

基于数据挖掘的地源热泵系统性能测试和分析

黄新江,刘成刚

(苏州科技大学 环境科学与工程学院,江苏 苏州 215009)

摘 要:以苏州市某办公楼地源热泵系统全年监测数据为基础数据,分析了地源热泵系统动态运行特性,并运用随机森林模型,探讨了夏季、冬季两种工况下各运行参数对预测机组性能系数的重要性。结果表明:该系统夏季机组平均性能系数为 3.92,冬季机组平均性能系数为 4.15,对机组性能系数影响最大的 3 个运行参数分别是部分负荷率、地源侧回水温度和用户侧回水温度。随机森林模型适用于地源热泵系统性能系数预测并且预测精度较高。

关 键 词:地源热泵;性能系数;随机森林;部分负荷率;运行特性

中图分类号:TU83 文献标识码:A DOI:10.16146/j.cnki.rndlge.2020.03.026

[引用本文格式]黄新江,刘成刚.基于数据挖掘的地源热泵系统性能测试和分析[J].热能动力工程,2020,35(3):181-186.
HUANG Xin-jiang,LIU Cheng-gang. Performance test and analysis of ground source heat pump system based on data mining[J]. Journal of Engineering for Thermal Energy and Power,2020,35(3):181-186.

Performance Test and Analysis of Ground Source Heat Pump System based on Data Mining

HUANG Xin-jiang,LIU Cheng-gang

(School of Environmental Science and Engineering,Suzhou University of Science and Technology,Suzhou,China,Post Code: 215009)

Abstract: Based on the annual monitoring data of the ground source heat pump system in an office building in Suzhou, the dynamic operation characteristics of the ground source heat pump system are analyzed, and the importance of each operating parameter in predicting the unit coefficient of performance under two kind of working conditions is discussed by using the random forest model. The results show that the average unit coefficient of performance is 3.92 in summer and 4.15 in winter. Partial load rate, backwater temperature at the ground source side and backwater temperature at the user side are the three operating parameters that have the greatest influence on the unit coefficient of performance. The random forest model can be used to predict the coefficient of performance with high accuracy.

Key words: ground source heat pump, coefficient of performance, random forest, partial load rate, operating characteristics

引 言

地源热泵技术已经成为国内建筑节能及暖通空调界的热门研究课题,并广泛应用于工程实践^[1]。如何改善热泵机组性能系数(COP)是地源热泵系统降低能耗,优化系统运行性能的关键问题。目前,针

对热泵机组 COP 的研究主要分为两类。一类是基于仿真模型得到的长期模拟数据分析,如刘昱等人通过 TRNSYS 软件模拟比较不同钻孔深度与钻孔间距时地源热泵系统性能差异^[2];Liang 等人从埋管换热器设计参数角度运用 TRNSYS 软件证明了克里格模型可以提高热泵机组运行性能^[3]。另一类是基于实际地源热泵系统的短期测试数据分析,如

收稿日期:2019-03-06; 修订日期:2019-04-03

作者简介:黄新江(1993-),男,江苏苏州人,苏州科技大学硕士研究生,E-mail:617292438@qq.com.

通讯作者:刘成刚(1958-),男,辽宁安台人,苏州科技大学教授,E-mail:cliu1977@163.com.

傅强等人运用多元线性回归模型分析了夏季工况下地源热泵机组 COP 的主要影响因素^[4];Magraner 等人对地源热泵试验台进行了全年实测,发现负荷率对机组 COP 有较大影响^[5];Sivasakthivel 等人采用正交试验法分析了影响热泵机组性能的参数^[6]。国内外学者的大量研究详细且全面地分析了热泵机组 COP 的影响因素,但缺乏对各个因素间的耦合关系以及共同影响机组 COP 时的权重关系研究。

本文通过对苏州市某地源热泵系统全年实测运行数据的分析,结合随机森林模型,对地源热泵机组 COP 进行了高精度的预测,探讨了各运行参数之间的相关性和对机组 COP 影响的重要程度。

1 工程信息

该空调工程位于江苏省苏州市某办公楼第 3 ~ 4 层,是以土壤为冷热源的地源热泵空调系统,空调面积 959 m²。配备两台型号为 MWH020 的水-水模块式热泵机组,单台机组名义制冷量为 69 kW,制热量为 58.3 kW,夏季供冷冬季供热。系统采用双 U 型地埋管,钻孔深度 50 m,共 40 口井。冷冻水泵冷却水泵均一用一备,工频运行,同时配置了一台冷却塔备用。

