A NOVEL ALGORITHM FOR LONG-TERM LOAD FORECASTING OF DISTRIBUTION NETWORKS UNDER REDEVELOPMENT CONDITIONS

Amir MOSHARI
NRI – IRAN
a.moshari@ec.iut.ac.ir

Zahra MADIHI
NRI – IRAN
zmadihi@nri.ac.ir

Niki MOSLEMI
NRI – IRAN
nmoslemi@nri.ac.ir

Davood JALALI
NRI – IRAN
djalali@nri.ac.ir

ABSTRACT

Multiple regression based methods are simple and popular trending methods for distribution long-term load forecasting which are employed in the cases that advanced land use data are not available. However, this method does not work properly for areas which are under redevelopment conditions. This paper proposes a novel algorithm, which can be employed in multiple regression based techniques in such conditions. This algorithm recognizes the redevelopment event in the small areas and shifts the S-shaped load growth curve to a suitable lower level using a decision making engine. The results show that employing this algorithm can significantly improve the accuracy of regression-based load forecasting methods under redevelopment conditions.

INTRODUCTION

In order to plan the efficient operation and economical capital expansion of an electric power distribution system, the system owner must be able to anticipate the need for power delivery - how much power must be delivered, and where and when it will be needed. Such information is provided by a spatial load forecast, a prediction of future electric demand that includes location (where) as one of its chief elements, in addition to magnitude (how much) and temporal (when) characteristics [1]. A vast variety of load forecasting methods have been applied to spatial electric load forecasting. Most are variations on one of two basic themes: Trending/Multi-Regression based, which involves extrapolating past load growth into the future; or Simulation/Land-Use based, which involves modeling the process of load growth itself [1]. Although simulation is the newer, and generally more effective, of the two, it requires large amounts of advanced data, i.e. economical, statistical and regional data, and therefore it is not applicable in the cases that such data are not available.
As the long-term load forecasting of distribution networks is a relatively practical matter there are much less researches and papers in this area compare with short-term load forecasting. Most of researches have employed simulation methods which needs advanced land use /economical data [2-5]. Trending methods work with historical load data, extrapolating past load growth patterns into the future. Despite the minimum needs of such methods to the advanced data, they are suffering from accuracy problem which becomes worse by extending the forecasting horizon. Some references have tried to proposed modifying techniques to improve the accuracy of trending methods [6-8]. Improved multiple regression (IMR) method is one of a successful trending based methods which employs various modifying algorithms such as load transfer coupling (LTC) and horizon year load estimation, to improve the accuracy of trending based load forecasting [1]. However, this method does not work properly for areas which are under redevelopment conditions.
In the existing power grids, almost 25% of load growth in large cities occurs due to the replacement of old land uses with high density land uses. This phenomenon, which is called "redevelopment", usually occurs in areas where the average building age is 35 years or older, but occasionally in areas only about half that old [1]. Generally, because the load densities of horizon year for under redevelopment areas are different from normal areas, multiple regression based techniques will not be able to predict the load pattern properly. Although the redevelopment is not a global issue in the existing power networks, but it will be a grid-wide problem due to the gradual transition from today's power networks to the smart grids. Specially, the emerging of the advanced transportation technologies like electric hybrid vehicles (EHVs) in the future will result in massive demand rises in the power networks. Though the smart grids will provide all the required data for advanced load forecasting methods and therefore using the regression based methods will not be reasonable anymore, in the meanwhile that the advanced communication and data management infrastructures are not implemented in the power networks, employing these methods will be continued. Therefore, it is necessary to extend them to be able to forecast the load in the redevelopment conditions. This paper proposes a novel algorithm, called "Redevelopment algorithm", which can be employed in multiple regression based techniques to forecast the load properly under redevelopment conditions.
The remaining parts of this paper are organized as follows. Section II describes the Improved multiple regression (IMR) method. Section III discusses the redevelopment issue and proposes a modifying technique to address it in trending based methods. The results are shown in section IV. Conclusions are drawn in Section V.

IMPROVED MULTIPLE REGRESSION METHOD

Improved Multiple Regression (IMR) is a trending based method which employs several modifying algorithms such as load transfer coupling (LTC) and horizon year load estimation, to improve the accuracy of trending based load forecasting [1]. The load growth in an area depends highly on the size of that area. The trend in large areas follows a straight line, while it will be an S-shaped curve for a small area (Fig. 1) [1]. The S curve has three distinct phases: The "dormant period", which is the time "before growth", when no load growth occurs; the "growth ramp" which shows a relatively rapid growth, because of new construction in the small area; and the saturated period, during which the small area is fully developed and load growth may continue, but at a very low level. What varies most among the thousands of small areas in a large utility service territory is the timing of their growth ramps.

