

文章编号: 1671-6612 (2021) 03-352-08

基于灰色多元线性回归的 上海某办公建筑空调能耗预测模型

豆晨昊 徐 峰

(上海交通大学船舶海洋与建筑工程学院 上海 200240)

【摘 要】 能耗预测可以为建筑管理人员提供空调系统未来运营期间的能耗额以及增长趋势, 以便制定相应的节能策略。主要讨论空调能耗的影响因素及各因素的影响程度, 将灰色模型和多元线性回归模型结合, 基于灰色多元线性回归建立空调能耗预测模型。首先, 以上海某办公建筑为例, 通过灰色模型对空调能耗的异常能耗值及室外天气温度、湿度的观测值进行修正, 降低数据的随机波动性, 得到灰色模型修正数据; 然后, 计算影响因素与空调能耗额之间的相关性系数, 讨论各因素对空调能耗的影响程度; 最后, 利用灰色模型修正数据建立灰色多元线性回归空调能耗预测模型, 并将空调能耗预测值和实际值相对比, 结果表明, 构建的灰色多元线性回归预测模型符合多元线性回归建模要求, 并具有较高的能耗预测精度。

【关键词】 灰色理论; 办公建筑; 灰色多元线性回归模型; 空调能耗
中图分类号 TU831 文献标识码 B

Air Conditioning Systems Energy Consumption Prediction Model of Office Buildings Based on Grey-multiple Linear Regression Theory

Dou Chenhao Xu Feng

(School of Naval Architecture, Ocean & Civil Engineering, Shanghai Jiaotong University, Shanghai, 200240)

【Abstract】 Energy consumption prediction model can provide building managers with energy consumption value and growth trend during the futural operation, so as to facilitate managers to make relevant energy-saving strategies. This paper mainly discusses the influence degree of influencing factors on air conditioning system and establishes the air conditioning energy consumption prediction model based on the combination of grey model and multiple linear regression model. Firstly, the grey model is used to modify the air conditioning energy consumption value of an office building in Shanghai and the observation value of outdoor weather temperature and humidity in 2018, which can reduce the random volatility of collected data. Secondly, calculating correlation coefficient between the influencing factors and air conditioning energy consumption, in order to explore the influence degree of the energy influencing factors. Finally, the grey multiple linear regression energy consumption prediction model is established by using the emendated data. This paper analyzes the influence degree of influencing factors on air conditioning energy consumption, and established grey multiple linear regression prediction model meets the requirements of multiple linear regression modelling and obtains high accuracy of energy consumption prediction.

【Keywords】 Grey theory; office buildings; grey-multiple linear regression model; air conditioning systems energy consumption

基金项目: 国家科技部国家重点研发计划子课题“基于 BIM 的绿色建筑运营成本控制技术(编号: 2018YFC0705901-04)”

作者简介: 豆晨昊(1995-), 男, 在读硕士研究生, E-mail: douchen hao1995@163.com

通讯作者: 徐 峰(1970-), 男, 硕士生导师, 副教授, E-mail: F.Xu@sjtu.edu.cn

收稿日期: 2020-08-17

0 引言

2018 年上海市住房和城乡建设管理委员会发布的《2018 年度上海市国家机关办公建筑和大型公共建筑能耗监测及报告分析》^[1] (以下简称《报告》) 的数据显示, 办公建筑年单位面积耗电量年耗电量为 97.4kWh/m², 其中供热季能耗约为 32kWh/m², 过渡季能耗约为 27.4kWh/m², 制冷季能耗约为 38kWh/m²。从整体能耗角度来看 (见图 1), 办公建筑能耗占所有建筑类型能耗的 28.1%, 属于上海市公共建筑能耗最高的建筑类型。

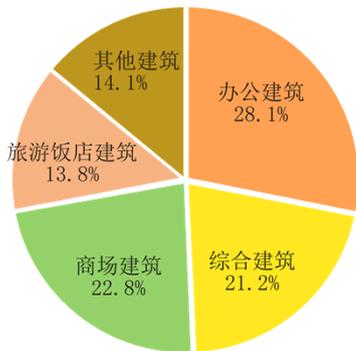


图 1 2018 年上海市不同类型建筑年耗能占比

Fig.1 Proportion of different types of buildings annual energy consumption in Shanghai, 2018

