基于曲线相似度的供热过程评价方法

孙志伟**1**，冯海波**1**，王富全**2**，马永军**1**，董亮亮**1**

（1. 天津科技大学计算机科学与信息工程学院，天津 300457；2. 天津市南开区人民政府供热办公室，天津 300457）

**摘　要**：随着企业信息化的开展和节能减排的规划，供热行业需要从节能角度量化供热过程管理，在满足居民供热的情况下尽可能节约能源。目前没有一种合适的评价方法从节能角度来评价供热过程。本文在供热采集数据的基础上利用数据分析，针对锅炉出水温度曲线提出了一种供热过程评价方法。该方法通过分析出水温度曲线与目标温度曲线的关系，提出从曲线趋势变化、平移、伸缩三个维度衡量两个曲线的一致程度，首先计算各个维度的相似度，然后加权融合为一个一致性结果用以评价锅炉操作人员的供热过程。实验结果证明了该评价方法的有效性，为锅炉管理人员量化供热管理指标提供了参考依据，有效促进了节能减排的管理。

**关键词：**节能减排； 曲线相似度；平移；伸缩

**中图分类号： TP3-05          文献标志码：                   文章编号：1672-6510（0000）00-0000-00**

**Heating Effect Evaluation Method based on Curve Similarity**

Sun Zhiwei1，Feng Haibo1，Wang Fuquan2, Ma Yongjun1，Dong Liangliang1

(1. College of Computer Science and Information Engineering, Tianjin University of Science & Technology, Tianjin 300457, China; Heating Office of Nankai District Tianjin People's Government, Tianjin 300457, China)

**Abstract:** In order to solve the large amount of historical data of boiler system utilization rate is low, poor heating effect, in the analysis of the relationship between reference temperature curve and the water temperature curve proposed a based on similarity curve of heating effect evaluation method, and gives the trend, translation, expansion three attributes similarity calculation formula, and finally the weighted fusion as a result to the heating effect evaluation of the boiler system. The experimental results show that the evaluation method is effective, which provides a decision-making basis for the boiler control personnel, and effectively promote the energy-saving and emission reduction.

**Key words:** heating system；curve similarity；evaluation method

随着供热行业自动化水平的不断提高，“十三五”规划节能减排的进行，我国的城市供热过程在一定程度上实现了自动控制，供热品质得到改善，能源利用率得到提高。很多学者已经利用数据挖掘等技术对锅炉系统历史数据分析。孙群丽等对锅炉运行数据进行关联规则挖掘，提供了几组在不同负荷及外部条件下的最优运行方式与参数控制[2]；郑斌祥等通过建立宝钢能源数据仓库对时序数据挖掘的离群挖掘、相似性挖掘、规则挖掘和趋势挖掘等算法进行了系统深入的研究[3]；路海昌等通过对时间序列进行相空间重构，建立了基于支持向量回归的时序数据预测模型，从而实现对锅炉输出参数的预测[4]。但是，以上的分析都是如何去控制锅炉，没有对锅炉供热的结果进行分析，用户是否正确应用了这种分析，各供热部门都希望综合各种情况从节能角度量化考核供热过程。而目前大量的供热锅炉运行现场依然是粗放管理，没有合适的方法对供热过程进行量化对已有的大量运行数据缺乏深层次分析与利用，缺乏切实可靠的量化评价体系，难以满足管理者实时掌握考评状况和调整运行策略的需求[1]。因此研究供热过程的评价方法对推进节能减排、减少锅炉运行成本都具有重要意义。

本文根据专家建议或预测的供热目标温度曲线，首先解释目标温度曲线的由来并分析目标温度曲线与出水温度曲线差异的多种可能情况，然后将其分解为趋势变化、平移和伸缩三个属性相似度并给出各个属性的相应计算方法，并最后融合为一个一致性指标来评价锅炉系统供热结果。通过实验分析，证明了此评价方法的有效性，为供热锅炉系统的相关管理人员量化管理提供了一种参考依据，避免仅以耗气量来衡量供热行为，需要区分不同的供热过程，在满足供热户室内温度的情况下尽可能节能。

