第五届泰迪杯
数据挖掘挑战赛

优秀作品

“泰迪杯”优秀论文

作品名称：中央空调系统的数据分析与控制策略

荣获奖项：特等奖

作品单位：山东科技大学

作品成员：唐静 曲文涛 张潇

指导老师：王新赠

中央空调系统的数据分析与控制策略

摘 要

随着经济社会的发展，中央空调的使用在大中型城市中越来越普遍，随之而来的是中央空调的智能控制与节能问题。我们从数学模型角度考虑系统动态实时的控制，并采用动态分析方法，对系统进行优化控制。

针对问题一，我们首先对数据进行预处理，利用三均值对数据的变化特征进行初步判断；然后，根据干球温度、湿球温度和湿度等变量的周期性变化特点，我们每隔七天选取一天的数据进行分析。再者，通过多元回归分析，得到冷却负载、系统效率、耗电量和各个可控变量与不可控变量之间的多元回归方程。最后，随机选取了部分数据，计算得到模型的均方误差小于 5 %，因此，模型具有较高的可靠性。

针对问题二，在时间、室外温度和湿度、冷却负载和设备状态已知的前提下，我们以各个设备的转速为决策变量，以中央空调系统效率和系统总耗电量为目标函数，以设备转速取值范围、设备状态参数取值、热稳定性范围为约束条件，建立多目标非线性规划模型。然后以 12 月 28 日的数据为基础，通过模拟退火算法求解最优解，并进行对比分析，得出总耗电量平均下降了 15.29%，系统效率平均下降了 19.73%。

针对问题三，在模型二的基础上，我们以各个设备的开关状态为决策变量，以中央空调系统效率和系统总耗电量为目标函数，建立多目标非线性规划模型。之后，我们仍以 12 月 28 日数据为基础进行求解，经过对比分析得到总耗电量平均下降了 18.34%，系统效率平均下降了 18.98%。

针对问题四，综合上述模型，我们以各个设备的状态和转速为决策变量，并加入系统设备开关次数和时间间隔的约束条件，完善优化模型。然后，我们使用最速下降法改进了模拟退火算法，大大提高了模型的求解效率。之后，我们随机选取了附件所给的数据进行优化，经过对比分析，得出总耗电量和系统效率显著地下降。

关键词：中央空调；多元回归分析；多目变非线性规划；模拟退火算法

泰迪杯优秀论文

Data analysis and control strategy of central air conditioning

Abstract

With the development of economy and society, the use of central air-conditioning is becoming more and more popular in large and medium-sized cities. In this paper, we consider the real-time control system from the mathematical model, and use the dynamic analysis method to optimize the system.

Firstly, we preprocess the data, and use the three means to judge the change of the data; then, according to the characteristics of dry bulb temperature, wet bulb temperature and humidity, we select the data every seven days. Furthermore, we can get the multiple regression equation among of the cooling load, the system efficiency, the total power consumption and the controllable variables and uncontrollable variables, by multiple regression analysis. At last, we randomly select some data, and calculate the mean square error of the model which is less than 5%.Therefore, the model has high reliability.

Secondly, in the case of time, outdoor temperature and humidity, cooling load and equipment state are known, we use the speed of each equipment as the decision variable, the central air conditioning system efficiency and the total system power consumption as the objective function, speed, equipment state parameters, thermal stability as constraints, to establish the multi-objective nonlinear programming model. Then, we based on the data of December 28 to get the optimal solution by simulated annealing algorithm, and the total power consumption is reduced by 15.29% and the system efficiency is decreased by 19.73% on average.

Thirdly, on the basis of model two, we build the multi-objective nonlinear programming model using the switching status of each device as the decision variable, and the central air conditioning system efficiency and the total system power consumption as the objective function. After that, we still solve the data on the basis of December 28, and compare and analysis, found that the total power consumption decreased by an average of 18.34%, the system efficiency decreased by an average of 18.98%.

Fourthly, based on the above model, we use the status and speed of each device as the decision variable, and add the constraint condition of the system state variable parameter to perfect the optimization model. In addition, we improved the simulated annealing algorithm by using the steepest descent method, which greatly improved the efficiency of the model. After that, we randomly selected the data in the annex to optimize, after comparative analysis, the total power consumption and system efficiency significantly decreased.

Keywords： Central Air-conditioning， Multiple Regression Analysis， Multi Variable Nonlinear Programming， Simulated Annealing Algorithm

泰迪杯优秀论文

目 录

1. 挖掘目标 1
1.1 挖掘背景 1
1.2 挖掘目标 1
1.3 研究现状 1
2. 分析方法与过程 1
2.1 流程分析 2
2.1.1 总体流程 2
2.1.2 变量分析 3
2.1.3 中央空调结构分析 4
2.2 具体步骤 5
2.2.1 数据预处理 6
2.2.2 符号说明 7
2.2.3 数据规律分析 7
2.2.4 基于回归分析的相关关系挖掘 11
2.2.5 当决策变量为设备转速时的优化模型 21
2.2.6 当决策变量为所有设备状态变量时的优化模型 25
2.6.7 当决策变量为所有可控变量时的优化模型 28
3. 结论 30
4. 参考文献 32

“泰迪杯”优秀论文

1. 挖掘目标

1.1 挖掘背景

随着我国社会经济的不断发展，城市进程不断加快，大量的现代化办公楼、商场、宾馆等大型建筑相继建成，这些建筑里面都越来越倾向于中央空调系统来实现室内温度和湿度的调节。尤其是近年来，随着“智慧城市”建设步伐的快速推进，如何围绕智慧城市建设实现中央空调的智能控制与节能，成为智慧城市建设中的重要课题之一。中央空调的优化控制策略研究对能源的节约有重要意义。

1.2 挖掘目标

根据题目中给出的中央空调的结构示意图和数据说明，在城市常年平均温度为 25 至 32 摄氏度之间，平均湿度为 85%左右的前提之下，解决问题。

首先，对数据进行深入的分析梳理，挖掘出数据变化的特征和规律，研究冷却负载、系统效率、耗电量与可控变量和不可控变量之间的关系模型，并检验模型的可靠性。其次，根据给出的时间、室外温度和湿度、冷却负载和设备状态信息，尝试给出通过条件可控变量的最优控制策略，并给出相应的系统总耗电量和系统效率。然后，根据给出的时间、室外温度和湿度、冷却负载和可控变量取值，试给出所有设备状态变量的最优控制策略，和相应的系统总耗电量和系统效率。最后，根据附件中给出的信息并结合对系统设备状态变量的要去，求解所有可控变量的最优控制策略。以及相应的系统总耗电量和系统效率。

1.3 研究现状

1987 年，Braun[1]通过对飞机场的中央空调系统进行优化控制研究，提出了基于部件和基于系统的两种模型，取得了较好的优化效果。

1993 年，MacArthur[2]对动态预测性模型制冷系统优化控制进行了一系列的研究，开发了多变量控制预测器，研究结果表明 EHC 在动态预测模型中的效果较好。

2004 年，kWork TaiChan[3]对空调系统负荷动态变化的节能问题进行了研究，利用 TRNSYS 空调系统模拟软件对冷水机建模分析，获得全年冷量分布。

2. 分析方法与过程

这里，我们首先对问题进行了分析，基于对问题的分析我们对数据进行预处理和降维，挖掘出各个变量的关联关系；又在不同的约束条件下建立出三个中央空调控制策略的优化模型，对其验证求解，并分析了模型验证结果。

泰迪杯优秀论文

2.1 流程分析

2.1.1 总体流程

这里，我们通过对问题进行详细的分析，得到如下总体处理流程。

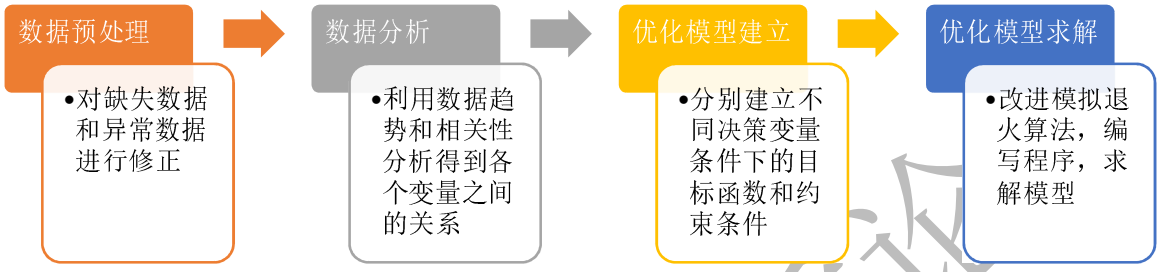


图 1 总体处理流程图

在这里，对优化模型建立过程，具体分析如下：

1) 当决策变量为冷水泵转速、冷凝转速和冷却塔风扇转速时的优化模型分析

这里，题目中要求在时间、室外温度和湿度、冷却负载和设备状态信息已知的条件下，通过调节冷水泵转速、 冷凝水泵转速和冷却塔风扇转速来获得最优控制策略。

本问题中，首先提出最优控制策略的两条原则：

（1）系统总耗电量应尽量小；

（2）系统效率应在一定范围内尽量小。

我们以上述两个原则作为目标函数建立双目标线性规划。对于约束条件，考虑不同设备状态的约束、转速范围约束、系统效率范围约束以及冷却负载等函数约束。在模型求解过程中，由于无法对多目标线性规划进行求解，因此，我们将双目标线性规划转化为单目标线性规划进行求解。

2) 当决策变量为所有设备状态变量时的优化模型分析

与第一个优化模型相比较而言，第一个优化模型是已知设备状态优化设备转速，而问题 3 中的优化模型则是已知设备转速优化设备状态。问题 2 与问题 3 的目标函数相同，但决策变量发生了变化。求解方法可以继续沿用问题 2 的方法。

3) 当决策变量为所有可控变量时的优化模型

问题 4 的优化问题，相比 2、3 问中已知信息减少，决策变量变成所有可控变量，

泰迪杯，优秀论

包括各个设备的状态和各个设备的转速。在问题 3 的基础上，目标函数没有发生变化，但是需要增加约束条件，来对问题的解空间缩减。这里，根据题目要求，增加约束条件如下：

- ✧ 每台设备的开/关以后两个小时内不可以关/开；
- ✧ 每台设备每台最多只能开关 6 次；
- ✧ 每台设备不能超负荷运转；
- ✧ 任何时间至少要开启一台设备；

将上述约束条件，抽象成数学模型，得到新的优化模型。

2.1.2 变量分析

附件中的数据共包含 88840 条记录，每个记录有 51 个字段，这些字段包括采集信息的时间、环境信息、系统状态参数、控制参数、系统采集信息和系统运行相关信息。这里，我们再次明确附件中 51 个字段的意义。

1) **采集时间**：采集时间是从 2016 年 10 月 4 日到 2016 年 12 月 29 日约三个月的数据，共有 74 天的数据。除去 2016 年 10 月 4 日和 2016 年 12 月 29 日的数据外，其它数据都是从每天的零点开始，每隔一分钟或者两分钟取一次数据，最长间隔不超过 5 分钟，数据非常的详尽。

