基于曲线相似度的供热过程评价

孙志伟**1**，冯海波**1**，王富全**2**，单渊博**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. 宏源热力，天津 300457）

**摘　要**：燃煤燃气供热是我国北方主要的供暖方式，而对供热锅炉历史数据的分析将有利于提高锅炉的供热效果。用历史气象数据计算一条参考曲线，分析实际出水温度与参考曲线的关系，建立了一个基于离散Fréchet距离并结合延时和温差的供热效果评价模型。结果表明此模型对于锅炉系统的供热效果有较好的评价结果，可以为锅炉管控人员提供决策支持和分析依据。

**关键词：**曲线相似度；离散Fréchet距离；延时度量

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

**Evaluation Model of Heat Supply Effect based on Discrete Fréchet Distance**

Sun Zhiwei1，Feng Haibo1，Wang Fuquan2, Shan Yuanbo1，Dong Liangliang1

(1. College of Computer Science and Information Engineering, Tianjin University of Science & Technology, Tianjin 300457, China)

**Abstract:** Heating boiler system is a nonlinear complex system, there are many problems in the practical application. Through the analysis of the boiler system in the running process of the characteristics of the data, combined with real-time acquisition of the meteorological conditions, established a evaluation model based on fréChet distance and increased delay metric to optimize fréChet for lack of time series data processing. Finally, and cluster analysis were combined to verify. The results show that the evaluation model for boiler system output parameters (water temperature) have better evaluation results can provide decision support and basis for the relevant personnel of the boiler control.

**Key words:** boiler control；data mining；Fréchet；cluster analysis；

我国是一个能源消费大国，同时也在能源匮乏的国家之列，目前从环保等各方面考虑，在“十二五”热电联产“煤改气”的过程中处理燃气锅炉系统时，供热信息化上已经有了很大进展，包括各种供热数据的采集，以及在此基础上的数据统计，能耗分析。所以如何根据已有的历史数据建立评价模型并进行数据分析，辅助锅炉管控人员决策，使锅炉系统提供较好的供热效果，进而减少过度的能源浪费，最终达到节能减排的目的，成为现实问题。

诸多专家学者已经对锅炉运行参数进行数据分析，利用数据挖掘等技术展开了一系列工作。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]。

本文根据专家建议或预测的供热目标参考曲线，利用离散Fréchet距离分析与热量的相似度，并从供热的及时性以及供热的供回水温差上考虑两条曲线间的关系，建立燃气锅炉供热效果的评价模型。通过实验分析，得出了多种控制模式的差异化对比结果，为燃气供热锅炉系统的相关管控人员提供了决策依据。

# 供热量与参考曲线分析

## 1.1 出水温度参考曲线

锅炉系统管控人员根据多年的运行经验，会制定一个标准作为锅炉出水温度的参考，通常为室外温度的线性相关曲线，我们称之为参考曲线。实际中会有更多的天气因素对供热用户的室温以及人对温度的感觉产生影响，而供热公司也需要根据用户的作息规律、生活习惯对供热行为作出调整。

为了更好的评价供热过程，有必要定义一个参考目标供热曲线，并把供热过程与参考目标供热曲先对照，评价供热行为。并给出可视化的结果。当然，条件许可还要及时根据供热户家里的室内温度的反馈。

参考目标供热曲线主要是根据专家经验（参考供热户室内温度的反馈）或根据统计分析数据进行预测得到。包括的因素主要有：室外温度、风速、日照、时间段偏移等，并依据供热公司多年的运行经验采用不同的偏移量作为锅炉出水温度参考曲线:



式中t为一天中的某个时刻，Tt、Wt、St、Pt分别为t时刻的室外温度、风速、日照和时间段偏移，a, b, c, d分别为四种条件因素的偏移值，结合供热地区，供热用户特点综合分析设定。**（时间段偏移，偏移是否听不明白，是否改成权值）**

## 1.2 出水温度及其参考曲线关系分析

将供热行喂鱼参考目标供热曲线进行对比，可以得出此系统的供热效果，为锅炉管控人员提供参考。

两个曲线之间的关系包含多种情况。最好的情况是两条曲线完全重合，及供热行为与预期行为完全一致。但这种情况概率较低。下面列出几种典型的情况。分别如图1、图2所示，（图1说明横向变化，图2说明纵向，图3给出两个方向都有变化，具体看怎么表述）

