

文章编号:1672-3961(2011)02-0114-05

# 基于电力需求响应的公共建筑基线负荷预测

马庆,李歧强\*

(山东大学控制科学与工程学院, 山东 济南 250061)

**摘要:**针对公共建筑基线负荷难以有效预测的问题,提出了一种基于模糊C-均值聚类预处理的人工神经网络预测方法。采用聚类算法,将大量的复杂历史数据集划分成多个群体的混合,每个群体对应单独的预测模型进行预测。该方法减少了培训数据,克服了标准方法数据量大和处理速度慢的缺点。将预测结果与标准的人工神经网络方法相比较,得到了较高的预测精度,能有效预测公共建筑基线负荷。

**关键词:**基线负荷预测;需求响应;模糊C-均值聚类;人工神经网络

中图分类号:TP715 文献标志码:A

## Public buildings baseline load forecasting based on demand response in electric power

MA Qing, LI Qi-qiang\*

(School of Control Science and Engineering, Shandong University, Jinan 250061, China)

**Abstract:** According to the problem that public buildings baseline load was hard to forecast effectively, a kind of artificial neural network forecasting method based on Fuzzy C-Means (FCM) optimization preprocesses was proposed. The complicated historical data were divided into a mixture of multiple populations, and each population was represented by a single forecasting model. This method could reduce ANN training data and overcame the disadvantage of very large data and slow processing speed. The forecasting result can have greater forecasting accuracy and effectively forecast the public buildings baseline load compared with the standard BP neural network.

**Key words:** baseline load forecasting; demand response; FCM algorithm; artificial neural network

## 0 引言

准确的建筑基线负荷预测可为电力市场决策提供指导,为定量评价各种电力需求响应程序对建筑负荷的减少程度提供依据。因建筑特性、客户特征、缩减策略以及需求响应事件特征的多样性,所以对于电力需求响应下建筑负荷减少的计算不是一件容易的事<sup>[1]</sup>。在电力需求响应事件驱动下,建筑设备运行人员会采用各种调节手段去削减电力峰值负荷。为确定实际峰值负荷减少的多少,需要对没有需求响应作用下的负荷进行估计,这就涉及到建筑

基线负荷,它是准确评价某种需求响应程序对负荷影响程度的关键,尤其是对那些直接从负荷减少获利的需求响应程序<sup>[2]</sup>。

长期以来,国内外学者对电力系统负荷预测和建筑空调系统负荷预测的理论和方法做了大量的研究,提出了各种各样的预测方法<sup>[3-8]</sup>。但由于国内启动电力需求响应程序比较晚,覆盖面比较小,因此相关文献比较少。国外的研究相对较多<sup>[9-18]</sup>,研究方法大致可分为两大类:一类是平均法,即采用需求响应事件前几天的小时负荷值进行线性拟合。这类方法仅对历史负荷数据进行统计、分析及运算,而没有考虑气象信息对基线负荷的影响。另一类是确定

收稿日期:2010-11-10

基金项目:山东省科学技术发展计划资助项目(2007GG20007004);山东省信息产业发展专项资助项目(2007R00065)

作者简介:马庆(1971-),男,山东海阳人,讲师,博士研究生,主要研究方向为电力需求响应与建筑节能。E-mail: maqing@sdu.edu.cn

\*通讯作者:李歧强(1964-),男,山东临沂人,教授,博士,博士生导师,主要研究方向为复杂系统优化调度的建模和算法。E-mail: qqli@sdu.edu.cn

天气模型的方法,即采用确定的公式,用当地小时温度去预测负荷<sup>[2]</sup>。这类方法虽然考虑了气象因素,但是一般不涉及气象以外的因素。这两类方法主要是采用了线性模型,比较成熟,算法简单,速度快。但在处理非线性问题时就显得无能为力,尤其对于日负荷差异较大的建筑,上述方法得到的预测结果不令人满意<sup>[1]</sup>。

本研究采用模糊 C-均值聚类方法,将大量复杂历史数据集划分成多个群体的混合;综合考虑当地天气和建筑特征等因素,利用人工神经网络对参与电力需求响应程序的公共建筑进行基线负荷预测,取得了较高的预测精度。

## 1 模糊 C-均值聚类算法

FCM 是目前最常用的聚类算法之一,算法描述如下:

假设  $X = \{x_i, i = 1, 2, \dots, n\} \subseteq R^s$  是  $s$  维向量空间的一个特征向量集,根据某种相似性度量,该集合被聚合成  $c$  个子集。 $v_i (i = 1, 2, \dots, c) \in R^s$  是第  $i$  个聚类的中心, $u_{ik} (i = 1, 2, \dots, c; k = 1, 2, \dots, n)$  是第  $k$  个样本对第  $i$  类的隶属度函数,且  $0 \leq u_{ik} \leq 1$  及  $0 < \sum_{k=1}^n u_{ik} < n$ , FCM 的目标函数表示为

$$J_m(U, v) = \sum_{k=1}^n \sum_{i=1}^c (u_{ik})^m \|x_k - v_i\|^2, \quad (1)$$

式中  $U = [u_{ik}]$  为模糊分类矩阵;  $\|x_k - v_i\|^2$  为第  $k$  个数据点与第  $i$  个聚类中心间的欧几里德距离;  $m \in [1, \infty]$  为影响隶属度矩阵模糊化程度的权重指数,其约束为

$$\sum_{i=1}^c u_{ik} = 1, \quad \forall k = 1, 2, \dots, n. \quad (2)$$

在约束式(2)下优化式(1)得:

$$u_{ik} = \frac{(1/\|x_k - v_i\|^2)^{1/(m-1)}}{\sum_{j=1}^c (1/\|x_k - v_j\|^2)^{1/(m-1)}}, \quad \forall i = 1, 2, \dots, c; k = 1, 2, \dots, n, \quad (3)$$

$$v_i = \frac{\sum_{k=1}^n u_{ik}^m x_k}{\sum_{k=1}^n u_{ik}^m}, \quad \forall i = 1, 2, \dots, c. \quad (4)$$

FCM 算法目标是使各个分类中的样本到聚类中心的加权距离平方和达到最小。在给出初始方案后,通过式(3)、(4)反复迭代,使得目标函数(1)达到极小<sup>[19]</sup>。

## 2 人工神经网络的 BP 算法

神经网络是由大量的、简单的处理单元(或称

神经元)互相连接而形成的复杂网络,它的特性和结构是由神经元的特性和它们的互连方式决定的。BP 网络是一种多层前馈型网络,它包含输入层、输出层和中间层,各层之间实现全连接。中间层有单层或多层,由于它们与外界没有直接连接,故也称为隐含层。BP 神经网络的输入层和隐含层神经元采用 Sigmoid 函数;输出层采用 Sigmoid 函数或线性函数,采用纠错规则进行学习。

BP 学习算法分两步进行,即正向传播和反向传播。正向传播即输入的样本从输入层经隐含单元一层一层处理后,传向输出层。在输出层把实际输出与期望输出比较后,如果两者不等,则进行反向传播。反向传播是把误差信号按原来正向传播的通路反向传回,并对每个隐含层的各个神经元的权系数进行修正,最后使得全局误差趋于极小值。相应计算公式为:

隐层节点的输出为:

$$y_h = f\left(\sum_{i=1}^{N_1} \omega_{ih} x_i + \theta_h\right),$$

其中  $y_h$  为隐层节点的输出;  $x_i$  为输入信号;  $\omega_{ih}$  为输入节点到隐层节点的连接权值;  $\theta_h$  为隐层节点的阈值。

输出层节点的输出为:

$$Z_j = f\left(\sum_{h=1}^{N_2} \omega_{hj} y_h + \gamma_j\right).$$

其中  $Z_j$  为输出节点的输出;  $\omega_{hj}$  为隐层节点到输出层的连接权值;  $\gamma_j$  为输出节点的阈值。

误差反向传播为:

$$E_k = \frac{1}{2} \sum_j (y_{jk} - o_{jk})^2.$$

其中  $E_k$  为第  $k$  个表征向量的误差;  $y_{jk}$  为第  $j$  个输出神经元的期望值;  $o_{jk}$  为第  $j$  个输出神经元的实际值。

## 3 基于 FCM 和 BP 神经网络的建筑基线负荷预测

本次预测结合济南历史室外小时干球温度和某办公建筑的实测历史小时负荷数据,将 FCM 分析和 BP 神经网络相结合,进行建筑基线负荷预测。为说明使用 FCM 和 BP 神经网络结合预测建筑基线负荷的有效性,本研究做了与标准 BP 神经网络性能的对比。

### 3.1 样本输入参数的选择

合理选择预测模型输入参数,既满足一定预测精度,又使得模型简约。此外,还要考虑实际工程的特点,参数多可能会使预测精度提高,但会给建模带

来困难和增加工程费用。

建筑物内负荷可以看成是建筑物系统在外扰和内扰综合作用的结果。外扰主要是指室外气象条件,在输入参数选择中主要考虑需求响应预测日前4 d的室外小时干球温度。内扰主要是指建筑物内人员、灯光及设备发热引起的负荷变化,在输入参数选择中主要考虑需求响应预测日前4 d的小时负荷。

