

SOH Estimation of Lithium-Ion Batteries Based on ANNs and Synthetic Data

J. Cabrera-García * J. A. Zepeda-Hernández *
M. de J. López-Pérez * J. E. Domínguez-Zenteno *

* *Tecnológico Nacional de México, I.T. Tuxtla Gutiérrez,
TURIX-Dynamics Diagnosis and Control Group, Carretera
Panamericana km 1080 S/N, Tuxtla Gutiérrez 29050, México;*(e-mail:
*M07270046@tuxtla.tecnm.mx, jose.zh@tuxtla.tecnm.mx,
manuel.lp@tuxtla.tecnm.mx, joaquin.dz@tuxtla.tecnm.mx*)

Abstract: Estimating the State Of Health (SOH) of Lithium-Ion Batteries (LIB) is essential for optimal performance in applications such as electric vehicles (EV) and energy storage systems. However, this battery's electrochemical parameter is impossible to measure directly. Many works have addressed estimating the battery SOH using different techniques. Recently, techniques such as machine learning have enabled data-driven predictions when a parameter cannot be measured. In this work, an artificial neuronal network (ANN) is proposed to estimate the battery's SOH for an EV. The ANN was trained at the beginning with Ames PCoE battery data provided by NASA and complemented with synthetic data from a LIB electrochemical model. The electrochemical model describes the battery behavior with specific SOH. Therefore, a parameter estimation algorithm such as Particle Swarm Optimization (PSO) was used to figure out the model parameters for each SOH to generate new profiles and add more information to train the ANN. Finally, the ANN was implemented with a candidate battery to demonstrate its effectiveness and performance in out-range LIB operating conditions in which it was trained.

Keywords: SOH, ANN, EV, Battery Model, PSO.

1. INTRODUCCIÓN

Actualmente, muchos sectores han incorporado baterías de iones de litio (LIB) en diversas aplicaciones, incluyendo teléfonos inteligentes, vehículos eléctricos (VE) y sistemas de almacenamiento de energía, entre otros (Chen et al., 2020). Esto se debe a que las LIB ofrecen una mayor densidad de energía y una vida útil más prolongada en comparación con otro tipo de baterías. Uno de los parámetros electroquímicos importantes de las LIB es su estado de salud (SOH), que indica la degradación de su capacidad de almacenamiento de energía a lo largo de su vida útil.

El envejecimiento de las LIBs es una de las principales causas de cambios en su SOH, ya que, con el tiempo los materiales internos y la electroquímica de la batería se degradan gradualmente afectando su capacidad. Los principales mecanismos de degradación de las baterías se dividen en tres tipos: pérdida de almacenamiento de litio, pérdida de material activo en el electrodo y aumento de la resistencia interna (Juantorena Polo, 2023). La estimación del SOH es esencial para aplicaciones basadas en LIB, debido a que este parámetro es indicativo de cuánto tiempo más se puede utilizar de forma segura y sin pérdida de potencia. Sin embargo, debido a sus características no lineales se dificulta obtener el valor preciso de este

parámetro. Por lo anterior, y a la falta de sensores, se recurre a métodos de estimación que utilizan otros parámetros electroquímicos medibles, tales como: voltaje, corriente y temperatura. En la literatura se describen diferentes métodos para estimar el SOH, cada uno con sus propias ventajas, desventajas y limitaciones.

La determinación del SOH se puede realizar mediante enfoques diferentes, tales como métodos experimentales, basados en modelos y basados en datos. Los primeros se llevan a cabo en laboratorios y en modo off-line (Benavides Flores, 2018), cuya desventaja principal es el uso de equipo especializado difícil de adquirir. En los segundos, se utilizan diferentes indicadores para diseñar un modelo y describir la degradación de la batería utilizando técnicas como el filtro de Kalman y sus extensiones, observadores, modelo de circuito equivalente, entre otros (Zhu and Fu, 2021). Se ha demostrado que estos métodos basados en modelos estiman con precisión el SOH de las LIB; sin embargo, a menudo se requiere un alto costo computacional y, en algunos casos, un gran número de variables. Por último, los métodos basados en datos realizan la estimación del SOH con los datos de ciclos de carga/descarga y parámetros que afectan la vida útil de la batería. En estos métodos, se utilizan *redes neuronales artificiales* (RNA) para estimar el SOH y se necesita conocer la correlación entre la operación y la degradación