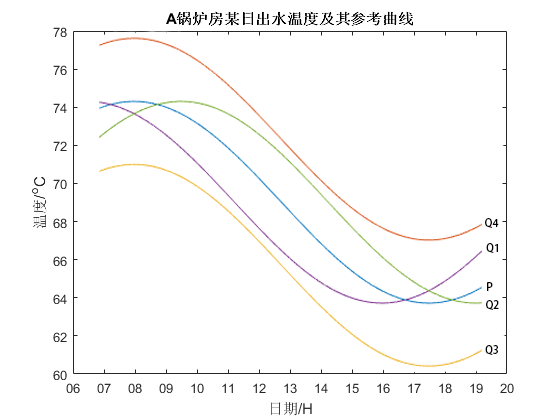


图1 参考曲线与出水温度关系图

两个曲线的对比我们借鉴曲线相似度来分析，目前对曲线相似度的研究及应用很多，列举相应的文献并说明。

相似度。两条曲线相似度相同，可以认为锅炉系统能按照天气变化、时间变化作出合理的调整，出水温度可以使用户在不同的天气下得到最合适的室温。对两条曲线相似度的分析，相似性较高的可以被认为提供了较好的供热效果。

延时。在天气变化后，锅炉系统才能进行调整并最终反馈在出水温度上，此过程中将会产生一定的延时。Q1和Q2的出水温度分别快于和慢于参考曲线P，显然延时较低更能满足用户的需求，因此两条曲线的时间延时也将作为供热效果的度量。

温差。锅炉系统能否达到预期的温度，以及在不需要提供太多热量时进行降温从而节约运行成本。Q3在极大值处低于参考曲线P，达不到用户所需，可能会引起用户投诉；Q4在极小值处高于参考曲线P，此时并不需要提供过多热量，增加了锅炉运行成本。通过分析两条曲线在极值点的差值可以为锅炉系统在的热效果提供依据，因此温差也是供热效果得重要度量。

# 供热效果评价

在这里我们考虑。。。。。。，首先采用。。。衡量曲线的相似度，然后从。。。。和。。。方面进一步分析曲线不相似的具体原因。

最后复合为一个度量为锅炉系统的供热效果作出评价。

## 2.1曲线相似性及离散Fréchet距离

相似度是用以度量两组数据变化趋势相似程度的一个数值度量。目前 Hausdorff 距离作为距离的测度被广泛应用于判断两个点集间的相似性，但Hausdorff并未考虑曲线的时序性，而时序在锅炉系统的数据中又是非常重要的参考因素，所以采用Fréchet距离来研究两个曲线相似性[5]。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[9]。

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[6]，曹凯等引入Frechet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[10]的定义 , 而朱洁等考虑了离散Frechet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[8]，收到了一定的效果。

离散Fréchet距离定义如下：

（1） 给定1个有n个峰值点的多边形链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个峰值点。

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



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



## 2.2 延时

延时定义为两条曲线的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。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极大值，取其时间的差值作为延时，然后计算次大值的延时，依次计算出一段时间内的延时集合，最后将这些延时的均值作为延时度量：



式中i为特征点的个数，P为参考曲线，Q为实际出水温度，则为参考曲线在第i个极大值的时刻。

### 2.3 温差

温差主要表现在参考目标温度较高时是否能满足采暖需求，温度较低时是否发生能源浪费。这里将分两方面考虑，曲线的极大值差和极小值差。

两条曲线极大值处的温差，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。参考曲线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。

两条曲线极小值处的温差，能够判断供热锅炉系统是否节约能源。参考曲线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i为特征点的个数，P为参考曲线，Q为实际出水温度，则为参考曲线在第i个极值点的温度。

## 2.4评价结果

表1 天津地区供热公司历史数据（部分）

将曲线的相似度、延时以及温差合并为一个复合度量来作为我们供热效果的评价模型:



式中α、β、γ分别为相似度、延时、温差的权值，ε1、ε2、ε3 分别为相似度、延时、温差的阈值，权重和阈值的确定需要结合具体锅炉房站点通过相应数据进行统计和分析得出。

本文在离散Fréchet距离的基础上加入延时及温差度量，作为供热锅炉出水温度和参考曲线的评价模型，将为锅炉系统的供热效果提供理论依据。

# 仿真实验

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，部分数据如表1所示。

现依据（1）式将数据计算出参考曲线，其中某日的A锅炉房及其参考曲线的对比如图2所示。

|  |
| --- |
| 出水温度/°C 回水温度/°C 气温/°C 日照/Lux |
| 65.88 40.87 6.00 522 66.13 40.86 6.00 522 66.25 40.88 5.90 516 66.38 40.88 5.80 519 66.63 40.88 5.70 516 66.75 40.88 5.60 521 66.88 40.88 5.60 526 66.86 40.87 5.60 522 66.76 40.88 5.60 521 66.88 40.88 5.70 516 |