对于建筑基线负荷预测,影响预测的预测因子形成模型输入参数样本。综上所述,输入参数样本值由需求响应事件天前4 d的小时负荷和室外小时干球温度8个特性指标构成。

### 3.2 训练样本输入参数的预处理

为了保证每个训练样本输入参数的权重相等,应该采取数值转换的方法对模型参数进行归一化处理,即把参数处理成0~1之间的数值。具体处理方法如下:

$$X_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}$$

其中  $X_i$  为归一化的输入值;  $x_i$  为原始输入值;  $x_{\min}$  为原始输入值的最小值;  $x_{\max}$  为原始输入值的最大值。

### 3.3 训练样本输入参数的FCM聚类

给定聚类数  $c$ 、权值数  $m$  及终止条件,选择初始的聚类中心,用FCM将训练样本输入参数进行模糊聚类,得到每个类的聚类中心以及每个特征向量对于每个类的隶属度。分别对这  $c$  个训练样本集用EP神经网络进行训练,得到  $c$  个BP神经网络预测模型。

### 3.4 判别分析

FCM聚类的目的是把训练样本分成若干类,相当于建立若干类判别函数,然后对待预测样本进行类别判断。在判别分析中,通常采用最大隶属度原则和择近原则。择近原则的定义为:

给定论域  $U$  上的模糊子集(模糊模式)  $A_1, A_2, \dots, A_n$  及另一个模糊子集(模糊对象)  $B$ , 若  $0 \leq i \leq n$ , 使  $N(A, B) = \max_{1 \leq j \leq n} N(B, A_j)$ , 则认为与  $A_i$  最贴近,  $B$  应规划为模式  $A_i$ , 这就叫做最大贴近度原则或择近原则。常用的贴近度有格贴近度和距离贴近度,本文采用距离贴近度,其计算公式为:

$$N(A, B) = 1 - \frac{1}{n} \sum_{k=1}^n |A(x_k) - B(x_k)|$$

贴近度  $N(A, B)$  越大,表明两个模糊子集的贴近度越好,越相似,反之则越差。因此,先是通过FCM聚类算法对已有的训练样本进行聚类得到各自的聚类中心,然后计算预测样本与各聚类中心的贴近度,取最大的贴近度值作为预测样本的分类依据<sup>[20]</sup>。

### 3.5 建筑基线负荷预测流程

根据上述分析,建筑基线负荷预测流程步骤如下:

(1) 首先对室外小时干球温度和建筑小时负荷数据进行归一化处理,形成训练样本和预测样本;

(2) 确定聚类类别数  $c$ 、权值数  $m$  以及终止条件,初始化模糊聚类矩阵  $U$ , 用式(3)、(4)反复迭代,得到满足精度要求的每个类的聚类中心以及每个特征向量对于每个类的隶属度;

(3) 得到  $c$  个训练样本子集,对样本子集都建立一个BP神经网络预测模型;

(4) 利用建立的判别函数来判别预测样本的类别,划分到已经分好的类中;

(5) 用每一类对应的BP神经网络预测模型来预测建筑基线负荷值,并输出结果。

### 3.6 算例分析

本次预测使用了济南市2005年6-8月12:00-16:00的室外小时干球温度和当地一座办公建筑的实测历史小时负荷数据,共有460组数据进行聚类分析,再用50组数据进行预测。部分数据见表1。将460组训练数据进行FCM聚类,得到聚类中心及其建筑基线负荷值如表2所示。对待预测的50组样本分别进行判别分析,可判别出每个样本的所属类别,再用该类别建立的BP神经网络预测模型进行预测,得到的结果如表3所示。将FCM的BP神经网络与标准的神经网络预测的标准偏差、期望偏差百分数与偏差系数相比较,结果如表4所示。预测时刻前4 d的历史室外干球温度分别用  $T_1, T_2, T_3, T_4$  表示,单位是摄氏度;预测时刻负荷及前4 d的历史负荷分别用  $K, K_1, K_2, K_3, K_4$  表示,单位是kW。

表1 预测时刻前4 d的室外干球温度和负荷及预测时刻负荷

Table 1 Outdoor dry-bulb temperatures and loads of the first four days and loads at forecasting time