a través del análisis físico o la evaluación de conjuntos de datos (Huang et al., 2017). Sin embargo, cuando el número de datos es limitado, se requiere la expansión del conjunto de datos para mejorar la eficacia. Una forma efectiva de abordar el problema es generando datos de manera sintética, que consiste en la creación de nuevos datos que emulan las características y propiedades de los datos reales, permitiendo aumentar el volumen y la diversidad del conjunto de datos original. Este enfoque es particularmente útil en el entrenamiento de la RNA (Torfi et al., 2022).

Dentro de la literatura, podemos encontrar que Driscoll et al. (2022) utiliza un conjunto de datos (dataset) generado por la *NASA Ames PCoE* (Song, 2019) para realizar el entrenamiento de una RNA y proponen un modelo de estimación de SOH simple a partir de perfiles de voltaje, corriente y temperatura. En ese artículo se utiliza el dataset mencionado y una RNA para realizar estimaciones de SOH por ciclos de carga/descarga de una batería candidata. En la literatura podemos encontrar algunos criterios para la arquitectura de una RNA (Zepeda H et al., 2015). Sin embargo no existe una regla sino criterios, para determinar el número de capas, neuronas y tipo de función de activación de las RNA. La validación de las estimaciones se realiza mediante el *error cuadrático medio* (MSE), cuyo valor se usa para cuantificar la precisión de las mismas.

Zhang et al. (2022) utilizan el mismo conjunto de datos proporcionado por la *NASA* y proponen una *red neuronal recurrente* (RNN). Sin embargo, al aumentar la escala de datos, surge el problema de la desaparición o explosión del gradiente. Para superar esta desventaja, proponen entrenar y experimentar con las siguientes arquitecturas de redes neuronales: *Red de Memoria a Corto y Largo Plazo* (LSTM), *Unidad Recurrente Controlada* (GRU), una combinación de algoritmos como CNN-LSTM y, finalmente, una *Red Neuronal Convolutiva Temporal* (TCN), con el objetivo de estimar el SOH de las LIB. De acuerdo con los valores obtenidos en la evaluación del MSE, la red TCN ofrece las mejores estimaciones.

Li et al. (2020) utilizan los mismos datos de los trabajos anteriores, sin embargo utilizan la combinación de NN de tipo LSTM junto con una *Estructura de Datos Árbol Sintético Abstracto* (AST). Evalúan el rendimiento de la red entrenada con la *raíz del error cuadrático medio* (RMSE).

En el análisis de los diferentes trabajos previos para estimar el SOH, se observa que los métodos basados en datos no requieren de un modelo para su análisis, sin embargo se utilizan datos históricos, como parámetros medibles durante su ciclo de carga/descarga. Además, se tiene la limitante de las condiciones a las cuales se obtuvieron dichos datos.

1.1 Motivación

La aplicación de LIBs en vehículos se ha incrementado en los últimos años y con ello la necesidad de estimar el SOH de las mismas. De los trabajos mencionados que utilizan el dataset de la NASA, se tiene la limitante que