2) **环境信息：**这里的环境信息包含干球室外温度、湿球室外温度以及相对湿度。对几个环境信息解释如下表所示。

表格 1 环境变量的解释

环境变量	解释说明
干球温度	是空气的真实温度，可以直接用普通的温度计测出，即我们通常意义中的温度。
湿球温度	湿球温度是空气中水蒸气达到饱和时的空气温度，通俗来说，湿球温度是当前环境中仅通过蒸发水分所能达到的最低温度。 湿球温度是用湿润的纱布的水不断向周围空气中气化吸热，使得温度计感温部的温度下降。
相对湿度	指空气中水汽压的百分比，湿空气的绝对湿度与相同温度下可能达到最大绝对湿度之比，即空气中的含水量[4]。

湿球温度与干球温度之差可以用来表示湿度，当湿度较小时，则表明空气中的水蒸气含量较低。所以单独给出湿球温度是没有意义的，必须要和干球温度结合起来看，同

“泰迪杯”优秀论文

时干球温度就是通常意义下的温度，由问题的背景可以知道，这些数据来自于一个热带地区的城市，常年平均温度在 25~32 摄氏度之间，平均湿度是 85%左右。

3) **系统状态参数：**冷水泵状态、冷凝水泵状态、冷却装置状态和冷却塔状态。这些都是属于可控变量，在对中央空调进行优化控制的时候，可以人为的对其进行改变，且都是 01 变量，0 状态表示此设备处于关闭状态，1 状态表示此设备处于开启状态。在后面求解时需要特别考虑 0-1 变量在自变量时如何应用。

4) **控制参数：**冷水泵转速、冷凝水泵转速、冷却塔风扇转速等，各个转速都有其频率范围，数据中都是将其化为了百分数形式。这些都属于可控变量，可人为进行调整，属于设备控制参数，转速与设备的效率和耗电量等有关系。

5) **系统采集信息：**主要指设备的耗电量和其他传感器的读数，比冷凝水进、出冷却装置的水温和流速。这些都是不可控变量，都属于因变量。比如设备的耗电量，设备的状态以及各个设备的转速都会影响设备的耗电量。再具体分析的时候，将其作为因变量进行分析。

6) **系统运行相关信息：**主要指耗电量、冷却负载和系统效率等内容。这几个变量都是不可控变量，也是作为因变量进行分析，与各个可控变量的取值有关系。对其解释如下：

表格 2 系统运行相关变量解释

变量名称	解释说明
耗电量	空调消耗的电量，从节约能源的角度来看，我们更希望中央空调的耗电量能够越小越好。
冷却负载	空调系统的输出功率。
系统效率	1RT 输出功率所消耗的电量，即单位功率所消耗的电量，所以理论上是希望其越小越好的。

2.1.3 中央空调结构分析

中央空调系统是由三套冷却装置、两个冷却塔、三个冷凝水泵和四个冷水泵组成的。其基本结构示意图如下所示：

“泰迪杯”优秀论文

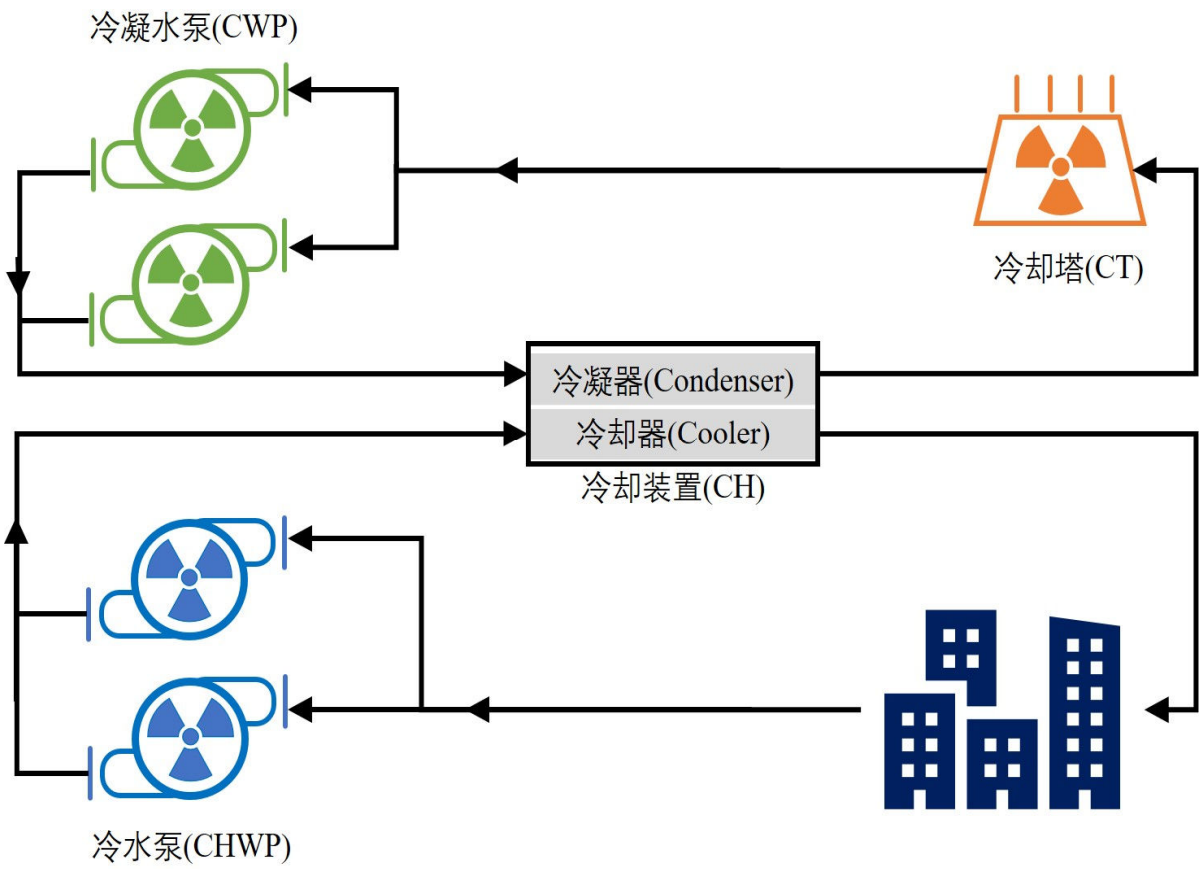


图 2 中央空调系统的基本结构示意图

从上图中可以简单分析出中央空调基本工作原理，其包含内循环和外循环两个热交换系统[5]。内循环中主要是利用了冷水泵、冷却装置对大楼内部空气进行降温和除湿，外循环主要应用冷凝水泵、冷却塔和冷却装置对内循环产生的热量进行降温发散。

冷凝输泵、冷却塔和冷水泵三个装置中都有风扇的存在，对应于数据中有一个装置的转速，风扇与各个设备的效率、功率以及耗电量等有很大的关系。在进行数据分析时，应该着重对其进行处理。根据中央空调的工作原理，冷却装置主要是承担内循环和外循环进行热量交换的地方，因此它的功率主要与流入流出冷却装置的水温差有关。

2.2 具体步骤

这里，我们首先对数据进行预处理，对数据中的异常值进行修正。并采用适当的方法对数据进行对数据降维，再挖掘其内在规律，分析出各个变量之间的关联关系。基于对优化问题的分析，建立中央空调控制策略优化模型，并采用交叉验证的方式对模型的准确性和实用性进行检验。

2.2.1 数据预处理

由于现实中的数据一般都是不完整、不一致的情况，数据的不完整性主要表现在数据缺失以及数据逻辑错误，通常情况下是由于人工填表的疏忽或者是数据录入电脑时不严谨造成的。为了使得数据挖掘结果更加准确，我们首先对全部的数据进行检查。

我们使用 R 语言对全部的数据进行统计分析，得到下图按行展示真实值和缺失值的矩阵图。

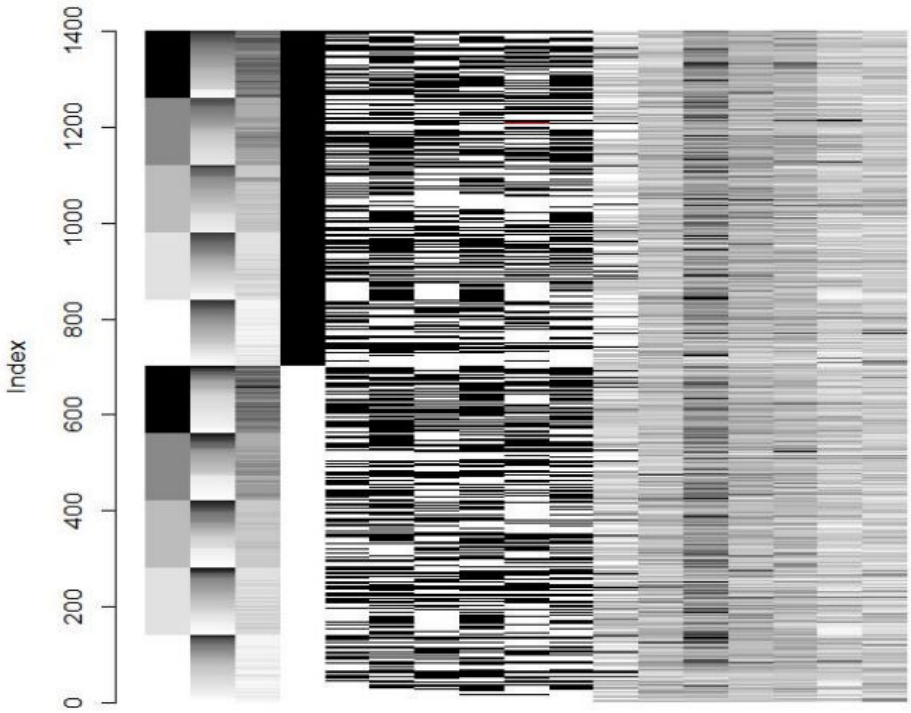


图 3 原始数据按行展示真实值和缺失值的矩阵图

上图中，浅色表示值小，深色表示值大，默认缺失值为红色[6]。因此，给出的数据中不存在缺失值，但在原始数据中存在取值不合理的现象，比如湿求温度室外的第 15428、42916、62304 的数据与周围值对比，或者偏高或者偏低，因此我们对其进修正。

对于缺失数据的处理，目前存在两种方法进行处理，一种是删除原有缺失数据的行，另一种是使用合理的值来替换缺失数据。这里，我们采用均值替换法对缺失的数据进行处理。以列为单位，取原始数据中剔除异常数据后的剩余数据的平均值，来对缺省数据的填补，即：

$$\bar{x}_j = \frac{\sum_{i=1}^{n} x_{ij} - x_{ij\max} - x_{ij\min}}{n-2}. \tag{1}$$

其中：$x_m^{'}$ 为缺失值。

对于数据不正常的异常数据，我们采取先删除，后添加的方法。预处理后的数据见，附件一。

2.2.2 符号说明

为了方便数据挖掘模型建立和求解，我们定义本文的符号说明如下表所示：

表格 3 模型符号说明

符号	解释说明
v_{chwp}	表示冷水泵转速（%）
v_{cwp}	表示冷凝水泵转速（%）
v_{ct}	表示冷却塔风扇转速（%）
db	表示干球温度（室外），单位^{o}C
wb	表示湿球温度（室外），单位^{o}C
rh	表示相对湿度（%）
W	表示总耗电量，单位千瓦
$loadsys$	表示系统的冷却负载，单位冷冻吨
$effsys$	表示系统效率，单位kw/RT

本文中用到的其它具体符号会在首次使用时加以说明。

2.2.3 数据规律分析

附件给出的数据量较大，全部用于分析其变化特征和规律比较困难。这里，我们先对数据进行降维，以便分析数据的大体趋势。我们对数据进行基本统计量分析，其描述数据基本特征主要为集中位置和分散位置，具体基本统计量特征如下：

