih, *Genetic Based Algorithm Approach to Long-Term Load Forecasting*, Six International Middle East Power Systems Conference (MEPCON98), Mansoura, Egypt, Vol.1, 1998, pp.237-241.
- [8] Po Hunng and H. C. Chang , *Large-Scale Economic Dispatch By Genetic Algorithms*, IEEE Trans. Power Systems, Vol.10, No. 4 Nov. 1995, pp. 1919-1927.
- [9] Reformate M., Kuffel E., Woodford D. and Pedrcyz W., *Application of Genetic Algorithms for Control Design in Power Systems*, Generation, Transmission and Distribution IEE- Proce., Vol.145- 4, July 1998, pp. 345-354.
- [10] F. Wen and H. Zhenxiang , *Fault Section Estimation in Electric Power Systems Using Genetic Algorithm*, Electric Power Systems Research Journal, Vol. 34, 1995, pp.165-172.
- [11] E.A. Abu-AlFeilat, *Performance Estimation Techniques for Power System Dynamic Stability using Least Squares, Kalman Filtering and Genetic Algorithms*, Proc. of SoutheastConf.2000, IEEE conf. Proc, pp.489-492.
- [12] G.S. Christensen and S.A. Soliman, *Optimal Filtering of Linear Discrete Dynamic Systems Based on Least Absolute Value Approximations*, Automatica, Vol. 26, No. 2, 1990.