

# 考虑气象因素累积效应的 电力系统 96 点日负荷预测方法

李汉巨

(广东电网有限责任公司东莞供电局, 广东 东莞 523008)

**摘要** 为了提升预测精度, 根据电力系统日负荷曲线特性以及不同时段负荷对气象因素敏感度的差异性, 将 96 点日负荷划分为 4 个时段分别预测。基于预测日的前一日, 以各个时段平均负荷变化为输出变量, 以预测日的日类型、月份、最高温度变化、最低温度变化和前一日的最高温度变化、最低温度变化为输入变量, 建立多元线性回归模型。通过该模型可计算预测日的平均负荷与前一日的平均负荷的比率。相关性分析表明相邻日的负荷曲线具备较高线性相关度, 因此基于前一日的负荷数据, 可以推算出预测日的负荷。以东莞地区负荷为例, 结果分析表明, 本文方法具备较高的稳定性和准确性, 优于传统方法。

**关键词:** 负荷预测; 电力系统; 日负荷; 气象

## The 96 point daily load forecasting method for power system considering the cumulative effect of meteorological factors

*Li Hanju*

(Dongguan Power Supply Bureau, Guangdong Power Grid Co., Ltd, Dongguan, Guangdong 523008)

**Abstract** In order to improve the prediction accuracy, according to the characteristics of daily load curve of electric power system and the sensitivity of different period load to meteorological factors, the daily load of 96 points is divided into four periods, respectively. Based on the day before the forecasting day, the multivariate linear regression models are set up by taking the average load change at each period as the output variable and taking daily type, month, maximum temperature change, minimum temperature change of the forecasting day and maximum temperature change, minimum temperature change of the day before the forecasting day as input variables. The model can calculate the ratio of the average load of the forecasting day to the average load of the previous day. The correlation analysis shows that the load curve of adjacent days has a high linear correlation, so the load of forecast day can be calculated based on the previous day's load data. Taking the load in Dongguan area as an example, the results show that the method has higher stability and accuracy, and is better than the traditional method.

**Keywords:** load forecasting; electric power system; daily load; meteorological

电力系统日负荷预测是提前一天对次日规定的所有时刻负荷值做出的估计, 本质上可看作为多维矢量预测, 其特点是要预测的数据多, 影响因素多且复杂<sup>[1-3]</sup>。日负荷主要影响因素有日类型、月份、气象因素(最高温度、最低温度)等。针对气象影响因素, 不同文献采取不同的处理策略。文献[4]将气象因素直接输入线性回归模型建模。文献[5-6]将

气象因素归一化后输入神经网络建模。文献[7]将气象因素模糊化处理后利用多元线性回归分析建模。文献[8]将气象因素模糊化处理后利用神经网络建模。文献[9]将温度敏感负荷与温度不敏感负荷分离, 并用负荷温度梯度来表达温度敏感负荷与温度的变化规律, 提出引入负荷温度梯度的负荷预测方法。长期来看, 由于一个地区的温度变化范围相对

固定, 比如一个地区每年最高温度都在  $35^{\circ}\text{C} \sim 37^{\circ}\text{C}$  之间, 而负荷的上升趋势幅度大, 因此温度对负荷的影响更多体现在短期的累积效应上。文献[10]指出, 由于超短期负荷预测的预测周期比较短, 而气象变化是缓慢的, 当以负荷曲线作为负荷预测的资料时, 天气的影响实际上就已包含在负荷曲线中了, 因此考虑气象的累积效应的短期影响, 应基于前一天的负荷和气象, 考虑气象变化和负荷变化之间关系, 这是本文采取的一种新的气象因素处理策略。

## 1 日负荷预测算法

考虑气象因素累积效应的电力系统 96 点日负荷预测的方法步骤如下。

### 1.1 相似性分析

96 点日负荷是指从零时开始每 15min 计量该时刻的负荷值, 形成每日 96 维负荷向量  $f_h$ , 即  $f_h = (f_1, f_2, \dots, f_{96})$ 。取两日负荷向量

$$f_{h1} = (f_{1,1}, f_{1,2}, \dots, f_{1,96})$$

$$f_{h2} = (f_{2,1}, f_{2,2}, \dots, f_{2,96})$$

其线性相关系数定义为

$$\rho(f_{h1}, f_{h2}) = \frac{\text{cov}(f_{h1}, f_{h2})}{\sqrt{D(f_{h1})} \cdot \sqrt{D(f_{h2})}}$$

式中, cov 为协方差, D 为方差。取东莞地区 2016 年 7 月 2 日至 2017 年 7 月 24 日的 96 点日负荷数据, 计算两日负荷向量的线性相关系数, 然后取平均值, 结果见表 1。

