工业化的兴起导致了对能源的大量需求，但化石能源的枯竭以及环境污染等问题迫使规划人员和决策者寻找替代能源[2,3]，许多可再生能源技术已经得到了广泛的发展。然而，大多数可再生能源，如太阳能和风能，本质上是间歇性的，依靠自然现象来发电，必须储存和按需使用，而智慧城市的发展也需要新的储能技术支持[4-10]。可充电锂离子电池作为一种储能技术，由于拥有更高的能量密度，更小的体积、更长的寿命、更大的容量等优点，作为新能源电动汽车的最佳选择，已经得到了广泛应用[11,12]。但同时锂电池的老化也带来了包括电动汽车的续航减少、动力不足、电池爆炸等问题，如图 1 所示。

如果能在电池老化之前对电池寿命进行预测的话，在避免上述问题的同时还可以为电池生产、使用和优化带来新的机遇[15]。例如，制造商可以加快电池单元开发周期，对电池进行分级，快速验证新的工艺等。同样，终端用户可以估计他们的电池寿命[16-18]。此外，电池预测能够使在电池完全老化之前进行二次回收。总之，对电池当前和未来状态的准确预测将为电池的制造、使用和优化带来巨大的机会【19、20、21】。

目前的电池寿命估计所采用的模型主要可以分为以下三种，等效电路模型(ecm)[22-23]、电化学模型[24-26]或数据驱动模型[27-32]。电化学模型和等效电路模型的准确性和鲁棒性有限。因此这两种模型并不是一个很好的可行解决方案。相反，数据驱动的方法有着一系列的优势，比如不需要了解电池内部的复杂化学反应，分析各种电池降解原理，没有复杂的建立电路的过程等。迄今为止，许多研究都使用机器学习工具来分析电池寿命预测估计。

随着近几年的研究展开，发现了电池数据集中带有噪声是不可避免的，这主要源于充放电过程中的环境干扰，如温度变化、湿度波动的影响。此外，大多公用数据集都是在实验条件下完全充放电测量得到，但实际情况下的电池充放电是不完全的，因此，研究带有噪声的电池数据集才更贴近实际生活，将模型推广到实际应用时才能保证其健壮性。

线性参数估计问题出现在信号处理等广泛的科学学科中[33-34]。它从线性(参数内)模型开始，其中表示可以测量或可以从其他测量推断或可以通过非线性变换计算的过程变量;所有变量都受到测量噪声的影响;包含表征流程变量的基础关系的参数。如[35]和[36]所示，在所有感兴趣的变量都具有参数线性关系且所有测量值都受到噪声污染的情况下，总最小二乘法是参数估计的最佳选择。

但实际情况中电池厂商提供的电池信息数据集来源不同，故由于温度、人为干扰以及传感器等造成的误差大不相同，此时便不能简单的假设数据集的噪声服从同一个分布。此时直接使用 TLS/OLS 并不能很好的建立电池寿命预测模型，故本文在建立线性模型计算电池的寿命时进行改进，对带有不同噪声分布的电池样本进行加权之后，使用 TLS/OLS 进行预测，经循环迭代能够准确的计算出噪声分布的标准差的同时建立适应不同噪声分布的预测模型对电池寿命进行预测，预测结果显示我们的方法比传统的 TLS/OLS 方法更好。

When dataset samples obey noise of different distributions, $y_i = X_i w + \epsilon_i$, we can know that our goal is to minimize the error when the samples obey different noises：

$$\min \sum_i (y_i - X_i w)^2,$$

To solve this problem, we give each sample a different weight:

$$\min \sum_i \alpha_i (y_i - X_i w)^2,$$

Assuming that the noise obeys a Gaussian distribution with zero mean and different variances: $\epsilon_i \sim N(0, \lambda_i^2 \sigma^2)$, then y_i obeys a Gaussian distribution $y_i \sim N(X_i w, \lambda^2 \sigma^2)$. The likelihood function is:

$$l(w) = \prod_i p(y_i|w) = \prod_i \frac{1}{\sqrt{2\pi}\lambda_i \sigma} \exp\left[-\frac{(y_i - X_i w)^2}{2\lambda_i^2 \sigma^2}\right],$$

Maximizing the likelihood function is equivalent to minimizing the objective function $\sum_i \frac{1}{\lambda_i^2}(y_i - X_i w)^2$, and we can get $\alpha_i = \frac{1}{\lambda_i^2}$ by comparison.