ésta sólo muestra el envejecimiento de las baterías bajo condiciones constantes (descarga de 2A). Sin embargo, es bien conocido que en aplicaciones automotrices, la descarga de las baterías no se realiza de manera constante y controlada. Por tal motivo, el uso de sólo esos datos reducirían la capacidad de estimación del SOH en condiciones diferentes a la del dataset. Por ello, para solventar esta limitante, se propone utilizar un conjunto de datos extendidos mediante el uso de datos generados sintéticamente a través de un modelo electroquímico. Cabe mencionar que la elaboración de dichos datos se realizaron off-line para su posterior aplicación a la RNA desarrollado en este trabajo. La técnica para la generación de los datos sintéticos fue de optimización por enjambre de partículas (PSO) debido a su simplicidad y rápida convergencia a los valores deseados (Wang et al., 2018). Este algoritmo entrega los parámetros del modelo de la LIB y son introducidos en un modelo de MATLAB (v. R2022a) que genera datos sintéticos en cada ciclo. Adicionalmente, se emplea una red neuronal BiLSTM con pocas épocas de entrenamiento para converger a los valores, lo que evita el sobre costo computacional, ya que el peso computacional se realiza fuera de línea y la estimación es realmente sencilla.

2. METODOLOGÍA

El SOH indica la degradación de una batería a lo largo de su vida útil. La definición más ampliamente utilizada del SOH se basa en la capacidad de la batería, y se define como:

$$SOH = \frac{Cap_{actual}}{Cap_{nominal}} 100\% \quad (1)$$

Donde: Cap_{actual} es la capacidad medida del estado actual de la batería y $Cap_{nominal}$ es la capacidad máxima de la batería posterior a su fabricación (nueva). Debido a los constantes ciclos de carga y descarga cuando el SOH de la batería cae por debajo del 80%, es recomendable reemplazarla, ya que el rendimiento y la confiabilidad disminuyen considerablemente (Driscoll et al., 2022). La premisa fundamental para estimar el SOH en los métodos basados en datos reside en la detección de variaciones en el voltaje de descarga, corriente de descarga, temperatura generada a partir de la descarga y el tiempo que lleva en descargarse la LIB, mediante una RNA reside en la detección de variaciones en el voltaje, corriente y temperatura en un periodo de tiempo que dura su descarga, permitiendo inferir el grado de degradación de la batería. Por lo cual, se propone entrenar una RNA para estimar el SOH en cada ciclo, lo cual implica la necesidad de contar con datos históricos específicos para entrenar la red. Dentro de estos datos, es necesario que se incluya la capacidad medida (Saha and Goebel, 2007). El dataset de la NASA incluye los datos históricos de 6 baterías evaluadas. Para este trabajo se considero aquella con nombre B0006 con características electroquímicas fáciles de adquirir en el mercado local. Ésta tiene una capacidad de 2 Ah y un voltaje nominal de 4 V. La información incluida en el dataset muestra que la batería fue sometida por 168 ciclos de carga/descarga. El ciclo de carga fue a una corriente constante de 1.5 A hasta alcanzar el voltaje de 4.2 V (ver Fig. 1). Mientras que para el ciclo de descarga, ésta se

descarga a una corriente de 2 A hasta que el voltaje cayó a 2.5 V (ver Fig. 2).

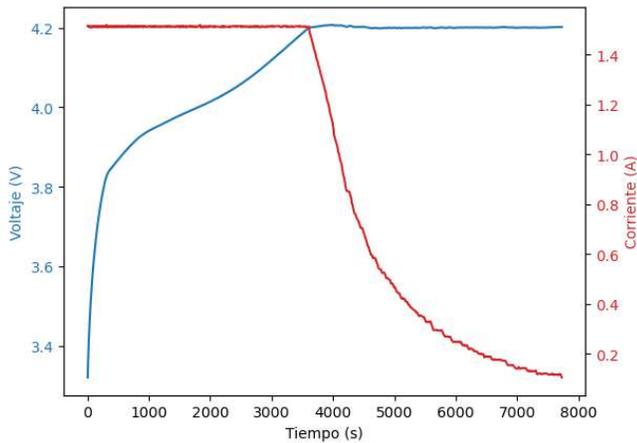


Fig. 1. Carga a corriente constante por ciclo.