为了把握系统各环节的能耗情况以及动态运行规律,数据趋势记录采用 Alerton 公司的 BACtalk 楼宇监控系统,配备 VLC-853 和 VLC-1600 两台现场控制器进行监控。监控内容主要包括地源热泵系统、新风机组和空调末端。系统数据采集频率 12 次/h。各运行参数监测点位置如图 1 所示。

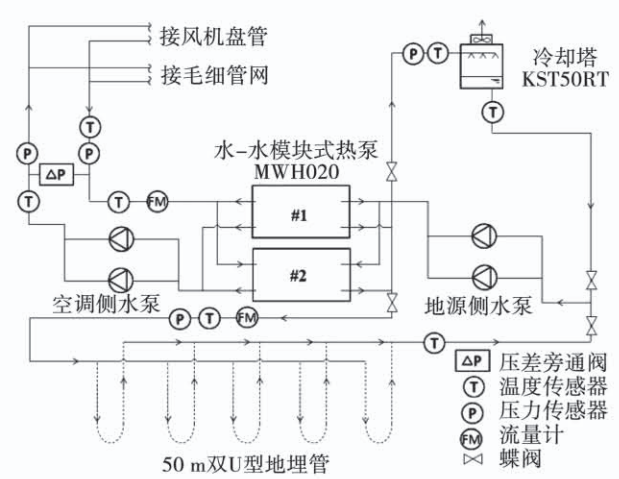


图 1 地源热泵系统监测点位分布示意图
Fig.1 Schematic diagram of monitoring point location of ground source heat pump system

2 研究方法

2.1 随机森林模型

随机森林(RF)是一种基于数据随机特征选择构建多决策树的集成学习方法^[7]。每棵树都是由一个引导样本生长的。在树的每个节点随机选取变量,在抽取的特征集中进行最优选择,最后的预测是通过汇总各决策树的预测结果得到的。这种方法不仅避免了对单个决策树的过度拟合,而且对决策树的随机抽样可以提高其准确性和泛化能力^[8]。地源热泵系统各运行参数耦合性较强,趋势记录时段内波动幅度较大,精确地数学模型很难建立,而 RF 相比于其他机器学习算法和传统的逻辑回归模型^[9],降低了森林中树之间的相关性,从而降低了错误率,能够有效地处理高度相关的数据,更适用于分析地源热泵系统数据。

2.2 模型构建和变量选取

RF 模型的构建过程中,预测性能很大程度依赖于变量的选取^[10]。为了保证变量选取完全独立于测试数据,降低泛化性,在模型训练过程中,随机森林模型使用的是 OOB 交叉验证算法,通过随机选取实验数据,将工作模式划分为训练和测试数据集,约选取数据中的 70% 用于模型的训练,30% 用于模型的测试。

选取夏季工况趋势记录区间为 2017 年 6 月 1 日~2017 年 8 月 31 日,冬季工况趋势记录时间为 2017 年 12 月 1 日至 2018 年 1 月 31 日。选取部分负荷率、用户侧供水温度、用户侧回水温度、地源侧供水温度、地源侧回水温度、室外干球温度和室外相对湿度 7 个运行参数作为预测机组 COP 的输入变量,如表 1 所示。其中 COP 与部分负荷率(PLR)的计算式为:

$$COP = \frac{Q_c}{\omega \cdot t} \tag{1}$$

$$Q_c = m_w \rho c_{w,p} (T_{chwr} - T_{chws}) \tag{2}$$

$$PLR = \frac{Q_c}{Q_{c,nom}} \tag{3}$$

式中: Q_c —机组制冷量, kW; ω —压缩机功率, kW; t —时间, s; m_w —用户侧供回水流量, m³/h; ρ —水的密度, kg/m³; $c_{w,p}$ —水的比热容, kJ/(kg · K); $Q_{c,nom}$ —热泵机组名义制冷量。

表 1 随机森林模型变量选取

参 数	数据来源	测量精度
机组性能系数 COP	—	—
部分负荷率 PLR	—	—
用户侧供水温度 T_{chws}	水道温度传感器	$\pm 0.10\text{ }^{\circ}\text{C}$
用户侧回水温度 T_{chwr}	水道温度传感器	$\pm 0.10\text{ }^{\circ}\text{C}$
地源侧供水温度 T_{cws}	水道温度传感器	$\pm 0.10\text{ }^{\circ}\text{C}$
地源侧回水温度 T_{cwr}	水道温度传感器	$\pm 0.10\text{ }^{\circ}\text{C}$
室外干球温度 T_o	温湿度记录仪	$\pm 0.21\text{ }^{\circ}\text{C}$
室外湿度 RH	温湿度记录仪	$\pm 3.5\%$

