基于曲线相似度的供热过程评价方法

孙志伟1，冯海波1，马永军1，王富全2，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：随着国家节能减排的规划，供热企业需要从节能角度对供热过程进行量化管理。而无线传感器等物联网技术在供热行业中的应用为量化管理提供了大量基础数据。如何分析这些基础数据为量化管理提供参考成为企业需要解决的问题。本文在供热采集数据的基础上针对锅炉出水温度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。分别计算各个维度的相似度，然后加权融合为一致性度量，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，促进了节能减排。

关键词：物联网；节能减排；曲线相似度；平移；伸缩

中图分类号：TP399 文献标志码：A 文章编号：1672-6510(0000)00-0000-00

**Evaluation Method of Heating Process Based on Curve Similarity**

SUN Zhiwei1, FENG Haibo1, MA Yongjun1, WANG Fuquan2, DONG Liangliang1

(1. College of Computer Science and Information Engineering, Tianjin University of Science & Technology, Tianjin 300457, China; 2. Heating Office of Nankai District Tianjin People's Government, Tianjin 300457, China)

**Abstract:** Along with the released national energy saving and emission reduction plan, the heat-supply enterprise need the quantitative management on the heating process from the energy saving perspective. And the application of the technology of internet of things on the heating industry such as the wireless sensors has also provided plenty of basic data for such quantitative management.While how to analyze these basic data to provide the reference for quantitative management becomes the main problem that the enterprise need to deal with it.Based on the collected heating data, this paper proposes an evaluation method on heating process by aiming to the temperature of boiler output water. Through analyzing the relationship between the temperature curve of output water and targets temperature curve, this method aims to measure the consistent degree of the two curves from the three dimensions such as the changes, translation and elasticity of curve trend.Calculate separately the similarity of each dimension and then integrate them into a consistent measurement in weighting, so as to appraise the heating process of the boiler operator. The experimental result proves the effectiveness of such evaluation method, and provide the reference basis for the quantitative management index on the heating process of boiler managing personnel, to promote the energy saving and emission reduction accordingly.

**Key words:** internet of things; energy saving and emission reduction; curve similarity; translation; elasticity

随着供热行业自动化水平的不断提高，“煤改气”和“十三五”规划节能减排的进行，我国的城市供热过程基本上实现了自动控制，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[1]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；路海昌等[2]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；岳晓忠[3]采用后向反馈BP神经网络理论和关联规则算法的数据挖掘方法，对锅炉实时运行数据进行分析，从而建立锅炉运行模型。以上的分析都是如何去优化控制锅炉，但实际操作人员由于各种原因较难掌握这些方法的使用，而供热公司也没有合适的方法对现场人员供热过程量化分析，即缺乏切实可靠的量化评价方法，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[4]。因此研究供热过程的评价方法对推进节能减排、降低运行成本都具有重要意义。

本文首先解释目标温度曲线的由来并分析目标温度曲线与实际出水温度曲线之间不同的多种可能情况，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度量来评价供热过程。通过实验分析，证明了此评价方法的有效性，为相关管理人员量化管理提供了一种参考依据，避免仅以燃气、水、电等能耗来衡量供热行为，而是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。

1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，根据室外温度会制定一个出水温度标准作为供热锅炉出水温度的参考，通常与室外温度为线性相关。在此基础上一段连续时间的出水温度就构成了目标温度曲线。随着节能减排的规划以及物联网技术的应用，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)、供热用户不同时间段的需求、回水温度、出水提前量等。其中天气条件将直接影响供热用户的采暖需求，回水温度反映了热量的利用情况，而供热公司也需要根据供热用户的作息规律、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，出水提前量指热水热交换后到用户家里的时间。可见，这是一个比较复杂的过程，需要考虑较多因素，主要因素如式(1)所示。