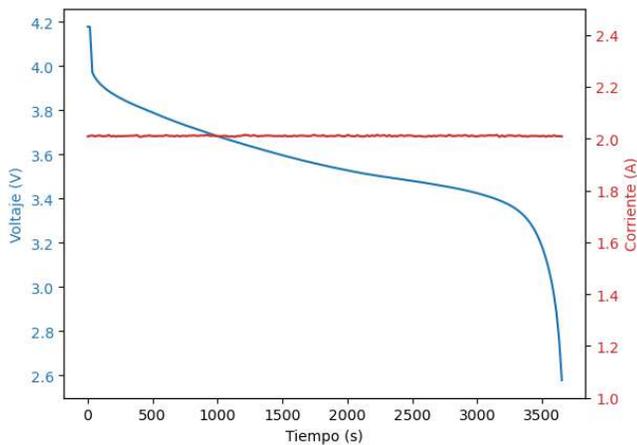


Fig. 2. Descarga a corriente constante por ciclo.

2.1 Generación de datos sintéticos

Desafortunadamente, el dataset de la NASA sólo proporciona información sobre la descarga de la batería a una única condición, una corriente constante de 2 A. Esta condición puede limitar la estimación del SOH por parte de la RNA, pues ésta última no cuenta con información suficiente para realizar la estimación en condiciones distintas a la tasa de descarga de 2A). Con el objetivo de incluir condiciones operativas que no están presentes en el dataset, se propone la generación de datos sintéticos, ya que la obtención de datos reales con diferentes corrientes de descarga implicaría degradar más baterías, lo cual es costoso e poco práctico. Por lo tanto, se propone generar datos sintéticos para corrientes de descarga de 0.5, 1 y 1.5 A. Esto permitirá evaluar el comportamiento de la batería bajo distintos escenarios operativos.

Para la generación de datos sintéticos, se utiliza el modelo electroquímico de una LIB 18650 (Plakhtii et al., 2020).

$$E_{BB} = E_0 - K \frac{Q}{Q - it} I_{BB} - K \frac{Q}{Q - it} it + Ae^{-Bit} \quad (2)$$

Donde:

E_{BB} Voltaje en los bornes de la batería (V)

E_0 Constante del voltaje (V)

K Resistencia de polarización (Ω) o constante de polarización ($\frac{V}{Ah}$)

Q Capacidad máxima de la batería (Ah)

it Capacidad extraída (Ah)

I_{BB} Dinámica de corriente de baja frecuencia (A)

A Exponencial del voltaje (V)

B Exponencial de capacidad (Ah)⁻¹

La ec. (2) representa el comportamiento de la batería acorde a sus parámetros electroquímicos y dependerá del envejecimiento de esta. El voltaje de la batería cambia respecto a la capacidad (energía) sustraída de la misma, la cual es dependiente de la corriente que se le demande. Es decir, la capacidad extraída es la integral de la corriente demandada por unidad de tiempo, la cual se representa en la ecuación siguiente:

$$it = \int_0^t I_{BB} dt \quad (3)$$

En la Fig. 3 se muestra el diagrama de SIMULINK de MATLAB que representa el modelo electroquímico de la batería utilizado para la generación de perfiles (datos sintéticos). Los parámetros por cada SOH fueron generados mediante un algoritmo de estimación por PSO.

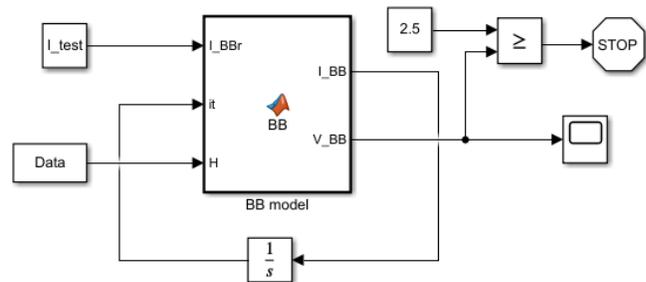


Fig. 3. Modelo de batería 18650.

