

基于数据统计特性的 GS-SVM 电池峰值功率预测模型

郑方丹¹, 姜久春¹, 陈坤龙¹, 韩智强², 姜婷婷³, 孙丙香¹

(1. 北京交通大学 国家能源主动配电网技术研发中心 北京电动车辆协同创新中心, 北京 100044;

2. 北京新能源汽车股份有限公司, 北京 102606; 3. 国家电网山东省电力公司电力科学研究院, 山东 济南 250002)

摘要: 以锰酸锂动力电池为研究对象, 对电池处于不同温度和荷电状态下的情况进行 10 s 峰值功率测试, 同时测量电池内阻。对实验测试得到的温度、荷电状态、内阻及峰值功率数据进行统计分析, 包括测试变量间的相关程度评估和共线性检测, 挖掘电池外特性参数与峰值功率数据间的统计关系。在此基础上, 提出采用基于网格搜索的支持向量机(GS-SVM)建立电池的峰值功率预测模型。验证结果表明所提模型预测精度高, 平均误差仅为 3.65%; 该模型训练时间短、响应速度快、操作性强, 可以实现对动力电池峰值功率的快速估计, 为电动汽车安全可靠运行提供有力保障。

关键词: 动力电池; 峰值功率; 相关分析; 共线性检测; GS-SVM; 电动汽车

中图分类号: TM 911; U 469.72

文献标识码: A

DOI: 10.16081/j.issn.1006-6047.2017.09.008

0 引言

全球经济的快速发展, 在提高人民生活水平的同时也增加了社会对能源的需求, 进而造成环境问题日益严重。新能源汽车的应用和推广旨在缓解能源短缺的压力, 改善传统燃油汽车存在的废气污染问题, 受到了社会各界的广泛关注。以锂离子电池为动力源的电动汽车以其污染少、噪声小、能源需求低等优点, 逐渐代替传统汽车成为未来交通车辆发展的主要方向^[1]。混合动力汽车 HEV (Hybrid Electric Vehicle) 采用内燃机和锂离子电池共同作为动力源, 通过整车管理系统实现多能量源的最优分配, 提高车辆的燃油效益和排放性能, 以达到省油和环保的目的。通常, 内燃机工作在 HEV 的匀速行驶状态, 而锂离子电池工作于车辆起步、加速或刹车状态, 为车辆提供或吸收瞬时的大功率。这就要求 HEV 车用电池必须是功率型电池, 能够承受大倍率的充放电电流。电池输出输入的峰值功率直接影响车辆的快速启动、加速和紧急制动能力, 进而影响整车运行的安全性和可靠性^[2]。因此, 基于电池管理系统 BMS (Battery Management System) 实现对锂离子电池峰值功率的在线估计是十分必要的。准确的电池峰值功率估计结果可以为整车系统的功率分配以及能量控制提供支持和参考, 不仅能够保证车辆的运行性能, 也能够保证车辆的燃油经济性。

目前, 对于锂离子电池峰值功率估计的相关研究主要集中于峰值功率的测试方法和峰值功率的预测算法。在实际应用中, 首先基于离线数据测试得到电池峰值功率数据表, 然后根据插值的结果进行在线预测^[3]。当离线测量电池峰值功率时, 常用测

试方法主要包括美国 FreedomCAR 项目提出的电池混合脉冲功率特性 HPPC (Hybrid Pulse Power Characteristic) 测试法和恒功率测试法 2 种^[4]。HPPC 测试法通过给电池加入 t_s 脉冲电流, 分析电压变化, 基于给定公式计算电池峰值功率。其中, 脉冲电流值的确定是需要通过多次测试数据拟合推断而得到。恒功率测试法则通过在某一恒定功率下对电池持续充电或放电 t_s , 以恰好达到电池工作电压的上限或下限。该恒定功率值需要通过调整功率设定值使测试时间逐渐逼近 t_s 而获得。据统计数据表明, 紧急加速持续时间在城市工况和高速工况下分别为 10 s 和 25 s, 紧急制动持续时间在上述 2 种路况下分别为 5~10 s 和 33 s^[5]。考虑到实际应用情况, 本文选取锂离子电池的 10 s 放电峰值功率典型值作为研究对象, 进行建模和在线估计。锂离子电池的峰值功率预测算法主要包括参数模型和非参数模型 2 种。前者通常由电池的等效电路推算输出功率的解析表达式, 再将测得的电流电压值等代入表达式进行计算; 后者将电池看作一个黑箱, 峰值功率作为输出, 其影响因素作为输入, 通过对大量测试数据进行训练, 实现对峰值功率的估计。非参数模型无需深入研究电池的内部机理, 而是借助于数据分析和机器学习算法, 因此又称为数据驱动模型。