办公建筑用电能耗结构 (见图 2) 中空调系统能耗占比 29.8%, 是办公建筑运营期间能耗大项, “十三五” 国家发展规划纲要中提出单位 GDP 能耗下降 20% 的节能目标, 其中建筑节能承担着巨大的节能减排任务, 构建空调能耗预测模型预测未来空调系统的能耗, 采取相应的措施降低空调系统能耗是很有必要的。

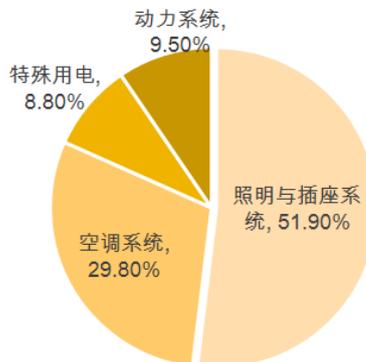


图 2 2018 年上海市办公建筑不同用电系统年耗能占比

Fig.2 Proportion of different power consumption systems of office buildings annual energy consumption in Shanghai, 2018

江亿^[2]等人研究了不同情况下空调系统的节能潜力, 表明空调系统在使用得当的条件下可以进一步降低能耗。空调能耗预测模型可以预测建筑未来能耗增长趋势, 研究空调系统运行能耗的特点以及影响因素, 为管理者开展建筑节能工作提供参考。

目前许多学者采用了不同的预测方法预测建筑能耗, 在智能算法方面, 王定奥^[3]等人利用 DeST 软件模拟建筑内多因素影响下的空调运行状态得出仿真数据, 然后构建多变量参数预测数据的 BP 神经网络对仿真数据进行拟合, 得出空调能耗预测模型。陈彦熹^[4]等人基于神经网络建立办公建筑的空调风系统和地源热泵空调预测模型, 该模型精度高, 而且需要输入的初始参数较少, 仅需建筑运行数据和气象参数即可得出能耗预测值。廖文强^[5]等人考虑到数据的时序性和非线性关系, 利用可以很好地适应这种特点的长短期记忆神经网络, 构建暖通空调系统能耗预测模型。由于利用传统的能耗分析软件建立的能耗预测模型不能很好地适用时刻改变的环境参数, Xu Yang^[6]等人考虑建筑内部环境参数, 利用 ELM (extreme learning machine) 算法构建了空调能耗预测模型。

在利用历史数据建模的方法上, 周芮锦, 潘毅群^[7]等人利用时间序列法预测和分析商业建筑能耗, 构建全年建筑总能耗的预测模型。马忠娇^[8]等人通过分析历史数据, 假定天气情况, 工作日与非工作日条件相同的条件下, 建立基于室外空气温度和照明插座用电量相似的组合权重的精确空调能耗预测模型。验证表明, 该模型实际误差在 5% 以内。杨柳^[9]等人首先利用 Energy Plus 分析办公建筑的能耗影响性因素, 然后利用多元线性回归法建立建筑能耗预测模型。

以上提到的能耗预测模型建模方法主要有时间序列法、人工神经网络法、相似工况预测法以及多元线性回归法。时间序列法是考虑到时间顺序排列的数据预测未来的方法, 但是这种方法局限性在于只适合中短期预测, 当外界因素发生变化时, 预测结果有时不符合实际规律^[10]。人工神经网络法是一种是模拟人脑神经元的智能算法, 可以适应复杂多变的情况, 具有自我学习的功能, 但局限性在于需要大量的数据, 并且无法研究模型中每一个参数的特点, 不能具体讨论每一个影响因素对因变量的影响程度大小。相似工况预测在历史日和预测日相

似的前提下具有较高的精度，构建的模型依赖于历史数据，如果未来的预测日和历史日相似度比较低，则会产生较大的偏差^[8]。

多元线性回归的方法可以充分考虑各个影响因素的特点，分析影响因素对空调能耗的影响程度。利用灰色多元线性回归模型对数据要求不高，无需大量的时间进行模型训练，省时省力，并且不过分依赖于历史数据。本文先以灰色理论修正数据，再分析各个影响因素与空调能耗之间的相关性，以及各个因素对空调能耗的影响程度，并将灰色理论和多元线性回归模型结合，以灰色模型修正过的数据构建基于灰色多元线性回归的空调能耗预测模型。