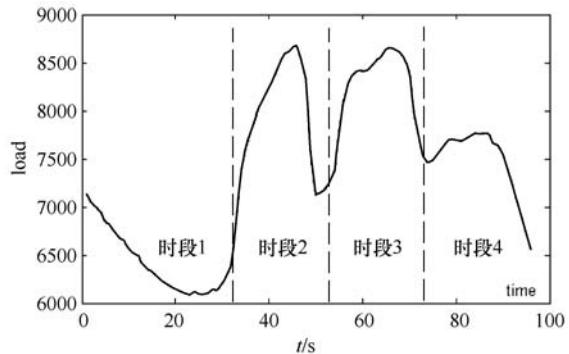
表 1 两日负荷向量的相关系数

两日相隔天数	样本数量	平均相关系数	排名	备注
0	360	0.9451	1	相邻两天
1	359	0.9389	3	
2	358	0.9363	4	
3	357	0.9350	5	
4	356	0.9348	6	
5	355	0.9329	7	
6	354	0.9396	2	隔一周相同日类型
7	353	0.9276	10	
8	352	0.9263	13	
9	351	0.9266	12	
10	350	0.9272	11	
11	349	0.9278	9	
12	348	0.9230	15	
13	347	0.9285	8	隔二周相同日类型
14	346	0.9232	14	

由表 1 可以看出, 相邻两日的相关系数最大。该结果表明, 相邻两日负荷向量具有很高的线性相关性, 因此可以利用前一日的负荷预测当日的负荷。

### 1.2 多时段划分

不同时段对外部因素(日类型、月份、气象)的敏感度不一样, 为提升预测精度, 分时段进行预测。根据东莞地区日负荷曲线特性, 如图 1 所示, 一般划为以下 4 个时段。



时段 1: 下半夜 (00:00—08:00, 32 点)

时段 2: 上午 (08:00—13:00, 20 点)

时段 3: 下午 (13:00—18:00, 20 点)

时段 4: 上半夜 (18:00—00:00, 24 点)

图 1 多时段划分

### 1.3 日特征量量化

研究表明, 电力系统负荷受到日类型、月份、气象等影响。这些当日影响因素称为日特征量。这些日特征量要输入模型, 需要进行量化处理。设  $w = (w_1, w_2, w_3, w_4, w_5, w_6, w_7)$  表示日类型向量, 日类型为星期一时,  $w_1 = 1$ , 其余分量为 0; 日类型为星期二时,  $w_2 = 1$ , 其余分量为 0; 以此类推。设  $y = (y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8, y_9, y_{10}, y_{11}, y_{12})$  表示月类型向量, 月类型为 1 月份时,  $y_1 = 1$ , 其余分量为 0; 月类型为 2 月份时,  $y_2 = 1$ , 其余分量为 0; 以此类推。

### 1.4 基于气象因素累积效应的多元线性回归模型

#### 1) 符号

表 2 符号意义

变量	含义	备注
$f_h(0)$	预测日的负荷向量	96 点
$f_{h1}(0)$	预测日时段 1 的负荷向量	32 点
$f_{h2}(0)$	预测日时段 2 的负荷向量	20 点
$f_{h3}(0)$	预测日时段 3 的负荷向量	20 点
$f_{h4}(0)$	预测日时段 4 的负荷向量	24 点
$f_{h1}(-1)$	预测日前 1 天时段 1 的负荷向量	32 点

(续)

变量	含义	备注
$mf_{h1}(0)$	预测日时段 1 点负荷的平均值	
$mf_{h2}(0)$	预测日时段 2 点负荷的平均值	
$mf_{h3}(0)$	预测日时段 3 点负荷的平均值	
$mf_{h4}(0)$	预测日时段 4 点负荷的平均值	
$mf_{h1}(-1)$	预测日前 1 天时段 1 点负荷的平均值	
$\max t(0)$	预测日最高温度	
$\min t(0)$	预测日最低温度	
$\max t(-1)$	预测日前 1 天最高温度	
$\min t(-1)$	预测日前 1 天最低温度	
$\min t(-2)$	预测日前 2 天最低温度	
$rf_{h1}(0)$	$\frac{mf_{h1}(0)}{mf_{h1}(-1)}$	预测日与前 1 天的平均负荷比
$r\max t(0)$	$\frac{\max t(0)}{\max t(-1)}$	预测日与前 1 天的最高温度比
$r\min t(0)$	$\frac{\min t(0)}{\min t(-1)}$	预测日与前 1 天的最低温度比
$r\max t(-1)$	$\frac{\max t(-1)}{\max t(-2)}$	预测日前 1 天与前 2 天的最高温度比
$r\min t(-1)$	$\frac{\min t(-1)}{\min t(-2)}$	预测日前 1 天与前 2 天的最低温度比

