

配电系统不良数据修正方法的研究

李慧

(北京机械工业学院计算机与自动化系,北京 100085)

摘要: 对 GM(1,1)基本模型中第一个数据不起作用的结论作了全新的简洁的推导,提出了计及第一个数的零加数 GM(1,1)模型,提高了原始数据的利用率。在此基础上,结合电力负荷呈连续性和日周期性变化的特性,提出了用于负荷数据修正的零加数 GM(1,1)组合预测方法。该方法通过从两个角度选取原始序列进行零加数建模,采用关联度的分析方法,将预测值进行线性组合。实验证明,零加数模型的预测精度优于原始模型的预测精度,且组合预测比单一序列预测在计算精度上有明显的提高。

关键词: 零加数 GM(1,1)模型; 关联度; 组合预测; 配电系统

中图分类号: TM72 文献标识码: A 文章编号: 1003-4897(2006)13-0042-04

0 引言

配电系统运行工况比较复杂,由测量装置所采集的负荷数据中不良数据与实际正常的突变数据混杂现象尤为突出。当所测量的负荷功率数据(包括有功和无功)作为可疑不良数据被剔除后^[1],还必须对其进行补充或修正,以保证系统的可观察性^[2]。本文借助配电网短期负荷预测的方法对可疑不良数据进行修正,即用预报数据替换可疑不良数据。

电力负荷系统是典型的灰色系统,因而灰色预测技术在电力系统有着广泛的应用^[3~5]。GM(1,1)模型是灰色理论中用于预测的最广泛的模型,它的算法建模过程简单,模型表达式简洁,求解速度快。但GM(1,1)模型实质是对除第一点外的原始数据序列作指数曲线拟合,而原始数据序列中的第一个数据不能起作用^[6],造成数据资源的浪费。本文对此结论作出全新的简洁的推导,并提出了计及第一个数据的零加数 GM(1,1)预测模型。该模型不仅能使 GM(1,1)模型建模需要的最小数据量由 4 个减少为 3 个,拓宽了 GM(1,1)模型的应用范围,而且提高了短数据序列 GM(1,1)模型的精度。同时,本文结合电力负荷随日呈连续性和日周期性变化的特性,从不同角度选取原始序列进行零加数 GM(1,1)建模,充分利用不同有用的信息,采用关联度分析方法将这些有用信息进行组合,得到用于负荷数据修正的计算模型。实验证明,该方法能提高短数据序列的预测精度,而且预测精度满足实际应用的要求。

1 零加数 GM(1,1)模型

1.1 传统 GM(1,1)模型^[3~5]

设原始序列为

$$X^{(0)}(k) = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\},$$

对 $X^{(0)}(k)$ 作一次累加生成(1-AGO)得生成数列

$$X^{(1)}(k) = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\},$$

其中: $x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i), k = 1, 2, \dots, n_0$

$X^{(1)}(k)$ 的紧邻均值生成序列

$$Z^{(1)} = \{z^{(1)}(1), z^{(1)}(2), \dots, z^{(1)}(n)\},$$

其中: $z^{(1)}(k) = [x^{(1)}(k) + x^{(1)}(k-1)]/2,$

$$k = 2, 3, \dots, n$$

称一阶线性微分方程

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = u, \quad t \in [0, \infty) \quad (1)$$

为灰色微分方程(即灰色 GM(1,1)模型)

$$x^{(0)}(k) + az^{(1)}(k) = u \quad (2)$$

的白化方程。

对式(1)的离散响应式作一次累减得到原始序列的预测模型为

$$\begin{cases} \hat{x}^{(0)}(1) = x^{(0)}(1) \\ \hat{x}^{(0)}(k) = (1 - e^{-a}) [x^{(0)}(1) - \frac{u}{a}] e^{-a(k-1)} \\ k = 2, \dots, n \end{cases} \quad (3)$$

式(3)中的系数 a 与常数项 u 是由式(2)表示的方程组采用最小二乘法作参数估计获得,即:

$$\hat{a} = \begin{bmatrix} \hat{a} \\ \hat{u} \end{bmatrix} = (B^T B)^{-1} B^T Y \quad (4)$$

