储能测试与评价

基于高斯混合回归的锂离子电池SOC估计

魏 孟,李嘉波,叶 敏,高康平,徐信芯 (长安大学公路养护装备国家工程实验室,陕西 西安 710064)

摘 要: 动力电池剩余电量的准确估算是电动汽车续航里程的重要依据和能量管理的基本前提,为降低电池系统因非线性,不平稳因素对荷电状态估计的不利影响。针对锂电池的状态数据采集存在非线性、不平稳以及外界干扰等问题,提出了基于高斯混合回归对荷电状态进行预测,以解决传统高斯过程模型含有异常数据和噪声的问题。利用 K-means 聚类算法与 EM 算法对高斯混合模型的超参数进行求解,然后采用高斯混合回归对输出的荷电状态进行预测。最后通过实验验证,并与高斯过程回归进行对比分析,验证了高斯混合回归算法在荷电状态估计过程中具有高精度和有效性。

关键词: 动力电池; 荷电状态; 高斯过程回归; 高斯混合回归

中图分类号: TM 911 文献标志码: A 文章编号: 2095-4239 (2020) 03-958-06

SOC estimation of Li-ion battery based on gaussian mixture regression

WEI Meng, LI Jiabo, YE Min, GAO Kangping, XU Xinxin

(Highway Maintenance Equipment National Engineering Laboratory, Chang'an University, Xi'an710064, Shaanxi, China)

Abstract: The state of charge (SOC) of a power battery must be accurately estimated as it determines the endurance mileage and is the basic premise for the energy management of electric vehicles. However, SOC estimation of battery systems is degraded by nonlinearity, instability, and other factors. Accordingly, the characteristic state data of a lithium battery contain nonlinearities, fluctuations, and external interference. This study proposes an SOC prediction method based on Gaussian mixture regression (GMR), which resolves the problems of abnormal values embedded in the state data and noise in the traditional Gaussian procession model (GPM). The hyper-parameters of the Gaussian mixture model are sequentially optimized by k-means clustering and an EM algorithm. The GMR predicts the SOC output. In an experimental validation and a comparative analysis of GMR and GPM, the GMR algorithm achieved superior prediction accuracy and effectiveness in SOC estimation.

Key words: power battery; state of charge; Gaussian process regression; Gaussian mixture regression

随着环境污染问题和能源短缺问题日益突出,电动汽车(EV)被认为是一种新的交通工具[1]。由于锂电池具有较高的能量密度、循环寿命长、自放电率低且成本较低等优点,因此在电动汽车领域广泛应用[2]。动力电池的荷电状态(state of charge,SOC)是动力电池管理系统的基本组成之一,精确

的SOC估计是保证动力电池安全、稳定、有效的工作前提^图。然而,由于动力电池的数据采集系统的不确定性,如温度引起的外界干扰、电磁干扰、紧急刹车等现象,使得采集的数据存在非线性、不平稳以及外界干扰问题。传统的高斯过程回归采用单个高斯分布难以准确的描述估计结果,因此提出基

收稿日期: 2019-10-14; 修改稿日期: 2019-11-27。

基金项目: 国家自然科学基金青年项目 (51805041), 中央高校专项资金 资助项目 (300102259204)。

第一作者及联系人: 魏孟 (1997—), 男, 博士研究生, 研究方向为新能源方向, E-mail: wm13484520242@163.com。

于高斯混合过程回归对锂离子电池的SOC进行估 计,并通过仿真验证该算法的有效性和高精度[4]。

目前,对于SOC的估计方法主要包括两大类: 基于模型的SOC估计和基于数据驱动的SOC估 计^[5]。基于模型的 SOC 估计起步较早,通过建立电 池的等效电路模型和状态方程,引入卡尔曼滤波算 法进行SOC估计[6]。提出卡尔曼滤波(KF)、扩展 式卡尔曼滤波(EKF)、自适应卡尔曼滤波 (AKF)、无迹卡尔曼滤波(UKF)以及各种滤波算 法之间的组合形式[7-8]的方法,提高了SOC的估计精 度。但是,基于模型的SOC估计方法需要建立精确 的电池模型。然而,基于数据驱动的SOC估计方法 避免建立精确的电池模型, 在实验测试电池的大量 离线数据的基础上,通过训练电池的电压、电流以 及温度等特性,建立电池的SOC映射模型。此方法 突破了原有的物理模型束缚,解决强非线性问题有 独特的优势且具有较高的估计精度的。人工神经网 络(ANNs)在数据驱动方法中最为广泛[10]。支持 向量机(SVM)在VC维空间理论的基础上,通过 结构化风险最小原理来训练模型,对于小样本、非 线性以及高维度等问题具有较好的适应能力凹。近 期,高斯过程回归(GPR)的研究引起人们的关 注, GPR 方法具备 ANNs 和 SVM 的优点, 但是单 高斯过程(GP)的分布假设对于非高斯条件以及 不准确的数据描述将影响结果的准确性。针对上述 问题,建立高斯混合模型 (GMM), GMM采用多 个单高斯密度函数来逼近任意形状的密度分布,能 对非线性、不稳定的数据的特性能进行很好的逼 近。因此,针对动力电池的数据采集系统的不确定 的问题,提出了基于高斯混合回归的 SOC 估计 方法[12-16]。