EM algorithm can solve the problem with hidden variables well. In this paper, we assume that the standard deviation of noise obeys the distribution as hidden variables, and improve the traditional OLS/TLS algorithm. The specific steps are as follows:

1. Initialize model coefficients w,b

2. Predict the battery life according to the model coefficient, and calculate the error between the predicted value and the real value to update the noise standard deviation of the three data sets.

$$1 \qquad\qquad \min \sum_i \frac{1}{\lambda_i^2}(y_i - X_i w)^2, \qquad\qquad ()$$

3. The samples are weighted according to the formula, and a new round of model coefficients are obtained by using TLS/OLS.

4. Repeat steps 2 and 3 until convergence.

The dataset, referred to as "Dataset", was generated by Severson et al., which consists of 124 commercial LiFePO4/graphite batteries cycled to EOL under fast-charging conditions. During the cycling test of these batteries, several important metrics, such as voltage, current, discharge capacity, temperature, impedance, charge time, etc., are measured in real time. Based on the availability of measurement data and domain expertise, three features in total are extracted for regression modeling, which are indexed by x_1, x_2 and x_3. Note that all these three features are available for Dataset . The feature names, physical meanings and their availabilities are summarized in Table 1. To reduce the nonlinearity of our modeling task, we take the logarithm for both the battery lifetime and the first feature x_1, following the common practice in the literature 错误!未找到引用源。. With these nonlinear transformations, we adopted a linear model template

$$\log(y) \approx w_1 \log(x_1) + w_2 x_2 + w_3 \quad (2)$$

for Dataset 1. To improve numerical stability, we normalize the predicted outcome $\log(y)$ and all features $\{\log(x_1), x_2, x_3, x_4, x_5\}$ so that they have zero mean and unit variance over the training dataset.

Table 1: Features for battery lifetime modeling

Feature Name	Description
x_1	Variance of the difference in the discharge capacity curves as a function of voltage between the 10-th and 100-th cycles
x_2	Slope of the capacity fade curve fitted by a linear function
x_3	Discharge capacity of the 2-nd cycle

数据集由三个不同来源的三个小数据集组成，我们对每个小数据集按照 9:1 比例划分训练集和测试集，再将训练集和测试集分别合并组成最终的训练集和测试集。实验重复 1000

次，每次运行均独立随机生成训练和测试数据集。为每种方法报告 1000 个 RMSE 值的中位数，以便误差度量不会因随机波动而产生强烈偏差。

图 3 显示了实验中不同噪声水平下四种方法 TLS、OLS、改进的 TLS(TLS_EM)和改进的 OLS(OLS_EM)的 RMSE，为了模拟现实情况，我们为三个小数据集添加的噪声水平不一致，但无论是 a,b,c,d 哪种噪声比例模式，都可以看出随着噪声水平的增加，OLS_EM 和 TLS_EM 优势明显，且在放大的图片里可以观察到 TLS_EM 比 OLS_EM 效果更好，结果说明我们提出方法的有效性。

图 4 显示了实验中不同训练集大小四种方法 TLS、OLS、改进的 TLS(TLS_EM)和改进的 OLS(OLS_EM)的 RMSE，为了模拟现实情况，我们为三个小数据集添加的噪声水平不一致，但无论是 a,b,c,d 哪种噪声比例模式，都可以看出，第一，随着训练集比例的增加，四种方法的 RMSE 都在降低，因为更多的训练样本是可用的，并且因此，更多的信息被结合用于模型训练。第二，OLS_EM 和 TLS_EM 的效果始终优于 OLS 和 TLS，且在放大的图片里可以观察到 TLS_EM 比 OLS_EM 效果更好，结果说明我们提出方法的有效性。第三，仅仅在训练集比例非常小（15%）时，TLS_EM 效果差于 OLS_EM，说明在绝大部分情况下 TLS_EM 比 OLS_EM 更准确。

