A COMBINED PRACTICAL SHORT TERM SPATIAL LOAD FORECASTING METHOD IN DISTRIBUTION SYSTEM

Lifeng LIU, Zhenyu MA
Zhejiang Hangzhou Electric Power Bureau –P.R.China
lioulf@yahoo.com.cn

INTRODUCTION

Spatial Load Forecasting (SLF) is the basic of distribution system planning. Both trending method and Land-Use based method have the same forecasting precision in short term load forecasting. A new method to estimate the feeder’s load growth is proposed in this paper, which uses the grey theory, curve fit and Geographic Information System (GIS). The proposed method describes in detail. The real instance shows that the proposed method has the better result than other trending method. It shows that the results are quite reasonable and realizable for this study. It is fit for the limited computer resource especially, where it could replace the Land-Use based method.

REVIEW

Spatial Load Forecasting (SLF) is not only the basic of distribution system planning, but also the key research problem of Geographic Information System (GIS) spatial information analysis function application. Distribution system planners usually use both Land-use based method and trending method [1] to determine the load spatial distribution. Land-use based method requires considerable amount of customer, land-use, demographic data, etc that are difficult for the utilities to collect quickly [2]. Although the Trending method has the lower precision than Land-use method, the precision of Trending method is equivalent to that of Land-use for short term forecasting application [2]. Trending method uses the past annual peak loads to extrapolate and forecasting the future peak loads, which usually uses the multiple regression curve to fit the past annual peak loads of every feeder, and then extended into the future. This is the chief advantage that trending method over Land-use based method, which only need historical annual peak loads for each feeder or substation that could be gotten by utilities easily and quickly. If we could improve the precision of Trending method, then we could get one practical SLF method. So, we propose one practical short term SLF method in this paper, which combines the merit of Grey theory, curve fit and GIS.

NEW METHOD DESCRIPTION

The proposed method in this paper includes four steps. Fig.1. is the flow chart for the proposed algorithm.

Fig.1 Flow chart for the algorithm

1. Horizon year load calculation
   feeders are divided into six classes according to power-supply area and growth rate

2. Grey relational grade clustering according to different feeder’s growth rate

3. Curve fit for these six classes feeders

4. Forecasting extend history load to horizon year load per feeder

Feeder’S h Year Load Estimation

Because the feeder load develops non-evenly, usually getting saturation in some years quickly, which means S type development. Although the whole distribution develop result is continue and smooth, in fact it is the combination of different non-evenly development for different feeder [3]. So it is difficult to extrapolate with this S type development. The purpose of using h year load estimate method is to eliminate the different time extrapolate effect, not improve the precision of long range load forecasting.

Reference [4] proposes one new method to estimate the h year load per feeder, which use the feeder’s real power-supply area and its growth rate to class the feeders. At first,
feeder’s power-supply area is possible to estimate accurately with the help of distribution GIS. The writer divides the utility service territory into a grid of square small areas, and save it as grid layer in GIS. Secondly the writer overlaps the grid layer with the actual feeders layer. So the actual power-supply area of every feeder could be acquired easily. Then feeders are broken into three sets based on the actual power-supply area equally. There are Large set (L), Small set (S), and Medium set (M). The L is the largest one third of the feeders. The S is the smallest one third of the feeders. The M is all other feeders. Within each set – L, M, S, the feeders are further classified by how fast they have grown during the last four years. So L feeders are classified into large size fast (LF) growing and large size slow (LS) growth. Similarly, M feeders are classified into medium size fast (MF) growing and medium size slow (MS) growth, and S feeders are classified into small size fast (SF) growing and small size slow (SS) growing set. So feeders have been broken into six classes as Table 1. Based on the result of [4], we could estimate h year load for every class.

<table>
<thead>
<tr>
<th>Growth Rate</th>
<th>power-supply area</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>SF, MF, LF</td>
</tr>
<tr>
<td>Low &amp; Avg.</td>
<td>SS, MS, LS</td>
</tr>
</tbody>
</table>

The method and its step described in detail in [4].