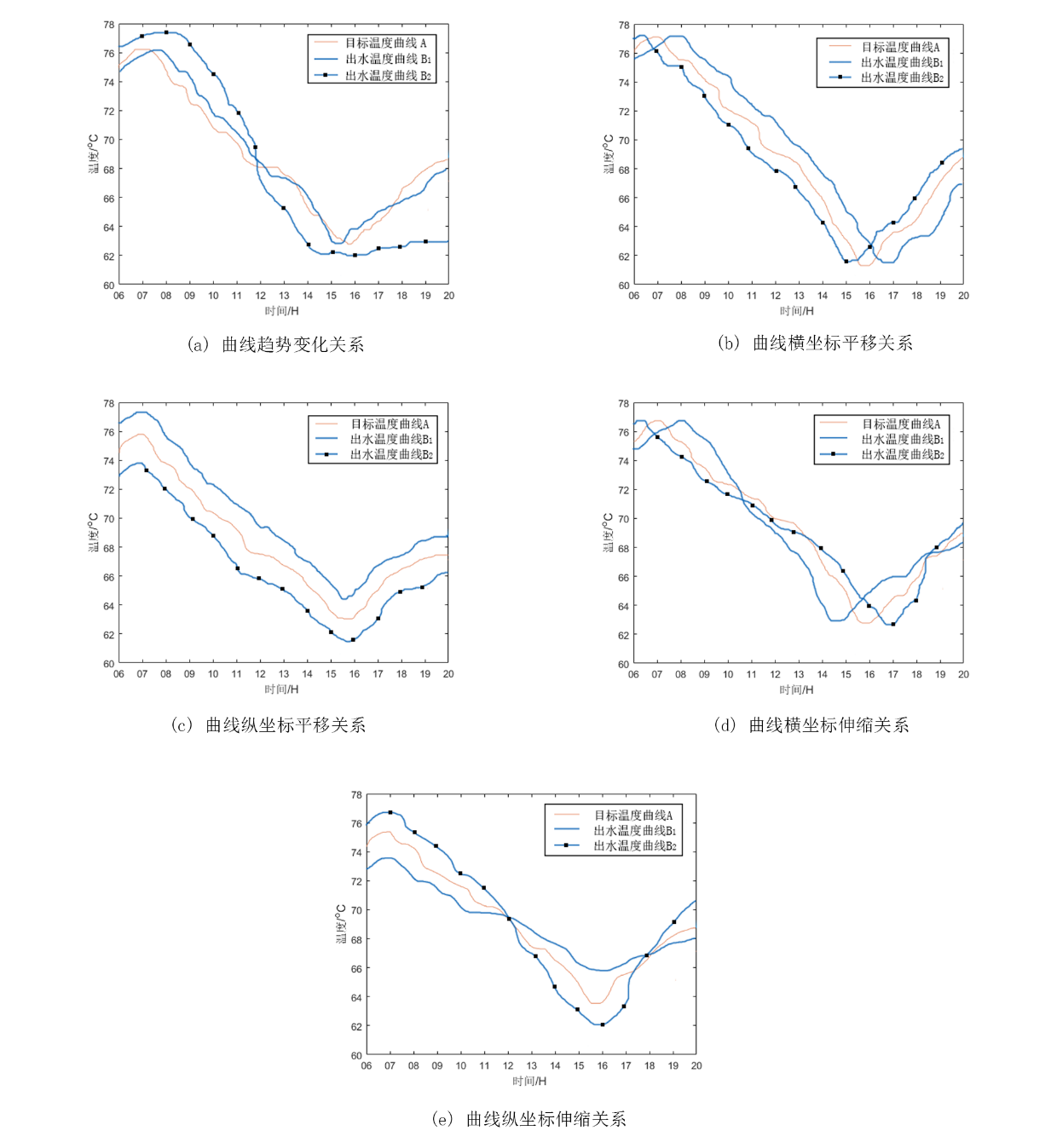
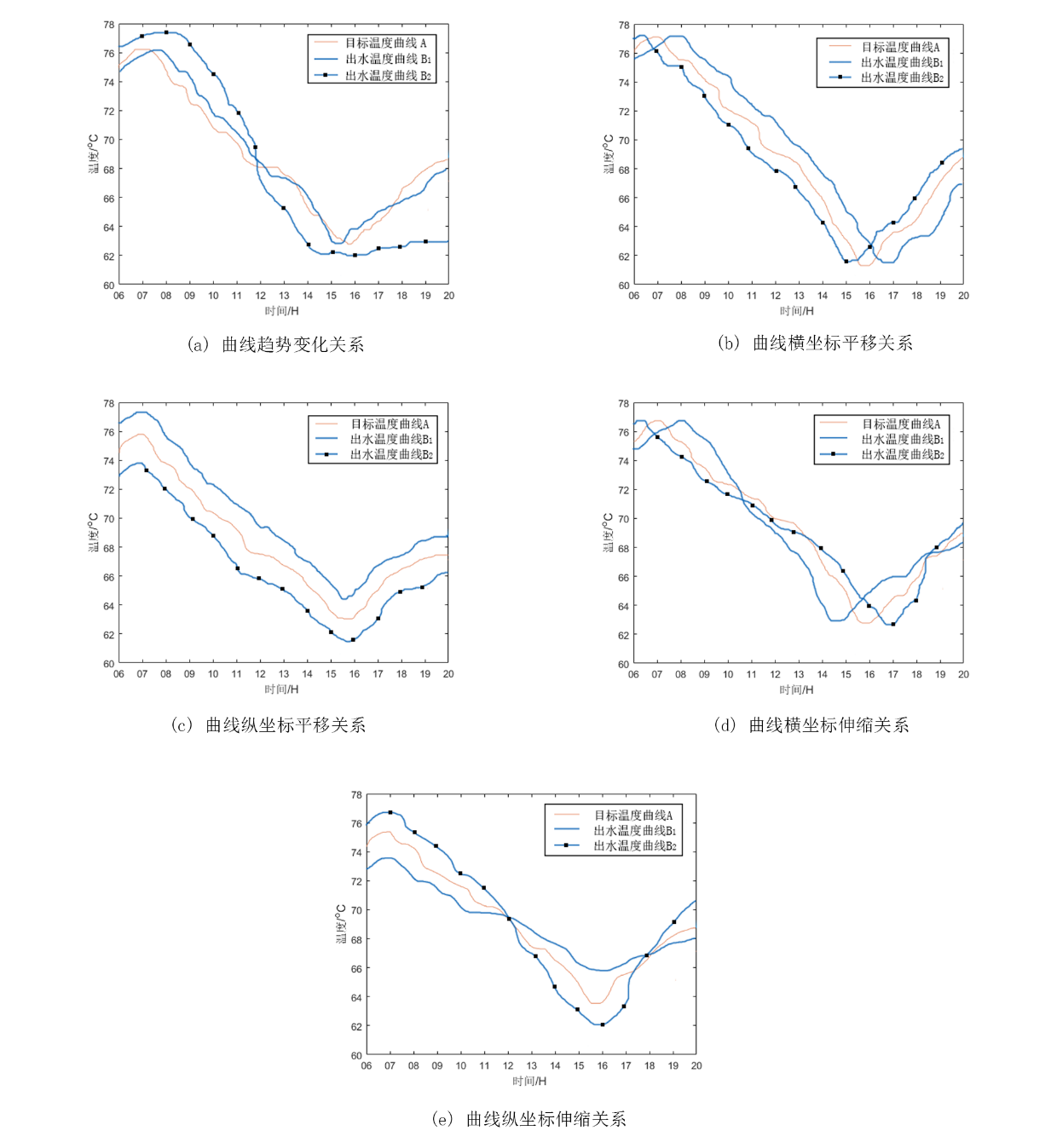
