

DOI: 10.3901/JME.2020.18.154

# 基于电池组模型和聚类算法的锂离子 电池组SOC不一致估计\*

向兆军<sup>1,2</sup> 胡凤玲<sup>3</sup> 罗明华<sup>1,2</sup> 方崇全<sup>1,2</sup> 胡晓松<sup>3</sup>

(1. 中煤科工集团重庆研究院有限公司 重庆 400039;

2. 瓦斯灾害监控与应急技术国家重点实验室 重庆 400037;

3. 重庆大学汽车工程学院 重庆 400044)

**摘要:** 限于单体的功率和能量, 电池组必须由成百上千的单体串并联而成, 同时依赖于有效的电池管理技术, 从而保证车辆行驶的动力性和经济性。然而制造误差和使用环境的不同导致电池单体存在不一致性, 这使得准确估计所有单体电池荷电状态(State of charge, SOC)变得困难, 从而导致电池发生过充和过放。因此, 融合电池组模型和聚类算法, 提出一种新的锂离子电池组 SOC 不一致估计方法, 在保证精度的同时极大地降低计算复杂度。基于充电数据特征利用二分 k-means 算法将众多单体分为不同等级, 减小需要考虑的对象; 建立考虑参数不一致的电池组模型, 用于估计单体 SOC; 从仿真数据和实车数据两方面验证估计结果的精度和计算复杂度, 并与现有模型对比, 结果表明所提方法单体 SOC 估计的误差在 0.03 以内, 计算效率提高 3~6 倍。

**关键词:** 锂离子电池; 不一致性; 聚类算法; 电池组模型; SOC 估计

**中图分类号:** TM912

## Estimation of SOC Inconsistencies in Lithium-ion Battery Packs Based on Battery Pack Modeling and Clustering Algorithm

XIANG Zhaojun<sup>1,2</sup> HU Fengling<sup>3</sup> LUO Minghua<sup>1,2</sup> FANG Chongquan<sup>1,2</sup> HU Xiaosong<sup>3</sup>

(1. China Coal Technology Engineering Group Chongqing Research Institute, Chongqing 400039;

2. State Key Laboratory of Gas Disaster Detecting, Preventing and Emergency Controlling, Chongqing 400037;

3. Department of Automotive Engineering, Chongqing University, Chongqing 400044)

**Abstract:** Limited by the power and energy of a single cell, a battery pack must consist of hundreds or even thousands of cells in series and parallel, and at the same time rely on an effective battery management system to ensure enough driving power and economy of vehicles. However, due to the manufacturing process and operating environment, there exist parameter inconsistencies among battery cells, which brings challenges to accurately estimating the state of charge (SOC) of all the cells in the pack, and results in over-charge and over-discharge. To estimate the SOC inconsistencies in lithium-ion battery packs, a new method is proposed by combing the battery pack modeling and the clustering algorithm, which can greatly reduce the computational complexity while ensuring accuracy. Based on the features of charging data, the bisecting k-means algorithm is adopted to categorize cells into different levels, reducing the objects that need to be considered. Then, a battery pack model considering parameter discrepancies is established to estimate SOC values of cells. The estimation accuracy and computational complexity have been verified by using both simulation and real-test data, and compared with existing models. The results indicate that the SOC estimation error of a single cell is less than 0.03 by using the proposed method, and the computational efficiency can be improved by 3-6 times.

**Key words:** lithium-ion battery; parameters inconsistency; clustering algorithm; battery pack model; SOC estimation

## 0 前言

随着全球经济快速发展, 车辆需求量剧增, 清

洁、安全、高效的新能源的开发、存储和利用技术成为当前社会的研究热点<sup>[1]</sup>。相比于传统内燃机车辆, 电动车辆以其经济环保的优势得到快速发展。动力电池技术作为电动车辆的关键技术之一, 为保证其能够在复杂的路况和环境下安全高效工作, 需要配置高效的电池管理系统(Battery management

\* 天地科技股份有限公司科技创新创业资金专项(2019-TD-QN036)和国家自然科学基金(51875054)资助项目。20200529 收到初稿, 20200703 收到修改稿

system, BMS)。电池内部荷电状态(State of charge, SOC)的估计, 作为 BMS 的关键技术之一, 可以防止电池在使用过程中发生过充或过放, 减小电池损害, 缓解里程担忧<sup>[2]</sup>。

为满足车辆行驶的功率、能量需求, 车用动力电池由众多的单体通过串联和并联而成。然而成组后电池不一致性使得电池组的 SOC 估计变得困难, 必须要估计每个单体的 SOC。因此, 为充分考虑单体间的差异性, 对电池组中的每个单体建模后估计其 SOC, 称为多单体模型 (Multi-cell model, MCM)<sup>[3]</sup>, 如电-热-老化耦合模型<sup>[4]</sup>, 等效电路模型<sup>[5]</sup>等, 模型精度高, 可以观察各单体的动态特性。但由于需要对每个单体模型进行参数辨识和状态估计, 极大地增加了计算时间。依据电池过充过放原理, ZHONG 等<sup>[6]</sup>在不同均衡策略下建立单体参数与电池组参数间的关系, 通过估计最先满充和满放的两个单体参数推出电池组 SOC。HUA 等<sup>[7]</sup>筛选容量最小的单体为最差的, 并基于带滞回的一阶 RC 模型估计其 SOC 作为电池组的 SOC。这种方法是利用充放电时最“弱”的两个单体建立模型代表电池组, 称为电压最大最小模型 (Voltage maximum and voltage minimum model, VVM), 但是对于最“弱”单体没有统一的定义, 且最多只能表示两个单体的状态。为了减小 MCM 的计算效率, 均值-差异模型 (Mean-difference model, MDM) 被提出。只考虑单体间 SOC 的差异 ( $\Delta SOC$ ), DAI 等<sup>[8]</sup>建立二阶 RC 模型代表电池组的均值模型, 采用双时间尺度的 EKF 估计均值 SOC 和  $\Delta SOC$ , DENG 等<sup>[9]</sup>用电化学模型表示均值模型, 构建 RBF 神经网络用于预测  $\Delta SOC$ 。进一步, ZHENG 等<sup>[10-12]</sup>同时考虑 SOC 和内阻差异, 用 Rint 模型表示差异模型, 能够改善前者的估计精度。PLETT 等<sup>[13]</sup>考虑 SOC, 内阻和容量三方面差异, 用“bar-delta filtering”的方式估计每个单体的 SOC 和 SOH, 但是由于容量是基于估计的 SOC 得出的, 计算准确性不高且计算量大。MDM 虽然综合了前述方法的优缺点, 但其差异模型的数量和电池组单体的数量一样多, 使得 BMS 的计算成本依旧很高。

