Article

The Prediction of Capacity Trajectory for Lead–Acid Battery Based on Steep Drop Curve of Discharge Voltage and Gaussian Process Regression

Qian Li [1,2], **Guangzhen Liu** [3], **Ji'ang Zhang** [4], **Zhan Su** [1,2], **Chunyan Hao** [1,2], **Ju He** [3] **and Ze Cheng** [4,*]

[1] Electric Power Research Institute of Tianjin Power System, Tianjin 300022, China; qian.li5@tj.sgcc.com.cn (Q.L.); zhan.su@tj.sgcc.com.cn (Z.S.); chunyan.hao@tj.sgcc.com.cn (C.H.)
[2] Tianjin Key Laboratory of Internet of Things in Electricity, Tianjin 300022, China
[3] STATE GRID Tianjin Electric Power Company, Tianjin 300022, China; guangzhen.liu@tj.sgcc.com.cn (G.L.); ju.he@tj.sgcc.com.cn (J.H.)
[4] School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China; ZhangJA@tju.edu.cn
* Correspondence: Chengze@tju.edu.cn

Abstract: In this paper, a method of capacity trajectory prediction for lead-acid battery, based on the steep drop curve of discharge voltage and improved Gaussian process regression model, is proposed by analyzing the relationship between the current available capacity and the voltage curve of short-time discharging. The battery under average charging is discharged for a short time, and the voltage drop of short-time discharging during equal time intervals, which has the highest relevance with presently available capacity, is extracted as the health feature (HF), and the ergodic method is used to search the optimal time interval. Then, a Gaussian process regression (GPR) model, which reflects the capacity degradation of lead–acid battery, is established (with the HF series as input and current available capacity series as output). Considering the complex trend of capacity trajectory, the rational quadratic covariance function is used as the kernel function of GPR model, and the conjugate gradient algorithm is used for optimization, in order to improve the nonlinear mapping ability of GPR. Finally, the experimental results of lead-acid batteries under different charging cut-off voltages and operating temperatures show that the proposed method can effectively predict the capacity change trajectory of lead–acid batteries with a small training sample, showing high prediction accuracy and wide applicability.

Keywords: lead–acid battery; health feature; improved Gaussian process regression; capacity prediction

1. Introduction

Lead–acid batteries have become important energy storage devices in DC substation, microgrid system, electric vehicle, and other fields, with stable output voltage, good safety performance, low operation cost, long service life, and superior environmental protection conditions [1]. With 50% of the share of the battery industry and 70% of the share of the rechargeable battery industry, the lead–acid battery has become the largest battery product in the world. Its technology has also gone through three main development stages: open battery, rich liquid maintenance-free battery, and valve-controlled sealed maintenance-free battery. Among them, the emergence of the valve-regulated lead–acid battery (VRLA battery) created a new era in the development of lead–acid battery, realized the reuse of oxygen in the process of internal chemical reaction of battery with AGM technology, and quickly occupied the market. The valve-regulated lead–acid battery has not only made a breakthrough in theoretical research, but it has also greatly improved in product types and performance. Therefore, as an important piece of energy storage equipment for backup power supply systems and the main power source for electric products, the VRLA battery is widely used in aviation and navigation, transportation, military communication,

power systems, and other fields. The RLA battery, as the main power source, has been used in automobiles, electric bicycles, scooters, etc., with a proportion of more than 60%. As a backup energy storage equipment, it is used in banks, electric power, hospitals, schools, shopping malls, and other places, accounting for about 21%. In addition, it is also widely used in all aspects of life, such as emergency lights, electric toys, and lighting power supplies, and plays an indispensable role in our production and life.

The lead–acid battery, used in substation, is the last safety barrier of DC system, and its performance is related to the operation safety of the whole power system. At present, the valve-regulated lead–acid battery (VRLA) is mainly used in substations [2], which is the core component of the DC power supply system in the station, whose state of health (SOH) is related to the safe operation of other equipment in the station. SOH is a measure of the current available capacity of the battery, which can be used as a direct reflection of the health status [3]. The low available capacity of the battery tends to indicate that the battery is close to the life threshold. In order to ensure operation safety, the battery should be maintained or replaced in time.

The current available capacity of the battery generally adopts the method of periodic discharge and capacity verification, but the time interval is long, generally once every half a year. The capacity verification time, through small current discharge, is generally about 10 hours, which cannot meet the practical needs of online detection of the SOH of battery. In order to facilitate online application, the available capacity of the battery should be estimated. The capacity estimation of energy storage battery usually needs to collect the health feature (HF) of the voltage and current curve during charge and discharge and combine the results with mathematical algorithms.

The battery generally works in the average charging state, so it is necessary to extract effective health features under the operating conditions of the battery. There are few studies on the capacity estimation of lead–acid batteries in the existing literature. Literature [4–7] uses the coup de fouet (CDF) phenomenon in the initial stage of battery discharge voltage to estimate the SOH. Reference [7] extracts the valley bottom voltage, peak voltage, valley bottom current, peak current, valley bottom temperature, and peak temperature of CDF health features as inputs and establishes battery degradation model through BP neural network, so as to realize the SOH estimation of the battery. However, the CDF phenomenon is relatively difficult to force to occur and requires the battery to be fully charged; and the experiment shows that the CDF will not occur under some average charge voltages. Literature [8] aimed at the research of lithium battery, which selects the duration time of equal discharge voltage drop intervals as HF, which has high estimation accuracy, good adaptability, and provides ideas for the capacity estimation of other types of energy storage batteries. With the increase of battery working time, the battery's internal electrochemical composition changes constantly. Due to the influence of nonlinear factors, such as charge, discharge current, and battery aging, the capacity degradation of battery presents a highly nonlinear characteristic, which makes it necessary to select an appropriate mathematical algorithm to map and model this characteristic. Literature [9] established the equivalent circuit model (ECM) of the lead–acid battery for SOH estimation, but the ECM parameters may not consider the influence of external conditions, such as temperature and average charging voltage; therefore, the adaptability of the ECM is poor, and the establishment process of the model is complex. Literature [10] uses the least squares support vector machine model (LSSVM) to reflect the aging of VRLA, which uses particle swarm optimization algorithm (PSO) to optimize the kernel function of LSSVM. However, the LSSVM algorithm performs poorly in small-scale sample modeling. In addition, the PSO algorithm takes a long time to optimize the hyperparameters and easily falls into the local optimal solution. Reference [11] takes lithium-ion battery as the research object and selects Gaussian process regression (GPR) to model the capacity regeneration phenomenon in the process of battery capacity degradation, which has high precision and small amount of calculation and can better adapt to various working conditions, reflecting the strong nonlinear mapping and small sample modeling ability.

In view of the reasons above, this paper proposes the capacity trajectory prediction of lead-acid battery, based on the steep drop curve of discharge voltage and Gaussian process regression, in order to realize the online capacity estimation of the battery under average charging. This method includes offline modeling and online estimation. In the offline modeling stage, the average charge battery is discharged for a short time, and the voltage steep drop curve for a short time is extracted; then, discharge the battery completely and check the capacity by ampere-hour method. After the discharge, the battery is fully charged, charge-discharge can be repeated many times, in order to accelerate the battery aging and obtain the cycle capacity data. The correlation analysis is carried out on the cycle data of voltage and available capacity, the voltage drops during equal time interval, which has the strongest correlation with the current capacity and is extracted as the health feature (HF); the HF sequence is used as the input and the capacity degradation sequence is used as the output to establish the Gaussian process regression model, reflecting capacity degradation. Rational quadratic covariance function and conjugate gradient algorithm are used to improve GPR model to enhance its nonlinear mapping ability. In the online application stage, the average charging battery is discharged for a short time to extract the corresponding HF, which is substituted into the established GPR model to output the estimation of the present capacity. The experimental results show that the proposed method has a high estimation accuracy and wide adaptability, which can also work well with small-scale training data.