泰迪杯，优秀论文

1) 样本均值描述了数据取值的**平均位置**，虽然计算简易，但易受到异常值的影响；

2) 中位数描述了**中心位置**的数据特征，比中位数大或小的数据个数大约为样本容量的一半[7]。若数据分布对称则均值与中位数比较接近。其受异常值的影响较小；

3) 方差描述数据取值的**分散性**，它是数据相对于均值的偏差平方的平均；

4) 极差更为简单地描述数据的分散性特征，且**四分位极差**在异常点判别中起到重要作用。

对于基本统计量均值和中位数，一方面，虽然均值与中位数都是描述数据集中位置的数字特征，但是均值用了全部信息，中位数只用了部分信息，通常情况下均值比中位数有效；另一方面，当数据有异常值时，中位数比较稳健。

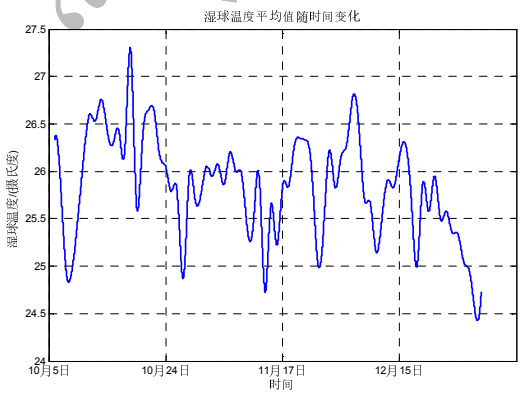
为了兼顾两方面，我们引入了三均值[8]的概念，其定义如下：

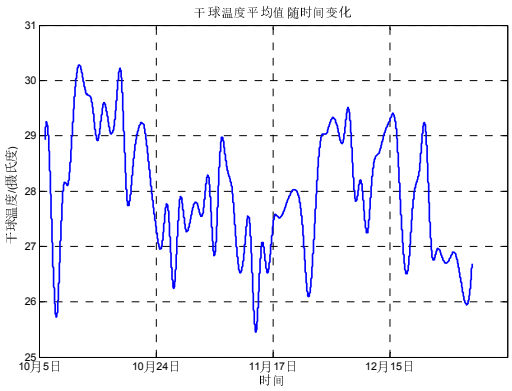
$$\hat{M}=\frac{1}{4}M_{0.25}+\frac{1}{2}M+\frac{1}{4}M_{0.75}.$$

其中 $\hat{M}$ 代表三均值，是上四分位数、中位数和下四分位数的加权平均，即分位数向量 $(M_{0.25},M,M_{0.75})$ 和权向量 $(0.25,0.5,0.25)$ 的内积。这里的四分位数是 p 分位数的特例，其定义如下：

$$M_p=\begin{cases}x_{([np]+1)}, np\text{不是整数},\\ \frac{1}{2}(x_{(np)}+x_{(np+1)}), np\text{是整数}.\end{cases}$$

因此，我们对附件数据进行处理，分别求出干湿球温度、相对湿度、流入流出冷却装置的水流速度、流入流出冷却装置的水温差、总耗电量、冷却负载和系统效率的三均值图。





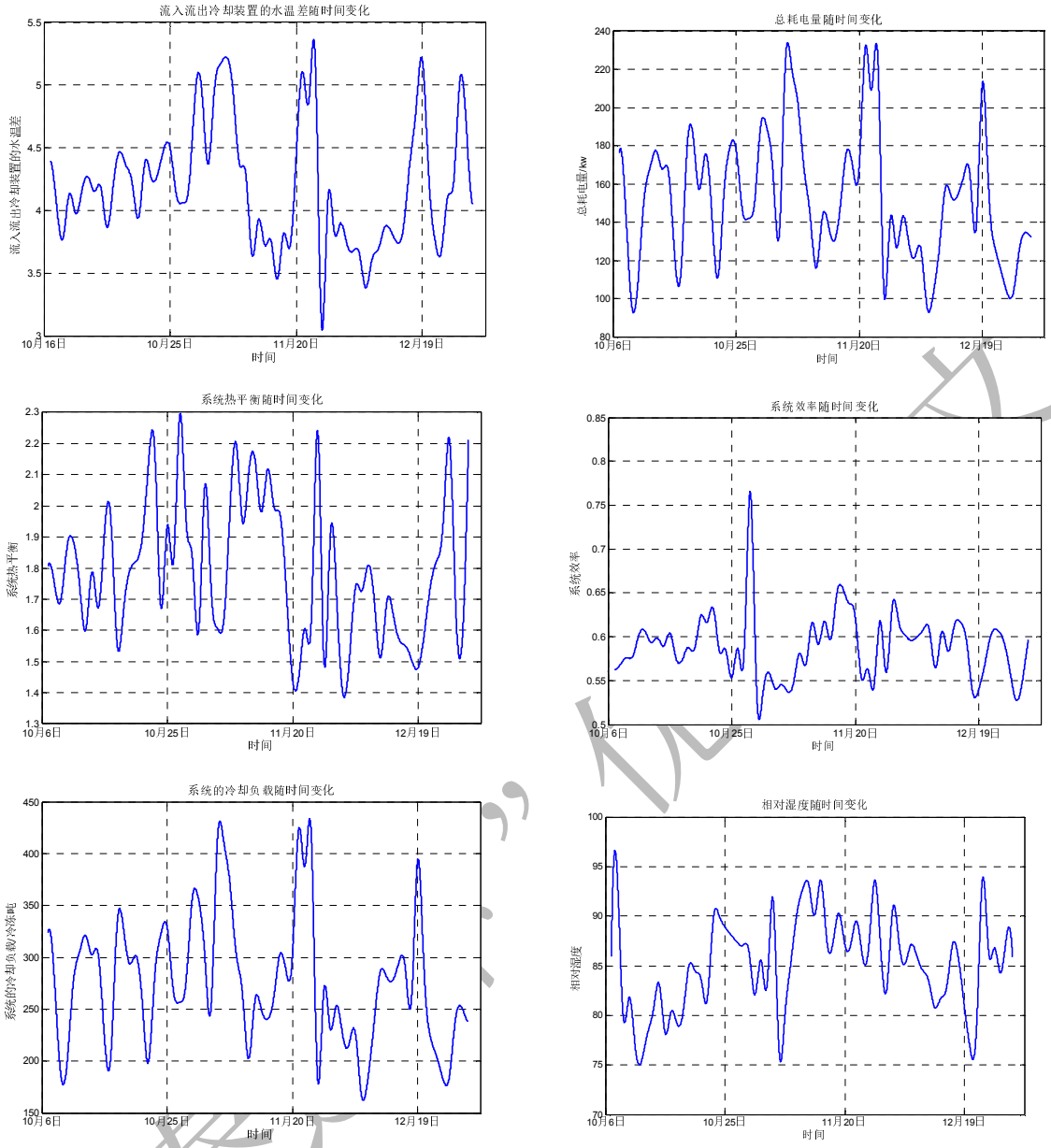


图 4 各个变量三均值图

根据上图可以看到，这些变量基本上都有一定的周期性，其中干球温度和湿球温度的周期性不明显，但基本上符合该地区的年平均气温规律，在 25 至 32 摄氏度之间震荡，其中湿球温度要略低于干球温度[9]。

为了更好的分析变量数据的周期性规律，我们随机抽取这些变量几个周的数据进行分析，画出以周为单位的连续七天的数据趋势图，又将其中的某一个周的数据拿出来，画出以天为单位的数据。

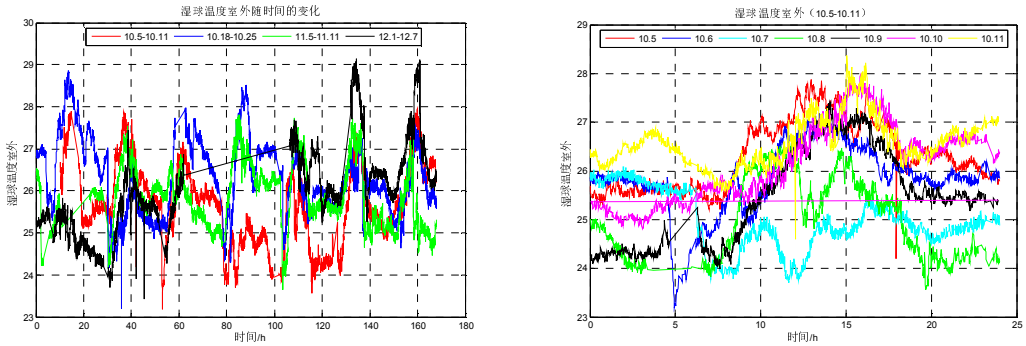


图 5 湿球温度周变化和日变化规律

上图分别是湿球温度（室外）一周和一天的变化曲线图。左图中描述了四个周的数据，红色曲线是 10.5-10.11 七天的数据曲线，蓝色为 10.18-10.25 七天的数据曲线，绿色为 11.5-11.11 七天的数据曲线，黑色为 12.1-12.7 七天的数据曲线。四周的数据曲线趋势大致相似，其中蓝色和绿色的曲线相似度最高，但并不是每一天都明显的重合。我们将其中 10.5-10.11 的数据进行更加细致的分析，画出右图的趋势图，可以看出在每天的同一时刻，即使数据没有完全重合，但是相差不大。其形状都是类似的，因此我们初步判断湿球温度是具有以日为周期的周期性。

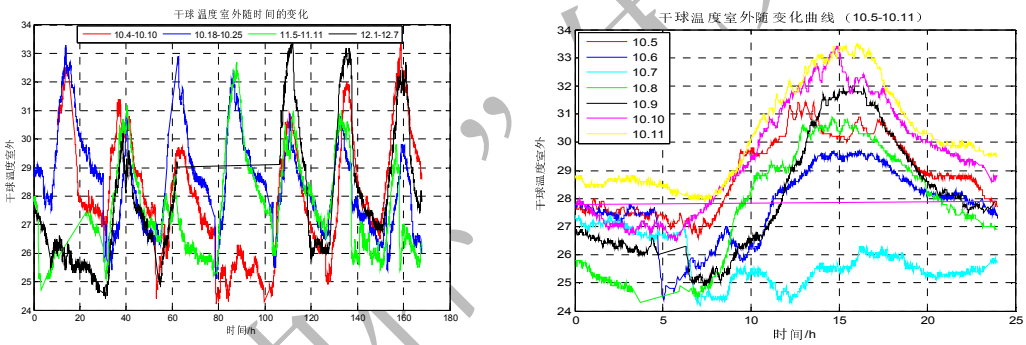


图 6 干球温度周和日变化规律

上图分别是干球温度（室外）一周和一天的变化。与湿求温度相同的分析，可以看出干球温度室外与时间具有明显的周期性关系。且干球温度与湿求温度对比来说，干球温度的变化幅度更大，具有能明显的峰值出现，且从横坐标来看，大约出现在中午。