在电动汽车的实际行驶过程中, 锂离子电池的峰值功率与电池的温度、荷电状态 SOC (State Of Charge)、端电压以及电流变化率等参数之间的关系呈现明显的非线性特征, 从而导致锂离子电池的峰值功率在线估计是十分复杂的。因此, 本文从电池的峰值功率测试数据角度出发, 利用统计学参量分析相关影响因素, 寻找数据间的统计规律和功率性能分布特征。在此基础上, 利用基于网格搜索的

支持向量机 GS-SVM(Grid Search-Support Vector Machine)对电池峰值功率进行建模,并利用实际测试数据对模型的准确性进行验证,最终实现电池峰值功率的在线估计。

1 峰值功率实验

锂离子电池在使用过程中的电压限制条件定义了该电池的 t 恒定充电/放电功率值。即电池在某一恒定功率 P 下持续充电/放电直至其限制工作电压为 U_{\max}/U_{\min} ,整个过程持续时间为 t ,那么称该恒定功率值 P 为电池的 t 充电/放电峰值功率^[6]。本文选取的是锂离子电池的 10 s 放电峰值功率,故取 $t=10$ s。测试采用恒功率法,通过不断更改恒功率放电的功率值 P 得到功率 $P-t$ 对应的关系曲线图,再由插值法求得电池的 10 s 峰值放电功率($P_{\max}, 10$ s)。

在实际应用中,电池的 10 s 峰值功率不仅受到电池 SOC、温度、电池内阻的影响,同时也受到充放电电流强度、充放电脉冲时间等其他因素的影响。全面分析考虑峰值功率的影响因素,并将其全部作为预测模型的输入,必将使得模型的数据量非常庞大、模型计算量大、运算速度慢,进而导致复杂度高、收敛速度慢等问题,并不利于峰值功率的在线估计。

因此,选取合适的峰值功率影响因素,在保证估计精度的前提下减少计算量是十分必要的。同时,在线获取模型输入量的影响因素的能力也是必要的,这样才能保证该算法的实际可行性。BMS 不仅可以在线测得电池温度,还可以由欧姆定律算得电池内阻值(本文中均指电池的欧姆内阻),此外电池的 SOC 值也是可以在线估算的。同一批次的电池在实际工况下面临的运行条件并不一致,如由于所处车内位置不同导致运行温度不同,由于串并联结构不同导致充放电倍率不同等。尽管经过相同的循环周期或运行里程,其内阻的变化值并不一致。但大量的实验数据说明,随着电池循环使用次数增加,电池的可用容量逐渐减少,内阻呈缓慢递增的趋势。电池的内阻在一定程度上可以反映电池所处的健康状态 SOH(State Of Health)。因此,本文选取电池温度、电池 SOC 以及电池内阻为影响峰值功率的主要因素,并进行相关性分析。

具体实验对象为 120 个锰酸锂电池,实验设备为 Digatron 电池测试系统(主功率模块)和高低温试验箱。实验时,首先随机将 120 个电池置于不同的温度点和 SOC 点,再根据直流放电内阻测量法^[7]测得电池内阻,最后通过恒功率测试以及数据拟合的方法得到电池的 10 s 峰值放电功率。其中,温度设置为 $-20\sim 50$ °C,每隔 10 °C 设一个温度点;SOC 设置为 0.1~0.9,每隔 0.1 设一个 SOC 点。