2. The Construction of HF

2.1. Data Source

The lead-acid battery used in this paper was a fixed, valve-regulated lead-acid battery GFMD-200C, produced by Shandong Shengyang power supply Co.Ltd, whose rated capacity is 200 Ah; the even average charging voltage at room temperature (25 °C) is 2.35 V. The experimental installment was produced by Digatron Company and consisted of a battery test system BST-600, battery comprehensive performance tester IBT300-018-2ME, incubator, and transmission line. The BTS-600 is a computer operation software that allows for writing charge and discharge instructions on the computer interface. The performance tester executes the charge and discharge instructions on the battery and transmits the collected data, such as voltage, current, and temperature, to the upper computer for storage and display. The incubator controls the working temperature to be constant. The experimental installation diagram is shown in Figure 1. Among them, C represents the rate, which is the ratio of the current to the rated capacity. Here 0.1 C = 20 Ah, which means it takes 10 h to fully discharge the battery.

Conduct charge-discharge cycle test on the battery to accelerate battery aging. The test steps are shown in Table 1. U_1 in the table is the charging cut-off voltage, that is, the average charging voltage of the battery. Considering that the batteries used in the substation are generally connected in series for average charging, the U_1 for each battery is different but is generally distributed between [2.3 V, 2.4 V]. In order to realize the SOH estimation of the battery, under different average charging voltages, U_1 = 2.3 V, 2.35 V and 2.4 V were selected for research. In addition, the temperature has a significant impact on the electrochemical performance of the battery. Generally, the battery rarely works at low temperatures. Therefore, two working temperatures: T = 25 °C and 50 °C were set. The charge-discharge cycle times under each working condition were 20~30 times, in order to obtain the charge-discharge voltage and current cycle data. Ampere-hour integration was performed on the discharge current, in order to obtain the current available capacity of each cycle. The available capacity degradation trajectory of the battery with the number of cycles is shown in Figure 2.

It can be seen from Figure 2 that the available capacity of the battery presents a complex degradation trajectory with the number of cycles. In the initial cycles, the capacity shows an upward trend. The main reason is that the charging current washes the electrode plate, dissolves the sulfide, and activates the battery [12]. With the increase of using times, the available capacity generally shows a downward trend, which is manifested in the irreversible aging of the battery [13]; there are still complex nonlinear fluctuations locally, which increase the difficulty of SOH estimation.

Table 1. The experimental steps.

	Steps
1	Charge with 0.1 C constant current until the terminal voltage rises to U_1 V.
2	Charge at the constant voltage of U_1 V until the current is less than 0.01 C.
3	Discharge at the constant current of 0.1 C until the terminal voltage drops to 1.8 V.
4	Stand for 30 min.
5	Repeat the steps above until the cycle life test is completed.

In Table 1, the constant current (CC)-constant voltage (CV) charging mode is adopted for battery charging [14]. The voltage in CC phase rises rapidly, most capacity is restored, and the current in CV phase decreases slowly. This method has high charging efficiency and reduces the impact on the battery. After the battery is fully charged, constant current discharge is conducted. The charge-discharge voltage and current curve for one cycle is shown in Figure 3. The discharge voltage curve of the battery, under various working conditions and different cycle times, is shown in Figure 4.
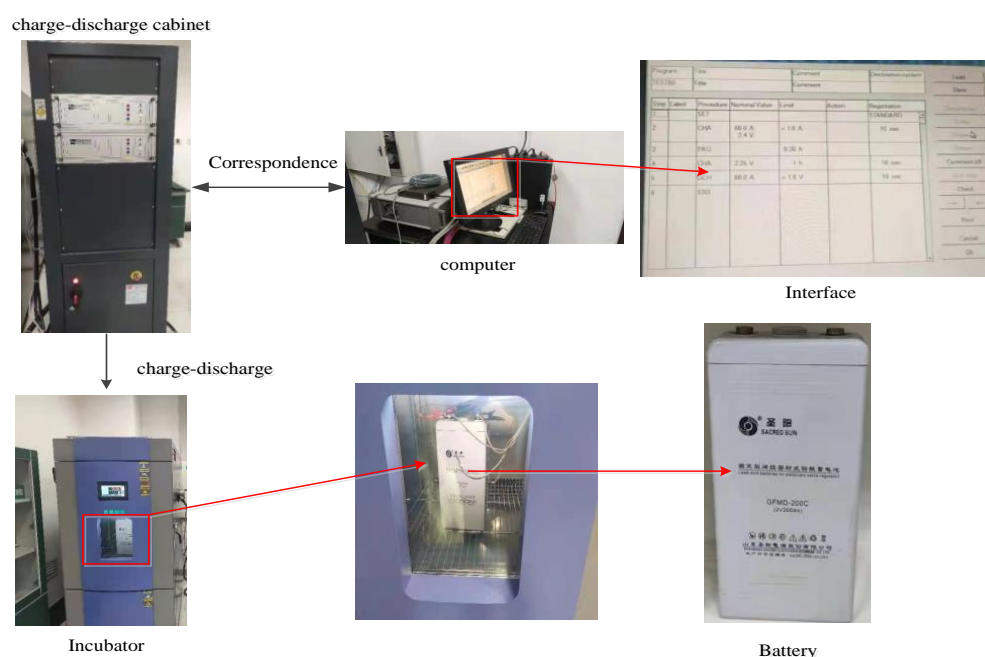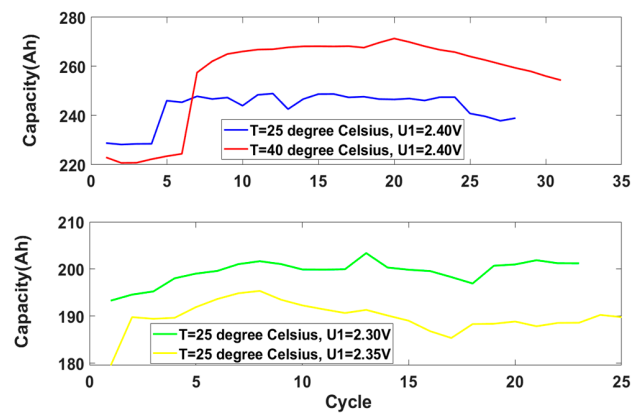

Figure 1. The experimental bench of lead-acid battery.

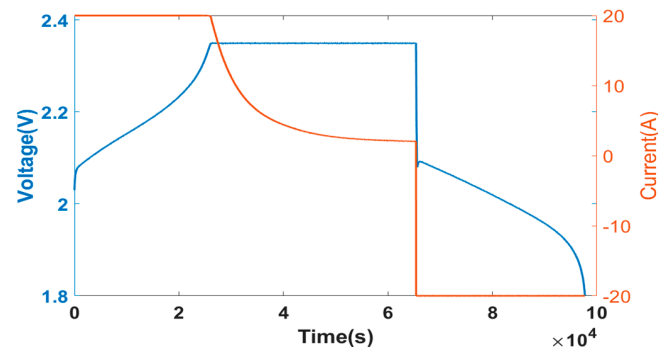**Figure 2.** The capacity degradation trajectory under different working conditions.

