

Estimativa de SOC em baterias íon-lítio de veículos elétricos: uma abordagem com dados de ciclo de operação e métodos híbridos de machine learning

Guilherme Mاتيولli
CECS

Universidade Federal do ABC
Santo André, Brasil
guilherme.matiolli@aluno.ufabc.edu.br

Priscila Maia da Silva
CECS

Universidade Federal do ABC
Santo André, Brasil
priscila.maia@aluno.ufabc.edu.br

Ricardo Caneloi dos Santos
CECS

Universidade Federal do ABC
Santo André, Brasil
ricardo.santos@ufabc.edu.br

Patricia Teixeira Leite Asano
CECS

Universidade Federal do ABC
Santo André, Brasil
patricia.leite@ufabc.edu.br

Ricardo Suyama
CECS

Universidade Federal do ABC
Santo André, Brasil
ricardo.suyama@ufabc.edu.br

Romulo Gonçalves Lins
CECS

Universidade Federal do ABC
Santo André, Brasil
romulo.lins@ufabc.edu.br

Resumo—Este trabalho propõe um método para estimativa do estado de carga (SOC) de baterias de íons de lítio utilizando redes neurais LSTM combinadas com pós-processamento por Filtro de Kalman Unscented (UKF). O modelo foi treinado com dados de perfis dinâmicos (DST) e validado em ciclos FUDS e US06 sob diferentes temperaturas. Os resultados demonstraram alta acurácia, com RMSE inferior a 2,2% e R^2 superior a 99%, mesmo em condições não vistas durante o treinamento. A aplicação do UKF aprimorou a estabilidade das predições, reduzindo erros máximos e suavizando flutuações. O tempo de execução obtido confirma a viabilidade do método para aplicações em tempo real, como sistemas embarcados, gêmeos digitais e plataformas cyber-físicas.

Palavras-chave—SOC, baterias, LSTM, UKF, veículos elétricos.

I. INTRODUÇÃO

O avanço da mobilidade elétrica, a ampla implementação de fontes de energia intermitentes e a crescente demanda por dispositivos móveis têm impulsionado significativamente o desenvolvimento de baterias com maior desempenho e confiabilidade. De acordo com o Stated Policies Scenario, que considera as políticas governamentais atualmente em vigor, estima-se que a capacidade global de produção de baterias aumente de 170 gigawatts-hora (GWh) por ano em 2019 para 1.500 GWh por ano até 2030. Já o Sustainable Development Scenario projeta uma demanda ainda mais elevada, atingindo 3.000 GWh anuais no mesmo período [1]. Nesse contexto, as baterias de íons de lítio (Li-ion) têm se consolidado como a tecnologia de armazenamento de energia mais promissora e de crescimento mais acelerado, especialmente no setor de veículos elétricos (EV), em razão de suas vantagens como alta densidade energética, carregamento rápido, longa vida útil, ausência de efeito memória, baixa taxa de autodescarga e reduzido impacto ambiental [2].

Para garantir elevados níveis de desempenho, autonomia e vida útil, sem comprometer a segurança operacional, os veículos elétricos empregam Sistemas de Gerenciamento de Baterias (Battery Management Systems – BMS). Esses sistemas são responsáveis por monitorar e controlar parâmetros essenciais, como o State of Charge (SOC) e o State of Health (SOH) [3]. O SOC, em particular, representa a fração da capacidade remanescente da bateria em relação à sua capacidade nominal [4]. No entanto, por se tratar de uma variável não diretamente observável por sensores, sua estimativa depende da aplicação de algoritmos especializados, que podem estar embarcados no próprio BMS ou integrados a plataformas online, como gêmeos digitais e sistemas cyber-físicos [5].

A estimativa do SOC é uma tarefa desafiadora, uma vez que essa variável apresenta comportamento dinâmico, altamente não linear e dependente da temperatura e da cinética físico-química da bateria [6]. Tradicionalmente, dois métodos principais têm sido empregados: o método baseado na tensão de circuito aberto (Open Circuit Voltage – OCV) [7] e a contagem de coulombs (coulomb counting) [8]. Embora de implementação simples, a contagem de coulombs tende a acumular erros ao longo do tempo devido à integração contínua da corrente. Por outro lado, o método OCV desconsidera efeitos importantes como a histerese, a impedância interna e os processos de difusão [9]. Como alternativa, métodos mais sofisticados vêm sendo explorados, como os modelos de circuito equivalente (ECM) e abordagens baseadas em dados [10]. No entanto, os ECM exigem processos complexos de identificação de parâmetros, o que se torna desafiador em baterias compostas por muitas células.

Entre os modelos baseados em dados utilizados para a estimativa do SOC destacam-se as máquinas de vetores de suporte (Support Vector Machines – SVM) [11] e as redes

neurais artificiais (NN) [12]. As redes neurais, em particular, constituem uma abordagem robusta para lidar com sistemas não lineares, como é o caso das baterias de íons de lítio. As NN são capazes de aprender a relação entre variáveis operacionais e o SOC, a partir de dados históricos coletados, a rede ajusta seus pesos internos e generaliza padrões para realizar estimativas em novos cenários operacionais [9].

Contudo, em aplicações reais de veículos elétricos, os perfis de condução podem apresentar ampla variabilidade, tanto entre diferentes ciclos padronizados quanto em situações cotidianas. Isso impõe um desafio adicional ao modelo preditivo, que deve ser capaz de generalizar corretamente para condições não observadas durante o treinamento. Portanto, a capacidade da rede neural em capturar características fundamentais do comportamento da bateria, e não apenas memorizar sequências específicas, é essencial para garantir sua aplicabilidade.