基于对现有研究的分析, 本文融合电池组模型和聚类算法, 提出一种新的锂离子电池组 SOC 不一致估计方法, 在保证精度的同时极大降低计算复杂度。首先对充电数据提取特征, 利用二分 k-means 算法将众多单体分为不同等级簇, 减小需要考虑的对象; 然后建立考虑参数不一致的电池组模型, 估计每簇的平均 SOC 和各单体 SOC, 并基于仿真数据和实车数据验证所提方法的准确性; 最后与

MDM、MCM 和 VVM 对比 SOC 的估计结果, 旨在体现本方法的优势。

## 1 电池组 SOC 不一致性估计

### 1.1 二分 k-means

对于车用动力电池而言, 在不能拆卸的前提下对成百的单体电池进行分组聚类只能使用实车数据。动力电池的实车数据可分为充电和放电两部分。放电数据主要受车主驾驶风格、行驶工况等的影响, 不可控和不确定性因素较多, 而充电时车辆处于静止状态, 较为稳定, 通常都能达到满充状态, 因此本文采用充电数据作为聚类算法的数据源。以仿真数据为例, 图 1 表示恒流充电下电池的电压变化曲线, 当端电压达到上截止电压后, 电流截止, 端电压迅速出现回降, 这主要是由于充电欧姆内阻导致的, 随后端电压开始缓慢下降直至稳定, 这主要是电池电化学反应中的极化反应引起的。因此, 从充电特性曲线中提取四个特征作为样本的特征矢量 ( $M_1, M_2, M_3, M_4$ ), 分别是充电初始端电压、充电末期终止端电压、充电结束后 1 s 内电压下降值、充电结束后 1~100 s 内电压下降值。

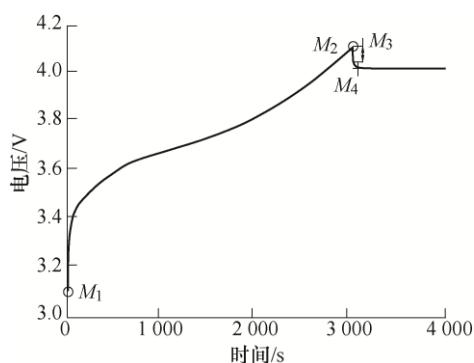


图 1 恒流充电下电池的端电压曲线

为克服 k-means 会收敛于局部最小值的缺点, 采用二分 k-means (Bisecting k-means) 算法对提取的样本特征进行聚类<sup>[14]</sup>。二分 k-means 的实现过程如图 2 所示, 算法的目标是最小化误差平方和 (Sum of squared error, SSE)。此外, 在聚类之前, 需要先将样本特征归一化处理, 消除数据之间量纲的影响。

在实际应用中, 聚类数目  $K$  一般是人为给定的, 不同  $K$  的选择将带来不同的聚类效果。本文通过将 SSE 与轮廓系数<sup>[15]</sup>二者结合作为优化目标确定最佳聚类数, 流程见图 3。SSE 越小, 表示样本的划分越精细, 聚类数  $K$  也就越大。当达到真实聚类数目时, 增大  $K$  值后 SSE 的变化趋于稳定, 这种通过

SSE 变化趋势确定最佳聚类数的方法称为手肘法。轮廓系数结合内聚度和分离度反映聚类结果的好坏,轮廓系数越大,表示聚类效果越好。

---

最大-最小标准化方法处理样本集  
 选定聚类簇数 $K$ ,将所有样本对象视为一个簇  
 当此时的簇数目小于 $K$ 时  
 对于每一个簇  
   计算SSE  
   在给定的簇上面进行k-means聚类( $K=2$ )  
   计算将该簇一分为二之后的整体SSE  
 选择使 SSE 最小的簇进行划分操作

---

图 2 二分 k-means 算法流程

---

确定最大的聚类数目  $K_{\max} = \lceil \sqrt{n} \rceil$  (向上取整)  
 $K=1:K_{\max}$   
   启动二分k-means算法  
   计算轮廓系数和SSE  
 选出轮廓系数值大且SSE变化肘部对应的 $K$ 值

---

图 3 确定最佳聚类数的流程

## 1.2 电池组建模和参数辨识

### 1.2.1 CMM 模型

HU 等<sup>[16]</sup>对比了 12 种电池模型,从计算复杂度和模型精度综合考虑来看,认为对于三元 NCM 电池而言,采用一阶 RC 模型最佳。在本文中无论是仿真实验或是实车采集的数据,电池的材料均为三元 NCM 电池,因此这里选择采用一阶 RC 等效电路模型作为电池组的均值模型结构。

图 4 给出了电池组均值模型的示意图,其中包括一个开路电压源、欧姆内阻、一个并联的 RC,这里的参数均代表电池组的平均水平。根据基尔霍夫定律,该模型的微分方程表达式如下<sup>[17]</sup>

$$\frac{dSOC_m}{dt} = -\frac{\eta I}{C_n} \quad (1)$$

$$\frac{dV_{p,m}}{dt} = -\frac{V_{p,m}}{R_{p,m}C_{p,m}} + \frac{I}{C_{p,m}} \quad (2)$$

$$V_{T,m} = V_{oc,m}(SOC_m) - V_{p,m} - I \cdot R_{o,m} \quad (3)$$

式中, $I$ 表示电路输入电流,其值为正表示放电,为负表示充电; $C_n$ 表示电池容量, $\eta$ 表示库仑效率,本文假设 $\eta=1$ , $SOC_m$ 表示电池组的平均 SOC; $R_{o,m}$ 、 $R_{p,m}$ 、 $C_{p,m}$ 分别表示欧姆内阻、极化内阻和极化电容, $V_{p,m}$ 表示回路 $R_{p,m}-C_{p,m}$ 的电压; $V_{T,m}$ 表示电池平均端电压, $V_{oc,m}$ 表示电池平均开路电压,通常与 SOC 有关。