Figure 3. Charge-discharge voltage and current curve for one cycle.

(**a**) U_1=2.3V, T=25 °C

(**b**) U_1=2.35V, T=25 °C.

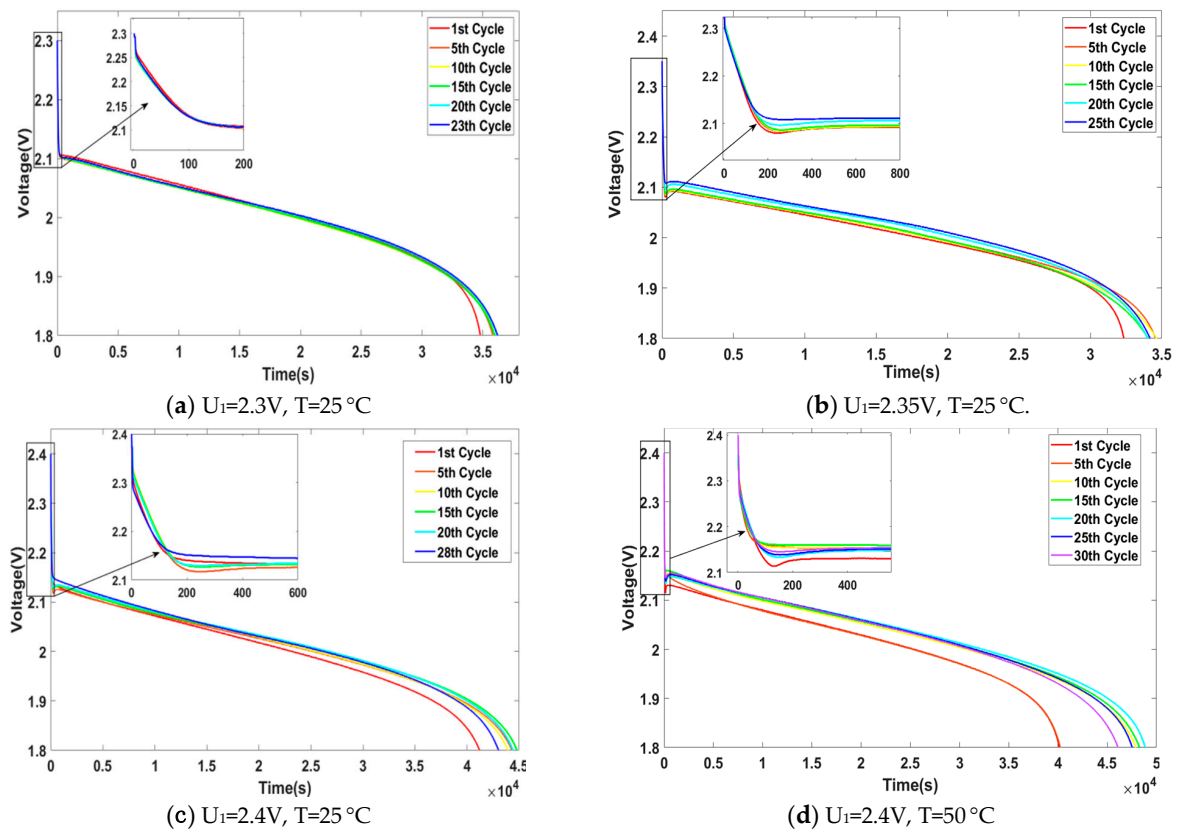(**c**) U_1=2.4V, T=25 °C

(**d**) U_1=2.4V, T=50 °C

Figure 4. The discharge voltage curve of the battery under various working conditions and different cycle times.

2.2. HF Construction

It can be seen from Figure 4 that the discharge voltage curve of the battery varies with the number of cycles under different working conditions, but all of them decrease rapidly at the beginning of discharge, then decrease slowly, and then decrease rapidly to the discharge cut-off voltage. According to reference [15], from the perspective of electrochemistry, at the beginning of discharge, the $PbSO_4$ produced by discharge reaction cannot crystallize immediately; so, the concentration of lead ion in electrolyte increased, causing the internal resistance of the battery to decrease, and the terminal voltage drops. The steep drop voltage data of the battery is easy to obtain by discharging for a short time on the average charging battery, with which the HF can be extracted to reflect the present health status, which can meet the engineering needs for the real-time monitoring of SOH, without affecting the normal operation of the battery. The research in reference [16–18] shows that the interval voltage in the charging and discharging process of energy storage battery has a high correlation with the present capacity. Inspired by this, this paper selects the steep drop voltage DV of short-time discharge during the time interval DT = [T_1, T_2] as HF, which is recorded as DV_DT.

For the given time interval DT = [T_1, T_2], the voltage interval is [$U_1(n)$,$U_2(n)$] during the time interval DT in the *n*th cycle (the initial discharge time is marked 0). The historical capacity degradation sequence is {S_i}, (*n* = 1, 2...N, N means total cycle numbers) and historical voltage drop sequence is {DV_DT(*n*)} ($U_1(n)-U_2(n)$ = DV_DT(*n*)). For different DT = [T_1, T_2], the Pearson relevance coefficient between {S_i} and {DV_DT(*n*)} is different, so the traversal method is used to determine the optimal time interval [T_1, T_2], during which the voltage drop sequence {DV_DT(*n*)} has the highest relevance with the capacity degradation sequence {S_i}. Therefore, we can estimate the capacity that is hard to measure with a health feature that is easy to extract, which embodies a transformation idea from complexity to simplicity. The boundary of traversal method is 10 min, which is in line with the real-time request.

Pearson and gray correlation coefficients (GRC) can be used to quantitatively evaluate the correlation between two variable sequences, as shown in Equations (1) and (2), respectively. In Equation (1), D_i means the HF under *i*th cycle and S_i is the present capacity under *i*th cycle, and *n* is the total cycle number, ρ is the resolution coefficient, generally equals 0.5. Pearson coefficient measures the correlation degree between two variable sequences from the perspective of linear correlation, whose value is between [−1, 1], and the greater the absolute value, the higher the correlation. Grey correlation coefficient (GRC) measures the correlation degree from the perspective of the geometric curve similarity of the two variable sequences, whose value is between [0, 1], and the closer it is to 1, the higher the correlation is.

The traversal method is adopted to search the optimal time interval DT, during which the steep drop voltage DV_DT has the highest relevance with present capacity. Figure 5a,c,e,g shows the pearson coefficient between the DV_DT sequence under different time interval [T_1, T_2] and the capacity degradation sequence, see Equation (1). As we can see from Figure 5a,c,e,g, the relevance is greatly affected by the different time interval. When U_1 = 2.3 V and T = 25 °C, DT = [240, 490]s and the highest relevance is −0.9228, the battery is marked as B1. When U_1 = 2.35 V and T = 25 °C, DT = [170, 600]s and the highest relevance is −0.8946, the battery is marked as B2. When U_1 = 2.4 V and T = 25 °C, DT = [170, 600]s and the highest relevance is −0.9147, the battery is marked as B3. When U_1 = 2.4 V and T = 50 °C, DT = [10, 20]s and the highest relevance is −0.9426, the battery is marked as B4. It can be seen that the longest discharge time is within 10 min, with good real-time performance, which will not affect the normal operation of the battery. Calculate the Pearson correlation degree and GRC of each battery respectively and the results are shown in Table 2.

$$P = \frac{\sum_i (D_i - \overline{D})(S_i - \overline{S})}{\sqrt{\sum_i (D_i - \overline{D})^2}\sqrt{\sum_i (S_i - \overline{S})^2}} \tag{1}$$

$$GRC = \frac{1}{n}\sum_{i=1}^{n}\frac{\min\left|S_i - D_i\right| + \rho\max_{\forall i}\left|S_i - D_i\right|}{\left|S_i - D_i\right| + \rho\max_{\forall i}\left|S_i - D_i\right|} \tag{2}$$

Table 2. Constant current pulse experiment steps.

