

文章编号:1005-4642(2022)10-0046-05

空气热机平衡点的快速判定方法

全泓达, 盖 磊

(中国海洋大学 物理与光电工程学院, 山东 青岛 266100)

摘 要:对空气热机实验数据中获取的转速、温差等进行 K 型均值聚类 and 最近邻节点分类分析, 获得空气热机是否达到平衡状态的快速判别方式, 以及该判别方式在实验数据中表现出的区分能力和准确度.

关键词:空气热机; 静态测量法; 平衡点; K 型聚类分析; KNN 分类

中图分类号: O414.1

文献标识码: A

DOI:10.19655/j.cnki.1005-4642.2022.10.008

热机是将热能转换为机械能的仪器. 法国青年工程师卡诺从理论上对理想热机的工作原理进行研究, 提出了一种理想的由 2 个等温及 2 个绝热过程的循环, 称为卡诺循环, 并在 1824 年提出了卡诺定理, 为后来热力学第二定律的建立奠定了基础, 为提高热机的效率指明了方向, 为热力学的发展做出了杰出的贡献^[1-3].

由于热机的整个装置处于动态平衡, 获取精准的静态状态数据需要较长的判断时间, 因此本文主要基于空气热机实验中验证卡诺定理的静态测量法, 探索了快速判断平衡点的方法.

1 实验设备及原理

1.1 空气热机实验仪

实验采用世纪中科的 ZKY-RJ 空气热机实验仪, 实验由空气热机电加热器进行加热, 并由空气热机测试仪得到冷端温度、热端温度、转速、气缸压力以及气缸体积等实验数据, 并经过实验通信器面板传回电脑进行数据显示.

1.2 静态测量法验证卡诺定理

热机每次循环从热源吸收的热量

$$Q_1 \propto \frac{\Delta T}{n}, \quad (1)$$

其中, n 为热机转速, ΔT 为冷热端温差. 因此热机

效率 η 可表示为

$$\eta = \frac{A}{Q_1} \propto \frac{A}{\frac{\Delta T}{n}} = \frac{nA}{\Delta T}, \quad (2)$$

其中, A 为气体对外所做的净功, 即 $P-V$ 图中曲线所围面积. 由于式中的变量均可以在实验中测量得到, 因此通过观测不同冷热端温度时的 $nA/\Delta T$ 与 $\Delta T/n$ 的关系即可验证卡诺定理.

测量式子(2)中的参量有 2 种方法: 静态测量法和动态测量法. 静态测量法是在电压档位一定时, 在系统稳定后测量得到的平均值, 可以表示在一定的冷热端温度时的关系. 热机稳定工作是指冷热端在一定的温度下, 保持一定的效率, 即在每次循环中热机对外做功近乎不变. 在一定输入功率下, 系统是否达到平衡取决于冷热端的温差和转速, 即通过电加热器可以使冷热端温差变大(视冷端温度不变或者变化很小), 而热机每次循环都将从热源吸收热量, 使冷热端温差下降. 所以二者的平衡即是静态测量法所需要的稳定状态.

实验中判断热机是否平衡需要长时间的数据记录, 当转速、冷热端温差在一定范围内上下浮动时(例如 $n < 1 \text{ r} \cdot \text{s}^{-1}$, $\Delta T < 5 \text{ K}$), 可判断空气热机达到平衡状态, 但判定平衡点的观测时间通常需要 10 min 甚至更长, 从而不利于实验教学.

收稿日期: 2022-03-13; 修改日期: 2022-06-08

基金项目: 教育部高等学校大学物理课程教指委(华东地区)2021 年度高等学校教学研究项目(No. 2021JZWH10); 中国海洋大学 2021 年度本科教育教学研究项目(No. 2021JY008)

作者简介: 全泓达(2000-), 男, 浙江温州人, 中国海洋大学物理与光电工程学院光电信息科学与工程专业 2019 级本科生. E-mail: quanhongda@stu.ouc.edu.cn

通讯作者: 盖 磊(1981-), 男, 山东烟台人, 中国海洋大学物理与光电工程学院高级实验师, 硕士, 从事大学物理实验教学与科研工作. E-mail: phylab@ouc.edu.cn

2 实验设计及步骤

- 1) 打开计算机、测试仪和通信器,加热电源;
- 2) 将电压调至 5 档,预热仪器使飞轮顺利转动(一般 $\Delta T > 80 \text{ K}$);
- 3) 将电压调回 1 档,开始计时计数,每隔 1 min 记录 1 次 $T_1, \Delta T, n$ 和 A ,直至各项数值在平衡点附近变化(记录时间约 1 h);
- 4) 将电压调至 2 档重复上述步骤 3)。