# 目标温度曲线与出水温度曲线分析

## 1.1 目标温度曲线

早期锅炉系统管理人员根据运行经验，会制定一个出水温度标准作为供热锅炉出水温度的参考，通常为室外温度的线性相关曲线，我们称之为目标温度曲线。而随着节能减排的规划，目标温度的定义需考虑各方面因素确定，主要包括天气条件（室外温度、风速、日照）、供热户不同时间段的需求，回水温度等。其中气象条件将直接影响用户的采暖需求，回水温度也反映了热量的利用情况，而供热公司也需要根据用户的作息规律、生活习惯、上班或在家的情况进行适度的调控，尽量节约能源。



式中t为时间，Tt、Wt、St、Pt分别为t时间的室外温度、风速、日照、偏移量，回水温度。目标温度曲线主要是根据专家经验综合以上（参考供热用户建筑）因素确定或对大量历史采集数据分析预测得到。

## 1.2 出水温度曲线与目标温度曲线相似度

将出水温度曲线（*A*）与目标温度曲线（*B*）进行一致性分析，可以得出此系统的供热效果，为锅炉管理人员对操作人员量化管理提供参考[5-6]。

两条曲线的一致性存在多种情况，如图1所示。其中图1-(a)所示，曲线B基本与曲线A相同，而曲线C的在11时后较曲线A的趋势有明显差异，表明曲线C基本没有按照曲线A进行相应的调整，我们认为两条曲线反映的供热调节趋势不一致；图1-(b)表现了曲线间的平移问题，在供热数据上，横坐标的平移代表了出水温度与目标温度调整的提前和延时情况，纵坐标的平移代表了出水温度与目标温度的温差情况，即供热温度相比目标温度偏高或偏低。其中曲线B在横坐标上较曲线A整体延后近两个小时，表明虽然趋势调整基本正确，但没有及时进行调整，曲线C在纵坐标上较曲线A高了近2摄氏度，表明供热温度偏高，造成了能源浪费；图1-(c)中，曲线B的最大值与最小值间横坐标差距相较于曲线A较小，说明该降温时间延后，而升温时间提前，造成了能源的浪费；而曲线C与曲线A相比，其最大值阈最小值间的纵坐标差距更小，在高温时未能达到用户需求，在低温时浪费热量。以上列出了曲线一致性的几种情况，每一种都存在相反的情况，并且实际的关系可能是这几种情况的综合体现。

综合上述分析，两条曲线间主要存在趋势变化、平移和伸缩三个属性上不同。因此，出水温度曲线及目标温度曲线的一致性评价问题可以分解为对这三种属性相似度的计算问题。在供热数据中，曲线间的差异也代表了不同的供热行为：趋势变化主要反映两者在整体形态上的一致性，即操作人员能否按照要求进行合理的调控，横向时间的偏移体现了温度调整的提前或延后，纵向温度的偏移体现了供热温度的偏高或偏低；伸缩体现了曲线在形态上调整的一致性。