高斯混合回归模型原理

1.1 高斯过程回归

GP是任意有限个随机变量均具有联合高斯分 部的集合, GP 由均值函数 $\mu_{(x)}$, 核函数k(x,x')。因 此GP定义为

$$f(x) \sim GP[(\mu_{(x)}, k(x, x'))] \tag{1}$$

假设噪声 $\varepsilon\sim N(0,\sigma_n^2)$, 在回归问题中, 考虑模 型y = f(x) + ε, 其中: y是观测值; x是输入向量; f是函数值。通常对数据进行处理取均值函数为0, 则可得y的先验分布和预测值y*的先验分布

$$y \sim N \left[0, \mathbf{k} \left(X, X \right) + \sigma_n^2 \right] \tag{2}$$

$$\begin{bmatrix} y \\ y_* \end{bmatrix} \sim N \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \boldsymbol{k}(X, X) + \sigma_n^2 \boldsymbol{I}_n & \boldsymbol{k}(X, \boldsymbol{x}_*) \\ \boldsymbol{k}(X_*, X) & \boldsymbol{k}(X_*, \boldsymbol{x}_*) \end{bmatrix} \right\}$$
(3)

式中, 输入X, 测试集 x_* , I_n 为n维单位矩阵; k(X,X)表示输入自身的协方差矩阵; $k(X,x_*)$ 表示 测试点与输入的协方差矩阵; $k(x_*,X)$ 表示输入与 测试点的协方差矩阵; $k(x_*,x_*)$ 表示测试集自身的 协方差矩阵。即得火,的后验分布为

$$P(y_*|X,y,x_*)\sim N(\mu_*,\Sigma_*)$$
(4)

$$\mu_* = \mathbf{k}(\mathbf{x}_*, X) \left[\mathbf{k}(X, X) + \sigma_n^2 I_n \right]^{-1} y \tag{5}$$

 $\Sigma_* = k(x_*, x_*) - k(x_*, X) \left[k(X, X) + \sigma_n^2 I_n \right]^{-1} k(X, x_*)$ (6) 式中, μ_* 为预测值的均值; Σ_* 为预测值的 方差。

高斯过程回归(GPR)常以平方指数协方差函 数如下

$$\mathbf{k}(x, x') = \sigma_f^2 \exp \left[-\frac{1}{2} (x - x') M^{-1} (x - x') \right]$$
 (7)

 $\theta = \{M, \sigma_n^2, \sigma_f^2\}$ 是高斯超参数,其中M = $\operatorname{diag}(l^2)$,l是方差尺度; σ_f^2 是信号方差。采用最大 似然估计函数对超参数进行优化见式(8)

$$L = \lg P(y|x,\theta) = \frac{1}{2} \lg \left[\left(\Delta \left(k + \sigma_n^2 I_n \right) \right) - \frac{1}{2} y^T \left(k + \sigma_n^2 I_n \right)^{-1} y - \frac{N}{2} \lg 2\pi \right]$$
(8)

针对超参数 θ 对似然函数L求偏导,最终得到 最优超参数。带入式(5)、(6)获得在 x_* 的预测值的 均值和方差。

1.2 高斯混合模型

设定动力电池数据是由n个独立的观测变量组 成。假设GMM每个样本均可由一个高斯分布表 示,从几何角度看,高斯混合模型为多个高斯分布 的加权平均,即用有限多个单高斯密度函数的加权 来逼近任意的连续分布[17]。概率密度函数为

$$P(x) = \sum_{k=1}^{k} \alpha_k N(\mu_k, \Sigma_k)$$
 (9)