1 模型建立

原始观测数据若存在较大的随机波动性，建立预测模型时不利于发现规律，且建立的模型预测准确性也不高，考虑到这个问题，先利用 GM(1,1)修正观测数据。

1.1 GM(1,1)模型

灰色模型多用于预测以及修正波动性较大的观测数据。在建模时，若原始数据受到外界扰动因素影响波动性较大，呈现的规律性不明显，则会降低模型的预测精度。适当修正原始数据可降低数据的随机波动性、减少误差，进而更能反映数据本身的规律以及预测的趋势，提高数据的预测精度。GM(1,1)模型在其他领域也得到广泛的应用，李俊峰^[11]利用 GM(1,1)模型估计上海市的发电量；宋晓华^[12]利用 GM (1,1) 模型修正原始序列，将 GM(1,1)模型和支持向量机 (SVM) 相结合，建立长期电量预测模型，将预测相对误差控制在 2.06%。

本文将 GM(1,1)模型与多元线性回归模型相结合，用 matlab 实现 GM(1,1)模型，再利用 GM(1,1)模型修正原始序列，降低原始序列的波动性，然后以修正的原始数据建立办公建筑空调系统灰色多元线性回归能耗预测模型。

若设原始数据序列为：

$$X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)\}$$

且 $X^{(0)}(h) \geq 0, h=1, 2, \dots, n$ ； $X^{(1)}$ 是 $X^{(0)}$

的一次累加序列。

$$X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), x^{(1)}(3), \dots, x^{(1)}(n)\}$$

一次累加方法为：

$$x^{(1)}(h) = \sum_{i=1}^h x^{(0)}(i); h=1, 2, 3, \dots, n \quad (1)$$

$Z^{(1)}$ 是 $X^{(1)}$ 的紧邻均值生成序列：

$$Z^{(1)} = \{z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(h)\}, h=1, 2, 3, \dots, n$$

其中，定义 $z^{(1)}(h)$ 为：

$$z^{(1)}(h) = \frac{1}{2} [x^{(1)}(h-1) + x^{(1)}(h)]; h=1, 2, 3, \dots, n \quad (2)$$

GM(1,1)所对应的一阶一元微分方程：

$$\frac{dx^{(1)}}{dt} + \alpha x^{(1)} = u \quad (3)$$

其中， α 、 u 为待解系数， α 为发展系数，

u 为灰色作用量。记 $\hat{\alpha} = \begin{pmatrix} \alpha \\ u \end{pmatrix}$ ，根据最小二乘法计

算 α 、 u 。定义 Y_n 、 B 为：

$$Y_n = \begin{pmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \mathbf{M} \\ x^{(0)}(n) \end{pmatrix} \quad B = \begin{pmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \mathbf{M} & \mathbf{M} \\ -z^{(1)}(n) & 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2}(x^{(1)}(1)+x^{(1)}(2)) & 1 \\ -\frac{1}{2}(x^{(1)}(2)+x^{(1)}(3)) & 1 \\ \mathbf{M} & \mathbf{M} \\ -\frac{1}{2}(x^{(1)}(n-1)+x^{(1)}(n)) & 1 \end{pmatrix}$$

利用最小二乘法计算： $\hat{\alpha} = (B^T B)^{-1} B^T Y_n$ ，

得出 α 、 u 。

将灰参数代入

$$\frac{dx^{(1)}}{dt} + \alpha x^{(1)} = u \quad (4)$$

得：

$$\hat{x}(t+1) = \left(x^{(1)}(1) - \frac{u}{\alpha} \right) e^{-\alpha t} + \frac{u}{\alpha} \quad (5)$$

将上述结果累减还原：

$$\hat{x}^{(0)}(t+1) = \hat{x}^{(1)}(t+1) - \hat{x}^{(1)}(t) \quad (6)$$

得到原始数据的修正值为：

$$\hat{x}^{(0)} = \{\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \hat{x}^{(0)}(3), \dots, \hat{x}^{(0)}(n)\}$$