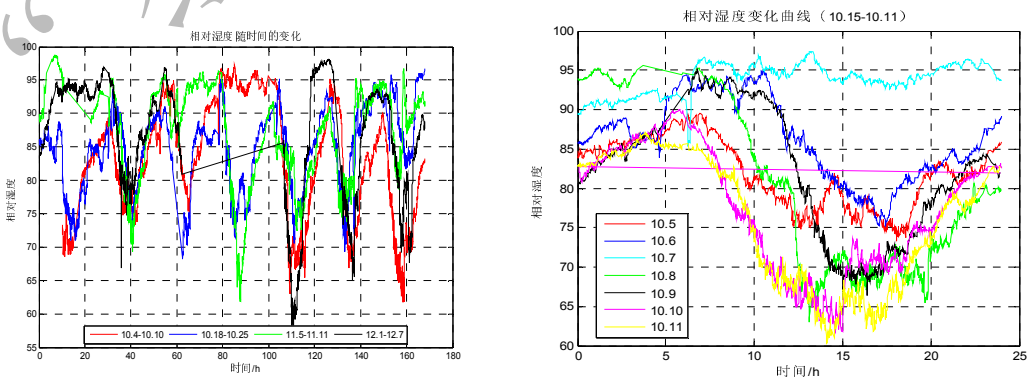
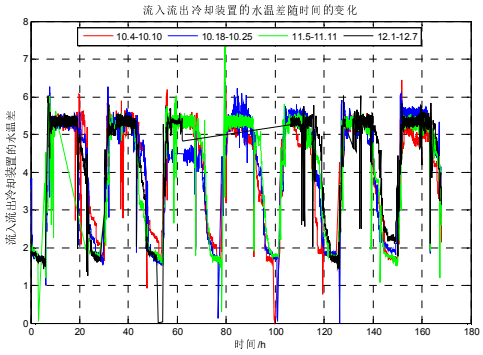


图 7 相对湿度周变化和日变化

同样的方法，我们做出相对湿度的变化曲线图，能够观察到 4 个周的相对湿度相对于时间的变化情况，也具有以天为单位的周期性关系。结合图 5、图 6 和图 7，我们发现一天中的干球温度的最高点出现在 15 时左右，湿球温度的最低点出现在 15 时左右，且此时的相对湿度是最低的，这与我们的分析相符合，湿求温度与干球温度的差值越大，则表明空气中的含水量越低，即相对湿度越低。



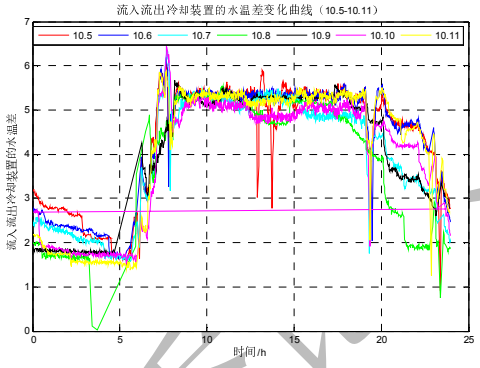


图 8 流入流出冷却装置水温差的周变化和日变化

我们观察上图流入流出冷却装置的水温差相对于时间的变化情况，可知流入流出冷却装置的水温差与时间具有明显的周期性关系。从由图中可以看到，水温差在 8-20 时之间的水温差变化很小，稳定在较高的水平。在 5 时和 20 时都存在突变，5 时水温差突然变大，可能是因为此时商场或者旅馆开始有工作人员进出，热量交换变多；在 20 时水温差突然减小，可能是因为此时商场面临关门等情况，人员变少，相应的热量交换减小。

2.2.4 基于回归分析的相关关系挖掘

1、基于聚类模型的设备分类

1) 聚类模型建立

在进行变量相关性分析时，设备状态是用 0-1 变量表示的，而 0-1 变量虽然可以作为回归变量分析，但是会影响回归效果。且由于数据量较大，也会影响回归效果，因此我们建立动态聚类模型，对设备状态进行聚类分析。

利用动态聚类方法的思想，我们将原始数据中三套冷却装置的状态、两个冷却塔的状态、三个冷凝水泵的状态、四个冷水泵的状态作为粗糙的初始分类，然后用最优分类准则进行修正，直到分类尽可能的合理为止，聚类的过程如下图。

泰迪杯，优秀论文

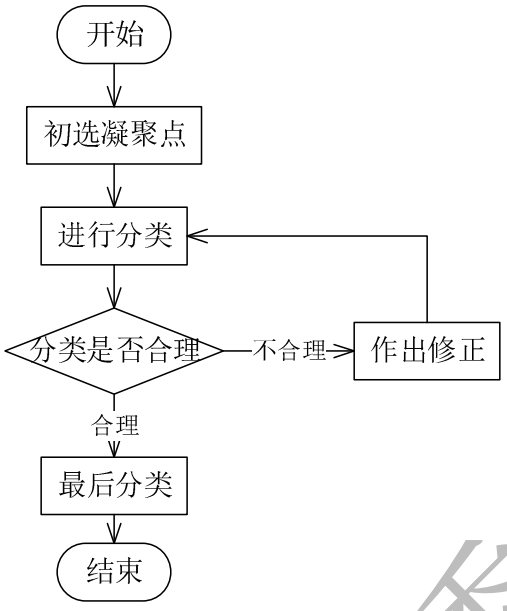


图 9 动态聚类过程图

根据上述动态聚类的流程，我们将其算法步骤归纳如下：

Step1. 确定分类个数 k。凝聚点是欲形成类的中心，因此我们三套冷却装置的状态、两个冷却塔的状态、三个冷凝水泵的状态、四个冷水泵的状态的样本作为凝聚点，把这些样本单独作为一类

Step2. 按照选择距离最近准则，将其余 n-k 个样本逐个并入最近凝聚点所代表的类，合并于已有的类中，实现分类。然后判断分类是否合理，若不合理，则应作出修正，然后再分类，一直循环到分类合理为止。

Step3. 把初始凝聚点距离较大的样本单独分为一个新类。

Step4. 对聚类结果进行分析，以判断其合理性。通过计算类间距离和类内距离之比，发现比值较大，说明各个类的类内凝聚度和类间区分度较高，因而聚类较为合理。

2) 聚类模型求解

针对上述流程，我们利用 SPSS 软件分析附件，得到聚类树状图如下。

优秀论文

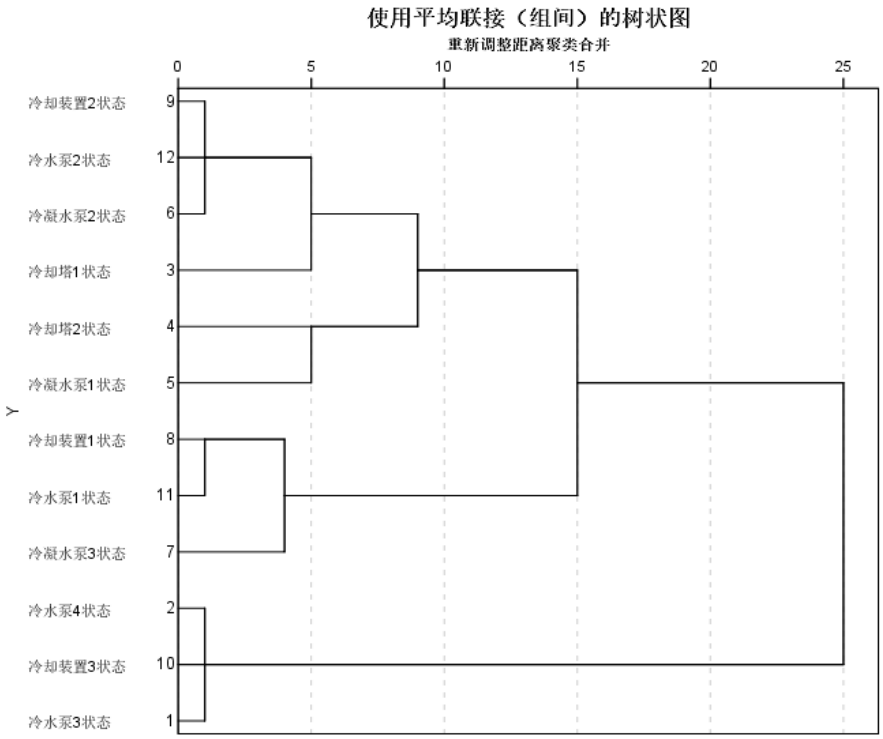


图 10 聚类分析树状图

从上图可以看出一共将设备分为了 4 类，下表为具体的聚类分类表。

表格 4 聚类分类表

阶	群集组合		系数	首次出现阶群集		下一阶
	群集 1	群集 2		群集 1	群集 2	
1	9	12	.000	0	0	5
2	2	10	.000	0	0	3
3	1	2	.000	0	2	11
4	8	11	1.000	0	0	6
5	6	9	5.000	0	1	8
6	7	8	87.500	0	4	10
7	4	5	102.000	0	0	9
8	3	6	105.333	0	5	9
9	3	4	179.250	8	7	10
10	3	7	317.222	9	6	11
11	1	3	554.444	3	10	0

由上表我们可以清楚的看出如何把设备状态聚成一类。结合表中的数据和聚类树状图，我们可以将 4 种 12 个设备分为以下 4 个组合：冷却装置 2、冷水泵 2、冷凝水泵 2、冷却塔 1 为一类，冷却装置 3、冷水泵 4、冷凝水泵 3 为一类，冷却装置 1、冷水泵 1、冷凝装置 3 为一类，冷却塔 2、冷凝水泵 1 为一类。

2、回归模型的一般形式

现实生活中引起被解释变量变化的因素并非仅只有一个解释变量，可能有很多解释变量。例如产出往往受各种投入要素，如资本、劳动、技术等的影响；销售额往往受价格和公司对广告费的投入的影响等。

所以在一元线性模型的基础上，提出多元线性模型[10](解释变量的个数≥ 2)。

表格 5 模型对比

多元回归模型	多元线性回归模型
含两个以上解释变量的回归模型	一个因变量与多个解释变量之间设定的是线性关系

多元线性回归模型的一般形式为：

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \ldots + b_k X_k. \tag{2}$$

其中，多元回归模型中的回归系数称为偏回归系数，其意义是在其它解释变量保持不变的条件下，该变量变化一个单位，被解释变量将平均发生偏回归系数大小的变动。

3、回归模型的误差检验

我们查阅文献，得到回归分析表如下所示：

表格 6 回归分析表

离差名称	平方和	自由度	均方差
回归	RSS	K	RSS/K(k 个自变量)
残差	ESS	n-k-1	ESS/n-k-1
总离差	TSS	n-1	

判定系数$R^2 = \frac{RSS}{TSS} = 1 - \frac{ESS}{TSS}$越大，自变量对因变量的解释程度越高，自变量引起的变动占总变动的百分比越高。

校正判定系数如下所示：

$$\bar{R}^2 = 1 - \frac{ESS/(n-k)}{TSS/(N-1)}. \tag{3}$$

当 k>1 时，$\bar{R}^2 < R^2$，且随着解释变量的增加两者的差距越来越大，也就是说校正的比未校正的判定系数增加的慢些。

“泰迪杯”优秀论文

- **回归方程的显著性检验**

Step1. 提出假设

原假设 $H_0: b_1 = b_2 = \ldots b_k = 0$；备择假设 $H_1: b_i$ 不全为 $0(i = 1,2,\ldots k)$.

Step2. 计算统计量 $F = \dfrac{RSS/K}{ESS/(n-k-1)} \sim F(k, n-k-1)$.

Step3. 查表得 $F_\alpha = F_\alpha(k, n-k-1)$.

