

电力短期负荷预测相似日选取算法

黎灿兵¹, 李晓辉¹, 赵 瑞², 李金龙³, 刘晓光¹

(1. 郑州大学电气学院, 河南省郑州市 450001; 2. 北京电力公司调度中心, 北京市 100031;
3. 华北电力大学电气与电子工程学院, 北京市 102206)

摘要:短期负荷预测是电力系统安全经济运行的基础,相似日选取的准确与否直接影响到短期负荷预测算法的精度。针对短期负荷预测的特点,提出一种能便于考虑各种因素影响的新算法。分析了气象、日类型等因素对负荷影响的常见规律,便于识别主导负荷变化的因素,建立了在短期负荷预测中选取相似日的新方法。理论和实例均表明,该方法适应性较强,能够通过历史数据分析从历史日中选取最合适的相似日,对提高短期负荷预测的精度具有较大价值。

关键词:短期负荷预测;相似日;气象因素;累积效应

中图分类号: TM715

0 引言

负荷预测是电力生产的基础^[1]。短期负荷预测针对未来一天到数天各时段的负荷预测进行研究^[2-4]。文献[5]将短期负荷预测分为 4 个研究阶段,揭示了负荷预测技术发展的方向,即越来越充分考虑气象因素的影响。近年来,空调负荷比重逐渐增加,气象因素对负荷的影响加大^[6-7],且出现的现象和规律也越来越复杂^[8-10]。文献[11]提出了一种规范处理各种相关因素的方法,便于考虑新增的影响因素和新规律。短期负荷预测一般可分为 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精度^[12]。因此,需要研究相似日选取算法以提高短期负荷预测的精度^[13-14],其中大量文献采用人工神经网络来选取相似日^[15-16],需要大量的样本进行训练以获得负荷变化规律,而对新出现的少量样本所体现的新规律反应不灵敏,气象变化时也不能准确把握负荷变化规律。

本文深入分析了常见因素对负荷影响的一般规律,提出了短期负荷预测中选取相似日的新算法。

1 算法基本设计

不同的条件下,影响负荷变化的因素不同。一般存在 1 个~2 个主导因素。举例说明如下:①在大部分地区,当最高气温超过 36℃时,气温是影响负荷变化的主导因素,其他因素对负荷的影响相对很小。如某日为星期六,最高气温 37℃,历史日中

只有上星期二最高气温超过 36℃,虽然日类型不同且相距十来天,但仍然极可能是与待预测日负荷水平最接近的历史日。②以居民负荷和商业负荷为主的大城市中心区域,在气温不太高或其他气象条件并非极端条件的情况下,日类型为主导因素。③一般情况下,如果日类型是特定节假日,则日类型为负荷变化的主导因素。④在冬季低温季节,最低气温可能成为主导因素。因此,成熟的短期负荷预测算法应能识别各种条件下影响负荷变化的主导因素,确保选取的相似日真正与待预测日相似。

在本算法中,设共考虑 N 种因素, D 个历史日。针对每种因素,分别计算各历史日与待预测日的相似度,获得相似度矩阵 $\mathbf{M}$ 如下:

$$\mathbf{M} = [m_{id}] \quad i = 1, 2, \dots, N; d = 1, 2, \dots, D \tag{1}$$

式中: $\mathbf{M}$ 为相似度矩阵; m_{id} 为每个历史日的每个因素与待预测日对应因素的相似度。

根据 $\mathbf{M}$ 计算各历史日与待预测日的总相似度。有 2 种常见计算方式:

1) 将历史日各因素的相似度相加,即

$$F_d = \sum_{i=1}^N m_{id} \tag{2}$$

式中: F_d 为第 d 个历史日与待预测日的总相似度。

采用式(2)计算各日相似度时,需解决各因素权重设定问题。如在夏季高温季节,最高气温是主导负荷变化的主要因素,应赋予最高气温较大权重。

2) 将历史日各因素的相似度相乘,即

$$F_d = \prod_{i=1}^N m_{id} \tag{3}$$

采用式(3)有如下优点:

1)能简单、自动地识别主导因素。主导因素可形象地理解为:某因素主导作用越强,其他因素对负荷的影响程度相对降低。当第*i*个因素的重要性增加时,满足下式:

$$\frac{\partial F_d}{\partial m_{jd}} \rightarrow 0 \quad j \neq i \tag{4}$$

式(4)的含义是,若第*i*个因素是主导因素,则其他因素对总相似度的影响程度下降。例如,最高气温是主导因素,则最高气温的变化会使总相似度的变化相对较大;其他因素变化难以明显改变总相似度排序。因此,能够避免某个历史日与待预测日在关键因素上差别很大,但由于其他因素很相似而被算法误选为相似日。

按式(4)定义总相似度时,有下式成立:

$$\frac{\partial F_d}{\partial m_{id}} = \prod_{j \neq i} m_{jd} \tag{5}$$

由式(5)可知,当考察第*i*个因素时,总相似度对其他因素的偏导数中,均含有因子*m_{id}*。当*m_{id}*较小时,第*i*个因素成为主导因素,而主导因素差距明显时,决定了2个日期不相似,下面举例说明。

例 1 设历史日、待预测日数据如表 1 所示,假设程序自动对 2 个历史日 A、B 与待预测日 D 的相似度进行对比分析。

表 1 主导因素分析
Table 1 Main factor analysis

日期	日类型	日期 距离/d	得分	最高气 温/(℃)	最低气 温/(℃)	相似度		
						日 类型	最高 气温	最低 气温
D	星期四			36.5	25			
A	星期四	7	0.65	28.0	25	1.0	0.3	1.0
B	星期六	5	0.80	36.5	28	0.4	1.0	0.7

表 1 中,分别计算 2 个历史日与待预测日各因素的相似程度。A 的累加总相似度为 2.95,累乘总相似度为 0.195; B 的累加总相似度为 2.9,累乘总相似度为 0.224。因此,采用累加,A 比 B 相似;采用累乘,B 比 A 相似。注意到,待预测日最高气温达 36.5℃,该情况下最高气温是主导因素,B 与 D 比 A 与 D 更相似,因此,累乘更符合实际情况。如果在某一主导因素上有明显差距,则可决定基本不相似。

2)可解决各因素的权重设定问题。调整各因素权重值,不影响历史日排序。因此,在参数自适应过程中,能够减少需要自适应的参数数量,降低计算量。

因此,本算法中采用累乘方式,且各因素的权重

值均为 1,即每个因素相似度的最大值为 1。

2 主要影响因素相似度计算方式

2.1 日类型

日类型是指工作日或周末。如星期一用“1”表示,其他类推。一般,周末负荷明显低于工作日,且负荷曲线形状也有明显区别。日类型相同,获得该因素最大相似度 1;同为工作日但日类型不同,给出一个相似度,如 0.7;等等。

2.2 是否为特殊日

特殊日指负荷明显不同的特定日期。主要包括一些节日,如元旦、春节等;另外,一些地区的灾害性天气、举办特定活动的日期也可认为是特殊日。特殊日的相似度计算可分 4 种情况分别给定相似度:同为同一种特殊日;都是特殊日但特殊日日期不同;同为非特殊日;一个特殊日、一个非特殊日。

2.3 日期距离

一般,离待预测日越近越相似。因此,在选择相似日时,有明显的“近大远小”规律。日期距离的相似度按下式计算:

$$m_{id} = \begin{cases} \beta^k & \beta^k \geq a \\ a & \end{cases} \tag{6}$$

式中:*m_{id}*为第*d*天日期距离因素的相似度;*β*为衰减系数,含义是历史日与待预测日的距离每增加 1 d 的相似度缩减比率,建议取值在 0.90~0.98 之间;*a*为该因素的最低相似度。

注意到,距离对负荷变化的影响具有明显的饱和效应,例如,历史日与待预测日相隔 1 d 或相隔 1 星期,其与待预测日的相似度有明显区别;但历史日与待预测日相隔 3 星期或 4 星期,其相似度则没有明显区别。一些地区,夏季可能出现持续 20 d 甚至更长时间的高温天气,然后出现降雨、突然降温,在近 20 d 没有相似日,而如果该相似度按照某一个衰减系数持续衰减,则距离远的日期没有入选的可能性,不能准确预测。因此,需设定该因素的相似度下限。

