基于阻抗的不均匀使用锂离子电池性能预测

（Impedance-based forecasting of lithium-ion battery performance amid uneven usage）

大多数关于电池健康预测的研究都集中在电池处于相同使用模式的研发环境中。然而，在实际操作中，细胞和周期之间的使用有很大的可变性，因此使预测具有挑战性。我们提出了电化学阻抗光谱测量与概率机器学习方法的结合。利用88个商业锂离子硬币电池通过多级充放电产生的数据集(循环之间电流随机变化)，我们表明，在未来循环协议和充电前立即进行的单一电化学阻抗谱测量的情况下，在不了解任何使用历史的情况下，未来的放电容量可以通过校准的不确定性进行预测。

这两种类型的预测方法可以细分为经验模型、基于物理的模型和数据驱动模型，有些模型是这些模型的混合。经验方法已被用于模拟长期容量衰减与幂律，但假设在电池寿命固定运行，并没有考虑电池在使用寿命开始时状态的内在差异。这些方法假设，如果以相同的方式操作，具有相同化学性质的所有细胞将以相同的方式褪色，但在实践中没有观察到这一点。在基于物理的方法中，电池要么使用内部物理和电化学过程的第一性原理分析进行机械建模，要么使用等效电路建模，将电池建模为由代表潜在电化学过程的电阻和电容组成的电路25,26。机械模型的目的是捕捉电池电压对外部施加电流的反应(或反之)，这可用于预测最佳充电方案17。

这类模型的参数需要针对每个单独的细胞进行更新，并且通常存在不可识别性——几组模型参数可以同样很好地解释观察到的数据，但对测试细胞或同一细胞在其生命的后期会做出截然不同的预测。电路参数可以用来预测在标准化使用条件下的容量退化20或模拟不同使用条件对电池组性能的影响30。然而，在一个分析模型中捕获每一个降解模式是具有挑战性的。此外，必须从一个周期到另一个周期学习每个细胞的一组新的模型参数，这使得推断一个一般的细胞到细胞模型具有挑战性。

纯数据驱动的预测方法使用原始数据作为机器学习算法的输入，预测长期容量衰减、阻力增加和剩余使用寿命31,32。数据驱动方法将机器学习应用于从充放电曲线中提取的特征，以预测放电容量19、剩余使用寿命22和容量突然衰减23、33。

在从充电/放电曲线中提取特征34和用于建模时间序列数据的机器学习方法方面的创新，使得预测的准确性得到了显著提高。

在这项工作中，我们试图确定是否存在一个足够信息的细胞健康标记，可以用来预测短期和长期的未来性能，在不均衡的历史和未来细胞使用，在充电前采集EIS光谱，可以预测下一个周期和更长期的电池容量，测试误差小于10%。

摘要：

然而，在实际操作中，细胞和周期之间的使用有很大的可变性，因此使预测具有挑战性。为了解决这一挑战，我们提出了电化学阻抗光谱测量与概率机器学习方法的结合。利用88个商业锂离子硬币电池通过多级充放电产生的数据集(循环之间电流随机变化)，我们表明，在未来循环协议和充电前立即进行的单一电化学阻抗谱测量的情况下，在不了解任何使用历史的情况下，未来的放电容量可以通过校准的不确定性进行预测。

引言：

电池性能的短11、17-20和长21 - 23时间尺度预测在电池预测中都很有意义。在短时间内，预测电池对特定充放电协议的反应可以用来制定最佳充电协议17。短期预测还包括SOH估计11,18 - 20:在这里，目标是在特定的、标准化循环下预测电池的电容或内阻协议。在很长一段时间内，研究的重点是预测剩余的使用寿命，即寿命的结束，以及电池寿命轨迹中的“膝点”，此时电池的退化会加速。