其中:

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \dots \\ x^{(0)}(n) \end{bmatrix}, \quad B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \dots & \dots \\ -z^{(1)}(n) & 1 \end{bmatrix}$$

由此可知,GM(1,1)模型建模中参数估计所需要的最少数据量应为4个。

1.2 零加数 GM(1,1)模型

文献[6]证明了GM(1,1)基本模型没有利用第一个数据进行建模的结论,下面给出简洁的推导,并提出了计及第一个数的零加数GM(1,1)模型。

1.2.1 推导过程

考虑新的非负序列

$$X_c^{(0)} = \{x^{(0)}(1) + c, x^{(0)}(2), \dots, x^{(0)}(n)\}$$

利用式(4)得到参数 a_c 和 u_c 的计算公式,即

$$T_c = [a_c, u_c]^T = (B_c^T B_c)^{-1} B_c^T Y \quad (5)$$

式中:

$$Y = [x^{(0)}(2) \ x^{(0)}(3) \ \dots \ x^{(0)}(n)] J_{(n-1)}^{-1} x_1$$

$$B_c = \begin{bmatrix} -\frac{1}{2} [x^{(1)}(1) + x^{(1)}(2)] - c & 1 \\ -\frac{1}{2} [x^{(1)}(2) + x^{(1)}(3)] - c & 1 \\ \dots & \dots \\ -\frac{1}{2} [x^{(1)}(n-1) + x^{(1)}(n)] - c & 1 \end{bmatrix}_{(n-1) \times 2}$$

$$B = \begin{bmatrix} 1 & 0 \\ -c & 1 \end{bmatrix}^T$$

将 B_c 的表达式代入到式(5)中,有

$$T_c = \left[\begin{array}{c} 1 \\ -c \end{array} \right] \begin{array}{c} 0 \\ 1 \end{array}^T (B^T B)^{-1} \begin{bmatrix} 1 & 0 \\ -c & 1 \end{bmatrix} \begin{array}{c} 1 \\ 0 \end{array}^T \begin{array}{c} 1 \\ -c \end{array} \right] B^T Y =$$

$$\left[\begin{array}{c} 1 \\ -c \end{array} \right] \begin{array}{c} 0 \\ 1 \end{array}^{-1} (B^T B)^{-1} \begin{bmatrix} 1 & 0 \\ -c & 1 \end{bmatrix} \begin{array}{c} 1 \\ 0 \end{array}^T \begin{array}{c} 1 \\ -c \end{array} \right] B^T Y =$$

$$B^T Y = \left[\begin{array}{c} 1 \\ -c \end{array} \right] \begin{array}{c} 0 \\ 1 \end{array}^{-1} (B^T B)^{-1} B^T Y = \left[\begin{array}{c} 1 \\ -c \end{array} \right] \begin{array}{c} 0 \\ 1 \end{array}^{-1} \begin{bmatrix} a \\ ac + u \end{bmatrix} =$$

此时,新序列的预测模型为

$$\begin{cases} \hat{x}_c^{(0)}(1) = x^{(0)}(1) + c \\ \hat{x}_c^{(0)}(k) = [x^{(0)}(1) - \frac{u}{a}] e^{-a(k-1)} (1 - e^a) \\ k=2, \dots, n \end{cases} \quad (6)$$

比较式(3)和(6)可知,拟合得到的序列 $\hat{x}^{(0)}$ 和 $\hat{x}_c^{(0)}$ 除第一项外,其它项都是相等的,这说明原始序列中的第一个数 $x^{(0)}(1)$ 的变化并不引起模型预测值 $\hat{x}^{(0)}(k)$ 的改变,即GM(1,1)基本模型实质是对

除第一项外的原始数据序列作指数曲线拟合,原始数据序列中的第一个数据未利用。

假如我们把原始序列的第一项 $x^{(0)}(1)$ 改为0,然后对 n 维序列 $\{0, x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)\}$ 建模,此时,数据处理、参数辨识、时间响应都变得简单。这样, n 维建模序列所需的观测数据由 n 个减少为 $(n-1)$ 个,从而使建模所需最小观测数据由4个减为3个,拓宽了GM(1,1)模型的应用范围。