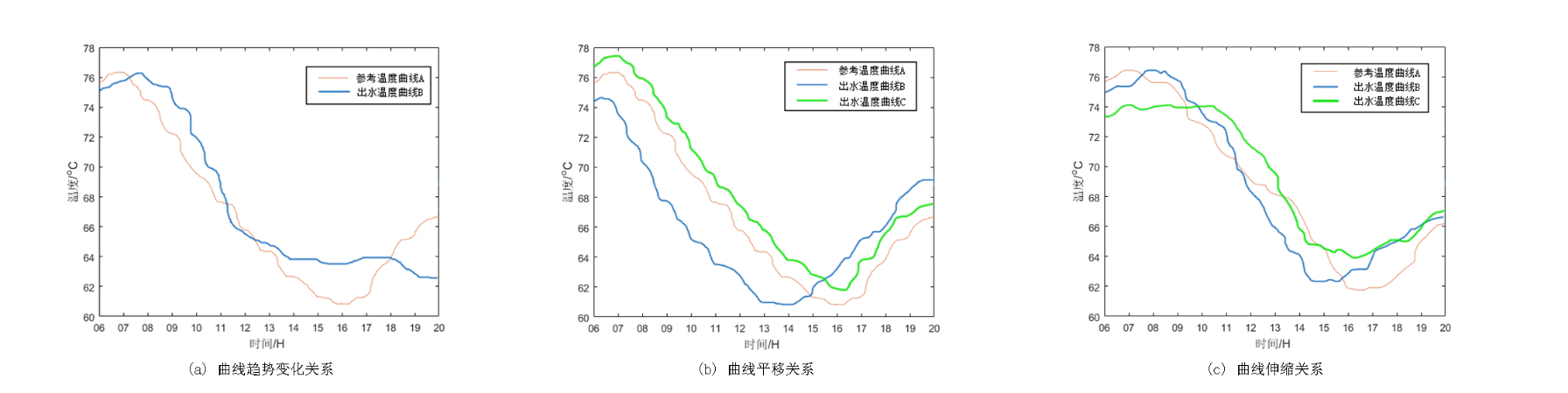
2 供热效果评价

图1 出水温度曲线与参考温度曲线关系图

## 2.1 出水温度曲线与目标温度曲线的趋势变化

两条曲线趋势变化相同，可以认为锅炉系统能综合考虑各种因素对出水温度做出合理的调整，使供热户在不同的条件下得到最合适的热量。对两条曲线趋势的分析，可以认为是曲线的相似性问题。

曲线的相似性测度一般有两种方法：距离测度法和相似性函数法[7]。相似性函数是用函数的方法来表征两曲线相似的程度，主要有夹角余弦和相关系数等方法，但由于供热数据受天气和地域影响较大，很难拟合成相应的曲线函数，而距离测度法主要有Euclidean距离、Minkowsky距离、Hausdorff距离、Frechet距离等[8]，其中，Hausdorff距离和Fréchet距离主要用来计算俩个点集间的相似性，但Hausdorff距离忽略了点集的时间序列问题，基于供热系统时序数据的特点，本文采用Fréchet距离作为曲线趋势变化属性相似度的度量，其优点在于充分考虑了曲线的连续性，非常适用于曲线间的相似性比较[9]。

Fréchet 距离由 M. Fréchet提出，描述了两质点分别沿着2条给定曲线以任意速度单向运动时，二者之间的最短距离。Axel Mosig和Michael Clausen曾将Fréchet 距离与变换群的交叉子集结合，应用到判别两条曲线的相似性上[10]，曹凯等引入Fréchet距离进行云规则推理，设计了一种智能地图匹配算法[11]。Eiter 和 Mannila在连续 Fréchet 距离的基础上提出了离散 Fréchet 距离[12]的定义 , 而朱洁等考虑了离散Fréchet距离的关键特征峰值点，减少了算法的复杂度并将其运用到了手写签名验证上[13]，收到了一定的效果。

离散Fréchet距离定义如下：

（1） 给定1个有n个至高点的多边形链P=<*P1,P2,P3,...,Pn*>，1个沿着P的*k*步，分割P的峰值点成为*k*个不相交的非空子集{Pi}*i=1,…,k*，使得Pi=<Pni-1+1,…,Pni>和1=n0<n1<…nk=n

（2） 给定2个多边形链A=<a1,…,am>，B=<b1,…,bn>，1个沿着A和B的组合步是1个沿着A的*k*步{*Ai*}i=1,…,k和1个沿着B的*k*步{Bi}i=1,…,k组成，使得对于1≤*i*≤*k*，且*Ai,Bi*中有1个恰好包含1个至高点。

（3） 1个沿着链A和B的组合步*W*={(*Ai,Bi*)}的花费(cost)为：



其中dist()为a,b间的欧氏距离，则链A和B间的离散Fréchet距离为：



## 2.2 出水温度曲线与目标温度曲线的平移问题

出水温度曲线与目标温度曲线的平移表现为两者横纵坐标的差异，在温度曲线中，横坐标的度量为时间，表现为出水温度调整的提前或延后问题；而纵坐标的度量为温度，体现在两条曲线的温差问题，具体为锅炉系统是否能按需达到预定温度以及在不需要较多热量时降低负荷节约能源。