Neste trabalho, propõe-se um método de estimativa do SOC baseado em redes neurais do tipo LSTM (Long Short-Term Memory), combinadas com o Filtro de Kalman Unscented (UKF). Para capturar a dinâmica temporal e não linear da bateria em condições reais de operação, a rede neural LSTM é alimentada com múltiplas medições históricas de corrente, tensão e temperatura, tendo o SOC como variável de saída. Visando aprimorar a consistência física das estimativas e mitigar ruídos ou flutuações indesejadas geradas pela rede, o UKF é empregado como técnica de pós-processamento, caracterizando uma abordagem híbrida. O modelo proposto foi treinado com dados provenientes de testes de estresse dinâmico (Dynamic Stress Test – DST) e validado utilizando perfis de condução padrão, como o Federal Urban Driving Schedule (FUDS) e o US06 Highway Driving Schedule. A principal contribuição deste trabalho está no desenvolvimento de um modelo híbrido baseado em redes neurais LSTM com pós-processamento com UKF, combinando precisão, robustez e baixo custo computacional. O modelo foi treinado com um conjunto limitado de dados reais e validado em ciclos e temperaturas não utilizadas no treinamento, demonstrando boa capacidade de generalização. Essa característica viabiliza sua aplicação em sistemas embarcados, gêmeos digitais e plataformas cyber-físicas, com foco em inferência em tempo real e resiliência frente à variabilidade operacional. Em relação a trabalhos relacionados, o modelo proposto difere de abordagens como: (i) a utilização de redes MLP com UKF [6], com menor capacidade de capturar dinâmicas temporais; (ii) o uso de LSTM com filtro adaptativo (ACKF) [36], de maior complexidade computacional; e (iii) a aplicação de restrição de saída em redes LSTM baseada no contador de Coulomb [14], que depende de suposições sobre a dinâmica da bateria. O modelo apresentado neste estudo adota uma estrutura causal com UKF, mantendo desempenho competitivo e maior viabilidade para aplicação prática.

O artigo está organizado da seguinte forma: A Seção 2 apresenta as principais técnicas e tecnologias usadas estimativa de SOC com uma abordagem de dados; A Seção 3 apresenta a metodologia do trabalho, desde a apresentação dos dados utilizados até o método híbrido desenvolvido para treinamento

e predição; A Seção 4 traz os resultados obtidos durante os teste e suas respectivas análises; E a Seção 5 apresenta a conclusão desse artigo.

II. REVISÃO BIBLIOGRÁFICA

Entre os métodos de estimativa do State of Charge (SOC), as abordagens baseadas em aprendizado de máquina (ML) têm se destacado por sua adequação a aplicações em engenharia [14]. Esses modelos realizam um mapeamento não linear direto entre variáveis mensuráveis da bateria e o SOC, aprendendo off-line a partir de dados operacionais, sem exigir modelos físico-químicos explícitos, o que tem atraído crescente atenção na última década [15].

Algoritmos como máquinas de vetores de suporte (SVM) [16], regressão por processo gaussiano (GPR) [17], redes neurais de retropropagação (BPNN) [18] e redes com funções de base radial (RBFNN) [19] vêm sendo aplicados. Dado o caráter sequencial da estimativa de SOC, redes neurais recorrentes (RNN) ganharam espaço [20]. Para superar a limitação das RNNs em capturar dependências de longo prazo, arquiteturas como LSTM [21], [22] e GRU [23] incorporaram mecanismos de controle de fluxo de informação (gating).

Modelos de ML geram estimativas independentes, sem necessidade de estado inicial, evitando o acúmulo de erro. No entanto, essa independência pode resultar em oscilações abruptas no SOC estimado, sobretudo diante de variações bruscas de corrente ou tensão. Tais flutuações comprometem a estabilidade e não são facilmente mitigadas via ajuste de hiperparâmetros. Para contornar essa limitação, técnicas de suavização com conhecimento físico ou filtros estatísticos, como o Kalman Filter, têm sido adotadas [6], [14].

Em [6], redes neurais profundas atingiram RMSE inferior a 2,5% com dados não vistos. No mesmo trabalho, uma rede MLP obteve erro máximo de 4,2%, reduzido para menos de 2,5% com UKF. Em [14], o modelo EI-LSTM teve erro máximo inicial de 3,2%, reduzido a 1,5% de RMSE após aplicação de restrição de saída (Constrained Output – CO), elevando a precisão e estabilidade.

Com os avanços em tecnologias digitais, soluções baseadas em gêmeos digitais, sistemas cyber-físicos (CPS) e BMS online têm ganhado espaço na estimativa em tempo real do SOC. Diferente de métodos offline, essas abordagens integram dados reais e modelos preditivos em ambientes distribuídos. Em [25], um sistema baseado em big data e CPS foi capaz de estimar SOC em tempo real com dados dinâmicos. Em [26], um gêmeo digital auto-adaptativo ajustou parâmetros com base no envelhecimento e variações de carga, aprimorando a confiabilidade. Já [27] propôs uma arquitetura distribuída baseada na colaboração entre nuvem, borda e dispositivos embarcados, com aprendizado contínuo centralizado em um gêmeo digital, destacando-se como caminho promissor para previsões precisas em aplicações veiculares.

III. METODOLOGIA

A. Descrição dos Dados

Neste estudo, foram utilizados dados experimentais disponibilizados pelo Center for Advanced Life Cycle Engineering (CALCE) da Universidade de Maryland [28], acessíveis em: <https://calce.umd.edu/battery-data>. O conjunto compreende registros de ciclos de carga e descarga de células LiFePO de 1,1 Ah, realizados em câmara climática com controle preciso de temperatura, sob os perfis dinâmicos DST, FUDS e US06. Os dados estão organizados por temperatura (0°C a 50°C) e perfil de condução, em arquivos estruturados contendo colunas de corrente (A), tensão (V), temperatura (°C), tempo (s) e SOC de referência (calculado por Coulomb Counting). A taxa de amostragem é de 1 Hz e o conjunto apresenta 20 ciclos de carga e descarga completos. Essa base é amplamente reconhecida por representar realisticamente os desafios operacionais de baterias tracionárias em veículos elétricos e tem sido utilizada na literatura para modelagem baseada em dados.

Os testes foram realizados com células LiFePO₄ de 1,1 Ah, submetidas a ciclos de carga e descarga em uma câmara climática com controle preciso de temperatura. Três perfis padronizados foram utilizados (DST, FUDS e US06), a característica de descarga da bateria para cada perfil e mostrada na Fig. 1. Esses testes foram conduzidos de forma cíclica de modo que o SOC seja levado de 100% até próximo de 0%, conforme mostrado na Fig. 2. Para cada temperatura a célula é testada com os três perfis de forma sequencial.