对于基于电路的模型，电路的参数可以拟合到电流-电压数据20,27，或电化学阻抗谱28,29。然后，电路参数可以用来预测在标准化使用条件下的容量退化20或模拟不同使用条件对电池组性能的影响30。然而，在一个分析模型中捕获每一个降解模式是具有挑战性的。此外，必须从一个周期到另一个周期学习每个细胞的一组新的模型参数，这使得推断一个一般的细胞到细胞模型具有挑战性。

除了充放电曲线，电化学阻抗谱(EIS)21、早期循环库仑效率38、电流中断39和声学飞行时间分析18、40等方法已被用于降解预测。

在这项工作中，我们试图确定是否存在一个足够信息的细胞健康标记，可以用来预测短期和长期的未来性能，在不均衡的历史和未来细胞使用。我们发现，在充电前采集EIS光谱，可以预测下一个周期和更长期的电池容量，测试误差小于10%。

当在与训练模型的周期条件相似的细胞上进行测试时，我们的模型达到了与最先进的预测模型相当的精度(8.2%的测试误差对8.8%的测试误差)，除了我们的模型在不访问任何历史数据的情况下进行预测，而之前的最先进的模型需要来自细胞周期轨迹的历史数据。此外，当外推到不同的操作温度时，我们的模型明显优于最先进的模型，实现了57%的测试误差减少(从34.2%到14.6%)。

我们观察到，我们的模型是数据高效的，只需要8个单元就可以达到小于10%的测试误差。至关重要的是，我们的方法对数据集的迁移是稳健的，在训练集的循环模式分布不同的数据集上，测试误差小于7%。此外，我们还证明，如果可用，使用基于历史容量-电压数据的附加特征可以增强状态表示，减少平均测试误差高达25%。我们的方法在电池制造商、平均使用模式和操作温度方面是稳健的。

此外，我们的工作提供了大量动态工作条件下细胞的循环数据，填补了公开可用数据的空白46。我们的工作集中在一组理想化的使用分布，而不是现实的驱动轮廓，以证明模型的可泛化程度。我们的工作脱离了NASA随机使用数据47，该数据随机循环电池50个周期，然后通过“参考”协议充电后测量下一个周期的放电容量。尽管基于这些数据已经建立了几个预测随机条件下退化的模型12,19,48，但单一协议对下一个循环放电容量的影响无法被分解，而且需要每几个周期就有一个参考充放电协议，这与典型的现场使用不一致。

结果：

Data generation：

我们生成了两个独立的数据集，对应于从两个不同的制造商购买的商用LiR硬币电池，这允许我们测试我们的方法对于电池制造商是否健壮。第一个数据集对应40个Powerstream LiR 2032硬币电池(标称容量1C = 35毫安时)。我们将24个电池置于随机选择的充放电电流序列中，在23±2°C下进行110-120个完整的充放电循环。每个循环包括电池状态的初始诊断，包括恒流EIS光谱的采集，然后是使用，包括充放电阶段。充电和放电分别由两级和一级恒流(CC)协议组成;电流分别在70-140 mA (2-4 C)、35-105 mA (1-3 C)和35-140 mA (1-4 C)的范围内随机选择。

为了测试模型对畴移的鲁棒性，我们在上面相同的条件下对剩下的16个电池进行循环，除了现在将所有电池和周期的放电电流固定在52.5mA (1.5 C)，而不是在每个周期随机改变放电电流。在使用第一个数据集来确认该方法可以成功预测未来几个周期的放电容量后，我们随后显著扩展了我们的分析，以探索该模型对电池制造商、使用模式和工作温度变化的稳健性。