样本	$T_1/$ ℃	$T_2/$ ℃	$T_3/$ ℃	$T_4/$ ℃	$K_1/$ kW	$K_2/$ kW	$K_3/$ kW	$K_4/$ kW	$K/$ kW
1	31.6	32.9	30.8	30.2	2 970	3 110	2 380	2 190	3 210
2	33.4	33.3	31.2	30.7	3 010	3 200	2 190	2 270	3 250
3	33.9	33.5	31.4	31.2	2 950	3 120	2 220	2 360	3 270
4	34.1	33.7	31.4	31.5	3 050	3 240	2 290	2 320	3 340
5	33.8	34.1	31.7	32.1	3 050	3 260	2 370	2 490	3 770
506	31.1	25.3	29.1	30.5	3 720	3 740	3 030	1 860	3 770
507	31.6	24.5	29.7	31.7	3 980	3 820	3 230	2 220	3 890
508	31.8	23.8	30.0	32.4	4 020	3 870	2 790	2 030	4 090
509	31.7	23.3	30.1	32.4	4 400	3 910	3 330	2 090	4 320
510	31.7	23.6	30.1	31.6	3 900	4 290	3 140	2 230	4 360

表2 聚类中心  
Table 2 Clustering center

$v_i$	$T_1/$ ℃	$T_2/$ ℃	$T_3/$ ℃	$T_4/$ ℃	$K_1/$ kW	$K_2/$ kW	$K_3/$ kW	$K_4/$ kW
$v_1$	28.5	30.2	23.1	26.1	3 800	2 390	2 720	3 450
$v_2$	35.5	34.1	36.5	38.0	3 610	3 160	1 930	2 160
$v_3$	31.1	22.3	25.2	31.4	4 060	4 070	3 340	3 390
$v_4$	27.8	28.8	30.7	30.2	3 110	3 190	3 440	3 000
$v_5$	30.8	32.4	31.7	33.2	3 070	4 170	4 150	4 400
$v_6$	31.0	31.8	24.6	25.4	3 500	4 000	2 910	2 880
$v_7$	34.2	34.3	32.4	34.0	4 400	4 090	3 270	3 350
$v_8$	25.5	27.3	23.8	26.3	2 980	2 280	2 990	2 020

表3 FCM-ANN 和 ANN 的预测结果  
Table 3 The forecasting result of FCM-ANN and ANN

$T_1/$ ℃	$T_2/$ ℃	$T_3/$ ℃	$T_4/$ ℃	$K_1/$ kW	$K_2/$ kW	$K_3/$ kW	$K_4/$ kW	$K/$ kW	FCM- ANN 预测 预测值	ANN 预测 预测值
32.6	33.3	30.4	31.5	2 970	3 200	2 220	2 360	3 120	3 126	3 129
33.4	33.3	31.4	31.2	3 110	3 090	3 020	2 360	3 250	3 245	3 258
27.8	28.8	31.7	33.2	2 510	3 190	3 570	3 340	3 240	3 243	3 230
31.8	23.8	30.0	32.4	4 020	3 870	2 700	2 030	4 090	4 092	4 098
31.7	23.3	30.1	32.4	4 400	3 910	3 330	2 090	4 320	4 328	4 330
31.7	23.6	30.1	31.6	3 900	4 290	3 140	2 230	4 360	4 365	4 366

表4 FCM-ANN 和标准 ANN 的预测误差结果比较  
Table 4 The forecasting error result comparison of FCM-ANN and standard ANN

预测算法	$\delta$	EEP/%	CV/%
FCM-ANN	38.6	3.28	4.32
标准 ANN	62.4	5.31	6.88

注: $\delta$ 为标准偏差;EEP为期望偏差百分数;CV为偏差系数百分数。

从表4中可以看到,在50组预测样本的预测结果中,FCM-ANN模型的预测精度均优于标准的ANN未聚类的预测值,从而证明了FCM-ANN模型的有效性。

## 4 结语

提出了一种基于FCM聚类算法的人工神经网络的公共建筑基线负荷预测方法。该系统采用聚类的方法,将大量的复杂历史数据集划分成多个群体的混合,每个群体对应单独的预测模型进行预测。采用该方法,不仅可以减少人工神经网络的培训数据,而且可以克服因构建人工神经网络模型时,所需大量数据和处理速度慢的缺点。将该系统应用到基线负荷预测中,其预测值与未聚类的标准人工神经网络预测值相比,比较结果说明本文所提出的预测方法能更精确的预测建筑基线负荷。

## 参考文献:

[1] COUGHLIN K, PIETTE M A, GOLDMAN C, et al. Estimating demand response load impact: evaluation of baseline load models for non-residential buildings in[R]. California: LBNL, Tech Rep, 2008:1-28.

[2] COUGHLIN K, PIETTE M A, GOLDMAN C, et al. Statistical analysis of baseline load models for non-residential buildings [J]. Building and Environment, 2009, 41: 374-381.