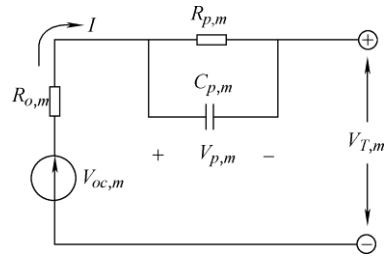


图 4 均值模型

电池的 OCV 与 SOC 之间存在着一定的非线性关系,文献[18]总结了五种表示 OCV-SOC 关系的函数表达,经过不同类型电池的试验对比可知利用多项式拟合的方式具有较高的拟合精度

$$V_{oc}(z) = \sum_{i=0}^p a_i z^i \quad (4)$$

式中, $z$ 为 SOC, $a_i$ 为多项式系数, $p$ 表示多项式最高阶数,本文选取最佳的拟合精度对应 $p=9$ 。

递归最小二乘算法(Recursive least square, RLS)被广泛应用于参数估计领域<sup>[19]</sup>,也经常用来在线辨识电池模型参数。对于 CMM 模型,定义 $V_{d,m} = V_{T,m} - V_a$ 作为系统测量值 $y_m$ ,并将式(2)和式(3)离散化后变换可得

$$V_{d,m}(k+1) = b_1 V_{d,m}(k) - R_{o,m} I(k+1) + (R_{o,m} b_1 - b_2) I(k) \quad (5)$$

$$\text{式中, } b_1 = \exp\left(-\frac{\Delta t}{R_{p,m} C_{p,m}}\right), \quad b_2 = R_{p,m} \cdot \left(1 - \exp\left(-\frac{\Delta t}{R_{p,m} C_{p,m}}\right)\right),$$

$\Delta t$ 为采样间隔。

那么,按照 RLS 辨识系统的标准形式, $\theta$ 和 $\varphi(k)$ 分别可表示为

$$\theta_m = (\theta_{m,1} \quad \theta_{m,2} \quad \theta_{m,3}) \quad (6)$$

$$\varphi_m(k) = (V_{d,m}(k-1) \quad -I(k) \quad I(k-1)) \quad (7)$$

式中, $\theta_{m,1}=b_1$ , $\theta_{m,2}=R_{o,m}$ , $\theta_{m,3}=R_{o,m}b_1-b_2$ 。通过 RLS 可求出 $\theta_m$ ,进一步可求解出参数 $(R_{o,m}, R_{p,m}, C_{p,m})$ 的值。

### 1.2.2 CDM 模型

考虑的不一致性参数不同,所建立的差异模型也是不同的。在仿真验证阶段使用只考虑 SOC 不一致的差异模型(CDM#1),实车验证阶段由于工况较为复杂,单体间差异性较大,使用考虑 SOC 和内阻不一致的差异模型(CDM#2),通过弥补内阻间的差异性增加 SOC 估计的精度。

图 5a 代表 CDM#1, SOC 间的差异带来 OCV 的不同,CDM#1 的模型表达式为<sup>[18]</sup>

$$\begin{cases} \Delta V_{T,i} = \Delta V_{oc,i}(\Delta SOC_i) \\ \Delta V_{T,i} = V_{T,i} - V_{T,m} \end{cases} \quad (8)$$

式中,  $\Delta V_{oc,i}$  表示第  $i$  个单体开路电压和平均开路电压的差值,  $\Delta V_{T,i}$  表示第  $i$  个单体端电压和平均端电压的差值,  $\Delta SOC_i$  表示第  $i$  个单体 SOC 和平均 SOC 的差值。

图 5b 代表 CDM#2, 将 SOC 差异和内阻差异同时作为差异变量, CDM#2 的模型表达式为<sup>[11]</sup>

$$\begin{cases} \Delta V_{T,i} = \Delta V_{oc,i}(\Delta SOC_i) - I \cdot \Delta R_{o,i} \\ \Delta R_{o,i} = R_{o,i} - R_{o,m} \end{cases} \quad (9)$$

式中,  $\Delta R_{o,i}$  表示第  $i$  个单体欧姆内阻和平均欧姆内阻的差值。

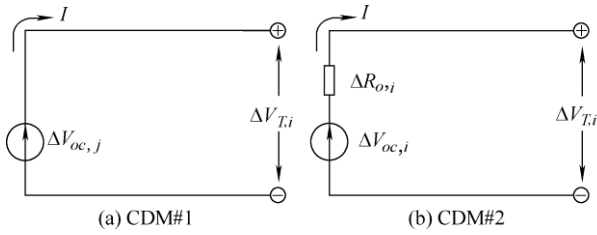


图 5 差异模型

由上述对模型的描述可知, 需要对 CDM#2 进行参数辨识。与 CMM 类似, 则有

$$y_i(k) = \Delta V_{T,i}(k) - \Delta V_{oc,i}(k) \quad (10)$$

$$\theta_i = \Delta R_{o,i} \quad (11)$$

$$\varphi_i = -I(k) \quad (12)$$

### 1.3 电池组 SOC 不一致估计

卡尔曼在 1960 年发表了用于解决离散数据线性滤波的卡尔曼滤波著作<sup>[20]</sup>, 其中的 EKF 在非线性动态系统的状态估计应用中, 具有更为简单的闭环自修正功能, 对存储和计算要求也较低, 因此采用 EKF 来实现对 SOC 的实时估计。

#### 1.3.1 CMM 的 SOC 估计

将电流  $I$  和均值电压  $V_{T,m}$  分别作为系统的输入和输出, 状态变量定义为  $\mathbf{x}_m = (\text{SOC}_m \ V_{p,m})^T$ , 状态方程式(1)~(3)离散化后表示为

$$\begin{pmatrix} \text{SOC}_{m,k} \\ V_{p,m,k} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \exp(-\Delta t / \tau) \end{pmatrix} \begin{pmatrix} \text{SOC}_{m,k-1} \\ V_{p,m,k-1} \end{pmatrix} + \begin{pmatrix} -\eta \Delta t / C_n \\ R_{p,m} (1 - \exp(-\Delta t / \tau)) \end{pmatrix} I_{k-1} \quad (13)$$

测量方程为

$$V_{T,m,k} = V_{oc,m,k}(\text{SOC}_{m,k}) - V_{p,m,k} - I_k R_{o,m} \quad (14)$$

式中,  $\tau = R_{p,m} C_{p,m}$ 。

设

$$\mathbf{A}_{m,k} = \begin{pmatrix} 1 & 0 \\ 0 & \exp(-\Delta t / \tau) \end{pmatrix} \quad \mathbf{C}_{m,k} = \left( \frac{dV_{oc,m}(\text{SOC}_m)}{d(\text{SOC}_m)} - 1 \right)$$

#### 1.3.2 CDM 的 SOC 估计

和 CMM 类似, CDM 的状态变量为  $\mathbf{x}_i = [\Delta \text{SOC}_i \ \Delta R_{o,i}]^T$ , 那么状态方程为

$$\Delta \text{SOC}_{i,k} = \Delta \text{SOC}_{i,k-1} \quad (15)$$

测量方程为

$$\text{CDM\#1: } g(\mathbf{x}_{i,k}) = \Delta V_{oc,i}(\Delta \text{SOC}_i) \quad (16)$$

$$\text{CDM\#2: } g(\mathbf{x}_{i,k}) = \Delta V_{oc,i}(\Delta \text{SOC}_i) - I_k \cdot \Delta R_{o,i} \quad (17)$$