2.4 最高气温

最高气温对相似度的影响需要考虑最高气温的累积效应和非线性 2 种现象。

2.4.1 最高气温的累积效应

最高气温的累积现象在空调负荷比重较大的地区十分明显,根本原因是人体的舒适度感觉存在惯性。连续高温时,即使气温已有所下降,还是习惯性感知为高温天气;在连续凉爽的情况下,即使气温显著攀升,但人对高温天气还处于适应过程中,空调负荷仍不大。最高气温的累积效应按下式处理:

$$T' = (1 - k)T + \frac{k(T_{-1} + \lambda T_{-2} + \lambda^2 T_{-3})}{1 + \lambda + \lambda^2} \quad (7)$$

式中: k 为累积系数; λ 为衰减系数, 即每往前推 1 d, 权重系数衰减的比率; T_{-i} 为待预测日前第 i 日的最高气温; T 为待预测日的最高气温; T' 为考虑最高气温累积效应修改后的待预测日的最高气温。

值得注意的是, 在不同情况下, 累积系数 k 应取不同值。例如, 当 $T \leq 24$ °C 时, 无论前一日气温多高, 空调已基本关闭, 之前的气温对该日没有明显影响; 同理, 当 $T \geq 37$ °C 时, 空调已基本满负荷, 之前的气温对该日也没有明显影响。可见, 累积系数应动态调整。基本原则是: 空调可开可不开, 或者设定温度可高可低的情况下, 累积效应最明显; 在空调负荷为 0 或满负荷的情况下, 累积效应不明显。

2.4.2 非线性

气温对负荷的影响存在明显的非线性现象, 应对气温进行调整, 使调整后的值与负荷的关系基本呈线性关系。体现为: ①气温较低时, 空调体现为升温负荷, 气温越低, 负荷越高; 而且随着气温的降低, 负荷增加速度提高, 直到饱和。②气温较高时, 空调负荷体现为降温负荷, 气温越高, 负荷越高; 而且随着气温的升高, 负荷增长加速, 直到饱和。③气温处于中间时, 气温对负荷变化基本没有影响。

综上所述, 气温对负荷的影响如图 1 所示。

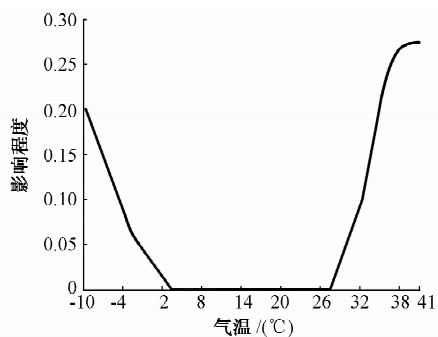


图 1 气温对负荷的影响规律

Fig. 1 The relationship between temperature and load

在对最高气温因素计算相似度时, 首先考虑最高气温累积效应进行修正处理, 然后采用类似图 1 的映射表, 将修正后的最高气温映射为一个修正值, 在映射中考虑气温对负荷影响的非线性效应, 最后根据映射值计算相似度。

2.5 湿度

湿度对负荷的影响依据气温而定。主要现象包括: 气温较高时(指 33 °C 或以上), 湿度越大, 负荷越高, 湿热天气使空调负荷快速上升; 气温低于 5 °C 时, 湿度越大, 也导致负荷越高。湿冷天气将导致“寒冷指数”上升, 人体舒适度下降, 空调负荷上升;

气温在居中的区间时, 湿度对负荷的影响较小。

考虑以上现象, 在评估湿度因素的相似度时, 在不同的气温条件下, 确定不同湿度对负荷的影响系数, 再根据影响系数计算相似度。

2.6 降雨量

降雨影响负荷的现象包括: ①降雨量较大、持续时间较长时, 将导致气温降低, 间接影响负荷变化。②降雨导致湿度增加。在夏季, 短暂降雨可使湿度显著增加, 形成湿热天气, 可能使负荷不降反升。在冬季气温较低时, 湿度增加使寒冷指数显著增加, 将导致空调负荷增加。③有较多的非统调径流式小水电的地区, 降雨导致相关流域流量增加, 小水电出力增加, 网供负荷降低。

