

一种改进的相似日负荷预测方法

负荷预测对电力系统的运行的可靠性和经济性意义重大。很多国内外学者都对负荷预测理论做过大量研究，提出了许多方法。提出一种基于误差预测调整的改进相似日预测方法，即通过对相似日预测产生的误差进行预测，调整原预测模型预测负荷，提高预测准确度。最后通过算例验证了该方法的科学性和实用性。

张春雷 李炎/国电南瑞科技股份有限公司

0 引言

在短期负荷预测方面，许多学者对负荷预测进行过深入研究，包括线性回归法、时间序列法、卡尔曼滤波法、人工神经网络法和相似日预测等^[1]。但由于大多数模型只考虑了影响预测结果的主要因素，而忽略了次要因素，造成了误差^[2]。本文提出的误差预测调整就是利用一定的预测方法，对误差进行预测，利用误差调整初始预测结果，调整后的最终结果就考虑了这些被忽略的因素，因而预测准确度会得到一定程度的提高。误差调整包含了预测算法对于当前样本未考虑到的次要因素，是对当前预测方法的补充。



张春雷/工程师

关键词/Keywords

负荷预测 ·
相似日 ·
误差调整 ·

1 数据预处理

系统负荷预测的历史数据来自于量测值，但由于量测方法、传输通道等原因，获取到的量测数据中有不少异常数据，需要进行基础的坏数据剔除及修复，否则会对预测结果造成不良影响。

数据横向比较法利用了负荷曲线中连续两个点不应该发生突变的性质来判断坏数据。如果连续两个点负荷变化超出阈值，则认为是坏数据。通过这样的

方法可以发现历史数据中的突变点，对其进行必要的修复。

1.1 主要原理

1) 选取与待处理日的日类型相同的历史负荷作为样本数据，将样本数据中相邻点负荷的历史最大变化率作为阈值。

2) 对于预测日的每个负荷点，只与当前时刻的上一个点进行比较，按时间顺序进行，一旦检测到坏数据就进行修复，并将其作为下一次比较的基础。

3) 采用同一日类型、同一时刻的历史负荷的线性组合作为调整值。

$$\begin{cases} \left| \frac{L(d,t)-L(d,t-1)}{L(d,t-1)} \right| > \alpha_{\max}(t,n), t \neq 0 \\ \left| \frac{L(d,t)-L(d-1,\infty)}{L(d-1,\infty)} \right| > \alpha_{\max}(t,n), t=0 \\ \alpha_{\max}(t,n) = \max \left\{ \frac{L(d-i,t)-L(d-i,t-1)}{L(d-i,t-1)} \right\} \end{cases} \quad (1)$$

其中， $i=1-n$ ，

$$L(d,t) = \lambda_1 L(d-1,t) + \lambda_2 L(d-2,t) + \dots + \lambda_m L(d-m,t) \quad (2)$$

$$\begin{cases} \lambda_j = \beta(1-\beta)^{j-1}, \beta \in (0,1) \\ \sum_{j=1}^m \lambda_j = 1 \end{cases} \quad (3)$$

式中， $L(d,t)$ 为第 d 天 t 时刻的负荷点； $L(d-1,\infty)$ 为第 $d-1$ 天的最后 1 个负荷点；阈值 $\alpha_{\max}(t,n)$ 为前 n 天内 t 时刻与 $t-1$ 时刻负荷变化率的最大值； i 为与当前日期的间隔天数； $L(d-m,t)$ 为第 $d-m$ 天 t 时刻的负荷； λ_m 为权值系数，用于表征第 $d-m$ 天 t 时刻负荷对第 d 天 t 时刻

负荷影响程度； β 为平滑系数。当 $\lambda_1 = \lambda_2 = \dots = \lambda_m$ 时，相当于取同一日期类型下同一时刻历史数据的平均值。

