

基于季节性时间序列模型的空调负荷预测

孙 靖 程大章

(同济大学社区信息化与智能建筑研究中心 上海 200092)

摘要 基于空调负荷预测的优化控制是解决冰蓄冷控制问题的理想途径。本文在分析国际建筑物空调负荷预测竞赛研究成果的基础上指出,利用季节性时间序列模型建模预测精度较高,且工程实施简便,特别适合于空调系统连续运行、负荷波动规律性较强的建筑物负荷预测。本文概要介绍利用季节性时间序列模型进行建模预测的理论和方法,并通过工程实例验证了建模方法的有效性。

关键词: 冰蓄冷 时间序列 负荷预测

中图分类号: TU831.2; TP18

Air-conditioning Load Prediction Based on Seasonal Time Series Methods

Sun Jing Cheng Dazhang

(Tongji University Shanghai 200092 China)

Abstract The optimal controlling method based on air-conditioning load prediction is an effective way to solve the problems of ITS (Ice Thermal Storage) system. On the basis of achievements of the two international load prediction tests, time series methods are useful in predicting the hourly load of the target building due to its convenience and high prediction accuracy characteristics. Then the theory and method to establish the time series model were discussed. At last, the validity of method mentioned above was verified by a project.

Keywords: Ice thermal storage system, time series, load prediction

1 引言

冰蓄冷技术作为一项“移峰填谷”、“平衡电网负荷”的技术,自 20 世纪 90 年代进入中国以来,一直受到能源部门的高度重视。冰蓄冷技术的社会意义是毋庸置疑的,但对于用户而言,更为关心的则是这一技术可以带来的经济效益。众所周知,冰蓄冷系统并不是一个节能系统,而只是起到负荷转移的功能,以平衡城市电网负荷。对于用户,则是充分利用当地的峰谷电价差,达到节约空调冷源系统能源费用的目的。如何合理地控制蓄冰设备在谷段电价时间的蓄冰量,以及在峰段和平段电价时间的融冰速度是冰蓄冷系统日运行费用的决定性因素。冰蓄冷系统的控制是涉及多学科的综合性问题,

由于实际空调负荷的不确定性,给控制策略的设计带来一定困难。经过多年的探索与研究,工程界普遍认为,基于空调负荷预测的冰蓄冷优化控制是解决冰蓄冷控制问题的理想途径。

国际上曾举行过两次大规模建筑物空调负荷预测竞赛,分别由美国供热、制冷和空调工程师协会 (ASHRAE) 及日本空气调和卫生工学会 (SHASE) 先后组织。竞赛的目的在于通过比较不同研究者利用的数学模型(对同一栋建筑物进行负荷预测),来确定各种负荷预测模型的准确性和适用程度。国内虽然尚无类似活动,但也有不少研究者从事建筑物能耗和负荷预测的研究工作。目前主要应用于负荷预测的技术有回归分析、时间序列分析及近年来发展起来的基于人工神经网络的负荷预测方法等。其中,时间序列分析和基于人工神经网络的负荷预测方法比较适合于建筑物空调负荷预测。

尽管国内外许多学者对建筑物空调负荷预测技术开展了大量的研究,但目前国内冰蓄冷工程中具有完善空调负荷预测模块的控制系统为数不多,大多数系统只是通过简单的类比前一天或前一年类似工作日的实测空调负荷状况对未来负荷进行预测,有的甚至由技术工人按经验人工控制。建筑物空调负荷预测技术在理论上和实践上存在极大的脱节,主要是由于部分负荷预测方法过于复杂,不适应于工程实际的需求。本文通过分析国际大规模建筑物空调负荷预测竞赛研究成果,认为利用季节性时间序列模型进行建模预测精度较高,且工程实施简便,特别适合于空调系统连续运行、负荷波动规律性较强的建筑物负荷预测,并利用季节性时间序列模型进行建模预测,通过工程实例验证了建模方法的有效性。

2 时间序列分析法在建筑物空调负荷预测中的适用性^[1,4,5]