The data set consists of three small data sets from three different sources. We divide each small data set into training set and test set according to the ratio of 9:1, and then merge the training set and test set to form the final training set and test set. The experiment was repeated for 1000 times, and the training and test data sets were generated independently and randomly for each run. Report the median of 1000 RMSE values for each method, so that the error measurement will not be strongly biased due to random fluctuations. Figure 3 shows the RMSE of four methods TLS, OLS, improved TLS(TLS_EM) and improved OLS(OLS_EM) under different noise levels in the experiment. In order to simulate the real situation, the noise levels we added to the three small data sets are different, but no matter which noise ratio mode is A, B, C and D, we can see that with the increase of noise levels, OLS_EM and TLS_EM have obvious advantages. Figure 4 shows the RMSE of four methods TLS, OLS, improved TLS(TLS_EM) and improved OLS(OLS_EM) with different training sets in the experiment. In order to simulate the real situation, the noise levels we added to the three small data sets are different, but no matter which noise proportion mode is A, B, C and D, it can be seen that, firstly, with the increase of the proportion of training sets, the four methods Secondly, the effects of OLS_EM and TLS_EM are always better than OLS and TLS, and it can be observed that TLS_EM is better than OLS_EM in the enlarged picture. The results show the effectiveness of our proposed method. Thirdly, TLS_EM is worse than OLS_EM only when the proportion of training set is very small (15%), which shows that TLS_EM is more accurate than OLS_EM in most cases.

本文考虑了当数据集噪声服从不同分布时的电池寿命预测问题，使用结合了 EM 思想的改进的 OLS 和 TLS 算法对其进行预测，预测结果显示，改进的方法效果显著。

In this paper, the problem of battery life prediction when the noise of data sets obeys different distributions is considered, and the improved OLS and TLS algorithms combined with EM idea are used to predict it. The prediction results show that the improved method is effective.

[1]Renewable energy technologies in Pakistan: Prospects and challenges

[2] Mirza UK, Ahmad N, Majeed T, Harijan K. Wind energy development in Pakistan. Renewable and Sustainable Energy Reviews 2007;11(9): 2179–90.

[3] Tiwari, GN, Ghosal, MK. Renewable Energy Resources: Basic Principles and Applications. Alpha Science Int'l Ltd., 2005. ISBN 1-84265-125-0

[4] B. Obama, Science 2017, DOI: 10.1126/science.aam6284.

[5]Porous Carbon Composites for Next Generation Rechargeable Lithium Batteries

Recent progress of magnetic field application in lithium-based batteries

[6] Raza MQ, Khosravi A. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. Renew Sustain Energy Rev 2015;50:1352e72.

[79] Xu J, Zhang R. CoMP meets smart grid: a new communication and energy cooperation paradigm. IEEE Trans Veh Technol 2013;64(6):2476e88.

[8]An energy matching method for battery electric vehicle and hydrogen fuel cell vehicle based on source energy consumption rate

[9] Wang G, Xu Z, Wen F, et al Traffic-constrained multiobjective planning of electric-vehicle charging stations.IEEE Trans Power Deliv 2013;28(4):2363e72.

[10] Hu J, Zheng L, Jia M, et al Optimization and model validation of operation control strategies for a novel dual-motor coupling-propulsion pure electric vehicle. Energies 2018;11.

[11] P. Poizot, S. Laruelle, S. Grugeon, L. Dupont, J. M. Tarascon, Nature 2000, 407, 496.

[12] B. Dunn, H. Kamath, J. M. Tarascon, Science 2011, 334, 928.

[13]Predicting the State of Charge and Health of Batteries using Data-Driven Machine Learning

[14]MACHINE LEARNING PIPELINE FOR BATTERY STATE OF HEALTH ESTIMATION

15Data-driven prediction of battery cycle life before capacity degradation

16. Peterson, S. B., Apt, J. & Whitacre, J. F. Lithium-ion battery cell degradation resulting from realistic vehicle and vehicle-to-grid utilization. J. Power Sources 195, 2385–2392 (2010).

17. Ramadesigan, V. et al Modeling and simulation of lithium-ion batteries from a systems engineering perspective. J. Electrochem. Soc. 159, R31–R45 (2012).