对式(16)和(17)线性化, 则有

$$\begin{aligned} C_{i,k} &= \left. \frac{\partial g}{\partial \mathbf{x}} \right|_{\mathbf{x}_{i,k} = \hat{\mathbf{x}}_{i,k}} = \left. \frac{d\Delta V_{oc,i}(\Delta \text{SOC}_i)}{d\Delta \text{SOC}_i} \right|_{\Delta \text{SOC}_i = \Delta \hat{\text{SOC}}_{i,k}} = \\ &= \left. \frac{d\Delta V_{oc,i}}{d\text{SOC}_i} \cdot \frac{d\text{SOC}_i}{d(\text{SOC}_i - \text{SOC}_m)} \right|_{\text{SOC}_i = \text{SOC}_m + \Delta \hat{\text{SOC}}_{i,k}} = \\ &= \left. \frac{d\Delta V_{oc,i}}{d\text{SOC}_i} \right|_{\text{SOC}_i = \text{SOC}_m + \Delta \hat{\text{SOC}}_{i,k}} \end{aligned}$$

CMM 和 CDM 的状态方程和测量方程都可以表示成 EKF 的标准形式。在同一时间尺度下, 运用 EKF 算法同步估计出  $\text{SOC}_m$  和  $\Delta \text{SOC}_i$ , 保证每一时刻估计值的精确度, 进一步可得出每个单体的 SOC。本文中差异模型的个数取决于二分 k-means 算法的聚类数  $K$ , 属于同一个簇中的单体由于具有相似的特性, 则假设为具有相同的 SOC。

## 2 仿真设计和实车数据采集

利用仿真数据和实车数据验证提出的电池组 SOC 不一致估计方法的有效性。仿真的方法可以模拟电池任何可能的工作状况, 通过设置初始 SOC 的不一致模拟电池组单体间存在的差异现象。而实车数据的来源更加具有不确定性, 数据的质量不高, 对实车数据的验证能够突出本方法的实际工程意义。

### 2.1 仿真搭建和方案设计

结合商业化软件 Autolion-st 和 Matlab Simulink(2017a)搭建 12 串的电组模型。仿真所用电池选择 18650 圆柱形三元 NCM111, 工作电压区间为 2.5~4.2 V, 采样时间间隔为 1 s, 标称容量为 2.15 A·h, 工作温度为 25 °C。针对该电池, 首先进行 OCV-SOC 曲线测试, 在 SOC 区间 0~100%上

每隔 5% 取一个 SOC 点进行 OCV 测试, 然后利用式(4)进行多项式拟合, 结果如图 6 所示, 曲线拟合度  $R^2=0.9991$ 。

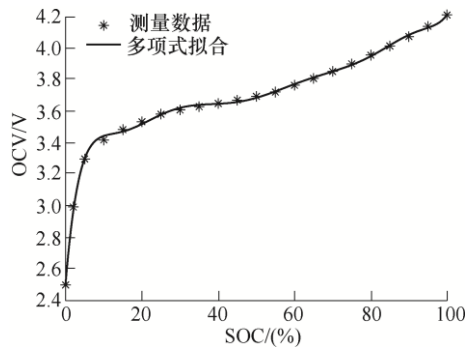


图6 仿真单体 OCV-SOC 曲线拟合

经过文献调研, 车用动力电池组的 SOC 通常呈现一定的分布, 本文将 12 个单体 SOC 的初始状态设置为常见的两组: 正态分布<sup>[21]</sup>和威尔分布<sup>[22]</sup>。正态分布的均值为 8.074 3 %, 方差为 0.05, 威尔分布的均值为 5.915 %, 偏度为 -0.342 6, 二者分布频率与概率密度函数如图 7 所示。然后分别对电池组进行 1C 充电, 采用城市工况 DST(Dynamic stress test)和高速工况 HWFET(Highway fuel economy test)来模拟电池的放电过程。为了便于区分, 定义初始 SOC 符合正态分布的电池组为 Group 1, 初始 SOC 符合威尔分布的电池组为 Group 2。

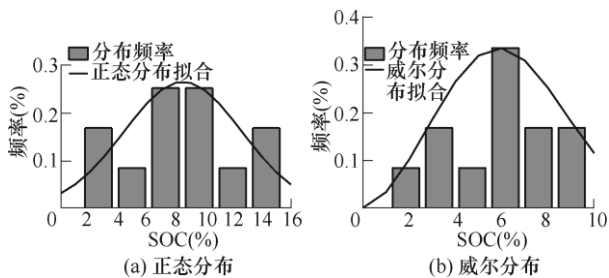


图7 初始 SOC 频率直方图

## 2.2 实车数据采集和预处理

新能源车辆数据采集通过车载终端将采集到的实时数据发送到企业平台, 能够采集不同种类车辆的动态数据、静态数据和相关部件数据。本文选择某款新能源纯电动乘用车搭载的三元 NCM523 软包电池, 标称容量  $37 \text{ A} \cdot \text{h}$ , 电池模组 2 并 86 串, 数据传输的时间间隔是 10 s。上传的数据中含有 86 组电压数据, 将并联的 2 个单体视为一个整体, 定义为  $S_1, \dots, S_{86}$ , 标称容量为  $74 \text{ A} \cdot \text{h}$ , 电池温度变化范围在  $14 \sim 31 \text{ }^\circ\text{C}$ 。电池在  $25 \text{ }^\circ\text{C}$  下进行 OCV-SOC 曲线测试, 多项式拟合后结果如图 8 所示, 曲线拟

合优度为  $R^2=0.9997$ 。

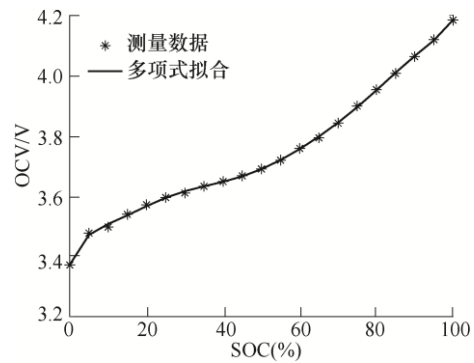


图8 NCM523 的 OCV-SOC 曲线拟合

鉴于实际运行工况的复杂性, 新能源汽车数据存在数据缺失和数据异常等质量问题, 所以本节利用缺失数据处理技术和异常数据剔除技术来处理数据<sup>[23]</sup>。选择一段完整的充放电工况数据, 经过预处理之后的电流和电压数据如图 9 所示。在充电初期电流不稳定, 这是因为电流从 0 跳变到 117 A 幅度太高, 不能迅速达到指定电流大小。但整体来说, 充电过程是多阶段恒流充电。