En cada generación del vector, se validan los parámetros del modelo y los valores estimados por el algoritmo PSO (Fig. 4).

2.2 Red BiLSTM

Las RNA están diseñadas para procesar datos secuenciales o temporales. Sin embargo, cuando se incrementan los datos y el número de neuronas en las capas, estas tienden a sufrir el fenómeno de desvanecimiento del gradiente al propagarse hacia atrás, volviéndose muy pequeños y haciendo que las primeras capas no aprendan (Campos-Taberner et al., 2020). Para solucionar esto, surgen las

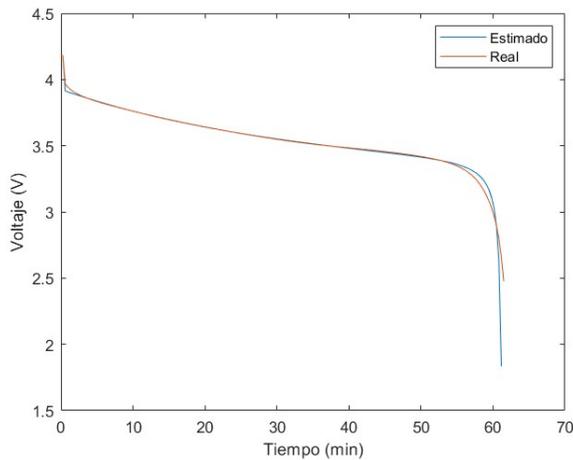


Fig. 4. Correlación del modelo y el PSO.

redes LSTM, que están diseñadas con un módulo que proporciona a la red memoria a corto y largo plazo y supera el problema de desvanecimiento del gradiente. La combinación de dos redes LSTM permite obtener una red BiLSTM. Al introducir datos en la red, es capaz de mantener dependencias temporales en estados pasados y estados futuros, como se muestra en la Fig. 5. La capa de salida se actualiza mediante el cálculo de la secuencia oculta hacia adelante \vec{h}_t , oculta hacia atrás \overleftarrow{h}_t y de salida y_t iterando sobre la capa hacia atrás, la actualización de la red neuronal se formula como:

$$\vec{h}_t = H(W_1 x_t + W_2 \vec{h}_{t-1} + \vec{b}), \quad (4)$$

$$\overleftarrow{h}_t = H(W_3 x_t + W_5 \overleftarrow{h}_{t-1} + \overleftarrow{b}), \quad (5)$$

$$y_t = W_4 \vec{h}_t + W_6 \overleftarrow{h}_t + b_y, \quad (6)$$

donde: H es la función de activación, \vec{h}_t , \overleftarrow{h}_t y y_t son vectores de propagación hacia adelante, hacia atrás y la capa de salida; $W_1, W_2, W_3, W_4, W_5, W_6$ son coeficientes de los pesos correspondientes; $\vec{b}, \overleftarrow{b}, b_y$ son los vectores de sesgo correspondiente.

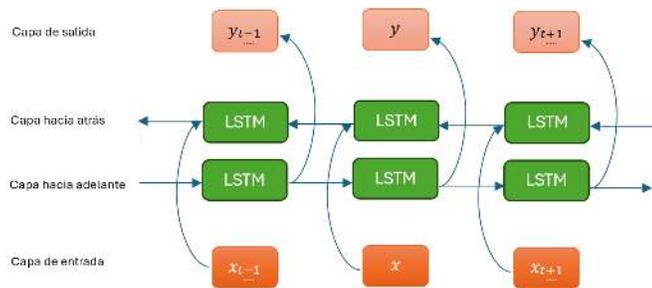


Fig. 5. Estructura de red neuronal BiLSTM.

2.3 Configuración y entrenamiento de la red BiLSTM

Con los datos sintéticos generados se realiza un preprocesamiento de datos para extraer características y normalización, posteriormente se dividen los datos 70% para entrenamiento del modelo y 30% para validación, en la

Tabla 1, se describen los hiperparámetros que se usaron para dicho entrenamiento y en la Fig. 6, se muestra la configuración de la red BiLSTM.