18. Waag, W., Fleischer, C. & Sauer, D. U. Critical review of the methods for monitoring of lithium-ion batteries in electric and hybrid vehicles. J. Power Sources 258, 321–339 (2014)

19Identifying degradation patterns of lithium ion batteries from impedance spectroscopy using machine learning

20. Severson, K. A. et al Data-driven prediction of battery cycle life before capacity degradation. Nat. Energy 4, 383–391 (2019).

21. Nuhic, A., Terzimehic, T., Soczka-Guth, T., Buchholz, M. & Dietmayer K. Health diagnosis and remaining useful life prognostics of lithium-ion batteries using datadriven methods. J. Power Sources 239, 680–688 (2013).

[22] Tianheng Feng, Lin Yang, Xiaowei Zhao, Huidong Zhang, and Jiaxi Qiang. Online identification of lithium-ion battery parameters based on an improved equivalent-circuit model and its implementation on battery state-of-power prediction. Journal of Power Sources, 281:192–203, 2015.

[23] D Andre, M Meiler, K Steiner, H Walz, T Soczka-Guth, and DU Sauer. Characterization of high-power lithium-ion batteries by electrochemical impedance spectroscopy. ii: Modelling. Journal of Power Sources, 196(12):5349–5356, 2011.

[24] Matthew J Daigle and Chetan Shrikant Kulkarni. Electrochemistry-based battery modeling for prognostics. 2013.

[25] Brian Bole, Chetan S Kulkarni, and Matthew Daigle. Adaptation of an electrochemistry-based li-ion battery model to account for deterioration observed under randomized use. Technical report, SGT, Inc. Moffett Field United States, 2014.

[26] Githin K Prasad and Christopher D Rahn. Model based identification of aging parameters in lithium ion batteries.

Journal of power sources, 232:79–85, 2013.

[27] Kristen A Severson, Peter M Attia, Norman Jin, Nicholas Perkins, Benben Jiang, Zi Yang, Michael H Chen, Muratahan Aykol, Patrick K Herring, Dimitrios Fraggedakis, et al Data-driven prediction of battery cycle life before capacity degradation. Nature Energy, 4(5):383, 2019.

[28] Bhaskar Saha, Kai Goebel, Scott Poll, and Jon Christophersen. Prognostics methods for battery health monitoring using a bayesian framework. IEEE Transactions on instrumentation and measurement, 58(2):291–296, 2008.

[29] Kai Goebel, Bhaskar Saha, Abhinav Saxena, Jose R Celaya, and Jon P Christophersen. Prognostics in battery health management. IEEE instrumentation & measurement magazine, 11(4):33–40, 2008.

[30] Xiaosong Hu, Jiuchun Jiang, Dongpu Cao, and Bo Egardt. Battery health prognosis for electric vehicles using sample entropy and sparse bayesian predictive modeling. IEEE Transactions on Industrial Electronics, 63(4):2645– 2656, 2015.

[31] Verena Klass, Mårten Behm, and Göran Lindbergh. A support vector machine-based state-of-health estimation method for lithium-ion batteries under electric vehicle operation. Journal of Power Sources, 270:262–272, 2014

[32] Peter M Attia, Aditya Grover, Norman Jin, Kristen A Severson, Todor M Markov, Yang-Hung Liao, Michael H Chen, Bryan Cheong, Nicholas Perkins, Zi Yang, et al Closed-loop optimization of fast-charging protocols for batteries with machine learning. Nature, 578(7795):397–402, 2020.

[33]Detection of Abrupt Changes of Total Least Squares Models and Application in Fault Detection

[34] S. Van Huffel, "Tls applications in biomedical signal processing," in Recent Advances in Total Least Squares Techniques and Error-in-Variables Modeling, S. Van Huffel, Ed. Philadelphia, PA: SIAM, 1997

[35] S. Van Huffel and J. Vandewalle, Frontiers in Applied Mathematics: The Total Least Squares Problem—Computational Aspects and Analysis. Philadelphia, PA: SIAM, 1991.

[36] S. Van Huffel, Ed., Recent Advances in Total Least Squares Techniques and Errors-In-Variables Modeling. Philadelphia, PA: SIAM, 1997.

What is the expectation maximization algorithm?