建筑物空调负荷预测技术的理论和实践脱节原因既有工程条件的限制,也由于有些研究在进行理论工作时未能充分地考虑到工程实际中存在的各种问题,设计的负荷预测模型过于复杂或不够稳定。

事实上,对于应用于实际工程的模型,稳定性是其最主要的性能指标,同时,数据的存储量以及相应的处理量也直接关系到模型的适用范围。理论上具有较高精度的预测模型在实际工程中由于各种误差的累积,其预测的准确度往往受到较大影响。这一点在日本空气调和卫生工学会(SHASE)组织的负荷预测竞赛(以下简称 BM Test)的研究成果中可以明显地发现。

根据 BM Test 的研究结论,在理论上,人工神经网络方法具有最高的精度,而时间序列模型则以输入参数及数据处理量少,模型收敛迅速而独具优势。而实际工程应用中,考虑人工神经网络模型各种输入参数的误差(如未来时刻室内外气象等参数在实际应用中也是未知的,需要通过预测获得,存在误差),在空调负荷规律性较强(主要是指不存在双休日空调系统停开情况)的建筑物中,时间序列模型的预测精度往往可以接近甚至达到人工神经网络模型的预测精度。

由此可见,时间序列分析法由于数据存储和处理量小,收敛速度快,实施简便、可靠而成为最适合冰蓄冷工程应用的预测方法,适用于空调系统连续运行、空调负荷规律性较强的建筑物(如医院、

宾馆、机场车站及住宅区等)空调负荷预测。

3 时间序列模型的基本形式及季节性时间序列模型^[2,3]

时间序列是按照时间顺序取得的一系列观测值,其一个本质特征就是相邻观测值之间的依赖性。时间序列分析所论及的就是对这种依赖性进行分析的技巧。

时间序列是一个时间参数离散的随机序列,若参数 t 表示时间,则随机序列 $\{X_t, t=0, \pm 1, \pm 2, \dots\}$ 就是一个时间序列。时间序列分析的任务就是根据这组随机序列的部分观测值来推断此序列总体的性质,建立随机动态模型(统计模型),并将这种模型用于各种应用领域,如预报等。

常用的时间序列分析方法分为平稳时间序列分析和非平稳时间序列分析两大类。平稳时间序列模型包括 AR 模型(自回归模型)、MA 模型(滑动平均模型)、ARMA 模型(自回归滑动平均模型)三类;非平稳时间序列模型主要包括 ARIMA 模型(求和自回归滑动平均模型)、Seasonal 模型(季节模型)两类。以上五种模型中,ARIMA 模型可以看成是时间序列模型的标准形式,其余模型均可视为 ARIMA 模型的特例或组合形式。其中 $AR(p)$ 、 $MA(q)$ 和 $ARMA(p, q)$ 模型仅适用于具有零均值的平稳序列建模和预测;ARIMA 模型适合于非零均值非平稳的时间序列;季节模型适合于具有周期特征的时间序列的分析与建模。

在冰蓄冷系统优化控制中,一般需要提前 24h 预测出下一工作日的逐时空调负荷,由于建筑物逐时空调负荷具有明显的周期特征,因此考虑采用季节模型进行分析建模。

对于季节模型,可以用一类疏系数 ARIMA 模型去描述:设 s 为一正整数,一个时间序列 $\{X_t, t=0, \pm 1, \pm 2, \dots\}$, 如果满足下列模型

$$F(B^s)\nabla_s^D X_t = Q(B^s)E_t$$

则称 X_t 是周期为 s 的季节性序列。

式中 X_t ——时间序列值

B ——推移算子

∇ ——差分算子

$$F(B^s) = 1 - F_1 B^s - F_2 B^{2s} - \dots - F_p B^{Ps}$$

$$Q(B^s) = 1 - Q_1 B^s - Q_2 B^{2s} - \dots - Q_q B^{Qs}$$

而 E_t 一般不必是白噪声,可设它是另一个 $ARIMA(p, d, q)$ 序列。

$$\mathbf{j}(B)\nabla^d \mathbf{E}_t = \mathbf{q}(B)\mathbf{e}_t$$

式中 \mathbf{e}_t ——平稳白噪声

B ——推移算子

∇ ——差分算子

$$\mathbf{j}(B) = 1 - \mathbf{j}_1 B - \mathbf{j}_2 B^2 - \dots - \mathbf{j}_p B^p$$

$$\mathbf{q}(B) = 1 - \mathbf{q}_1 B - \mathbf{q}_2 B^2 - \dots - \mathbf{q}_q B^q$$