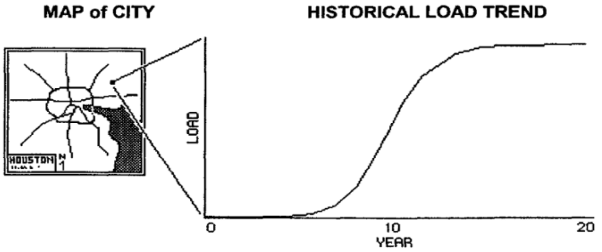

Fig. 1: Load growth trend for small areas [1].

Regression-based methods involve extrapolating past load growth into the future. There are a wide number of polynomials that he could use for the curve fit, but among them, the most suitable for small area load forecasting is the four-term cubic log equation using seven years of historical load data[1]:

(1) $l_n(t) = a_n + b_n.log(t) + c_n.log(t)^2 + d_n.log(t)^3 \quad ,t = 1,2,...,7$

Which, $l_n(t)$ is the annual peak load estimate for substation n for year t (which beginning with t =1 for the first year of load history) and a_n, b_n, c_n, and d_n are the polynomial coefficients which should be estimated using regression. Multiple regression curve fitting would begin with a parameter matrix seven elements high (for the seven years of data) by four elements wide (for the four coefficients to be determined). For example, if the cubic equation is used instead of cubic log equation of (1), the parameter matrix will be as follows:

(2)
$$P_n = \begin{array}{c c} year & \begin{array}{cccc} 1 & t & t^2 & t^3 \end{array} \\ \begin{array}{c} 1\\2\\3\\4\\5\\6\\7 \end{array} & \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 81 \\ 1 & 5 & 25 & 125 \\ 1 & 6 & 36 & 216 \\ 1 & 7 & 49 & 343 \end{bmatrix}_{7\times 4} \end{array}$$

The annual peak loads of Substation n for the past seven years are placed into a seven elements vector:

(3)
$$L_n = \begin{bmatrix} L(1) \\ L(2) \\ \vdots \\ L(7) \end{bmatrix}_{7\times 1}$$

The coefficients a_n, b_n, c_n, and d_n, that best fit the polynomial to the load history, are determined as follows:

(4)
$$C_n = \begin{bmatrix} \hat{a}_n \\ \hat{b}_n \\ \hat{c}_n \\ \hat{d}_n \end{bmatrix} = \left[P_n^T.P_n\right]^{-1}.P_n^T.L_n$$

The classical regression based load forecasting have poor results when there are load transfers between feeders and also in some cases they cannot properly simulate the S-shape of the load growth trend. These problems are addressed in IMR method using load transfer coupling (LTC) regression and horizon load estimation technique. A polynomial curve fit to historical data can yield quite different results, depending on exactly where in the S curve pattern of load growth the load history happens to lie [1]. One way to reduce this type of over-extrapolation is to use horizon year loads, i.e. to estimate future load values and put it into the data set to be curve fit (treated like the historical data). In fact, using this technique, the extrapolation problem will be treated as an interpolation problem. Fig. 2 shows that using estimated horizon load can improve the extrapolation process.

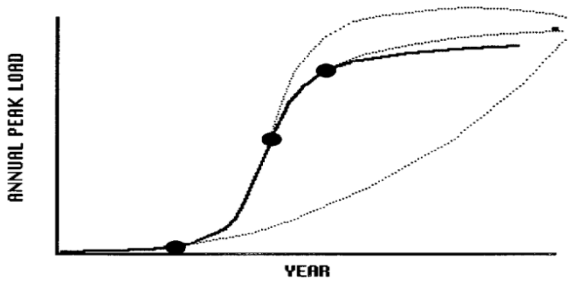

Fig. 2: the effect of horizon load estimation on curve fitting.

In IMR method, there are three methods to estimate the horizon load: based on feeders length, based on customers type, and based on k-means clustering of small area characteristics. For each area, the best method will be determined based on the previous performance of each method and the available data.

The basic multiple regression curve fitting method is usually overreact wildly to load transfers. However, a modification, called load transfer coupling (LTC) regression, can substantially reduce errors. LTC can accommodate situations when the direction of the transfer (to or from the substation) is not known with certainty, and when no idea of the amount of the load transfer is available. This makes it practical [1]. It is possible to solve the multiple regression curve fit (i.e., determine the coefficients) for two or more small areas simultaneously. Consider small areas n and m, corresponding to the two substations each with a seven year load history. In this case, the equation (4) will be modified as follows:

(5) $$C_{n,m} = \left[P_{n,m}^T.R.P_{n,m}\right]^{-1}.P_{n,m}^T.R.L_{n,m}$$

(6) $$P_{n,m} = \begin{bmatrix} P_n & 0 \\ 0 & P_n \end{bmatrix}$$

(7) $$L_{n,m} = \begin{bmatrix} L_n \\ L_m \end{bmatrix}$$

If R is equal to the identity matrix, then it makes no impact on the resulting solution, and the coefficients obtained are identical to those can be found separately. But, if the elements corresponding to the load transfer between *n* and *m* areas are changed from 0 to 1, the load transfer effects will be lessened during regression process [1].

