锂电池作为新能源在航空航天、汽车能源领域被广泛应用的同时，也因为使用和维护不当带来许多安全事故。对锂电池进行寿命预测研究能够提前做好电池替换、维护、二次利用的准备且大大减少安全事故，是当今研究热点。然而现实中收集到的电池循环信息大多都含有噪声，噪声对电池寿命的预测精度造成了不可忽略的影响。本文考虑了当输入特征中含有噪声的情况下，使用完全自适应噪声集合经验模态分解方法对从电池循环早期数据中提取的的健康因子进行降噪处理，使用降噪处理之后的数据预测精度比未降噪的数据预测精度有明显提升。本文不仅仅从电池早期循环数据中提取与电池寿命相关的健康因子同时处理了数据噪声，提升了电池寿命的预测精度，对锂电池寿命预测研究有着重大意义。

引言：锂电池作为新能源的一种，在化石能源迅速消耗的今天作为其替代，不仅能解决能源消耗殆尽，还能保护环境。锂电池因为其体积小、续航时间长、 成本低、使用寿命时间长等原因，被广泛使用在汽车、航空航天等领域。

然而锂电池的广泛使用也带来了一些新的问题，锂电池的不当使用造成电池加速老化，存在较大的安全隐患。随时电池的多次循环使用其容量减少和内阻增加导致性能恶化，容易出现内短路、热失控等问题，导致设备故障甚至带来灾难性事故。锂电池寿命预测是电池管理系统中的一个重要组成部分，它可以通过预测电池寿命来确定何时需要更换电池或对电池进行维护。虽然经过多年的优化改造，但对于锂电池寿命的预测仍然存在不足，是当今研究的热点。

电池寿命预测可以帮助用户提前了解电池的寿命状况，避免电池寿命到期后出现电池失效或存在安全隐患的情况。能提高电池系统的可靠性和安全性；同时通过对电池寿命进行预测，可以提前发现电池故障，及时对电池进行维护或更换，从而降低电池系统的维护成本；通过对电池寿命进行预测，可以优化电池的充电和放电策略，从而提高电池系统的续航里程和电池寿命；锂电池寿命预测技术对于新能源汽车的发展至关重要，它可以提前预测电池寿命，让用户更加放心地使用新能源汽车，同时也可以降低新能源汽车的维护成本，提高其使用寿命。