因此, 降雨可能导致负荷升高或降低, 根据气温、是否具有较多小水电等外部环境而定; 而且, 同样大小的降雨量, 降雨强度和起始时间不同, 对负荷的影响程度也不同。在算法中还需要根据如下具体情况考虑: 在绝大部分地区, 气象部门对降雨量有实测数据, 但预报只能以特大暴雨、暴雨、大雨、中雨、小雨、阵雨等文字描述; 气象部门基本不能提供降雨开始的具体时间。

本算法考虑降雨量的基本方式如下:

1) 首先确定降雨是令负荷升高还是降低。

2) 根据对负荷影响的 2 个方向, 按照降雨量大小排序, 依次确定各种降雨形式对负荷影响的程度。例如: ①特大暴雨、大暴雨的降雨量较大, 在高温或低温季节, 都会使负荷较大程度降低, 影响系数是绝对值较大的负数; ②在高温季节, 除了阵雨外, 其他降雨都将导致负荷降低, 因此, 根据降雨大小分别设定影响系数, 且影响系数均为负; ③在低温季节, 降雨可引起负荷小幅度增加, 但增加幅度较小, 且不同地区不同; ④在高温季节, 短暂降雨(阵雨), 使湿度增加, 对负荷有小幅度的促进增长效果, 但增长幅度有限。

2.7 最低气温

最低气温对负荷的影响主要在冬季, 最低气温越低, 空调负荷的比重上升。注意到冬季空调负荷的比重显著低于夏季, 因为降温负荷主要是用电空调, 但取暖可能通过燃煤集中供暖等更有效的方式。因此, 最低气温对空调负荷的影响可参照最高气温对空调负荷的影响确定, 但相应的调整系数要小一些, 此处从略。

3 算法的参数自适应性

3.1 参数自适应的方式和策略

第 2 节分析了各主要因素的相似度计算方式,

需要根据各地区、各季节的具体情况,调整参数,使之自动适应具体地区、具体时段的负荷特性。本文算法中参数自动调整的方式与策略如下。

1)每个因素的权重值不调整。按照式(3)计算每个历史日的相似度,而每个因素权重的设定,对各个历史日的总相似度排序不构成影响,因此,每个因素的权重都是1。

2)每个因素计算相似度的函数不调整。第2节分析了7个相关因素相似度函数的基本框架,这些函数在自适应过程中各项参数不变,不参与训练。原因如下:①相似度函数基本上较全面地考虑了这些因素影响负荷变化的主要规律。②每个因素的相似度函数参数较多,若全部进行参数自适应训练需要大量的样本,而大量样本中必然是历史时间较长,导致不能反映最新的负荷发展规律。因此,参与自适应训练的参数越少,算法反应越灵敏。③需要训练的参数多会导致自适应训练的计算量过大。

3)每个因素设定拟合指数。为了准确反映不同地区、不同季节各个相关因素对负荷的影响程度,通过自适应训练确定各个因素的影响指数,如下:

$$F_d = \prod_{i=1}^N (m_{id})^{R_n} \tag{8}$$

式中: R_n 为第 n 种因素的影响指数。

可通过调整 R_n 来调整各个因素的影响程度, R_n 越大,该因素越重要。同时, R_n 越大,该因素不同历史日与待预测日的相似度差距会扩大,在该因素上不相似的历史日入选为相似日的可能性下降,表明该因素在当前情况下主导负荷变化的程度增加。因此, R_n 的大小表明各因素对负荷的影响程度。

每种影响因素各有一个对应的影响指数,共 N 个参数对应 N 个影响指数,若 N 个影响指数同步增长或减小,不影响相似度的排序,自由度为 $N-1$ 。可指定一个因素的影响指数为1,如日类型,对剩余 $N-1$ 个影响指数逐一优化,这样计算量较小且能反映最新负荷变化规律。

3.2 参数自适应流程

参数自适应过程是一个优化过程,各因素的影响指数是决策变量,目标是使各历史日与待预测日的相似度排序,与各历史日与待预测日的负荷水平的接近程度的排序尽可能一致。本算法中设定参数自适应的目标函数为如下3个比率之和:①选取的第1相似日是真实最相似日的比率;②选取的2个相似日是历史日真实相似度前2名的比率;③第3相似日和第4相似日类推的比率。