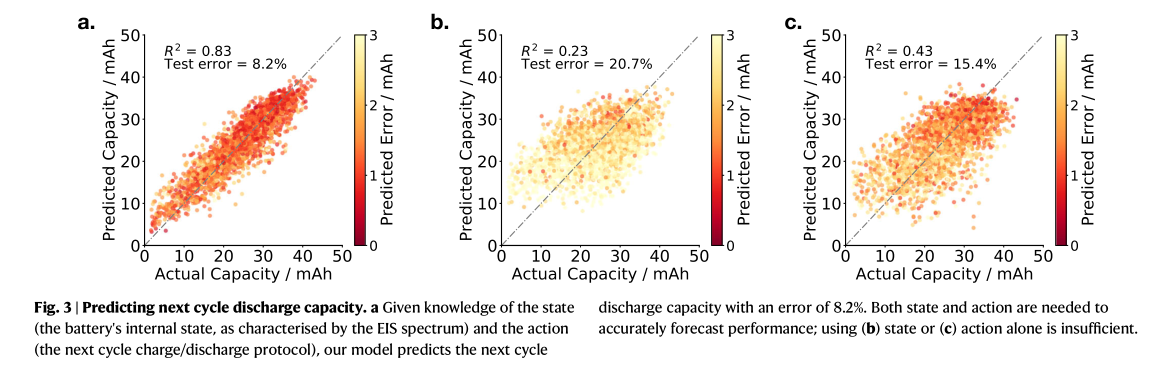
为了实现这一目标，我们在更广泛的使用模式下，从第二个制造商RS Pro(名义容量40毫安时)循环额外的48个电池。在这种情况下，每个电池再次进行100次两级CC充电和一级CC放电的循环，在每个循环开始时随机选择三个速率。然而，我们现在通过为每个电池设置不同的电流分布，来复制不同电池用户彼此有不同的平均使用模式，但仍然表现出随机的周期与周期行为的场景，从而使这个问题更具挑战性。在这些电池中，有16个也在35°C的较高工作温度下循环。

使用EIS进行容量预测：

(使用模型预测下一个循环放电容量，分别测试了只考虑状态向量、未来动作、还有两者结合，发现两者结合效果最好。然后再推广到不仅是下一次循环甚至更远的循环次数)

如果我们应用特定的充放电剖面，我们首先考虑我们想要预测下一个循环放电容量的设置，对于一个使用历史(包括例如，周期或日历年龄，或历史容量-电压数据)完全未知的电池。我们将问题框定为一个回归任务，训练一个概率机器学习模型来学习映射，Qn = f(sn, an)，具有不确定性估计，其中sn是第n个循环开始时的电池状态，an是未来的动作(第n个循环充放电协议)，Qn是循环结束时测量到的放电容量。电池状态矢量sn是由实数(Z re)和虚数(Z im) co mp on n t s o the在57个频率下测量的阻抗串联而成，ω1，。ω57，在0.02Hz-20kHz范围内;sn =½Z再保险ðω1ÞZ imðω1Þ,:::,Z再保险ðω57Þ,Z imðω57Þ?。动作矢量是由第n个循环的充放电电流串联而成。

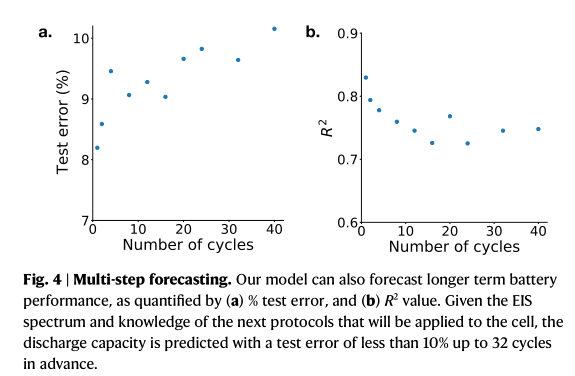
图3说明了我们模型的准确性。将状态和动作同时作为输入，预测下一轮流量，平均误差为8.2%。重要的是，我们发现状态和动作(图3a)都是预测未来单元性能的必要条件:如果只使用状态(图3b)或动作(图3c)作为输入，测试误差大约翻倍至20.7%和15.4%。