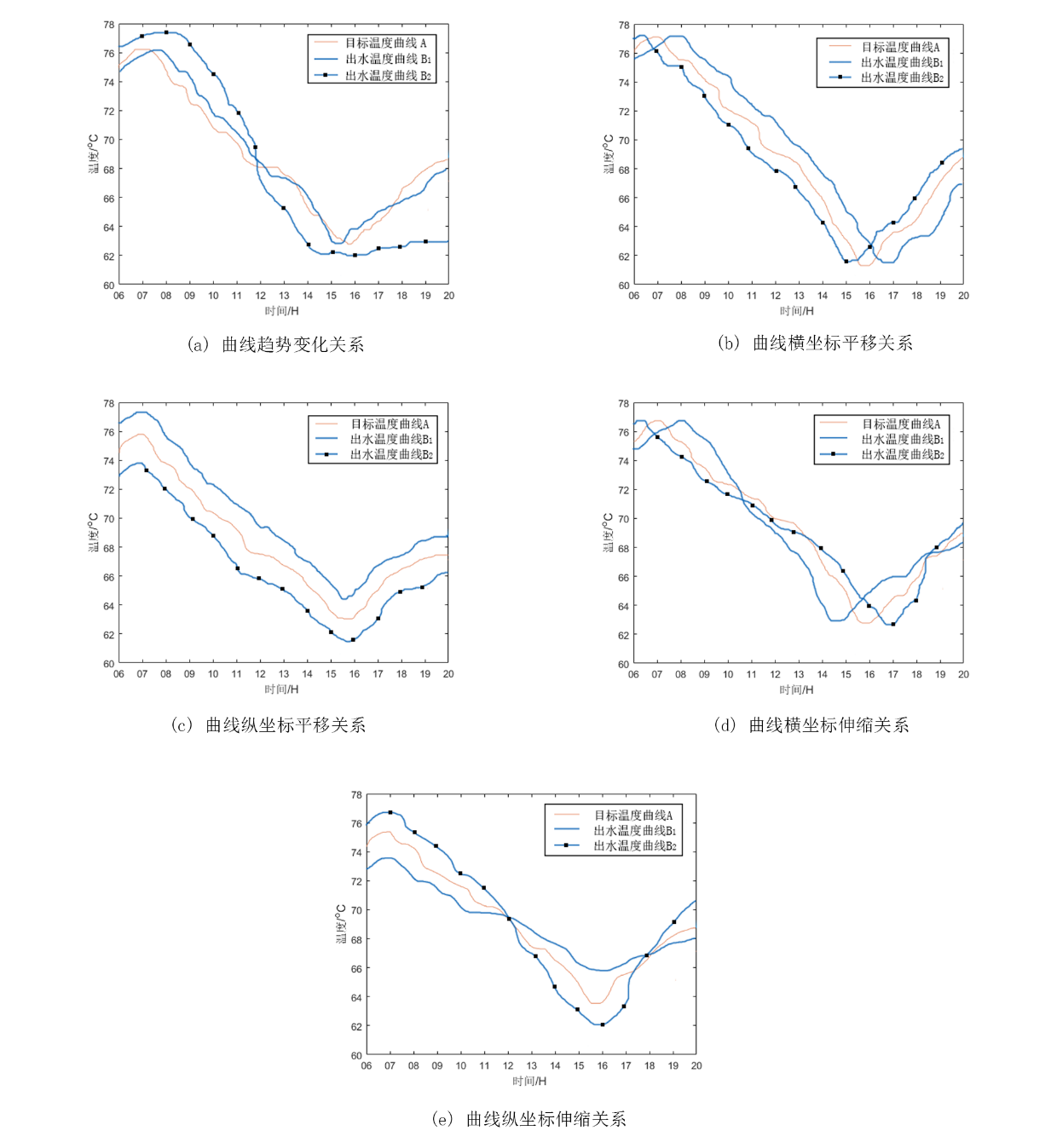
式中：为时间；分别为时间的室外温度、风速、日照、偏移量、回水温度、出水提前时间。