1.2 方法优点

采用该横向数据比较法进行异常数据调整具有以下优势。

1) 可以有效避免基准值为异常数据的情况，通过即时的调整，能够识别出连续的数据丢失和连续的数据异常。

2) 同一日期类型下同一时刻的历史负荷一般波动不大，采用这种线性组合方式调整异常数据，使调整结果更加可信。

3) 采用相邻点的负荷变化率作为评判异常数据的标准，在一定程度上避免了由基准值本身的误差引起的误判。

4) 选取与待处理日日期类型相同的历史负荷作为参考数据，可以避免由参考数据本身的特性差异引起的误判。

2 相似日的选取

相似日选取的是否准确会直接影响到负荷预测的准确性。由于采用了误差调整的预测方法，其中的误差预测同样基于相似日算法，因此相似日选择是否合适同时也会直接影响到误差预测的正确性，进一步影响最终的预测结果。

2.1 时间因子匹配系数

负荷日的时间距离和类型是影响相似日选取的重要因素。其中，时间距离指的是历史日与预测日的相差天数。显然，历史日与预测日相差的天数越多，二者的相似度越小。因此，选取相似日有着“近大远小”的规律。

日类型主要是指工作日或周末。一般，工作日的负荷水平明显高于周末，而且负荷的变化趋势也有明显区别。除了工作日和周末的区别，工作日的每一天的日类型也是不一样的。比如星期一用1表示，星期二用2表示，其他类推。如果历史日和预测日的日类型相同，则该因素的相似度最大，设置为1；如果同样是工作日或周末，但日类型不同，给出一个较大的相似度，比如0.7等。

时间因子匹配系数需要能够量化历史日与预测日在日类型和时间距离上的相似程度。根据以

上原因，采用如下的函数表达式来计算第 t 个历史日与预测日的时间匹配系数 $\delta(1)$ ，即

$$\delta(t) = S\beta^t \quad (4)$$

式中，当历史日和预测日的日类型相同时， $S=1$ ；当历史日和预测日同为工作日或周末但日类型不同时， $S=0.7$ 。 β 为相似度衰减比率，一般取 $0.90 \sim 0.98$ 。

2.2 气象因子匹配系数

气象因素在相似日选取中也占有重要的位置。一般来说，夏季气温越高负荷越大，冬季气温越低负荷越大，历史日与待预测日的气象条件越相似，则负荷水平也越接近。这里可以采用气象因子的匹配系数来量化历史日气象数据与预测日气象数据的相似度。将气象因素特征向量化，再用灰色关联分析法来计算气象因子匹配系数。

设 X_0 、 X_i ($i=1, \dots, 1$) 分别为预测日及历史日的气象因素特征向量；预测日的气象因素特征向量 $X_0 = (X_0(1), \dots, X_0(m))$ ，其中 m 为气象因素的个数；第 i 个历史日的气象因素特征向量 $X_i = [X_i(1), \dots, X_i(m)]$ ；预测日与第 i 个历史日在第 k 个特征的差异 $\Delta_i(k)$ 为

$$\Delta_i(k) = |X_i(k) - X_0(k)| \quad (5)$$

将差异归一化后的值为

$$\Delta'_i(k) = \frac{\Delta_i(k) - \min_i \Delta_i(k)}{\max_i \Delta_i(k) - \min_i \Delta_i(k)} \quad (6)$$

则预测日与第 i 个历史日在第 k 个特征的关联系数为

$$\zeta_i(k) = \frac{\min_i \min_k \Delta'_i(k) + \rho \max_i \max_k \Delta'_i(k)}{\Delta'_i(k) + \rho \max_i \max_k \Delta'_i(k)} \quad (7)$$

式中， $\min \Delta_i(k)$ 为第一级最小差，表示在第 i 个历史日中找与预测日差异最小的特征值； $\min \min \Delta_i(k)$ 为第二级最小差，表示所有历史日中差异最小的特征值； $\max \max \Delta_i(k)$ 为第二级最大差，表示所有历史日中差异最大的特征值； ρ 是一个介于0与1之间的常数，一般取 $\rho = 0.5$ 。综合各个特征的关联系数，得到预测日与第 i 个历史日的关联度 r_i 为

$$r_i = \frac{1}{m} \sum_{k=1}^m \zeta_i(k) \quad (8)$$

由关联系数计算公式可知，气象因素差异越小，关联系数越大，历史日与预测日的关联度也越大。取关联度的值作为气象因子匹配系数。

3 预测流程

在电力系统中，通常将系统负荷定位为时间序列。将时间域表示为

$$t \in D^- + D^+ \tag{9}$$

式中， D^- 为历史时域，该时域负荷设为 $\{Y(t)\} (t \in D^-)$ ，它是建模的依据； D^+ 表示预测时域，目标就是要预测出这个时域的值，设为 $\{Y(t)\} (t \in D^+)$ 。