关于锂电池寿命预测的研究已经取得了很多进展，但现有的锂电池寿命的研究体系主要集中在两个方面，基于经验的方法和基于性能的方法。

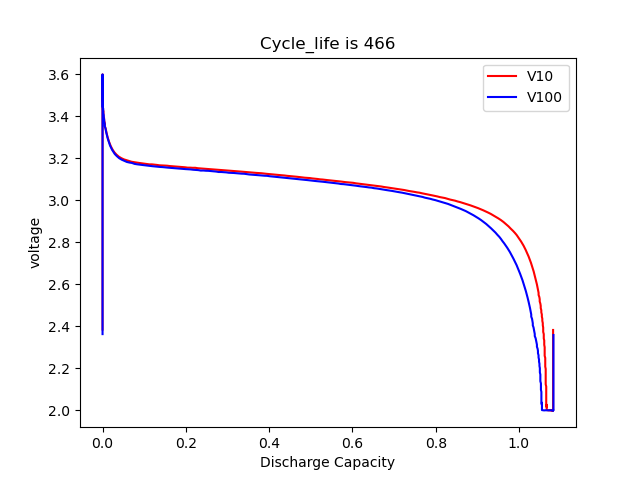
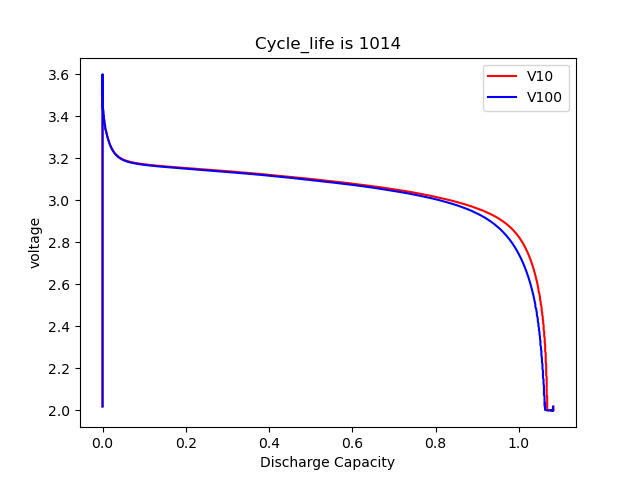
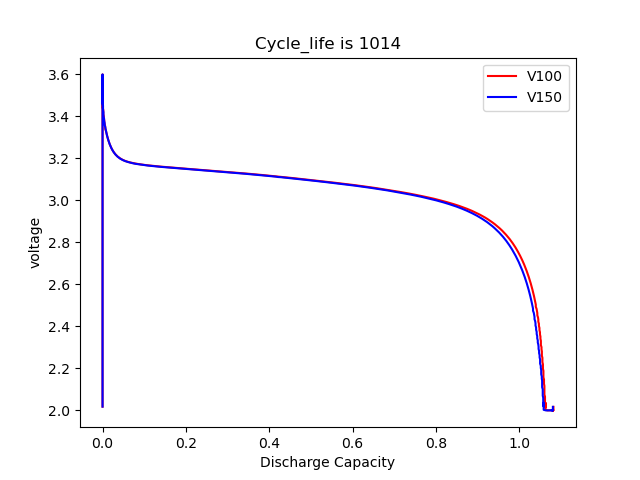
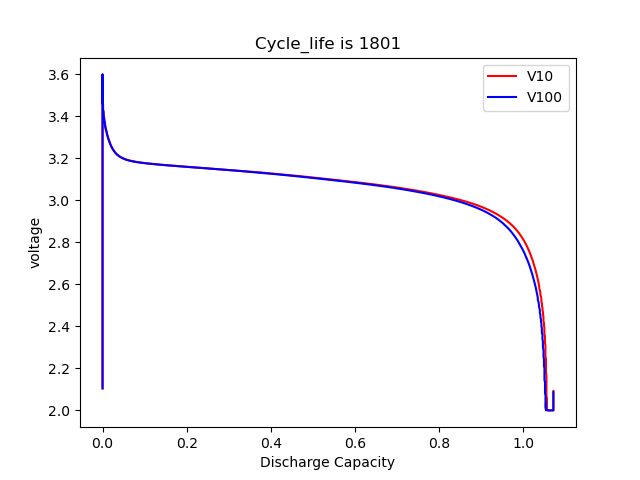
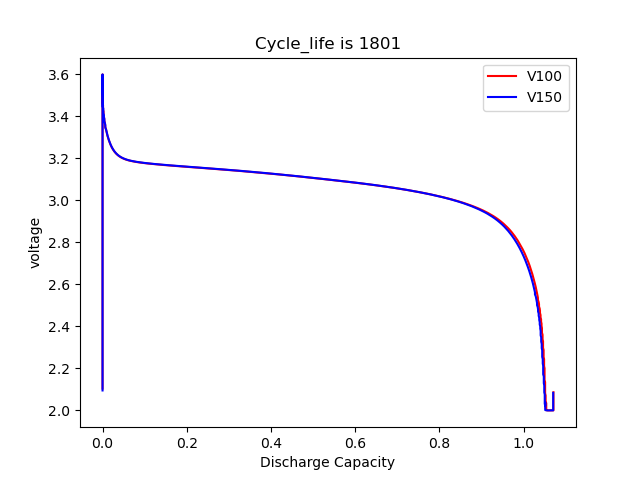
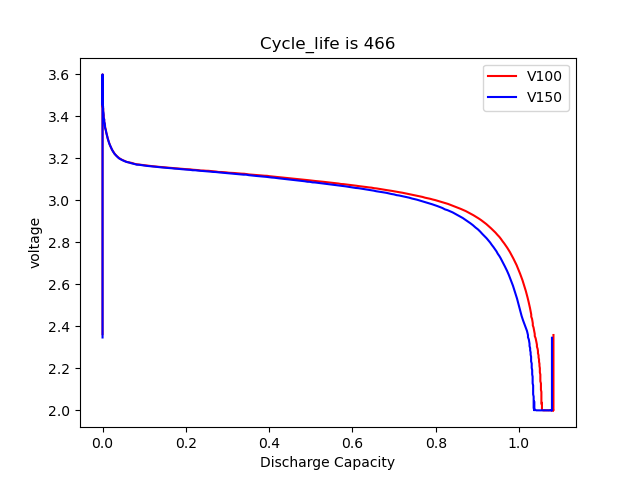
基于经验的方法：循环周期数法、安时法和加权安时法、面向事件的老化累积方法。但基于经验的方法对数据的依赖性太强，以至于换一个类型的电池，经验模型效果不好，鲁棒性差。