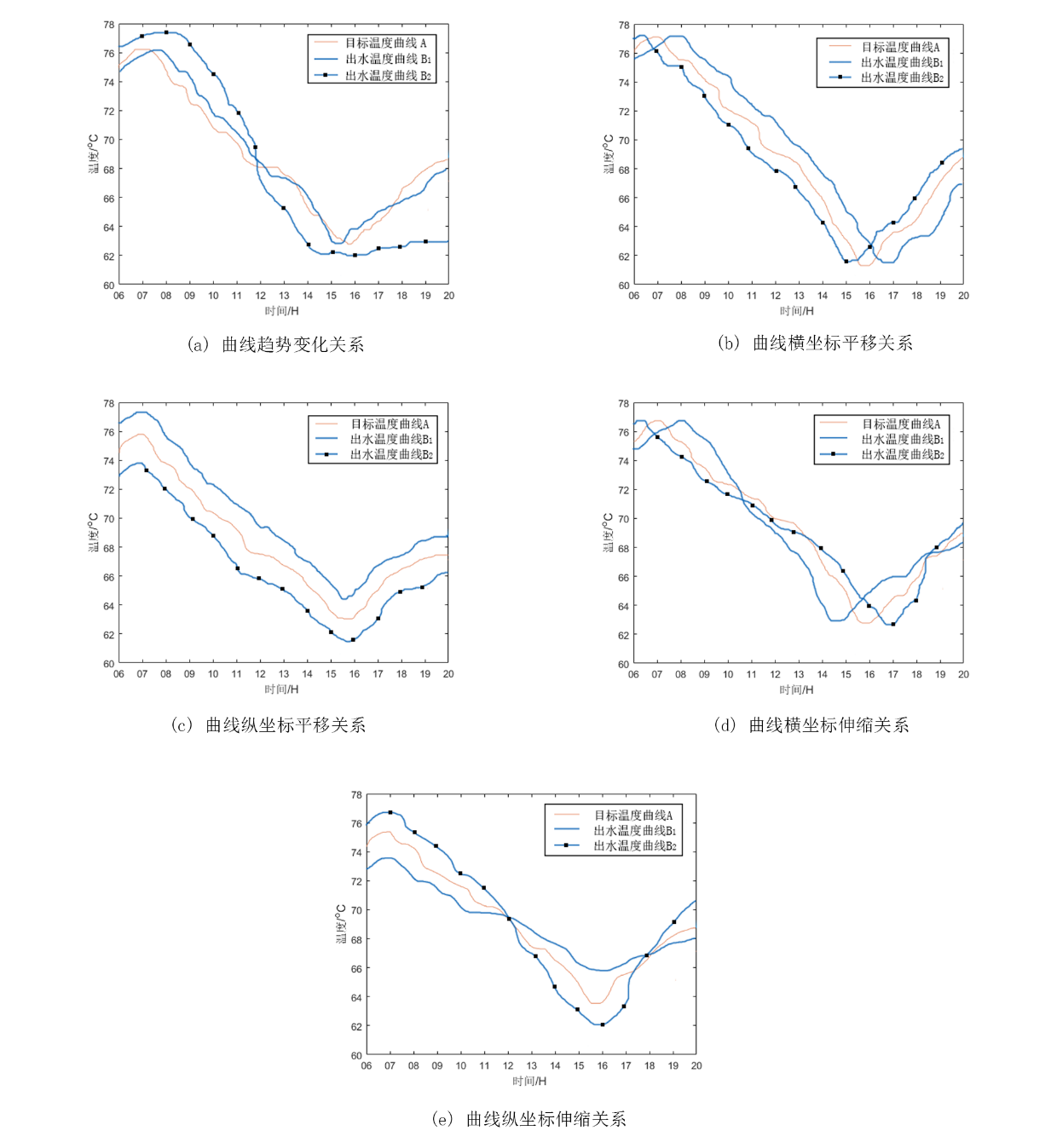
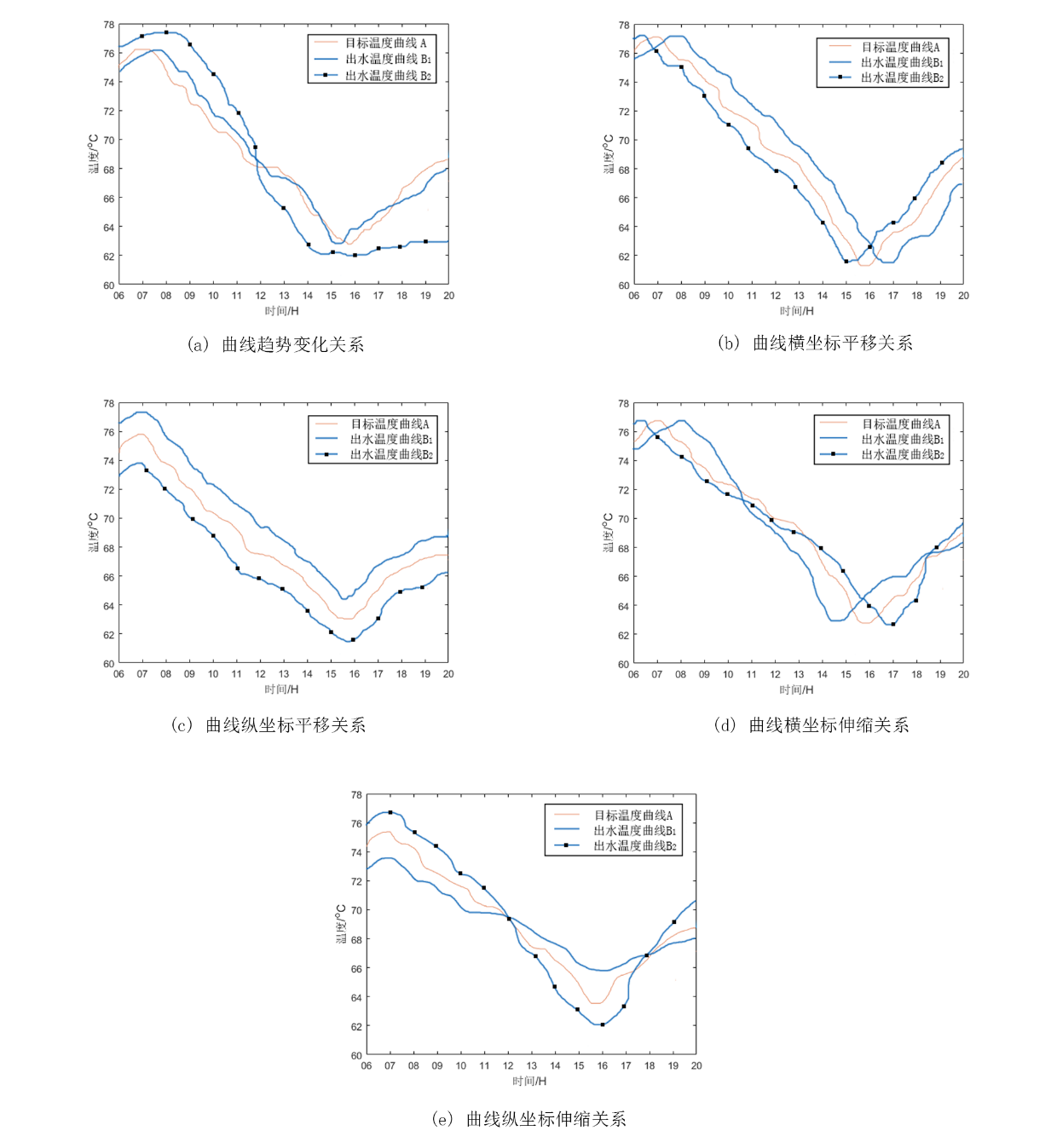
目标温度曲线主要是根据专家经验综合以上因素并参考供热用户建筑计算或对大量历史采集数据分析预测得到。