Step4. 检验

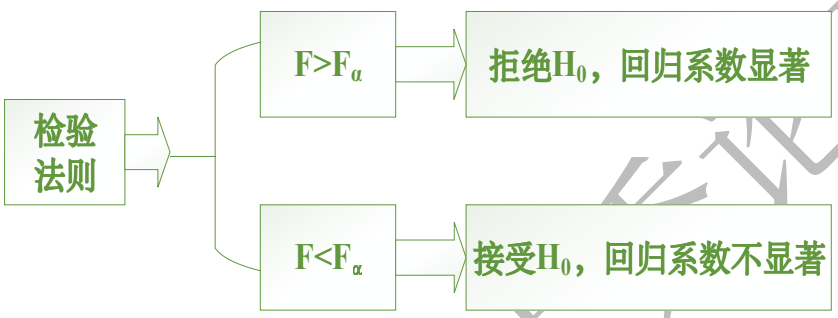


图 11 回归方程检验

- **回归系数的显著性检验**

Step1. 提出假设

原假设 $H_0: b_1 = b_2 = \ldots b_k = 0$；备择假设 $H_1: b_i$ 不全为 $0(i = 1,2,\ldots k)$.

Step2. 计算统计量 $T_i = \dfrac{\hat{b}_i}{s(\hat{b}_i)}, i = 1,2\ldots\ldots k$.

Step3. 查表得 $t_{\frac{\alpha}{2}} = t_{\frac{\alpha}{2}}(n-k-1)$.

Step4. 检验

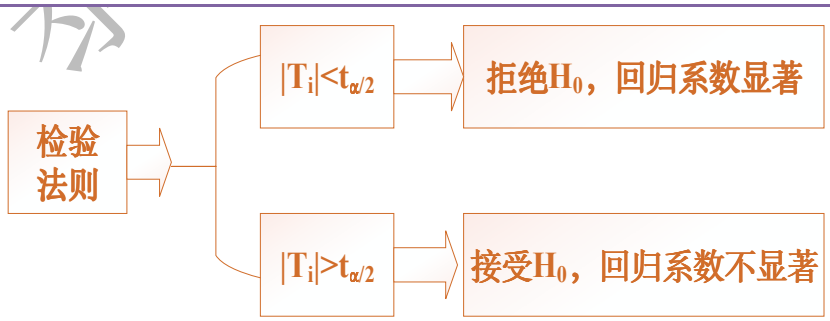


图 12 回归系数检验

- **变量的确定**

在进行回归分析的时候，需要找出各个因变量和自变量。这里，因变量我们设定为是不可控变量，结合问题二三四中求解的量，我们将功率、冷却负载、总耗电量作为因变量进行分析。其它的可控变量作为因变量进行分析。

依次对各个因变量和自变量建立多元线性回归模型，并用 SPSS 直接求解结果。

4、变量的回归分析

1) 功率为因变量的回归分析

首先，我们定义设备的平均功率如下：

$$\bar{P}=\sum_{i=1}^{n}x_i\cdot P_i/\sum_{i=1}^{n}x_i. \tag{4}$$

其中，$\bar{P}$ 表示平均功率，x_i 表示设备 i 的状态，P_i 表示设备 i 的功率。

◆ **冷凝水泵功率**

通过对冷凝水泵功率与各个自变量的回归分析，得到如下系数结果。

表格 7 冷凝水泵平均功率的回归分析

模型	非标准化系数		标准系数	t	Sig.	共线性统计量	
	B	标准 误差	试用版			容差	VIF
(常量)	1.150	0.471		2.440	0.015		
冷凝水泵转速	0.074	0.002	0.904	49.046	0.000	0.845	1.184
相对湿度	0.012	0.005	0.041	2.241	0.025	0.845	1.184

通过上表，我们可以得出多元回归方程为：

$$\bar{P}_{cwp}=1.15+0.074\cdot v_{cwp}+0.012\cdot db. \tag{5}$$

其中，$\bar{P}_{cwp}$ 表示**冷凝水泵平均功率。**

通过对回归系数进行 t 检验，可得**冷凝水泵**转速、相对湿度、常量的显著性检验的 P 值分别为 0、0.025、0.015，在显著性水平 α 为 0.05 下，均能通过显著性检验。并且，回归分析的 R 方为 0.889，拟合效果较好。

◆ **冷却塔功率**

对冷却塔和各个自变量进行回归分析，得到下表中的结果。

表格 8 冷却塔平均功率的回归分析

模型	非标准化系数		标准系数	t	Sig.	共线性统计量	
	B	标准 误差	试用版			容差	VIF
(常量)	1.677	0.248		6.757	0.000		
冷却塔风扇转速	0.135	0.001	0.974	97.505	0.000	0.972	1.029
相对湿度	0.024	0.003	0.085	8.502	0.000	0.972	1.029

通过上表，我们可以得出多元回归方程为：

$$\bar{P}_{ct} = 1.677 + 0.135 \cdot v_{ct} + 0.024 \cdot rh . \quad (6)$$

其中，$\bar{P}_{ct}$ 为**冷却塔平均功率**。

通过对回归系数进行 t 检验，可得**冷却塔**转速系数、相对湿度系数、常量的显著性检验的 P 值均接近 0 在显著性水平 α 为 0.01 下，均能通过显著性检验。并且，回归分析的 R 方为 0.964，拟合效果较好。

◆ **冷水泵功率**

我们对冷水泵功率和各个自变量进行回归分析，得到下表中的结果

表格 9 冷水泵平均功率的回归分析

模型	非标准化系数		标准系数	t	Sig.	共线性统计量	
	B	标准 误差	试用版			容差	VIF
(常量)	1.781	0.176		10.121	0.000		
冷水泵转速	0.082	0.001	0.961	81.970	0.000	0.956	1.046
相对湿度	0.010	0.002	0.056	4.742	0.000	0.956	1.046

通过上表，我们可以得出多元回归方程为：

$$\bar{P}_{chwp} = 1.781 + 0.082 \cdot v_{chwp} + 0.010 \cdot rh . \quad (7)$$

其中，$\bar{P}_{chwp}$ 为**冷水泵平均功率**。

通过对回归系数进行 t 检验，可得**冷水泵**转速系数、相对湿度系数、常量显著性检验的 P 值均接近 0 在显著性水平 α 为 0.01 下，均能通过显著性检验。并且，回归分析的 R 方为 0.951，拟合效果较好。

◆ **冷却装置功率**

我们对冷却装置的功率和各个自变量进行回归分析，得到下表中的结果。

表格 10 冷却装置平均功率的回归分析

模型	非标准化系数		标准系数	t	Sig.	共线性统计量	
	B	标准 误差	试用版			容差	VIF
(常量)	-73.149	12.896		-5.672	0.000		
流入流出冷却装置的水温差	17.005	0.378	0.760	45.035	0.000	0.463	2.162
冷水泵转速	0.958	0.030	0.466	32.099	0.000	0.624	1.602
冷却塔风扇转速	-0.285	0.027	-0.140	-10.382	0.000	0.724	1.380
湿球温度室外	6.063	0.851	0.183	7.123	0.000	0.199	5.024
干球温度室外	-2.589	0.431	-0.160	-6.005	0.000	0.186	5.390
冷凝水泵转速	-0.059	0.022	-0.049	-2.713	0.007	0.403	2.484

通过上表，我们可以得出多元回归方程为：

$$\bar{P}_{ch} = -73.149 + 17.005 \cdot \Delta T + 0.958 \cdot v_{chwp} - 0.285 \cdot v_{ct} - 0.059 \cdot v_{cwp} + -2.589 \cdot db + 6.063 wb \,. \quad (8)$$

其中，$\bar{P}_{ch}$ 为**冷却装置平均功率。**

通过对回归系数进行 t 检验，可得各变量系数和常量的显著性检验的 P 值在显著性水平 α 为 0.01 下，均能通过显著性检验。并且，回归分析的 R 方为 0.951，拟合效果较好[11]。

2) 总耗电量分析

这里，我们不对总耗电量和自变量进行回归分析，直接根据总耗电量和功率之间的关系，可以得到总耗电量的计算公式如下：

$$W = \sum_{i=1}^{n} P_{ct,i} \times x_{ct,i} + \sum_{i=1}^{n} P_{chwp,i} \times x_{chwp,i} + \sum_{i=1}^{n} P_{cwp,i} \times x_{cwp,i} + \sum_{i=1}^{n} P_{ch,i} \times x_{ch,i}. \quad (9)$$

即总耗电量是三个冷却装置、两个冷却塔、三个冷凝水泵和四个冷水泵的功率相加，得到总耗电量的值。

3) 冷却负载为因变量的回归分析

以冷却负载作为因变量，与各个自变量进行回归分析，得到下表中的系数结果。

“泰迪杯”优秀论

表格 11 冷却负载的回归分析

模型	非标准化系数		标准系数	t	Sig.	共线性统计量	
	B	标准 误差	试用版			容差	VIF
(常量)	-232.401	16.977		-13.689	0.000		
冷却装置状态	135.375	4.717	0.422	28.702	0.000	0.216	4.620
冷水泵转速	2.287	0.091	0.218	25.232	0.000	0.628	1.592
冷却塔状态	79.474	5.346	0.231	14.866	0.000	0.195	5.141
冷凝水泵转速	0.834	0.065	0.136	12.866	0.000	0.419	2.384
冷却塔风扇转速	0.736	0.090	0.071	8.220	0.000	0.631	1.586
冷水凝泵状态	26.277	3.747	0.137	7.014	0.000	0.123	8.103
干球温度	2.650	0.656	0.032	4.038	0.000	0.743	1.346

通过上表，我们可以得出多元回归方程为：

$$loadsys = -232.401 + 135.375 \cdot x_{ch} + 2.287 \cdot v_{chwp} + 79.474 \cdot x_{ct} + 0.834 \cdot v_{cwp} + 0.736 \cdot v_{ct} + 26.277 \cdot x_{cwp} + 2.65 \cdot db \quad (10)$$

通过对回归系数进行 t 检验，可得各变量系数和常量的显著性检验的 P 值均接近 0，在显著性水平 α 为 0.01 下，均能通过显著性检验。并且，回归分析的 R 方为 0.983，拟合效果较好。

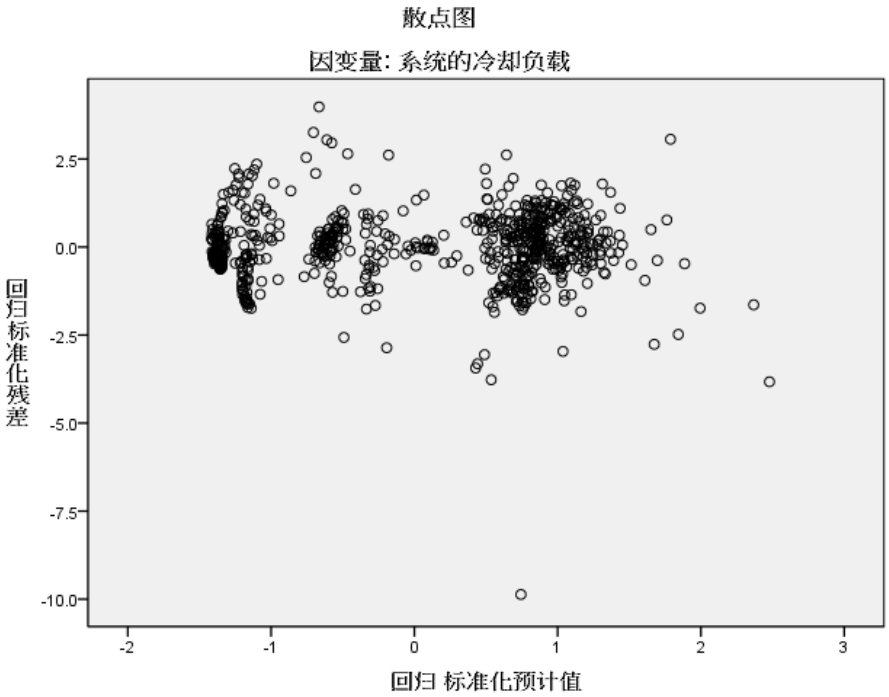


图 12 回归残差

通过上图我们可以看出，回归分析的残差大多分布正在-2.5~2.5 之间，即在横坐标

上下呈带状分布，因此拟合效果较好。

4) 系统效率分析

系统效率的单位是 kw/RT，kw 是功率的单位，RT 是冷却负载的单位，因此可以将系统效率定义为功率与冷却负载的商。

$$effsys = W / loadsys. \tag{11}$$

其中，W 是系统总功率，$loadsys$ 是冷却负载。

根据这系统功率与冷却负载的回归结果，可以分析得出系统效率与各个设备的转速、各个设备的状态以及流入流出冷却装置的水温差有关。

5) 误差分析

为了检验模型的可靠性，我们定义如下均方误差[12]：

$$\sigma^2 = \frac{\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2}{\sum_{i=1}^{n} y_i^2}.$$