2.3 重要度评价方法

随机森林模型为每个输入变量提供了一个重要度的概念,该参数值用 ΔMSE_i 表示,它表示当第 i 个输入变量被忽略时,输出变量的均方误差(MSE)增加的百分比。误差增加百分比越大,表明与其他输入变量相比,该输入变量对预测 COP 的变化更重

要。输入变量对预测 COP 的重要性可直接基于 ΔMSE 的值进行排序。

$$\Delta\text{MSE}_i = \frac{\text{MSE}_{-i} - \text{MSE}}{\text{MSE}} \times 100\%$$

(4)

式中: MSE_{-i} —第 i 个变量被忽略后的均方误差。

3 结果与讨论

3.1 输入变量相关性诊断

表 2~表 3 为不同模式下变量之间的皮尔森相关系数。COP 与单一输入变量之间的相关系数在 0.029 0~0.437 0 之间,低于 0.8,说明两者不具有强相关性。此外,对于相同的输入变量的相关系数的符号可能在不同模式下发生变化,这说明 COP 与输入变量间可能不存在线性相关关系。而且根本无法识别出与 COP 变化具有最强相关性的前几个输入变量。由于皮尔森相关系数只能反映变量间的线性关系,想要研究 COP 和输入变量之间的非线性关系,需要基于随机森林模型进一步研究。

表 2 变量间的皮尔森相关系数(制冷工况)

参数	COP	PLR	T_{chws}	T_{chwr}	T_{cws}	T_{cwr}	T_o	RH
COP	1.0000	0.3047	-0.4370	-0.3602	0.0667	0.1786	0.0290	0.3148
PLR	0.3047	1.0000	-0.6703	-0.1674	0.7132	0.2872	0.1931	0.0429
T_{chws}	-0.4370	-0.6703	1.0000	0.8438	-0.5334	-0.3471	0.0096	0.3216
T_{chwr}	-0.3602	-0.1674	0.8438	1.0000	-0.1929	-0.2535	0.1524	-0.1008
T_{cws}	0.0667	0.7132	-0.5334	-0.1929	1.0000	0.8038	0.5709	-0.6409
T_{cwr}	0.1786	0.2872	-0.3471	-0.2535	0.8038	1.0000	0.6817	-0.6814
T_o	0.0290	0.1931	0.0096	0.1524	0.5709	0.6817	1.0000	-0.8704
RH	0.3148	0.0429	0.3216	-0.1008	-0.6409	-0.6814	-0.8704	1.0000

表 3 变量间的皮尔森相关系数(制热工况)

参数	COP	PLR	T_{chws}	T_{chwr}	T_{cws}	T_{cwr}	T_o	RH
COP	1.0000	0.0956	0.2393	0.2292	-0.2431	-0.3181	-0.1407	0.2214
PLR	0.0956	1.0000	0.5080	0.0904	-0.4591	-0.2178	-0.0880	0.0248
T_{chws}	0.2393	0.5080	1.0000	0.9037	-0.1688	-0.0760	0.0839	0.0146
T_{chwr}	0.2292	0.0904	0.9037	1.0000	0.0330	0.0204	0.1408	0.0278
T_{cws}	-0.2431	-0.4591	-0.1688	0.0330	1.0000	0.9401	0.2838	-0.0160
T_{cwr}	-0.3181	-0.2178	-0.0760	0.0204	0.9401	1.0000	0.2889	-0.1464
T_o	-0.1407	-0.0880	0.0839	0.1408	0.2838	0.2889	1.0000	-0.9522
RH	0.2214	0.0248	0.0146	0.0278	-0.0160	-0.1464	-0.9522	1.0000

3.2 模型验证

考虑到运行参数在记录过程中,数据可能会包含一定的噪声和瞬态条件下的突变,导致在数据集中存在潜在的异常值,从而导致模型预测出现异常值。因此引入确定系数^[11]来检验随机森林模型预测地源热泵机组 COP 的精确度,具体结果如表 4 所示。确定系数 r^2 的值在测试数据集中均保持在 90% 以上,说明预测 COP 和实际 COP 的一致性较好。由此可见,随机森林模型用于地源热泵机组 COP 预测准确性较高。