1.2 目标温度曲线与出水温度曲线相似度

将目标温度曲线(A)与出水温度曲线(B)进行一致性分析，可以分析出锅炉房供热过程是否严格按照要求供热，及满足节能要求，为锅炉管理人员对操作人员量化管理提供参考。

两条曲线的一致性存在多种情况，如图1所示。其中图1(a)中曲线B1基本与曲线A相同，而曲线B2的在12时后较曲线A的趋势有明显差异，表明曲线B2基本没有按照曲线A进行相应的调整，即两条曲线反映的供热调节趋势不一致；图1(b)表现了曲线间的横向平移问题，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B1在横坐标上较曲线A整体提前，而曲线B2在横坐标上较曲线A整体延后，表明虽然趋势调整基本正确，但调节的及时性存在问题；图1(c)表现了曲线间的纵向平移问题，曲线B1在纵坐标上较曲线A低了近2摄氏度，表明供热温度偏低，未能达到供热户供暖需求，而曲线B2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(d)中，曲线B1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B2的最大值与最小值间横坐标差距大于曲线A，说明降温时间提前而升温时间延后，未能达到供热户需求，容易引发客诉；图1(e)中，曲线B1与曲线A相比，其最大值与最小值间的纵坐标差距更小，在高温时未能达到供热户需求，在低温时浪费热量，而曲线B2最大值与最小值间的纵坐标差距更大，在高温时浪费热量，在低温时未能达到供热户需求。



图1 出水温度曲线与目标温度曲线关系

**Fig.1 The relationship between output water temperature curve and target temperature curve**

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，出水温度曲线与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；伸缩体现了曲线在升温和降温调整的不同。

2 供热效果评价

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[5]。

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[6]。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、Fréchet距离等[7]，其中，Hausdorff距离和Fréchet距离主要用来计算两个点集间的相似性，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[8]。

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。Mosig等[9]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，曹凯等[10]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。Eiter等[11]在连续Fréchet距离的基础上提出了离散Fréchet距离的定义，而朱洁等[12]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，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，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，使得和。

(2)给定2个多边形链，，1个沿着*A*和*B*的组合步是1个沿着*A*的*k*步和1个沿着*B*的*k*步组成的，使得对于，中有1个恰好包含1个至高点。

(3)1个沿着链*A*和*B*的组合步的花费(cost)为

其中*dist*()为*a*与*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[13]。这里先获得两条曲线的最值，然后按照最值将每条曲线单独划分为多个单调区间，对于单调增区间最小值为初始特征点对，否则最大值为初始特征点对，然后在各区间根据初始特征点对的类型计算下一个特征点对，取其时间的差值作为时间差异。依次计算出每个时间段的差异集合，最后将这些差异的均值作为差异度量。

式中：为特征点的个数，*m*为每个单调区间的点数，A为目标温度，B为出水温度，则为目标温度曲线在第*j*个区间温度为第*i*大值的时间，供热控制较差时可在合理时间段取值。

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为

式中：A为目标温度，B为出水温度，为目标温度曲线最小值点的温度。

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锅炉系统在时间上是否按统一节奏对锅炉系统进行调控，对天气情况的变化是否做出与时间一致的操作；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，最值差可以完整地表现整体的数据跨度[14]，将横纵坐标的最值差比作为两条曲线的伸缩比，能较好地反映数据整体的特点，对其横向和纵向伸缩比的计算方式如式(5)和式(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纵坐标。

2.4 评价结果

将三种属性相似度的共五个度量加权融合为出水温度曲线和目标温度曲线的一致性度量：

式中、、、、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，，可通过数据统计及最小二乘法得出[15]。、、、、分别为五个属性相似度的阈值。

本文通过分析出水温度曲线与目标温度曲线间的一致性，分别给出曲线的趋势变化、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，用来对锅炉供热过程进行评价。

3 实验分析

利用天津某供热公司提供的采集数据，对2015—2016供热季中两个供热站的数据进行了分析，依据供热习惯及节能需求，目前很多供热单位晚十点半左右开始维持在某一较低温度，早四五点左右开始升温，即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。由于物联网系统的稳定性等原因会导致缺失值、异常值等离群点的出现，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。

依据式(1)专家根据运行经验确定一天的目标温度调整规律，其中某日的甲锅炉房出水温度曲线及其温度曲线的对比如图2所示。从图中可以看出，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，供热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，所以升高锅炉出水温度，为用户提供更多热量。对数据预处理后，将甲锅炉房的出水温度曲线和乙锅炉房出水温度曲线分别与目标温度曲线进行一致性分析，依据式(2)—式(6)计算出三种属性相似度共五个度量，计算某七天的结果如表1和表2所示。

