

doi:10.11835/j.issn.1674-4764.2018.04.011



锅炉热水供热系统能耗机器学习诊断模型

邬棋帆^a, 于军琪^a, 郑佳蕾^b

(西安建筑科技大学 a. 信息与控制工程学院; b. 土木工程学院, 西安 710055)

摘要: 供热系统能耗诊断是一个难点。提出一种基于机器学习算法的能耗诊断标准模型结构, 通过聚类或分类算法, 从采集的诊断数据中筛选出节能特性较好的运行数据, 基于回归模型建立能耗诊断模型对被诊断数据进行诊断。研究发现: 1) 经 K-means 聚类筛选数据并基于贝叶斯正则化训练的人工神经网络建立拟合模型, R 值分别达到了 0.975 6、0.970 5、0.921 4 和 0.910 1, 模型拟合度较高; 2) 模型经过 3 个被诊断数据集验证, 经过筛选的数据建立诊断模型, 节能率指标分别 10.7%、17% 和 4%, 累积误差指标达到了 -149 498.67、-86 526、-4 052.27 kW, 诊断效果优于未经聚类的数据建立的模型; 3) 对诊断结果进行分析, 发现供热系统二次换热端热水流量变频节能控制措施节能效率较低, 一次供热端热水流量人工调节是造成能耗过高的主要原因。这种数据建模诊断的方式是基于输入、输出变量之间的物理响应关系而不受数据时间特性的影响。

关键词: 机器学习; 能耗诊断; 供热系统; 人工神经网络

中图分类号: TU111.195.2; TP183 **文献标志码:** A **文章编号:** 1674-4764(2018)04-0071-10

Energy consumption diagnosis methodology model of boiler hot water heating system

Wu Qifan^a, Yu Junqi^a, Zheng Jiale^b

(a. School of Information and Control Engineering; b. School of Civil Engineering,
Xi'an University of Architecture and Technology, Xi'an 710055, P. R. China)

Abstract: To solve the energy diagnosed problem of boiler hot water heat supply, a energy consumption diagnosed method based on machine learning algorithm was proposed, firstly filtrating the data which has better energy-conservation performance from all data based on clustering or classification method. Then based on the regression model, the informant data had been tested has been test. though Through the four case study, these conclusion conclusions can be gained: 1) The R value of model which is built by artificial neutral neural network(ANN), which was trained by Bayesian regularization method based on the data

收稿日期: 2017-07-11

基金项目: 教育部归国留学科研基金(教司留[2014]1685); 陕西科技厅科技国际合作项目(2014kw17); 安徽省教育厅自然科学基金研究项目(KJ2016A814)

作者简介: 邬棋帆(1992-), 男, 主要从事智能建筑研究, E-mail: wuqifan_0854@163.com。

于军琪(通信作者), 男, 教授, 博士生导师, E-mail: junqiyu@126.com。

Received: 2017-07-11

Foundation item: The Scientific Research Foundation for Returned Scholars from Overseas, Ministry of Education (No. [2014]1685); The Science and Technology International Cooperation Project of Shaanxi Science and Technology Department (No. 2014kw17); The Natural Science Research Project of Anhui Education Department (No. KJ2016A814)

Author brief: Wu Qifan(1992-), main research interests: intelligent building, E-mail: wuqifan_0854@163.com.

Yu Junqi(corresponding author), professor, doctoral supervisor, E-mail: junqiyu@126.com.

clustered by K-means algorithm was is up to the 0.976, 0.970 5, 0.921 4, 0.910 1; 2) though the test by the three data gather Model validated with 3 diagnosed datasets, the energy conservation ratio were are 10.7%, 17%, 4%, the accumulation error has been is up to the $-149\,498.67$, $-86\,526$, $-4\,052.27$ kW, the effect of new model is better than before; 3) the artificial control of first heating supply is the mainly reason, which cased the high heating energy. The model based on the physical response between input and output variable, which has higher robustness in time series can be widely employed in energy consumption diagnosed of boiler hot water supply system, and by the developing of data technology, the model based on the data machine learning can supply some idea ideas for the similar system.

Keywords: machine learning; energy consumption diagnosed; heating supply system; artificial neural network

建筑能耗占社会总能耗的 25%~30%^[1],在建筑能耗组成中,供热、制冷能耗占比达 50%以上^[2]。中国北方地区大多都采取市政热水集中供热的方式^[3],而南方地区主要是混合型供热方式^[4],其主要形式包括锅炉热水二次换热^[5]、市政蒸汽^[6]以及空调供热等。对比常见制冷供热方式制冷供热能效比,由于锅炉热水二次换热供热能效比较低^[7],并且运行成本也较高,因此,对于夏天采用中央空调制冷,冬天采取自烧锅炉热水二次换热系统供热的建筑来说,对锅炉进行能耗诊断对降低建筑能耗和能源成本有着重要意义。

从数据获取方式的角度来说,能耗诊断主要分为线上和线下诊断两种模式^[8]。线上诊断主要基于数据动态性能自适应的筛选并建立诊断模型,本质上是一个动态数据筛选和建模的过程^[9],线下诊断主要是基于已有数据的模型辨识,本质上是基于过去数据对未来或者其他数据集模型预测诊断的过程^[10]。数据处理方法主要包括基于能耗指标^[11]、基于数据挖掘^[12]、基于统计学^[13]以及基于能耗软件模拟^[14]。然而,这些方法都只针对能耗数据而并未考虑到实际工况变化对能耗的影响,因此,都存在一定的局限性。