[3] 崔航,王冕,罗贵明,等.基于扩展时序距离法和自适应控制的超短期负荷预测[J].电力系统自动化,2009,33(15):38-42.

CUI Hang, WANG Mian, LUO Guiming, et al. Ultra-short term load forecasting based on adaptive control and extended time-series distance method [J]. Automation of Electric Power Systems, 2009, 33(15):38-42.

[4] 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.

LI Canbing, LI Xiaohui, ZHAO Rui, et al. A novel algorithm of selecting similar days for short-term power load forecasting [J]. Automation of Electric Power Systems, 2008, 32(9):69-73.

[5] 畅广辉,刘涤尘,熊浩.基于多分辨率SVM回归估计的短期负荷预测[J].电力系统自动化,2007,31(9):37-41.

CHANG Guanghui, LIU Dichen, XIONG Hao. Short term load forecasting based on multi-resolution SVM regression [J]. Automation of Electric Power Systems, 2007, 31(9):37-41.

[6] 何大四,张旭.改进的季节性指数平滑法预测空调负荷分析[J].同济大学学报:自然科学版,2005,33(12):1672-1676.

HE Dasi, ZHANG Xu. Analysis of air conditioning load prediction by modified seasonal exponential smoothing model [J]. Journal of Tongji University: Natural Science, 2005, 33(12):1672-1676.

[7] 万迪光.不确定输入对电力负荷预测的影响[J].山东大学学报:工学版,2004,34(5):55-58.

WAN Diguang. Effect of uncertain inputs on electric load forecasting [J]. Journal of Shandong University: Engineering Science, 2004, 34(5):55-58.

[8] 高荣,刘晓华.基于小波变换的支持向量机短期负荷预测[J].山东大学学报:工学版,2005,35(3):115-118.

GAO Rong, LIU Xiaohua. Short-term load forecasting method based on support vector machine combined with wavelet transform [J]. Journal of Shandong University: Engineering Science, 2005, 35(3):115-118.

[9] INKINEN S J, NEUVO Y. Base line estimation using

- sparse nonlinear operators [J]. *Signal Processing*, 2002(82): 811-817.
- [10] BUKVIC S, SPASOJEVIC D. An alternative approach to spectrum base line estimation [J]. *Spectrochimica Acta; Part B*, 2005(60): 1308-1315.
- [11] REZA D, SAEED A. Demand response regulation modeling based on distribution system asset efficiency [J]. *Electric Power Systems Research*, 2011(81): 667-676.
- [12] WANG Jianhui, CARY N B, HU Zhaoguang, et al. Demand Response in China [J]. *Energy*, 2010(35): 1592-1597.
- [13] SEZGENA O, GOLDMANA C A, KRISHNARAOB P. Option value of electricity demand response [J]. *Energy*, 2007(32): 108-119.
- [14] CAPPERS P, GOLDMAN C, KATHAN D. Demand response in U. S. electricity markets: empirical evidence [J]. *Energy*, 2010(35): 1526-1535.
- [15] ALBADI M H, EISAADANY E F. A summary of demand response in electricity markets [J]. *Electric Power Systems Research*, 2008(78): 1989-1996.
- [16] KIM J H, SHCHERBAKOVA A. Common failures of demand response [J]. *Energy*, 2011(36): 873-880.
- [17] HSU C I, WEN Y H. Reliability evaluation for airline network design in response to fluctuation in passenger demand [J]. *Omega*, 2002(30): 197-213.
- [18] AAZAMI R, AFLAKI K, HAGHIFAM M R. A demand response based solution for LMP management in power markets [J]. *Electrical Power and Energy Systems*, 2011(11): 1-8.
- [19] 徐巍, 陈祥光, 彭红星, 等. 基于模糊 C-均值聚类与支持向量机的 PMV 指标预测系统 [J]. *系统工程理论与实践*, 2009, 29(7): 119-124.
- XU Wei, CHEN Xiangguang, PENG Hongxing, et al. PMV index forecasting system based on fuzzy C-means clustering and support vector machine [J]. *System Engineering-Theory & Practice*, 2009, 29(7): 119-124.
- [20] 王旭平, 王汉功, 陈小虎. 基于模糊 C 均值算法的齿轮泵故障诊断技术 [J]. *机床与液压*, 2009, 37(10): 269-272.
- WANG Xuping, WANG Hangong, CHEN Xiaohu. Diagnosis technics of gear pump fault based on the Fuzzy C-means clustering algorithm [J]. *Machine Tool & Hydraulics*, 2009, 37(10): 269-272.

(编辑:陈斌)