基于性能的方法：基于模型的方法，基于数据驱动的方法，基于融合模型的方法。基于模型的方法大多建立在电池电化学机理的研究基础上，通过分析电池的退化现象或构建等效 电路而建立数学模型，常用的模型有基于 Wiener 过 程的模型［6，7］，Thevenin 模型［8，9］，自回归模型［10］，单粒子 (single particle, SP) 模型， 伪二维 (pseudo-two-dimensions, P2D) 模型等。其中基于数据驱动的方法不需要了解锂电池内部的物理化学反应，仅仅根据电池的历史循环数据就能很好的预测锂电池的寿命，成为当下研究热点。举例，支持向量机［14］Deng Y W，Ying H J，E J Q，et al． Feature parameter extraction and intelligent estimation of the State-ofHealth of lithium-ion batteries［J］． Energy，2019，176: 91 － 102．，神经网络［15 ～ 17］［15］ 吴忠强，尚梦瑶，申丹丹，等． 基于 BSA-ＲELM 的纯 电动汽 车 锂 离 子 电 池 SOC 估 计［J］． 计 量 学 报和相关向量机［18，19］( relevance vector machine，ＲVM)［18］ Song Y C，Liu D T，Hou Y D，et al． Satellite lithiumion battery remaining useful life estimation with an iterative updated ＲVM fused with the KF algorithm［J］． Chinese Journal of Aeronautics，2018，31( 1) : 31 － 40．［19］ Yang W A，Xiao M H，Zhou W，et al． A hybrid prognostic approach for remaining useful life prediction of Lithium-Ion batteries ［J］． Shock and Vibration， 2016，2016: 1 － 15等都用于预测锂电池的寿命。其中关于使用神经网络的方法主要是集中于直接使用神经网络进行预测还有就是使用神经网络的时序模型，不需要电池的其他信息，只需要知道电池寿命的历史放电容量就能对电池有一个比较准确的预测。文献［37］考虑了数据少时 存在预测精度低问题，对长短期记忆神经网络进行了改进，将该方法与 BP 和传统长短期记忆神经网络进行了对比分析，发现改进的长短期记忆神经网络预测结果误差最小; ［37］ 李瑞津，刘斌，张学敏，等． 基于改进 LSTM 的变电站铅酸电池寿 命 预 测［J］． 电 池，2020，50 ( 6 ) : 560-564．基于融合的方法：主要分为模型与数据驱动的方法和多种数据驱动方法的结合，包括主要涉及电池充放电截止 电压和标称电压，利用径向基神经网络修正无迹卡尔曼的状态估计，从而预测电池寿命的研究。［51］梅枭央． 基于融合算法的锂离子电池剩余使用寿 命预测［D］． 武汉: 华中科技大学，2019．在基于多模型融合和无迹粒 子滤波的电池寿命预测［52］作者使用电压、电流、温度 和 EIS 数据等，考虑了不同种电池衰减情况对电池寿命进行预测。［52］ 王帅． 数据驱动的锂离子电池剩余寿命预测方法 研究［D］． 哈尔滨: 哈尔滨工业大学，2017．