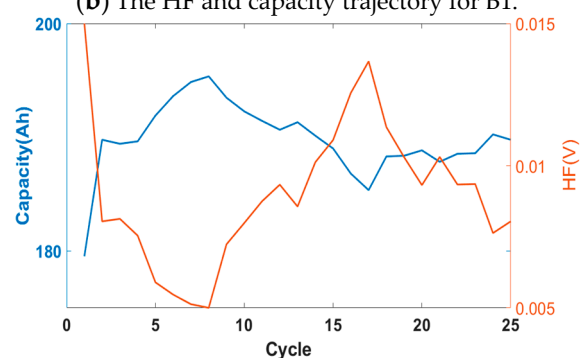| Battery	$[T_1, T_2]$s	Pearson	GRC
B1: T = 25 °C, U_1 = 2.3 V	[240, 490]	−0.9228	0.8334
B2: T = 25 °C, U_1 = 2.35 V	[170, 600]	−0.8946	0.7861
B3: T = 25 °C, U_1 = 2.4 V	[170, 600]	−0.9147	0.8126
B4: T = 50 °C, U_1 = 2.4 V	[10, 20]	−0.9426	0.8590
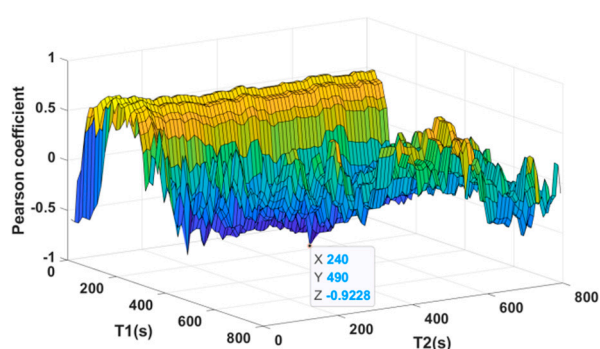
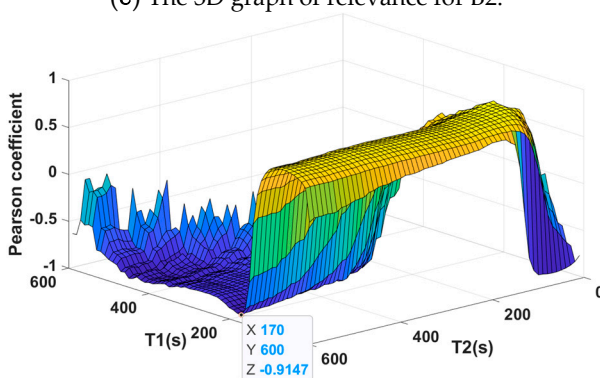(**a**) The 3D graph of relevance for B1.
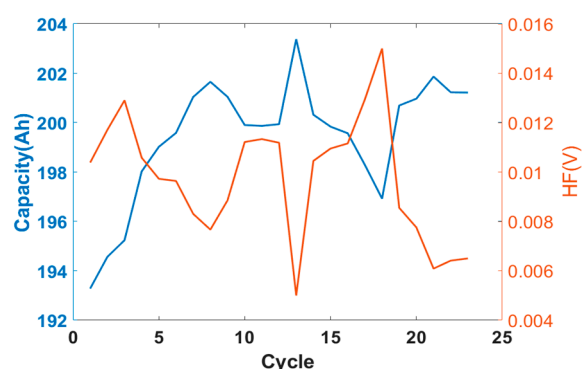
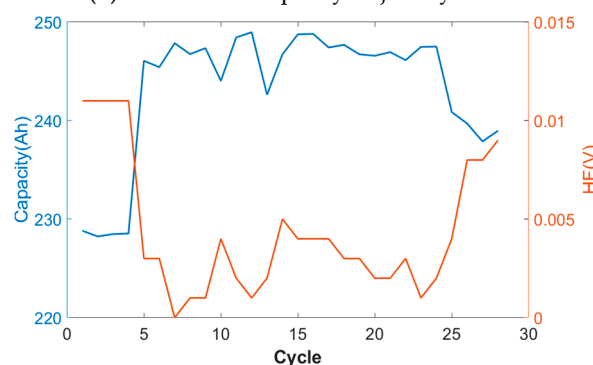(**b**) The HF and capacity trajectory for B1.
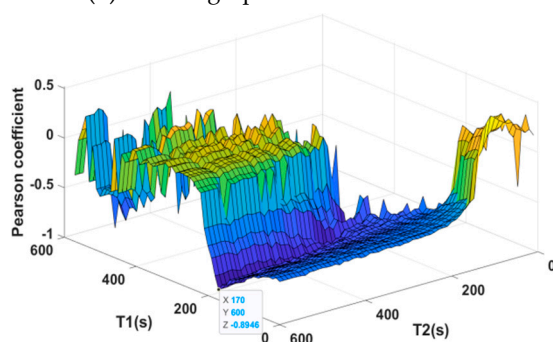
(**c**) The 3D graph of relevance for B2.
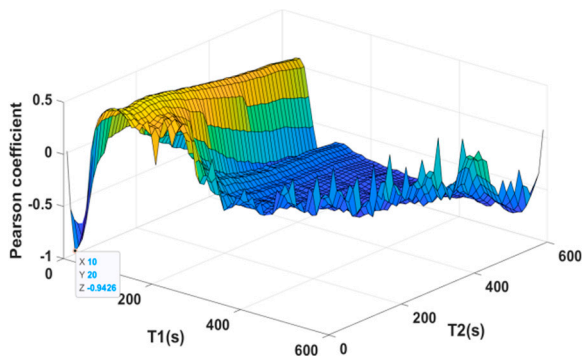

(**d**) The HF and capacity trajectory for B2.
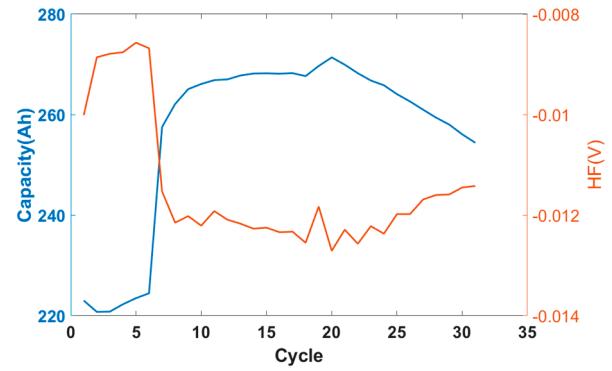

(**e**) The 3D graph of relevance for B3.

(**f**) The HF and capacity trajectory for B3.

Figure 5. *Cont.*

(**g**) The 3D graph of relevance for B4.

(**h**) The HF and capacity trajectory for B4.

Figure 5. The 3D graph of relevance and HF-capacity trajectory for each battery.