1.2.2 零加数 GM(1,1)模型

为了充分利用已有信息,本文提出零加数GM(1,1)的模型,即在原始数据序列前加一个零,形成新数据序列,再建立GM(1,1)模型。

设原始非负序列 $X^{(0)}$ 前添加一个数 $A=0$,生成新序列为

$$X_1^{(0)} = \{0, x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$$

对这个新序列进行GM(1,1)基本建模,可得到生成数据模型为

$$\begin{cases} \hat{x}_1^{(0)}(0) = 0 \\ \hat{x}_1^{(0)}(k) = [0 - \frac{u_1}{a_1}] e^{-a_1(k-1)} (1 - e^{a_1}) \\ k=2, \dots, n \end{cases} \quad (7)$$

其中: a_1 和 u_1 的求解方法类似公式(5)。

关于原始序列长度 n 的取值,文献[7]给出了分析结论:实际应用中,考虑到在满足预测精度的前提下,尽可能少地利用数据资源,取 $n=4$ 为最佳。所以,本文算例中原始序列长度为 $n=4$ 。

2 基于零加数 GM(1,1)的组合预测方法

电力负荷呈连续性和日周期性的变化特性,从不同角度选取原始序列进行零加数GM(1,1)建模,可以充分利用不同有用的信息。本文采用关联度的分析方法,将这些有用的信息进行组合,得到用于负荷数据修正的计算模型。

2.1 原始序列的选取

就所研究的电力负荷数据而言,具有明显的连续性和日周期性,某时刻的负荷受其邻近的负荷影响最大,连续几天内同一时刻的负荷大小相近。为此,本文针对一般和特殊两种情况下的数据序列进行分析。

一般情况:若待预测数据的前 n 个时刻的负荷数据齐全,而后 n 个时刻的负荷数据不足于3个(缺数据),这种情况比较常见。此时,选取下面两组不同的时间序列。序列一是由邻近的前 n 个时刻的负荷数据构成,记为 $\{x_1(k) | k=1, 2, \dots, n\}$, $\hat{x}_1(n+1)$

为预测数据。序列二是由邻近的前 n 天内同一时刻的负荷数据构成, 记为 $\{x_2(k) | k=1, 2, \dots, n\}$, $\hat{x}_2(n+1)$ 为预测数据。

特殊情况: 若待预测数据的前后时刻存在连续的 n 个数据 ($n \geq 3$), 这种情况为特例。此时, 选取下面两组不同的时间序列。序列一同样是由邻近的前 n 个时刻的负荷数据构成, 记为 $\{x_1(k) | k=1, 2, \dots, n\}$, $\hat{x}_1(n+1)$ 为预测数据。序列二是由邻近的后 n 个时刻的负荷数据构成, 记为 $\{x_2(k) | k=1, 2, \dots, n\}$, $\hat{x}_2(n+1)$ 为预测数据。

无论是序列 $x_1(k)$ 还是序列 $x_2(k)$, 其中元素的排列规则满足: $x_i(1)$ 代表距离待预测数据最远的负荷数据, $x_i(n)$ 代表距离待预测数据最近的负荷数据, 以此类推。

2.2 基于关联度分析的组合过程

关联度是灰色理论处理数据的方法之一, 是分析系统中各因素关联程度的一种方法^[8]。关联度的基本思想, 是根据曲线间相似程度来判断关联程度, 实质是几种曲线间几何形状的分析比较, 即认为几何形状越接近, 则关联度越大, 同时说明对应的预测模型越优, 拟合误差越小。

记 $x_i = \{x_i(1), x_i(2), \dots, x_i(n)\}$ 为实际序列, $\hat{x}_i = \{\hat{x}_i(1), \hat{x}_i(2), \dots, \hat{x}_i(n)\}$ 为预测序列, $r_i(k)$ 为曲线 x_i 与 \hat{x}_i 在 k 点的关联系数, ρ_i 为预测曲线 \hat{x}_i 与实际曲线 x_i 的关联度, 则有

$$r_i(k) = \frac{\min_i \min_t |x_i(k) - \hat{x}_i(k)| + 0.5 \max_i \max_t |x_i(k) - \hat{x}_i(k)|}{|x_i(k) - \hat{x}_i(k)| + 0.5 \max_i \max_t |x_i(k) - \hat{x}_i(k)|} \quad (8)$$