其中，y_i 为样本的观察，$\hat{y}_i$ 为样本的实验值。在工业生产中，当 $\sigma^2 < 0.1$ 时，我们认为模型具有稳定的可靠性。

我们使用降维后的数据进行回归分析，再通过对剩余数据我们的模型进行求解，并求模型所得结果与真实数据之间的均方误差，可以得到以下结果：

表格 12 误差分析表

变量	$\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2$	$\sum_{i=1}^{n} y_i^2$	σ^2
总耗电量	851444.66	24475166.32	0.035
系统冷却负载	955.53	19624.59	0.049
系统效率	0.03	0.86	0.037

由上图可以看出，总耗电量、系统冷却负载、系统效率的方差均小于 5%，因此，我们可以认为我们的模型通过了误差检验。

“泰迪杯”优秀论文

2.2.5 当决策变量为设备转速时的优化模型

（1）优化模型建立

基于问题的分析，0 我们从目标函数、决策变量、约束条件等方面对模型进行说明。

1) 目标函数的确定

这里，根据我们提出的两个原则，确定出目标函数。

- 系统总耗电量尽可能小
- 系统效率在正常范围内尽可能小

首先，本着能够保证温度适宜的前提下，应该尽可能的减少能源的消耗。而根据题目出中给出的关于系统效率的定义：单位功率消耗的电量，则应该让系统效率尽可能小，但是系统效率也不能无限小，因为空调运转势对空气降温做工一定会消耗电量，因此只能是让系统效率在一定范围内尽可能小。其次，对于系统总的耗电量来说，肯定是越小越好。

将其表示为如下数学公式：

$$\begin{array}{l}\min\ W = f\left(v_k, x_i\right) \\ \min\ effsys\end{array} \tag{12}$$

其中，W 表示总耗电量，它是设备状态与设备转速的函数；v_k 是设备的转速，$k=1,2,3$；x_i 是设备的状态参数，$i=1,2,\cdots,12$；$effsys$ 为系统效率。

2) 决策变量分析

本文题中影响目标函数的变量是各个设备的转速，因此各个设备的转速是决策变量，我们以冷水泵转速、冷凝水泵转速和冷却塔风扇转速三个变量为决策变量。为方便表示，不妨记为v_k，$k=1,2,3$。

3) 约束条件分析

这里，我们从设备的转速约束、设备状态约束、系统稳定约束三个方面进行分析。

- **设备转速约束**

冷水泵转速、冷凝水泵转速和冷却塔风扇转速三个转速均有范围限制，因此，应有$0<v_k<100, k=1,2,3$。

- **设备状态约束**

由于设备的状态分为开和关，我们以 0 表示设备处于关闭状态，以 1 表示设备处于工作状态，则设备状态x_i为 0-1 变量，应满足$x_i=0\ or\ 1, i=1,2,\cdots,12$。

泰迪杯优秀论文

➢ **系统稳定约束**

为了能够保证设备正常工作，系统热平衡不应该过大，通过查阅相关文献，我们得出以下约束：$hbsys < 5\%$。同时，每台设备均不能超负荷运转，即设备的实际功率应小于额定功率 $P_i \le P_{i,\max}$。

综上所述，我们得到如下的多目标线性规划模型：

$$\begin{aligned}
&\min\ W = f\left(v_k, x_i\right)\\
&\min\ effsys\\
&s.t.\begin{cases}
loadsys = g\left(v_k, y_i\right)\\
effsys = W / loadsys\\
hbsys = h\left(v_k, y_i\right)\\
hbsys < 5\%\\
P_i \le P_{i,\max}, i = 1,2,\ldots,12\\
0 < v_k < 100, k = 1,2,3\\
x_i = 0\ or\ 1, i = 1,2,\cdots,12
\end{cases}
\end{aligned} \tag{13}$$

其中，W 表示总耗电量，可由公式（5）-（9）求得；$effsys$ 为系统效率；$loadsys$ 为系统冷却负载；$hbsys$ 为系统热平衡；P_i 为设备 i 的实际功率，$P_{i,\max}$ 为设备 i 的额定功率；v_k 是设备的转速，$k = 1,2,3$；x_i 是设备的状态参数，$i = 1,2,\cdots,12$。

（2）模型求解算法

这里所建立的优化模型为多目标线性规划问题，在现有的方法中无法直接求解，因此我们将其转化为单目标线性规划求解，这里采用将目标函数转化为约束条件的方法。

对于目标函数系统效率尽量小这一目标函数，针对空调运转必然有系统效率这一特点，我们将该目标函数转化为约束条件：系统效率大于等于系统效率的最小值，求解时系统效率的最小值我们通过对求取附件中数据的系统效率的最小值来表示[14]。

从而，得到如下简化后的模型：

泰迪杯，优秀论文

$$\min\ W = f\left(v_k, x_i\right)$$
$$s.t.\begin{cases} loadsys = g\left(v_k, y_i\right) \\ effsys = W / loadsys \\ effsys \geq effsys_{\min} \\ hbsys = h\left(v_k, y_i\right) \\ hbsys < 5\% \\ P_i \leq P_{i,\max}, i = 1,2,...,12 \\ 0 < v_k < 100, k = 1,2,3 \\ x_i = 0\ or\ 1, i = 1,2,\cdots,12 \end{cases} . \tag{14}$$

针对上述简化后的最优控制策略模型，我们采用比较成熟的概率算法——模拟退火算法，它是基于 Monte Carlo 迭代求解策略的一种启发式随机搜索算法。应用模拟退火法求解出优化模型的最优解，即目标函数的最小值。

模拟退火算法是目前一种通用概率算法，用来在一个大的搜寻空间内寻找问题的最优解。其思想源于固体的退火过程，即将温度加热至足够高的温度，再缓慢冷却。升温时，固体内部粒子随温度上升变为无序状，内能增大，而缓慢冷却时粒子又趋向于有序，从理论上讲，如果冷却过程足够缓慢，那么冷却中任一固体都能达到热平衡，而冷却到低温时将达到这一低温下内能的最小状态。

将模拟退火算法应用于最优化问题求解时，将简化后的目标函数作为内能 E，利用 Java 语言编写程序（见附件二），具体的算法步骤如下：

Step1. 初始化各参数：设置初始温度 T、初始解状态、每个 T 值的迭代次数 L；

Step2. 对于 K=1,2,3…L。重复 Step3 到 Step5；

Step3. 随机出各个设备的转速，满足约束条件，则对结果进行检验

Step4. 若初始解大于新解，则更新初始解，否则继续下一次迭代；

Step5. 迭代次数达到或者解已经满足目标函数时，输出结果；否则而转到 Step3。

（3）模型求解结果分析

为了对模型进行检验，我们以附录中 2016 年 12 月 28 日数据为基础，通过模拟退火算法对该条件下的中央空调进行优化控制，得到 12 月 28 日各个设备的转速、系统效率和总耗电量，部分数据如下表所示（见附件三）。

表格 13 最优控制策略下的转速、系统效率和总耗电量

采集时间	冷水泵转速	冷凝水泵转速	冷却塔风扇转速	优化后系统效率	优化后总耗电量
12/28/2016 0:00	1.7079	14.3735	99.5874	0.4849	32.9708
12/28/2016 0:01	2.1373	77.1165	78.1749	0.5706	38.4254
12/28/2016 0:02	1.6457	45.1421	90.5559	0.5187	35.2912
12/28/2016 0:03	2.5844	93.3231	71.6261	0.5796	40.0520
12/28/2016 0:04	0.0257	31.2971	60.8005	0.5627	39.4529
12/28/2016 0:06	0.2226	40.1118	97.0384	0.4889	33.4312
12/28/2016 0:07	1.0589	84.4804	93.4927	0.5087	35.4974
12/28/2016 0:08	4.9531	26.5154	98.7868	0.5635	38.8306
12/28/2016 0:10	2.0982	91.3273	95.7305	0.5319	35.7712