现实中的数据大多含有噪声，使用带有噪声的电池循环数据对电池寿命进行预测精度较差，本文将提取电池循环的前150周期信息作为电池寿命健康因子HI，使用了完全自适应噪声集合经验模态分解降噪方法进行降噪，获得新的输入特征。结果表明，使用了本文提取的HI并进行降噪之后的输入特征进行预测有着更低的均方根误差。本文可以仅使用电池早期循环数据就能有着良好的预测精度，不需要了解电池内部的物理化学原理仅使用电池早期的放电信息就能对电池寿命进行预测得到较高的精度。在实际中获取更多的电池循环信息，故能使用早期的循环数据进行预测对于锂电池寿命预测的实际应用有着重大意义。

数据集介绍：由124个在快速充电条件下循环失效的商用锂离子电池组成。这些锂离子磷酸盐（LFP）/石墨电池由A123系统公司（APR18650M1A）制造，在设置为48°C的强制对流温度室中的30通道Arbin LBT恒电位仪上的水平圆柱形夹具中循环。 电池的标称容量为1.1 Ah，标称电压为3.3 V。这项工作的目标是优化锂离子电池的快速充电。因此，此数据集中的所有单元格都使用一步或两步快速充电策略收费。此策略的格式为“C1（Q1）-C2”，其中 C1 和 C2 分别是第一和第二个恒流步骤，Q1 是电流切换的充电状态 （SOC， %）。第二个电流步骤以80%SOC结束，之后电池以1C CC-CV充电。上限和下限截止电位分别为3.6 V和2.0 V，符合制造商的规格。这些截止电位对于所有电流步骤都是固定的，包括快速充电;经过一些循环后，电池可能会在快速充电期间达到上限截止电位，从而导致显着的恒压充电。所有细胞在4C下放电。本文将三个批次的电池全部提取，并按照9:1划分训练测试集。所有电池均采用一步或两步充电策略进行循环。充电时间从~8到13.3分钟（0-80%SOC）不等。每个策略通常测试两个细胞，除了3.6C（80%）。在充电期间和放电后分别达到1%SOC后放置1分钟和80秒休息。我们循环到标称容量的 80% （0.88 Ah）。在每次测试开始时执行初始 C/10 周期。恒压阶跃的截止电流在充电和放电时均为C/50。红外测试的脉冲宽度为 30 ms。

特征提取：我们提取了电池循环的具体信息，发现在早期周期中，能力退化可以忽略不计;但是，电压-容量放电曲线发生了变化。为了捕捉循环早期电位与循环寿命之间的关系，通过放电电压曲线测量了几个特征，如图1所示。图1(a)显示了循环寿命为534次的电池放电数据，包括了第10个循环周期的放电电压和放电容量曲线图，第100个循环周期的放电电压和放电容量曲线图。图1(a)显示了循环寿命为534次的电池放电数据。图1(a)显示了循环寿命为534次的电池放电数据。我们可以明显发现，随着电池循环寿命的增加，第10-100循环电压曲线的面积差越来越小，同理在100-150也是如此。我们提取了电池循环的两个特征，分别是V10-100，V100-150并用斯皮尔逊相关系数进行分析，分析结果显示，V10-100、V100-150和电池寿命都有强相关性，且相比之下V100-150和电池寿命有着更强的相关性。

