

数据驱动技术在建筑能耗模拟中的应用研究*

张 涛, 王雯翡, 成雄蕾, 王飒枫

(中国建筑科学研究院有限公司, 北京 100013)

摘 要: 随着城市化快速发展以及人类对舒适度的要求不断提高, 建筑能源消费增长迅速, 建筑能耗建模和预测对于解决建筑节能问题至关重要。在综合智能化的基础设施发展的背景下, 数据驱动技术依赖于数据分析和机器学习, 为建筑能耗的预测提供灵活的方法。对传统建筑能耗建模方法进行了回顾, 分析了不同模拟软件的优缺点和适应性。对近年来出现的数据驱动建筑能耗模型方法进行了分析, 重点考虑输入数据特性和数据预处理对算法的精确度影响。最后分析了数据驱动方法的局限性和未来发展方向。

关键词: 数据驱动; 建筑能耗模拟; 机器学习; 负荷预测

中图分类号: TU111.1

文献标识码: A

DOI: 10.15913/j.cnki.kjycx.2020.16.072

1 引言

建筑用能占全球能源消费的 30%, 约 1/3 的二氧化碳排放与其相关^[1]。显然, 建筑节能已成为建筑可持续发展的热点, 开展建筑节能工作十分必要^[2]。建筑能耗模拟涉及到机理复杂的热湿传递过程, 是一个复杂非线性、多扰动系统, 针对建筑能耗预测问题相关学者提出了多种预测方法, 如图 1 所示, 包括物理模型法、基于统计学的方法。

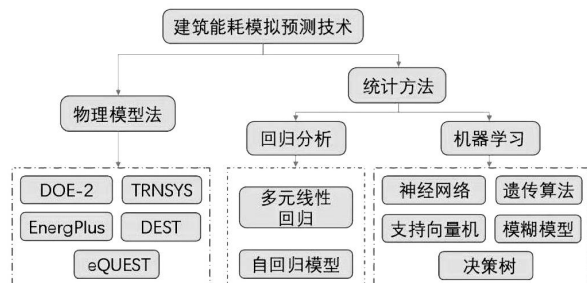


图 1 建筑能耗建模与预测方法的汇总分类

基于统计的方法包括回归法和数据驱动法, 多元线性回归法建立显式的数学表达式模型来预测建筑能耗, 其优点是模型结构简洁, 缺点是预测精度较差, 可移植性不强^[3-4]。数据驱动法包括支持向量机法 (Support Vector Machine, SVM)、人工神经网络法 (Artificial Neural Network, ANN)、遗传算法 (Genetic Algorithm, GA)、决策树 (Decision Tree, DT) 等。

本文首先针对传统建筑能耗建模方法进行回顾, 分析不同模拟软件的优缺点和适应性。然后针对近年来出现的数据驱动建筑能耗模型方法进行分析, 重点考虑输入数据特性和数据预处理对算法精确度的影响。最后分析数据驱动方法的局限性和未来发展方向。

2 建筑能耗模拟分析软件

一般认为 DEST、EnergyPlus 等在建筑负荷计算表现优良, EQUEST、TRNSYS 等更适合系统模拟^[5]。而对于有明显的温度分层和对流传热的高大空间建筑, 一般通过引入新的传热推导理论和数学方法进行分析研究, 如考虑高大空间气流和温度分布, 再结合模拟软件分析获得能耗参数。以 EnergyPlus 软件为例, 允许用户实现建筑的几何建模、材料参数设置和其内暖通空调系统特性设置, 允许用户设置设计日和年度模拟, 对于设计日的模拟, 用户需设置如最大值和最小值干球温度、湿球温度等典型日的参数。对于年度的模拟需要提供如干/湿球温度、太阳辐射、风速/风向等天气参数。软件计算结果包括区域温度、建筑能源消费、冷/热负荷等^[6]。建筑能耗模拟预测模型框架如图 2 所示。

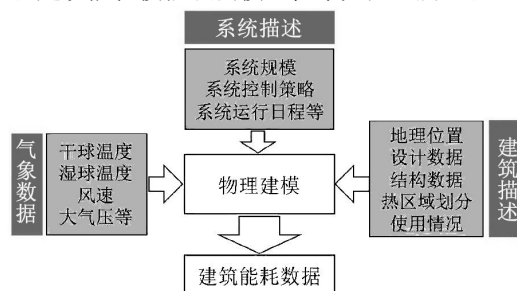


图 2 建筑能耗模拟预测模型框架

3 数据驱动技术

3.1 回归分析预测法

回归分析预测法是通过在两上或两个以上的自变量与一个因变量的相关分析, 建立预测模型进行预测的方法。当自变量与因变量之间存在线性关系时, 称为多元线性回归分析。多元线性回归预测方法建立显式的数学表达式模型来预

* [基金项目] 国家重点研发计划项目“基于全过程的大数据绿色建筑管理技术研究与示范”(编号: 2017YFC0704200)