It can be seen from Table 2 that the absolute value of the Pearson coefficient of each battery is more than 0.9 and GRC is more than 0.78, indicating the selected health feature DV_DT has strong correlation with SOH, which is suitable for different working conditions of battery and easy to extract for online estimation of present available capacity. Figure 5b,d,f,h shows the variation trend of HF sequence under the optimal time interval and current capacity sequence of each battery with the number of cycles. The HF sequence represented by the orange curve can be extracted online, while the current capacity degradation sequence represented by the blue curve is difficult to obtain. Because the change trends of them are relatively similar, the capacity degradation sequence can be modeled and mapped with the HF sequence, which reflects a transformation idea from complexity to simplicity. The improved Gaussian process regression algorithm is used to model HF and capacity degradation trajectory in the next content.

3. The Capacity Trajectory Prediction, Based on Improved Gaussian Process Regression

3.1. Gaussian Process Regression

Gaussian process regression (GPR) [19] is a flexible and nonparametric model, which can give the point estimation and uncertainty expression of the predicted value at the same time, and has the generalization ability for small sample and strong nonlinear prediction problems.

Given the training dataset S = [X, Y], X is the input dataset and $X = \{x_i\}_{i=1}^{n}, x_i \in R^d$, and $X \in R^{d \times n}$, the corresponding output function is $Y = \{y_k\}_{k=1}^{n}, Y \in R^{n \times 1}$. The regression aims to establish the nonlinear map between training dataset X and Y $\left(f(.) : R^d \mapsto R \right)$, when the new testing input x_* is given, the most promising prediction $f(x_*)$ can be output.

The input matrix $(x_1, x_2, \ldots, x_n)$ and the corresponding output vector $f(X) = \{f(x_1), \ldots, f(x_i)\}$ can form a set of random variables, and any finite dimensional combination in $f(X)$ follows the joint Gaussian distribution, which forms the random process called Gaussian process, which can be described by Equation (3).

$$f(x) \sim GP(m(x), k(x_i, x_j)) \tag{3}$$

In Eqution (3), $m(x) = E(f(x))$ is the mean function of $f(x)$, and GP means the *n*-dimensional Gaussian distribution, and n is the dimension of training set vector.

$k(x_i, x_j) = E[(f(x_i) - m(x_i))(f(x_i) - m(x_j))]$ means covariance function.

In practical application, only the function value containing noise can be obtained, so the Gaussian process regression model shown in Equation (4) is established.

$$y = f(x) + \varepsilon \tag{4}$$

where $y = [y_1, y_2, y_3, \ldots, y_k]^{\mathrm{T}}$ is the output vector, ε is noise vector, which follows the Gauss distribution of 0-means, shown in Equation (5).

$$\varepsilon \sim N\left(0, \sigma_n^2 \boldsymbol{I}\right) \tag{5}$$

Therefore, if $f(\boldsymbol{x})$ is a Gaussian process, the set of finite observations $\boldsymbol{y}$ is a Gaussian process, expressed as Equation (6).

$$\boldsymbol{y} \sim GP\left(m(\boldsymbol{x}), k(\boldsymbol{x}_{\mathrm{i}}, \boldsymbol{x}_{\mathrm{j}}) + \sigma_n^2 \boldsymbol{I}\right) \tag{6}$$

At present, for a new set of testing sample $(\boldsymbol{x}_*, f(\boldsymbol{x}_*))$, $f(\boldsymbol{x}_*)$ and $\boldsymbol{y}$ follows the joint Gaussian distribution, see Equation (7).

$$\boldsymbol{y} \sim GP\left(m(\boldsymbol{x}), k(\boldsymbol{x}_{\mathrm{i}}, \boldsymbol{x}_{\mathrm{j}}) + \sigma_n^2 \boldsymbol{I}\right) \tag{7}$$

According to Bayesian regression method, the prediction value $f(\boldsymbol{x}_*)$ through GPR model follows the Gauss distribution with mean $\overline{f(\boldsymbol{x}_*)}$ and variance $\mathrm{cov} f(\boldsymbol{x}_*)$, shown in Equations (8)–(10), respectively.

$$f(\boldsymbol{x}_*) \mid \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{x}^* \sim GP(\overline{f(\boldsymbol{x}_*)}, \mathrm{cov} f(\boldsymbol{x}_*)) \tag{8}$$

$$\overline{f(\boldsymbol{x}_*)} = k(\boldsymbol{x}, \boldsymbol{x}_*)^{\mathrm{T}} [k(\boldsymbol{x}, \boldsymbol{x}) + \sigma_n^2 \boldsymbol{I}]^{-1} \boldsymbol{y} \tag{9}$$

$$\begin{aligned} \mathrm{cov} f(\boldsymbol{x}_*) = \\ k(\boldsymbol{x}_*, \boldsymbol{x}_*) - k(\boldsymbol{x}, \boldsymbol{x}_*)^{\mathrm{T}} [k(\boldsymbol{x}, \boldsymbol{x}) + \sigma_n^2 \boldsymbol{I}]^{-1} k(\boldsymbol{x}, \boldsymbol{x}_*) \end{aligned} \tag{10}$$

Not only can GPR output the point prediction value for the testing input, but can also give the uncertainty expression of prediction results, which enhances the reliability of prediction results and is determined by $\mathrm{cov} f(\boldsymbol{x}_*)$ in Equation (10), and a 95% confidence interval is shown in Equation (11).

$$\left[\overline{f(\boldsymbol{x}_*)} - 2 \times \sqrt{\mathrm{cov} f(\boldsymbol{x}_*)}, \overline{f(\boldsymbol{x}_*)} + 2 \times \sqrt{\mathrm{cov} f(\boldsymbol{x}_*)} \right] \tag{11}$$

GPR is the common machine learning algorithm, which is widely used in the field of the health prognosis and SOH estimation of power battery, for more details of the mathematical derivation and application of GPR, see reference [20,21].

3.2. GPR Model Optimized by Kernel Function and Conjugate Gradient Algorithm

From the derivation process of GPR, the predicted value can be expressed linearly by the covariance function. Therefore, the key to the regression modeling of Gaussian process lies in the selection of covariance function and the optimization of hyperparameters in the kernel function. The capacity degradation trajectory of the battery presents strong nonlinear, so the rational quadratic covariance function is selected to map the capacity trajectory nonlinearly, as shown in Equation (12).

$$k(\boldsymbol{x}_i, \boldsymbol{x}_j) = \sigma_f^2 \exp\left[1 + \frac{(\boldsymbol{x}_i - \boldsymbol{x}_j)^{\mathrm{T}} M(\boldsymbol{x}_i - \boldsymbol{x}_j)}{2\alpha} \right]^{-\alpha} \tag{12}$$

where $M = diag[\lambda^{-2}]$, hyperparameter vector $\boldsymbol{\theta} = (\lambda, \sigma_f, \alpha)$, where λ means the vector weight, α is the scaling factor and δ_f is variance factor, which can be optimized by maximum likelihood method. Firstly, take the negative logarithm of the marginal likelihood function of the hyperparameter, as shown in Equation (13).

$$\begin{aligned} L(\boldsymbol{\theta}) &= -\log p(\boldsymbol{y} | \boldsymbol{x}, \boldsymbol{\theta}) \\ &= -\tfrac{1}{2} \boldsymbol{y}^{\mathrm{T}} \beta - \tfrac{1}{2} \log \alpha - \tfrac{n}{2} \log(2\pi) \end{aligned} \tag{13}$$

In (13), $\alpha = k(\boldsymbol{x}_i, \boldsymbol{x}_j) + \sigma^2 \boldsymbol{I} \beta = (k(\boldsymbol{x}_i, \boldsymbol{x}_j) + \sigma^2 \boldsymbol{I})^{-1} \boldsymbol{y}$.