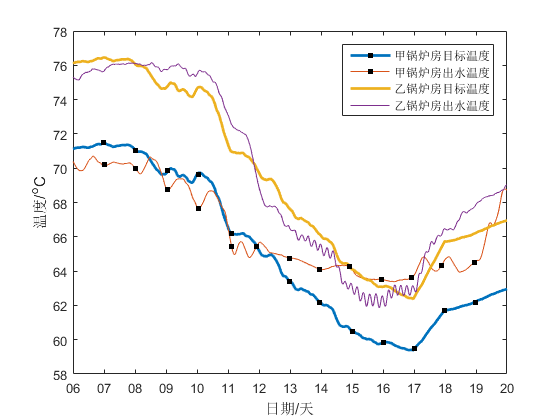


图2 甲、乙锅炉房某日出水温度及其目标温度

**Fig. 2 Output water temperature and target temperature of A and B boiler one day**

表1甲锅炉房五个属性相似度七天计算结果

**Tab. 1 Calculation results of five attribute similarities of A boiler roomin seven days**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| No. | *F* |  |  |  |  |
| 1 | 7.51 | 16.00 | -3.54 | 0.87 | -1.86 |
| 2 | 7.83 | 41.40 | -2.62 | 0.97 | -1.26 |
| 3 | 4.84 | 27.00 | -2.38 | -0.94 | -0.62 |
| 4 | 11.61 | 11.00 | -5.68 | 0.26 | -1.45 |
| 5 | 8.99 | 17.82 | -4.30 | -0.51 | -1.37 |
| 6 | 5.70 | 23.73 | -1.77 | 0.97 | -1.84 |
| 7 | 5.04 | 48.40 | -0.49 | 0.6 | -2.80 |

表2乙锅炉房五个属性相似度七天计算结果

**Tab. 2 Calculation results of five attribute similarities of B boiler room in seven days**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| No. | *F* |  |  |  |  |
| 1 | 2.46 | 2.24 | -0.21 | -0.27 | -0.26 |
| 2 | 2.07 | 12.59 | 0.41 | 0.36 | 0.02 |
| 3 | 3.64 | 1.71 | -0.69 | -0.25 | 0.29 |
| 4 | 3.85 | -2 | -1.06 | -0.08 | 0.24 |
| 5 | 4.52 | 6.2 | -0.37 | -0.19 | 0.23 |
| 6 | 4.03 | 6.06 | -0.76 | -0.32 | 0.21 |
| 7 | 2.97 | 10.15 | 0.32 | -0.74 | 0.06 |

对数据库中的所有数据进行统计分析，利用最小二乘法并结合锅炉供热系统相关管理人员历史经验，确定(7)式中的复合度量参数值分别为：。依据式(7)将五个属性相似度加权融合为一个评价结果，经过对历史数据分析，给出阈值12,42,6,1,3，其一周计算结果如表3和图3所示：

表3 甲、乙锅炉房一致性七天计算结果

**Tab. 3 Calculation results of the consistency of A and B boiler room in seven days**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 甲 | 0.68 | 0.75 | 0.54 | 0.73 | 0.67 | 0.62 | 0.65 |
| 乙 | 0.16 | 0.20 | 0.21 | 0.20 | 0.23 | 0.25 | 0.28 |

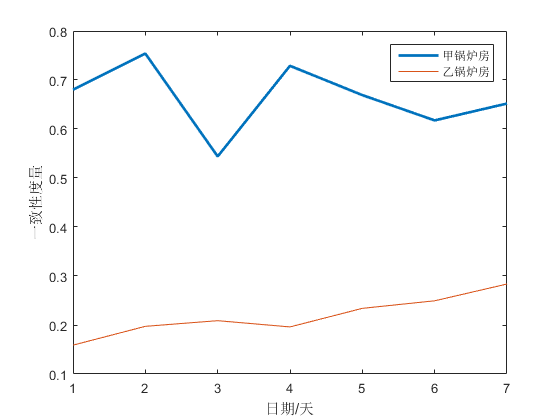


图3 甲、乙锅炉房一致性七天对比

**Fig.3 The consistency of A and B boiler room inseven days**

图3可以看出：依据本评价方法，乙锅炉房的出水温度与目标温度一致性更高，查询原始采集数据可知，甲锅炉房的单位耗气量为10.6，乙锅炉房的单位耗气量为9.2，因此算法是有效的。

由图3可知，第二天和第四天的一致性度量基本相同，进一步分析其原因，对数据按照式(8)进行离差标准化后，直观的对比图如图4和图5所示。

甲锅炉房第二天的纵向伸缩基本一致，趋势变化和纵向平移相较于第四天数据更优，横向平移和横向伸缩相对较差，表现为时间延时较小能及时为用户提供热量，但随目标温度的调控较差。乙锅炉房第二天的趋势变化、纵向平移和纵向伸缩相较于第四天数据更优，而横向平移和横向伸缩相对较差，供热上表现为能为用户提供足够热量，随目标温度合理调控，但其时间延时相对较大。乙锅炉房五个度量皆优于甲锅炉房，供热上表现为乙锅炉房可以按照目标温度进行及时的适度调控，其时间延时较低，相较于甲锅炉房，能在高温时达到供热户需求，低温时能及时降温节约能源。

实际运行中乙锅炉房工作质量较高，根据天气、用户生活习惯、回水温度等各种因素积极调整运行参数，采取自动控制系统来调节出水温度；而甲锅炉房只是按照室外温度进行人工控制，没有精细化供热，其时效性表现较差。所以乙锅炉房的出水温度曲线在与目标温度曲线的一致性上更为接近。两个锅炉房供热户的投诉都很少，则以上数据说明在保证用户室内温度的情况下，乙锅炉房在一定程度上减少了煤气消耗，节约了能源。