若令 $W_t = \nabla^d \nabla_s^D X_t$ ，则

$$\mathbf{j}(B)\mathbf{F}(B^s)W_t = \mathbf{q}(B)\mathbf{Q}(B^s)\mathbf{e}_t \quad (1)$$

式(1)称为乘积型季节性模型，其阶次常用 $(p, d, q) \times (P, D, Q)_s$ 表示。

4 建筑物空调负荷预测模型的建立^[6,7]

建立具体的建筑物空调负荷预测模型首先要根据历史数据的特点进行模式识别，确定采用哪种时间序列模型进行负荷预测建模；其次按照一定的准则确定模型的具体阶次并进行参数估计；然后对建立的模型进行考核，确认模型的有效性；最后将模型应用于负荷预测。

以下仅对各步骤的实现基本思路进行介绍，具体公式推导可参阅参考文献[3]，计算功能可由 SAS、SPSS 等专业统计软件实现。

4.1 模型的模式识别

模型的模式识别主要是通过对空调负荷历史数据进行各种处理后，作图观察完成的。

(1) 首先做历史数据的时序图，观察是否具有明显的周期性，若其周期为 s ，则进行季节差分，否则转下一步。

(2) 消除数据的周期性后，计算其自相关函数和偏相关函数。如果是截尾的或拖尾的（即被负指数控制的），说明序列已平稳下来，可以确定其模型类型；若自相关函数与偏相关函数中至少有一个既不是截尾的，也不是拖尾的，说明处理后的样本仍不平稳，可进行差分处理，然后再求其自相关函数和偏相关函数，再进行判别，直到序列平稳下来（一般差分次数少于两次）。

4.2 模型的定阶与参数估计

理论上，时间序列模型的建模应该先定阶，然后进行参数估计，但在实际应用中，由于模型定阶除了试算外没有更好的方法，所以往往是对多种模型结构分别进行参数估计，然后对建立的模型用统一的指标去衡量，以确定其中的最优模型。

(1) 模型的参数估计方法很多，包括矩估计、条件最小二乘估计、无条件最小二乘估计、最大似然估计等。其中，最大似然估计的估计效果较好，且以此为基础的模型定阶方法成熟可靠。但由于这种方法计算复杂、计算量大，在以往的应用中受到限制。随着计算机技术的发展，许多统计应用软件（如 SAS、SPSS 等）都具有相应的计算模块，几分钟内就能完成对成千上万数据的统计、计算。

(2) 对于试算模型的评判，可采用由日本统计学家 Akaike 提出的 AIC 准则。AIC 准则起源于 Kullback-Lelibler 信息量，是信息论与统计学的重要研究成果。运用 SAS 软件进行参数估计时，其输出结果中直接就包含了该模型的 AIC 值，因此只需比较试算各模型的 AIC 值，选取其中最小的作为预测模型。

4.3 模型的考核

在定阶和参数估计以后，还需对建立的模型进行考核，以检验模型的有效性。其基本做法是检验模型误差 e_t 是否为白噪声，若检验认为 e_t 是白噪声，则建模获得通过，否则要重新进行定阶与参数估计。考核某序列是否为白噪声的方法很多，采用 c^2 检验法、D 检验法、W 检验法等均可。

4.4 模型的还原与预测

由于 SAS 软件输出的实际上是平稳化处理后的序列的预测模型，因此要获得未来时刻的空调负荷预报，还需对预测值进行还原。还原后的预测模型即可应用于空调负荷预测。

5 建筑物空调负荷预测算例

下面以上海某医学中心为例，介绍模型的具体构建方法。在建模过程中，利用医学中心 2001 年的空调负荷实测数据进行建模，然后对 2002 年的空调负荷进行预测，与实测数据比较，以评价模型的有效性。