这证明了电池的内部健康和外部选择的用途在确定实现的电池性能中的重要性。

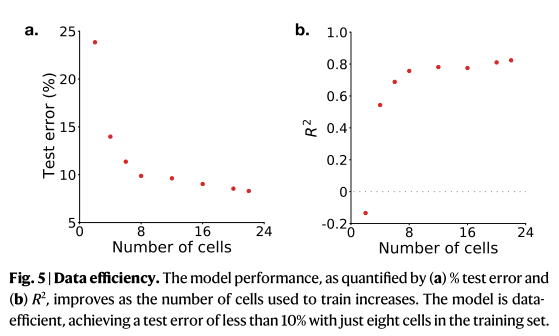
对于优化充电和重新分类等应用，重要的是电池寿命轨迹模型不仅可以预测下一个循环的放电容量，而且可以预测未来几个循环的容量49,50。利用这种思想，我们接下来研究当我们将模型推进到更远的未来产能预测时，模型的预测准确性如何变化。在每种情况下，输入包括第n个周期开始时的状态表示的串联，sn, 还有“动作”向量和a(n.....n+j包括所有将应用在循环n和循环n +j之间充电和放电电流.

图4显示了决定系数R2随j的变化情况。不出所料，随着预测区间的增大，模型的准确性普遍下降。然而，当预测未来40个周期时，模型仍然达到R2 = 0.75.



数据效率和对域漂移的鲁棒性：

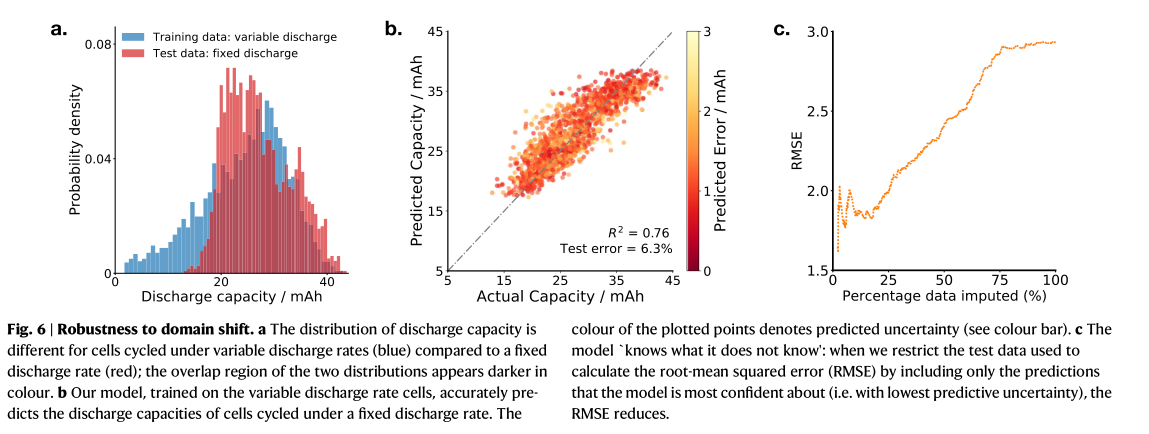
接下来，我们通过调查数据效率和模型的可泛化性来测试我们方法的鲁棒性。为了测试数据效率，我们测量性能如何随着用于训练模型的单元数的增加而变化。如图5所示，当单元数从2个增加到22个时，测试误差从23.8%显著降低到8.2%。尽管如此，该模型明显具有数据效率，只需要8个单元就可以获得小于10%的测试误差



模型可泛化性的一个重要测试是研究领域分布变化时的模型准确性，即当模型部署在与训练数据不同的设置中时12。

这对于在现场部署是很重要的，因为该方法需要对可能与训练数据不同的驾驶模式具有鲁棒性8。我们通过循环来自同一制造商的另外16个电池来测试模型的稳健性，但现在调整循环协议，将每个电池的整个寿命周期的放电电流固定在1.5C。我们使用一个只训练细胞在其生命周期中受随机放电电流的模型，来预测受固定放电的细胞的下一个周期放电能力。为了说明训练和测试数据集的差异，放电能力的分布如图6a所示。