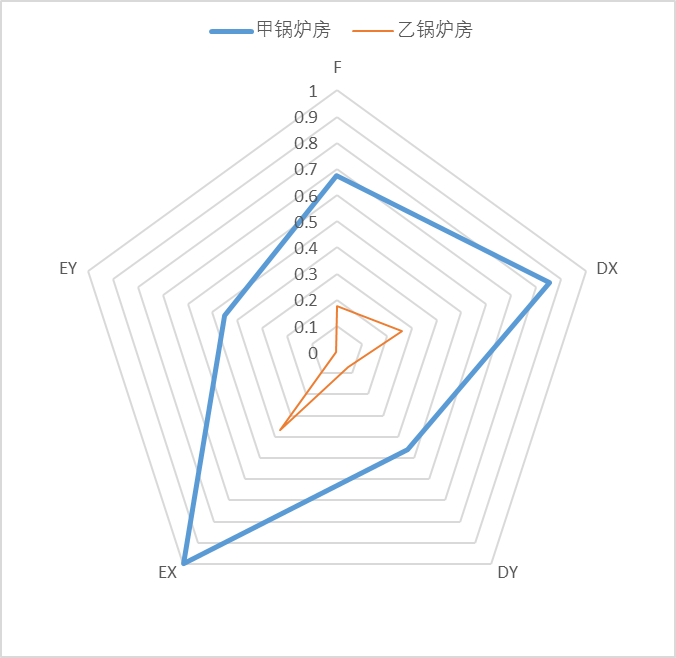


图4 甲、乙锅炉房第二天属性相似度对比

**Fig. 4 The attribute similarities comparison of A and Bboiler room on second day**

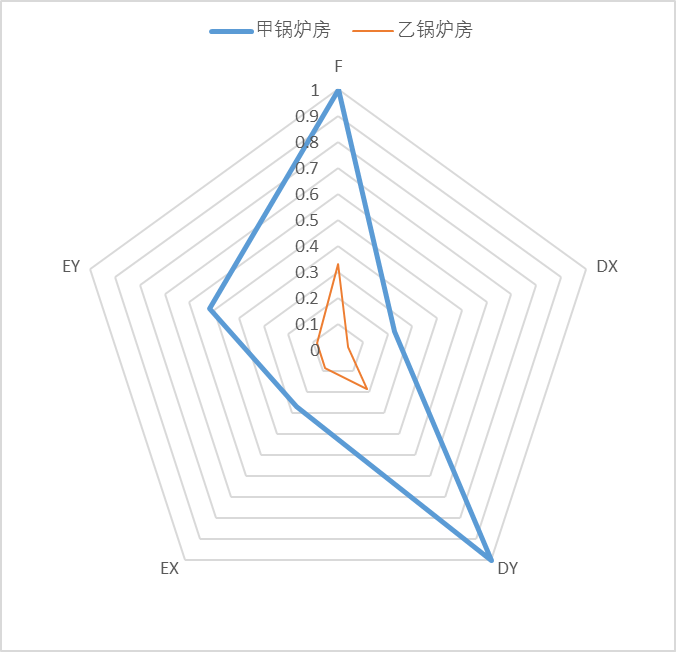


图5 甲、乙锅炉房第四天属性相似度对比

**Fig. 5 The attribute similarities comparison of A and B boiler room on fourth day**

4 结 论

为了更好地量化管理，达到满足供热的前提下节能减排的目的，本文提出了一个锅炉供热过程的评价方法。对锅炉房的出水温度曲线与目标温度曲线进行一致性分析，分解为趋势变化、平移和伸缩三种属性相似度分析其一致性，给出每个属性相似度的计算方式并加权融合为一个一致性度量。然后利用天津某供热公司的供热数据，对2015—2016供热季的部分数据进行分析，通过本文提出的供热行为评价方法将甲、乙锅炉房进行了对比分析，结果表明，本文提出的评价方法可以较好的区分不同的供热行为，为锅炉供热系统的相关管理人员提供量化考核及其物联网系统参数调整的参考依据，对节能减排，精细化供热具有重大意义。在以后的研究工作将在此基础上，将对实时采集数据进行分析，为及时调整供热行为提供实时建议和参考。

参考文献

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

[2] 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.

[3] 岳孝忠. 基于锅炉运行优化的数据挖掘平台研究与实现[D]. 西安：电子科技大学, 2012.

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

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

[6] 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.

[7] 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.

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

[9] Mosig A, Clausen M. Approximately matching polygonal curves with respect to the Fréchet distance[J]. Computational Geometry, 2005, 30(2): 113-127.

[10] 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.

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

[12] 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.

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

[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]王福昌, 曹慧荣, 朱红霞. 经典最小二乘与全最小二乘法及其参数估计[J]. 统计与决策, 2009(1):16-17.