3 数据处理

3.1 数据处理方法的原理

3.1.1 K 型均值聚类

K 型均值聚类(K-means clustering)是经典的无监督机器学习算法,用于无目标值的数据集^[4]。K 型算法主要是通过迭代来对数据集进行分类。首先在样本中选取任意 K 个样本作为聚类中心,利用所选取的聚类中心计算各样本与中心间的距离;然后根据获得的最小计算距离重新划分对象。距离一般通过欧式距离 D_δ 进行度量,表达式如下:

$$D_\delta(x_i, x_j) = \sqrt{(x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2 + \dots + (x_{ip} - x_{jp})^2}, \quad (3)$$

其中, $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})$ 和 $x_j = (x_{j1}, x_{j2}, \dots, x_{jp})$ 表示 p 维欧几里德空间中的 2 点。重新计算获得每个样本的均值后,迭代至目标函数值不变或者小于指定阈值。目标函数为平方误差准则函数,表示为

$$\sigma_i = \sqrt{\frac{\sum_{i=1}^{N_i} (x_i - c_i)^2}{|C_i| - 1}}, \quad (4)$$

其中, C_i 为数据对象数量, c_i 表示第 i 个簇中心。

3.1.2 KNN 分类

最近邻节点分类(K-nearest neighbor, KNN)方法是有监督的弱机器学习算法^[5],其通过计算预测值与现有样本的闵可夫斯基距离(Minkowski distance)来确定预测所属类别。

闵式距离的表达式为:

$$D_M(x, y) = \sqrt[p]{\sum_{i=1}^m (|x_i - y_i|)^p}. \quad (5)$$

当 $p \rightarrow \infty$ 时,闵式距离转化为切比雪夫距离(Chebyshev distance);当 $p=2$ 时,闵式距离转化

为欧几里得距离;当 $p=1$ 时,闵式距离转化为曼哈顿距离(Manhattan distance)。通过闵式距离,可较好表示 n 维欧几里德空间中 2 点相对距离。

KNN 算法中最为重要的参量为判断类别时采用的样本个数,若选取的样本数量过小,容易受到异常点的影响;若选取的样本数量过大,就会受到样本均衡的影响,所以选取合适的样本数对 KNN 模型具有重要作用。

本文对实验数据进行 KNN 分类处理,以获得能较好判断热机是否到达平衡的模型,通过该模型能准确快速地判断空气热机是否达到平衡,从而减少实验时间。

3.2 实验数据处理

为使实验结果具有普适性,消除仪器和档位(加热电压)的偶然性,本文对 2 台仪器 3 个档位进行实验,部分原始数据如表 1 所示。

表 1 实验原始数据(部分)

T_1/K	$\Delta T/\text{K}$	$n/(\text{r} \cdot \text{s}^{-1})$	W_{out}/J	t/min
457.7	137.7	7.9	0.173 9	1
457.6	137.7	8.0	0.174 0	2
459.1	138.9	8.1	0.174 5	3
461.1	140.5	8.2	0.174 5	4
461.5	140.7	8.2	0.175 1	5
462.4	141.5	8.2	0.176 3	6
464.2	143.2	8.3	0.175 9	7
465.2	144.1	8.3	0.176 0	8
467.0	145.6	8.4	0.176 1	9
469.2	147.6	8.5	0.176 8	10

由于各项指标的数据单位以及数值大小不同,所以需要对各类数据进行标准化。假设进行分析的指标变量有 m 个,分别为 z_1, z_2, \dots, z_m ,共有 n 个评价对象,第 i 个评价对象的第 j 个指标的取值为 a_{ij} ^[6]。将各指标值转化为标准化值有

$$\tilde{a}_{ij} = \frac{a_{ij} - \mu_j}{s_j}, \quad (6)$$

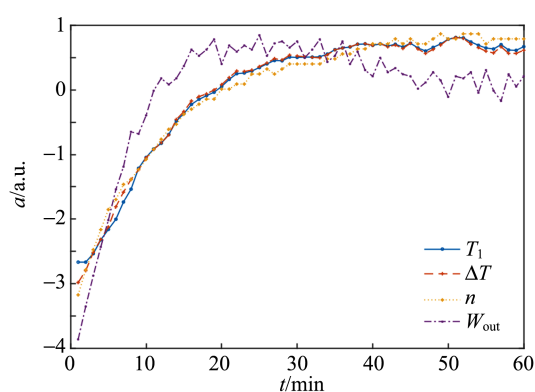
其中 $i, j = 1, 2, \dots, m$ 。 $\mu_j = \frac{1}{m} \sum_{i=1}^m a_{ij}$; $s_j =$