机器学习是一个集合的概念,是对数据内在本质的一个挖掘,是把有效信息从大量的数据集中萃取出来的过程^[15];是把已有算法与实际数据相结合的产物,也是大数据背景下基于计算机强大计算能力对算法的一种再创新和应用^[16]。随着计算机科学的飞速发展,机器学习越来越多地应用到人工智能^[17]、数据挖掘^[18]、模式识别^[19]、图像处理^[20]等领域,在实际工程中得到越来越多的应用^[21]。文献[22-23]分别讨论了机器学习算法在电力系统能耗诊断中的应用。本文利用机器学习算法建立针对锅炉二次换热供热系统的能耗诊断模型,并做相应的

案例研究,为机器学习算法在锅炉热水节能诊断中的应用提供一个标准模型结构。

1 能耗诊断模型

锅炉热水供热系统是锅炉烧热水经过二次换热把能量从一次供热端传递到用户末端的过程。从锅炉热水供热系统节能的定义上,在满足热负荷需求的情况下,尽可能减少能耗是节能优化的目标。然而,影响热负荷的因素有很多,除了室外气候条件,还包括建筑物本体的体形架构,建筑物的外围护结构等。这些影响因子之间所满足的物理关系是非线性的,无论是基于机理建模还是基于一般非线性优化的方法都难达到诊断的目的。基于机器学习的能耗诊断模型是一种线下诊断模型,诊断的原理如图 1 所示。由于锅炉热水二次换热供热系统是一个多变量复杂系统,因此,诊断数据的质量对模型建立以及能耗诊断效果起决定性作用,理想情况下更希望用节能数据对不节能数据进行诊断,这样诊断的结果才能有一定的实用价值;其次,变量的选取也会影响模型拟合,倘若所选变量不完全,模型拟合度较低(即 R 值低),诊断本身就没有意义,因而,变量的选取也是数据质量的一个层面。因此,诊断模型建立主要包括:变量选取、数据采集、判别与筛选以及模型的建立。

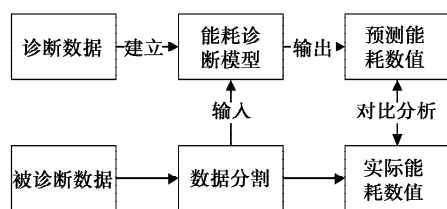


图 1 线下诊断模型

Fig. 1 Offline diagnostic model

变量选取是对系统初步基于原理的定性分析的

过程,一般选取的原则为:

- 1)系统运行与能耗相关性较大的作为输入变量,与能耗直接相关的作为输出变量。
- 2)对有节能控制的系统(如气候补偿器、变频器)应把影响节能效果的被控制变量(如热水流量)作为输出变量,控制参变量作为输入变量(如室外温度、末端温度等)。
- 3)为了模型拟合的效果,输入变量要尽量包含所有影响热负荷或者能耗的参变量。

针对锅炉二次换热供热系统,表 1 给出了常见的输入、输出可选取变量。变量的系统层次不同,能耗诊断也可分为系统层的诊断和设备层的诊断,前者主要集中于能耗数据的数值诊断,后者侧重于设备运行的故障诊断^[24]。而本文模型参变量选择主要针对能耗数值的诊断,这种诊断本质上是节能控制中被控制变量对于工况条件的响应,而模型反映了响应的映射关系,能耗数据则是这种响应的直观表现,且这种直观表现是基于外界条件(比如室外温度、供回水温度等)而与时间序列无关^[25],从原始数据中筛选出有效的节能数据并不影响诊断模型本身的正确性。

表 1 常选取的变量

Table 1 Frequently selected variable

系统层次	输入变量	输出变量
热负荷环境 影响因子	室外温度,湿度,末端温度,太阳辐照度	
系统运行参数	一次供回水温度,二次供回水温度,一次热水流量,二次热水流量	供热量,一次热水流量,二次热水流量
设备运行参数	换热板参数,锅炉运行参数	

1.1 数据判别和处理

数据筛选是基于影响因子对数据进行筛选过程,从节能的角度看,在约束条件下(设备、工况、人员等),用尽量少的热量去满足相应的热负荷被认为是最为理想的节能状态^[26]。由于室外温度对热负荷的影响最大^[27],因此,把室外温度与供热量相关性系数 R 作为数据筛选初级指标较为合理。数据筛选流程如图 2 所示,若相关性系数 R 大于 -0.2 ,可以认为供热量与热负荷脱节,处于不节能状态,而不能作为诊断数据; R 介于 -0.2 到 -0.5 之间的数据,可以认为数据中部分是节能的,需要把节能数据从数据集中筛选出来;如果 R 值小于 -0.5 ,可以认为它具有很好的节能特性,可直接用于诊断。常用

筛选算法有概率模型筛选算法^[28]、无监督聚类算法^[29]及监督学习分类算法^[30],筛选获得数据的评价应考虑:

- 1) R 是否小于 -0.5 ,即是否筛选出适宜建模的诊断数据。
- 2)筛选出来的数据占采集总数据集的比例 η 应大于 30% ,才能保证被筛选数据的普遍性和代表性,从而保证诊断的可靠性。
- 3)节能数据量是否足够大,即诊断数据量 N 必须大于被诊断数据量 $2N'$ 作为标准。

如果符合要求的数据集很多并对诊断数据没有很高的节能特性要求,且都满足 $R < -0.5$,可以适当减少分类数来保证数据的量。由于变量个数过多会影响模型拟合的速度,因此,原则上若实际情况需要,可以用变量降维算法^[31-32]。