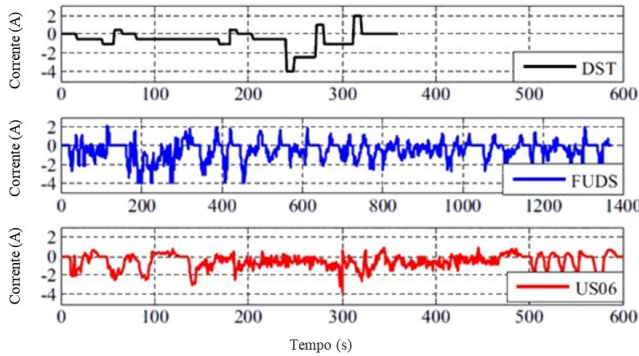


Fig. 1. Característica de consumo de corrente para os três perfis de teste (DST, US06 e FUDS) a 30°C

- Dynamic Stress Test (DST): dados com perfil com variações abruptas e rápidas na corrente, incluindo eventos de carga regenerativa, usado para treinar os modelos;
- Federal Urban Driving Schedule (FUDS): dados que representam longos trajetos em tráfego urbano, com frequentes variações de aceleração e frenagem regenerativa usados para avaliar o modelo;
- US06 Highway Driving Schedule (US06): dados que simulam condições de condução em rodovia com acelerações intensas e alta demanda energética, as frenagens regenerativas são tem baixa intensidade, usados para avaliar o desempenho do modelo.

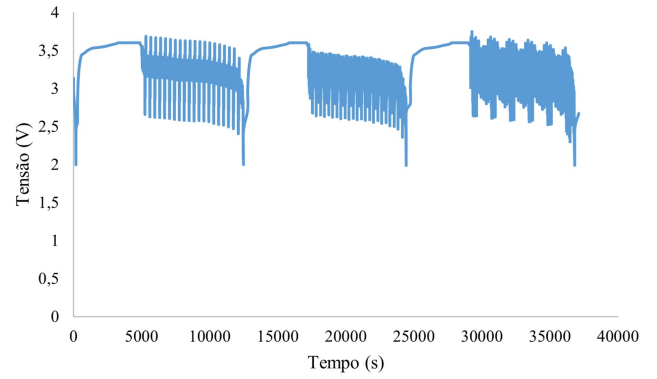


Fig. 2. Ciclos de teste DST, US06 e FUDS (respectivamente) em temperatura de 30°C

Todos os ensaios foram conduzidos nas temperaturas de 0 °C, 10 °C, 20 °C, 25 °C, 30 °C, 40 °C e 50 °C, com intervalo de amostragem de 1 segundo, conforme os procedimentos definidos no Battery Test Manual for Electric Vehicles do Idaho National Laboratory [29]. Cada amostra contém registros de corrente (A), tensão (V), temperatura (°C) e tempo (s). O valor de referência do SOC foi obtido por integração da corrente (método de Coulomb Counting), uma vez que o sistema de medição é muito preciso e sabe-se a condição inicial da bateria [6]. Em todos os testes o SOC inicia-se em 100% e termina próximo a 0%, os testes foram conduzidos de forma que a diminuição do SOC seja linear ao longo do tempo, conforme Fig. 3.

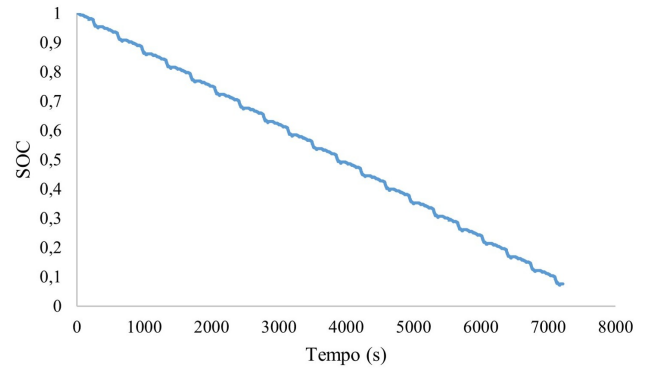


Fig. 3. Evolução do SOC durante o perfil US06 à 30°C

O perfil DST foi empregado exclusivamente no treinamento do modelo por abranger diferentes padrões de estresse dinâmico, enquanto os ciclos US06 e FUDS foram reservados para validação, a fim de avaliar a capacidade de generalização do modelo frente a dados não vistos. Essa estratégia de separação foi adotada por autores da área [6] [14].

B. Pré-processamento dos Dados

Antes da aplicação dos modelos de aprendizado de máquina, foi realizada uma etapa de pré-processamento dos dados com o objetivo de assegurar consistência temporal, integridade das variáveis e adequação estatística para modelagem.

Os dados utilizados compreendem registros de testes DST realizados sob diferentes temperaturas controladas: 10°C, 20°C, 30°C e 40°C. Para o treinamento dos modelos, os dados foram empilhados verticalmente, ou seja, concatenados mantendo-se a ordem sequencial de cada ciclo, mas respeitando os blocos por temperatura. Cada segmento referente a uma dada temperatura foi mantido intacto, preservando sua coerência interna ao longo do tempo [6] e [14].

Para alimentar a rede LSTM, os dados foram organizados em uma estrutura de janelas deslizantes temporais, conforme Fig. 4. Cada instância de entrada consiste em uma sequência de $n=50$, que representa 50 segundos de operação contínua. Essa abordagem captura a dinâmica do sistema ao longo do tempo, sendo eficaz para modelar fenômenos dependentes de histórico, como a estimativa de SOC em baterias LiFePO₄.

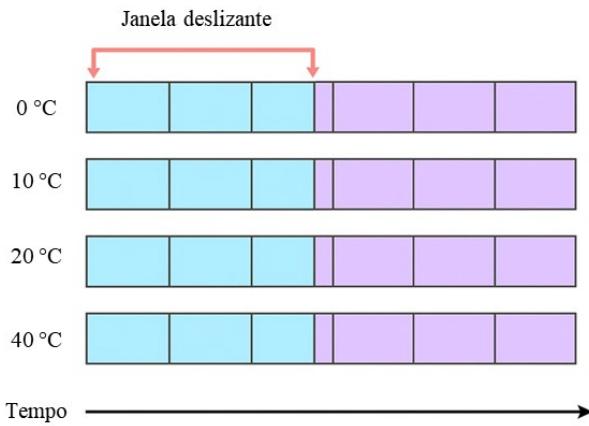


Fig. 4. Janelas deslizantes por segmentos de temperatura durante o treinamento com DST

Importante ressaltar que as janelas deslizantes foram limitadas ao interior de cada segmento de temperatura, não sendo permitido que transitem de um bloco térmico para outro. Essa decisão evita a introdução de descontinuidades abruptas nos dados de entrada, uma vez que mudanças de temperatura acarretam alterações na cinética da reação da bateria. Assim, cada sequência de entrada respeita a homogeneidade térmica e garante que o modelo aprenda os padrões de SOC de forma consistente.