这里，我们将优化后的系统总耗电量和系统效率与原来的数值进行对比，得出总耗电量平均下降了 15.29%，系统效率平均下降了 19.73%，得到如下对比图。

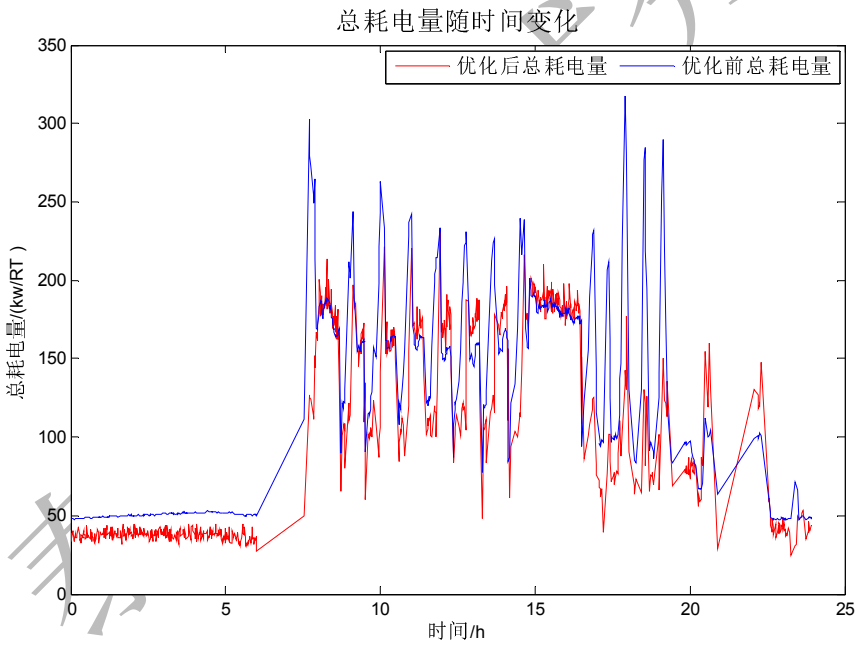


图 13 总耗电量随时间变化图

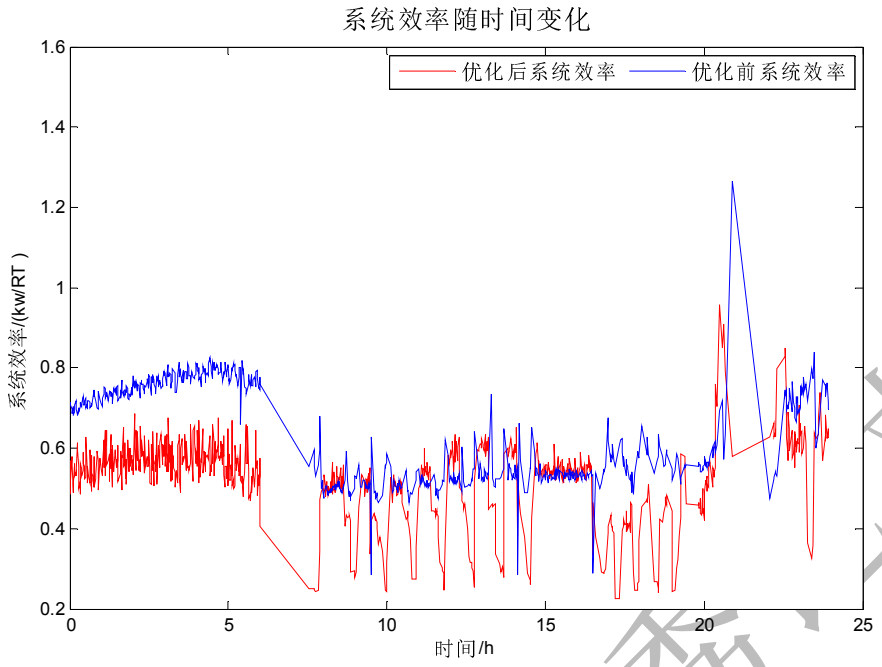


图 14 系统效率随时间变化图

红色为优化后的系统效率，蓝色是优化前的系统效率，优化后的系统耗电量和系统效率都是要低于优化之前的数据，且表中的各个设备的状态也是可见的。因此验证了模型的可靠性与合理性。

2.2.6 当决策变量为所有设备状态变量时的优化模型

（1）优化模型建立及求解

基于问题的分析，我们对决策变量进行了修改，得到如下多目标线性规划模型：

$$\min\ W = f\left(x_i, v_k\right)$$

$$s.t.\begin{cases} loadsys = g\left(v_k, y_i\right) \\ effsys = W / loadsys \\ effsys \geq effsys_{\min} \\ hbsys = h\left(v_k, y_i\right) \\ hbsys < 5\% \\ P_i \leq P_{i,\max}, i = 1,2,\ldots,12 \\ 0 < v_k < 100, k = 1,2,3 \\ x_i = 0\ or\ 1, i = 1,2,\cdots,12 \end{cases} . \tag{15}$$

其中，$effsys_{\min}$ 为系统效率的最小值。

针对上述模型，我们修改模拟退火算法[15]的代码（见附件四），在已知时间、室外温度和湿度、冷却负载和冷水泵转速、冷凝水泵转速和冷却塔风扇转速取值的情况下，

确定设备的状态参数。

算法流程图如下：

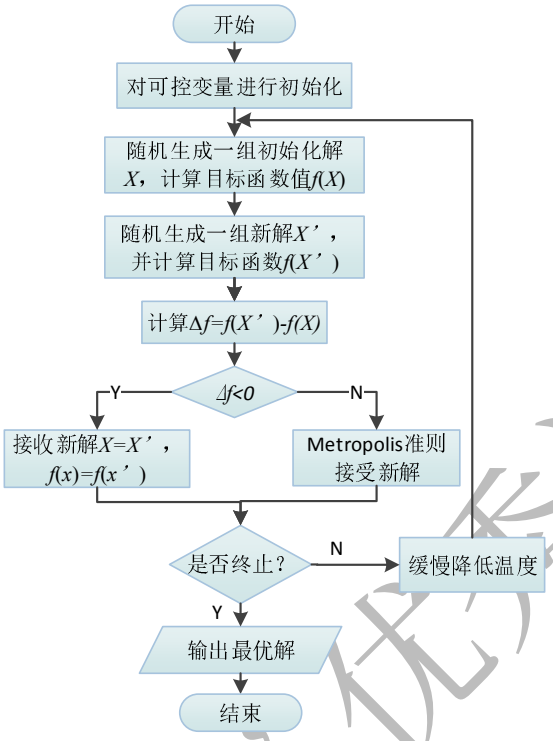


图 15 模拟退火算法流程图

（2）模型结果分析

为了对模型进行检验，我们仍然以 12 月 28 日已知时间、室外温度和湿度、冷却负载和冷水泵转速、冷凝水泵转速和冷却塔风扇转速等变量的数据作为已知量，通过模拟退火算法对该条件下的中央空调进行优化控制，得到 12 月 28 日各个设备的状态、系统效率和总耗电量，部分数据如下表所示（见附件五）。

表格 14 最优控制策略下的设备状态

冷却装置			冷却塔		冷水泵				冷凝水泵		
1	2	3	1	2	1	2	3	4	1	2	3
0	0	0	1	0	0	0	1	1	0	0	0
0	0	0	0	0	0	0	1	1	0	0	0
0	0	0	1	0	0	0	1	0	0	0	0
0	1	0	1	0	0	0	1	0	0	1	0
1	0	0	0	0	0	0	1	0	1	0	0
0	0	0	1	0	0	1	1	0	0	0	0
0	0	0	1	0	0	1	1	0	0	0	0
0	0	0	1	0	0	1	1	0	0	0	0
1	0	0	0	0	0	1	1	0	1	0	0

“泰迪杯”优秀论文

这里，我们将优化后的系统总耗电量和系统效率与原来的数值进行对比，得出总耗电量平均下降了 18.34%，系统效率平均下降了 18.98%，得到如下对比图。

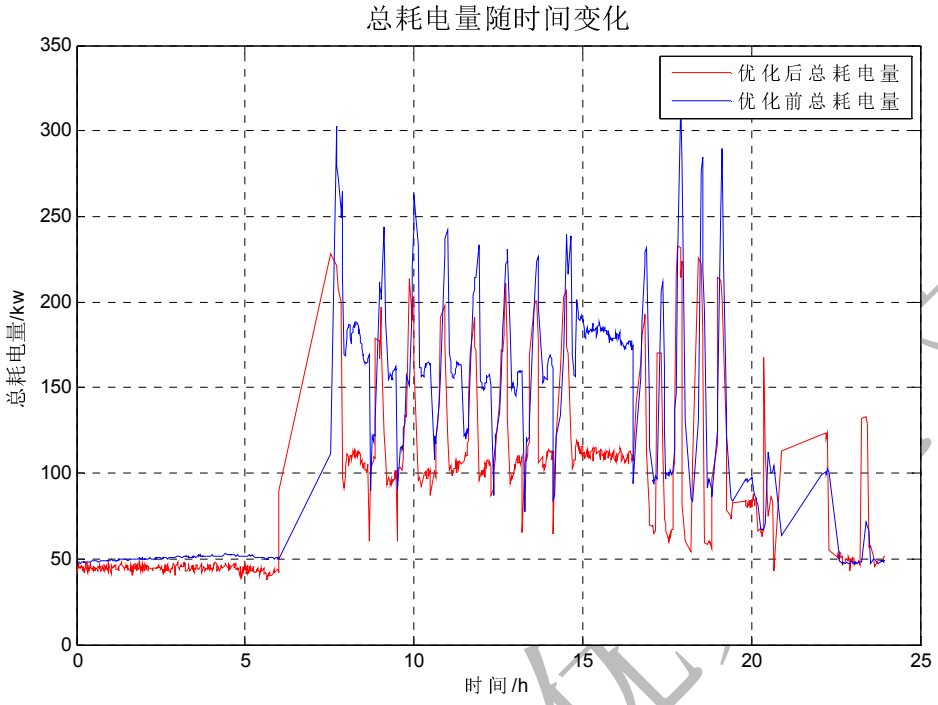


图 16 总耗电量随时间变化图

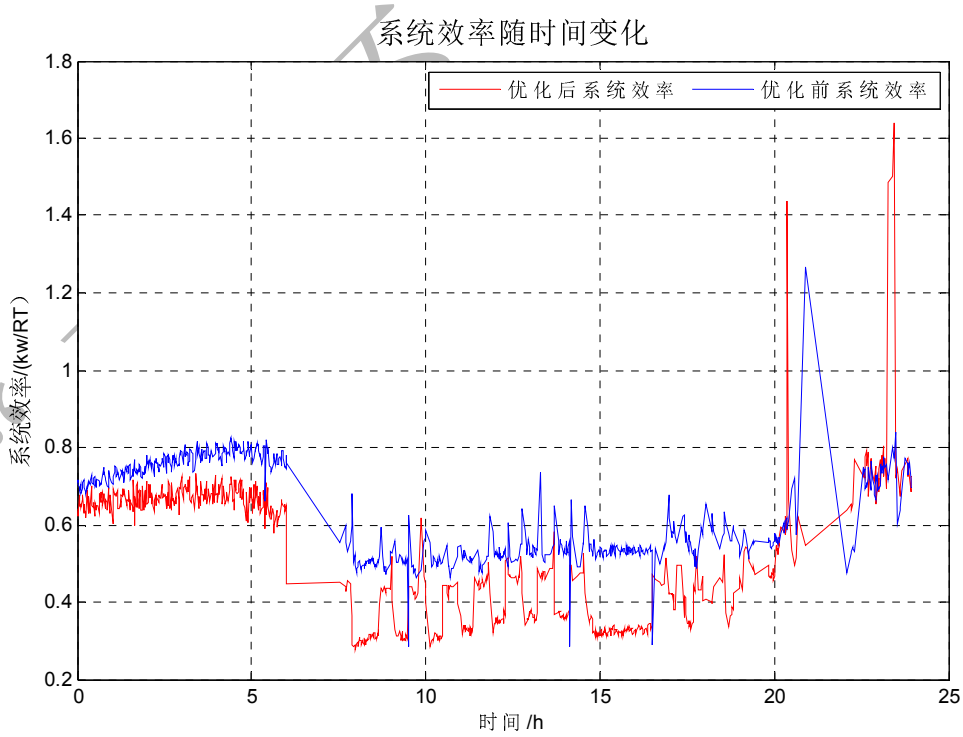


图 17 系统效率随时间变化图

红色为优化后的系统效率，蓝色是优化前的系统效率，优化后的系统耗电量和系统效率都是要低于优化之前的数据，且表中的各个设备的状态也是可见的。因此验证了模型的可靠性与合理性。

2.6.7 当决策变量为所有可控变量时的优化模型

（1）优化模型建立及求解

问题 4 与前两问相比，问题中的已知信息减少，并且需要优化决定的参数有设备状态、设备转速，因此可将问题 4 视为是问题 2 和问题 3 的结合。

在问题 4 中，决策变量应改为设备状态和设备转速，也应增加以下几个约束条件：

- 每台设备的开/关以后 2 个小时以内不可以关/开；
- 每台设备每天最多只能开关 6 次；
- 每台设备不能超负荷运转
- 任何时间都至少要开启一台设备。从而，可以得到以下改进后的多目标线性规划模型：

$$
\begin{aligned}
&\min\ W = f\left(v_k, x_i\right)\\
&\min\ effsys\\
&s.t.\begin{cases}
loadsys = g\left(v_k, y_i\right)\\
effsys = W / loadsys\\
\Delta t = t_{i.2} - t_{i,1} > 2, i = 1,2,...,12\\
P_i \le P_{i,\max},\\
\sum_{t=0}^{24} x_i \le 6, i = 1,2\cdots,12\\
\sum_{i=1}^{12} x_i \ge 1, i = 1,2\cdots,12\\
hbsys = h\left(v_k, y_i\right)\\
hbsys < 5\%\\
0 < v_k < 100, k = 1,2,3\\
x_i = 0\ or\ 1, i = 1,2,\cdots,12
\end{cases}
\end{aligned}
\tag{16}
$$