## 2) 模型

以预测日的  $w$  取值、预测日的  $y$  取值、 $r\max t(0)$ 、 $r\min t(0)$ 、 $r\max t(-1)$ 、 $r\min t(-1)$  为输入变量，分别以  $mf_{h1}(0)$ 、 $mf_{h2}(0)$ 、 $mf_{h3}(0)$ 、 $mf_{h4}(0)$  为输出变量，建立以下多元线性回归模型。

$$\begin{aligned} & rf_{h1}(0), rf_{h2}(0), rf_{h3}(0), rf_{h4}(0) = \\ & b_{11} + \sum_{i=1}^7 b_{2i} w_i + \sum_{i=1}^{12} b_{3i} y_i + \\ & b_{41} r \max t(0) + b_{42} r \min t(0) + b_{43} r \max t(-1) + b_{44} r \min t(-1) \end{aligned} \quad (1)$$

其中

$$\mathbf{b} = (b_{11}, b_{21}, \dots, b_{27}, b_{31}, \dots, b_{312}, b_{41}, b_{42}, b_{43}, b_{44})$$

为待定系数向量，通过多元线性回归分析分别确定时段 1、时段 2、时段 3、时段 4 模型的系数。

## 1.5 日负荷预测

模型 (1) 建立后，分别预测时段 1、时段 2、时段 3、时段 4 的点负荷值。

$$f_{h1}(0) = f_{h1}(-1) * rf_{h1}(0)$$

$$f_{h2}(0) = f_{h2}(-1) * rf_{h2}(0)$$

$$f_{h3}(0) = f_{h3}(-1) * rf_{h3}(0)$$

$$f_{h4}(0) = f_{h4}(-1) * rf_{h4}(0)$$

96 点日负荷最终预测结果为

$$\mathbf{f}_h(0) = [f_{h1}(0), f_{h2}(0), f_{h3}(0), f_{h4}(0)]$$

以上述预测结果为基础，利用文献[11]的算法对其他点进行预测，得出整日负荷曲线。

## 1.6 算法验证

算法准确性的影响因素有两个：①多元线性回归模型的准确性；②相邻两日的相关系数。多元线性回归模型越准确，相邻两日的相关系数越接近 1，算法准确性越高。设时刻  $i$  ( $i=1, 2, \dots, 96$ ) 负荷  $f_i$  的预测值为  $f_i^c$ ，则时刻  $i$  负荷预测准确率定义为

$$r_i = \left( 1 - \left| \frac{f_i - f_i^c}{f_i} \right| \right) \times 100\%$$

则时段 1、时段 2、时段 3、时段 4 的负荷预测准确率依次为

$$\begin{aligned} R_1 &= \frac{\sum_{i=1}^{32} r_i}{32}, \quad R_2 = \frac{\sum_{i=33}^{52} r_i}{20} \\ R_3 &= \frac{\sum_{i=53}^{72} r_i}{20}, \quad R_4 = \frac{\sum_{i=73}^{96} r_i}{24} \end{aligned}$$

预测日负荷预测准确率为

$$R = \frac{32R_1 + 20R_2 + 20R_3 + 24R_4}{96}$$

## 2 案例分析

### 2.1 负荷数据

96 点负荷数据来源于东莞供电局负荷预测系统，数据范围从 2016 年 7 月 3 日至 2017 年 7 月 31 日。

建模数据范围：2016 年 7 月 3 日—2016 年 7 月 24 日。

预测数据范围：2016 年 7 月 25 日—2016 年 7 月 31 日，共 7 天，该数据不参与建模。

### 2.2 外部数据

从天气网 ([www.tianqi.com](http://www.tianqi.com)) 获取东莞地区每日最高温度、最低温度等气象数据。

### 2.3 数据处理

通过对变量进行线性相关分析，发现时段 1 的平均负荷比率与  $w_4, w_5, w_6, w_7$  相关性不大，因此建立时段 1 的多元线性回归模型的数据结构为

$[rf_{h1}(0); w_1, w_2, w_3, y_1, y_2, \dots, y_{12}, rmaxt(0), rmint(0), rmaxt(-1), rmint(-1)]$

式中,  $rf_{h1}(0)$ 为模型输出变量, 其余变量为模型输入变量。数据集示例见表 3。

建立时段 2、时段 3、时段 4 的多元线性回归模型的数据结构为 (以时段 2 为例)