式中, α_k 为第k个高斯分量的权重; μ_k 为第k个高斯分量的均值; Σ_k 为第k个高斯分量的协方 差;k表示GMM中单个高斯分量的个数,因此k个 高斯分量进行 GMM 模型的超参数为 Q= $\{\{\alpha_1,\mu_1,\Sigma_1\},\{\alpha_2,\mu_2,\Sigma_2\},...,\{\alpha_n,\mu_n,\Sigma_n\}\}$

最大期望 (EM) 算法是高斯混合模型中进行

求解最优超参数应用最为广泛的一种方法。EM 算法由 E-step 和 M-step 交替组成,通过迭代运算,最终获得 GMM 模型的超参数。

E-step

$$P^{(m)}(i|x_n,Q) = \frac{\alpha_i^{(m)} N \left[x_n | \mu_i^{(m)}, \Sigma_i^{(m)}\right]}{\sum_{k=1}^{M} \alpha_k^{(m)} N \left[x_n | \mu_k^{(m)}, \Sigma_k^{(m)}\right]}$$
(10)

M-step

$$\alpha'_{i} = \frac{\sum_{n=1}^{N} P^{(m)}(i|x_{n},Q)}{N}$$
 (11)

$$\mu'_{i} = \frac{\sum_{n=1}^{N} P^{(m)}(i|x_{n},Q)x_{n}}{\sum_{n=1}^{N} P^{(m)}(i|x_{n},Q)}$$
(12)

$$\Sigma_{i}^{t} = \frac{\sum_{n=1}^{N} P^{(m)}(i|x_{n},Q) x_{n}^{2}}{\sum_{n=1}^{N} P^{(m)}(i|x_{n},Q)} - \mu_{i}^{2}$$
(13)

1.3 高斯混合回归

通过建立 GMM 模型,利用 EM 算法对高斯混合的超参数求解得出每组高斯的权重 α ,均值 μ 以及方差 Σ 。根据高斯条件和高斯分布的线性组合进行高斯混合回归^[18]。动力电池的数据由输入数据和输出数据两部分组成:输入训练集数据x包括,电流A和电压V,期望输出数据y,即为 SOC。给定x、y,则混合高斯联合密度

$$f_{XY}(x,y) = \sum_{i=1}^{k} \alpha_{i} \phi(x,y,\mu_{i},\Sigma_{i})$$
 (14)

$$\mu_{i} = \begin{bmatrix} \mu_{jx} \\ \mu_{iy} \end{bmatrix} \Sigma_{j} = \begin{bmatrix} \Sigma_{jxx} & \Sigma_{jxy} \\ \Sigma_{iyx} & \Sigma_{iyy} \end{bmatrix}$$
 (15)

其中条件分布的均值和方差分别为

$$\mu_{j(x)} = \mu_{jy} + \sum_{jyx} \sum_{jxx}^{-1} (x - \mu_{jx})$$
 (16)

$$\Sigma_{j(x)} = \Sigma_{jyy} - \Sigma_{jyx} \Sigma_{jxx}^{-1} \Sigma_{jxy}$$
 (17)

回归表达式为

$$\mu_{(x)} = \sum_{j=1}^{k} \alpha_{j(x)} \mu_{j(x)}$$
 (18)

$$v_{(x)} = \sum_{j=1}^{k} \alpha_{j(x)} - \left[\mu_{j(x)}^{2} + \sum_{j=1}^{2}\right] \left[\sum_{j=1}^{k} \alpha_{j(x)} \mu_{j(x)}\right]^{2}$$
(19)

基于高斯混合回归算法(GMR)的核心思想 是利用有限多个单高斯密度函数,计算每簇的权 重,均值和方差进行拟合逼近任意连续分布。该算 法的流程如表1所示。

2 实验验证及分析

选用磷酸锂电池,在城市道路循环工况

表1 基于GMR的电池SOC估计

Table 1 Battery SOC estimation based on GMR

算法:基于GMR的电池SOC估计

Step1.输入训练集数据

Step2. 通过 K-means 进行聚类

Step3.建立电流,电压和 SOC 的联合密度 GMM 模型

$$P(x) = \sum_{k=1}^{k} \alpha_k N(\mu_k, \Sigma_k)$$

Step4. 利用 EM 算法得到模型的最优参数

$$Q = \{\{\alpha_1, \mu_1, \Sigma_1\}, \{\alpha_2, \mu_2, \Sigma_2\}, ..., \{\alpha_n, \mu_n, \Sigma_n\}\}$$