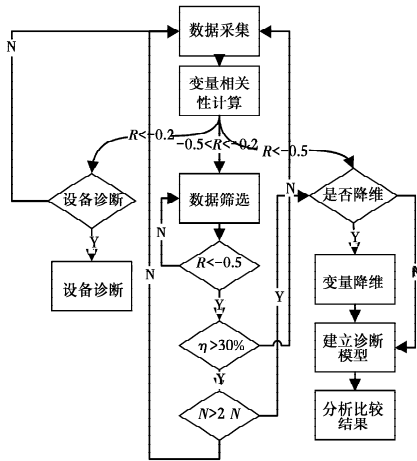


图 2 数据筛选流程图
Fig. 2 Data filter flow chart

1.2 模型建立

二次换热供热系统的诊断模型建立采用多变量回归拟合的方式,常用的多元回归算法有最小二乘回归^[33]、人工神经网络^[34]、LS-SVM 回归^[35]、回归树^[36]等。由于在选择回归模型时既要考虑数据的数量、质量,也要考虑所诊断的目标变量,因此,并不存在统一的模型选择标准。本文案例研究中采用 K-means 聚类^[37]的数据筛选算法和基于贝叶斯正则化训练的人工神经网络回归模型。

1.3 诊断结果评价指标

主要选取以下 3 个指标作为节能特性的判定:

- 1)供热量与室外温度的相关性 R 。 R 数值越大,越节能。
- 2)MAP 指标。

$$\text{sumerror} = \sum_{i=1}^n (E_{\text{预测}}^i - E_{\text{实际}}^i)$$

(1)

式中： $E_{预测}^i$ 和 $E_{实际}^i$ 分别表示预测的供热量和实际供热量； n 表示数据量，负值表示不节能，正值表示节能。

3) 节能率 η 。

$$\eta = - \frac{\sum_{i=1}^n E_{实际}^i}{n} \times 100\%$$

(2)

节能率反映相对的节能效果，正值表示被诊断数据不节能，负值表示被诊断数据节能。

2 案例研究

2.1 系统介绍及数据采集

以某医院锅炉二次换热供热系统为研究案例，如图 3 所示。从能量流动角度讲，能量是由锅炉热水经过两次换热传递给末端，第一次换热是一次供热端与二次供热端经过换热板换热，介质都是水，二次换热是二次供热端热水与末端风机盘管中的空气换热，风机盘管再把热量传递给空间。因此，决定能量传递大小的因素是一次供热端的热水流量和二次供热端热水流量，一次供热端供回水温差以及二次供热端供回水温差。从节能控制角度，可以直接控制的变量只有一次供热端和二次供热端的热水流量，而在本文中，只有二次供热端有变频泵根据室内温度、末端温度、供回水温度这 3 个方面控制热水流量，因此，这些相关变量都应考虑到。表 2 所示是系统参数、3 个区域的采暖面积、采暖热源、末端、锅炉型号以及数据采集的变量和数据量等信息，诊断数据是 2015-12-15 到 2016-03-15，每 2 h 为步长的数据，被诊断数据是 2016-02-17 至 2016-02-22，每 0.5 h 为步长的数据，在这段时间内，二次供热端关闭变频节能控制措施。经过滤波诊断数据和被诊断数据量分别为 1 100 和 194 个。

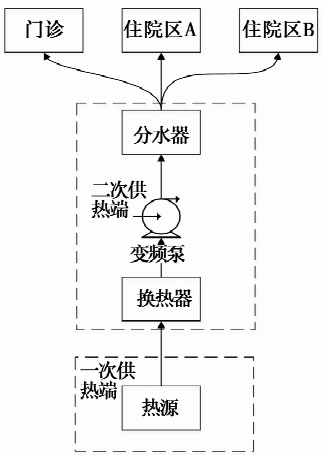


图 3 医院供热系统

Fig. 3 Heating system of hospital

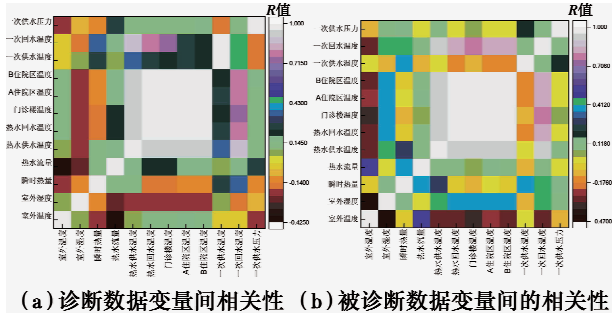
表 2 系统介绍和采集的数据变量

Table 2 Data variables introduced and collected data by the system

项目	住院区 A	住院区 B	门诊部	变量编号	变量名称
建造年份	2011 年	2011 年	2011 年	1	室外温度/℃
建筑面积	17 303 m ²	37 295 m ²	32 011 m ²	2	一次供水温度/℃
采暖面积	17 303 m ²	37 295 m ²	32 011 m ²	3	一次回水温度/℃
采暖热源	燃气锅炉热水	燃气锅炉热水	燃气锅炉热水	4	一次供水压力/MPa
采暖末端	FCU+新风	FCU+新风	FCU+新风	5	二次供水温度/℃
运行时间	24 h	24 h	08:00—18:00	6	二次回水温度/℃
设计总负荷	8 705 kW			7	瞬时热量/kW
锅炉参数(3 台)	工作压力 1 MPa;使用压力 1 MPa。额定出力 4.2 t/h;介质温度 115 ℃。			8	二次热水流量/(m ³ ·h ⁻¹)
节能控制变量	二次供热端热水流量			9	门诊区末端温度/℃
二次最大热水流量	750 m ³ /h			10	住院区 A 末端温度/℃
被诊断数据	197			11	住院区 B 末端温度/℃
诊断数据	1 100			12	室外湿度/%