REDEVELOPMENT ALGORITHM

In most metropolitan areas, a portion of load growth occurs due to redevelopment. In this case, old buildings are torn down and replaced by newer, taller, and denser construction, generally with a great deal more demand for electric power. This condition usually occurs in areas where the average building age is 35 years or older, but occasionally in areas only about half that old. Fig. 3 shows the peak load history in an area that actually experienced three of these transitions in only a 60-years period. Generally, because the load densities of horizon year for under redevelopment areas are different from normal areas, multiple regression based techniques will not be able to predict the load pattern properly.

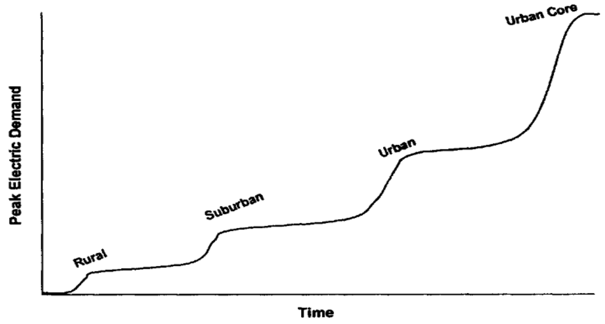

Fig. 3: Load growth pattern during redevelopment.

As mentioned before, although the redevelopment is not a global issue in the existing power networks, but it will be a grid-wide problem due to the gradual transition from today's power networks to the smart grids. Specially, the emerging of the advanced transportation technologies like electric hybrid vehicles (EHVs) in the future will result in massive demand rises in the power networks. It may also change the nature of redevelopment events and shorten the time interval between two transitions. Though the smart grids will provide all the required data for advanced load forecasting methods and therefore using the regression based methods will not be reasonable anymore, in the meanwhile that the advanced communication and data management infrastructures are not implemented in the power networks, employing these methods will be continued. Therefore, it is necessary to extend them to be able to forecast the load in the redevelopment conditions.

As can be seen above, just like the normal conditions, the load growth in each redevelopment transition follows the S-shaped curve. In other words, the load growth curve of these areas is almost the same as the curve of the normal small areas which have been shifted up. Prediction of redevelopment event in a small area needs advanced data such as regional economic indices, penetration rate of new technologies, land use information, etc. However, the regression based load forecasting methods are usually employed when such data are not accessible.

This paper proposes a novel algorithm, called "Redevelopment algorithm", which can be employed in multiple regression based techniques to forecast the load properly under redevelopment conditions. This algorithm recognizes the redevelopment event in the small areas and shifts the S-shaped load growth curve to a suitable lower level. For each small area, it uses the load average (LA_i), the load growth (GL_i), the growth ratio (GR_i), the average of horizon year load estimation in the whole region (HL_T), the residual value ($R^{LTC}{}_i$) [1], and occurring redevelopment in neighboring small areas in previous years ($Red^i{}_j$), as the inputs, and predicts/recognizes the redevelopment event based on some pre-established rules in a decision making process. At first, the algorithm calculates the values of input parameters. For the small area *i*, the load growth of year *t* can be calculated as follows:

(8) $$GL_i(t) = \frac{L_i(t) - L_i(t-1)}{L_i(t)}$$

The growth ratio is defined as (9):

(9) $$GR_i = \frac{GL_i(7)}{GL_i(7) + GL_i(6) + GL_i(5)}$$

The residual value for year *t* is the difference between the actual load $L_i(t)$ and the modified load based on LTC regression $\hat{L}_i^{LTC}(t)$, as follows:

(10) $$R_i^{LTC}(t) = \frac{L_i(t) - \hat{L}_i^{LTC}(t)}{L_i(t)}$$

The average load up to year *T* is defined as follows:

(11) $$LA_i(T) = \frac{\sum_{t=1}^{T} L_i(t)}{T}$$

The parameter Red^i_j determines that if in the previous years a redevelopment event has occurred in neighbor j of the small area i (=1) or not (=0).

Then, the algorithm checks the following conditions for the last year of the historical load (t = 7) and predicts/recognizes the redevelopment based on their weighted combination:

- $GR_i > GR^R \quad ,GR^R \in (0.5,1)$
- $GL_i(7) > GT^R \quad ,GT^R \in (0.1,0.5)$
- $R_i^{LTC}(7) > E^{LTC} \quad ,E^{LTC} \in (0,0.05)$
- $\dfrac{L(7) - LA_i(6)}{LA_i(6)} > LR^R \quad ,LR^R \in (1.2,2)$
- $\dfrac{L(7) - HL_T}{HL_T} > LR^H \quad ,LR^H \in (1.2,2)$
- $\sum\limits_{j=neighbors} Red^i_j > EN^R \quad ,EN^R \in (1,5)$

The constants (GR^R, GT^R …) are determined based on trial and error and expert knowledge. Since the algorithm employs several rules and decide based on a weighted combination of them, the value of these constants is not strict. The ranges which are shown for these constants are determined using sensitivity analyses.