In this paper, a conjugate gradient algorithm [22] is proposed to optimize the hyperparameters of GPR model. The conjugate gradient method mainly constructs the gradient direction at the position of the initial parameter point, so that the solution parameters approach step by step along the gradient direction and converge to the minimum point of the objective function quickly.

Calculate the gradient $\Delta L(\boldsymbol{\theta}_k)$ of Equation (13), and the conjugate search direction can be calaculated by Equation (14).

$$d_k = \left\{ \begin{array}{l} -\triangle L(\boldsymbol{\theta}_k) + \beta_{k-1}d_{k-1}\, k > 0 \\ -\triangle L(\boldsymbol{\theta}_k)\, k = 0 \end{array} \right\} \tag{14}$$

where $\beta_{k-1} = \frac{\Delta L(\boldsymbol{\theta}_k)^{\mathrm{T}} \cdot \Delta L(\boldsymbol{\theta}_k)}{\Delta L(\boldsymbol{\theta}_{k-1})^{\mathrm{T}} \cdot \Delta L(\boldsymbol{\theta}_{k-1})}$, and step α_k can be determined by solving the following optimation problem, see Equation (15). We can try some different random α to calculate the corresponding $L(\boldsymbol{\theta}_k + \alpha d_k)$ values, and select α corresponding to the minmum of the $L(\boldsymbol{\theta}_k + \alpha d_k)$ value as α_k.

$$\min_{\alpha \geq 0} L(\boldsymbol{\theta}_k + \alpha d_k) \tag{15}$$

New hyperparameter point can be obtained by Equation (16).

$$\min_{\alpha \geq 0} L(\boldsymbol{\theta}_k + \alpha d_k) \tag{16}$$

Repeat (14)–(16), until Equation (17) is met, the optimal solution $\boldsymbol{\theta}*$ can be acquired.

$$\|\Delta L(\boldsymbol{\theta}_k)\| = \|(\frac{\partial L}{\partial \lambda}, \frac{\partial L}{\partial \sigma_f}, \frac{\partial L}{\partial \alpha})\| \leq 1 \times 10^{-5} \tag{17}$$

$\|\cdot\|$ means the modulus of the vector $(\frac{\partial L}{\partial \lambda}, \frac{\partial L}{\partial \sigma_f}, \frac{\partial L}{\partial \alpha})$, and 1×10^{-5} is the empirical threshold, which can be other smaller numbers. When the norm is less than 1×10^{-5}, we think the gradient vector $\Delta L(\boldsymbol{\theta}^*)$ is small enough, $L(\boldsymbol{\theta}^*)$ reaches minimum value, and the corresponding hyperparameter $\boldsymbol{\theta}^* = (\lambda, \sigma_f, \alpha)$ is optimal.

After the improvement of kernel function and conjugate gradient method, the nonlinear mapping ability of Gaussian process regression algorithm is significantly enhanced.

The flow chart of the lead–acid battery capacity trajectory prediction method, based on the steep drop voltage curve and improved Gaussian process regression, is shown in Figure 6. The method is divided into two parts: offline modeling and online estimation. The offline part carries out correlation analysis on the historical data of battery aging experiment, searches the optimal time interval DT, and trains the Gaussian process regression model with HF under DT and capacity degradation sequence. In the online application, the average charging battery is discharged for a short time, and HF is extracted and substituted into the established GPR model to predict the present available capacity, and the uncertainty expression of the predicted value can be given.
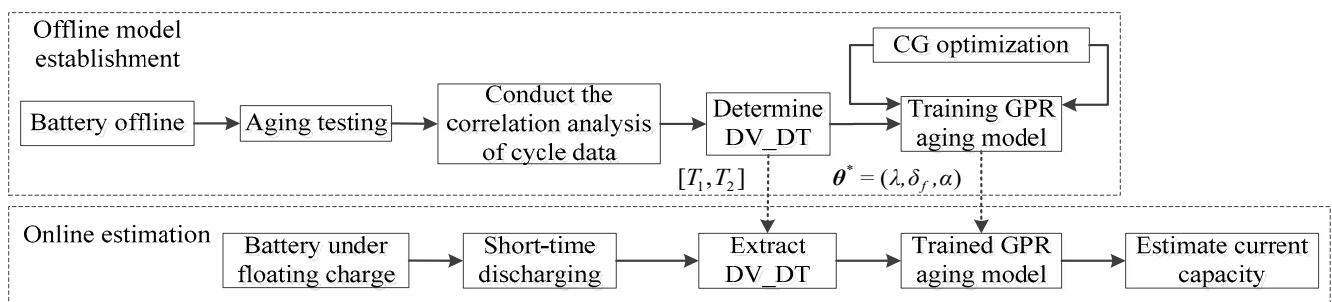

Figure 6. The flow chart of the proposed method.

## 4. Experimental Verification and Analysis

The historical operation data before the prediction starting point (SP) of B1, B2, B3 and B4 are selected as the training set to search the optimal time interval DT to extract the health feature DV_DT and establish GPR model. SP selects 50% of the total number of cycles. The available capacity of each battery after SPth cycle is estimated online. The estimation results and relative error percentage are shown in Figure 7. In this section, relative percentage error (RPE), mean absolute error (MAE) and root mean square error (RMSE) are used for quantitative evaluation. The specific expression is

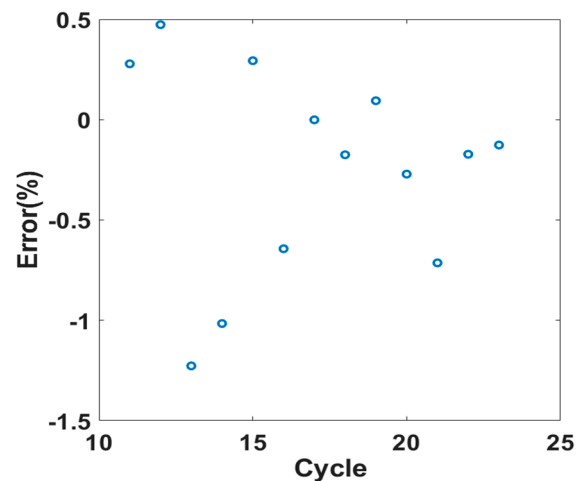$$\text{RPE} = \frac{x_i - \hat{x}_i}{x_i} \times 100\% \tag{18}$$