2.2 数据处理

如图 4 为采集变量之间的相关性热点图。其中，瞬时热量与室外温度的相关性 R 为 $-0.272\ 3$ ，因此，属于第 2 种情况，应采用 K-means 算法进行聚类，由于案例中对诊断数据并没有要求很高的节能特性，根据模型以 R 是否小于 -0.5 作为诊断数据的选择标准，从图 4(b)也可以看出，室外温度与瞬时热量的相关性为 $-0.093\ 52$ ，节能性较低。



(a) 诊断数据变量间相关性 (b) 被诊断数据变量间的相关性

图 4 变量间相关性热点图

Fig. 4 Hot spot map of Inter-variable correlation

图 5 为诊断数据经过 K-means 算法聚类后,聚类数据集中室外温度与瞬时热量的相关性及数据量,数据集 3 相关性 R 值达到了 $-0.539\ 2$,数据量达到了 412,满足诊断数据的要求,可以作为诊断数据对被诊断数据进行诊断。

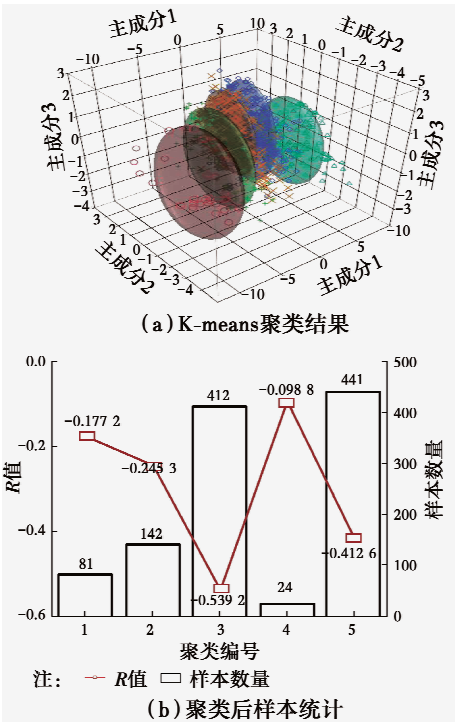


图 5 K-means 聚类结果

2.3 模型建立

模型选取输入变量应为能耗的影响因子,相应的响应变量应为供热量值与节能控制中的被控制变量热水流量。神经网络回归模型的拓扑结构如图 6 所示,采用 3 层神经网络,输入层有 10 个神经元,表示 10 个输入变量,隐含层有 25 个神经元,其激活函数为 sigmoid 函数,输出层有 1 个神经元,表示 1 个输出变量,输出函数为线性函数。人工神经网络常用的训练算法有误差反向传播法(BP)、共轭梯度法、牛顿法、L-M 法,由于这些算法本质上是基于误差-权重空间中误差沿空间场的梯度下降最快的原理,对误差-权重空间的平滑性(训练样本的质量)敏感性较高而易陷入局部最优和欠拟合^[38]。其次,能耗诊断的准确与否与神经网络模型的泛化性相关性较大,而贝叶斯正则化方法把误差和网络结构复杂度都作为训练的惩罚函数,因此,能有效提高神经网络的泛化性能^[39],基于以上原因,采用贝叶斯正则化训练方法对神经网络进行训练,既保证模型的准确性,也保证模型的泛化性。同时,为了保证训练结果的有效性,训练集(train)、测试集(test)按 7:3 的比

例,采用随机分割的方法^[40]获得。表 3 所示为 4 个诊断模型结构所选取的数据集及模型的拟合度 R 值,以供热量作为响应变量拟合度的 R 值高于以二次端热水流量为响应变量模型的 R 值。但是,总体 R 值都高于 0.9,模型拟合度较高,具有较好的诊断可靠性。

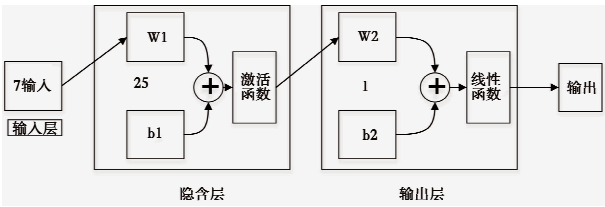


图 6 神经网络结构

表 3 回归模型结构

模型编号	输入变量编号	输出变量编号	模型拟合 R 值	诊断目标	模型拟合数据集	训练算法	迭代次数
D3	1~6, 9~12	7	0.975 6	供热量	No. 3		445
2	1~6, 9~12	7	0.970 5	供热量	所有	贝叶斯正则化训练算法	472
D3-flux	1~6, 9~12	8	0.921 4	二次热水流量	No. 3		569
2-flux	1~6, 9~12	8	0.910 1	二次热水流量	所有		673

2.4 诊断结果

图 7 所示为模型 D3 和模型 2 对聚类后数据集 1、数据集 4 和被诊断数据进行诊断的结果,可以发现,模型 D3 预测供热量与室外温度相关性系数分别达到了 $-0.308\ 7$ 、 $-0.471\ 5$ 和 $-0.203\ 0$,累积误差 MAP 值分别为 $-149\ 498.67$ 、 $-86\ 526$ 、 $-4\ 052.27\ \text{kW}$,节能率为 10.7%、17%和 4%,都优于模型 2;且从图中可以看出,模型 2 预测和实际供热量高于设计负荷线(图中 X 轴方向虚线)采样点比例更大,趋势更明显,而模型 D3 预测结果相对较好。图 7(b)所示为被诊断数据模型 2 预测值与实际值的相关性图,相关性高达 0.95,这说明处于模型 2 变频节能控制状态下运行的能耗水平与非变频运行状态下较为相似,节能率指标 η 也只有 1%左右。因此,变频节能控制并未达到应有的节能效果。同理,对数据集 1 和数据集 4 的诊断验证也可发现,总数据集预测结果与实际相近,与室外温度的相关性为 $-0.259\ 2$ 和 -0.1153 ,节能率指标只有 0.4%