表 4 两种工况下的确定系数

Tab.4 The determination coefficient under two working conditions

运行模式	$r^2/\%$		数据集预测离群值百分比/%	
	训练	测试	训练	测试
制冷工况	96.03	92.1	8.11	10.01
制热工况	97.69	93.5	7.23	9.56

3.3 输入变量重要度分析

表 5 列出了两种运行工况下输入变量对于预测机组 COP 的误差增加率以及相比于其他输入参数的重要度排名,重要度由高到低对应排名为 1~7,不同的运行工况下相同输入变量的重要度排名不同。制冷工况下对于预测机组 COP 最重要的前 3 个变量分别为:部分负荷率、地源侧回水温度、地源侧供水温度。制热工况下对于预测机组 COP 最重要的前 3 个变量分别为:部分负荷率、用户侧回水温度、地源侧回水温度。

表 5 变量重要度测量结果

Tab.5 Results of variable significance measurement

输入变量	制冷工况		制热工况		综合排序
	Δ MSE	重要度排序	Δ MSE	重要度排序	
PLR	60.54	1	75.64	1	1
T_{cwr}	45.14	2	37.60	3	2
T_{chw}	41.63	4	54.07	2	3
T_{cws}	44.65	3	24.79	5	4
T_{chws}	21.05	6	26.75	4	5
T_o	31.30	5	16.64	6	6
RH	27.52	7	11.56	7	7

3.3.1 部分负荷率

显然,部分负荷率的变化在地源热泵系统全年

运行当中对预测机组 COP 影响最大。图 2 和图 3 给出了夏季和冬季地源热泵机组 COP 和负荷率的配对关系。虽然两个参数间并没有强线性相关关系,但随着负荷率的逐渐增大,热泵机组 COP 呈现明显的整体上升趋势。数据显示,该系统在夏季负荷率低于 50% 工况下的工作时间占趋势记录时间的 66.3%,而冬季供暖模式下,系统负荷率低于 50% 的工作时间仅占 38.6%,夏季机组平均 COP 为 3.92,冬季机组平均 COP 为 4.15。表明该系统夏季长期处于较低的负荷率下运行,导致热泵机组运行效率相对较低,不利于热泵系统的长期稳定运行,而冬季负荷率相对稳定,机组 COP 较高,可见实际运行过程中根据机组部分负荷率的变化及时作出运行策略的改变是提高机组 COP 的关键,比如在满足制冷需求的前提下,可关闭一台制冷机组或者减少压缩机启动台数。