### 2.2.1时间差异

时间差异定义为两条曲线的n个同一维度上特征点间的时间差均值。而特征点的确定会对度量结果产生较大影响，考虑到供热锅炉数据的特点，在同一维度上很难找到成对的特征点[14]。这里将两条曲线的峰值点作为我们的特征点对，分别计算出两条曲线的最大值，取其时间的差值作为延时，然后计算次大值的差异，依次计算出一段时间内的差异集合，最后将这些差异的均值作为差异度量：



式中i为特征点的个数，A为目标温度，B为出水温度，则为目标温度曲线在第i大值的时间。

### 2.2.2 温度差异

温度差异简称温差，主要表现在供热温度较低时是否能满足采暖需求，反之是否发生能源浪费。这里将分两方面考虑，曲线的最大值差和最小值差。

两条曲线最大值处的温差，能够判断供热效果，出水温度是否能按需达到采暖需求，为供热用户提供足够的热量。目标温度较高时通常意味着用户在家或者气象条件不能提供较多自然热量，所以需要锅炉系统高负荷运转提供充足热量，也可以一定程度上减少客诉。

两条曲线最小值处的温差，能够判断供热锅炉系统是否节约能源。目标温度较低时通常意味着，用户家中无人，或者室外温度等气象条件能提供较多自然热量，故锅炉系统需要降低运行负荷减少热量，节约能源降低运行成本。

综合上述两个方面将两条曲线的温差定义如下：



式中n为特征点的个数，A为目标温度，B为出水温度，为目标温度曲线在第i个最小值点的温度。

## 2.3 出水温度曲线与目标温度曲线的伸缩问题

出水温度曲线与目标温度曲线的伸缩问题，在其横坐标上表现为供热行为在时间上是否按统一节奏对锅炉系统进行调控，对各因素的变化是否做出时间一致的操作；在其纵坐标上表现为锅炉系统对温度控制的灵敏度，在高低温转换时可以及时达到预期温度。

借助离差标准化的思想，最差可以完整的表现整体的数据跨度[15]，将横纵坐标的最差比作为两条曲线的伸缩比，能较好的反映数据整体的特点，对其横向和纵向的伸缩比的计算方式如式（5）和（6）所示。





式中为目标温度曲线最大值的横坐标，为出水温度曲线最小值的横坐标。为目标温度曲线最大值的横坐标，为出水温度曲线最小值的横坐标。

## 2.4 评价结果

将三种属性相似度共五个度量加权融合为出水温度曲线和目标温度曲线的相似度：



式中ω1、ω2、ω3、ω4、ω5分别为趋势、横向平移、纵向平移、横向伸缩、纵向伸缩属性相似度的权值，ω1+ω2+ω3+ω4+ω5=1，可通过数据统计及最小二乘法得出[13]。、、、、分别为五个属性相似度的阈值，现根据历史数据分析，给出默认值为40,550,25,1.5,1.5，实际使用时应根据历史数据分析得到。

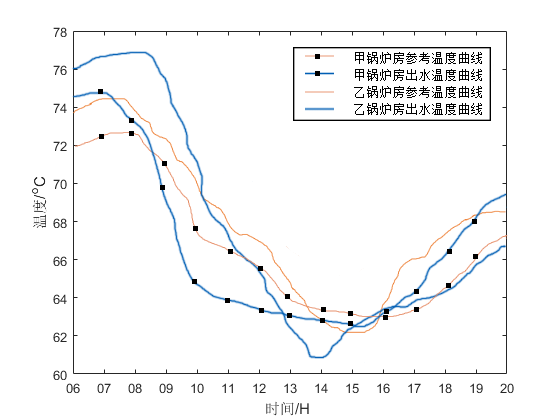
本文通过分析出水温度曲线与目标温度曲线间的相似度，分别给出曲线的趋势、平移、伸缩三种属性相似度的计算方式，并将其加权融合为一个评价结果，用来评价锅炉系统的供热效果。