和1.8%,诊断效果并不明显。对比发现,经过K-means筛选出来的数据节能诊断效果较好。

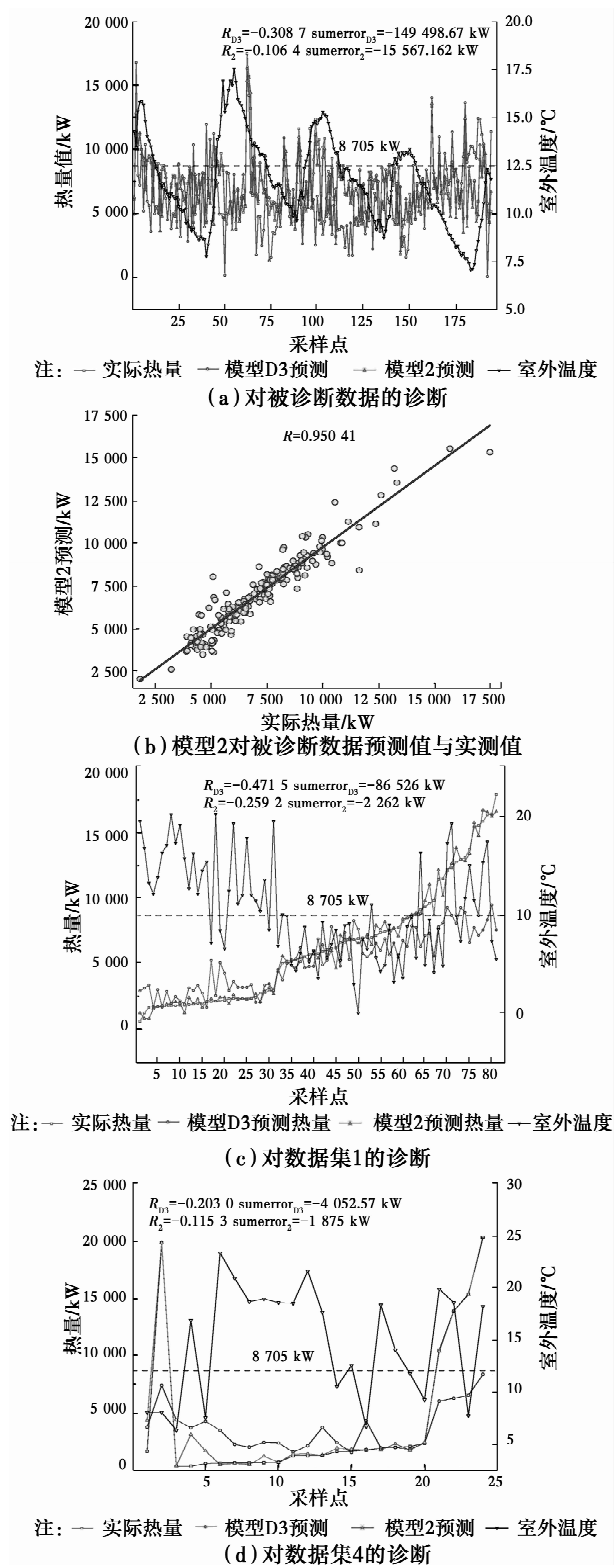


图7 模型D3和模型2诊断结果

Fig. 7 Diagnostic results of Model D3 and Model2

图8所示为模型D3-flux和模型2-flux诊断的结果,从图8(a)中可以看出,由于被诊断数据关闭了变频泵控制,因此,水流量波动较为平稳,随室外

温度变化波动不大。模型2-flux预测热水流量虽然有一定的波动,但是无论均值、方差以及累计误差都与实际热水流量相近,对比模型D3-flux和室外温度发现,当室外温度升高热负荷降低时,模型D3-flux预测热水流量也有明显下降,方差为17320.67,波动性相对于模型2-flux较大,累计误差MAP为21778.86 m³/h,η值为18.7%,被诊断数据不节能。图8(b)所示为通过预测的二次热水流量和实际供回水温差计算出来的理论供热量,计算式为

$$Q = CM\Delta T \quad (3)$$

式中: C 为水的比热容; M 为热水流量; ΔT 为供回水温差,由于比热 C 是常量,为了方便起见,本文用 $M\Delta T$ 作为替代。从图7(b)中可以看出,模型D3-flux相关性系数 R 要优于模型2-flux,MAP指标达到了-66380.237 m³·°C/h,节能率 η 为19.7%,高于模型2-flux的1.2%。对比诊断结果,被诊断数据不节能,模型D3-flux诊断效果要优于模型2。

