

基于距离矢量法的短期负荷预测方法研究

短期负荷预测是电力企业调度运行的基础,科学合理的负荷预测是进行错峰的必要条件。提出了在负荷特性分析的基础上进行改进的距离矢量法的聚类方法,并在此基础上提出了点对点加权求值法的负荷预测方法。实际系统的应用效果验证了该方法的有效性和实用性。

黄晓彤¹ 陈文炜¹ 王劲锋¹ 陈娅楠² 廖志伟^{2/1}. 广州供电局有限公司

2. 华南理工大学电力学院

0 引言

短期电力负荷预测主要是预报未来几个小时,一般是1天到几天的电力负荷数据^[1-3]。已研究的负荷预测方法主要分为传统预测方法和人工智能算法。传统负荷预测方法主要有线性回归分析法^[4]、时间序列法^[5-6]等。人工智能方法主要包括专家系统、模糊逻辑方法和人工神经网络方法^[7]等。

短期负荷预测可分为2个基本过程:一是选取相似日;二是根据相似日负荷进行加权、外推来预测待预测日负荷。相似日选取的好坏直接影响预测准确度^[9-10],其是根据温度、湿度等气象条件和节假日因素选取,相关性越高说明为同类型日的可能性越大,然而每种负荷影响因子的影响作用不一,难以准确把握负荷变化规律。

基于以上分析,本文提出一种新的利用距离矢量法来实现聚类分析的方法,将负荷特性的距离作为参量,把给定数据集划分成多个组或簇群,使得每组内部的数据矢量较为相似,组间的数据矢量则差别较大,从而得到该负荷的聚类曲线,再通过待预测日的工作日类

型进行点对点负荷预测。

1 距离矢量法

距离矢量法是一种基于划分的动态聚类算法,采用距离作为相似性的评价指标,即认为两个对象的距离越近,其相似性就越大。该类算法认为簇是由距离靠近的对象组成的,将获取紧凑且独立的簇作为最终目标。

目前已研究的聚类算法大多数采用欧式距离来度量样本之间的相似性。而诸如数据值突变等异常情况将引起欧式距离变大,将造成负荷无法准确归类。故本文引进相似性点数阈值,比较负荷矢量中不同时刻点的距离,并得出两个矢量满足距离阈值的点的个数,与相似性点数阈值进行比较,来度量样本之间的相似性。

距离矢量法的聚类结果主要取决于样本容量、距离阈值与点数阈值、待分类特征矢量参与分类的次序和初始特征矢量的选取。

距离阈值与点数阈值的选取,通常是采用试探法,选用不同的阈值及矢量输入次序来试分类,并对聚类结果进行检验,即用聚类准则函数。例如,计算每一聚类中心与该类中最远样本点的矢量距离,或计算类间方差,用这些结果指导及校验阈值与输入次序的重选。最后对各种方案的划分结果进行比较,选取最好的一种聚类结果。

1.1 条件、目标及约定

把待分类的负荷集合划分为 $\{C_1, C_2, C_3, \dots, C_n\}$,其中为历史日负荷集合,为选取的历史日的天数,这里



黄晓彤/工程师

关键词/Keywords

短期负荷预测 ·
距离矢量法 ·
点对点加权求值法 ·

国家自然科学基金(00000000),国家高技术研究发展计划(863计划)(2008AA000000)。

$C_i = D$, C_i 为其中一天的负荷, 一天96个负荷值表示为矢量 $C_i = (C_{i1}, C_{i2}, \dots, C_{i96})$ 。

聚类分析的目标是把负荷集合 D 中相似的曲线划分到1个类别里边, 重新划分为 m 个特征类 $\{U_1, U_2, U_3, \dots, U_m\}$ 。其中, U_j 为聚类后的类别, m 为聚类后的特征类数。选定相似性距离阈值 I , 相似性点数阈值 p 。

距离矢量法将一个对象最多只归于一个聚类, 每个聚类都是全体对象集 D 的一个子集。它把待归类矢量 C_i 与已有的特征类 U_j 进行矢量相减, 当得到的矢量值96个点中满足负荷阈值 I 的点数不小于 m 时, 将 C_i 归属到 U_j 类中, 并相应地修改 C_i 类的特征矢量, 若在特征类中找不到相似曲线, 则 C_i 作为新的一类加入到特征类中。定义函数

$$g(C_i) = \sum_{k=1}^{96} f(I - |C_{ik} - U_{jk}|) - m \quad (1)$$

$$f(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (2)$$

当 C_i 满足条件 $g(C_i) \geq 0$, C_i 与 U_j 相似。

1.2 算法原理步骤

处理流程如下:

- 1) 在 D 取一个待归类矢量作为第一个特征矢量。例如取第一天, 令 $U_1 = C_1$ 。
- 2) 选取下一个待归类矢量, 若已到最后一个, 则结束计算, 输出所有特征矢量。
- 3) 将待归类矢量 C_i 依次与特征矢量集中的矢量进行相似性判断, 若 C_i 与 U_j 满足相似性条件, 则停止比较, 令 $U_j = C_i + U_j / 2$, 若不满足, 令 $U_{k+1} = C_i$, U_k 为特征矢量集最后一个矢量, 返回第2)步。