Jing Song, G., Wen Wang, Q. On the weighted least-squares, the ordinary least-squares and the best linear unbiased estimators under a restricted growth

curve model. *Stat Papers* **55**, 375–392 (2014). https://doi.org/10.1007/s00362-012-0483-9

B. De Moor and J. Vandewalle, "A unifying theorem for linear and total linear least squares," in IEEE Transactions on Automatic Control, vol. 35, no. 5, pp. 563-566, May 1990, doi: 10.1109/9.53523.

Energy is the foundation of all science and engineering technology, without which the human world will be difficult to operate [1]. The rise of industrialization has led to a large demand for energy, but the depletion of fossil energy and environmental pollution have forced planners and decision makers to look for alternative energy [2,3], and many renewable energy technologies have been widely developed. However, most renewable energy sources, such as solar energy and wind energy, are intermittent in nature and rely on natural phenomena to generate electricity, so they must be stored and used on demand [4-5]. At the same time, smart cities will become the future model of people [6,7], and electric vehicles can solve the energy-saving development and environmental pollution problems of smart cities

well, and the development of new energy electric vehicles has become a global consensus [8,10]. As an energy storage technology, rechargeable lithium-ion batteries have been widely used as the best choice for new energy electric vehicles because of their higher energy density, smaller volume, longer life and larger capacity [11,12]. However, lithium batteries will deteriorate with time, which is manifested in the loss of battery capacity and the increase of impedance [13]. Therefore, the rechargeable lithium-ion battery can not only promote the development of electric vehicles, but also inevitably produce a series of problems, such as reduced battery life and insufficient power, etc. As time goes by, the aging of lithium-ion batteries may cause safety accidents, as shown in Figure 1.
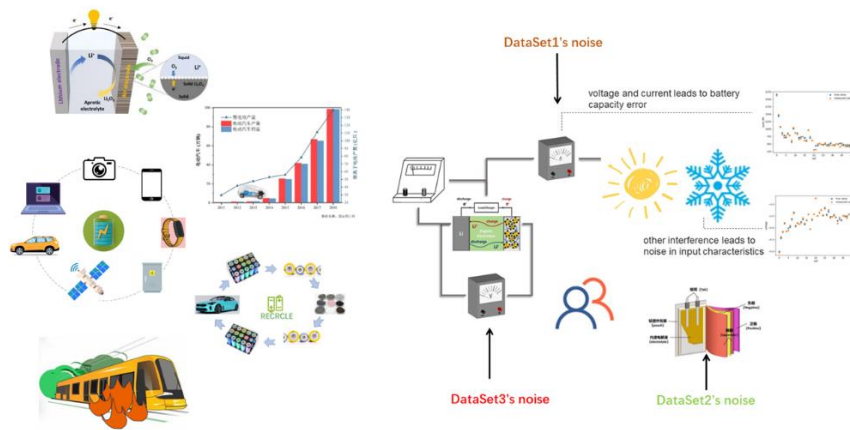

Figure1: Lithium battery applications and hidden dangers
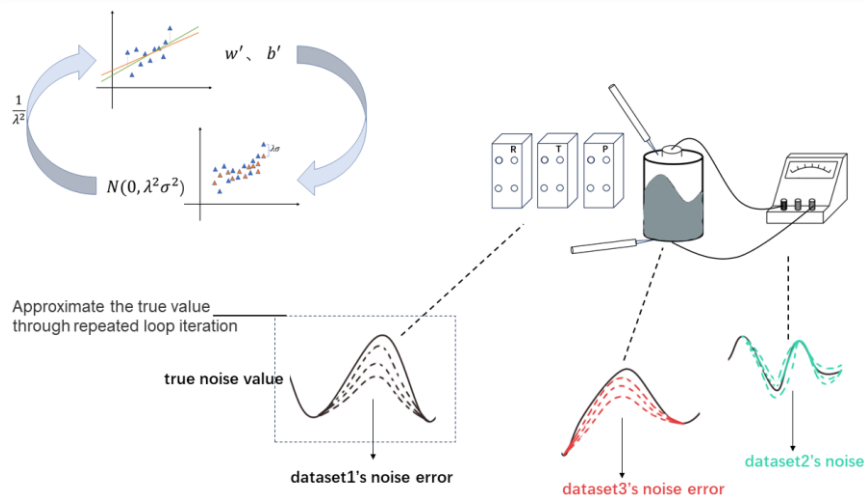