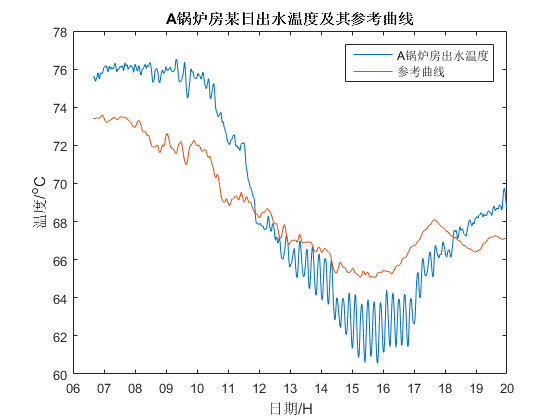


图2 A锅炉房某日出水温度及其参考曲线

参考曲线在上午7时达到极大值，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。

然后将A站的出水温度曲线与B站的出水温度曲线分别与参考曲线进行分析，依据（2）式、（3）式、（4）式计算出相应的离散Fréchet距离、延时及温差度量，其结果如图3所示：**（三维散点图或者雷达图）**

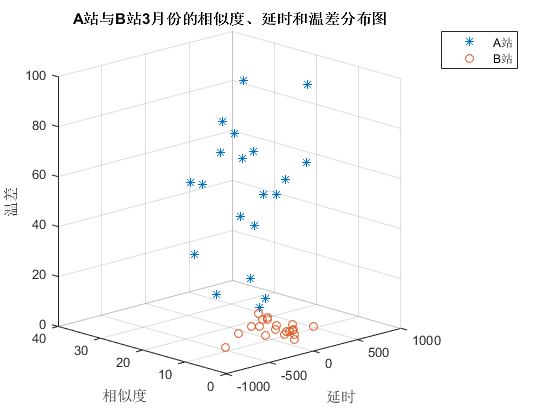


表2 A站与B站3月份耗气量数据表

图3 A站与B站3月份的相似度、延时和温差分布图

由此图可知，B站在离散Fréchet距离、延时以及温差上较A站有更好的表现，在对数据库中的所有数据进行分析，并通过大量的实验验证以及锅炉供热系统相关管控人员历史经验，确定（7）式中的复合度量参数值分别为：

α = 0.6, β = 0.2, γ = 0.2, ε1 = 200, ε2 = 400, ε3 = 10

依据（7）式将上文所述离散Fréchet距离和延时度量以及温差度量复合成为一个评价结果，如图4所示：

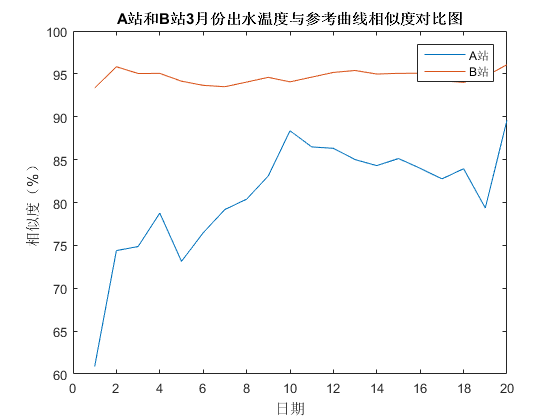
此图可以看出B站整体上较A站相比在我们的评价模型上表现更为优秀，取出一天的相应数据作为特征点加以验证，如表2所示**（单位问题）**

图4 A站与B站3月份出水温度与参考曲线评价结果对比图

|  |
| --- |
| 锅炉房 总耗气量/m3 供热面积/m2 单位耗气量/°C |
| A 3551056.00 369901.66 9.6 B 561134.00 53441.33 10.5 |

根据数据可知，A站的单位耗气量大于B站。实际运行中B站采取了自动控制系统来调节出水温度，在考虑气象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B站在与参考曲线的相似度、延时及温差上更为接近，而A站只是按照室外温度进行人工控制，其时效性表现较差。而其实际意义是B站一定程度上减少了资源浪费，节约了能源，这将对锅炉供热系统的相关管控人员提供极大的参考依据，对节约资源浪费具有重大意义。

增加评价结果的表述，评价值小，说明具体的原因，以某种可视化方式表达。根据情况最少要有两组结果。

# 结论

为了提高锅炉系统的供热效果，达到节能减排的目的，本文建立了一个锅炉供热效果的评价模型。将离散Fréchet距离和延时以及温差度量作为评价因子，分析锅炉系统出水温度曲线和参考曲线的关系，可以很好将不同供热锅炉系统按其供热效果区分出来。利用天津地区某供热公司提供的历史数据，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模型将A站和B站进行了实验，结果表明，本文提出的评价模型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] HELMUT ALT, MICHAEL GODAU. COMPUTING THE FRÉCHET DISTANCE BETWEEN TWO POLYGONAL CURVES[J]. International Journal of Computational Geometry & Applications, 1995, 5(1):75-91.

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

[7] Holm L, Sander C. Mapping the protein universe.[J]. Science, 1996, 273(5275):595-602.

[8] 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.

[9] Fréchet M M. Sur quelques points du calcul fonctionnel[J]. Rendiconti Del Circolo Matematico Di Palermo, 1906, 22(1):1-72.

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

[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] 江亿, 唐孝炎, 倪维斗,等. 北京PM2.5与冬季采暖热源的关系及治理措施[J]. 中国能源, 2014, 36(1):7-13.