建模及预测流程如图1所示。

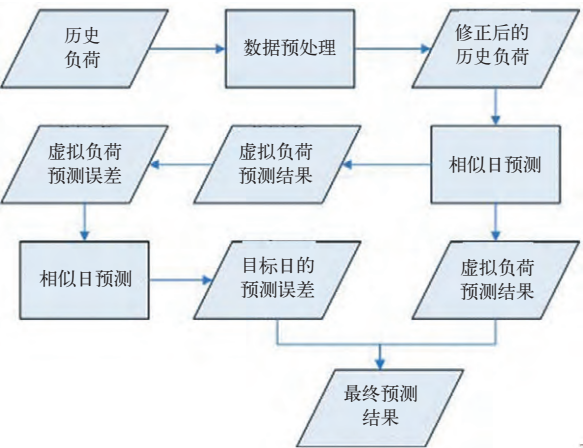


图1 预测流程

4 算例

采用该方法对某省2013年8月2日~15日的系统负荷进行预测结果如下表所示。

表 预测结果对比

日期	相似日预测准确度 (%)	考虑误差调整的相似日预测准确度 (%)
2013-08-02	96.65	97.89
2013-08-03	96.21	98.02
2013-08-04	96.36	97.45
2013-08-05	95.48	96.56
2013-08-06	97.01	97.21
2013-08-07	96.23	97.83
2013-08-08	96.69	97.79
2013-08-09	97.23	98.19
2013-08-10	96.97	98.34
2013-08-11	97.23	98.82
2013-08-12	96.13	97.47
2013-08-13	97.43	97.86
2013-08-14	95.87	97.54
2013-08-15	96.64	97.67

根据上表可以看出，在相似日预测方法的基础上考虑误差调整可以有效提高负荷预测的准确率，平均提高了1.18%

5 结束语

提出了利用预测误差对相似日预测结果进行调整的方法，阐明了这种方法的基本思路和作用的原理。利用数据横向比较法修正异常的历史样本，避免由于样本原因导致计算出来的误差与实际偏差过大，影响误差的调整效果。根据日类型、时间距离及气象因素选择合适的相似日。利用算例进行对比分析，验证了该方法对提高预测准确度的有效性。当前对于调整结果时所用误差的生成方式处理比较简单，后续在这一方面还可以做进一步的研究。

参考文献

[1] 牛东晓. 电力负荷预测技术及其应用[M]. 北京:中国电力出版社, 1998.

[2] 周明, 严正, 夏清, 等. 含误差预测校正的ARIMA电价预测新方法[J]. 中国电机工程学报, 2004, 24(12): 63-68.

[3] 张艳杰, 刘耀年, 祝滨, 等. 基于模糊聚类识别及统计相关的短期负荷预测[J]. 电网技术, 2002, 26(11): 32-36.

[4] 程松. 一种新的基于相似负荷日的短期负荷预计方法[J]. 江苏电机工程, 1999, 18(4): 28-22.

[5] 程旭, 康重庆, 夏清, 等. 短期负荷预测的综合模型[J]. 电力系统自动化, 2000, 24(13): 42-44.

[6] 李妮, 江岳春, 黄珊, 等. 基于累积式自回归平均传递函数模型的短期负荷预测[J]. 电网技术, 2009, 33(8): 93-97.

(收稿日期: 2015-05-26) EA

作者简介

张春雷 (1981-), 男, 工程师, 主要研究方向为电力系统继电保护科研与开发工作。

李炎 (1982-), 男, 工程师, 主要研究方向为电力系统负荷预测科研及开发工作。