2 实验数据的统计分析

2.1 实验结果

根据实验结果得到 120 组测试数据,每组数据均包含温度、SOC、内阻以及峰值功率 4 个变量。为了形象化地表示这 120 组数据,将 4 个变量的关系用图 1 表示。图中,横坐标表示温度;纵坐标表示峰值功率;颜色表示内阻(Ω),颜色越深表示内阻越大;气泡表示 SOC,气泡越大表示 SOC 越大。可以看出,随着温度的升高或内阻的减小或 SOC 的增加,峰值功率随之增大;当温度等于 50 °C 且 SOC 等于 0.9 时,峰值功率达到最大值;温度是影响电池峰值功率的重要因素,当温度为某一定值时,峰值功率与 SOC 呈明显线性关系,峰值功率随着气泡的变大而逐渐增大(图 1 中垂直方向的气泡排列)。

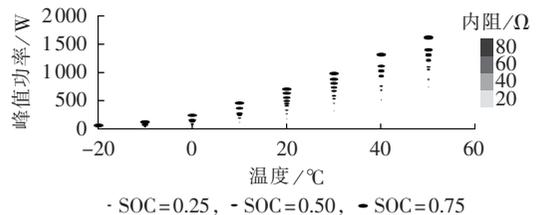


图 1 电池温度、SOC 和内阻与峰值功率的关系
Fig.1 Relationship between battery temperature, SOC, internal resistance and 10 s peak power

由图 1 可知,SOC、内阻以及电池峰值功率三者之间存在一定的规律性。当温度确定时,电池的 SOC 和内阻同时影响了电池峰值功率的大小,主要表现为峰值功率随着内阻的增大而减小。但不同温度下这种规律关系并不是完全一致的,电池内阻的变化并不是独立的,温度影响了三者之间规律变化的定量关系。实际上,这些变量之间是相互耦合的,并不存在绝对的线性关系。因此,为了进一步分析实验结果的统计特性,需要对变量间的相关程度进行分析。

2.2 变量的相关分析

利用统计学方法可以研究一个随机变量与多个随机变量之间的相关关系,本文中自变量组的变量为电池温度、电池 SOC 和电池内阻,因变量组的变量为电池的 10 s 峰值功率。为研究电池的 10 s 峰值功率与电池温度、SOC 以及内阻之间的关系,采用统计学理论中的典型相关分析法对实验数据进行相关性评估。图 2 为电池温度、SOC 和内阻与电池的 10 s 峰值功率的散点图。

虽然散点图能够直观地展现变量之间的统计关系,但严格意义上,这样的结果并不精确。变量间的密切程度可以通过相关系数来体现。相关系数的计算方法有多种,比较常用的有 Pearson 简单相关系数、Spearman 等级相关系数以及 Kendall 相关系数等^[8]。Pearson 相关系数用来衡量 2 组数据集之间的线性

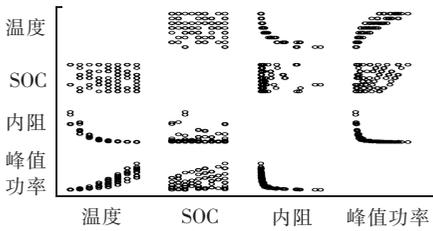


图 2 电池温度、SOC、内阻和峰值功率的散点图矩阵
Fig.2 Scatterplot matrix of battery temperature, SOC, internal resistance and 10 s peak power

相关程度,相关系数的绝对值越接近 1,说明相关性越强;相关系数的绝对值越接近于 0,则说明相关性越弱。Spearman 等级相关系数用来计算数据之间的等级关系,主要针对有次序差别的数据,并不常用于分析连续数据。Kendall 相关系数又称和谐系数,通常用以表示多列等级变量的相关程度,适用于多属性数据集。在 HEV 的实际运行过程中,电池工作的环境条件是随机的,包括电池所处的温度、电池的 SOC 等。同时,每个电池的老化程度也是不同的,电池的内阻值很大程度上就反映了电池衰退的程度即电池的健康状态。通过统计分析可知,电池实际工作温度、SOC、电池内阻均呈正态分布^[9],并且由图 1 可知,在电池的低温、常温和高温情况下均存在随着电池 SOC 增加,电池的内阻减少,同时峰值功率短时增加的情况,因此本文采用 Pearson 相关系数计算得到四者之间的线性相关程度,如表 1 和表 2 所示。