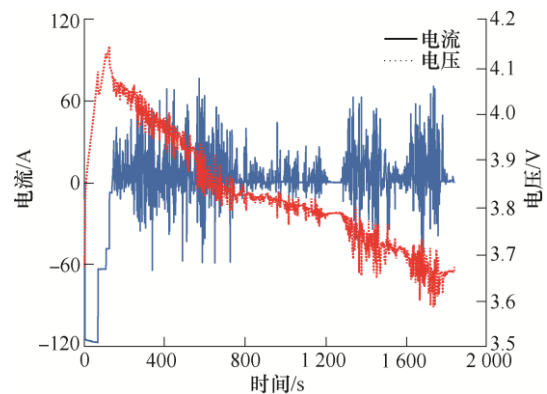


图9 S1 工况

本文选取的这段实车数据记录中, 车辆的行驶里程已有 115 947 km, 电池存在一定的老化现象, 需要重新标定此时的电池容量。电池静置足够长的时间后的端电压可视为电池的 OCV, 选择电池放电时的初始电压作为第一时刻的 OCV。由图 9 可知在 1 200 s 时电流开始为零并持续一段时间, 选择静置的最后一个时刻对应的端电压作为此刻的 OCV, 从而得到两个时刻的 OCV 和 SOC, 进一步得到电池容量, 如图 10 所示, 平均容量为  $64.052 5 \text{ A} \cdot \text{h}$ 。在已知容量的前提下, 通过安时积分可得到每个时刻下的 SOC, 作为下面电池组 SOC 不一致估计的参考值。



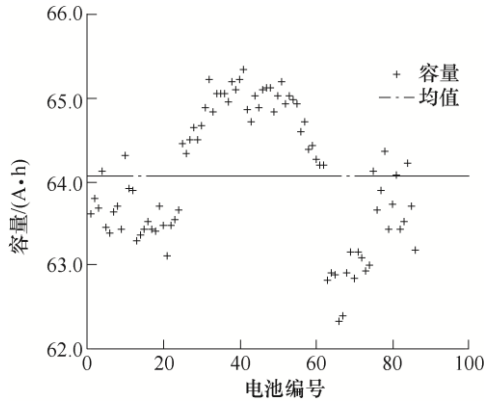


图 10 电池容量(两个一组)

### 3 结果和分析

#### 3.1 仿真数据结果分析

对充电数据提取特征向量后形成  $12 \times 4$  的样本集，然后通过二分 k-means 算法分别对两种分布形式聚类分组。SSE 和轮廓系数的变化如图 11 所示。从图中可知，随着  $K$  值的增加，SSE 不断降低，当  $K$  从 3 取到 4 时，SSE 的下降幅度明显降低，显然， $K=3$  处是肘部法的肘部；另一方面，轮廓系数的变化和  $K$  值没有显示线性关系，但  $K=3$  处的轮廓系数最大。虽然图 11b 中在  $K=6$  时的轮廓系数值高于  $K=3$  处，但是结合 SSE 的变化应该取  $K=3$  最佳。因此两组电池最佳聚类数均选择  $K=3$ ，聚类结果见表 1。

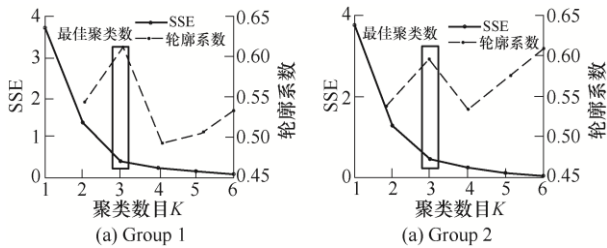


图 11 不同聚类数目下的 SSE 和轮廓系数

表 1  $K=3$  时的聚类结果

簇	电池编号(Group 1)	电池编号(Group 2)
第一簇	Cell 1、Cell 3、Cell 5、Cell 6、Cell 8、Cell 10、Cell 11	Cell 7、Cell 8、Cell 9、Cell 10、Cell 11、Cell 12
第二簇	Cell 2、Cell 7、Cell 9	Cell 1、Cell 2、Cell 3、Cell 5
第三簇	Cell 4、Cell 12	Cell 4、Cell 6

将两组的电池聚类之后对各簇取均值来代表该簇的平均水平，同时认为该簇的单体都等同于该簇的均值水平。由于仿真只考虑了 SOC 的差异，所以差异模型采用 CDM#1，结合 RLS 和 EKF 联合估计模型参数和 SOC。图 12 表示 DST 工况下 Group 1 的估计结果，平均 SOC 的实时估计值和估计误差见图 12a 和 12b，算法的初始平均 SOC 设置为 95%，估计的平均 SOC 能够从一个错误的初始值快速收敛，估计

误差在  $\pm 1\%$  以内，平均 SOC 的 MAE 和 RMSE 分别为 0.003 和 0.004 7。图 12c 表示差异模型中  $\Delta SOC$  的估计结果，为了更好地观察各差异模型估计结果的收敛性，初始  $\Delta SOC$  的绝对值设置为 8%。图 12d 表示各簇的代表 SOC 估计值与各单体 SOC 估计值，与参考 SOC 对比后其估计误差在  $\pm 3.5\%$  范围内。