1.2 相关性分析

相关性分析主要目的是研究变量间关系的密切程度^[13], 在统计分析中, 常利用相关系数定量地描述两个变量线性关系的紧密程度, 相关系数取值区间为[-1,1]。线性相关系数绝对值越大则表明两变量之间的相关性越强。本文采用 Pearson 相关分析, 来研究温度和湿度与空调能耗额之间的相关性。

Pearson 相关系数:

线性相关是用来度量具有线性关系的两个变量之间相关密切程度以及相关方向, 线性相关系数又称 Pearson 相关系数, 也称积差相关系数。

(1) P 值 (也称显著性值或 sig 值), 判断两个变量之间是否存在显著相关关系。

如果 $P < 0.05$, 则表明两者之间具有相关性。 $P < 0.01$ 使用 2 个*号表示; $P < 0.05$ 使用 1 个*号表示。

(2) 相关系数: 判断两者之间相关性的强弱。

相关系数绝对值	相关性强弱
0.8-1.0	极强相关
0.6-0.8	强相关
0.4-0.6	中等程度相关
0.2-0.4	弱相关
0.0-0.2	极弱相关或无相关

相关系数为:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 (y_i - \bar{y})^2}} \quad (7)$$

其中, n 为样本容量, x_i 和 y_i 为所检验相关性的两变量对应的样本值

t 检验统计公式为:

$$t = \frac{\sqrt{n-2}}{\sqrt{1-r^2}} \quad (8)$$

即: t 统计量服从自由度为 $n-2$ 的 t 分布。

若 $t > t_{\frac{\alpha}{2}}$, 则认为变量在 α 水平下相关。

若 $t < t_{\frac{\alpha}{2}}$, 则认为变量在 α 水平下不相关。

1.3 多元线性回归模型

1.3.1 多元线性回归模型建模的步骤

在线性回归分析中, 影响因变量变化的因素包含两个或两个以上的自变量, 且因变量与自变量之间存在线性关系, 则称为多元线性回归分析。

此时:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_n x_{in} + \varepsilon_i \quad (9)$$

$(i = 1, 2, 3, \dots, K, n)$

其中, y 为可观测的随机变量, $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ 为 $(n+1)$ 个未知参数, $(x_1, x_2, x_3, \dots, x_n)$ 和 y 为 n 次观测的 n 组观测值, $(x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}, y_i)$, ε_i 为随机误差, 且满足以下假设:

(1) 正态性假设: 服从均值为 0, 方差为 σ^2 的正态分布。

(2) 无偏性假设: ε_i 的条件期望值为 0。

(3) 独立性假设: ε_i 之间相互独立, 满足 $Cov(\varepsilon_i, \varepsilon_j) = 0, (i \neq j)$ 。

(4) 同共方差性假设: ε_i 所有的方差都相同。

此处, 设 $\varepsilon_i = 0$ 。

多元线性回归方程就是利用现有的观测数据, 求得 $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ 的估计值 $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_n$ 即得出多元线性回归方程。

$$\text{令: } Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & K & x_{1m-1} \\ 1 & x_{21} & x_{22} & \dots & M & x_{2m-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & K & x_{nm-1} \end{pmatrix},$$

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{m-1} \end{pmatrix}$$

以矩阵形式表示即为:

$$Y = X\beta \quad (10)$$

用最小二乘法估计回归方程系数 β :

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (11)$$

1.3.2 多元线性回归模型的检验

利用 SPSS22.0 建立多元线性回归分析模型, 得出模型后需要检验分析模型是否符合建模要求,

检验方法主要分为相关系数检验、F 检验和 t 检验。

具体而言：

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_n x_{in} + \varepsilon_i, (i=1,2,3,\dots,n) \quad (12)$$

假设 $H_0: \beta = \beta_1 = \beta_2 = \dots = \beta_n = 0$

假设 $H_1: \beta_i (i=1,2,3,\dots,n)$ 不全为 0。

若假设 $H_0: \beta = \beta_1 = \beta_2 = \dots = \beta_n = 0$ 成立，表明 y_i 与 x_{in} 之间不存在任何线性关系，则建立模型无效。

(1) 相关系数的检验

$$R_2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 / p}{\sum_{i=1}^n (y_i - \bar{y})^2 / (n - p - 1)} = \frac{SSR/p}{SST/n - p - 1} \quad (13)$$