Step5.输入测试数据,并计算相应高斯分量测试集的后验概率

$$\mu_{j(x)} = \mu_{jy} + \sum_{jyx} \sum_{jxx}^{-1} (x - \mu_{jx}) \sum_{j(x)} = \sum_{jyy} - \sum_{jyx} \sum_{jxx}^{-1} \sum_{jxy}^{-1} \sum_{jxy}^{-1$$

Step6.根据GMR,输出预测结果

$$\mu_{(x)} = \sum_{j=1}^{k} \alpha_{j(x)} \mu_{j(x)} \quad v_{(x)} = \sum_{j=1}^{k} \alpha_{j(x)} \left[\mu_{j(x)}^{2} + \Sigma_{j}^{2} \right] - \left[\sum_{j=1}^{k} \alpha_{j(x)} \mu_{j(x)} \right]^{2}$$

(urban dynamometer driving schedule,UDDS)和动态应力测试工况(dynamic stress test,DST)下获取实验数据,验证GMR方法的有效性,并与GPR进行对比。为获取实验所需的数据,搭建电池测试平台,设计测试程序得到电池的特性测试结果。采用EVT500-500设备测试系统以及配套的上位机,恒温箱,Fluke数据记录仪和相关配套的充放电设备在20℃下进行数据采集,分别取3000样本作为训练确定模型的最优参数。由于磷酸锂电池的数据在采集的过程中,数据不仅会随着温度的变化而发生波动而且数据采集过程中存在电磁干扰、紧急刹车等情况,使得电池的SOC曲线最终呈现非线性、不平稳趋势。图1所示为UDDS工况下的SOC数据,如图2所示为DST工况下的SOC数据。

为了对 GPR、GMR 的预测性能进行评估,采 用均方根误差(root mean square error,RMSE)和

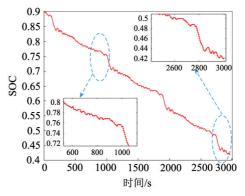


图 1 UDDS 工况下的 SOC 数据 Fig.1 SOC under UDDS conditions

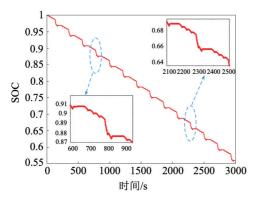


图 2 DST工况下的SOC数据 Fig.2 SOC under DST conditions

绝对百分比均值误差(mean absolute percentage error,MAPE)两个指标。RMSE和MAPE的计算式如下

$$RMSE = \sqrt{\frac{\sum_{1}^{n} (Y_{e} - Y_{i})^{2}}{n}}$$
 (20)

$$MAPE = \frac{1}{n} \sum_{i}^{n} \frac{\left| Y_{e} - Y_{i} \right|}{\left| Y_{i} \right|}$$
 (21)

式中, Y_i 是 SOC 的实际值; Y_c 是 SOC 的估计值;n为测试样本的个数。

分别在UDDS,DST工况下,将采集的3000个电压和电流数据作为GPR的训练集并将其所对应的3000个SOC值作为测试数据,分别采用GPR和GMR对SOC进行估计结果:如图3所示基于GPR的SOC的预测结果(UDDS),图4为GPR和GMR对SOC的预测结果对比(UDDS),图5为基于GPR的SOC的预测结果(DST),图6为GPR和GMR对SOC的预测结果对比(DST)。

分析 GPR 和 GMR 对 SOC 的预测结果,根据图

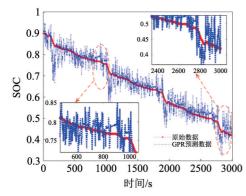


图 3 基于 GPR 的 SOC 的预测结果(UDDS) Fig.3 The result of SOC prediction based on GPR (UDDS)

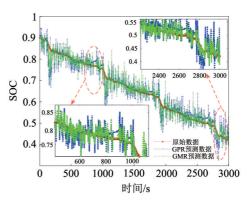


图 4 GPR和GMR对SOC的预测结果对比(UDDs)
Fig.4 Comparison of GPR and GMR prediction results for SOC(UDDS)

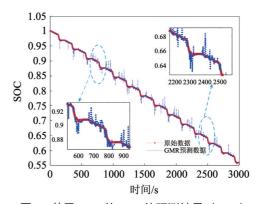


图 5 基于 GPR 的 SOC 的预测结果(DST) Fig.5 The result of SOC prediction based on GPR (DST)