Tabla 1. Hiperparámetros del modelo de red neuronal.

Hiperparámetro	Descripción
Capa de entrada	4 neuronas
Capas LSTM ocultas	4 (capas bidireccionales)
Neuronas por capa LSTM	[139, 115, 74, 16]
Función de activación de las LSTM	Sigmoidal
Capa de salida (Dense)	1
Función de activación de la capa de salida (Dense)	Lineal
Épocas	50

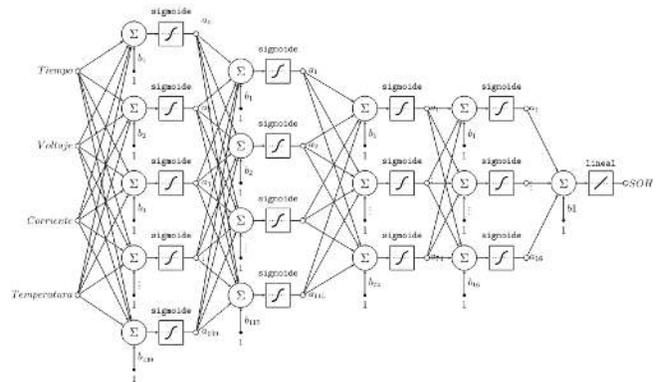


Fig. 6. Configuración de red BiLSTM.

3. RESULTADOS

Una vez que se ha generado los parámetros del modelo electroquímico de la batería en cada uno de los ciclos de descarga, se realiza una comparación de la degradación de la capacidad reportada en el dataset proporcionado por la NASA y la capacidad estimada, (Figura 7).

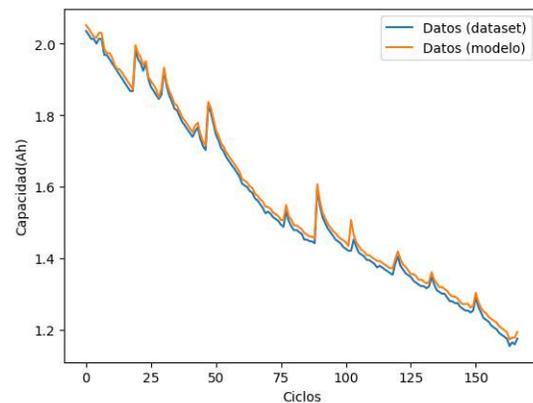
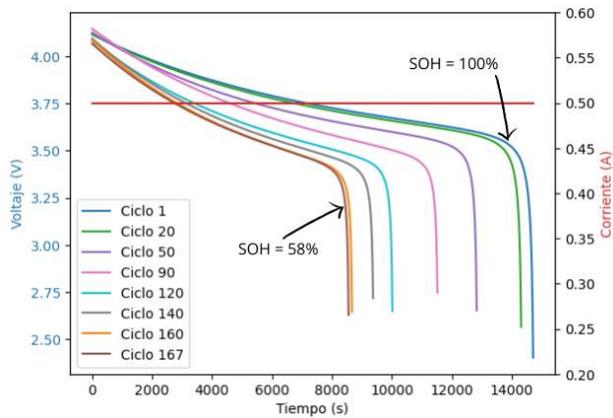


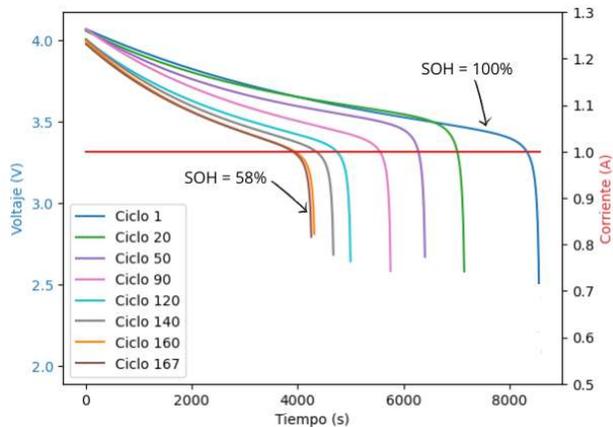
Fig. 7. Degradación de batería B0006, tasa de descarga 2 A.