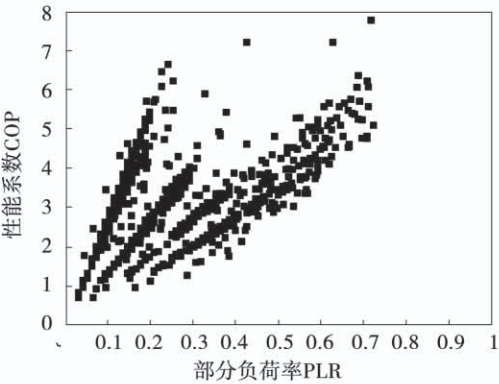


图 2 制冷模式机组 COP 与负荷率

Fig.2 COP and part load rate of cooling mode unit

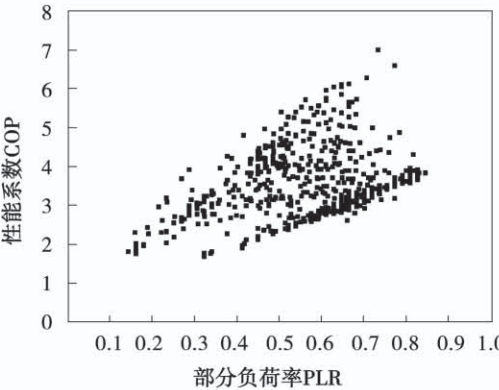


图 3 制热模式机组 COP 与负荷率

Fig.3 COP and part load rate of heating mode unit

3.3.2 地源侧回水温度

从重要度排名中可以发现,夏季工况下地源侧回水温度对机组 COP 影响的重要度优先于用户侧回水温度,排名第二,而冬季工况下则相反。从制冷循环角度分析,制冷模式下地源侧回水温度和供暖模式下用户侧回水温度均属于冷凝器侧入口流体温度,该侧流体温度的变化间接影响了热泵机组的冷凝温度,说明无论何种运行模式下,热泵机组的冷凝器侧流体温度变化对预测机组 COP 的重要度要高于蒸发器侧流体温度。从地埋管换热器角度分析,地埋管换热器作为地源热泵系统的核心部件,其换热性能能够直接影响热泵机组 COP。图 4 和图 5 给出了夏季和冬季地源侧供回水平均温度变化趋势,数据显示,夏季工况下地源侧回水平均温度为 28.7℃,换热平均温差为 2.1℃。8 月地源侧回水平均温度为 28.4℃,较 6 月升高 4.8℃,8 月供回水平均温差为 2.6℃,较 6 月降低 0.6℃,8 月机组平均 COP 为 3.41,较 6 月降低 1.05。冬季趋势记录时段内地源侧供水平均温度为 14.5℃,回水平均温度为 16.6℃,供回水平均温差 2.1℃。虽然夏季埋管侧流体换热增大,地埋管换热器换热量升高,但换热流体平均温度升高,导致机组冷凝温度升高,运行性能下降,而冬季埋管换热器相对夏季更为稳定,这是因为苏州位于夏热冬冷地区,该工程夏季冷负荷约为冬季热负荷的 1.6 倍,地源热泵系统全年运行导致土壤内部出现了冷热不平衡现象,产生了一定量的热堆积^[12],因此冬季工况下地埋管换热器换热更为稳定,对于机组运行性能的影响更小。实际工程中回水温度往往较难控制,可通过改变供水温度间接控制回水温度,达到改善机组 COP 的目的。

3.3.3 室外温度

输入变量室外干球温度是重要度排名前五的变量中唯一的不可控变量,由当地气候条件决定。而地源热泵系统相比于传统的风冷式系统优点在于利用土壤作为冷热源,土壤良好的蓄热能力使得地埋管换热器受外界气温变化影响较小。该地源热泵系统的埋管深度为 50 m,属于浅埋式地埋管,从图 4 和图 5 反映出地源侧流体平均温度与室外干球温度变化关系表明,夏热冬冷地区浅埋式地埋管受冬季

室外温度变化影响较小,相比于夏季有更为良好的表现。同时,也从侧面解释了表 5 中夏季工况下室外干球温度预测 COP 变化的重要度高于冬季工况的原因。

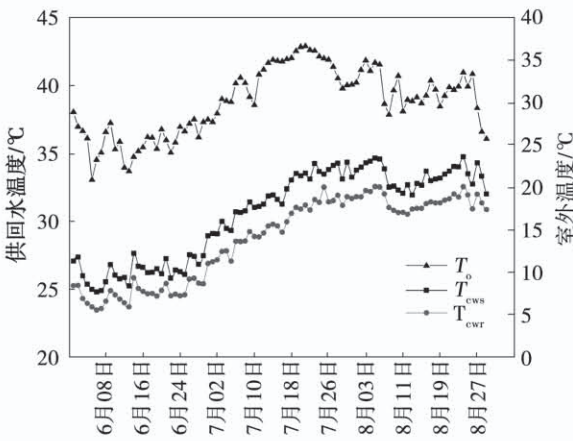


图 4 夏季地源侧供回水温度与室外温度
Fig. 4 Supply and return water temperature of source side and outdoor temperature in summer

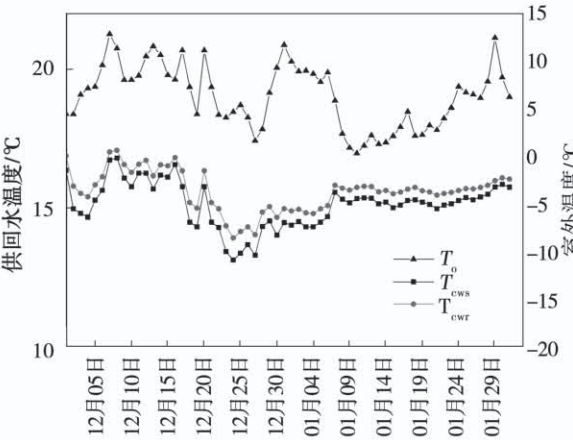


图 5 冬季地源侧供回水温度与室外温度
Fig. 5 Supply and return water temperature of source side and outdoor temperature in winter

此次研究分析的 7 个参数均属于既有地源热泵系统的动态运行参数,由 BACtalk 监控系统记录的热泵运行时的瞬时值,而地源热泵机组 COP 的变化往往由设计参数和运行参数等众多参数共同决定。设计参数诸如地源侧流体流量、用户侧流体流量、水泵功率等由于在系统运行中是定值,所以并未应用于探讨预测机组 COP 的重要度问题当中。