表 3 时段 1 建模数据结构

日期	$rf_{h1}(0)$	$w_1$	$w_2$	$w_3$	$y$	$rmaxt(0)$	$rmint(0)$	$rmaxt(-1)$	$rmint(-1)$
2017-7-24	0.79457	1	0	0	$y_7=1$	1.06897	1.00000	0.93545	1.00000
2017-7-23	0.99689	0	0	0	$y_7=1$	0.93549	1.00000	0.96875	1.00000
2017-7-22	1.04043	0	0	0	$y_7=1$	0.96875	1.00000	0.96970	1.00000
2017-7-21	1.06495	0	0	0	$y_7=1$	0.96970	1.00000	1.06452	0.96154

表 4 多元线性回归模型系数<sup>b</sup>

模型	1	$w_1$	$w_2$	$w_3$	$w_4$	$w_5$	$w_6$	$w_7$	$y_1$	$y_2$	$y_3$	$y_4$
时段 1	0.87139	-0.05570	0.16898	0.02356	—	—	—	—	-0.03128	0.05133	0	0.00420
时段 2	0.84023	0.29544	0.06702	0.04369	0.04697	0.05097	0	-0.13413	-0.01830	0.06384	0.01473	0.01394
时段 3	0.80486	0.32628	0.06430	0.04440	0.04309	0.05649	0	-0.13471	-0.01764	0.06153	0.01358	0.01069
时段 4	0.86325	0.30649	0.08357	0.05264	0.05235	0.0566	0	-0.08516	-0.02206	0.04788	0.01016	0.00644
模型	$y_5$	$y_6$	$y_7$	$y_8$	$y_9$	$y_{10}$	$y_{11}$	$y_{12}$	$rmaxt(0)$	$rmint(0)$	$rmaxt(-1)$	$rmint(-1)$
时段 1	-0.00080	0.03295	-0.00240	0.00150	0.00438	-0.01790	-0.00070	0.00702	0.05598	-0.09097	0.21828	-0.07423
时段 2	-0.01480	0.06085	0.01474	0.01636	0.01818	0	0.01131	-0.00273	0.18080	-0.06011	0.08312	-0.10121
时段 3	-0.01770	0.05140	0.01342	0.01337	0.01492	0	0.00850	-0.00715	0.25416	-0.07064	0.05402	-0.10046
时段 4	-0.01540	0.03925	0.00696	0.00401	0.00309	0	0.00609	-0.00865	0.24042	-0.08234	0.01405	-0.10068

通过表 4 得出东莞地区日负荷特性如下。

结论 1: 相邻日的负荷变化与星期六这个日类型无关。

结论 2: 时段 2—4 的相邻日的负荷变化与十月份这个月类型无关。

结论 3: 时段 1—3 的相邻日负荷变化对预测日最高温度变化的敏感度比预测日最低温度变化、前天最高温度变化、前天最低温度变化要大得多。

$[rf_{h2}(0); w_1, w_2, \dots, w_7, y_1, y_2, \dots, y_{12}, rmaxt(0), rmint(0), rmaxt(-1), rmint(-1)]$

式中,  $rf_{h2}(0)$ 为模型输出变量, 其余变量为模型输入变量。

## 2.4 多元线性回归模型建立

利用 Matlab 求出各个多元线性回归模型系数如下。

表 3 时段 1 建模数据结构

结论 4: 时段 1—3 的相邻日负荷变化对前天最低温度变化的敏感度比预测日最低温度变化要大得多。

结论 5: 时段 1 的相邻日负荷变化对前天最高温度变化的敏感度比预测日最高温度变化、预测日最低温度变化、前天最低温度变化要大得多。

## 2.5 预测结果

预测数据范围为 2016 年 7 月 25 日—2016 年 7 月 31 日, 共 7 天, 该数据不参与建模。其预测准确率如下。

表 5 7 天预测准确率

预测日	时段 1	时段 2	时段 3	时段 4	全天平均值
2017-7-31	0.974705851	0.952268752	0.967961331	0.932788052	0.9581469
2017-7-30	0.959567476	0.967031015	0.972127911	0.953416606	0.96220142
2017-7-29	0.915602088	0.981694889	0.975315177	0.978104134	0.95743716
2017-7-28	0.956009600	0.995945083	0.987699037	0.948512096	0.96905708
2017-7-27	0.993165224	0.981728786	0.983591618	0.984323538	0.98657771
2017-7-26	0.952564369	0.988946512	0.983765379	0.981885428	0.97397446
2017-7-25	0.943086918	0.983869033	0.990683729	0.986792498	0.97242559
7 天平均值	0.956385932	0.978783439	0.980163455	0.96654605	0.96854576