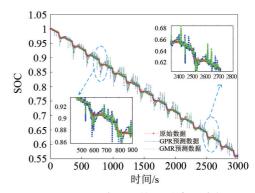


图 6 GPR和GMR对SOC的预测结果对比(DST) Fig. 6 Comparison of GPR and GMR prediction (DST)

3、图 5 可以得出,由于动力电池采集的数据具有 非线性、不稳定和外界干扰现象,因此实际的 SOC 数据曲线并不是相对平稳的直线,而呈现不同斜率 的下降曲线,甚至出现明显的斜率下降阶段。传统 的 GPR 在上述斜率下降较明显的阶段估计偏差较 大。针对采集数据存在的问题,采用有限多个单高 斯来逼近任意连续分布,根据图 4、图 6 可以得出, GMR的预测精度有所提高,尤其是在SOC斜率突降阶段较传统的GPR预测精度明显提高。为了进一步比较GPR和GMR的预测性能,分别在UDDS、DST工况下,计算出两种算法的预测值和真实值之间的误差图: 图 7 为 GPR和 GMR的误差对比(UDDS),图 8 为 GPR和 GMR的误差对比(DST)。

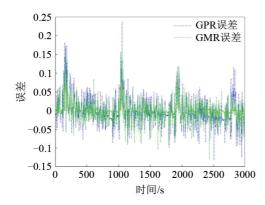


图 7 GPR和GMR的误差对比(UDDS) Fig.7 The comparison of error between GPR and GMR (UDDS)

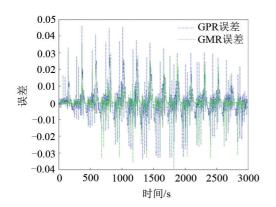


图 8 GPR和GMR的误差对比(DST)
Fig.8 The comparison of error between GPR and
GMR (DST)

为了验证 GMR 方法的有效性,本文将GMR和GPR的预测结果进行比较。表2描述了GMR和GPR在训练集上的预测性能表。由表2可知,在UDDS工况下,GMR的均方根误差为2.37%,较GPR误差降低了29%,GMR的绝对百分比均值误差为2.23%,较GPR误差降低了41%。在DST工况下,GMR的均方根误差为0.64%,较GPR误差降低了30%,GMR的绝对百分比均值误差为2.22%。

表 2 GMR和 GPR 在训练集上的预测性能
Table 2 Prediction performance of GMR and GPR
on training

方法	RMSE/%	MAPE/%
GPR (UDDS)	3.34	3.79
GMR (UDDS)	2.37	2.23
GPR (DST)	0.91	2.24
GMR (DST)	0.64	2.22

3 结 论

为了提高SOC的估计精度,提出高斯混合回归的方法进行锂电池的SOC估计。由于锂电池的状态数据采集中存在的非线性、不平稳以及外界干扰问题,传统的单个高斯过程难以解决此问题,在分析高斯过程回归的基础上利用有限多个单高斯密度函数拟合逼近任意连续分布,通过求解多个单高斯密度函数的权值、均值、方差,根据GMR对锂离子电池的SOC进行预测。通过实验仿真,将GMR的估计精度与GPR相比较,GMR估计精度提高了41%,验证了GMR算法的高精度性。

参考文献

- [1] CHENG C, RUI X, SHEN W. A lithium-ion battery-in-the-loop approach to test and validate multi-scale dual H infinity filters for state of charge and capacity estimation[J]. IEEE Transactions on Power Electronics, 2018(99): doi: 10.1109/TPEL.2017.2670081.
- [2] MENG J, LUO G, FEI G. Lithium polymer battery state-of-charge estimation based on adaptive unscented Kalman filter and support vector machine[J]. IEEE Transactions on Power Electronics, 2015, 31 (3): 2226-2238.
- [3] ZHANG Y, RUI X, HE H, et al. Lithium-ion battery pack state of charge and state of energy estimation algorithms using a hardware-inthe-loop validation[J]. IEEE Transactions on Power Electronics, 2017, 32(6): 4421-4431.
- [4] SHENG H, XIAO J, WANG P. Lithium iron phosphate battery electric vehicle state-of-charge estimation based on evolutionary gaussian mixture regression[J]. IEEE Transactions on Industrial Electronics, 2016, 64(1): 544-551.
- [5] 盛瀚民. 计及数据不可靠性的动力电池组 SOC 估计方法研究[D]. 成都: 西南交通大学, 2017.
 - SHENG Hanmin. A study on power battery SOC estimation considering date uncertainty[D]. Chengdu: Southwest Jiatong University, 2017.
- [6] 程晨, 顾伟, 褚建新, 等. 基于 Sage-Husa 自适应滤波算法的锂电池荷电状态估计[J]. 南京工业大学学报(自然科学版), 2016, 38(3): 126-130.
 - CHENG Chen, GU Wei, CHU Jianxin, et al. State of charge estimation of lithium batteries based on Sage-Husa adaptive filter algorithm[J].