4 算例分析

表2为华中某市2007年6月的部分数据。

表2 2007年6月的部分负荷及气象因素
Table 2 Load and meteorological data in the June, 2007

日期	$T_m/(^{\circ}\text{C})$	$T_a/(^{\circ}\text{C})$	$T_z/(^{\circ}\text{C})$	J/mm	$S/(%)$	L/MW
6月8日	29.8	30.0	25.6	0.0	50	2 930.20
6月9日	32.2	28.8	26.8	0.0	59	2 835.31
6月10日	32.8	26.3	25.4	0.9	67	2 553.86
6月11日	35.3	27.4	22.9	0.0	57	2 697.59
6月12日	35.8	25.7	23.1	7.7	69	2 574.97
6月13日	32.3	24.3	23.4	2.6	80	2 507.03
6月14日	28.3	24.2	21.1	0.0	65	2 408.51
6月15日	31.1	25.2	21.1	0.0	62	2 512.58
6月16日	32.7	27.3	22.7	0.0	57	2 527.61
6月17日	33.8	28.1	24.7	0.0	57	2 516.09
6月18日	36.8	28.0	26.0	0.0	65	2 707.84
6月19日	36.4	25.9	23.4	16.5	84	2 639.17
6月20日	37.4	23.1	21.1	15.3	81	2 397.13
6月21日	36.5	25.2	22.9	0.0	73	2 489.73
6月22日	35.3	26.9	23.9	0.0	73	2 722.39
6月23日	31.1	25.2	22.4	52.1	80	2 505.17

注: T_m 为最高气温; T_a 为平均气温; T_z 为最低气温; J 为降雨量; S 为相对湿度百分数; L 为平均负荷。

算法程序从45个历史日中寻找相似日,但因6月份的相似日很少发生在较长时间之前,因此表2只列出最近一段时间的详细数据。

预测6月23日平均负荷,算法所选取的相似日分别是6月15日、6月21日、6月16日。6月15日为最相似日的原因是,当前一段时期温度比较高,是主导因素,6月15日气温与待预测日最接近;6月16日入选的原因是最高气温差别不大,且日类型相同;6月21日入选的原因是平均气温相同且湿度差别不大;等等。实际情况表明,这3个相似日与待预测日的平均负荷差别均在1%以内。

5 结语

选择合理的相似日,是提高短期负荷预测精度的有效途径。本文提出的短期负荷预测中相似日的选取方法,深入分析了各种因素的影响规律,自动识别主导因素,且具有自适应能力,从而能够选取真实相似的历史日。采用该方法选取相似日,对提高短期负荷预测的水平具有较大的价值。

参考文献

[1] 刘晨晖. 电力系统负荷预报理论与方法. 哈尔滨: 哈尔滨工业大学出版社, 1987.

[2] ABDEL-AAL R E. Short-term hourly load forecasting using abductive networks. IEEE Trans on Power Systems, 2004, 19(1): 164-173.