3 短期负荷预测

经研究发现, 负荷数据主要聚类为工作日类型和节假日类型, 而且温度等气象条件相对类似, 故选取近期一个月的数据作为预测样本集, 根据待预测日工作日类型, 可对待预测日进行96点负荷预测。

正常日点对点加权求值法的预测思路为: 待预测日96点的负荷值可由同类的同一时段的负荷

值加权平滑结果求得, 所谓点对点, 是指逐时刻进行。步骤如下:

- 1) 求样本集各时的负荷和

$$P_t = \sum_{i=1}^n P_{it}, t = 1, 2, \dots, T \quad (3)$$

- 2) 求样本集各日各时负荷系数比

$$\lambda_{it} = P_{it} / P_t \quad (4)$$

- 3) 进行96点负荷预测

$$P_{0t} = \sum_{i=1}^n \lambda_{it} P_{it} \quad (5)$$

4 案例分析

对某电网负荷数据采用负荷影响因子数据库同类型日和本文方法两种方法进行预测比较。

4.1 聚类

4.1.1 方法一

选取某变电站6月负荷数据来做负荷预测, 其中, 前29天为历史日, 30号为待预测日, 选取其中的10条进行展示如图1所示。经过分析, 得出簇类。

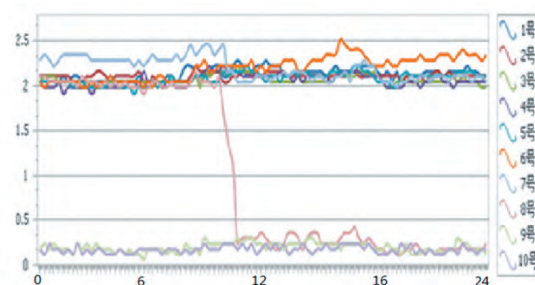


图1 某电网2012年6月负荷值

根据距离矢量法对历史数据进行分类, 得到分类效果图如图2所示具体日期分类表格1所示。

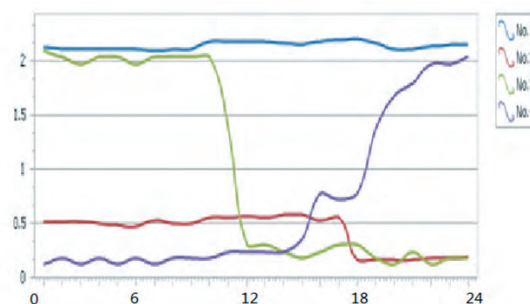


图2 聚类分析图

表1 聚类分析日期表

聚类类型	日期
第一类	1, 2, 3, 4, 5, 6, 7, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29
第二类	9, 10, 11, 12, 13
第三类	8
第四类	14

4.1.2方法二

对于历史日的日特征量，进行无量纲化处理。对于温度，以30℃为界限，30℃以下作线性处理，刻度为0.1，30℃以上作刻度为0.2的线性映射处理；对于星期类型，周一~周五分别映射0.1~0.5，周六、周日映射为3，以此来加大区分工作日和节假日的不同，对于天气类型，分为晴、多云、阵雨、大雨和雪，分别映射为0.5，1，2，3，4。根据以上映射规则，建立负荷影响因素映射如表2所示。

表2 负荷影响因素映射表

时间	最高温度 / ℃	映射值	最低温度 / ℃	映射值	星期类型	映射值	天气类型	映射值
6/1	31	0.2	24	2.4	周五	0.5	多云	1
6/2	33	0.6	25	2.5	周六	3	多云	1
.....								
6/21	32	0.4	25	2.5	周四	0.4	大雨	3
6/22	31	0.2	26	2.6	周五	0.5	大雨	3
6/23	30	0.1	26	2.6	周六	3	阵雨	2
.....								
6/28	35	1	26	2.6	周四	0.4	多云	1
6/29	35	1	24	2.4	周五	0.5	多云	1

待预测日日负荷特征向量为 U_i ，历史日日负荷特征向量为 U_j 。

根据式（4）计算相关度得出7号、8号、15号、19号、20号和28号为同类型日。

$$\rho_{ij} = \frac{\sum_{i=1}^n u_i u_j}{\sqrt{\sum_{i=1}^n u_i^2 \sum_{i=1}^n u_j^2}}$$

(6)