Figure 2: Improved TLS/OLS algorithm

The degradation rate of batteries is affected by dynamic operating conditions. If the battery life

can be predicted before aging, it will bring new opportunities for battery production, use and

optimization [15]. For example, manufacturers can speed up the development cycle of battery cells, grade batteries, and quickly verify new processes. Similarly, end users can estimate their battery life [16-18]. In addition, battery prediction enables secondary recovery before the battery is completely aged. In a word, the accurate prediction of the current and future state of the battery will bring great opportunities for the manufacture, use and optimization of the battery [19, 20 and 21]. At present, the models used in battery life estimation can be mainly divided into the following three types: equivalent circuit model (ecm)[22-23], electrochemical model [24-26] or data-driven model [27-32]. The accuracy and robustness of electrochemical model and equivalent circuit model are limited. Therefore, these two models are not a good and feasible solution. On the contrary, data-driven method has a series of advantages, such as no need to understand the complex chemical reactions inside the battery, analysis of various battery degradation principles, no complicated process of establishing circuits, etc. So far, many studies have used machine learning tools to analyze battery life prediction and estimation.

With the development of research in recent years, it is found that noise in battery data set is inevitable, which mainly comes from environmental interference during charging and discharging, such as temperature change and humidity fluctuation. In addition, most public data sets are measured under experimental conditions, but the actual battery charge and discharge is incomplete. Therefore, it is closer to real life to study the battery data set with noise, and its robustness can be guaranteed when the model is extended to practical application. The problem of linear parameter estimation appears in a wide range of scientific disciplines such as signal processing [33-34]. It starts with a linear (in-parameter) model, which represents process variables that can be measured or inferred from other measurements or calculated by nonlinear transformation; All variables are affected by measurement noise; Contains parameters that represent the basic relationship of process variables. As shown in [35] and [36], the total least square method is the best choice for parameter estimation when all the variables of interest have parameter linear relations and all the measured values are polluted by noise.

However, in the actual situation, the data sets of battery information provided by battery manufacturers come from different sources, so the errors caused by temperature, human interference and sensors are very different. At this time, it is impossible to simply assume that the noise of the data sets obeys the same distribution. At this time, directly using TLS/OLS can't establish a battery life prediction model. Therefore, this paper improves the linear model to calculate the battery life. After weighting the battery samples with different noise distributions, TLS/OLS is used to predict the battery life. Through cyclic iteration, the standard deviation of noise distribution can be accurately calculated, and a prediction model suitable for different noise distributions can be established to predict the battery life. The prediction results show that our method is better than the traditional TLS/OLS method.

When dataset samples obey noise of different distributions, $y_i = X_i w + \epsilon_i$, we can know that our goal is to minimize the error when the samples obey different noises：

$$\min \sum_i (y_i - X_i w)^2 \qquad (1)$$

To solve this problem, we give each sample a different weight:

$$\min \sum_i \alpha_i (y_i - X_i w)^2 \qquad (2)$$

Assuming that the noise obeys a Gaussian distribution with zero mean and different variances: $\epsilon_i \sim N(0, \lambda_i^2 \sigma^2)$, then y_i obeys a Gaussian distribution $y_i \sim N(X_i w, \lambda^2 \sigma^2)$. The likelihood function is:

$$l(w) = \prod_i p(y_i|w)$$
$$= \prod_i \frac{1}{\sqrt{2\pi}\lambda_i\sigma} \exp\left[-\frac{(y_i - X_iw)^2}{2\lambda_i^2\sigma^2}\right] \quad (3)$$

Maximizing the likelihood function is

EM algorithm can solve the problem with hidden variables well. In this paper, we assume that the standard deviation of noise obeys the distribution as hidden variables, and improve the traditional OLS/TLS algorithm, The algorithm flow is shown in Figure 2. The specific steps are as follows:

1. Initialize model coefficients w,b

2. Predict the battery life according to the model coefficient, and calculate the error between the predicted value and the real value to update the noise standard deviation of the three data sets.

$$\min \sum_i \frac{1}{\lambda_i^2}(y_i - X_iw)^2 \quad (4)$$

3. The samples are weighted according to the formula, and a new round of model coefficients are obtained by using TLS/OLS.