**Grey Relational Grade Clustering Method for Feeder Load Growth Rate Class**

We could use S type development curve to denote the load growth of feeders, as Fig.2. Also we know that different feeder has different S type development curve. So we need use clustering method to determine different feeder’s load growth characteristic automatically. Clustering method divides data into some classes. Each class has the similar characteristic. There are many clustering method that use recently, e.g. Artificial Nerve Network method, fuzzy clustering method, K-mean value clustering method [5], etc. The widely used method in feeder’s load forecasting is K-mean value clustering method [5], which divides feeders into groups. Each group contains the feeders that have the similar growth characteristic, and would use the average growth curve of these feeders as the group template. Shown as Fig.3, the thin solid line denote load growth rate curve of each feeder for one group, while the thick solid line denotes the load template of this group. Then we could use this load template to forecasting the feeder’s load growth that has the similar growth characteristic.

According to least value of square error rule, we could use K-mean value method to acquire clustering result. The function is

\[ J = \sum_{i=1}^{k} \sum_{x \in R_i} \left\| x^j - w_i \right\|^2 \]  

where \(w_i\) is clustering center of \(R_i\), \(x^j\) is sample data.

K-mean value method implies to use sample data to denote the actual feeder growth curve. K-mean value method only uses the discrete point value to cluster, while not the actual curve to cluster. So K-mean method could not use the similarity of curve and the shape of curve fully. As we known, the feeders data are few and difficult to collect, which means that the data could be used has the grey characteristic. Also for data save manually, if we only use K-mean method without improvement, then we only have the roughness result. Grey relational grade clustering method fits for the situation that no matter few or enough data, the orderliness or chaos data [6], which could be used easily. Then we propose to use grey relational grade clustering method to deal with the different growth rate of different feeder.

Grey relational grade clustering method uses the curve similarity to judge the relational grade between curves. In fact, if some curves have the similar shape, then these curves have the similar development, and they have the biggest relational grade. Shown as Fig.4, the similarity grade of curve 1 and curve 2 is bigger than that of curve 1 and curve 3. Then the relational grade of curve 1 and curve 2 is bigger, while the relational grade of curve 1 and curve 3 is smaller. If \(x_0\) is the reference sequence, \(x_i\) is compared sequence, where \(i\) is from one, two, until \(m\). And,

\[ x_0 = \{x_0(1), x_0(2), \ldots, x_0(n)\} \]
\[ x_i = \{x_i(1), x_i(2), \ldots, x_i(n)\} \]

so

\[ \zeta_{ij}(k) = \frac{\min \{x_0(k) - x_i(k)\} + \rho \max \{x_0(k) - x_i(k)\}}{\max \{x_0(k) - x_i(k)\} + \rho \min \{x_0(k) - x_i(k)\}} \] (2)
is the relational coefficient between curve \( x_0 \) and curve \( x_i \) at \( k \) point, where \( p \) is distinguish coefficient, \( p \in [0,1] \), normally \( p = 0.5 \).

Similarly, then we could get the relational grade \( \gamma_i \) corresponding to the curve \( x_i \) and curve \( x_0 \) is

\[
\gamma_i = \frac{1}{n} \sum_{k=1}^{n} \zeta_i(k)
\]

if \( \Delta_{\text{max}} = \max_k \max_i |x_0(k) - x_i(k)| \)

\( \Delta_{\text{min}} = \min_k \min_i |x_0(k) - x_i(k)| \)

\( \Delta_i = |x_0(k) - x_i(k)| \)

then equ.(2) changes as

\[
\zeta_i(k) = \frac{\Delta_{\text{min}} + p\Delta_{\text{max}}}{\Delta_i + p\Delta_{\text{max}}}
\]

for sample data \( x_i, i \leq j, i, j = 1,2,\Lambda, m \), we could use equ.(3) to calculate the relational grade \( r_{ij} \) between \( x_i \) and \( x_j \). Then we get matrix

\[
A = \begin{bmatrix}
  r_{11} & r_{12} & \Lambda & r_{1m} \\
  r_{21} & \Lambda & r_{2m} \\
  \vdots & \vdots & \ddots & \vdots \\
  r_{m1} & \ldots & \ldots & \Lambda & r_{mm}
\end{bmatrix}
\]

where \( r_{ii} = 1, i = 1,2,\Lambda, m \).

Definition1: Matrix \( A \) defines as character vector relational matrix.

The critical value \( r \in [0,1] \), normally \( r > 0.5 \). Both \( x_i \) and \( x_j \) have the same characteristic if \( r_{ij} \geq r \) ( \( i \neq j \) ).