4.2 负荷预测

通过两种负荷预测方法进行负荷预测，其预测曲线和实际日曲线，对比分析图如图3所示。

4.3 误差分析比较

由于日负荷预测的误差指标较多，本文选取均方根相对误差指标作为以上算法预测效果判断的依据。

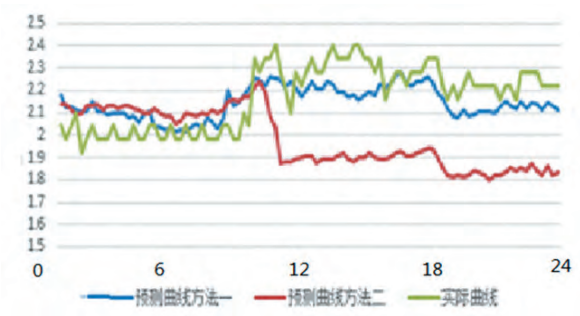


图3 预测曲线和实际曲线对比图

均方根相对误差 对相对误差进行了平方，加强了数值大的误差在指标中的作用，从而提高了该指标的灵敏度，计算公式为

$$EMSE = \left(\frac{Y_i - Y_i'}{Y_i} \right)^2$$

(7)

式中， Y_i 为实际负荷； Y_i' 为负荷预测值。

对以上预测方法进行误差分析,具体如图4所示。

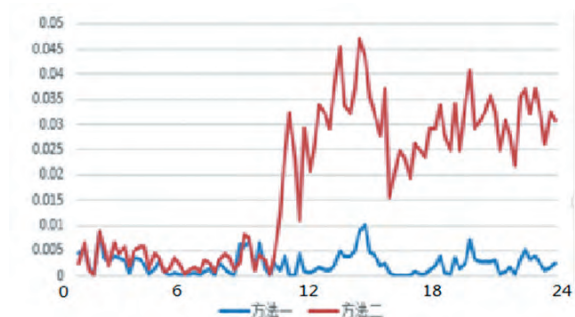


图4 误差曲线(%)分析图

从图3和图4可以直观地看出,相对实际负荷曲线,方法一的误差都在1%以内,绝大部分预测点的预测误差分布在0%~0.2%之间,基本上所有点的预测精度可以控制在0.7%以内,比方法二的预测准确度高很多。

以上分析图表可知:

1) 距离矢量法的预测准确度控制在0.7%以内,历史日负荷曲线是各种负荷特性影响因子的综合作用结果,直接从负荷曲线的距离来进行聚类,避免了各种影响因子的权重不一。

2) 日特征量法里的最高温度,最低温度特征量没有考虑到温度的累加效应,而且在映射过程中没有明确的规定,映射规则不一,预测的结果也不一样。再者,预测日的温度和天气类型都是基于预测结果,本身就具有不确定性。

实验结果表明,通过距离矢量法进行聚类,再通过点对点确定历史日与待预测日两者结合的方法是对同类型日法的一种改进,具有比相似日法更高的准确度。

5 结束语

短期负荷预测方法预测时间短,预测结果的准确度对电力系统实际应用有重大意义。本文通过比较基准日与历史日的96点负荷数据来进行聚类分析,进行聚类,在此基础上利用点对点加权求平均法对广州南沙地区6月份负荷数据做后验比较。通过误差计算,发现这种方法具有较高的准确度和较强的工程实用价值。采用这种方法减少了相似日方法对人的经验的依赖,同时也避免了建立负荷特性影响因素数据库以及每天更新数据

库的麻烦。其不但方便维护和扩展,而且可以更大提高负荷预测的准确度。

参考文献

- [1] 张志明,金敏.基于灰关联分段优选组合模型的短期电力负荷预测研究[J].电工技术学报,2009,24(6):115-120.
- [2] 周德强,武本令.灰色BP神经网络模型的优化及负荷预测[J].电力系统保护与控制,2011,39(21):65-68.
- [3] 鞠平,浆巍,赵夏阳,等.96点短期负荷预测方法及其应用[J].电力系统保护与控制,2011,39(21):65-69.
- [4] 谷子,唐巍.电力短期负荷时间序列混沌相空间重构参数优选法[J].中国电机工程学报,2006,26(14):18-23.
- [5] 王凌,刘波.微粒群优化与调度算法[M].北京:清华大学出版社,2008.
- [6] Fatih Tasgetiren M, Pan Quan-Ke, Suganthan P N, et al. A discrete differential evolution algorithm for the no-wait flowshop scheduling problem with total flowtime criterion[C]. Proceedings of the 2007 IEEE Symposium on Computational Intelligence in Scheduling .New York, USA: ACM, 2007: 251-258.
- [7] 陈刚,周杰,张雪君,等.基于BP与RBF级联神经网络的日负荷预测[J].电网技术,2009,33(12):101-105.
- [8] 牛东晓,谷志红,邢棉,等.基于数据挖掘的SVM短期负荷预测方法研究[J].中国电机工程学报,2006,26(18):6-12.
- [9] 黎灿兵,李晓辉,赵瑞,等.电力短期负荷预测相似日选取算法[J].电力系统自动化,2008,32(9):69-73.
- [10] 莫维仁,张伯明,孙宏斌,等.短期负荷预测中选择相似日的探讨[J].清华大学学报:自然科学版,2004,44(1):106-109.

(收稿日期:2015-05-26) EA

作者简介

黄晓彤(1969-),男,工程硕士,工程师,主要研究方向为电力设备运行与控制。