表 1 相关系数计算结果

Table 1 Calculative results of correlation coefficients

影响因素	相关系数			
	电池温度	电池 SOC	电池内阻	峰值功率
电池温度	1.000	-0.079	-0.784	0.878
电池 SOC	-0.079	1.000	-0.097	0.310
电池内阻	-0.784	-0.097	1.000	-0.606
峰值功率	0.878	0.310	-0.606	1.000

表 2 显著性水平计算结果

Table 2 Calculative results of significance level

影响因素	显著性水平			
	电池温度	电池 SOC	电池内阻	峰值功率
电池温度	—	0.485	0	0
电池 SOC	0.485	—	0.390	0.005
电池内阻	0	0.390	—	0
峰值功率	0	0.005	0	—

根据表 1 和表 2,可以分析出电池温度、SOC、内阻以及 10 s 峰值功率两两之间的相关程度以及显著性水平。计算出 Pearson 相关系数后,用假设检验的方法对该结果的显著性水平进行检验。Pearson 相关分析中的原假设为变量之间不存在线性关系。显著性水平指的是第一类错误(原假设为真但被拒绝)发生的概率,即变量之间实际不存在线性相关性却被误判成线性相关的概率。所以显著性水平的值越小,说明原假设不成立的概率越大,即变量间存在线

性相关的概率越大。通常,当显著性水平小于 0.05 时,则认定拒绝原假设。换言之,显著性水平小于 0.05 说明变量之间相关性显著。如电池温度和电池内阻之间的相关系数是 -0.784,可知两者相关程度很大,又由于统计检验的显著性水平为 $0 < 0.05$,说明相关性非常显著,且为负相关。又如电池内阻和峰值功率之间的相关系数是 -0.606,统计检验的显著性水平为 $0 < 0.05$,说明二者相关性较大,相关性显著,也为负相关。同理可以分析出其他变量之间的相关关系。根据实验设计可知,实验过程中电池的温度和 SOC 是人工设定的,并不具有相关性。而相关性检测结果也说明了这一点,温度和 SOC 的相关系数为 -0.079,显著性水平为 $0.485 \gg 0.05$,因此二者之间不存在相关关系,两变量相互独立。

2.3 输入变量的共线性检测

共线性指的是线性回归模型中输入变量之间存在高度相关关系,从而导致模型失真或估计结果不准确。在实际的工程应用中,解析的回归预测模型通常要求输入变量之间不存在很强的依赖关系,以确保可靠的预测结果。当输入变量之间完全没有线性关系时,则称之为正交的,此时采用线性回归模型就可以准确地实现模型预测。当输入变量之间的相关性较弱时,即使变量不正交也不会显著地影响分析结果。但是,当输入变量之间具有很强的线性相关性,即变量之间的共线性很强时,就会显著地影响回归的结果。这是因为在线性回归的模型假设中,每个变量的系数的实际意义是指当其他预测输入变量的值保持不变,而此输入变量的值变动一个单位时,输出变量值的变动量。当预测输入变量之间的相关性比较大时,不可能准确地估计某一个变量真实的效应,从而影响了预测的精度。这种估计系数的不稳定性表现为:当改变一个样本点时,估计的系数可能会有很大的改变。易得这样的模型是缺乏稳健性的。

在本实验中,所有电池的测试均是在给定 SOC 和温度的情况下进行的。由于不能人为设定内阻只能通过测试获取,故内阻与 SOC 或温度之间存在共线性的可能。又根据上述相关性分析结果,内阻极有可能与其他变量之间存在严重的共线性,因此,在建立回归预测模型前需要对输入变量的共线性进行检测。

主成分分析 PCA(Principal Component Analysis),又称为主分量分析,是统计学中常用的一种方法,可以用来观察数据的共线性^[10]。基于数据降维的思想,将高维数据在低维线性空间中做正交投影,使得投影数据的方差最大化,从而将多变量转化为少变量,降低空间维数,进而反映最主要的数据信息^[11]。当