5.1 数据的平稳化处理

首先判别医学中心的空调负荷实测值是否具有周期性。选取 2001 年 8 月 1 日至 2001 年 8 月 7 日一周的逐时负荷实测数据作为样本，绘制时序图，如图 1 所示。

由图 1 可以清楚地看出，医学中心的空调负荷实测值具有明显的周期性，其周期为 24h，即以一天为周期波动。因此，需进行周期为 24 的季节差分。并对差分后的序列再作时序图。由图 2 可见，经季节差分后的序列基本消除了周期性。

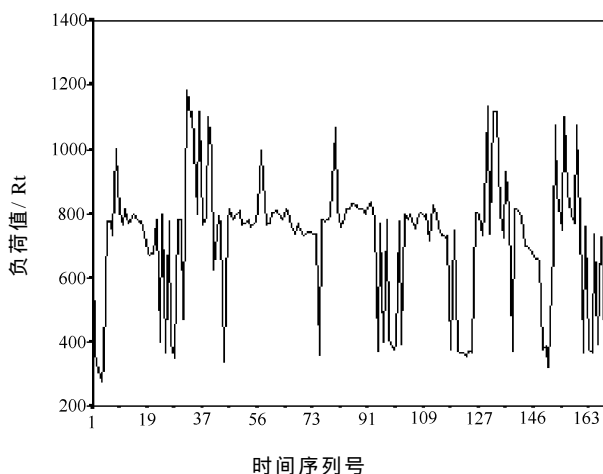


图 1 空调负荷时序图

Fig.1 Sequence chart of the air-conditioning load

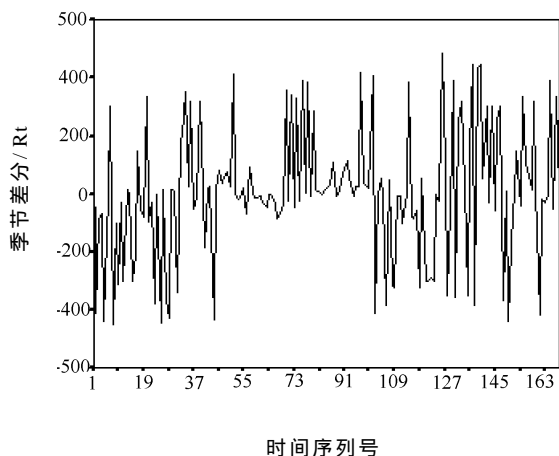


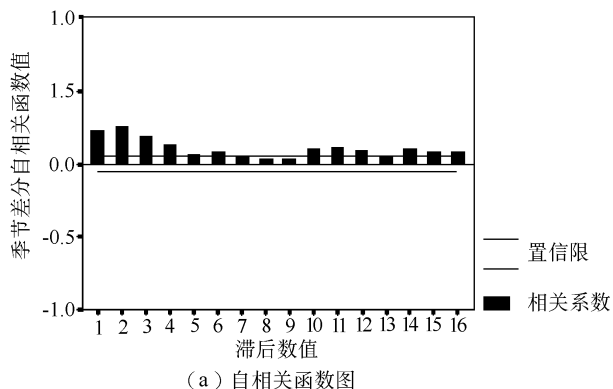
图 2 空调负荷经季节差分后时序图

Fig.2 The sequence chart after seasonal difference

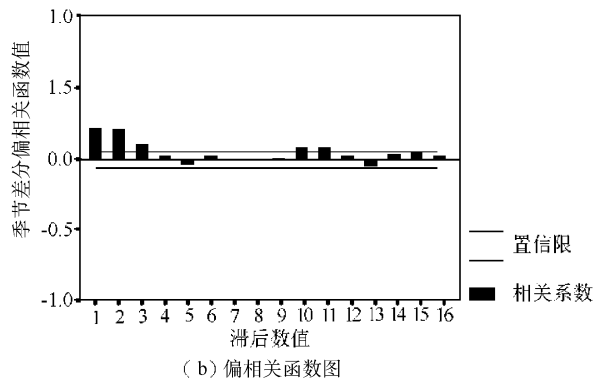
下一步通过自相关函数和偏相关函数判别序列的稳定性。选取 2001 年 7 月 1 日至 2001 年 8 月 31 日两个月的实测数据作为样本，对经过周期为 24 的季节差分后的数据做其自相关函数及偏相关函数图。