As variáveis foram normalizadas utilizando a técnica de *Min-Max normalization*, com os seguintes intervalos: Entradas (X): $[-1, 1]$; Saída (Y , SOC): $[0, 1]$.

Os parâmetros de normalização foram obtidos exclusivamente a partir do conjunto de treinamento, sendo posteriormente aplicados ao conjunto de validação e teste para evitar vazamento de dados (data leakage). Além disso, foi adicionada uma variável de contextualização do passado: a média móvel da tensão, calculada com uma subjanela interna de 5 amostras. Esta variável adicional contribui para estabilizar flutuações locais e facilitar a detecção de tendências do SOC, conforme relatado em [14].

C. Arquitetura da Rede LSTM e Estratégia de Treinamento

Para estimar o SOC de baterias de íons de lítio a partir de séries temporais multivariadas, foi implementada uma rede neural recorrente do tipo Long Short-Term Memory (LSTM). A entrada do modelo tem formato tridimensional ($N_{amostras}, N_{janelas}, N_{variáveis}$), onde cada janela de entrada contém 50 amostras sucessivas de quatro variáveis: corrente (A), tensão (V), temperatura (°C) e média móvel da tensão (V). A escolha por uma única camada LSTM com 30 unidades ocultas se baseia no equilíbrio entre capacidade de modelagem temporal e custo computacional reduzido. A saída da LSTM é conectada a uma camada densa com um único neurônio e função de ativação sigmoide, adequada para a previsão de valores normalizados entre 0 e 1, como o SOC. A arquitetura do modelo é representada na Fig. 5, esse tipo de arquitetura é eficaz em problemas com dependência sequencial de longo prazo, como aqueles associados à resposta dinâmica de baterias submetidas a diferentes perfis de carga e descarga [16], [30].

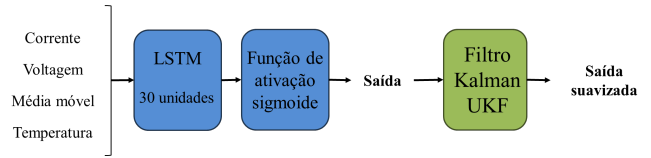


Fig. 5. Arquitetura simplificada da rede usada no treinamento do modelo

Antes do treinamento, os dados foram embaralhados aleatoriamente. Essa aleatorização foi realizada com o intuito de evitar que a ordem dos ciclos experimentais, que variam por temperatura e padrão de estresse, influenciasse negativamente no aprendizado do modelo. Após o embaralhamento, os dados foram divididos em dois subconjuntos, sendo 80% destinados ao treinamento e 20% reservados para teste.

O treinamento é conduzido por até 150 épocas, com tamanho de lote (batch size) igual a 64. Foi utilizada a função de perda do tipo erro quadrático médio (MSE), apropriada para tarefas de regressão, e o otimizador adotado foi o Adam, que é robusto em problemas com gradientes não estacionários e dados ruidosos.

Para prevenir o sobreajuste e reduzir o tempo computacional, foi adotada uma estratégia de parada antecipada (early stopping), monitorando a perda no conjunto de treinamento. O critério estabelecido foi de paciência de 10 épocas consecutivas sem melhora, sendo que, ao final do processo, os melhores pesos observados durante o treinamento são automaticamente restaurados. Essa abordagem contribuiu para a estabilidade do treinamento e para a melhoria da generalização do modelo em dados não vistos.

A definição da arquitetura e dos hiperparâmetros foi realizada com base em testes exploratórios e fundamentada na literatura recente sobre o uso de redes LSTM em problemas de estimativa de SOC, como demonstrado em [14] e [30].

D. Pós-processamento – Suavização das Predições com Filtro de Kalman Não Linear (UKF)

Com o objetivo de aprimorar a estabilidade e a qualidade das predições do estado de carga (SOC) geradas pela rede LSTM, foi realizada uma etapa de pós-processamento utilizando um Filtro de Kalman Não Linear do tipo Unscented (UKF). A aplicação deste filtro visa reduzir oscilações locais, corrigir desvios pontuais e proporcionar uma curva de SOC mais contínua e condizente com a dinâmica real da bateria.

O UKF é adequado para cenários em que tanto o modelo de evolução do estado quanto o modelo de observação podem apresentar não linearidades [30]. No presente caso, embora o modelo de evolução tenha sido assumido como linear e estacionário, a natureza das predições provenientes de uma rede neural justifica a escolha de um filtro mais robusto.

O modelo de estado adotado neste trabalho considera que o valor real do SOC evolui suavemente ao longo do tempo. Essa hipótese foi inicialmente representada por um modelo determinístico e estacionário, onde o estado atual é igual ao estado anterior, ou seja, $x_k = x_{k-1}$, sendo x_k o SOC no instante k . A função de medição foi modelada como uma observação direta do estado, expressa como $z_k = x_k$, em que z_k corresponde à predição fornecida pela rede LSTM.

Para refletir a presença de incertezas tanto no modelo quanto nas observações, essas equações foram generalizadas segundo a formulação estocástica clássica do Filtro de Kalman. Assim, o modelo de estado passou a ser representado por $x_k = x_{k-1} + w_k$ e o modelo de observação por $z_k = x_k + v_k$, onde w_k representa o ruído do processo e v_k representa o ruído de medição. Ambos são assumidos como ruídos aleatórios com distribuição normal e média zero, com covariâncias $Q = 10^{-4}$ e $R = 10^{-2}$, respectivamente.

Com base nessa modelagem, o UKF realiza a estimativa corrigida do estado a partir da geração de pontos sigma que capturam a distribuição do SOC. Esses pontos são propagados pelo modelo e utilizados para calcular uma média ponderada que incorpora a incerteza associada ao processo e à observação. A aplicação do filtro foi feita de forma causal, ou seja, considerando apenas informações atuais e passadas, o que assegura sua viabilidade para implementações embarcadas e aplicações em tempo real.