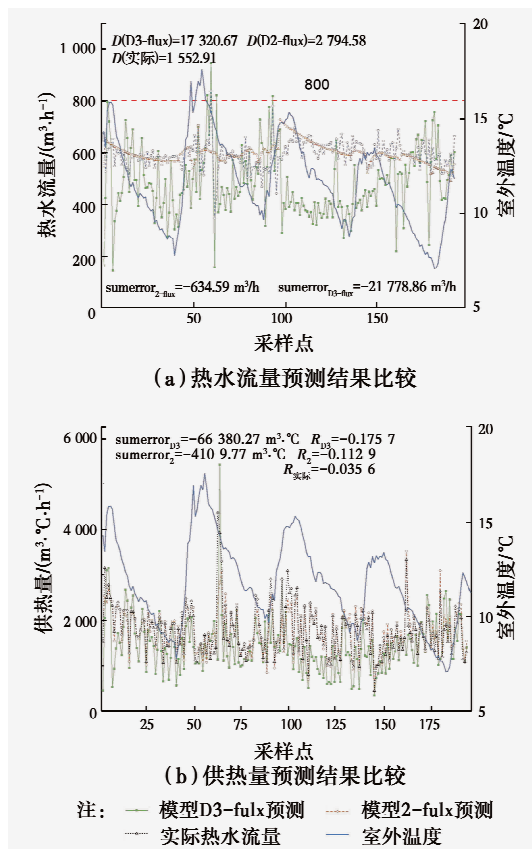


图8 模型D3-flux和模型2-flux诊断结果

Fig 8 Diagnostic results of model D3-flux and model 2-flux

2.5 结果分析与讨论

4个模型本质上是对节能控制结果的有效性进行诊断。从能量守恒的角度,能量是通过热水两次换热进行传递的,除了满足末端的空间热负荷外,换

热板损失和路途管道损失也是能耗主要的组成部分。从锅炉二次供热原理图可以看出,系统主要有两个回路,以第2个回路为例,由于有变频泵的存在,因此,节能控制的原理是基于热负荷进行流量调节,而在控制反馈的参变量只有室内外温度等气候因素和二次端的供回水温差,其中,环境条件只能反映实际的环境热负荷,但能量供给的多少是由一次端的供热量和换热板效率决定,前者是通过人工手动调节供水压力来控制,具有一定的盲目性和滞后性;而后者效率决定因素是换热板的结构和换热两侧的供回水温度。换热板结构是固定的,因此,换热两侧供回水温度就成了主要影响变量。图9所示为一次端供水压力和换热板热量损失,一次换热板损失热量由式(4)计算^[41-42]。

$$Q = KF \Delta t_m \quad (4)$$

式中: K 为传热系数; F 为换热板接触面积; Δt_m 是对数。平均温差由式(5)计算。

$$\Delta t_m = \frac{(T_1 - t_2) - (T_2 - t_1)}{\ln \frac{T_1 - t_2}{T_2 - t_1}} \quad (5)$$

式中: T_1 是二次端热水供水温度; T_2 是二次端热水回水温度; t_1 是一次端供水温度; t_2 是一次端回水温度,负值表示损失,由于 F 和 C 是常量,为了简单起见,用 Δt_m 替代供热量进行分析。由图9可以看出,当从2016-02-11开始人为降低一次平均供水压力,换热板热量损失也明显降低,因此,一次端供水压力合理调节也是提高换热板能效的有效方式。

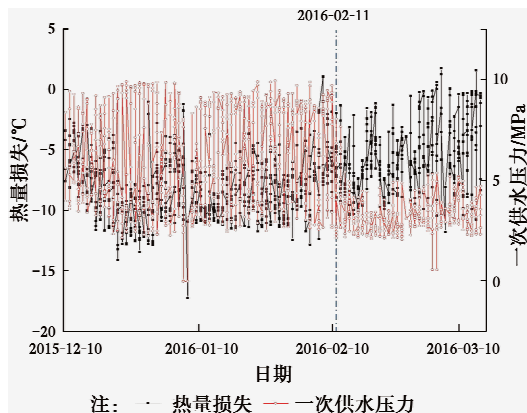


图9 一次供水压力和一次换热板热量损失

Fig. 9 Heat loss of primary water supply pressure and primary heat exchange plate

二次供热端虽然是变频泵控制热水流量来达到节能的目的,但是,真正供给能量的是一次供热端,因此,二次供热端变频节能控制对能量供给只起缓冲作用,节能效果有限。从诊断结果可以发现,关闭变频泵和不关闭变频泵在能耗表现上没有太大差

异,因此,对于系统本身节能应增加针对一次供热端的热水流量的节能调节措施。图10为日累计热量值与室外温度的走势图,可以发现,无论室外温度如何变化,日供热量仍然保持一个平稳且高能耗的供热状态。关闭节能泵的5天(图中虚线所示),虽然能耗数值略微高于相似气温条件的能耗数值,但是,劣势并不明显,气温与关闭变频泵相近的几天(图中X轴方向实线和虚线之间的区域)平均相对节能率只有1.7%,也印证了分析结果。因此,从所有数据中筛选出节能数据对于诊断意义重大,这种模型对比也是对诊断数据和被诊断数据的一个交叉检验。

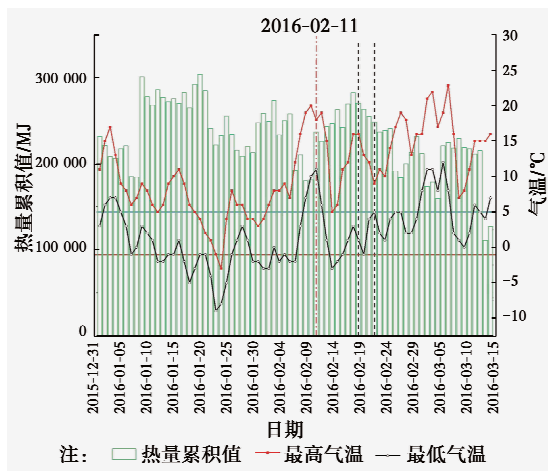


图10 热量累积值与室外温度

Fig. 10 Heat accumulation and outdoor temperature

3 结论

1) 经过筛选的数据能耗诊断效果要优于未经筛选的数据,经过数据集1、数据集4和被诊断数据检验结果与室外温度相关性 R 值达到了-0.3087、-0.4715和-0.2030,累积误差MAP值也分别达到了-149498.67、-86526、-4052.27 kW,节能率也达到了10.7%、17%和4%,都优于未经筛选数据建立的模型2。

2) 经案例研究发现,一次供热端供水压力人工调节的滞后性、盲目性是导致能耗高、能效低的主要原因。只针对二次供热端热水流量采取节能控制措施,而不对一次供热端采取相应的节能控制措施所达到的节能效果有限,经实际能耗数据验证,这个结论客观有效。