原始数据存在多重共线性时,通过 PCA 方法可以将原始的高维数据投射到低维空间而几乎不损失原有数据中存在的信息^[12]。本研究中,原始数据处于一个三维(温度、SOC 和内阻)的特征空间,为观察原始数据的共线性程度,利用 PCA 方法对该特征空间进行转换,得到第一主成分、第二主成分、第三主成分的方差贡献率分别为 53.3%、41.5%、5.2%。故 95% 的信息可以用前 2 个主成分表示,第三主成分实际上并没有提供足够的信息。因此,可以认定原始数据存在比较严重的共线性,输入变量之间并不是相互独立而是相互耦合的。即输入变量之间存在较强的依赖性,且输入变量和输出变量(峰值功率)不是呈完全的线性关系,故不能直接建立解析的线性回归模型,否则存在很高的过拟合风险,无法保证准确可靠的峰值功率预测结果。

3 电池峰值功率预测的模型

3.1 算法及原理

由上述分析可知,无法直接采用线性回归方法建立电池峰值功率预测的模型,而利用数据驱动方法则可以很好地解决这一问题。数据驱动方法将输入输出的内部关系简化为一个黑匣子,利用机器学习算法直接从测试数据中挖掘隐含的电池状态关系或信息,从而预测电池的峰值功率。本文选取的支持向量机 SVM(Support Vector Machine)法是基于统计学理论发展而来的机器学习算法,具有较强的学习泛化能力,尤其适用于小样本数据的识别和估计^[13]。

SVM 模型的确定主要依赖于以下 5 个参数:核函数参数 g 、支持向量的数目、非零 Lagrange 乘子、目标函数偏置量 b 、惩罚系数 c ^[14]。其中,核函数参数 g 以及惩罚系数 c 是需要事先确定的,而其他参数则是在模型训练过程中由解二次规划问题得到^[15]。对于 SVM 的核函数而言,其参数 g 通常采用试算的方式得到。预先选取 g 为某一固定值,然后进行模型计算,根据计算结果以及经验来调整 g 的取值,直至得到预测精度最高的最为满意的计算结果,同时评估选取合适的惩罚系数 c ,从而得出两参数的最优值。惩罚系数 c 指对超出误差样本的惩罚程度, c 值越大则对样本的约束性越强,计算过程越复杂,计算时间越长。 c 值太小会导致模型的准确性和模型精度下降,使得计算结果脱离实际,不具说服力。

针对上述算法计算效率低、寻找全局最优值困难的缺点,考虑引入网格搜索 GS(Grid Search)的方法来获得 c 和 g 的全局最优值。相比其他 2 种常用的参数优化方法:遗传算法 GA(Genetic Algorithm)和混沌优化算法 COA(Chaos Optimization Algorithm),GS 算法更适合小样本预测模型。GS 算法的搜索范

围大,优势在于可以同时搜索 2 个参数值,得到的最优参数能够保证最高的分类准确率^[16]。此外,各组参数在并行计算过程中是相互解耦的,提高了运行效率。基于上述分析,在对电池数据的统计特性进行深入分析的前提下,本文采用 GS-SVM 方法建立电池峰值功率的预测模型。

3.2 模型的训练

对于实验获得的 120 组数据,随机抽取其中的 80 组作为训练样本。利用 GS 算法对核函数参数 g 和惩罚系数 c 进行参数寻优,设定 g 和 c 的搜索范围均为 $2^{-8} \sim 2^8$,步长为 0.5。选取 $c-g$ 二维网络上的每一个参数对 (c, g) 作为模型参数,得到均方误差 MSE (Mean Square Error)的等高线图,如图 3 所示。GS 算法得到的最佳模型参数 (c, g) 值为 $(16, 0.0625)$,此时 MSE 值最低,为 0.00381。