其中， p 和 $(n-p-1)$ 为自由度， $SST=SSR+SSE$

判断条件： R^2 值越接近于 1，说明 y, x 之间的相关性越强。若 R^2 为 0，则说明 y 与 x 之间无关。

(2) F 检验

$$F = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 / p}{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n - p - 1)} = \frac{SSR/p}{SSE/(n-2)} \quad (14)$$

$$= \frac{R^2}{1-R^2} (n-2) : F(1, n-2)$$

确定显著性水平 α 后，得出临界值 $F_\alpha(p, n-p-1)$ 。

判断条件：

若 $F > F_\alpha(p, n-p-1)$ ，则拒绝原假设假设 H_0 ，认为多元线性回归在显著性水平 α 总体显著。

若 $F < F_\alpha(p, n-p-1)$ ，则接受原假设假设 H_0 ，认为多元线性回归在显著性水平 α 总体不显著。

(3) t 检验

$$t = \frac{\hat{\beta}_1}{\frac{\sigma}{\sqrt{\sum_{j=1}^n (x_{ij} - \bar{x}_{i1})^2}}} : t(n-p-1) \quad (15)$$

其中，

$$\hat{\beta}_1 : N\left(\frac{\sigma^2}{\sum_{j=1}^n (x_{ji} - \bar{x}_1)^2}\right) \quad (16)$$

确定显著性水平 α 后，得出临界值 $t_{\frac{\alpha}{2}}(n-p-1)$ 。

判断条件：

若 $|t| > t_{\frac{\alpha}{2}}(n-p-1)$ ，则拒绝原假设假设 H_0 ，

认为多元线性回归在显著性水平 α 总体显著。

若 $|t| < t_{\frac{\alpha}{2}}(n-p-1)$ ，则接受原假设假设 H_0 ，

认为多元线性回归在显著性水平 α 总体不显著。

2 案例应用

本文以上海某办公建筑为例建立灰色多元线性回归模型，验证灰色理论与多元线性回归模型结合的可行性。由于每年过渡期间空调使用较少，而供冷季空调使用量更加集中，分析供冷季空调能耗与天气温度、湿度之间的关系更能体现出空调系统运行的特点。节假日、休息日以及下班期间空调系统处于关闭状态，此期间的能耗数据不具有研究代表性，故本案例采用 2018 年的 6、7、8 月的上海某办公建筑的除去节假日、周末休息日的正常工作期间每天该建筑上班时间的 7 点-17 点的空调系统能耗额，以及上海 2018 年室外天气温度和湿度的历史数据。表 1 为截取分析数据观测值的某一天的数据。

表 1 截取分析数据某一天的分析数据

Table 1 Analysis data of one day from original data					
时间	空调系统 (kWh)	冷热站 (kWh)	空调末端 (kWh)	温度 (°C)	湿度 (%)
2018/7/12 7:00	157.05	123.16	33.89	28	83
2018/7/12 8:00	338	265.63	72.38	29.1	78
2018/7/12 9:00	237.69	185.59	52.11	28.8	81
2018/7/12 10:00	333.7	257.62	76.09	29.7	77
2018/7/12 11:00	242.23	187.11	55.12	30.7	71
2018/7/12 12:00	337.98	260.78	77.2	31.1	68
2018/7/12 13:00	233.15	182.46	50.69	31.7	68
2018/7/12 14:00	271.74	216.02	55.72	32	65
2018/7/12 15:00	266.4	211.13	55.27	31	69
2018/7/12 16:00	265.24	209.61	55.63	30.9	70
2018/7/12 17:00	49.22	38.39	9.83	30.7	71

注：(1) 该建筑的上班时间为 8:30-17:00；(2) 8-9 时的空调能耗比 7 时显著增高，这是由于员工上班后空调启动时的能耗比正常运行时能耗大。

2.1 修正观测数据

由于空调开启时的能耗异常值以及建筑内人员的行为因素导致了数据的波动性变大，为了增加

模型的预测精度, 首先利用 GM(1,1)对原始数据序列进行数据修正处理, 得到原始数据的修正值, 此处截取某一天的空调能耗数据为例, 如表 2 所示。

表 2 GM(1,1)模型修正某一天空调能耗值

Table 2 Amended air-conditioning system energy consumption value of one day by GM(1,1) model