模型对固定流量数据集的预测精度如图6b所示。值得期待的是，该模型在这个域移位数据集上的测试误差仅为6.3%，对应于R2 = 0.76。



我们的模型还输出预测不确定性，这表明模型对其预测质量的确定程度。在域转移设置中，尤其重要的是，模型“知道它不知道的东西”,并估计可能获得高误差的数据点的高预测不确定性。我们可以通过观察当数据点的数量减少到只包括模型最有把握的数据点时，平均测试误差如何变化来测试模型估计其不确定度的能力。如果一个模型能够成功地估计其确定性水平，那么平均测试误差应该随着数据比例的减少而减少，只包括最有把握的预测点。图6c的均方根误差(RMSE)减少了3.2%，因为数据的比例从100%减少到最可信的25%，这表明我们的模型已经知道它应该对哪些预测有信心。

状态表示的比较：

在演示了EIS光谱捕获电池状态的能力之后，我们现在将这种电池健康表示与文献中使用的其他方法(包括最先进的基于特征的方法22,51)进行了基准测试，并考虑EIS光谱中是否存在可用于增强电池状态的附加特征。用来预测或估计电池SOH的简单方法包括使用以前的循环放电容量，或循环开始后的容量吞吐量。更先进的方法包括提取历史容量-电压放电曲线的特征，如图1所示。提取这些特征的最先进的方法是由Severson等人实现的，而最近Attia等人和Paulson等人使用的特征提取方法是相同的37,51。我们对基于eis的方法相对于那些最先进的功能的表现进行了基准测试。

此外，我们还评估了以等效电路模型(ECM)的形式加入物理解释是否能改善预测。我们使用广泛实现的兰德尔斯电路模型，包括一个串联电阻，与一个并联的电阻与一个电容和一个华宝阻抗元件连接，以及更复杂的扩展兰德尔斯电路，它在兰德尔斯电路中增加了一个额外的电阻-电容并联组合。ECM拟合到频谱(以一个相关的计算成本)，我们使用提取的参数作为状态表示，而不是原始的EIS数据。

总的来说，我们在基准测试中考虑了以下特性:

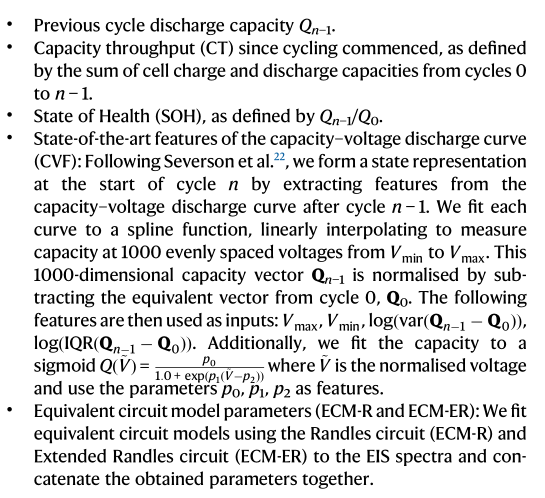
前一次循环放电容量Qn−1。

•循环开始以来的容量吞吐量(CT)，由从循环0到n−1的电池充放电容量之和定义。

•运行状况(SOH)，定义为Qn−1/Q0。

•容量-电压放电曲线(CVF)的最新特征:根据Severson et al22，我们通过从n - 1周期后的容量-电压放电曲线中提取特征，形成了在第n周期开始时的状态表示。我们将c - h曲线与样条函数拟合，线性插值以测量在V min到V max之间的1000个均匀间隔电压下的容量。T h i的1000维容量向量Qn−1通过从周期0 Q0减去等价向量归一化。然后使用以下特征作为输入:V max, V min, lo ðvarð qn ?1 ?Q0ÞÞ、日志ðIQRðQn吗?1 ?Q0ÞÞ。此外，我们将容量拟合为s型图Qð VÞ = p0 1:0 + expðp1ð~V?p2ÞÞ，其中~V为归一化电压，使用参数p0, p1, p2作为特征。