模型介绍：

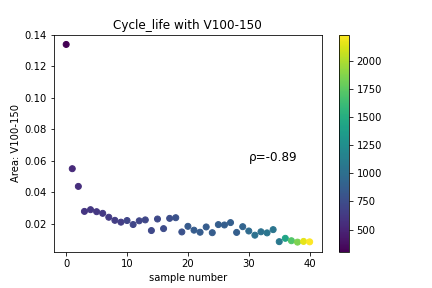
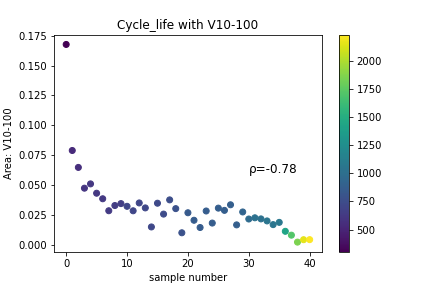
斯皮尔曼系数:由于皮尔逊相关系数的使用条件相当苛刻：两组变量必须是连续数据、呈现正态分布，且两者间必须成线性关系。故在此文中我们选择斯皮尔曼相关系数对输入特征和电池寿命结果进行分析。 它是衡量两个[变量](https://baike.baidu.com/item/%E5%8F%98%E9%87%8F?fromModule=lemma_inlink)的[依赖性](https://baike.baidu.com/item/%E4%BE%9D%E8%B5%96%E6%80%A7/3927846?fromModule=lemma_inlink)的非参数指标。它利用[单调](https://baike.baidu.com/item/%E5%8D%95%E8%B0%83/9753519?fromModule=lemma_inlink)方程评价两个统计变量的相关性。 如果数据中没有重复值， 并且当两个变量完全单调相关时，斯皮尔曼相关系数则为+1或−1，两个变量越靠近1说明其越靠近正相关，spearman越靠近-1，说明两者负相关关系越强。原始数据依据其在总体数据中平均的降序位置，被分配了一个相应的等级。Spearman计算公式为：

，其中，表示排序位置的差值，代表电池容量，代表HI的排序位置

，其中n代表样本容量，就是我们要计算的斯皮尔曼相关系数。

根据公式，我们就能计算出特征和电池寿命的相关性，本文提取了两个关于电池循环的特征，

分别是电池的第十次循环和第一百次循环的电压—放电容量曲线面积差作为特征F1，电池的第十次循环和第一百次循环的电压—放电容量曲线面积差作为特征F2,；我们首先将数据集的前41个数据样本提取出来，再根据电池的循环寿命长短进行排序，这样能得到更直观的结果。经过计算可以发现图1是F1和电池寿命的相关性，相关系数为-0.78，图2是F2和电池寿命的相关性，相关系数为-0.89。（-0.899444523235248，-0.7842766999664926）由上述两幅图可知，相较于跨度为100个周期的V10-100，跨度为50个周期的V100-150与电池寿命有着更高的相关度。且对于一些循环周期数短的电池，获取150个循环周期对比200个周期对电池有更小的损耗，能够从更早的电池循环数据中获取信息对电池寿命进行预测，是我们研究的目标。



降噪方法：

EMD:

对于信号首先有如下假设：（1）任何信号都是由若干本征模态函数组成的；（2）各个本征模态函数即可是线性的，也可是非线性的，各本征模态函数的局部零点数和极值点数相同，同时上下包络关于时间轴局部对称；（3）在任何时候，一个信号都可以包含若干本征模态函数，若各模态函数之间相互混叠，就组成了复合信号；

EMD将输入信号分解为几个本证模函数和一个残差组成，即由下列公式组成：

其中表示输入信号， 表示的本征模函数， 表示残差

EMD分解步骤：

1.寻找信号想全部极值点，通过三次样条曲线将局部极大值点连成上包络线，将局部极小值点连成下包络线。上、下包络线包含所有的数据点。

2.  由上包络和下包络线的平均值，得出

3.原始信号减去均值包络线，得到中间信号：,该过程称之为筛分

4.判断该中间信号是否满足IMF的两个条件,如果满足,该信号就是一个IMF分量;如果不是，以该信号为基础，重新做1~4的分析。即继续进行“筛分”直至分解k次后分解得到的信号满足IMF条件得到原始信号的第一个IMF分量

代表原始信号中最高频的IMF分量，从原始信号中减去得到剩余分量。IMF分量的获取通常需要若干次的迭代。

5.对进行上述“筛分”处理，可得到第二个IMF分量，再将减去后获得剩余分量。如此分解下去，直到最后一个残余信号无法继续分解为止，完成对信号×(t)的EMD过程:

...