由图 3 可见，尽管其自相关函数收敛较慢，但已基本趋于平稳。再进行一次差分进行比较。

由图 4 可见，经两次差分后的序列，其自相关函数基本上是一步截尾的，而偏相关函数呈拖尾趋势(即被负指数控制)，可以判定此时序列已属于平稳序列。



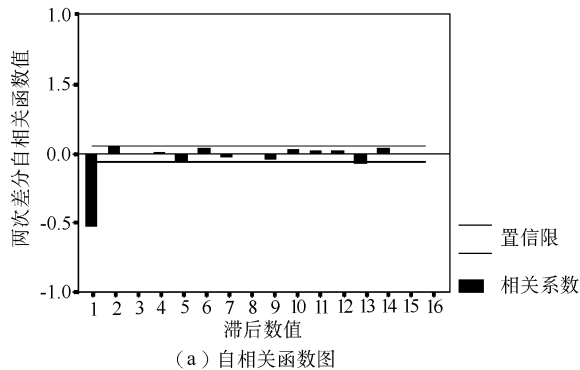
(a) 自相关函数图



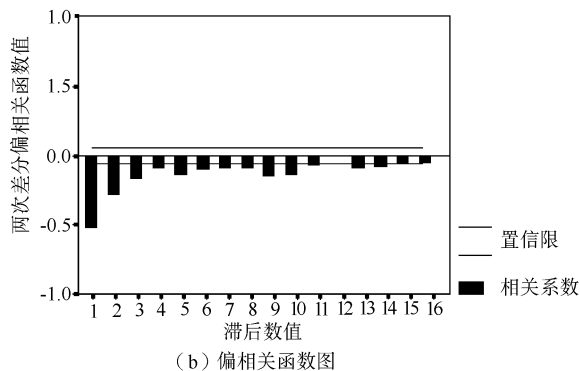
(b) 偏相关函数图

图 3 空调负荷 (季节差分后) 相关函数图

Fig.3 The autocorrelations function chart after seasonal difference



(a) 自相关函数图



(b) 偏相关函数图

图 4 空调负荷 (两次差分后) 相关函数图

Fig.4 The autocorrelations function chart after second difference

以上分析说明,该医学中心空调逐时负荷具有明显的周期性,但最多经两次差分处理可转化成平稳序列,因此可考虑使用单纯的季节模型或乘积型季节模型进行建模。

5.2 模型的定阶及参数估计

模型的定阶及参数估计,可调用 SAS 软件中 PROC ARIMA 过程的 ESTIMATE 命令,利用 AIC 准则进行模型的定阶及参数估计。

在定阶过程中选取了多种 p,q 的各种阶数形式进行试算,用 AIC 准则寻求最优模型。

以下列出由 PROC ARIMA 过程计算得到的一些比较优秀的模型结构及 AIC 值:

- (1) $q=(1,24)$, AIC=19950.13 ;
- (2) $q=(24)$, $p=(24)$, AIC=20018.47 ;
- (3) $q=(1,24)$, $p=(1)$, AIC=19580.88 ;
- (4) $q=(1,2)(24,48)$, $p=(1)(24)$, AIC=19370.01 ;
- (5) $q=(1,2,3)(24)$, $p=(1,2,3)(24)$,
AIC=19352.36 ;
- (6) $q=(1,2,3,4)(24)$, $p=(1)$, AIC=19348.87 ;
- (7) $q=(1,2,3,4)(24)$, $p=(1,2,3,4)(24)$,
AIC=19449.41。