$i=1, 2; \quad k=1, 2, \dots, n$

$$\rho_i = \frac{1}{n} \sum_{k=1}^n r_i(k) \quad (9)$$

利用式 (8) 和 (9), 分别计算序列 \hat{x}_1 和 \hat{x}_2 各自的关联度 ρ_1 和 ρ_2 。然后, 由关联度得到两个序列的预测值在组合中的权重。设 ω_1 和 ω_2 分别为 \hat{x}_1 和 \hat{x}_2 在组合方法中的权重, 则有

$$\omega_1 = \frac{1}{\rho_1 + \rho_2}, \quad \omega_2 = \frac{2}{\rho_1 + \rho_2} \quad (10)$$

利用权重对各序列得到的预测值进行线性组合, 得到最终预测值为

$$\hat{x}(n+1) = \sum_{i=1}^2 \omega_i \hat{x}_i(n+1) \quad (11)$$

2.3 建模步骤

步骤一: 采用文献 [1] 中提出的不良数据辨识方法判定出不良负荷数据的位置。

步骤二: 判断需修正的 (不良) 负荷数据属于哪一种预测数据类型, 从而生成对应的原始时间序列一和序列二。

步骤三: 利用式 (7), 即零加数 GM (1, 1) 建模法, 对序列一和序列二分别进行预测, 得到各自的预测值 $\hat{x}_i(k)$, $i=1, 2; k=1, 2, \dots, n, n+1$ 。

步骤四: 利用公式 (8)、(9), 分别计算两个序列下的关联度。

步骤五: 利用公式 (10) 计算两个时间序列在组合中的权重。

步骤六: 利用式 (11) 得到最终的修正值。

3 算例分析

本文用所提出的基于零加数 GM (1, 1) 组合预测方法编制了用于不良数据修正的程序, 并用某变压器 2004 年 12 月 6 日的有功负荷数据进行测试来验证本文所提出的修正方法的优越性。已知该配变在当天有完整的 24 点负荷记录 (见表 1)。为了突出本文方法的优越性, 算例分析中省略了步骤一的不良数据辨识过程, 直接假设不良数据的位置已知, 主要任务进行修正方法的对比测试。

表 1 2004 年 12 月 6 日整点时刻的负荷功率

Tab 1 Load power on December 6, 2004

时间	有功 /kW	时间	有功 /kW
0:00	72.0073	12:00	73.8680
1:00	68.4721	13:00	67.5418
2:00	65.4950	14:00	64.7508
3:00	69.7745	15:00	71.4491
4:00	57.1221	16:00	65.4950
5:00	61.7737	17:00	79.2639
6:00	65.8672	18:00	76.8450
7:00	64.1926	19:00	79.8221
8:00	65.1229	20:00	80.5663
9:00	80.5663	21:00	81.4967
10:00	82.7991	22:00	70.7049
11:00	87.2647	23:00	64.5647

3.1 零加数模型与基本模型在预测精度上的对比测试

利用测试系统 1 来验证零加数模型的预测精度优于基本模型的预测精度。

测试系统 1: 假设一天中存在两个不良数据, 分别发生在 4 点和 6 点、7 点和 8 点、10 点和 11 点、21 点和 23 点四种情况。

利用两种组合预测方法对测试系统 1 中发生在 4 点、7 点、10 点和 21 点的四个不良数据分别进行修正。由于分别发生在上述时刻的不良数据的后 n

个负荷数据不足于 3 个,所以需选取一般情况下的数据序列一和序列二进行建模,修正结果列于表 2。

表 2 两种预测方法对测试系统 1 中不良数据的修正结果比较

Tab 2 Revised results comparison of two methods with bad data in the 1st test system