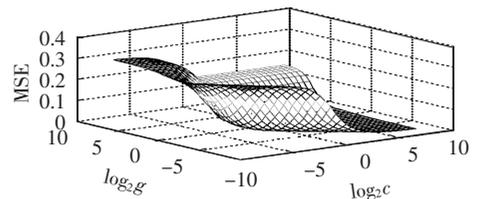


图 3 参数 g 和 c 的选择结果图

Fig.3 Grid selection result for parameters g and c

对 80 组电池样本进行训练,得到的 10 s 峰值功率实际值和估计值的对比如图 4 所示。可以看出,模型具有很高的精度,模型估计值与实际值之间相差很小,对应的 2 条曲线的吻合度非常好。通过数据分析可得,训练样本的估计值与实际值的平均误差为 41.69 W,是最大峰值功率的 2.61%;平均相对误差为 28.7%;最大误差为 88.35 W,占最大峰值功率的 5.53%。由于电池的峰值功率值一般都比较大,数量级在千瓦左右,因此个别的极小值(数量级在十瓦左右)会拉大平均误差,造成相对误差较高甚至大于 100% 的情况,如图 5 所示。但实际上,整个模型的精度较高,估计值与实际值之间基本有很强的吻合度,并且变化趋势极为相同。该 GS-SVM 模型具有很高的判定系数 COD(Coefficient Of Determination)。COD 值越接近 1,说明该回归模型拟合数据程度越好,而计算结果 $COD=0.9878$,也验证了这一点。

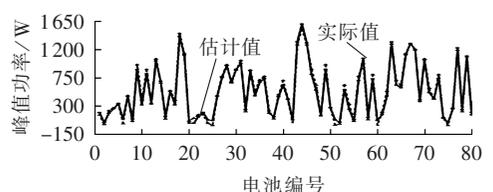


图 4 模型训练样本估计值与实际值对比

Fig.4 Comparison between measured value and estimated value for model training sample

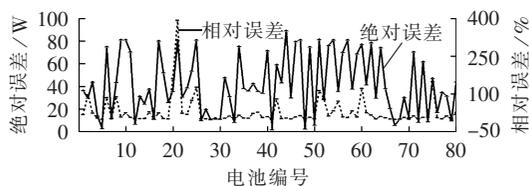


图 5 模型训练的绝对误差和相对误差
Fig.5 Absolute error and relative error of model training results

3.3 模型的验证

对于实验获得的 120 组数据,除去 80 组训练样本外,将剩余的 40 组作为预测样本。用本文建立的 GS-SVM 模型对样本数据进行预测,得到的峰值功率预测值与实际值的对比如图 6 所示。可见预测值与实际值的跟随程度相当好,变化趋势也保持一致,2 条曲线吻合度比较高。通过计算可得,预测值与实际值的平均误差与最大误差分别为 59.16 W 和 189.53 W,占最大峰值功率的 3.65%和 11.70%。根据统计分析结果可得,模型的估计峰值功率与实际峰值功率的相对误差在 20% 以内的电池数量为 20 个,占抽取电池总数的 50%;相对误差在 10% 以内的电池数量为 16 个,占抽取电池总数的 40%;相对误差在 5% 以内的电池数量为 8 个,占抽取电池总数的 20%。其中,相对误差大的情况均出现在低温低 SOC 区域,原因是此时电池的峰值功率实际值很小,即便预测的绝对误差很小,相对误差仍然会很大,如图 7 所示。从图中也可以看出,相对误差较大的值多分布于峰值功率低(低温低 SOC)区域,电池预测结果稍有偏差就会造成相对误差值的激增,即峰值功率值较小的部分其相对误差较大。当实际峰值功率高于 300 W 时,对应的相对误差急剧减小。

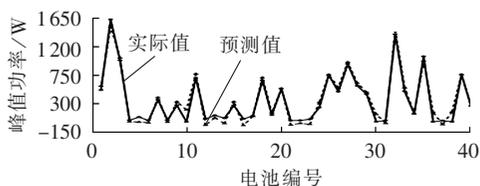


图 6 模型预测样本估计值与实际值对比
Fig.6 Comparison between measured value and estimated value for model forecasting sample

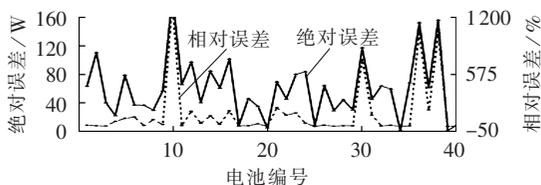


图 7 模型验证的绝对误差和相对误差
Fig.7 Absolute error and relative error of model validation results