时间	空调能耗 观测值(kWh)	GM(1,1)模型 修正值 (kWh)
2018/7/12 7:00	157.05	216.95
2018/7/12 8:00	338	217.05
2018/7/12 9:00	237.69	217.14
2018/7/12 10:00	333.7	217.24
2018/7/12 11:00	242.23	217.34
2018/7/12 12:00	337.98	217.44
2018/7/12 13:00	233.15	217.55
2018/7/12 14:00	271.74	217.65
2018/7/12 15:00	266.4	217.75
2018/7/12 16:00	265.24	217.85
2018/7/12 17:00	49.22	218.06

此时输入空调能耗的观测值:

$$X^{(0)} = \begin{bmatrix} 157.4, 345.73, 239.42, 334.46, \\ 254.72, 346.4L \end{bmatrix}$$

输出的空调能耗修正值为:

$$\hat{X}^{(0)} = \begin{bmatrix} 262.34, 262.47, 262.59, 262.71, \\ 262.83, 262.95L \end{bmatrix}$$

观察这一天的数据, 空调能耗在 8 时、10 时、12 时和 17 时中有突变值, 主要是因为 8:30、9:30 员工上班相继开启空调, 造成这些时段空调能耗突然增高。11:20 是员工的午餐时间, 办公楼内办公区域的空调仍然开启, 但是办公楼餐厅人流量增加, 局部过密的人流导致空调能耗也相应会增加。在 17:00 左右员工陆续下班关闭空调, 所以空调能耗骤降。由于这些外界影响因素导致了空调能耗值产生了异常值, 增加了数据的波动性, 若这些异常的空调能耗值参与空调能耗预测模型建模, 对于其他正常时间段空调能耗的预测是有影响的, 但经 GM(1,1)模型优化异常的空调能耗值, 削减了过高的异常值, 可以提高空调能耗预测模型的精度。

2.2 相关性分析

2.2.1 散点图分析

利用 SPSS22.0 绘制散点图, 通过散点图可以直观地看出变量之间的相关性程度。散点图结果如图 3、4 所示。

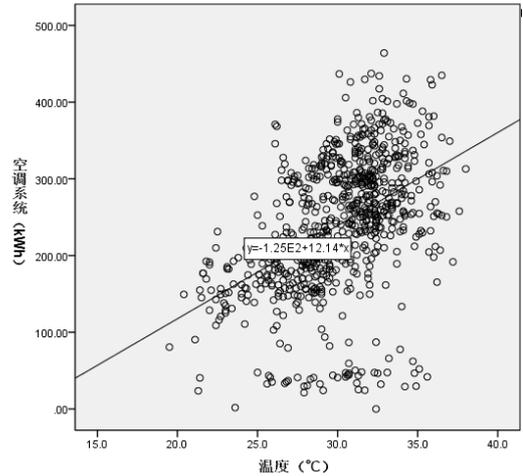


图 3 温度与空调系统能耗之间的散点图

Fig.3 Scatter diagram of temperature and air conditioning system energy consumption

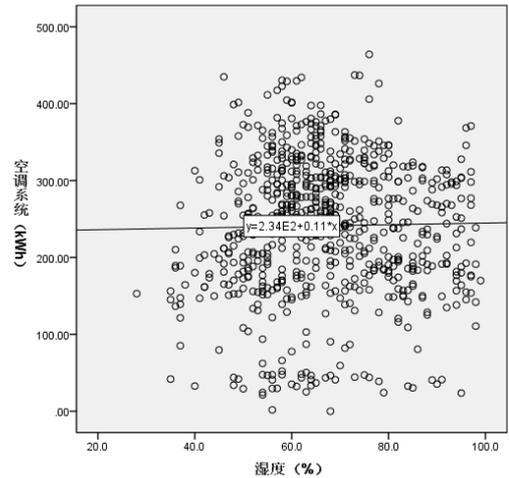


图 4 湿度与空调系统能耗之间的散点图

Fig.4 Scatter diagram of humidity and air conditioning system energy consumption

从空调系统能耗与室外温度、湿度的散点分布图可以看出制冷季期间的温度明显与空调系统能耗呈现正相关性, 但是湿度与空调系统能耗之间的相关性并不明显。

2.2.2 统计学分析

本文采用 Pearson 相关性分析, 为了对比灰色模型修正前、后数据的特点, 本文用 SPSS22.0 软件对温度与空调能耗、湿度与空调能耗进行相关性

分析,修正前后的结果对比如表 3 所示。

表 3 修正数据前后温度与空调能耗之间的相关性分析结果对比

Table 3 Comparison of prior to and after modification correlation analysis results between temperature and air conditioning energy consumption