Definition2: Clustering class for critical value \( r \) defines as grey relational clustering of the character vector \( r \).

Then we could use grey relational clustering method to clustering the feeder’s load growth curve by adjust the critical value to acquire the different classes number.

**TEST ANALYSIS**

The new proposed method in this paper is also applied to one actual system. The system comprises 81 feeders operating at 110 and 10 kV, with a peak load (summer) of 271.5 MVA. The system includes 40 km², and 4 substations at 110kV. Load history of the system is from 1994 to 2000. Annual system load growth rate (1994 to 2000) averaged 14.3%.

**Test Method and Error**

The comparison test in this paper includes:

1. Basic method—cubic polynomial fit, use equ.(1) that propose in [4] and six years of load data.
2. Improved basic method—cubic polynomial fit, use equ.(1), equ.(6) that propose in [4], six years of load data and one \( h \) year load data.
3. Improved basic method—cubic polynomial fit, use equ.(1), equ.(7) that propose in [4], six years of load data and one \( h \) year load data.
4. \( GM(1,1) \) method [6].
5. K-mean value method and experienced formula [5].
6. The new method in this paper—combine merit of Grey theory, curve fit and GIS.

Average Absolute Error (AAE) is used to assess the error in this paper [7],

\[
AAE = \frac{1}{S} \sum_{n=1}^{S} \left| l_n^+(2000) - l_n^-(2000) \right|
\]

\[
\left| l_n^+(2000) - l_n^-(2000) \right|
\]

where \( l_n^+(t) \) is \( r \)th forecasted load for feeder \( n \), \( l_n(t) \) is \( r \)th load for feeder \( n \). AAE is average of mis-forecasting load growth rate per feeder.

**Test Results**

<table>
<thead>
<tr>
<th>Test system</th>
<th>Method</th>
<th>AAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>23.99%</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>19.95%</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>12.63%</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>11.41%</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>9.33%</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>9.01%</td>
<td></td>
</tr>
</tbody>
</table>
TABLE 2 is the load forecast accuracy of six forecast methods in 2000. The method ⑥ that proposed in this paper has the best result than other trending methods. For the test system belongs to an area that its growth rate co-exists fast and slow speed, the different load growth could be distinguished correctly by suing grey rational grade.

Test Results Analysis

Although there is only one test, there also have some useful test results.

① For we only collect the test data from 1994 to 2000, and the data of 2000 is the tested data, which means that we only have six years history data. Within test ① and ②, and reference [4], the result of using recent six year data is better than use few year data when we use one h year load.

② In step 2, different critical value $r$ could get different clustering classes number. The numbers are range from one to nine. The best result could be gotten when we adopt clustering classes five.

CONCLUSIONS

SLF is not only the basic of distribution system planning, which would affect the planning precision, but also the key research problem of GIS spatial information analysis function application. A new method to estimate the feeder’s load growth is proposed in this paper, which uses the grey theory, curve fit and GIS.

1 The feeders load growth curves are analyzed and clustered by using grey relational grade clustering method, instead of using K-mean value method, which depends on the integrality and consistent of data heavily. K-mean value method only uses the discrete point value to cluster, while grey rational grade uses the curve to cluster, which would be more reasonable, for we need use the whole trend of the curve to extrapolate.

2 The real instance shows that the proposed method has the better result than other traditional trending method. It shows that the results are quite reasonable and realizable for this study. It is fit for the limited computer resource especially, where it could replace the Land-Use based method.

REFERENCES


Liu Lifeng was born in Shannxi, China, on Aug., 28, 1973. He received B.S. (1995), M.S. (1998), and Ph.D degrees (2001) in Electrical Engineering from Chongqing University, China. From April 2002, he was with Hangzhou Electric Power Bureau, Zhejiang, China. His special fields of interest included DA, DMS, and GIS.

Ma Zhenyu was born in Hangzhou, China, on Apr., 8, 1972. He received associate degrees ( 1992 ) in automation of electric power systems from Nanjing electric power junior college, and entered the raining course of master degree ( 2000 ) in computer science from Zhejiang University, China. From Jul., 1992, he was with Hangzhou Electric Power Bureau, Zhejiang, China. His special fields of interest included IT, DMS, and SCADA.