由 AIC 定阶准则,取 $q=(1,2,3,4)(24)$, $p=(1)$ 模型,即认为经两次差分后的序列 W_t 服从下列模型

$$(1-j_1B)W_t = (1-q_1B-q_2B^2-q_3B^3-q_4B^4)(1-Q_1B^{24})e_t \quad (2)$$

同时,获得 $j_1, q_1, q_2, q_3, q_4, Q_1$ 的最大似然估计是: $\hat{j}_1=0.22202$; $\hat{q}_1=0.98882$; $\hat{q}_2=0.23698$; $\hat{q}_3=0.04696$; $\hat{q}_4=0.08584$; $\hat{Q}_1=0.92588$ 。

5.3 模型的考核检验

利用 c^2 检验法对医学中心 2001 年 7 月 1 日至 8 月 31 日的预测模型的残差进行检验,检验结果如表 1 所示。

表 1 预测模型残差 c^2 检验表

Tab.1 c^2 verification								
m	6	12	18	24	30	36	42	48
p 值	0.998	0.582	0.407	0.091	0.086	0.058	0.085	0.116

取显著水平 $\alpha=0.05$, 各 p 值均大于 0.05, 故接受 H_0 , 认为 e_t 是白噪声, 模型考核通过。

5.4 模型的预测及评价

利用 SAS 软件中 PROC ARIMA 过程的 FORCAST 命令,对医学中心 2002 年 7 月 1 日至 8

月 31 日的空调负荷做提前 24h 的预测,并通过与实测值的比较评价模型预测的优劣。

图 5 为医学中心 2002 年 7 月某周的空调负荷时序图。从图中可以看出,负荷模型的预测值虽然不能全部准确地显示空调负荷的实际值,但基本描绘出了空调负荷变化的趋势,对冰蓄冷系统的优化控制具有指导意义。

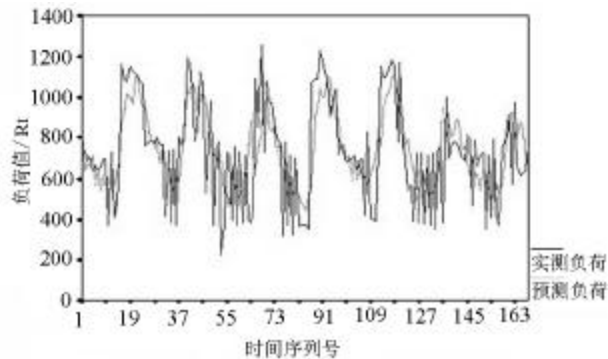


图 5 空调负荷时序比较图

Fig.5 Sequence comparison between the predicted load and the actual load

定义评价指标——预测负荷逐日累计均方误差

$$r_{\text{day}} = \frac{\sqrt{\frac{1}{N} \sum_{k=1}^N (\tilde{Q}_k - Q_k)^2}}{Q_{\text{max}}} \quad (3)$$