E. Métricas de Avaliação

A avaliação da performance dos modelos de predição de SOC foi realizada por meio de métricas estatísticas típicas de regressão contínua. A primeira métrica, o erro quadrático médio (RMSE), calcula a raiz da média dos quadrados das diferenças entre valores previstos e reais, sendo sensível a grandes desvios e útil para medir a precisão absoluta das estimativas [34].

O erro absoluto médio (MAE), por sua vez, representa a média dos desvios absolutos e oferece uma visão mais equilibrada dos erros ao longo do tempo, sendo menos influenciado por outliers [34].

Também foi utilizado o erro absoluto máximo (MAXE), que indica o maior erro individual entre previsão e valor real,

sendo crítico para aplicações embarcadas ou de controle, onde desvios pontuais podem comprometer a segurança [34].

Por fim, o coeficiente de determinação (R^2) quantifica o quanto da variabilidade dos dados reais é explicado pelo modelo. Valores próximos de 1 indicam boa aderência entre a predição e os dados [34].

Essas métricas foram aplicadas tanto às predições brutas da LSTM quanto às estimativas suavizadas pelo UKF, permitindo comparar a acurácia direta da rede com os ganhos em estabilidade e confiabilidade trazidos pela suavização.

IV. RESULTADOS E DISCUSSÕES

Como mencionado na metodologia, o treinamento da rede LSTM descrita foi realizado apenas com dados dos perfis de condução DST (Dynamic Stress Test) para temperaturas de 10°C, 20°C, 30°C e 40°C, concatenadas em um único conjunto de dados. A avaliação do modelo LSTM-UKF treinado se deu em dados de US06 e FUDS, permitindo avaliar a capacidade de generalização do modelo frente a diferentes perfis de condução veicular.

O treinamento e teste foram realizados no Google Colab, em ambiente Linux, com CPU Intel Xeon @ 2.20 GHz (1 núcleo, 2 threads) e 12 GB de RAM. O algoritmo foi implementado em Python e executado no ambiente Google Colab. O código-fonte está disponível publicamente em [35] e acessível em: <https://colab.research.google.com/drive/19FA2liTqbwNn6UNoWx4a1lCzYRKQaDfL?usp=sharing>. Para a configuração de hardware utilizada, os tempos de processamento foram os seguintes:

- Tempo de predição total: 0,189 ms
- Tempo de suavização total: 0,273 ms
- Tempo total: 0,427 ms

Segundo [31] e [32], a resposta para diagnóstico de falhas em veículos elétricos deve ocorrer na ordem de milissegundos por evento, especialmente para eventos críticos como thermal runaway. Assim, o método proposto é considerado viável para implementações em tempo real em plataformas, gêmeos digitais e sistemas cyber-físicos, como BMS online.

A. Treinamento

O treinamento com os dados de DST apresentou uma rápida convergência, alcançando seu melhor desempenho na época 128 de 150. As perdas observadas durante o treinamento estão apresentadas na Fig. 6. Ao avaliar o modelo com dados de teste do ciclo DST, os resultados obtidos foram: RMSE de 0,7%, MAXE de 3,4% e R^2 de 99,9%. Esses valores demonstram que o modelo foi capaz de aprender a dinâmica da bateria sob diferentes condições de temperatura.

B. Teste e Generalização

Para validar o poder de generalização do modelo LSTM-UKF durante as predições, utilizou-se perfis de condução FUDS e US06 em diferentes temperaturas. Esses perfis, significativamente distintos do DST, não foram vistos durante o treinamento. De modo a avaliar o desempenho da rede LSTM e o desempenho do filtro UKF as métricas de avaliação

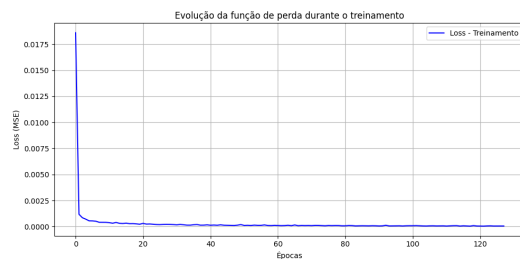


Fig. 6. Perdas durante o treinamento do modelo LSTM

serão mostradas de forma separada: inicialmente com a rede neural LSTM isolada e, em seguida, com o modelo composto por LSTM e o filtro UKF na saída (LSTM-UKF), sendo demonstrados em Tabela I e Tabela II.

Tabela I
DESEMPENHO DA LSTM E LSTM-UKF NO CICLO FUDS

Métrica	10°C	20°C	30°C	40°C	Média
RMSE LSTM	2,7%	1,9%	1,7%	1,8%	2,0%
RMSE LSTM-UKF	2,5%	1,6%	1,2%	1,1%	1,6%
MAXE LSTM	9,8%	8,5%	6,9%	7,2%	8,1%
MAXE LSTM-UKF	6,5%	5,1%	3,7%	3,1%	4,6%
R ² LSTM	98,8%	99,4%	99,5%	99,5%	99,3%
R ² LSTM-UKF	99,0%	99,6%	99,7%	99,8%	99,5%

Tabela II
DESEMPENHO DA LSTM E LSTM-UKF NO CICLO US06

Métrica	10°C	20°C	30°C	40°C	Média
RMSE LSTM	2,6%	2,3%	2,0%	2,0%	2,2%
RMSE LSTM-UKF	1,5%	0,9%	0,8%	0,9%	1,0%
MAXE LSTM	9,3%	8,7%	7,5%	7,2%	8,2%
MAXE LSTM-UKF	4,5%	3,2%	2,5%	2,6%	3,2%
R ² LSTM	98,9%	99,2%	99,4%	99,3%	99,2%
R ² LSTM-UKF	99,7%	99,9%	99,9%	99,9%	99,8%

As tabelas demonstram os resultados de teste para o ciclo FUDS e US06 respectivamente. A análise delas evidencia que o modelo treinado apresentou bom desempenho, mesmo em ciclos de condução não vistos anteriormente. O RMSE manteve-se abaixo de 2,2% e o R² acima de 99% na maioria dos casos, embora o MAXE ainda tenha se mantido em torno de 8%, indicando oscilações localizadas. Essas oscilações são típicas em saídas de redes neurais [14], principalmente na região de SOC intermediário, onde a não linearidade da bateria é mais intensa, conforme mostrado na Fig. 7.