时间	真值	修正值		相对误差	
		原始模型	零加数模型	原始模型	零加数模型
4:00	57.12	65.07	63.44	-13.91%	-11.06%
7:00	64.19	58.72	58.37	8.53%	9.07%
10:00	82.80	75.03	74.05	9.39%	10.57%
21:00	81.50	75.39	78.93	7.49%	3.15%
平均				9.83%	8.46%

从表 2 看出,原始模型中预测的最大和平均的相对误差分别为 13.91%和 9.83%,而零加数模型中预测的最大和平均的相对误差分别为 11.06%和 8.46%。这也说明零加数模型比基本模型的预测精度高。

3.2 组合预测与单序列预测在预测精度上的对比测试

利用测试系统 2 来验证组合预测精度优于单序列预测精度。

测试系统 2:假设一天中只有一个不良数据,分别发生在 4 点、7 点、10 点、21 点四种情况。

由于分别发生在上述时刻的不良数据前后存在连续的 n 个数据,所以需选取特殊情况下的数据序列一和序列二进行建模。针对测试系统 2 中四种情况下的不良数据,采用三种不同模式进行预测。表 3 列出了三种模式下的修正结果比较。

模式一:对序列一进行零加数 GM(1,1)建模;

模式二:对序列二进行零加数 GM(1,1)建模;

模式三:利用关联度分析方法对模式一和模式二的预测结果进行组合。

表 3 三种模式下的修正结果比较

Tab 3 Revised results comparison in three modes

时间	真值	修正值			相对误差		
		模式一	模式二	模式三	模式一	模式二	模式三
4:00	57.12	68.93	56.57	62.69	-20.7%	0.96%	-9.75%
7:00	64.19	61.66	76.72	68.53	3.94%	-19.52%	-6.76%
10:00	82.80	76.06	80.27	78.09	8.14%	3.05%	5.69%
21:00	81.50	86.92	73.97	79.93	-6.65%	9.24%	1.93%
平均					9.86%	8.19%	6.03%

从表 3 看出,三种模式下预测的最大和平均的相对误差分别为 20.7%和 9.86%、19.52%和 8.19%、9.75%和 6.03%。由此得出,模式三(即组合预测模型)是最佳方案。

另外,比较表 2 中的零加数模型和表 3 中的模

式三所预测的最大和平均的相对误差,可以看出,选取同一天前后时刻的负荷数据(即模式三)进行预测,其预测效果最佳。也就是说,越靠近待预测时刻的负荷数据对预测精度的影响越大。

4 结论

1) 对小样本数据序列建模的情况下,采用加数模型拓宽了 GM(1,1)模型的应用范围,提高了数据利用率和预测精度。

2) 采用单一的灰色模型,预测精度不高;通过选取不同的时间序列,采用组合灰色模型的方法,预测精度有显著提高。

3) 预测时刻的负荷大小受临近负荷的影响最大,尽可能地选取同一天前后时刻的负荷数据对其进行预测,效果最佳。

参考文献:

- [1] 李慧,杨明皓. 小波分析在电力系统不良数据辨识中的应用[J]. 继电器, 2005, 33(3): 10-14, 20
LI Hui, YANG Ming-hao. Application of Wavelet Analysis to Bad Data Identification for Power System [J]. Relay, 2005, 33(3): 10-14, 20
- [2] 诸骏伟. 电力系统分析[M]. 北京:中国电力出版社, 1995.
CHU Jun-wei. Power System Analysis [M]. Beijing: China Electric Power Press, 1995.
- [3] 高燕. 灰色理论在配电网短期负荷预测中的应用(硕士学位论文)[D]. 北京:中国农业大学, 2004.
GAO Yan. Application of Gray Theory on Short Load Forecasting in Distribution Network, Thesis [D]. Beijing: China Agricultural University, 2004.
- [4] 曹国剑,黄纯,隆辉. 基于 GM(1,1)改进模型的电网负荷预测方法[J]. 电网技术, 2004, 28(13): 50-53.
CAO Guo-jian, HUANG Chun, LONG Hui. Load Forecasting Based on Improved GM(1,1) Model [J]. Power System Technology, 2004, 28(13): 50-53.
- [5] 牛东晓,曹树华,赵磊,等. 电力负荷预测技术及其应用[M]. 北京:中国电力出版社, 1998
NIU Dong-xiao, CAO Shu-hua, ZHAO Lei, et al. Power Load Forecasting Technology and Application [M]. Beijing: China Electric Power Press, 1998
- [6] 李留藏,许闻天,蔡相展. 灰色系统 GM(1,1)模型的讨论[J]. 数学的实践与认识, 1993, (1): 15-22
LI Liu-cang, XU Wen-tian, CAI Xiang-zhan. Discussion of Gray System GM(1,1) Model [J]. Mathematics in Practice and Theory, 1993, (1): 15-22