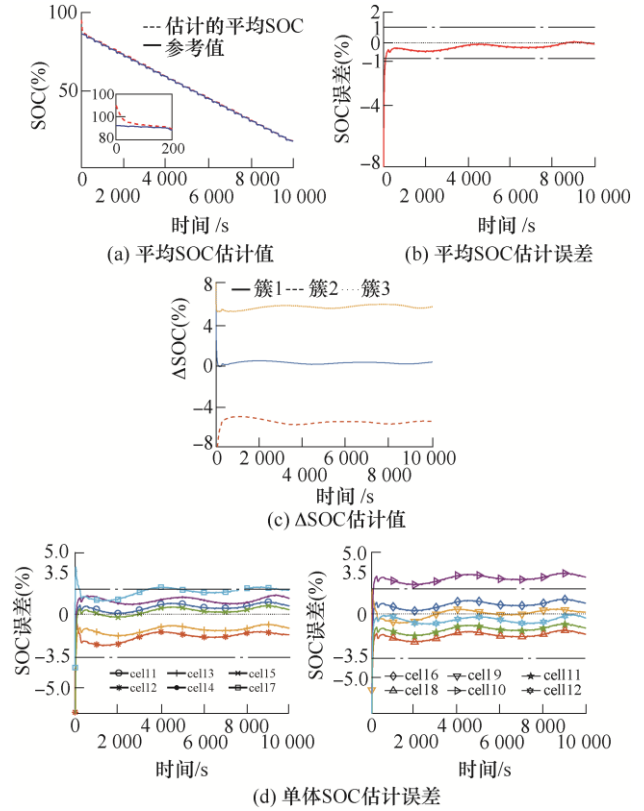
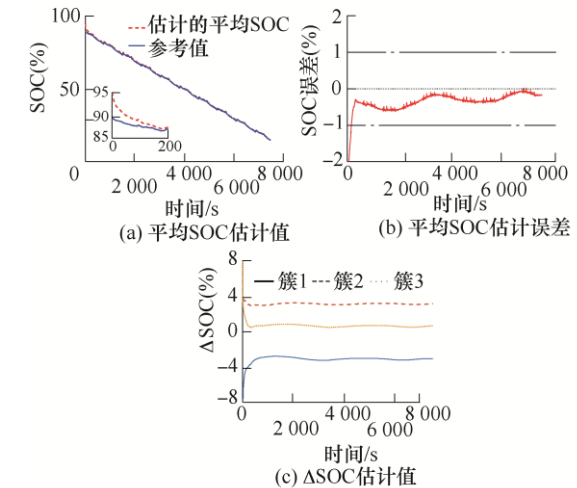


图 12 DST 工况下 Group 1 SOC 的估计结果

图 13 表示 HWFET 工况下 Group 2 的估计结果，平均 SOC 的估计误差在  $\pm 1\%$  以内，SOC 的 MAE 和 RMSE 分别为 0.003 6 和 0.004 9，各单体最终的估计误差在  $\pm 2.5\%$  范围内。由于两组初始 SOC 设置不同，Group 2 的单体间 SOC 差异相比于 Group 1 小，其估计出的各单体 SOC 误差相应也较低。



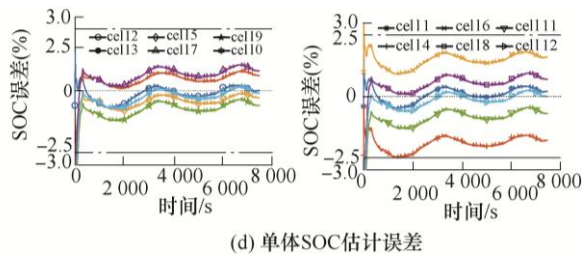


图13 HWFET工况下 Group 2 SOC 的估计结果

### 3.2 实车数据结果分析

实车数据的充电过程是多阶段恒流方式,由图9可知,电流在变化的一刻电压出现突降,且从第一阶段换至第二阶段时的压降最大。因此,为了使得聚类的结果更加精确,实车数据的特征提取增加2个特征,提取第一阶段的终止电压和10 s内的压降,组成 $86 \times 6$ 的样本集用于聚类算法。聚类后的结果如图14所示,在初时轮廓系数虽然较大,但是SSE也大,显然这是不合理的,最终选择 $K=7$ 作为最佳聚类数。

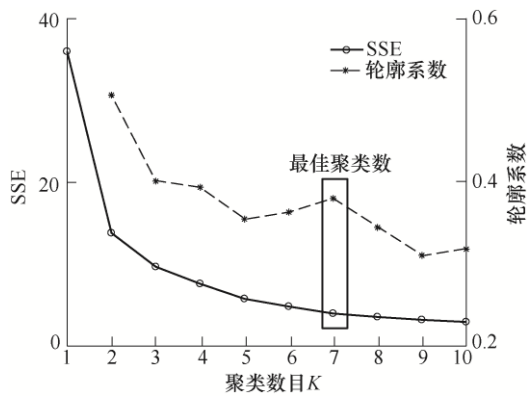


图14 实车工况下的SSE和轮廓系数

基于实车数据进行电池组SOC不一致估计时,使用差异模型CDM#2来估计单体和均值间的差异。由于容量差异估计首先需要精确估计SOC差异,同时也会带来极大的计算量<sup>[10]</sup>,所以暂不考虑容量差异,选择容量的均值 $64 \text{ A} \cdot \text{h}$ 作为每个电池的容量。图15表示实车数据下基于CMM模型估计的结果,分别表示电池组的平均电压、平均SOC、平均欧姆内阻、平均极化内阻和平均极化电容。估计的平均电压相对误差在1%以内,相对误差均值为0.1059,平均SOC误差在 $\pm 1.5\%$ 范围内。实车数据估计的误差相比于仿真数据的误差稍微大一些,这也说明实车数据来源的不确定性和复杂性会影响模型估计的精度。此外,模型参数平均欧姆内阻随着放电的进行缓慢增加,参数在1200 s和1600 s左右有些微波动,而对应的电流值在这段时间变化较小,而较小的电流变化不利于参数的在线估计。

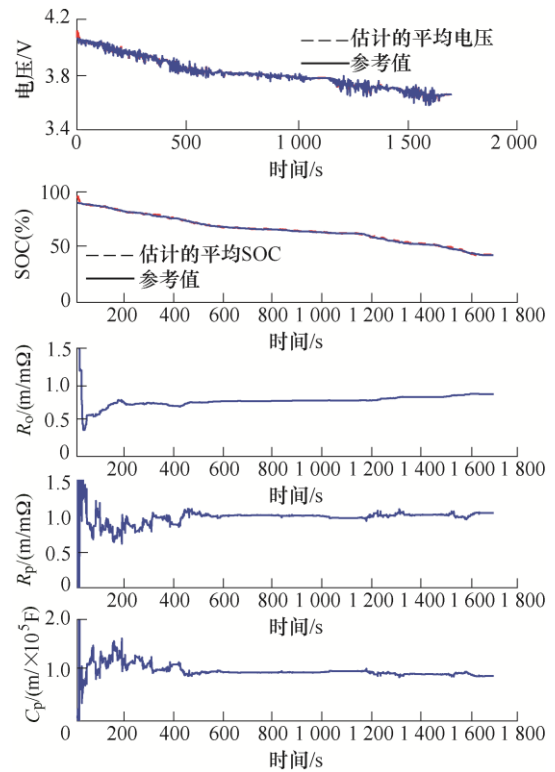
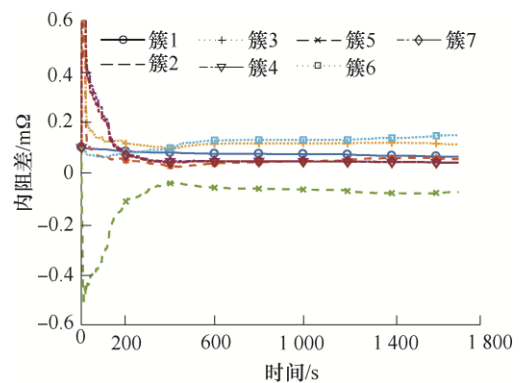