1) *Suavização com UKF*: Tratando as oscilações como ruídos, aplicou-se um filtro UKF para suavizar as previsões. A aplicação do UKF resultou em melhorias claras nas métricas de desempenho, conforme pode ser visto na Tabela III.

A aplicação do UKF, que lida melhor com dinâmicas não lineares do sistema, permitiu reduções significativas no MAXE e aumentos no R², comprovando a efetividade da abordagem. A Fig. 8 demonstra a menor oscilação dos valores previstos para SOC, principalmente na região de SOC médio.

Tabela III
DESEMPENHO APÓS APLICAÇÃO DO UKF (VALORES MÉDIOS)

Ciclo	RMSE	MAXE	R ²
FUDS	1,6%	4,6%	99,5%
US06	1,0%	3,2%	99,8%

C. Predição em Temperatura Não Treinada (25°C)

Para validar ainda mais a capacidade de generalização, o modelo foi testado com dados a 25°C, temperatura não utilizada no treinamento, como pode ser visto na Tabela IV.

Tabela IV
DESEMPENHO A 25°C

Métrica	FUDS	US06
RMSE LSTM	2,5%	3,1%
RMSE LSTM-UKF	1,0%	1,6%
MAXE LSTM	9,1%	11,5%
MAXE LSTM-UKF	2,8%	5,1%
R ² LSTM	99,0%	98,5%
R ² LSTM-UKF	99,8%	99,6%

Mesmo sem treinamento específico para esta temperatura, o modelo apresentou bom desempenho, reforçando sua capacidade de generalizar para diferentes condições.

D. Comparação de abordagens

Foram comparados os resultados experimentais do LSTM-UKF com outros métodos de estimativa de SOC com a finalidade de comprovar seu bom desempenho. Para seguir de forma mais rigorosa, foram selecionados dois modelos para comparação, NN-UKF [6] e LSTM-ACKF [36] são modelos construídos usando os mesmos conjuntos de dados de treinamento e teste. Especificamente, esses dois modelos são todos treinados usando os conjuntos de dados DST em temperaturas diferentes do conjunto de dados público do CALCE e testados usando os conjuntos de dados US06 e FUDS. Os resultados da comparação são mostrados na Tabela V.

Tabela V
DESEMPENHO COMPARATIVO: LSTM-UKF, NN-UKF [6] E LSTM-ACKF [36]

Ciclo	T (°C)	RMSE LSTM	RMSE NN	RMSE ACKF	MAXE LSTM	MAXE NN	MAXE ACKF
US06	10	1,50	2,10	0,80	4,50	2,70	2,20
US06	20	0,90	0,90	—	3,20	1,30	—
US06	30	0,80	1,30	0,90	2,50	1,90	1,70
US06	40	0,90	1,00	1,20	2,60	1,50	2,70
US06	Méd.	1,00	1,33	0,97	3,20	1,85	2,20
FUDS	10	2,50	0,50	1,60	6,50	1,40	3,50
FUDS	20	1,60	0,50	—	5,10	1,60	—
FUDS	30	1,20	0,90	1,20	3,70	1,30	2,90
FUDS	40	1,10	0,60	1,80	3,10	1,00	3,20
FUDS	Méd.	1,60	0,63	1,53	4,60	1,33	3,20

Ao comparar o RMSE do modelo LSTM-UKF com o LSTM-ACKF é possível verificar que ambos apresentaram desempenho próximos em ambos os ciclos, com leve superioridade do LSTM-ACKF. Ao comparar com o NN-UKF, é

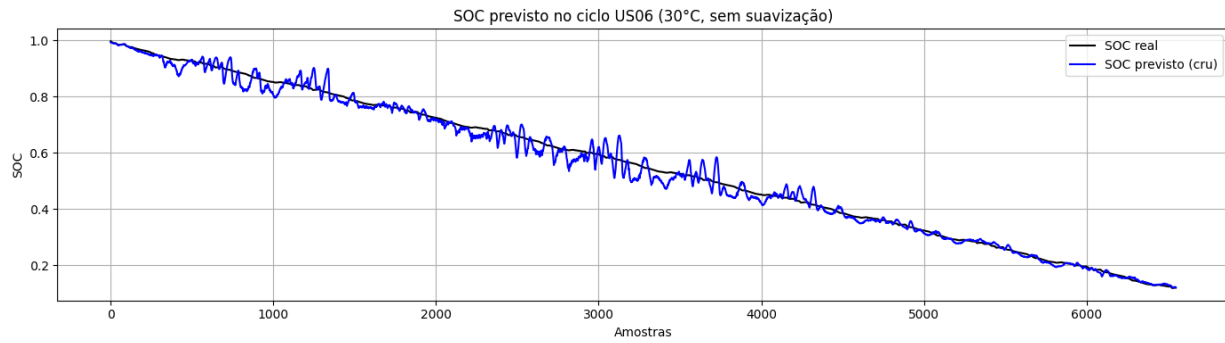


Fig. 7. SOC real e previsto pelo modelo para um ciclo US06 a 30°C, sem suavização

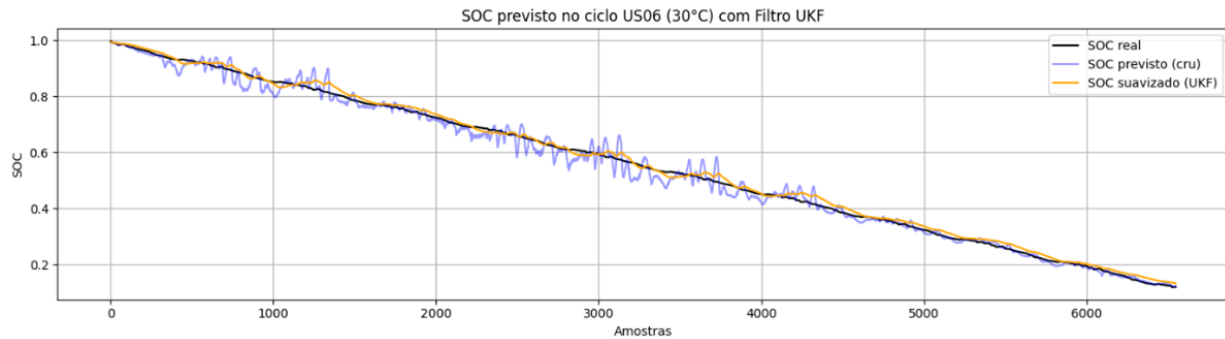


Fig. 8. SOC real e previsto pelo modelo para um ciclo US06 a 30°C, com suavização

possível verificar que o LSTM-UKF apresentou desempenho inferior no ciclo FUDS e superior no ciclo US06. Em ambos os modelos de comparação o MAXE foi inferior, sugerindo que ajustes adicionais no filtro UKF adotado podem melhorar ainda mais o desempenho. Em linhas gerais, o modelo LSTM-UKF proposto apresentou desempenho semelhante a outros comparados, de modo que apresenta-se como uma alternativa viável para implementação prática.