$$\sqrt{\frac{1}{m-1} \sum_{i=1}^m (a_{ij} - \mu_j)^2}.$$

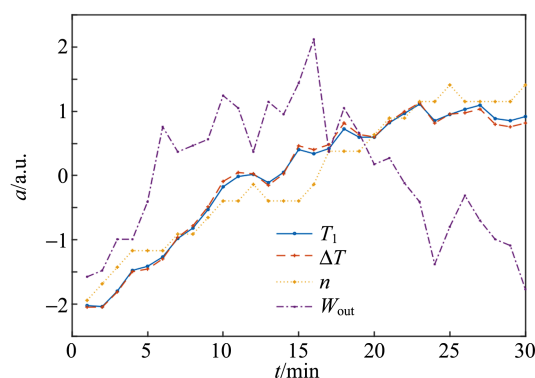
由式(6)可获得各项指标标准化后随时间的变化图线,如图 1 所示。根据卡诺定理,热机在稳定运行时,其在 1 个循环内的效率仅跟冷热端温

度有关,而在多次循环中,转速与循环次数有关,那么是否平衡也与转速有关.以下将不考虑对外做功对平衡判断的影响.需要补充说明的是,为了获取仅与上述因素有关的热机平衡判定方法,而非对比仪器间差异,所以无需控制仪器 A 和 B 都在电压档 1 和 2 进行实验测量.

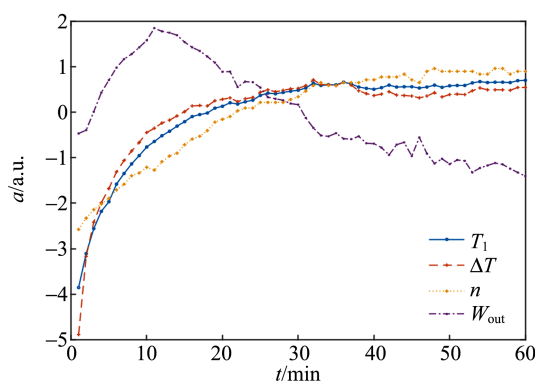
利用得到的标准化数据,进行 K 型聚类^[7],将欧几里德距离作为 K 型聚类的判断标准,将数据分为平衡状态与非平衡状态,由此可以获得较为客观公正的平衡点,图 2 为聚类后的轮廓图.