6.当变成单调函数后，剩余的成为残余分量。所有IMF分量和残余分量之和为原始信号

CEEMDAN：

假设为经过EMD分解之后得到的第个本证模态分量，CEEMDAN分解得到的第个本征模态分量为，为满足标准正态分布的高斯噪声信号，是加入噪声的次数，为白噪声的标准差，为待分解信号，CEEMDAN分解步骤如下：

1)将高斯白噪声加入到待分解信号得到新信号，其中.对新信号进行EMD分解，得到第一阶本征模态分量

2)对产生的N个模态分量进行总体平均就得到CEEMDAN分解的第1个本征模态分量:

3)计算去除第─个模态分量后的残差:

4)在中加入正负成对高斯白噪声得到新信号，以新信号为载体进行EMD分解，得到第一阶模态分量，由此可以得到CEEMDAN分解的第2个本征模态分量:

5)计算去除第二个模态分量后的残差:

6）重复上述步骤，直到获得的残差信号为单调函数，不能继续分解，算法结束。此时得到的本征模态分量数显为K，则原始信号被分解为:

线性回归：

线性回归模型定义为：，其中，表示模型系数，表示样本数目，线性回归就是找到合适的模型系数使尽可能的贴近实际值y。线性回归的目标函数是，将其展开并对W求导，令导数等于零进行矩阵变换能得到解析解，这就是利用最小二乘法求解线性回归的结果。

实验结果：我们首先提取了特征F2，即第100个循环和第150个循环中放电容量和电压曲线之间的面积差，加入噪声之后，利用完全自适应噪声集合经验模态分解对该特征进行降噪分析得到新的输入特征，使用线性回归对新生成的数据进行预测，结果如下：

|  |  |  |
| --- | --- | --- |
| **MAE** | | |
| Noise(\*0.05) | LR | LR\_CEEMDAN |
| 0.1 | 0.1178 | 0.0927 |
| 0.5 | 0.1263 | 0.0958 |
| 0.9 | 0.1779 | 0.1556 |
| 1 | 0.1712 | 0.1443 |

|  |  |  |
| --- | --- | --- |
| **RMSE** | | |
| Noise(\*0.05) | LR | LR\_CEEMDAN |
| 0.1 | 0.1528 | 0.1256 |
| 0.5 | 0.1648 | 0.1184 |
| 0.9 | 0.2095 | 0.1960 |
| 1 | 0.2172 | 0.1824 |

由表可知，利用线性模型对加入噪声的输入特征进行预测，预测结果不如使用完全自适应噪声集合经验模态分解对特征进行降噪处理的结果。在带有噪声的情况下，使用CEEMDAN方法降噪提高了模型预测精度。

总结：

锂电池作为新能源在汽车以及航空航天方面的普及一方面很大程度上缓解了能源消耗另一方面也带来了新的问题，对于锂电池的寿命预测成为亟需解决的研究热点。本文利用了锂电池充放电的早期循环数据，发现了早期循环数据特征和电池寿命有着密不可分的联系，且考虑了数据带有噪声的情况下，使用完全自适应噪声集合经验模态分解对带有噪声的数据进行分解产生新的不会被噪声影响的输入特征，并根据这个发现建立起了锂电池早期循环数据和电池寿命预测的模型，利用数据驱动的方法对锂电池寿命进行精准预测。不需要了解锂电池内部的物理化学反映，只需电池早期的充放电数据，仅使用数据驱动的方法就能预测电池寿命，对于电池的损耗少，对锂电池寿命预测的研究有着重大意义。