东莞供电局负荷预测系统日负荷预测的基本算法见表 6, 这些算法不考虑日类型、月类型和气象因素。

表 6 东莞供电局负荷预测系统日负荷预测基本算法

序号	基本算法
1	点对点法
2	均荷倍比平滑法
3	谷荷倍比平滑法
4	峰荷倍比平滑法
5	均荷重叠曲线法
6	谷荷重叠曲线法
7	峰荷重叠曲线法
8	变化系数法
9	一元线性回归
10	新息点对点
11	均荷新息倍比平滑法
12	谷荷新息倍比平滑法
13	峰荷新息倍比平滑法
14	均荷新息重叠曲线法
15	谷荷新息重叠曲线法
16	峰荷新息重叠曲线法
17	新息变化系数法
18	低频平均频域分量法
19	低频平滑频域分量法
20	低频相关频域分量法
21	余量相关频域分量法
22	低频平均新息频域分量法
23	低频平滑新息频域分量法
24	低频相关新息频域分量法
25	余量相关新息频域分量法

东莞供电局负荷预测系统基于上述的基本算法采取“自适应训练”的预测策略, 即所有基本算法都对预测日的前一天进行日负荷预测, 选择准确率最高的基本算法用于预测日的日负荷预测。该系统的预测结果见表 7。

表 7 东莞供电局负荷预测系统 7 天预测准确率

预测日	全天平均值
2017-7-31	0.8798
2017-7-30	0.8831
2017-7-29	0.8532
2017-7-28	0.8897
2017-7-27	0.9043
2017-7-26	0.8952
2017-7-25	0.9816
7 天平均值	0.8981

从表 7 中可以看出, 本文算法在稳定性和准确性方面比东莞供电局负荷预测系统要好。

### 3 结论

首先通过线性相关性分析, 得出相邻两日的负荷向量线性相关系数最大, 因此利用过去一天的负荷预测当天的负荷提供了理论保证。其次是过去一天的气象因素累积效应已体现在过去一天的负荷曲线上, 因此当天的气象因素累积效应就是气象因素的变化对负荷变化产生的作用。另外, 日负荷的不同时段对气象因素累积效应、日类型、月类型的敏感度不一样, 因此考虑分时段建模。最后分时段建立了时段平均负荷变化与预测日的日类型、月份、最高温度变化、最低温度变化和前一日的最高温度变化、最低温度变化的多元线性回归关系, 提升了负荷预测的稳定性和准确性。

### 参考文献

- [1] 谢宏, 陈志业, 牛东晓, 等. 基于小波分解与气象因素影响的电力系统日负荷预测模型研究[J]. 中国电机工程学报, 2001, 21(5): 5-10.
- [2] 陈亚, 李萍. 基于神经网络的短期电力负荷预测仿真研究[J]. 电气技术, 2017, 18(1): 26-29.
- [3] 高赐威, 李倩玉, 苏卫华, 等. 短期负荷预测中考虑积温效应的温度修正模型研究[J]. 电工技术学报, 2015, 30(4): 242-248.
- [4] 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.
- [5] 张颖, 张瑞, 高中文. 基于相似日的神经网络短期负荷预测方法[J]. 黑龙江大学自然科学学报, 2003, 20(2): 59-61, 65.
- [6] 黎祚, 周步祥, 林楠. 基于模糊聚类与改进 BP 算法的日负荷特性曲线分类与短期负荷预测[J]. 电力系统保护与控制, 2012, 40(3): 56-60.
- [7] 李滨, 覃芳璐, 吴茵, 等. 基于模糊信息粒化与多策略灵敏度的短期日负荷曲线预测[J]. 电工技术学报, 2017, 32(9): 149-159.
- [8] 陈根永, 史敬天, 毛晓波, 等. 考虑温度积累效应的短期电力负荷预测方法研究[J]. 电力系统保护与控制, 2009, 37(16): 24-28.
- [9] 程卓. 引入负荷温度梯度的负荷预测方法研究[J]. 电气技术, 2015, 16(10): 13-16, 44.
- [10] 蔡佳宏, 刘俊勇. 超短期负荷预测中相似日的选择方法[J]. 华北电力大学学报, 2006, 33(1): 38-41.
- [11] Han-Ju Li. A data forward stepwise fitting algorithm based on orthogonal function system[J]. ITM web of conferences 12, 04013(2017).

收稿日期: 2017-10-18

### 作者简介

李汉巨 (1979-), 男, 广东省湛江市人, 硕士, 高级工程师, 主要从事数据挖掘和负荷预测工作。