其中，$t_{i.2}$ 为第 i 个设备关闭的时间，$t_{i,1}$ 为第 i 个设备开启的时间；P_i 为第 i 个设备实际功率，$P_{i,\max}$ 为第 i 个设备额定功率。

上述模型中的决策变量较多，因此我们采用最速下降法对模拟退火算法进行改进。

最速下降法一个最优化算法，通常也称为负梯度法。函数 $j(a)$ 在某点 a_k 的负梯度

“泰迪杯”优秀论文

$-\nabla j(a_k)$是一个向量，其方向是$j(a)$减少最快的方向。最速下降法是以负梯度方向为搜索方向的，最速下降法越接近目标值，步长越小，前进越慢。

梯度下降法的计算过程就是沿梯度下降的方向求解极小值，其迭代公式为

$$a_{k+1} = a_k + \rho_k \hat{s}^{(k)}$$

其中，$\hat{s}^{(k)}$代表梯度负方向，ρ_k表示梯度方向上的搜索步长。

我们利用 Java 编写代码，实现模拟退火算法（见附件六）。

（2）模型结果分析

为了对模型进行检验，我们以附件中已知时间、室外温度和湿度、冷却负载等变量的数据作为已知量，通过模拟退火算法对该条件下的中央空调进行优化控制，得到各个设备的状态、转速、系统效率和总耗电量的数据，见附件七。我们将优化后的系统总耗电量和系统效率与原来的数值进行对比，得出总耗电量平均下降了 41.55%，系统效率平均下降了 41.56%。

我们仍然是使用 12 月 28 日一天内已知时间、室外温度和湿度、冷却负载等变量的数据作为已知量，求解出各个设备开关状态和各个设备的转速，并将优化后的系统效率和功率进行输出。

我们将原始的中央空调系统效率和总功率与优化后的进行对比，得到下图。

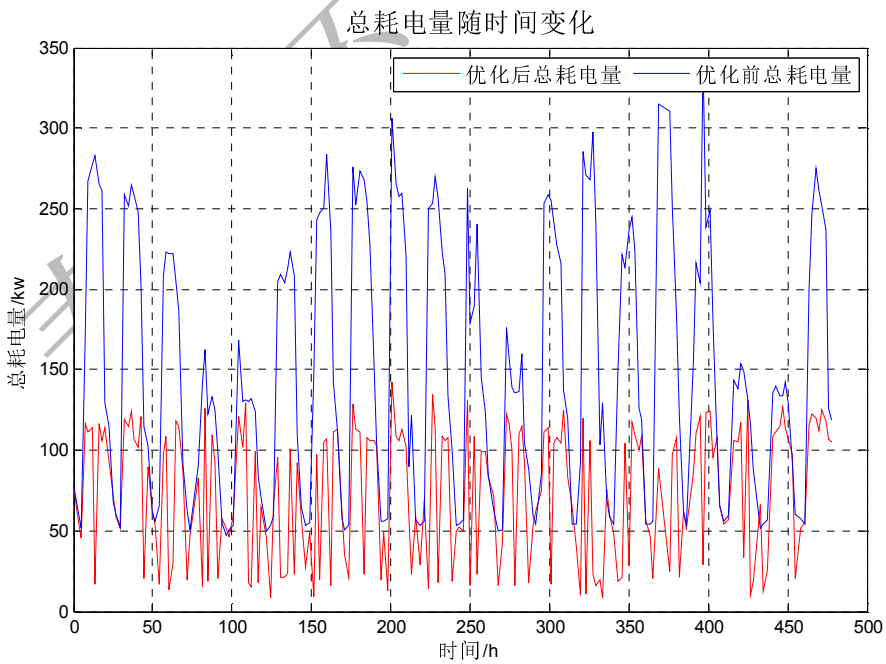


图 18 总耗电量随时间变化图

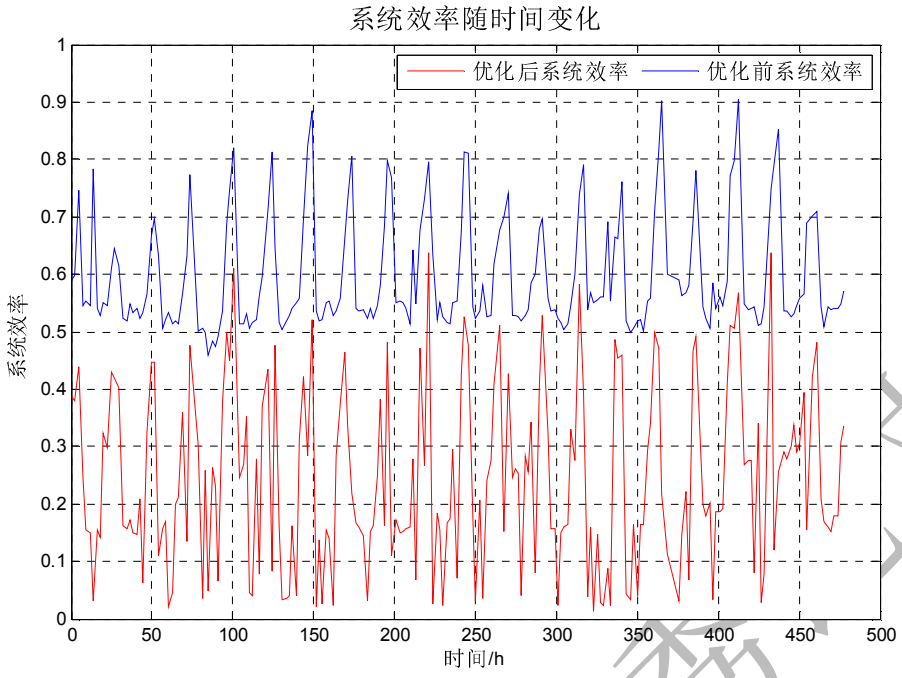


图 19 系统效率随时间变化图

红色为优化后的系统效率，蓝色是优化前的系统效率，优化后的系统耗电量和系统效率都是要低于优化之前的数据，且表中的各个设备的状态也是可见的。因此验证了模型的可靠性与合理性。

3. 结论

本文首先对数据进行预处理，将异常值和缺失值用均值进行修正，然后引入三均值的概念，绘制出干球温度、湿球温度、相对湿度、流入流出装置的水温差、总耗电量、冷却负载和系统效率以天为单位的三均值曲线图，对其趋势进行初步的判断，发现这些曲线基本上都具有日周期性。

为了对更加具体的数据进行趋势分析，我们挑出这几个变量的四个周的数据画出其以周为单位的趋势图和以日为单位的趋势图，发现这几个变量都具有日周期性，且干球温度与湿求温度之差最大出现在 15 时左右，相对湿度的最低值出现在 15 时左右，验证了我们之前的分析。我们为了对数据进行降维，因此我们对各个设备的开关状态进行动态聚类分析，得到 12 个设备的 4 个分类组合：冷却装置 2、冷水泵 2、冷凝水泵 2、冷却塔 1 为一类，冷却装置 3、冷水泵 4、冷凝水泵 3 为一类，冷却装置 1、冷水泵 1、冷凝装置 3 为一类，冷却塔 2、冷凝水泵 1 为一类。

我们将分成一类的几个设备的状态是相同的挑选出来作为应用于回归的数据。通过回归分析，我们得出冷凝水泵与冷凝水泵的转速和相对湿度都呈正相关，冷却塔功率与冷却塔风扇转速和相对湿度呈正相关，冷水泵功率与冷水泵转速和相对湿度呈正相关，

同理可以得到其它几个自变量与各个因变量之间的关系。此外，我们用剩余数据对所建立的回归模型进行了检验，验证了模型的准确性。

针对文中的三个优化问题，我们建立了以空调系统功率最小和空调系统总耗电量最小为目标函数的多目标非线性规划，根据决策变量的不同，我们也建立了不同的约束条件。

第二问中，以冷凝水泵、冷水泵、冷却装置和冷却塔的风扇转速为决策变量，我们考虑了设备转速的取值范围、设备状态参数取值、热稳定性范围为约束条件，并且应用 Java 编写程序实现模拟退火算法，以 12 月 28 日的数据为基础，求解出最优解与原始数据进行对比分析，得到总耗电量平均下降了 15.29%，系统效率平均下降了 19.73%。

第三问中，以冷凝水泵、冷水泵、冷却装置和冷却塔的开关状态为决策变量，以设备转速的取值范围、设备状态参数取值、热稳定性范围建立约束条件，应用 Java 编写程序实现模拟退火算法，以 12 月 28 日数据为基础进行优化求解，对比分析得到优化后总耗电量平均下降了 18.34%，系统效率平均下降了 18.98%。

第四问中，以冷凝水泵、冷水泵、冷却装置和冷却塔的开关状态和转速为决策变量，在第二、三问中的约束条件前提下，加入设备开关次数和时间间隔等的约束条件，求解的时候引入了最速下降法对模拟退火算法进行改进，提高了算法的效率。以附件中的数据为基础，进行优化求解，得出总耗电量平均下降了 41.55%，系统效率平均下降了 41.56%。

“泰迪杯”优秀论文

4. 参考文献

[1] J.E.Braun,et al. Metholologies for optimal control of chilled water systems without storage [J]. ASHRAE Transactions, 89-6-5(RP-539).

[2] J.MacArthur. A novel predictive strategy for cost-optimal control in buildings[J]. ASHRAE Transactions, 99(1), 1993.

[3] Work Tai Chan and Fu Wing Yu. How Chillers React to Building Loads [J]. ASHRAE Journal. 2004.

[4] 刘雪峰. 中央空调冷源系统变负荷运行控制机理与应用研究[D].华南理工大学,2012.

[5] 郭雪平. 基于能效优化中央空调节能控制系统的应用研究[D].西安石油大学,2016.

[6] 陈孜虎. 中央空调末端盘管及水系统变频节能运行控制研究和实现[D].重庆大学,2010.

[7] 罗坚. 地源热泵中央空调运行优化控制系统设计[D].湖南大学,2010.

[8] 张青. 中央空调系统节能运行控制方法研究[D].东南大学,2016.

[9] 周洪煜,陈孜虎,高鹏飞. 中央空调系统节能运行改造的控制策略与方案[J]. 计算机测量与控制,2009,(10):1977-1980.

[10] 李巧. 中央空调系统末端设备节能优化控制策略研究[D].广州大学,2015.

[11] 刘春蕾,孙勇,王利民,陈忠海. 中央空调系统节能运行控制的优化模型[J]. 低压电器,2008,(24):21-23+26.

[12] 赵波峰. 基于负荷预测的中央空调系统自适应优化节能控制[D].武汉科技大学,2012.

[13] 杨建明. 中央空调系统用电设备改变频控制节电运行[J]. 甘肃科技纵横,2008,(03):50+123.

[14] 吴万庆. 多 Agent 技术在中央空调节能控制中的应用研究[D].重庆大学,2007.

[15] 龚明启. 中央空调系统动态运行节能优化策略研究[D].广州大学,2006.

泰迪杯优秀论文