After the above event recognition, the shifting level of the load curve is determined by averaging the load of the years 1 to 6 ($LA_i(6)$). After shifting the load curve to a proper level and forecasting the load using the normal IMR method, the historical and forecasted loads will be restored to their original level.

SIMULATION RESULTS

This algorithm was employed for a distribution test network which has been established in Niroo Research Institute [13]. This network consists of 219 small areas. The following figures show two examples of the result of IMR based load forecasting with and without employing redevelopment algorithm, for under redevelopment small areas in the test network. Fig. 4 shows that the mean absolute percentage error (MAPE) without using the redevelopment algorithm is 28.51%, while it is 3.95% using redevelopment algorithm. Also Fig. 5 shows another example in which the MAPE without using the redevelopment algorithm is 24.11%, and it is 4.10% using redevelopment algorithm. As can be seen, the estimation of the horizon year load is very poor without using redevelopment algorithm. The results showed that employing this algorithm can improve the averaged load forecasting error for under redevelopment areas by 17% over the test network.

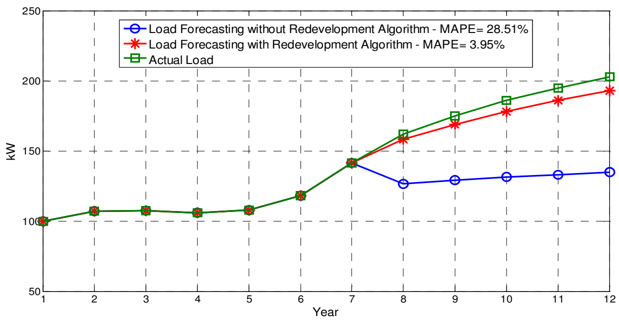

Fig. 4: the effects of the proposed algorithm on small area load forecasting under redevelopment condition.

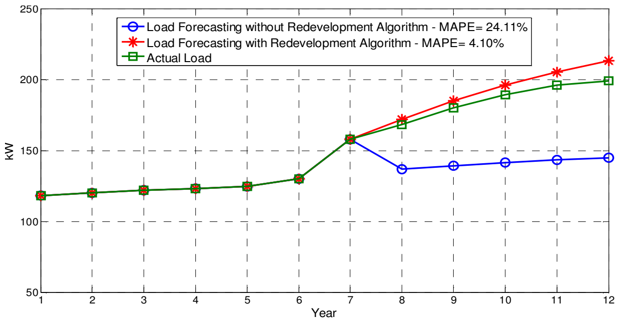

Fig. 5: the effects of the proposed algorithm on small area load forecasting under redevelopment condition.

REFERENCES

[1] H. L. Willis, *Spatial Electric Load Forecasting* - Second edition, Marcel Dekker Inc., 2002.

[2] L. Zhou, T. Wei, M. Fan, Z. Zhang, 2008, "Study on the practical spatial load forecasting based on geographic information system", *China Inter. Conf. on Electricity Dist. (CICED 2008)*, 1-4.

[3] J.D. Melo, E.M. Carreno, A. Padilha-Feltrin, 2012, "Multi-Agent Simulation of Urban Social Dynamics for Spatial Load Forecasting", *IEEE Trans. on Power Systems*, Vol. 27, Issue 4, 1870 - 1878.

[4] H.S. Penton, E. McKinney, 2012, "Geotypical Growth-based Load Forecasting: An introduction to an innovative approach", *IEEE PES (T&D)*, 1-8.

[5] M.M. Dalvand, S. Azami, H. Tarimoradi, 2008, "Long-term load forecasting of Iranian power grid using fuzzy and artificial neural networks", *43rd International Universities Power Engineering Conference*, 1-4.

[6] S. Carevic, T. Capuder, M. Delimar, 2010, "Applications of clustering algorithms in long-term load forecasting", *Proceedings of IEEE International Energy Conf. and Exhibition (EnergyCon)*, Vol. 1, 688-693.

[7] H. Daneshi, M. Shahidehpour, A.L. Choobbari, 2008, "Long-term load forecasting in electricity market ", *Proceedings of IEEE International Conf. on Electro/Information Tech.*, Vol. 1, 395-400.

[8] F. Kavehnia, H. Keivani, A. Mohammadi, 2007, "Long term demand forecast for an electricity distribution network", *Australasian Universities Power Engineering Conf. (AUPEC 2007.)*, 1-6.