基于 GS-SVM 模型比较训练样本和预测样本,计算训练过程和验证过程模型的 COD、平均误差和

最大误差,得到的结果如表 3 所示。模型的 COD 均在 0.97 以上,平均误差小于 60 W。

表 3 GS-SVM 模型在训练和验证过程中的误差比较
Table 3 Error comparison between training and validation for GS-SVM model

计算过程	COD	平均误差/W	最大误差/W	绝对误差标准差/W
模型的训练	0.9878	41.69	88.35	49.11
模型的验证	0.9747	59.16	189.53	72.61

以上分析说明本文建立的峰值功率预测的 GS-SVM 模型的精度和准确率均较高,其对于高峰值(高于 300 W)功率场合的适用能力远远高于低峰值功率场合,这为后续的峰值功率预测方法的适用性和对比研究提供了指导和参考。

4 结论

本文针对 HEV 的动力电池,分析影响电池峰值功率的主要因素,基于统计学原理研究了实验数据的统计特性。定量分析的结果表明,温度、SOC、内阻作为输入变量存在严重的共线性,它们对峰值功率的影响是相互耦合的,不存在独立的线性相关性。在此基础上,提出了 GS 和 SVM 相结合的数据驱动方法,建立 GS-SVM 模型对电池峰值功率进行预测。结果表明模型预测值对实验测量值的跟随程度很高,平均误差仅为 3.65%。本文提出的峰值功率预测模型是从数据角度出发,与基于机理型的预测模型相比,不需要电池的内部参数信息,保证了峰值功率预测的高效率。此外,该模型在保证低误差、高精度的基础上,还具有很强的可操作性,推广应用价值高,可以为电动汽车的安全可靠运行提供有力保障。

参考文献:

- [1] 袁笑颜,王雪. 中国电动汽车产业发展背景及可行性分析[J]. 中国电子商务,2009,11:15.
YUAN Xiaoyan,WANG Xue. Analysis of electric vehicle's background and feasibility in China[J]. Chinese Electronic Commerce, 2009,11:15.
- [2] 蒲松林,周念成,郭胤. 基于因子回归分析和分层雷达图法的电动汽车运行状态评估[J]. 电力自动化设备,2015,35(6):44-51.
PU Songlin,ZHOU Niancheng,GUO Yin. Electric vehicle operational status evaluation based on factor regression analysis and layered radar map method[J]. Electric Power Automation Equipment,2015,35(6):44-51.
- [3] 孙丙香,高科,姜久春. 基于 ANFIS 和减法聚类的动力电池放电峰值功率预测[J]. 电工技术学报,2015,30(4):272-280.
SUN Bingxiang,GAO Ke,JIANG Jiuchun. Research on discharge peak power prediction of battery based on ANFIS and subtraction clustering[J]. Transactions of China Electrotechnical Society, 2015,30(4):272-280.
- [4] Idaho National Engineering & Environmental Laboratory. Freedom car battery test manual for power-assist hybrid electric vehicles:

- DOE/ID-11069[S]. Idaho Falls, USA:INEEL,2003.
- [5] 胡宇. 电动车动力电池功率状态预测研究[D]. 哈尔滨:哈尔滨理工大学,2012.
- HU Yu. Prediction status of peak power of battery on HEV[D]. Harbin:Harbin University of Science and Technology,2012.
- [6] 郭宏榆. HEV用锂离子电池功率能力预测方法研究[D]. 北京:北京交通大学,2013.
- GUO Hongyu. Research on power capability prediction method of HEV lithiumion battery[D]. Beijing:Beijing Jiaotong University,2013.
- [7] WAAG W,KÄBITZ S,SAUER D U. Experimental investigation of the lithium-ion battery impedance characteristic at various conditions and aging states and its influence on the application[J]. Applied Energy,2013,102(2):885-897.
- [8] HASTIE T,TIBSHIRANI R,FRIEDMAN J. The elements of statistical learning:data mining,inference,and prediction[M]. Berlin, Germany:Springer,2009.
- [9] ZHENG Fangdan,JIANG Jiuchun,SUN Bingxiang. Capacity estimation of large-scale retired Li-ion batteries for second use based on support vector machine[C]//2014 IEEE 23rd International Symposium on Industrial Electronics(ISIE). Istanbul,Turkey:IEEE, 2014:1628-1634.
- [10] CHATTERJEE S,HADI A S. Regression analysis by example [M]. Hoboken,USA:John Wiley & Sons,2015.
- [11] 刘宝英,杨仁刚. 基于主成分分析的最小二乘支持向量机短期负荷预测模型[J]. 电力自动化设备,2008,28(11):13-17.
- LIU Baoying,YANG Rengang. Short-term load forecasting model based on LS-SVM with PCA[J]. Electric Power Automation Equipment,2008,28(11):13-17.
- [12] 苑津莎,尚海昆. 基于主成分分析和概率神经网络的变压器局部放电模式识别[J]. 电力自动化设备,2013,33(6):27-31.
- YUAN Jinsha,SHANG Haikun. Pattern recognition based on principal component analysis and probabilistic neural networks for partial discharge of power transformer[J]. Electric Power Automation Equipment,2013,33(6):27-31.
- [13] NELLO C,TAYLOR J S. An introduction to support vector machine and other kernel-based learning methods[M]. New York, USA:Cambridge University Press,2000.
- [14] 欧阳庭辉,查晓明,秦亮. 含核函数切换的风电功率短期预测新方法[J]. 电力自动化设备,2016,36(9):80-86.
- OUYANG Tinghui,ZHA Xiaoming,QIN Liang. Short-term wind power prediction based on kernel function switching[J]. Electric Power Automation Equipment,2016,36(9):80-86.
- [15] 王兴玲,李占斌. 基于网格搜索的支持向量机核函数参数的确定[J]. 中国海洋大学学报(自然科学版),2005,35(5):859-862.
- WANG Xingling,LI Zhanbin. Identifying the parameters of the kernel function in support vector machines based on the grid-search method[J]. Periodical of Ocean University of China (Natural Science Edition),2005,35(5):859-862.
- [16] 纪昌明,周婷,向腾飞. 基于网格搜索和交叉验证的支持向量机在梯级水电系统随随机调度中的应用[J]. 电力自动化设备, 2014,34(3):125-131.
- JI Changming,ZHOU Ting,XIANG Tengfei. Application of support vector machine based on grid search and cross validation in implicit stochastic dispatch of cascaded hydropower stations[J]. Electric Power Automation Equipment,2014,34(3):125-131.

作者简介:



郑方丹

郑方丹(1990—),女,福建宁德人,博士研究生,研究方向为动力电池成组应用和管理技术(E-mail:fdzheng@bjtu.edu.cn);

姜久春(1973—),男,吉林四平人,教授,博士研究生导师,博士,研究方向为新能源技术(E-mail:jjcjiang@bjtu.edu.cn);

孙丙香(1979—),女,吉林长春人,副教授,博士研究生导师,博士,通信作者,研究方向为能量存储系统的应用研究(E-mail:bxsun@bjtu.edu.cn)。

Peak power prediction model for batteries based on data statistical characteristic and GS-SVM

ZHENG Fangdan¹,JIANG Jiuchun¹,CHEN Kunlong¹,HAN Zhiqiang²,LOU Tingting³,SUN Bingxiang¹

(1. National Active Distribution Network Technology Research Center,Beijing Collaborative Innovation Center of Electric Vehicles,Beijing Jiaotong University,Beijing 100044,China;2. Beijing Electric Vehicle Co.,Ltd.,Beijing 102606,China;3. State Grid Shandong Electric Power Research Institute, Ji'nan 250002,China)

Abstract: Tests for 10 s peak power and internal resistance are carried out in different temperatures and SOC(State Of Charge) values for lithium manganese batteries. In order to obtain the relationship between external characteristic parameters and peak power data,the temperature,SOC,internal resistance and peak power data obtained from the test are statistically analyzed,including the correlation evaluation and collinear detection among test variables. Based on the analytical results,it is proposed to establish peak power predict model based on GS-SVM(Grid Search-Support Vector Machine). The verification shows that the established model has high prediction accuracy and its mean error is only 3.65%. With short training time,fast calculation speed and strong feasibility,the established model can rapidly estimate the peak power of power battery and ensure the reliable operation of electric vehicles.

Key words: power battery; peak power; correlation analysis; collinear detection; GS-SVM; electric vehicles