测建筑能耗,其优点在于模型结构简洁,易于建立;不足之处在于对非线性系统需要人为确定部分模型参数,预测精度较差,可移植性不强。多重共线性指在多元线性回归方程中,自变量之间有较强的线性关系,这种关系若超过了因变量与自变量的线性关系,则回归模型的稳定性受到破坏,回归系数估计不准确。

3.2 机器学习方法

机器学习预测建筑能耗需要基于建筑历史数据的训练,可较好反映建筑实际运行能耗特征。K-nearest neighbors (KNN) 是一种模式聚类技术,该方法给定一组历史观测值(能源消费和其他变量),首先创建集群,针对用户定义的特性构造峰值负载、平均值、日负荷变化的大小、日消耗(积分)等。计算每个时间序列的特征后进行分类。使用简单的欧几里德距离表征能源消耗的差异,通过观测数据与集群进行比较,确定两个标准的亲密程度,进而进行预测。该方法预测精度高,且适用于具有小时时间步长的短期预测。如文献[7]基于机器学习的 K-means 和 KNN 算法,对上海某酒店建立记忆网络模型,利用处理后的能耗数据预测逐时能耗数据。预测结果表明,能耗异常数据识别与修复方法能准确识别并修复建筑能耗异常数据,提高能耗预测效果。

建筑能源模型分为短期、中期和长期能源预测,影响结果的相关因素包括人为因素、社会因素和环境气候等,预测十分困难。机器学习模型能基于先前记录的数据准确地预测能耗,目前使用最为广泛的智能模型大多是基于人工神经网络(ANN)发展而来^[8]。SVM 是一种广泛应用于信号处理、最优控制和系统中建模的统计学习方法,适用于有限样本回归,具有良好的回归泛化能力。SVM 的思想是选择适当的非线性映射函数,映射输入变量的低维空间到高维特征空间。

4 建筑能耗输入数据影响

输入数据集具有的不同的特性,直接影响建模和预测精度。基于文献调研,数据来源主要分为三大类:真实的数据、模拟的数据、基准数据,分别占所研究数据的 64%、20%和 16%。实际数据直接从调控中心进行读取或现场调查,模拟数据是基于物理的模型计算而得,例如使用 EnergyPlus、TRNSYS 等工具获得。基准数据来自公开提供的数据集。建筑能耗输入数据集中的特征可分为四大类:天气数据等所有与室外条件有关的数据,建筑物的描述和室内条件环境,居住者行为,时间序列的建筑物运行数据。

5 结论

本文分析了主要建筑能耗建模和预测技术,从传统的建

筑模拟程序,到回归模型、统计回归模型、KNN、决策树、支持向量机、人工神经网络等方法。分析不同方法的优缺点,尽管数据驱动方法提供了非常广泛的工具来建模,且预测模型能适应许多不同情况的能源消耗,但是针对不同建筑物的类型、可用数据等,均需具体问题具体分析,利用不同方法求解。目前针对具体不同类型的建筑,采用特定预测方法开展预测研究较多。而通过将数据驱动能耗预测模型嵌入到建筑能效管理平台中,进一步节约资源、提高能源效率的研究并不多见,未来这方面的研究具有重要意义。

参考文献:

- [1] BOURDEAU M, ZHAI X Q, NEFZAOU E, et al. Modeling and forecasting building energy consumption: a review of data-driven techniques [J]. Sustainable Cities and Society, 2019(48): 101533.
- [2] 吴盼红. 知识与数据驱动的建筑电能需求智能预测方法研究 [D]. 济南: 山东建筑大学, 2019.
- [3] 刘倩, 杨昌智, 胡攀. 长沙市商场建筑能耗影响因素与预测模型构建研究 [J]. 建筑科学, 2017, 33(12): 96-102.
- [4] 周峰, 张立茂, 秦文威, 等. 基于 SVM 的大型公共建筑能耗预测模型与异常诊断 [J]. 土木工程与管理学报, 2017, 34(6): 80-86.
- [5] 孙一晁, 吕浩宇, 陈毅兴, 等. 基于 EnergyPlus-Python 联合模拟和强化学习算法的室内环境控制优化 [J]. 建设科技, 2019(24): 52-58.
- [6] NETO A H, FIORELLI F A S. Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption [J]. Energy and Buildings, 2008, 40(12): 2169-2176.
- [7] 高英博, 顾中煊, 罗淑湘, 等. 能耗预测导向的建筑能耗异常数据识别与修复 [J]. 科学技术与工程, 2019, 19(35): 298-304.
- [8] DEB C, EANG L S, YANG J, et al. Forecasting diurnal cooling energy load for institutional buildings using artificial neural networks [J]. Energy and Buildings, 2016(121): 284-297.

作者简介: 张涛(1987—), 男, 博士, 工程师, 研究方向为建筑节能技术、能源互联网。

[编辑: 严丽琴]