4. Repeat steps 2 and 3 until convergence.

The dataset, referred to as "Dataset", was generated by Severson et al. 错误!未找到引用源。, which consists of 124 commercial LiFePO4/graphite batteries cycled to EOL under fast-charging conditions. During the cycling test of these batteries, several important metrics, such as voltage, current, discharge capacity, temperature, impedance, charge time, etc., are measured in real time. Based on the availability of measurement data and domain expertise, three features in total are extracted for regression modeling, which are indexed by x_1, x_2 and x_3.

Note that all these three features are available for Dataset . The feature names, physical meanings and their availabilities are summarized in Table 1. To reduce the nonlinearity of our modeling task, we take the logarithm for both the battery lifetime and the first feature x_1, following the common practice in the literature 错误!未找到引用源。. With these nonlinear transformations, we adopted

equivalent to minimizing the objective function $\sum_i \frac{1}{\lambda_i^2}(y_i - X_iw)^2$, and we can get $\alpha_i = \frac{1}{\lambda_i^2}$ by comparison.

a linear model template

$$\log(y) \approx w_1\log(x_1) + w_2 x_2 + w_3 \quad (5)$$

for Dataset .To improve numerical stability, we normalize the predicted outcome $\log(y)$ and all features $\{\log(x_1), x_2, x_3, x_4, x_5\}$ so that they have zero mean and unit variance over the training dataset.

Table 1: Features for battery lifetime modeling

Feature Name	Description
x_1	Variance of the difference in the discharge capacity curves as a function of voltage between the 10-th and 100-th cycles
x_2	Slope of the capacity fade curve fitted by a linear function
x_3	Discharge capacity of the 2-nd cycle

The data set consists of three small data sets from three different sources. We divide each small data set into training set and test set according to the ratio of 9:1, and then merge the training set and test set to form the final training set and test set. The experiment was repeated for 1000 times, and the training and test data sets were generated independently and randomly for each run. Report the median of 1000 RMSE values for each method, so that the error measurement will not be strongly biased due to random fluctuations.
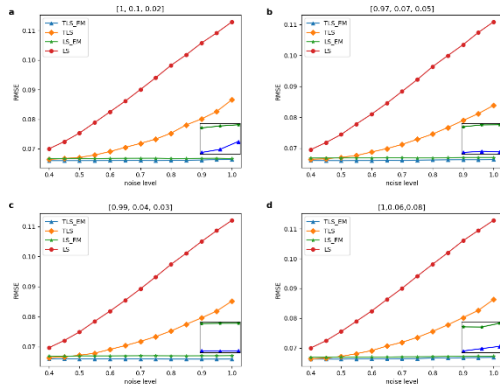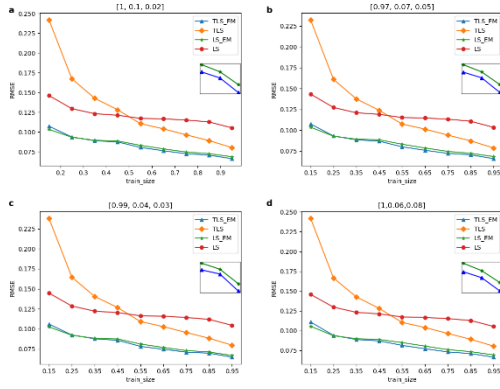
Figure 3: RMSE with increased noise level

Figure 4: RMSE with increased training set proportion

Figure 3 shows the RMSE of four methods TLS, OLS, improved TLS(TLS_EM) and improved OLS(OLS_EM) under different noise levels in the experiment. In order to simulate the

real situation, the noise levels we added to the three small data sets are different, but no matter which noise ratio mode is A, B, C and D, we can see that with the increase of noise levels, OLS_EM and TLS_EM have obvious advantages.