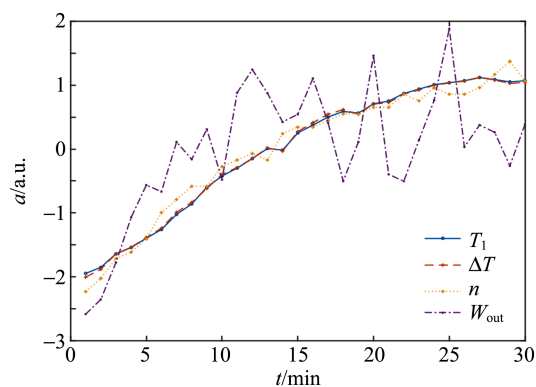
(a) 仪器 A 电压档 1



(b) 仪器 A 电压档 2

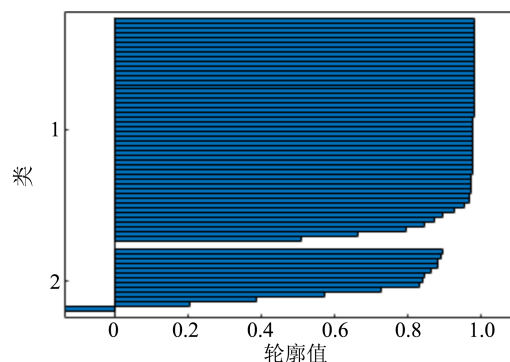


(c) 仪器 B 电压档 1

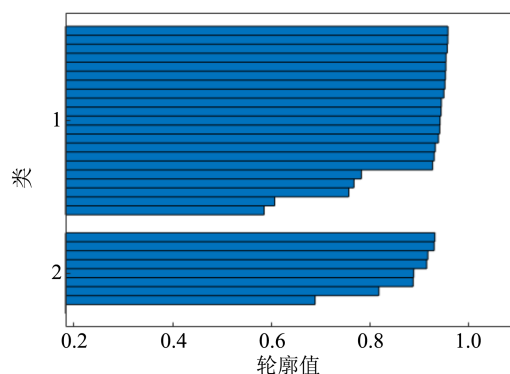


(d) 仪器 B 电压档 3

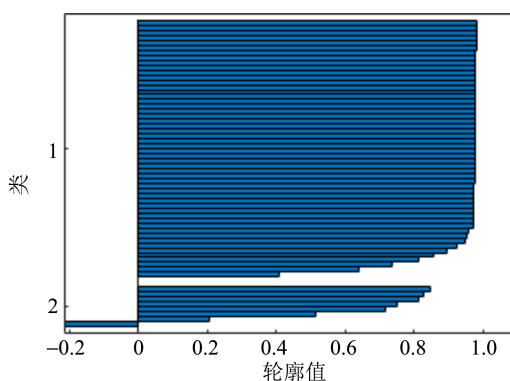
图 1 实验测量量随时间变化曲线



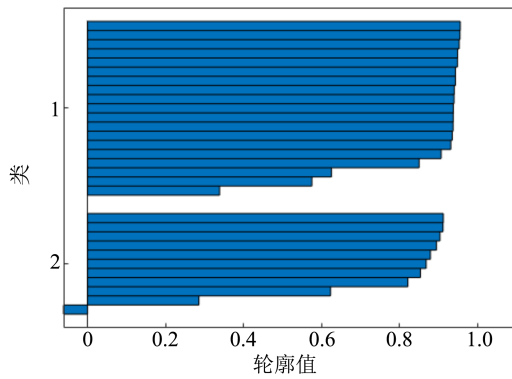
(a) 仪器 A 电压档 1



(b) 仪器 A 电压档 2



(c) 仪器 B 电压档 1



(d)仪器 B 电压档 3

图 2 K-means 聚类后轮廓值图

图 2 中簇 1 为聚类后认定的平衡状态,簇 2 为聚类后认定的非平衡状态. 通过对轮廓图的分析,可以发现聚类效果较为显著. 为了让实验数据更具有信服力,本次实验记录了长时间(远超热机平衡所需时间)的实验数据,所以有大量的平衡状态数据点,而该类型的点能很好地被聚集在一起,形成较为明显的的数据边际,说明平衡与非平衡之间存在一定的衡量指标.

然后利用 KNN 进行分类,将所有数据进行整合,并将加热电压值作为模型建立的特征值. 图 3 和图 4 为模型的训练结果.

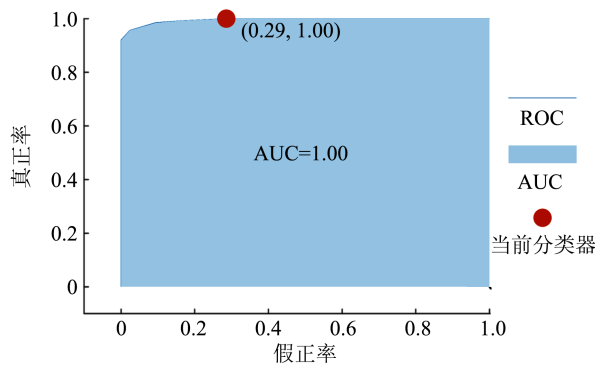


图 3 模型 ROC 曲线与 AUC

整体利用 5 折交叉验证获得的模型准确率达到 93.3%,且为了防止过拟合,预留部分数据进行验证,最后效果较好,基本满足预期需求.

图 3 为模型的 ROC 曲线(Receiver operating characteristic),其中 AUC^[8](Area under curve)接近 1,远在超假正率(False positive rate, FPR)等于真正率(True Positive Rate, TPR)的直线上方,接近 1 个完美的分类器.

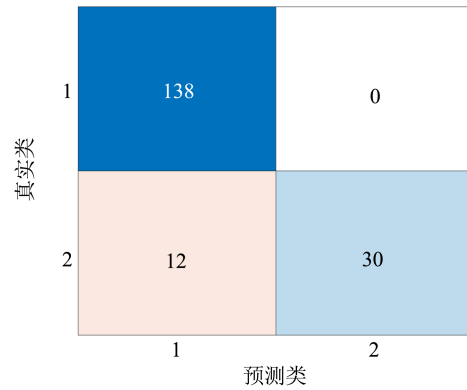


图 4 模型混淆矩阵

图 4 为模型的混淆矩阵^[9],从矩阵可以看出误判断数量较少,整体能比较好地对平衡与非平衡状态进行分类.

基于以上模型,再次进行 5 组实验,在不同仪器不同电压档位下各测量 60 个数据,并将得到的数据放入模型预测系统是否处于平衡状态. 为了避免仪器因转速过高而触发保护机制,将最高电压档位设定为 4 档.

图 5 是模型对于系统是否达到平衡状态判断,横坐标表示仪器和电压档位,上下阈值为 K-means 聚类算法得出的平衡点的最大值和最小值,模型所判断出的平衡点落在该范围越多说明模型判定准确性越好. 可以看出,模型判断的平衡点基本都集中在平衡范围内.