4 结 论

(1) 该地源泵系统夏季机组平均 COP 为 3.92, 地源侧回水平均温度为 28.7℃, 换热平均温差为 2.1℃; 冬季机组平均 COP 为 4.15, 地源侧回水平均温度为 16.6℃, 换热平均温差为 2.1℃。

(2) 随机森林模型精准地预测了地源热泵机组 COP 的变化, 确定系数达到 93.5%, 相比于皮尔森相关系数, 随机森林模型能够更深入分析系统各运行参数间非线性相关关系, 有利于区分参数影响机组 COP 变化的重要度。

(3) 部分负荷率是对机组 COP 变化影响最重要的因素, 地源侧回水温度次之, 但制热工况下用户侧回水温度的重要度要高于地源侧回水温度, 即地源热泵系统实际运行中针对不同负荷率下的运行调整以及提高对冷凝侧温度的控制将更有利于改善机组 COP, 提高机组运行性能。

参考文献:

- [1] 徐 伟, 刘志坚. 中国地源热泵技术发展展望[J]. 建筑科学, 2013, 29(10): 26-33.
XU Wei, LIU Zhi-jian. Development and prospect of ground source heat pump technology in China[J]. Building Science, 2013, 29(10): 26-33.
- [2] 刘 昱, 李慧星, 冯国会. 超低能耗建筑中土壤源热泵系统的应用分析[J]. 太阳能学报, 2018, 39(10): 2691-2698.
LIU Yu, LI Hui-xing, FENG Guo-hui. Application of ground source heat pump system in ultralow energy consumption building[J]. ACTA Energiae Solaris Sinica, 2018, 39(10): 2691-2698.
- [3] LIANG P, DI Q, XU L, LI Y. Optimization on the performance of ground heat exchangers for GSHP using Kriging model based on MOGA [J]. Applied Thermal Engineering, 2017 (118): 480-489.
- [4] 傅 强, 李峥嵘, 钱必华, 等. 上海某办公楼地埋管地源热泵系统夏季运行性能分析[J]. 暖通空调, 2015, 45(12): 88-94.
FU Qiang, LI Zheng-rong, QIAN Bi-hua, et al. Analysis on operat-

ing performance of the ground-coupled heat pump system in summer for an office building in Shanghai [J]. HV&AC, 2015, 45(12): 88-94.

- [5] MAGRANER T, MONTERO A, QUILIS S, et al. Comparison between design and actual energy performance of a HVAC-ground coupled heat pump system in cooling and heating operation [J]. Energy and Buildings, 2010, 42(9): 1394-1401.
- [6] SIVASAKTHIVEL T, MURUGESAN K, THOMAS H R. Optimization of operating parameters of ground source heat pump system for space heating and cooling by Taguchi method and utility concept [J]. Applied Energy, 2014, 116(3): 76-85.
- [7] WANG J, ZHOU P, HUANG G, et al. A Data mining approach to discover critical events for event-driven optimization in building air conditioning systems [J]. Energy Procedia, 2017, (143): 251-257.
- [8] 余坤勇, 姚 雄, 邱祈荣, 等. 基于随机森林模型的山体滑坡空间预测研究[J]. 农业机械学报, 2016, 47(10): 338-345.
YU Kun-yong, YAO Xiong, QIU Qi-rong, et al. Landslide spatial prediction based on random forest model [J]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(10): 338-345.
- [9] Meyer D, Leisch F, Hornik K. The support vector machine under test [J]. Neurocomputing, 2003, 55(1): 169-186.
- [10] 王 盼, 陆宝宏, 张瀚文, 等. 基于随机森林模型的需水预测模型及其应用[J]. 水资源保护, 2014, 30(1): 34-37.
WANG Pan, LU Bao-hong, ZHANG Han-wen, et al. Water demand prediction model based on random forests model and its application [J]. Water Resources Protection, 2014, 30(1): 34-37.
- [11] Yu F, Ho W, Chan K, et al. Critique of operating variables importance on chiller energy performance using random forest [J]. Energy Build, 2017, (139): 653-664.
- [12] 赵 耀, 饶政华, 伊若璇, 等. 夏热冬冷地区地源热泵长期运行对土壤热平衡的影响 [J]. 可再生能源, 2013, 31(10): 74-79.
ZHAO Yao, RAO Zheng-hua, YI Ruo-xuan, et al. Effects of long-term operation of ground-source heat pump on soil heat balance in hot summer and cold winter region [J]. Renewable Energy Resources, 2013, 31(10): 74-79.

(丛 敏 编辑)