		温度 (°C)		等效电能耗(kWh)	
		优化前	优化后	优化前	优化后
空调 能耗	Pearson 相关	1	1	0.489**	0.999**
	显著性(双尾)			0.000	0.000
	N	829	829	829	829
温度	Pearson 相关	0.489**	0.999**	1	1
	显著性(双尾)	0.000	0.000		
	N	829	829	829	829

注: **相关性在 0.01 层上显著(双尾)。

结果分析:数据修正后, Pearson 相关系数由原来的 0.489 上升到 0.999,修正后的数据明显比修正前的数据与空调能耗有更明显的相关性,呈极强相关。这是由于 GM(1,1)修正后的数据削减了异常能耗值,降低了数据的随机波动性,进而增强空调能耗与温度之间的相关性。

表 4 修正数据前后湿度与空调能耗之间的相关性分析结果对比

Table 4 Comparison of prior to and after modification correlation analysis results between humidity and air conditioning energy consumption

		湿度 (°C)		等效电能耗 (kWh)	
		优化前	优化后	优化前	优化后
空调 能耗	Pearson 相关	1	1	-0.151**	0.447**
	显著性(双尾)			0.000	0.000
	N	829	829	829	829
湿度	Pearson 相关	-0.151**	0.447**	1	1
	显著性(双尾)	0.000	0.000		
	N	829	829	829	829

注: **相关性在 0.01 层上显著(双尾)

结果分析:原始数据中每天中午时湿度最低的时候,但是此时空调能耗却比较高,所以导致了呈现非常弱的负相关的错误相关性归分析结果。优化后的数据分析结果呈现的是中等相关,符合空调系

统运行能耗的特点。从 Pearson 相关系数来看,数据修正后的湿度与空调能耗之间有更明显的相关性,呈中等程度相关。

温度与空调能耗之间的 Pearson 相关系数要大于湿度与空调能耗之间的相关系数,表明温度与空调能耗之间的相关性更紧密,温度对空调能耗的影响程度要大于湿度。

2.3 建立灰色多元线性回归模型

经 GM (1.1) 修正后的数据结合多元线性回归模型,建立灰色多元线性回归空调能耗预测模型,并对灰色线性回归模型的拟合程度进行检验,相关系数检验、F 检验和 t 检验结果分别如下表所示。

表 5 R² 检验

Table 5 R² Test

模型	R	R ²	调整后 R ²	标准偏斜度错误
	0.999 ^a	0.999	0.999	0.88057

注: 预测值: (常数) 温度、湿度。

表 6 F 检验

Table 6 F Test

模型	平方和	df	平均值平方	F	显著性
回归	526291.437	2	263145.718	339362.356	0.000 ^b
残差	640.490	826	0.775		
总计	526931.927	828			

注: 应变数: 空调能耗; 预测值: (常数) 温度、湿度。

表 7 t 检验

Table 7 t Test

模型	非标准化系数		标准化系数		T	显著性
	B	标准错误	Beta			
(常数)	-537.627	1.692			-317.694	0.000
湿度	-0.149	0.028	-0.007		-5.421	0.000
温度	25.894	0.035	1.003		736.861	0.000

注: 应变数: 空调能耗。

以上结果显示灰色多元线性回归方程拟合结果良好,得到灰色多元线性回归方程为:

$$Y = 25.894 \times t - 0.149 \times h - 537.627 \quad (17)$$

式中, Y 为空调能耗预测值; t 为温度; h 为湿度。

本文以上海某办公建筑 6、7、8 月空调能耗观测值建立多元线性回归空调能耗预测模型,求出与 9 月份实际观察值之间的差值以及差值百分比,从

而验证灰色理论与多元线性回归模型结合方法的合理性以及模型预测精度。下表是截取的某一时间段的对比值。

表 8 模型检验

Table 8 Model checking

时间	空调能耗	空调能耗	差额 (kWh)	误差
	实测值 (kWh)	预测值 (kWh)		
2018/9/3 7:00	169.89	167.48	-2.41	-1.42%
2018/9/3 8:00	377.83	226.98	-150.85	-39.93%
2018/9/3 9:00	272.23	258.79	-13.44	-4.94%
2018/9/3 10:00	285.48	301.42	15.94	5.58%
2018/9/3 11:00	294.93	312.07	17.14	5.81%
2018/9/3 12:00	330.95	325.91	-5.04	-1.52%
2018/9/3 13:00	326.19	312.96	-13.23	-4.05%
2018/9/3 14:00	320.08	310.38	-9.7	-3.03%
2018/9/3 15:00	298.04	315.41	17.37	5.83%
2018/9/3 16:00	309.86	326.06	16.2	5.23%
2018/9/4 7:00	181.65	169.71	-11.94	-6.57%
2018/9/4 8:00	324.69	177.63	-147.06	-45.29%
2018/9/4 9:00	204.79	194.36	-10.43	-5.09%
2018/9/4 10:00	236.45	224.03	-12.42	-5.25%
2018/9/4 11:00	285.34	262.67	-22.67	-7.95%
2018/9/4 12:00	299.82	274.22	-25.6	-8.54%
2018/9/4 13:00	301.45	279.40	-22.05	-7.31%
2018/9/4 14:00	309.42	282.88	-26.54	-8.58%
2018/9/4 15:00	311.24	285.02	-26.22	-8.42%
2018/9/4 16:00	309.56	282.43	-27.13	-8.76%

需要注意的是, 根据办公楼运营时间, 每天早上 8 点-9 点是员工上班时间, 此时会陆续开启空调, 开启空调时电量消耗比普通运行模式情况下大, 所以造成了预测能耗与实际能耗之间的差值较大。结果显示, 除去空调开启时间段能耗差异值较大, 其余时间预测结果显示有较高的精度。将 GM (1,1) 模型修正后的数据与多元线性回归法相结合, 建立的灰色多元线性回归模型拟合精度较高, 符合建模要求。

3 结论

本文是根据逐时的温度和湿度数据建立灰色多元线性回归模型, 在实际使用空调过程中, 由于空调系统不正当使用等因素导致了空调系统能耗的浪费, 管理者可以根据天气气象参数得出空调能耗预测值, 然后根据预测值设定本建筑空调系统的能耗阈值, 建立节能规章制度, 减少不必要的能耗浪费。

(1) 根据相关性分析结果, 温度与空调能耗之间的相关性系数明显高于湿度与空调能耗之间的相关性系数, 说明温度对空调系统能耗的影响程度要高于湿度。

(2) 经过灰色理论修正过的温度、湿度与空调能耗之间的相关性明显增强, 灰色理论降低数据的随机波动性的效果是显著的。并且根据灰色模型优化前后的数值拟合方程效果对比, 灰色模型优化后的数据在建立多元线性回归模型时拟合精度要高于原始数据的, 说明降低外界因素对于数据造成的波动可以提高模型的预测精度。

(3) 灰色多元线性回归空调能耗预测模型有较高的预测精度, 即灰色理论与多元线性回归模型相结合的方法是可行的。

参考文献:

- [1] 2018 年上海市国家机关办公建筑和大型公共建筑能耗监测及分析报告[R].上海:上海市住房和城乡建设管理委员会,2018.
- [2] 江亿,薛志峰.北京市建筑用能现状与节能途径分析[J].暖通空调,2004,(10):17-20.
- [3] 王定奥,刘清倦,戴伟,等.基于 BP 神经网络的空调能耗预测与监控系统[J].现代电子技术,2019,42(22):140-144.
- [4] 陈彦熹,刘建华,李旭东,等.基于 ANN 的绿色办公建筑 HVAC 系统运行能耗预测[J].建筑节能,2017,45(10):1-5.
- [5] 廖文强,王江宇,陈焕新,等.基于长短期记忆神经网络的暖通空调系统能耗预测[J].制冷技术,2019,39(1):45-50,54.
- [6] X Yang, et al. A forecasting method of air conditioning energy consumption based on extreme learning machine algorithm[C]. 2017 6th Data Driven Control and Learning Systems (DDCLS), Chongqing, 2017:89-93.

(下转第 450 页)