图15 实车数据下均值模型估计结果

差异模型CDM#2相比于CDM#1多了 $\Delta R_o$ ,所以需要同时估计 $\Delta R_o$ 的值,图16a和16b分别表示差异模型下的 $\Delta R_o$ 和 $\Delta \text{SOC}$ 估计结果。 $\Delta R_o$ 的初始值均设置为 $0.1 \text{ m}\Omega$ , $\Delta \text{SOC}$ 的初始值大小为8%,在联合估计算法下二者均能从错误初始值收敛, $\Delta R_o$ 收敛较慢与估计的平均内阻收敛慢有关。图17表示各单体SOC的估计误差,误差范围为 $-3.5\% \sim 3.5\%$ ,同样满足行业5%的标准要求。此外,图中显示估计误差在400~1200 s较小,而在两端出现较大误差值,一方面是前期模型参数和状态参数初始值选择不准确引起的,另一方面是在1200 s和1600 s时出现的小电流变化造成模型参数估计不准确进而影响SOC估计的精确性。

(a)  $\Delta R_o$ 的估计值

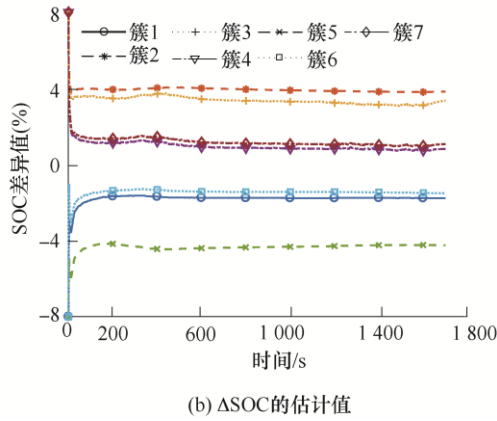


图 16 实车数据下差异模型的估计结果

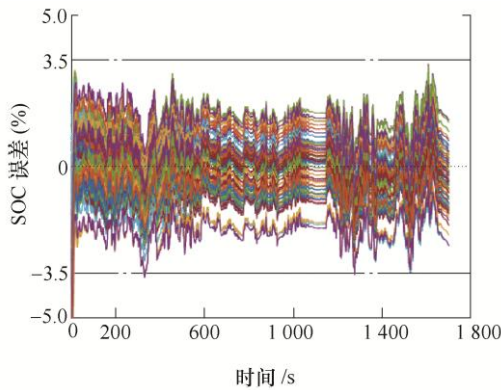


图 17 实车数据下 SOC 的估计结果

### 3.3 模型对比分析

为了进一步凸显本方法的优势,将现有研究中的电池组模型(CMM+CDM, MCM, VVM),以及本文所提模型(Bik+CMM+CDM)从估计精度和计算复杂度两方面进行对比。所有模型和算法的运行环境为 3.4 GHz Inter(R) Xeon(R) E3-1231 CPU 以及 16 GB RAM。

图 18 分别表示 DST 工况、HWFET 工况和实车工况下 SOC 估计的 RMSE,横坐标分别表示模型 [Bik+CMM+CDM, CMM+CDM, MCM, VVM],实车工况下用方均根误差的最大值、最小值和平均值表示。其中 MCM 的精度最高, VVM 的精度最差,本文提出的方法精度适中,但 RMSE 都在 0.03 以内。另外可以看到 Group 2 的 RMSE 整体较 Group 1 低,这说明初始状态的分布情况会影响电池组的估计结果。

表 2 分别展示了仿真数据和实车数据下各电池组的计算耗时,从整体上来说, MCM 的计算复杂度最高,因为它需要对每个单体进行估计, VVM 的耗时最低,只计算了最大和最小的两个电池,本文的方法用时仅比 VVM 稍微高一点点。Bik+CMM+CDM 模型是在 CMM+CDM 上的改进,牺牲

了一部分估计精度,但是仿真中计算效率提高了近 3 倍。随着电池组规模的扩大,计算效率的优势更加显著,可以看到 86 串的车电池组计算效率提高近 6 倍,比 MCM 提高大约 10 倍,耗时 0.665 4 s。因此,本文所提出的 SOC 不一致估计方法能够满足行业的精度要求,却能够极大降低计算复杂度,从而减轻对 BMS 的运行负担。

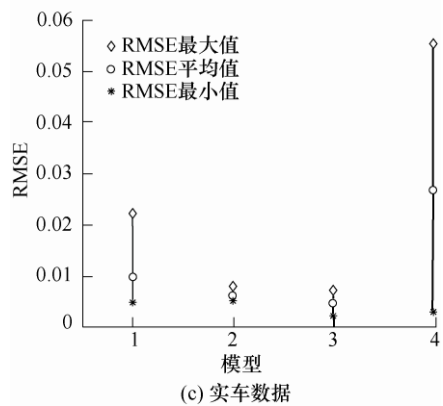
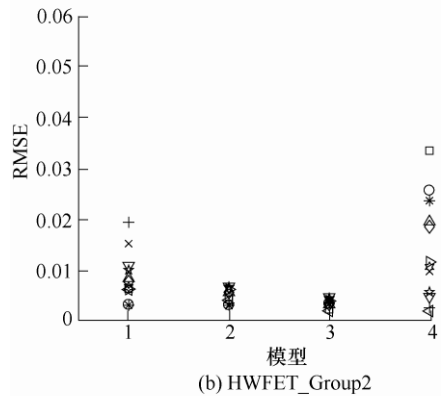
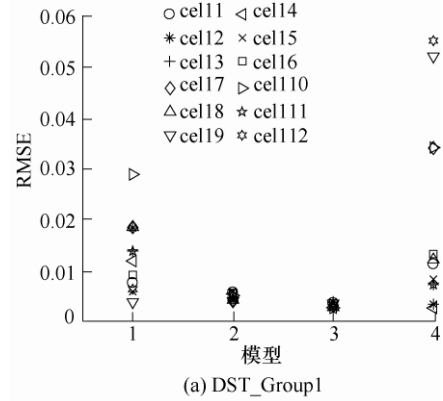