式中 Q_k, \tilde{Q}_k ——实测和预测的日累积负荷值
 Q_{max} ——实测日累积负荷最大值
 N ——整个预测期间的天数

计算结果该模型的逐日累计均方误差为 0.043。

图 6 为医学中心 2002 年 7 月 1 日至 8 月 31 日的空调负荷散点图,同时,计算出预测值与实测值之间的相关系数 $R=0.759$ 。

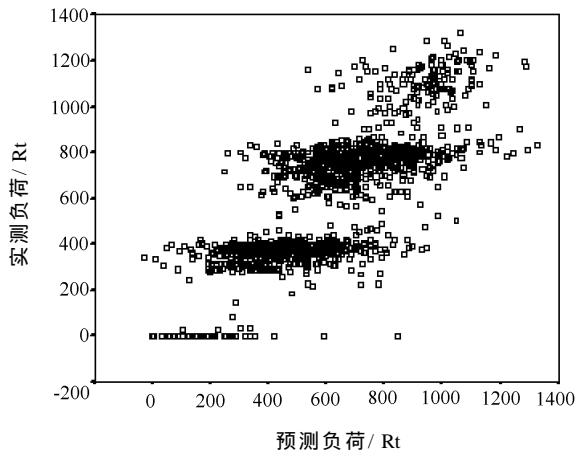


图 6 空调负荷散点比较图

Fig.6 Scatterplots of the air-conditioning load

一般工程应用中认为,预测与实际值之间的相关系数达到 0.75 以上即为有效预测,对控制策略具有指导意义;相关系数达到 0.85 以上即认为预测具有相当高的精度。而目前工程中常用前一天实测负荷通过比例处理获得预测负荷的方法。其预测值与实测负荷的相关系数只能达到 0.35~0.40 左右,这样的预测精度在工程上的意义已经不大了。

有研究者曾利用人工神经网络方法对此医学中心进行过负荷预测,其预测值与实测值的相关系数为 0.733。虽然所采用的数据不同,但说明人工神经网络方法也无法使预测精度获得明显提高,且输入参数和实现过程都较时间序列模型要复杂。

6 模型的应用

通过模型获得次日逐时空调负荷的预测值后便可应用于冰蓄冷系统优化控制。冰蓄冷系统优化控制的核心就是以各时段的预测负荷为基础,合理安排和分配峰段及平段电价时间内制冷机组直接供冷和蓄冷装置融冰供冷之间的比例,使二者能最经济地满足空调负荷的需求。因此,负荷预测精度对系统的控制效果有很大影响,同时,控制效果也是检验模型有效性的最终标准。

这三项指标均为评价冰蓄冷系统控制策略优劣的常用标准,其中日运行费用是冰蓄冷系统实施优化控制策略的主要目的;减少制冷机组启停次数可以延长机组的使用寿命;剩余冰量是冰蓄冷系统负荷转移能力发挥程度的衡量指标,并影响下一工作周期的充冷效率。

将第 5 节中的建立的模型应用于医学中心冰蓄冷系统控制,设计基于此模型的优化控制策略,并与原控制方案进行比较,结果显示,改进后的控制策略可节约日运行费用 5%~10%,平均减少制冷机组启停次数 1.637 次/天,平均剩余冰量减少 93Rt,各项指标均得到较大的改善,控制效果令人满意。

7 结论

基于空调负荷预报的冰蓄冷优化控制是解决冰

蓄冷系统优化控制问题的有效途径。

季节性时间序列模型在理论上可以达到较高的预测精度,且工程实施简便,对数据的存储和处理要求低,收敛迅速,特别适合于空调系统连续运行、空调负荷规律性较强的建筑物(如医院、宾馆、机场车站及住宅区等)空调负荷预测。

利用专业统计软件 SAS、SPSS 等可以方便地建立建筑物空调负荷预测模型,其预测精度较高,预测结果对冰蓄冷系统优化控制具有较强的指导作用。

参考文献

- 1 陈沛霖,曹叔维,郭建雄. 空气调节负荷计算理论及方法. 上海: 同济大学出版社, 1987
- 2 George E.P.Box, Gwilym M.Jenkins, Gregory C.Reinsel 著, 顾岚主译. 时间序列分析预测与控制. 北京: 中国统计出版社, 1997
- 3 范金城, 梅长林. 数据分析. 北京: 科学出版社, 2002
- 4 Kawashima M. Hourly thermal load prediction for the next 24 hours by ARIMA, LR and an artificial neural network. ASHRAE Trans, 1995
- 5 Harbel J S, Thamilsaran S. Predicting hourly building energy use: the great energy predictor shootout :measuring retrofit savings-overview and discussion of results. ASHRAE Trans, 1996
- 6 卢纹岱. 统计分析软件 SPSS for Windows 从入门到精通. 北京: 电子工业出版社, 1997
- 7 高惠璇等编译. SAS 系统 SAS/ETS 软件使用手册. 北京: 中国统计出版社, 1997

作者简介

孙 靖 男, 1977 年生, 博士研究生, 主要从事智能建筑、信息技术及管理领域相关课题的研究。

程大章 男, 1951 年生, 教授, 主要从事智能建筑、城市信息化工程方向的教学与研究。