V. CONCLUSÃO

Este trabalho propôs um método híbrido para a estimativa do SOC de baterias de íons de lítio, combinando redes neurais LSTM com pós-processamento via filtro UKF. A abordagem demonstrou boa capacidade de generalização, mesmo quando aplicada a ciclos de condução e condições térmicas não vistas durante o treinamento. A aplicação do UKF como etapa de suavização foi de grande importância para minimizar flutuações locais nas previsões da LSTM, resultando em estimativas mais estáveis e mais próximas à dinâmica real da bateria. Os resultados mostraram que o modelo atingiu RMSE inferior a 2,2% e R^2 superior a 99% nos testes com ciclos FUDS e US06, validando sua precisão e robustez. Tais resultados estão alinhados ao desempenho de modelos desenvolvidos na literatura recente.

A adaptação do método a temperaturas intermediárias não incluídas no treinamento reforçou sua capacidade de extrapolação prática, aspecto fundamental para aplicações em veículos elétricos operando sob condições dinâmicas variadas.

Além da acurácia, o tempo de execução do modelo, inferior a 0,5 ms, indica viabilidade para implementação em sistemas embarcados, plataformas de gêmeos digitais e sistemas cyber-físicos para inferência em tempo real.

Conclui-se que a combinação de redes neurais LSTM com técnicas de filtragem como o UKF representa uma solução equilibrada entre precisão, robustez e viabilidade computacional para a estimativa de SOC em baterias de veículos elétricos. Trabalhos futuros poderão explorar estratégias de adaptação online do modelo e integração com plataformas de edge computing para aprimorar a resposta em ambientes operacionais complexos, e novas abordagens de ajuste fino do filtro UKF podem gerar um desempenho ainda melhor.

REFERÊNCIAS

- [1] International Energy Agency, "Global EV Outlook 2020: Entering the decade of electric drive?," IEA, Paris, 2020. [Online]. Available: <https://www.iea.org/reports/global-ev-outlook-2020>
- [2] K. Luo, X. Chen, H. Zheng, and Z. Shi, "A review of deep learning approach to predicting the state of health and state of charge of lithium-ion batteries," *J. Energy Chem.*, vol. 74, pp. 159–173, 2022, doi: 10.1016/j.jechem.2022.06.049.
- [3] A. Jose and S. Shrivastava, "Evolution of Electrical Vehicles, Battery State Estimation, and Future Research Directions: A Critical Review," *IEEE Access*, 2024.
- [4] B. Yang, J. Wang, P. Cao, T. Zhu, H. Shu, J. Chen, J. Zhang, and J. Zhu, "Classification, summarization and perspectives on state-of-charge estimation of lithium-ion batteries used in electric vehicles: A critical comprehensive survey," *J. Energy Storage*, vol. 39, 2021, Art. no. 102572, doi: 10.1016/j.est.2021.102572.