- Journal of Nanjing Tech University (Natural Science Edition), 2016, 38 (3): 126-130.
- [7] 田茂飞, 安治国, 陈星, 等. 基于在线参数辨识和 AEKF 的锂电池 SOC 估计[J]. 储能科学与技术, 2019, 8(4): 745-750.
 - TIAN Maofei, AN Zhiguo, CHEN Xing, et al. SOC estimation of lithium battery based online parameter identification and AEKF[J]. Energy Storage Science and Technology, 2019, 8(4): 745-750.
- [8] LI B, BEI S. Estimation algorithm research for lithium battery SOC in electric vehicles based on adaptive unscented Kalman filter[J]. Neural Computing and Applications, 2018: doi: 10.1007/s00521-018-3901-7.
- [9] YUE L, CHATTOPADHYAY P, XIONG S, et al. Dynamic data-driven and model-based recursive analysis for estimation of battery state-ofcharge[J]. Applied Energy, 2016, 184: 266-275.
- [10] 苏振浩, 李晓杰, 秦晋, 等. 基于 BP 人工神经网络模型的动力电池 SOC 估计方法[J]. 储能科学与技术, 2019, 8(5): 868-873.
 - SU Zhenhao, LI Xiaojie, QIN Jin, et al. SOC estimation method of power battery based on BP artificial neural network[J]. Energy Storage Science and Technology, 2019, 8(5): 868-873.
- [11] 戴庚, 耿诗尧. 基于最小二乘支持向量机的 SOC估计方法[J]. 信息与电脑(理论版), 2018, 39(2): 36-37.
 - DAI Geng, GENG Shiyao. SOC estimation based on least squares support vector machines[J]. China Computer & Communication, 2018, 39(2): 36-37.
- [12] MEI C, SU Y, LIU G, et al. Dynamic soft sensor development based on Gaussian mixture regression for fermentation processes[J]. Chinese Journal of Chemical Engineering, 2016, 25(1): 116-122.
- [13] SAHINOGLU G O, PAJOVIC M, SAHINOGLU Z, et al. Battery state

- of charge estimation based on regular/recurrent gaussian process regression[J]. IEEE Transactions on Industrial Electronics, 2017(99): doi: 10.1109/TIE.2017.2764869.
- [14] LIU D, PANG J, ZHOU J, et al. Prognostics for state of health estimation of lithium-ion batteries based on combination Gaussian process functional regression[J]. Microelectronics Reliability, 2013, 53 (6): 832-839.
- [15] 何志昆, 刘光斌, 赵曦晶, 等. 高斯过程回归方法综述[J]. 控制与决策, 2013(8): 1121-1129.
 - HE Zhikun, LI Guangbin, ZHAO Xijing, et al. Overview of Gaussian process regression [J]. Control and Decision, 2013(8): 1121-1129.
- [16] 李嘉波, 魏孟, 叶敏, 等. 基于高斯过程回归的锂离子电池 SOC 估计 [J]. 储能科学与技术, 2019, 9(1): 131-137.
 - LI Jiabo, WEI Meng, YE Min, et al. SOC estimation of lithium-ion batteries based on Gauss process regression[J]. Energy Storage Science and Technology, 2019, 9(1): 131-137.
- [17] 周亚同, 赵翔宇, 何峰, 等. 基于高斯过程混合模型的大气温湿度预测 [J]. 农业工程学报, 2018, 34(5): 219-226.
 - ZHOU Yatong, ZHAO Xiangyu, HE Feng, et al. Atmospheric temperature and humidity prediction of Gaussian process mixed model [J]. Transactions of the Chinese Society of Agricultural Engineering, 2018, 34(5): 219-226.
- [18] 杨飞, 乔铁柱, 庞宇松, 等. 基于改进的高斯混合回归的球磨机料位软测量[J]. 现代电子技术, 2018, 41(5): 153-158.
 - YANG Fei, QIAO Tiezhu, PANG Yusong, et al. Soft measurement for ball mill fill level based on improved Gaussian mixture regression[J]. Modern Electronics Technique, 2018, 41(5): 153-158.