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|x_i - \hat{x}_i| \tag{19}$$

$$\text{RMSE} = \sqrt{\frac{1}{N}(\sum_{i=1}^{N}(x_i - \hat{x}_i)^2)} \tag{20}$$
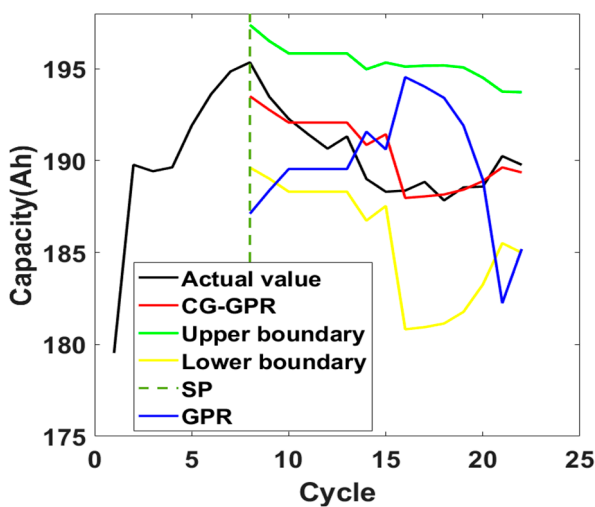
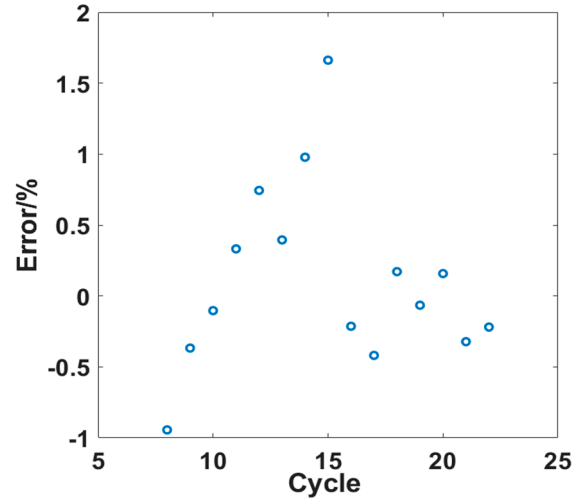where x_i and $\hat{x}_i$ are real and estimated value of capacity under ith cycle, respectively.
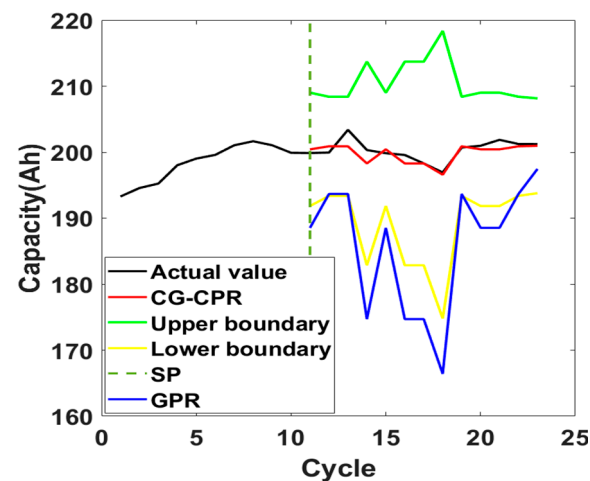
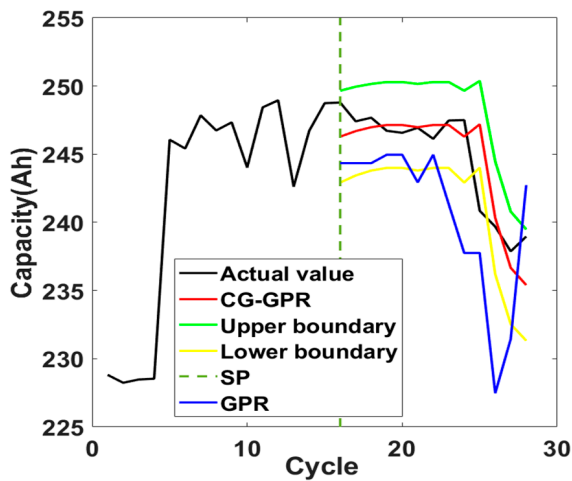(**a**) Estimation results of B1

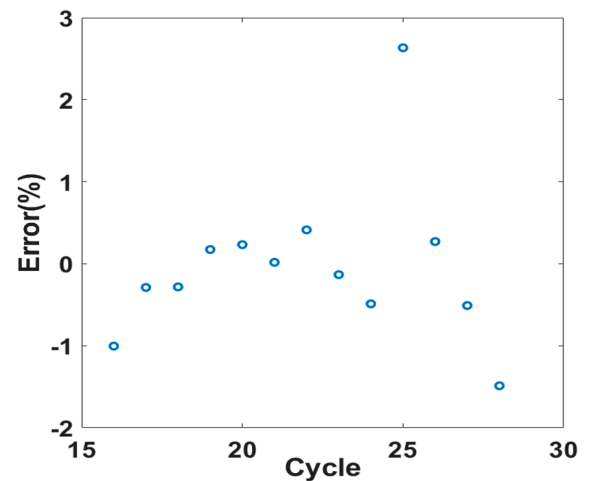(**b**) Estimation errors of B1

(**c**) Estimation results of B2

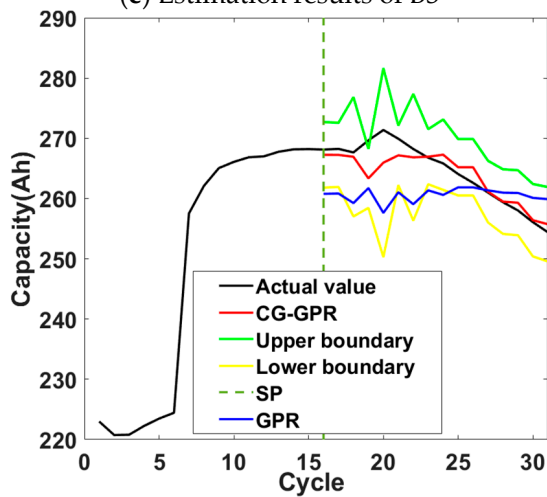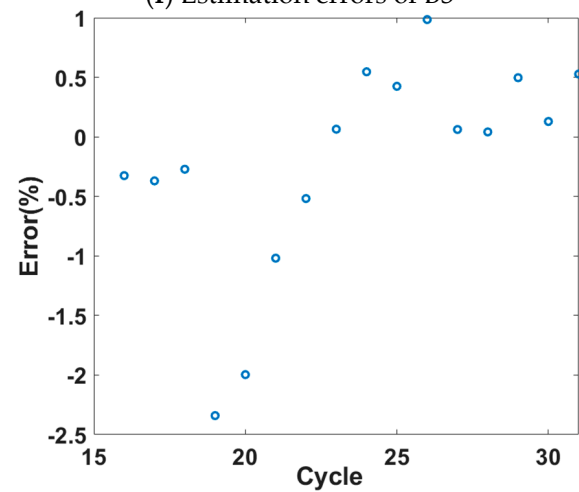(**d**) Estimation errors of B2

Figure 7. *Cont.*

(**e**) Estimation results of B3

(**f**) Estimation errors of B3

(**g**) Estimation results of B4

(**h**) Estimation errors of B4

Figure 7. The estimation results of capacity degradation trajectory and relatively error for each battery.

The prediction results of capacity trajectory for each battery with the improved Gaussian process regression (CG-GPR) model is satisfying, the relative error percentage of most points is within 1%, and the error of a few points is about 2%, and the error of very few points is up to 3%. Table 3 shows the MAE and RMSE calculation results of each battery, the calculation results of B1 and B2 are within 1.5 (Ah), and the error of B3 and B4 is no more than 2.5 (Ah), Quantitative results show that the proposed method has high estimation accuracy and good adaptability. The green and yellow curves respectively represent the 95% upper and lower confidence limits of the prediction results, and the true values of the capacity trajectory are within the confidence interval, indicating that the prediction results have high reliability. The blue curve shows the estimation results of the GPR model without the conjugate gradient method, whose kernel function is linear kernel function, which is far worse than that of the improved Gaussian process regression model. It shows that the strong nonlinearity of the lead–acid battery capacity trajectory puts forward higher requirements for the hyperparameters, and the conventional GPR algorithm cannot effectively fit and map this trend, causing the divergence of prediction results. CG algorithm is used to optimize the kernel function, so as to fully mine the correlation information of training data, thus enhancing the nonlinear fitting ability of GPR algorithm.