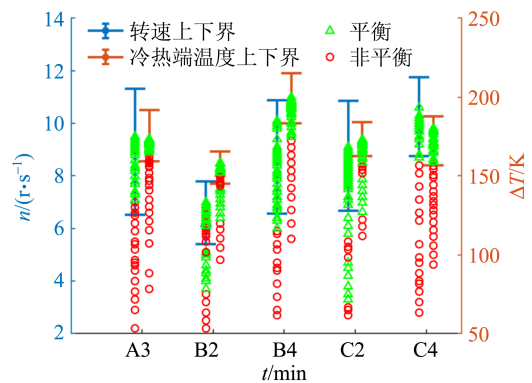


图 5 不同仪器、不同电压档位下系统平衡状态的判断结果

图 6 为 C 仪器在电压档为 4 时的转速随时间的判断结果,可看出系统在平衡状态与非平衡状态间的转换位置处于曲线拐点,与真实进入平衡点时间不超过 2 min,说明该模型判断效果较好.

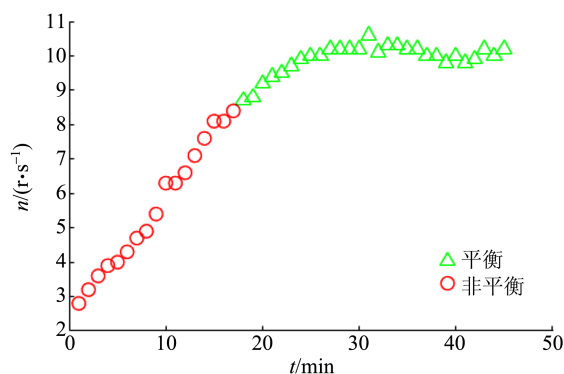


图 6 C 仪器在电压档为 4 时转速随时间的判断结果

4 结束语

本文通过对空气热机实验测得的数据进行分析,获得了基于 KNN 的模型分类器,通过多种指标对模型进行衡量分类都有较好的预测结果.其中,误分类点处于平衡状态与非平衡状态的模糊界线处,从错误预测进入正确预测所需时间不大于 2 min,因此误分类相较于接近 10 min 的主观判断有了大幅度的提升,并具备一定的实用性.将每一时刻获得的数据输入该模型,能得到较好的预测结果.另外,通过 K-means 聚类算法可以对因测量精度、环境影响、仪器误差、动态平衡等因素带来的数据波动提供较好的参考中心值,从而使得真实值的选取更加准确.

A fast method to determine the equilibrium point of air engine

QUAN Hong-da, GAI Lei

(College of Physics and Optoelectronic Engineering, Ocean University of China,
Qingdao 266100, China)

Abstract: In order to obtain a fast discriminatory method for whether the air engine reached equilibrium, the K-means clustering and K-nearest neighbor (KNN) analysis were performed on several factors gained from the experimental data of air engine, such as the rotation rate and temperature difference. The discriminatory method showed discrimination ability and accuracy in experimental data.

Key words: air engine; static measurement; equilibrium point; K-means cluster analysis; KNN classification

参考文献:

- [1] 马洪良,王春涛,陆江. 基于热效率实验仪的热机模式最佳负载测量[J]. 物理实验,2004,24(8):26-27,30.
- [2] 由存,戎张泽,刘真,等. 基于空气热机的热机效率分析及在教学中的应用[J]. 内燃机与配件,2018(7):243-245.
- [3] 刘铭基,滕华强. 空气热机原理的应用设想[J]. 仪表技术,2016(8):26-28,32.
- [4] 王千,王成,冯振元,等. K-means 聚类算法研究综述[J]. 电子设计工程,2012,20(7):21-24.
- [5] Peterson L E. K-nearest neighbor [J]. Scholarpedia, 2009,4(2):1883.
- [6] 司守奎,孙兆亮. 数学建模算法与应用[M]. 北京:国防工业出版社,2015:231.
- [7] Thinsungnoena T, Kaoungkub N, Durongdumronchaib P, et al. The clustering validity with silhouette and sum of squared errors [J]. Learning, 2015,3(7):44-51.
- [8] Huang J, Ling C X. Using AUC and accuracy in evaluating learning algorithms [J]. IEEE Transactions on Knowledge and Data Engineering, 2005,17(3):299-310.
- [9] 赵存秀. 基于混淆矩阵的分类器性能评价指标比较[J]. 电子技术与软件工程,2020(13):146-147.

[责任编辑:郭 伟]