图 18 SOC 估计结果对比

表 2 不同电池组模型计算复杂度对比

工况	模型			
	Bik+CMM+C DM	CMM+C DM	MCM	VVM
DST(Group 1)	1.303 3	3.389 6	5.661 4	0.978 2
DST(Group 2)	1.550 2	4.231 5	6.547 7	1.147 1
HWFET(Group 1)	0.835 4	2.074 2	4.021 2	0.712 7
HWFET(Group 2)	0.927 7	2.331 1	4.382 0	0.770 6
实车	0.665 4	3.514 9	6.946 3	0.161 8



## 4 结论

(1) 采用的二分 k-means 算法能够极大地减少建立的差异模型数量, 结合聚类结果的误差平方和轮廓系数变化确定最佳聚类数目, 为保证 SOC 估计精度奠定基础。

(2) 仿真设置 SOC 分布为正态分布和威尔分布两组, 各单体 SOC 百分比误差范围分别为  $\pm 3.5\%$  和  $\pm 2.5\%$ 。相比于 DST 工况, HWFET 工况下 SOC 估计百分比误差波动更大。这说明 SOC 分布和行驶工况均会影响电池的估计精度。

(3) 将本文所提方法与基于 CMM+CDM、基 MCM 和基于 VVM 的 SOC 估计方法对比, 本文提出的方法精度适中, 但 SOC 的估计误差都在 0.03 以内。同时相比于 CMM+CDM, 实车电池组使用所提方法进行电池组 SOC 不一致估计的计算效率提高近 6 倍, 比 MCM 提高大约 10 倍。

## 参 考 文 献

- [1] ARMAROLI N, BALZANI V. Towards an electricity-powered world[J]. *Energy & Environmental Science*, 2011, 4(9): 3193-3222.
- [2] GARCHE J, JOSSEN A, DÖRING H. The influence of different operating conditions, especially over-discharge, on the lifetime and performance of lead/acid batteries for photovoltaic systems[J]. *Journal of Power Sources*, 1997, 67(1-2): 201-212.
- [3] FENG F, HU X, HU L, et al. Propagation mechanisms and diagnosis of parameter inconsistency within Li-Ion battery packs[J]. *Renewable and Sustainable Energy Reviews*, 2019, 112: 102-113.
- [4] CORDOBA-ARENAS A, ONORI S, RIZZONI G. A control-oriented lithium-ion battery pack model for plug-in hybrid electric vehicle cycle-life studies and system design with consideration of health management[J]. *Journal of Power Sources*, 2015, 279: 791-808.
- [5] SUN F, XIONG R, HE H, et al. Model-based dynamic multi-parameter method for peak power estimation of lithium-ion batteries[J]. *Applied Energy*, 2012, 96(3): 378-386.
- [6] ZHONG L, ZHANG C, HE Y, et al. A method for the estimation of the battery pack state of charge based on in-pack cells uniformity analysis[J]. *Applied Energy*, 2014, 113: 558-564.
- [7] HUA Y, CORDOBA-ARENAS A, WARNER N, et al. A multi time-scale state-of-charge and state-of-health estimation framework using nonlinear predictive filter for lithium-ion battery pack with passive balance control[J]. *Journal of Power Sources*, 2015, 280: 293-312.
- [8] DAI H, WEI X, SUN Z, et al. Online cell SOC estimation of Li-ion battery packs using a dual time-scale Kalman filtering for EV applications[J]. *Applied Energy*, 2012, 95: 227-237.
- [9] DENG Y, XIONG F, YANG B, et al. State-of-charge inconsistency estimation for li-ion battery pack using electrochemical model[C]//Chinese Automation Congress, 2018: 6959-6964.
- [10] ZHENG Y, GAO W, OUYANG M, et al. State-of-charge inconsistency estimation of lithium-ion battery pack using mean-difference model and extended Kalman filter[J]. *Journal of Power Sources*, 2018, 383: 50-58.
- [11] ZHENG Y, OUYANG M, LU L, et al. Cell state-of-charge inconsistency estimation for LiFePO<sub>4</sub> battery pack in hybrid electric vehicles using mean-difference model[J]. *Applied Energy*, 2013, 111: 571-580.
- [12] LI J, KLEE BARILLAS J, GUENTHER C, et al. Multicell state estimation using variation based sequential Monte Carlo filter for automotive battery packs[J]. *Journal of Power Sources*, 2015, 277: 95-103.
- [13] PLETT G L. Efficient battery pack state estimation using bar-delta filtering[C]//EVS24 International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium, 2009: 1-8.
- [14] KRISHNA B V, SATHEESH P, SUNEEL KUMAR R. Comparative study of k-means and bisecting k-means techniques in wordnet based document clustering[J]. *International Journal of Engineering and Advanced Technology*, 2012, 1(6): 1-4.
- [15] ROUSSEEUW P J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis[J]. *Journal of Computational and Applied Mathematics*, 1987, 20: 53-65.
- [16] HU X, LI S, PENG H. A comparative study of equivalent circuit models for Li-ion batteries[J]. *Journal of Power Sources*, 2012, 198: 359-367.
- [17] LIAW B Y, NAGASUBRAMANIAN G, JUNGST R G, et al. Modeling of lithium-ion cells: A simple equivalent-circuit model approach[J]. *Solid State Ionics*,

- 2004, 175(1-4): 835-839.
- [18] HU X, LI S, PENG H, et al. Robustness analysis of state-of-charge estimation methods for two types of Li-ion batteries[J]. Journal of Power Sources, 2012, 217: 209-219.
- [19] CAMPI M. Performance of RLS identification algorithms with forgetting factor: A mixing approach[J]. Journal of Mathematical Systems Estimation and Control, 1997, 7(1): 29-54.
- [20] PLETT G L. Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs: Part 1. Background[J]. Journal of Power Sources, 2004, 134: 252-261.
- [21] ZHOU L, ZHENG Y, OUYANG M, et al. A study on parameter variation effects on battery packs for electric vehicles[J]. Journal of Power Sources, 2017, 364: 242-252.
- [22] ZHANG C, JIANG Y, JIANG J, et al. Study on battery pack consistency evolutions and equilibrium diagnosis for serial-connected lithium-ion batteries[J]. Applied Energy, 2017, 207: 510-519.
- [23] HAN J, PEI J, KAMBER M. Data mining: Concepts and techniques[M]. Amsterdam: Elsevier, 2011.
- 
- 作者简介: 向兆军, 男, 1987 年出生, 助理研究员。主要研究方向为检测技术与自动化装置。  
E-mail: qepyou@126.com  
胡晓松(通信作者), 男, 1983 年出生, 博士, 教授。主要研究方向为电气化车辆动力系统建模与控制。  
E-mail: xiaosonghu@ieee.org