- [3] ESPINOZA M, JOYE C, BELMANS R, et al. Short-term load forecasting, profile identification, and customer segmentation: a methodology based on periodic time series. *IEEE Trans on Power Systems*, 2005, 20(3): 1622-1630.
- [4] FAN S, CHEN L. Short-term load forecasting based on an adaptive hybrid method. *IEEE Trans on Power Systems*, 2006, 21(1): 392-401.
- [5] 康重庆,周安石,王鹏,等.短期负荷预测中实时气象因素的影响分析及其处理策略. *电网技术*, 2006, 30(7): 5-10.
KANG Chongqing, ZHOU Anshi, WANG Peng, et al. Impact analysis of hourly weather factors in short-term load forecasting and its processing strategy. *Power System Technology*, 2006, 30(7): 5-10.
- [6] KARAKI S H. Weather sensitive short-term load forecasting using artificial neural networks and time series. *International Journal of Power and Energy Systems*, 1999, 19(3): 251-256.
- [7] 谢宏,陈志业,牛东晓.短期电力负荷预测的数据主成分分析. *电网技术*, 2000, 24(1): 43-46.
XIE Hong, CHEN Zhiye, NIU Dongxiao. Principle component analysis of data processing for short-term load forecasting. *Power System Technology*, 2000, 24(1): 43-46.
- [8] LAMEDICA R, PRUDENZI A, CACIOTTA M, et al. The influence of meteorological parameters on Italian electric hourly load: the selection of variables of the ANN training set for short term load forecasting// *Proceedings of Electrotechnical Conference*, Vol 3, May 13-16, 1996, Bari, Italy: 1453-1456.
- [9] 赵锡平,宋岱,张国庆,等.山东省气温与最大负荷和用电量的关联性分析. *电网技术*, 2004, 28(17): 37-40.
ZHAO Xiping, SONG Dai, ZHANG Guoqing, et al. Research on correlativity among air temperature, maximum load and power consumption in Shandong Power Grid. *Power System Technology*, 2004, 28(17): 37-40.
- [10] 张梅,陈玉光,韩家福,等.辽阳地区6~8月耗电量与气象条件关系及预报. *气象与环境学报*, 2006, 22(2): 62-64.
ZHANG Mei, CHEN Yuguang, HAN Jiafu, et al. Relationship between electric power consumption and meteorological conditions and its forecast from June to August in Liaoyang. *Journal of Meteorology and Environment*, 2006, 22(2): 62-64.
- [11] 康重庆,程旭,夏清,等.一种规范化的处理相关因素的短期负荷预测新策略. *电力系统自动化*, 1999, 23(18): 32-35.
KANG Chongqing, CHENG Xu, XIA Qing, et al. A new unified approach to short-term load forecasting considering correlated factors. *Automation of Electric Power Systems*, 1999, 23(18): 32-35.
- [12] 鞠平,王俊锴,赵夏阳.96点短期负荷预测方法及其应用. *电力系统自动化*, 2001, 25(22): 32-36.
JU Ping, WANG Junkai, ZHAO Xiayang. Ninety-six points short-term load forecasting—theory and applications. *Automation of Electric Power Systems*, 2001, 25(22): 32-36.
- [13] 金海峰,熊信良,吴耀武.基于相似性原理的短期负荷预测方法. *电力系统自动化*, 2001, 25(23): 45-48.
JIN Haifeng, XIONG Xingen, WU Yaowu. Short-term load forecasting based on analogous theory. *Automation of Electric Power Systems*, 2001, 25(23): 45-48.
- [14] 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. *清华大学学报:自然科学版*, 2004, 44(1): 106-109.
MO Weiren, ZHANG Boming, SUN Hongbin, et al. Method to select similar days for short-term load forecasting. *Journal of Tsinghua University: Natural Sciences*, 2004, 44(1): 106-109.
- [15] TOMONOBU S, HITOSHI T, KATSUMI U, et al. One-hour-ahead load forecasting using neural network. *IEEE Trans on Power Systems*, 2002, 17(1): 113-118.
- [16] HENRIQUE S H, CHARLOS E P, REINALDO C S. Neural networks for short-term load forecasting: a review and evaluation. *IEEE Trans on Power Systems*, 2001, 16(1): 44-55.

黎灿兵(1979—),男,通信作者,博士,讲师,主要研究方向:电力负荷预测、规划、信息化与新能源。E-mail: licanbing@gmail.com

李晓辉(1982—),男,硕士研究生,主要研究方向:电力负荷预测及规划。E-mail: lixiaohui_lxh@gmail.com

赵 瑞(1980—),男,助理工程师,主要研究方向:电力负荷预测与调度管理。

A Novel Algorithm of Selecting Similar Days for Short-term Power Load Forecasting

LI Canbing¹, LI Xiaohui¹, ZHAO Rui², LI Jinlong³, LIU Xiaoguang¹

(1. Zhengzhou University, Zhengzhou 450001, China; 2. Beijing Electric Power Corporation, Beijing 100031, China;

3. North China Electric Power University, Beijing 102206, China)

Abstract: Short-term load forecasting is the basis of safe and economical operation of power grids. The accuracy of selecting similar days directly influences the degree of accuracy of the short-term load forecasting algorithm. After an in-depth analysis of the common laws of such factors as the meteorological factors and the day type that impact power loads, a new algorithm of selecting similar days for the short-term load forecast is proposed. The algorithm can distinguish between the key factors in many cases. The simulation results demonstrate that it can select the most suitable similar day from a large number of historical data to improve the accuracy of short-term load forecasting with relatively strong adaptability.

Key words: short-term load forecasting; similar day; meteorological factors; cumulative effect