- [5] K. Zhao, Y. Liu, W. Ming, Y. Zhou, and J. Wu, "Digital Twin-Driven Estimation of State of Charge for Li-ion Battery," in Proc. IEEE Int. Energy Conf. (ENERGYCON), Riga, Latvia, 2022, pp. 1–6, doi: 10.1109/ENERGYCON53164.2022.9830324.
- [6] W. He, N. Williard, C. Chen, and M. Pecht, "State of charge estimation for Li-ion batteries using neural network modeling and unscented Kalman filter-based error cancellation," *Int. J. Electr. Power Energy Syst.*, vol. 62, pp. 783–791, 2014, doi: 10.1016/j.ijepes.2014.04.059.
- [7] M. A. Roscher and D. U. Sauer, "Dynamic electric behavior and open-circuit-voltage modeling of LiFePO₄-based lithium ion secondary batteries," *J. Power Sources*, vol. 196, no. 1, pp. 331–336, 2011, doi: 10.1016/j.jpowsour.2010.06.098.
- [8] K. S. Ng, C. S. Moo, Y. P. Chen, and Y. C. Hsieh, "Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries," *Appl. Energy*, vol. 86, no. 9, pp. 1506–1511, 2009, doi: 10.1016/j.apenergy.2008.11.021.
- [9] D. Darbar and I. Bhattacharya, "Application of Machine Learning in Battery: State of Charge Estimation Using Feed Forward Neural Network for Sodium-Ion Battery," *Electrochem*, vol. 3, no. 1, pp. 42–57, 2022, doi: 10.3390/electrochem3010003.
- [10] W. Waag, C. Fleischer, and D. U. Sauer, "Critical review of the methods for monitoring of lithium-ion batteries in electric and hybrid vehicles," *J. Power Sources*, vol. 258, pp. 321–339, 2014, doi: 10.1016/j.jpowsour.2014.02.064.
- [11] J. C. Alvarez Anton, P. J. Garcia Nieto, C. Blanco Viejo, and J. A. Vilan Vilan, "Support vector machines used to estimate the battery state of charge," *IEEE Trans. Power Electron.*, vol. 28, no. 12, pp. 5919–5926, 2013.
- [12] M. Ismail, R. Dlyma, A. Elrakaybi, R. Ahmed, and S. Habibi, "Battery state of charge estimation using an Artificial Neural Network," in Proc. IEEE Transp. Electrific. Conf. Expo (ITEC), Chicago, IL, USA, 2017, pp. 342–349, doi: 10.1109/ITEC.2017.7993295.
- [13] M. Wu, M. Sun, F. Zhang, L. Wang, N. Zhao, J. Wang, and W. Huang, "A fault detection method of electric vehicle battery through Hausdorff distance and modified Z-score for real-world data," *J. Energy Storage*, vol. 60, p. 106561, 2023, doi: 10.1016/j.est.2022.106561.
- [14] J. Chen, Y. Zhang, J. Wu, W. Cheng, and Q. Zhu, "SOC estimation for lithium-ion battery using the LSTM-RNN with extended input and constrained output," *Energy*, vol. 262, p. 125375, 2023, doi: 10.1016/j.energy.2022.125375.
- [15] M. Hossain Lipu, M. Hannan, A. Hussain, A. Ayob, M. H. Saad, T. F. Karim, and D. N. How, "Data-driven state of charge estimation of lithium-ion batteries: Algorithms, implementation factors, limitations and future trends," *J. Clean. Prod.*, vol. 277, Art. no. 124110, 2020, doi: 10.1016/j.jclepro.2020.124110.
- [16] A. Manoharan, K. Begam, V. R. Aparow, and D. Sooriamoorthy, "Artificial neural networks, gradient boosting and support vector machines for electric vehicle battery state estimation: A review," *J. Energy Storage*, vol. 55, Art. no. 105384, 2022, doi: 10.1016/j.est.2022.105384.
- [17] I. Babaeiyazdi, A. Rezaei-Zare, and S. Shokrzadeh, "State of charge prediction of EV Li-ion batteries using EIS: A machine learning approach," *Energy*, vol. 223, Art. no. 120116, 2021, doi: 10.1016/j.energy.2021.120116.
- [18] X. Mao, S. Song, and F. Ding, "Optimal BP neural network algorithm for state of charge estimation of lithium-ion battery using PSO with Levy flight," *J. Energy Storage*, vol. 49, Art. no. 104139, 2022, doi: 10.1016/j.est.2022.104139.
- [19] G. Zhang, B. Xia, J. Wang, B. Ye, Y. Chen, Z. Yu, and Y. Li, "Intelligent state of charge estimation of battery pack based on particle swarm optimization algorithm improved radical basis function neural network," *J. Energy Storage*, vol. 50, Art. no. 104211, 2022, doi: 10.1016/j.est.2022.104211.
- [20] X. Feng, J. Chen, Z. Zhang, S. Miao, and Q. Zhu, "State-of-charge estimation of lithium-ion battery based on clockwork recurrent neural network," *Energy*, vol. 236, Art. no. 121360, 2021, doi: 10.1016/j.energy.2021.121360.
- [21] X. Ren, S. Liu, X. Yu, and X. Dong, "A method for state-of-charge estimation of lithium-ion batteries based on PSO-LSTM," *Energy*, vol. 234, Art. no. 121236, 2021, doi: 10.1016/j.energy.2021.121236.
- [22] E. Almaita, S. Alshkoor, E. Abdelsalam, and F. Almomani, "State of charge estimation for a group of lithium-ion batteries using long short-term memory neural network," *J. Energy Storage*, vol. 52, Art. no. 104761, 2022, doi: 10.1016/j.est.2022.104761.
- [23] M. Jiao, D. Wang, and J. Qiu, "A GRU-RNN based momentum optimized algorithm for SOC estimation," *J. Power Sources*, vol. 459, Art. no. 228051, 2020, doi: 10.1016/j.jpowsour.2020.228051.
- [24] S. El Fallah et al., "Advanced State of Charge Estimation Using Deep Neural Network, Gated Recurrent Unit, and Long Short-Term Memory Models for Lithium-Ion Batteries under Aging and Temperature Conditions," *Appl. Sci.*, vol. 14, no. 15, p. 6648, 2024, doi: 10.3390/app14156648.
- [25] Y. Zhang, L. Wang, and C. Xu, "Big data driven vehicle battery management method: A novel cyber-physical system approach," *J. Energy Storage*, vol. 32, p. 101972, 2020.
- [26] H. Wang, Y. Gao, H. Peng, and S. Li, "Development of self-adaptive digital twin for battery monitoring and management system," *J. Power Sources*, vol. 612, p. 228484, 2024.
- [27] Y. Tang, Y. Yu, Y. Wang, and W. Li, "Digital twin and cloud-side-end collaboration for intelligent battery management," *J. Saf. Environ.*, vol. 42, pp. 53–63, 2021.
- [28] CALCE Battery Research Group, "Battery Data Sets," University of Maryland. [Online]. Available: <https://calce.umd.edu/battery-data>. [Accessed: 24-Apr-2025].
- [29] J. P. Christophersen, "Battery Test Manual for Electric Vehicles, Revision 3," Idaho National Laboratory, INL/EXT-15-34184, Jun. 2015.
- [30] G. Chen, W. Peng, and F. Yang, "An LSTM-SA model for SOC estimation of lithium-ion batteries under various temperatures and aging levels," *J. Energy Storage*, vol. 84, 2024, Art. no. 110906, doi: 10.1016/j.est.2024.110906.
- [31] Z. Sun, Z. Wang, P. Liu, Z. Qin, Y. Chen, Y. Han, P. Wang, and P. Bauer, "An Online Data-Driven Fault Diagnosis and Thermal Runaway Early Warning for Electric Vehicle Batteries," *IEEE Transactions on Power Electronics*, vol. 37, no. 10, pp. 12636–12648, Oct. 2022, doi: 10.1109/TPEL.2022.3173038.
- [32] J. Zhao, X. Feng, M.-K. Tran, M. Fowler, M. Ouyang, and A. F. Burke, "Battery safety: Fault diagnosis from laboratory to real world," *Journal of Power Sources*, vol. 598, Art. no. 234111, 2024, doi: 10.1016/j.jpowsour.2024.234111.
- [33] D. C. Montgomery and G. C. Runger, *Applied Statistics and Probability for Engineers*, 7th ed. Hoboken, NJ, USA: Wiley, 2018.
- [34] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York, NY, USA: Springer, 2009.
- [35] G. Mantioli, "SOC Estimation with LSTM and UKF – Google Colab," Google Colaboratory, 2024. [Online]. Available: <https://colab.research.google.com/drive/19FA2liTqbwNn6UNoWx4a1lCzYRKQaDfL?usp=sharing>.
- [36] Y. Tian, R. Lai, X. Li, L. Xiang, and J. Tian, "A combined method for state-of-charge estimation for lithium-ion batteries using a long short-term memory network and an adaptive cubature Kalman filter," *Applied Energy*, vol. 265, p. 114789, Mar. 2020, doi: 10.1016/j.apenergy.2020.114789.