3) 这种数据建模诊断方式是基于输入、输出变量之间的物理响应关系,而不受数据时间特性的影响,因此,在实际应用中具有较好的泛化性和实用性,能有效对锅炉热水供热系统进行诊断。

4)由于本文提出的模型是通过数据“学习”系统运行模式,因此,数据的质量、数量及其采集的难易程度也会影响模型的普适性。但随着云技术、大数据平台技术的发展,建立类似的系统运行标准数据库并基于更大数据量、更复杂学习模式(比如深度学习)能极大提升模型的泛化性和准确度。

参考文献:

- [1] Key world energy statistics [R]. International Energy Agency, 2014.
- [2] 梁境,李百战,武涌. 中国建筑节能现状与趋势调研分析[J]. 暖通空调, 2008(7):29-35.
LIANG J, LI B Z, WU Y. The analysis of the current situation and trend of building energy saving China research [J]. Journal of HVAC, 2008(7):29-35. (in Chinese)
- [3] 于德森. 我国供热行业特性与政府规制研究[D]. 武汉:武汉理工大学, 2012.
YU D M. Research on the characteristics and government regulation of heating industry in China [D]. Wuhan: Wuhan University of Technology, 2012. (in Chinese)
- [4] 金毅诚. 中国供热计量收费现状与思考[J]. 建设科技, 2010(8):70-72.
JIN Y C. Present situation and thinking of heating metering and charging in China [J]. Construction Technology, 2010(8):70-72. (in Chinese)
- [5] 王德静. 有关区域热水锅炉冬季采暖运行方案的设计方案[J]. 中国科技信息, 2013(7):177-178.
WANG D J. Design scheme of heating operation scheme for regional hot water boiler in winter [J]. 2013(7):177-178. (in Chinese)
- [6] 钱鹏. 过程管理在市政供暖锅炉改造工程中的应用与实践[D]. 长春:吉林大学, 2015.
QIAN P. Application and practice of process management in municipal heating boiler reconstruction project [D]. Changchun: Jilin University, 2015. (in Chinese)
- [7] ELMGAARD B, OMMEN T S, MARKUSSEN M, et al. Integration of space heating and hot water supply in low temperature district heating [J]. Energy and Buildings, 2016, 150: 255-264.
- [8] HONG T, KOO C, KIM J, et al. A review on sustainable construction management strategies for monitoring, diagnosing, and retrofitting the building's dynamic energy performance: Focused on the operation and maintenance phase [J]. Applied Energy, 2015, 155: 671-707.
- [9] MAO W T, WANG J W, HE L, et al. Online sequential prediction of bearings imbalanced fault diagnosis by extreme learning machine [J]. Mechanical Systems and Signal Processing, 2017, 83: 450-473.
- [10] SEERA M, LIM C P, ISHAK D, et al. Offline and online fault detection and diagnosis of induction motors using a hybrid soft computing model [J]. Applied Soft Computing, 2013, 13: 4493-4507.
- [11] GENG Z Q, GAO H C, WANG Y Q, et al. Energy saving analysis and management modeling based on index decomposition analysis integrated energy saving potential method: Application to complex chemical processes [J]. Energy Conversion and Management, 2017, 145: 41-52.
- [12] FAN C, XIAO F. Assessment of building operational performance using data mining techniques: A case study [J]. Energy Procedia, 2017, 111: 1070-1078.
- [13] BELUSSI L, DANZA L. Method for the prediction of malfunctions of buildings through real energy consumption analysis: Holistic and multidisciplinary approach of energy signature [J]. Energy and Buildings, 2012, 55: 715-720.
- [14] 李刚,冯国会,王丽,等. 严寒地区农村住宅节能改造能耗模拟[J]. 沈阳建筑大学学报(自然科学版), 2012(5):884-890.
LI G, FENG G H, WANG L, et al. Energy consumption simulation of rural residential energy saving transformation in severe cold area [J]. Journal of Shenyang Construction University (Nature Science), 2012(5):884-890. (in Chinese)
- [15] ZHOU L N, PAN S M, WANG J W, et al. Machine learning on big data: Opportunities and challenges [J]. Neurocomputing, 2017, 237(10):350-361.
- [16] SULEIMAN D, AL-ZEWAIIRI M, NAYMAT G. An empirical evaluation of intelligent machine learning algorithms under big data processing systems [J]. Procedia Computer Science, 2017, 113(8):539-544.
- [17] DANGLADE F, PERNOT J. A priori evaluation of simulation models preparation processes using artificial intelligence techniques [J]. Computers in Industry, 2017, 91: 45-61.
- [18] KAVAKIOTIS L, TSAVE O, SALIFOGLU A, et al. Machine learning and data mining methods in diabetes research [J]. Computational and Structural Biotechnology Journal, 2017, 15: 104-116.
- [19] SEOL J W, YI W J, CHOI J, et al. Causality patterns and machine learning for the extraction of problem-action relations in discharge summaries [J]. International Journal of Medical Informatics, 2017, 98: 1-12.