Table 3. The error results.

Battery	CG-GPR		GPR	
	MAE	**RMSE**	**MAE**	**RMSE**
B1: T = 25 °C, U_1 = 2.3 V	0.8480	1.1188	14.027	16.201
B2: T = 25 °C, U_1 = 2.35 V	0.9004	1.2023	3.9304	4.5728
B3: T = 25 °C, U_1 = 2.4 V	1.4826	2.2358	4.6756	5.6188
B4: T = 50 °C, U_1 = 2.4 V	1.6884	2.4319	5.6811	6.6806

5. Conclusions

Aiming at the problems of difficulty in health feature extraction and strong nonlinearity of the capacity degradation trajectory of the lead–acid battery, a capacity trajectory prediction method of lead–acid battery, based on drop steep discharge voltage curve and improved Gaussian process regression, is proposed in this paper. Experiments were carried out on batteries with different average charging voltages and temperatures, in order to verify the effectiveness.

In the model establishment stage, the health feature DV_DT, constructed under different working conditions, has a high correlation with the capacity degradation trajectory. The traversal method searches the optimal time interval, in order to maximize the Pearson coefficient of HF and capacity degradation sequence, which has strong operability. The introduction of the conjugate gradient optimization algorithm enhances the nonlinear mapping ability of GPR, so as to adapt to the complex trend of battery capacity degradation trajectory. In the online estimation stage, the corresponding HF can be extracted by a short-time discharge of the average charging battery, without affecting the normal operation of the battery, and the HF is substituted into the established GPR model to output the point estimation and 95% probability estimation of the present available capacity. The estimation accuracy is high, which shows that the proposed method can effectively realize the online estimation of presently available capacity and the SOH of the lead–acid battery.

Future research work can be carried out from the following aspects: (1) the SOH estimation of battery pack; (2) the adaptability of the method remains to be verified under more complex and changeable working conditions.

However, the proposed method also has the following shortcomings. As we can see from the Figure 4, the discharge voltage curves, under different conditions, are quite different; so, the suitable data base for various temperatures and voltages established in offline stage may take a long time. When the offline experiments have been completed and the data set has been established, the online SOH estimation of lead–acid battery would be very convenient and efficient.

Author Contributions: Conceptualization, Q.L. and J.Z.; methodology, J.Z.; software, G.L.; validation, Q.L., Z.S. and C.H.; formal analysis, J.H.; writing—review and editing, Z.C.; funding acquisition, Q.L. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhang, Y.; Ali, A.; Li, J.; Xie, J.; Shen, P.K. Stereotaxically constructed graphene/nano lead composite for enhanced cycling performance of lead-acid batteries. *J. Energy Storage* **2021**, *35*, 102192. [CrossRef]
2. Seguel, J.L.; Seleme, S.I. Robust Digital Control Strategy Based on Fuzzy Logic for a Solar Charger of VRLA Batteries. *Energies* **2021**, *14*, 1001. [CrossRef]

3. Shahriari, M.; Farrokhi, M. Online State-of-Health Estimation of VRLA Batteries Using State of Charge. *IEEE Trans. Ind. Electron.* **2021**, *60*, 191–202. [CrossRef]

4. Phillip, E.; Pascoe, A.H. Standby power system VRLA battery reserve life estimation. *IEEE Trans. Energy Convers.* **2018**, *33*, 6422–6429.

5. Yuan, S.; Cheng, L. Estimation of SOH of battery based on Coup de fouet. *Battery* **2018**, *55*, 65–68.

6. Pascoe, P.E.; Sirisena, H.A.; Anbuky, H. Coup de fouet based VRLA battery capacity estimation. In Proceedings of the First IEEE International Workshop on Electronic Design, Test and Applications Christchurch, Christchurch, New Zealand, 29–31 January 2002; pp. 149–153.

7. Yuanshikui. Research on Online Diagnosis Method of VRLA Battery Performance. Master's Thesis, Dong Nan University, Nanjing, China, 2018.

8. Liu, D.; Zhou, J.; Liao, H. A health indicator extraction and optimization framework for lithium-ion battery degradation modeling and prognostics. *IEEE Trans. Syst. Man Cybern. Syst.* **2015**, *45*, 915–928.

9. Sadabadi, K.K.; Ramesh, P.; Tulpule, P.; Guezennec, Y.; Rizzoni, G. Model-based state of health estimation of a lead-acid battery using step-response and emulated in-situ vehicle data. *J. Energy Storage* **2021**, *36*, 102353. [CrossRef]

10. Zhuang, H.; Xiao, J. VRLA Battery SOH Estimation Based on WCPSO-LVSVM. *Mech. Mater.* **2014**, *628*, 396–400. [CrossRef]

11. Zhou, D.; Zheng, W.; Chen, S. Research on state of health prediction model for lithium batteries based on actual diverse data. *Energy* **2021**, *230*, 120851. [CrossRef]

12. Guo, P.; Cheng, Z.; Yang, L. A data-driven remaining capacity estimation approach for lithium-ion batteries based on charging health feature extraction. *J. Power Sources* **2019**, *412*, 442–450. [CrossRef]

13. Kwiecien, M.; Badeda, J.; Moritz, K.; Huck, K. Determination of SoH of Lead-Acid Batteries by Electrochemical Impedance Spectroscopy. *Appl. Sci.* **2018**, *8*, 873. [CrossRef]

14. Wen, Y.; Wang, L.; Cheng, C. Study on the influence factors of the cycle life of lead-acid battery in DC system. *Iop Conf. Ser. Earth Environ. Sci.* **2021**, *772*, 12039. [CrossRef]

15. Nandanwar, M.; Kumar, S. Charge coup de fouet phenomenon in soluble lead redox flow battery. *Chem. Eng. Sci.* **2016**, *154*, 61–71. [CrossRef]

16. Li, J.F.; Wang, L.X.; Lyu, C. A method of remaining capacity estimation for lithium-ion battery. *Adv. Mech. Eng.* **2013**, *5*, 154831. [CrossRef]

17. Zhou, Y.P.; Huang, M.H.; Chen, Y.P. A novel health indicator for on line lithium-ion batteries remaining useful life prediction. *J. Power Sources* **2016**, *321*, 1–10. [CrossRef]

18. Patil, M.A.; Tagade, P.; Hariharan, K.S. A novel multistage support vector machine based approach for li ion battery remaining useful life estimation. *Appl. Energy* **2015**, *159*, 285–297. [CrossRef]

19. Liu, K.; Hu, X.; Wei, Z. Modified Gaussian Process Regression Models for Cyclic Capacity Prediction of Lithium-Ion Batteries. *IEEE Trans. Transp. Electrif.* **2019**, *5*, 1225–1236. [CrossRef]

20. Fan, L.; Wang, P.; Cheng, Z. A remaining capacity estimation approach of lithium-ion batteries based on partial charging curve and health feature fusion. *J. Energy Storage* **2021**, *43*, 103155. [CrossRef]

21. Zhou, D.; Yin, H.; Xie, W. Research on Online Capacity Estimation of Power Battery Based on EKF-GPR Model. *J. Chem.* **2019**, *2019*, 1–9. [CrossRef]

22. Sakai, H.; Iiduka, H. Sufficient Descent Riemannian Conjugate Gradient Methods. *J. Optim. Theory Appl.* **2021**, *190*, 130–150. [CrossRef]