(下转第 57 页 continued on page 57)

及研究 [J]. 继电器, 2005, 33 (8): 68-70

NI Wei-dong, LI Rui-sheng, LI Zheng-feng Experimentation and Research on Optical Fiber Channel in Current Differential Protection System [J]. Relay, 2005, 33 (8): 68-70

- [3] 郑玉平, 金华锋. 线路纵联保护采用光纤通信一些问题的探讨 [A]. 2005第十届全国保护和控制学术研讨会论文集. 2005. 13-18

ZHENG Yu-ping, JIN Hua-feng Some Issues of Line Pilot Relays Using Fiber Communication Network [A]. Proceedings of the 10th National Academic Conference for Electric Power System Protection and Control Nanjing: 2005. 13-18

2005. 13-18

- [4] G 703, Telecommunication Standardization Sector of ITU [S].

收稿日期: 2006-01-11; 修回日期: 2006-03-02

作者简介:

尹成群 (1953 -), 男, 教授、硕士生导师, 主要研究方向为应用电子和信号与信息处理;

杨贵 (1976 -), 男, 硕士研究生, 主要研究方向为信号与信息处理。E-mail: guiyu@sina.com

Tests and studies of protection optical fiber channel simulation

YN Cheng-qun, YANG Gui

(School of Electric and Communicational Engineering, North China Electric Power University, Baoding 071003, China)

Abstract: New problems emerges with the wide application of optical fiber current differential protection in the power system. Without the simulation tests on various bit-error conditions of optical fiber channel, the problems cannot be tested and verified including the channel time delay, operation behavior and performance on the bit-error condition in optical fiber current differential protection. This paper designs a simulation and test instrument based on the real application of the protection and introduces its functions, hardware design and application.

Key words: optical fiber current differential protection; time delay; bit error

(上接第 45 页 continued from page 45)

- [7] 李慧. 配电系统负荷数据处理方法的研究 (博士学位论文) [D]. 北京: 中国农业大学, 2005.

LI Hui Research on Load Data Processing Method in Distribution Systems, Doctoral Dissertation [D]. Beijing: China Agricultural University, 2005.

- [8] 肖国泉, 王春, 张福伟. 电力负荷预测 [M]. 北京: 中国电力出版社, 2001.

XIAO Guo-quan, WANG Chun, ZHANG Fu-wei Power

Load Forecasting [M]. Beijing: China Electric Power Press, 2001.

收稿日期: 2005-12-05; 修回日期: 2006-01-20

作者简介:

李慧 (1976-), 女, 博士, 讲师, 主要研究方向为配电系统状态估计、负荷预测算法研究, 小波分析理论及其应用。

E-mail: lhbxy21@sina.com

On bad data correcting for distribution systems

LI Hui

(Department of Computer and Automation, Beijing Mechanical Industry College, Beijing 100085, China)

Abstract: The paper proves compactly the conclusion that the first datum of original data lists doesn't work in the course of building GM (1, 1) model. Further, a zero addend GM (1, 1) model that takes into account the first datum is developed. It improves the utilization ratio of raw data. Based on the means, combining with the characteristic of continuity and day periodicity changing of power loads, a zero addend GM (1, 1) combined forecasting approach is developed to correct load data of distribution systems. This method selects two original data lists to forecast two values, and combines them to obtain final result by relative-degree analysis approach. Test results indicate that the precision of zero addend model is superior to of primitive model, and the precision of combined forecasting is higher than that of single list forecasting.

Key words: zero addend GM (1, 1) model; relational grade; combined forecasting; distribution system