•等效电路模型参数(ECM-R和ECM-ER):我们使用兰德尔电路(ECM-R)和扩展兰德尔电路(ECM-ER)将等效电路模型与EIS谱拟合，并将获得的参数连接在一起。



我们注意到，与EIS特征相比，使用前面提到的前四个特征形成状态表示需要访问历史电流电压数据，至少在之前放电的整个过程中，对于某些特征，在整个电池寿命中。然而，他们受益于不需要设备来测量EIS光谱的优势，这带来了相关的财务和时间成本。使用ECM参数(从EIS谱中提取)形成状态表示具有相关的计算成本，可以认为是原始EIS数据的降维形式。ECM通常面临的另一个问题是非唯一性，即ECM参数的多个不同组合通常可以同样好地解释特定的EIS谱。

表1显示了状态表示如何影响测试误差和模型拟合优度。在所有情况下，在给定下一个循环协议和所选状态表示的情况下，训练模型预测下一个循环放电容量。为了询问特性的相对重要性，我们首先考虑只使用EIS(不包括协议)和只使用协议(不包括EIS)的基线。也许毫不奇怪，电池的退化是当前状态和未来充放电协议的功能。因此，同时使用EIS和协议的性能明显优于仅使用EIS或仅使用协议的性能。

我们现在将我们的分析扩展到探索我们的方法在改变电池制造商、调整操作温度和调整平均使用模式方面的鲁棒性。我们在另一家制造商RS Pro的新一批32个商用LiR硬币电池(标称容量为1 C = 40 mAh)上重复了我们的实验，只是我们现在通过将不同的电池分组置于四种不同的使用分布中的一种，使问题明显更具挑战性。这些使用分布情况载于补充表1。

我们通过两种方式来衡量模型的准确性:首先，我们考虑这样一种情况:模型暴露在与测试集受到相同协议分布(随机分裂)的细胞中;其次，更有挑战性的情况是，模型只在受三种循环协议分布影响的细胞上进行训练，并在受不同循环协议影响的其余八个细胞上进行测试。这是一项困难得多的任务，因为测试单元的平均使用量与训练单元的平均使用量非常不同——这是一个测试，看模型是否可以推断出不同的平均使用量，而不仅仅是不同的周期与周期之间的使用

不同状态表示的结果如表2所示，这两种情况下，训练/测试的分割是随机的，以及分割被划分为不同的使用模式。对从第二个制造商购买的电池进行了比较观察:即，当使用EIS光谱特征和放电曲线(CVF)形成的特征形成状态表示时，可以做出最准确的预测。正如预期的那样，当使用来自一些细胞的数据进行训练时，模型的表现明显更好，这些细胞暴露在与模型所测试的细胞相似的循环模式分布中。然而，该模型在支架分裂场景下仍保持性能，在这种设置下，当使用EIS谱和放电曲线的特征形成状态表示时，而不是仅使用放电曲线的特征，测试误差减少了30%。

这些额外的结果进一步证明，如果可用，EIS谱和放电曲线都可以作为电池内部状态的信息标记，但它们是相互补充的。

接下来，我们验证了该模型对于外部工作温度的变化是稳健的。我们在35°C下循环了另外16个细胞，并测试了在室温下循环的细胞数据训练的模型。表3显示，我们的模型可以推断在这些较高温度下运行的电池，但当电池不在相同温度下运行时，EIS谱在描述电池状态方面发挥着特别重要的作用。当仅使用放电曲线特征表征状态时，模型的测试误差为34.2%，而同时使用EIS谱和放电曲线特征时，模型的测试误差为14.0%。这进一步证明了EIS信号包含相对于充放电曲线的附加信息，并支持EIS隐式跟踪温度的假设