- [20] TSAFTARIS S A, MINERVINI M, SCHARR H. Machine learning for plant phenotyping needs image processing [J]. *Trends in Plant Science*, 2016, 21(12): 989-991.
- [21] TRAVIESO C M, FODOR J, ALONSO J B. Special issue on information processing and machine learning for applications of engineering [J]. *Neurocomputing*, 2015, 150: 347-348.
- [22] YILDIZ B, BILBAO J I, SPROUL A B. A review and analysis of regression and machine learning models on commercial building electricity load forecasting [J]. *Renewable and Sustainable Energy Reviews*, 2017, 73: 1104-1122.
- [23] HAN Y M, ZHU Q X, GENG Z Q, et al. Energy and carbon emissions analysis and prediction of complex petrochemical systems based on an improved extreme learning machine integrated interpretative structural model [J]. *Applied Thermal Engineering*, 2017, 115 (3): 280-291.
- [24] 李萍, 卢洪波, 刘士祥, 等. 锅炉故障诊断综述[J]. *科技信息*, 2009(4): 584.
- LI P, LU H B, LIU S X, et al. The reviews of diagnosis of boiler [J]. *Science Information*, 2009(4): 584. (in Chinese)
- [25] YILDIZ B, BILBAO J I, SPROUL A B. A review and analysis of regression and machine learning models on commercial building electricity load forecasting [J]. *Renewable and Sustainable Energy Reviews*, 2017, 73: 1104-1122.
- [26] 江亿. 我国建筑节能战略研究[J]. *中国工程科学*, 2011(6): 30-38.
- JIANG Y. Study on building energy efficiency strategy in China [J]. *Chinese Engineering Science*, 2011(6): 30-38. (in Chinese)
- [27] 张晓彤, 刘金祥, 陈晓春, 等. 区域建筑冷(热)负荷的影响因素敏感性分析及预测方法研究[J]. *建筑科学*, 2013(8): 77-82.
- ZHANG X T, LIU J X, CHEN X C, et al. Sensitivity analysis and prediction method study on influence factors of regional building cold (heat) load [J]. *Architectural Science*, 2013(8): 77-82. (in Chinese)
- [28] SUN S L, SHAW-TAYLOR J, MAO L. PAC-Bayes analysis of multi-view learning [J]. *Information Fusion*, 2017, 35: 117-131.
- [29] STIMPFLING T, BELANGER N. Extensions to decision-tree based packet classification algorithms to address new classification paradigms [J]. *Computer Networks*, 2017, 122: 83-95.
- [30] STIMPFLING T, BELANGER N. Fault classification in power systems using EMD and SVM [J]. *Ain Shams Engineering Journal*, 2017(8): 103-111.
- [31] DONG Y, QIN S J. A novel dynamic PCA algorithm for dynamic data modeling and process monitoring [J]. *Journal of Process Control*, 2017, 56(6): 148-163.
- [32] 吴晓婷, 闫德勤. 数据降维方法分析与研究[J]. *计算机应用研究*, 2009(8): 2832-2835.
- WU X T, YAN D Q. Analysis and research of data reduction method [J]. *Computer Application Research*, 2009(8): 2832-2835. (in Chinese)
- [33] 张景阳, 潘光友. 多元线性回归与 BP 神经网络预测模型对比与运用研究[J]. *昆明理工大学学报(自然科学版)*, 2013(6): 61-67.
- ZHANG J Y, PAN G Y. Comparison and application of multiple linear regression and BP neural network prediction mode [J]. *Journal of Kunming University of Science and Technology*, 2013(6): 61-67. (in Chinese)
- [34] HANIEF M, WANI M F, CHAROO M S. Modeling and prediction of cutting forces during the turning of red brass (C23000) using ANN and regression analysis [J]. *Engineering Science & Technology An International Journal*, 2017, 20(3): 1220-1226.
- [35] ZHANG Y Y, WEI H, YANG Y D, et al. Forecasting of dissolved gases in oil-immersed transformers based upon wavelet LS-SVM regression and PSO with mutation [J]. *Energy Procedia*, 2016, 104: 38-43.
- [36] CHEN W, XIE X S, WANG J L, et al. A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility [J]. *Catena*, 2017, 151: 147-160.
- [37] 王千, 王成, 冯振元. K-means 聚类算法研究综述[J]. *电子设计工程*, 2012(7): 21-24.
- WANG Q, WANG C, FENG Z Y, et al. A survey of K-means clustering algorithms [J]. *Electronic Design Engineering*, 2012(7): 21-24. (in Chinese)
- [38] 刘冉, 卢本捷. 神经网络训练算法的对比及应用[J/OL]. *软件*, 2011, 32(10): 29-31, 34. (2012-01-08). <http://kns.cnki.net/kcms/detail/12.1151.TP.20120108.2053.004.html>.
- LIU R, LU B J. Comparison and application of neural network training algorithm [J/OL]. *Software*, 2011, 32(10): 29-31, 34. (2012-01-08). <http://kns.cnki.net/kcms/detail/12.1151.TP.20120108.2053.004.html>. (in Chinese)
- [39] 魏东, 张明廉, 蒋志坚, 等. 基于贝叶斯方法的神经网络非线性模型辨识[J]. *计算机工程与应用*, 2005(11): 5-8, 11.
- WEI D, ZHANG M L, JIANG Z J, et al. Neural

- network non-linear modelling based on bayesian methods [J]. Computer Engineering and Applications, 2005(11):5-8,11. (in Chinese)
- [40] 李晓菲. 数据预处理算法的研究与应用[D]. 成都:西南交通大学, 2006.
- LI X F. Research and application of data preprocessing algorithm [D]. Chengdu: Southwest Jiaotong University, 2006. (in Chinese)
- [41] 陈进富. 换热器传热过程的分析[J]. 西南石油学院学报, 1994(2):115-122.
- CHENG J F. Analysis of heat transfer process in heat exchanger [J]. Journal of Southwest Petroleum Institute, 1994(2):115-122. (in Chinese)
- [42] 娄载强, 杨冬, 翟建修, 等. 换热器的热损失分析方法介绍[J]. 制冷与空调, 2013(1):94-96.
- LOU Z Q, YANG D, ZHAI J X, et al. Heat loss analysis method of heat exchanger [J]. Refrigeration and Air-Conditioning, 2013(1):94-96. (in Chinese)
- (编辑 王秀玲)