Figure 4 shows the RMSE of four methods TLS, OLS, improved TLS(TLS_EM) and improved OLS(OLS_EM) with different training sets in the experiment. In order to simulate the real situation, the noise levels we added to the three small data sets are different, but no matter which noise proportion mode is A, B, C and D, it can be seen that, firstly, with the increase of the proportion of training sets, the four methods Secondly, the effects of OLS_EM and TLS_EM are always better than OLS and TLS, and it can be observed that TLS_EM is better than OLS_EM in the enlarged picture. The results show the effectiveness of our proposed method. Thirdly, TLS_EM is worse than OLS_EM only when the proportion of training set is very small (15%), which shows that TLS_EM is more accurate than OLS_EM in most cases.

In this paper, the problem of battery life prediction when the noise of data sets obeys different distributions is considered, and the improved OLS and TLS algorithms combined with EM idea are used to predict it. The prediction results show that the improved method is effective.

[1] M.A. Chaudhry, et al. ISO4, 13(6), 1657-1662 (2009).
[2] Mirza UK, Ahmad N, et al. ISO4,11(9),2179–90(2007).
[3] G.N. Tiwari, M.K. Ghosal, Renewable Energy Resources: Basic Principles and Applications, 2005.
[4] Barack Obama.*Science***355**,126-129(2017).
[5] LIU H, LIU X, et al. Advanced Energy Materials, 2017, 7(24): 1700283.
[6] P. Poizot, S. Laruelle, et al. Nature 2000, 407, 496.
[7] B. Dunn, H. Kamath, et al. Science 2011, 334, 928.
[8] Raza MQ, Khosravi A. Renew Sustain Energy Rev 2015;50:1352e72.
[9] Xu J, Zhang R. IEEE Trans Veh Technol 2013;64(6):2476e88.
[10] H. Xiong, H. Liu, et al. International Journal of Hydrogen Energy, vol. 44, no. 56, pp. 29733-29742, 2019.
[11] Wang G, Xu Z, et al. IEEE Trans Power Deliv 2013;28(4):2363e72.
[12] Hu J, Zheng L, et al . Energies 2018;11.
[13] Ng, MF., Zhao, J., Yan, Q. et al. Nat Mach Intell 2, 161–170 (2020).
[25] Brian Bole Moffett Field United States, 2014.

[14] Roman, D., Saxena, S., Robu, V. et al. Nat Mach Intell 3, 447–456 (2021). [15] Severson, K.A., Attia, P.M., Jin, N. et al. Nat Energy 4, 383–391 (2019).
16. Peterson, S. B., et al. J. Power Sources 195, 2385–2392 (2010).
17. Ramadesigan, V. et al .J. Electrochem. Soc. 159, R31–R45 (2012).
18. Waag, W., Fleischer, C. & Sauer, D. U. J. Power Sources 258, 321–339 (2014)
19 Zhang, Y., Tang, Q., Zhang, Y. et al. Nat Commun 11, 1706 (2020).
20. Severson, K. A. et al .Nat. Energy 4, 383–391 (2019).
21. Nuhic, A., Terzimehic, T., et al. J. Power Sources 239, 680–688 (2013).
[22] Tianheng Feng, Lin Yang, et al. Journal of Power Sources, 281:192–203, 2015.
[23] D Andre, M Meiler,et al. Journal of Power Sources, 196(12):5349–5356, 2011.
[24] Daigle,et al. "Electrochemistry-based Battery Modeling for Prognostics." (2013).

[26] Githin K Prasad Journal of power sources, 232:79–85, 2013.

[27] Kristen A Severson Nature Energy, 4(5):383, 2019.

[28] Bhaskar Saha IEEE Transactions on instrumentation and measurement, 58(2):291–296, 2008.

[29] Kai Goebel, IEEE instrumentation & measurement magazine, 11(4):33–40, 2008.

[30] Xiaosong Hu, IEEE Transactions on Industrial Electronics, 63(4):2645– 2656, 2015.

[31] Verena Klass, Journal of Power Sources, 270:262–272, 2014

[32] Peter M Attia, et al. Nature, 578(7795):397–402, 2020.

[33]Detection of Abrupt Changes of Total Least Squares Models and Application in Fault Detection

[34] S. Van Huffel, Philadelphia, PA: SIAM, 1997

[35] S. Van Huffel, Frontiers in Applied Mathematics. PA: SIAM, 1991.

[36] S. Van Huffel, Ed., Philadelphia, PA: SIAM, 1997.