# 实验分析

本文利用天津地区供热公司提供的历史数据，对2015-2016年供热季的数据进行了分析，依据采暖节能需求，供热公司在夜间会将锅炉系统的出水温度维持在较低的水平，所以只对每天6:00至20:00的数据进行分析。由于自动采集系统的稳定性等原因会导致缺失值，错误值等离群点的出现，基于时序数据的特点，采用拉格朗日插值法来对数据进行预处理，以此作为实验的初始数据集。

现依据（1）式将数据计算出目标温度，其中某日的甲锅炉房出水温度曲线及其目标温度曲线的对比如图2所示。

图2 甲、乙锅炉房某日出水温度曲线及其参考温度曲线



目标温度曲线在上午7时左右达到最大值，为一天最冷的时间，用户在此时间段在家居多，所以需要提供较高热量。随着室外温度、日照等气象条件的提高，出水温度将逐渐降低，而在下午4时开始，供热公司考虑到更多用户将会回到家中，且室外温度和日照逐渐降低，所以调控锅炉出水温度增加偏移量，为用户提供更多热量。

图3 甲、乙锅炉房某日五个属性相似度对比

然后将甲锅炉房的出水温度曲线与乙锅炉房的出水温度曲线分别与目标温度曲线进行相似度计算，依据（2）-（5）式计算出三种属性相似度共五个度量，部分结果如表1和表2所示：

图4 甲、乙锅炉房一周相似度对比图

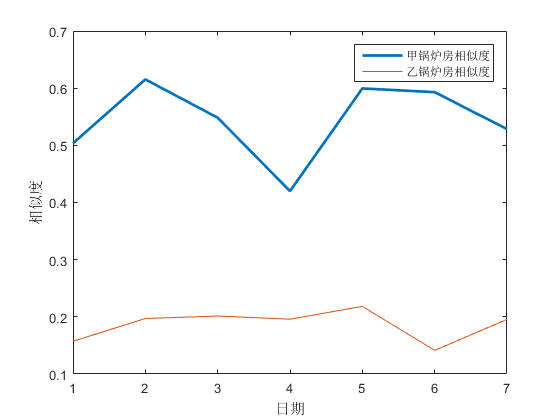


表3 甲、乙锅炉房相似度一周计算结果

表2 乙锅炉房五个属性相似度一周计算结果

表1 甲锅炉房五个属性相似度一周计算结果

|  |
| --- |
| No. F DX DY EX EY |
| 1 14.72 316.38 17.27 0.12 1.36 2 31.15 556.14 5.86 1.15 0.27 3 33.00 211.49 11.63 0.92 0.39 4 13.35 217.59 16.07 0.98 0.21 5 13.14 257.45 20.49 1.04 1.33 6 26.74 170.68 22.06 0.31 1.27 7 18.70 252.85 12.58 0.96 0.93 |

由表1和表2可知，乙锅炉房在趋势变化(F)、平移（DX、DY）以及伸缩（EX、EY）上较甲锅炉房有更好的结果，对数据进行离差标准化后，直观的对比如图3所示。

对数据库中的所有数据进行统计分析，并结合锅炉供热系统相关管控人员历史经验利用最小二乘法，确定（6）式中的复合度量参数值分别为：

ω1 = 0.3,ω2 =0.175,ω3 =0.175,ω4=0.175,ω5 = 0.175

|  |
| --- |
| No. F DX DY EX EY |
| 1 15.22 23.71 2.76 0.03 0.11 2 09.55 36.31 6.79 0.02 0.55 3 11.65 42.53 6.55 0.26 0.21 4 10.63 64.59 1.19 0.38 0.37 5 14.53 25.21 4.98 0.46 0.11 6 07.25 22.85 3.40 0.06 0.42 7 12.32 45.83 5.85 0.23 0.18 |

依据（6）式将五个属性相似度加权融合为一个评价结果，其部分结果如表3以及图4所示：

|  |
| --- |
| 1 2 3 4 5 6 7 |
| 甲 0.50 0.62 0.55 0.42 0.60 0.60 0.53 乙 0.16 0.20 0.20 0.20 0.22 0.14 0.20 |

图4可以看出乙锅炉房在本文提出的评价模型上表现更为优秀，查询数据可知，甲锅炉房的单位耗气量为10.6 m3/m2，乙锅炉房的单位耗气量为9.5 m3/m2。实际运行中乙锅炉房采取了自动控制系统来调节出水温度，在考虑天气条件的同时还按照用户的作息规律在相应的时间段设置了温度偏移量，所以乙锅炉房的出水温度曲线在与目标温度曲线的相似度上更为接近，而甲锅炉房只是按照室外温度进行人工控制，管理较为粗放，其时效性表现较差。而其实际意义是乙锅炉房一定程度上减少了资源浪费，节约了能源，并为用户提供了较好的供热效果，这将对锅炉供热系统的相关管控人员提供最大的参考依据，对节能减排具有重大意义。

# 结论

为了提高锅炉系统的供热效果，达到节能减排的目的，本文提出了一个锅炉供热效果的评价方法。对锅炉房的出水温度曲线与目标温度曲线进行相似度分析，分解为趋势变化、平移和伸缩三种属性相似度，给出每个属性相似度的计算方式并加权融合为一个相似度度量。利用天津地区某供热公司的供热数据，对2015年供热季的数据进行分析，通过本文提出的供热效果评价方法将甲、乙锅炉房进行了实验，结果表明，本文提出的评价方法可以将人工控制和自动控制的锅炉房区分出来，并判断出较好的供热效果，这为供热锅炉系统的相关人员提供很好的决策依据，对锅炉系统节能减排具有较大意义。此外，锅炉系统的回水温度也是供热效果的评价因素之一，将出水和回水温度结合起来构建评价模型将会在今后的工作中继续完善。

**参考文献**

[1] 江亿, 彭琛, 胡姗. 中国建筑能耗的分类[J]. 建设科技, 2015(14):22-26.

[2] 张珊. 供热锅炉绩效评价及优化系统的研究[D]. 大连海事大学, 2013.

[3] 郑斌祥. 基于数据仓库的时序数据挖掘研究[D]. 上海交通大学, 2002.

[4] 路昌海, 刘贵松, 张明琤. 基于支持向量回归的锅炉出水温度时间序列预测[J]. 区域供热, 2014(6):18-22.

[5] 徐志, 李栋, 刘挺,等. 微博用户的相似性度量及其应用[J]. 计算机学报, 2014, 37(1):207-218.

[6] 高兴.基于特征信息的测井曲线相似度算法研究与应用[D]. 东北石油大学, 2013.

[7] 张宇, 刘雨东, 计钊. 向量相似度测度方法[J]. 声学技术, 2009, 28(4):532-536.

[8] 郑丽萍, 李光耀, 梁永全,等. 本体中概念相似度的计算[J]. 计算机工程与应用, 2006, 42(30):25-27.

[9] HELMUT ALT, MICHAEL GODAU. COMPUTING THE FRÉCHET DISTANCE BETWEEN TWO POLYGONAL CURVES[J]. International Journal of Computational Geometry & Applications, 2011, 5(1):75-91.

[10] Axel Mosig, Michael Clausen. Approximately matching polygonal curves with respect to the Fréchet distance ☆[M]// STACS 2001. Springer Berlin Heidelberg, 2010:63-74.

[11] 曹凯, 唐进君, 刘汝成. 基于Fréchet距离准则的智能地图匹配算法[J]. 计算机工程与应用, 2007, 43(28):223-226.

[12] Eiter T, Mannila H. Computing discrete Fréchet distance. See Also[J]. See Also, 1994, 64(3):636-637.

[13] 朱洁, 黄樟灿, 彭晓琳. 基于离散Fréchet距离的判别曲线相似性的算法[J]. 武汉大学学报:理学版, 2009, 55(2):227-232.

[14] 李光, 吴祈宗. 基于结论一致的综合评价数据标准化研究[J]. 数学的实践与认识, 2011, 41(3):72-77.

[15] 蔡启林, 寿晓峰. 供暖热负荷廷时曲线及其应用[J]. 区域供热, 1991(2):1-10.