Se generan 167 ciclos de datos sintéticos de descarga y se almacenan en un archivo .csv para posteriormente entrenar la red neuronal BiLSTM. En la Fig 8 se muestran

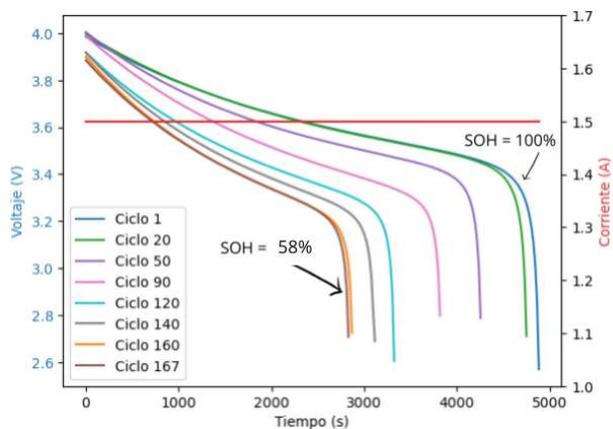
los perfiles de descarga 0.5, 1 y 1.5 A. Conforme el SOH va disminuyendo se observa que la batería pierde capacidad, afectando la potencia y energía.



(a) Descarga constante a 0.5 A



(b) Descarga constante a 1 A



(c) Descarga constante a 1.5 A

Fig. 8. Perfil de descarga sintético

Entrenada la RNA con la configuración mencionada en la tabla 1 se imprimen las métricas de evaluación obtenidas y se presentan en la Tabla 2.

En la evaluación de la RNA entrenada se verifica si la red puede inferir en datos de corriente de descarga que están

Tabla 2. Métricas de evaluación obtenida.

Métrica	valor obtenido
MSE	0.01408
RMSE	0.11866
MAE	0.06913
R^2	0.98588

fuera del dataset proporcionando valores de corriente de la Tabla 3. Los resultados obtenidos se observan en la Fig 9.

Tabla 3. Evaluación de RNA entrenada.

Corriente de descarga	Ciclo	SOH
0.5 A	100	70 %
0.7 A	100	71 %
1 A	100	71 %
1 A	20	97 %
1.3 A	20	97 %
1.5 A	20	97 %
1.5 A	60	81 %
1.8 A	60	81 %
2 A	60	81 %

4. CONCLUSIONES

La demanda variable de corriente en las baterías y la limitación del conjunto de datos de la NASA a una sola tasa de descarga requieren de la expansión del conjunto de datos para estimar el SOH en diferentes condiciones, para lograrlo, se ha implementa un modelo electroquímico de LIB 18650, para generar datos sintéticos en cada ciclo de descarga. Esto proporciona una diversidad de datos que representan diferentes condiciones de operación, permitiendo que la RNA aprenda y generalice de manera efectiva. Además, la generación de datos sintéticos evita la necesidad de degradar físicamente baterías, para obtener datos en una amplia gama de condiciones de descarga. Esto no solo reduce los costos asociados con la adquisición de datos y el mantenimiento de las baterías, sino que también minimiza el riesgo de daños a los recursos físicos y riesgos ambientales.

El empleo de un algoritmo de identificación de parámetros como el PSO es relevante para estimar los parámetros de un modelo electroquímico de la batería en cada ciclo de descarga. Los resultados obtenidos confirman que los datos sintéticos generados a partir del modelo electroquímico destacan a que la red BiLSTM exhiba una notable capacidad de generalización al estimar valores de SOH incluso para datos que no fueron parte del conjunto de entrenamiento.

Cabe mencionar que para este trabajo se hizo uso de una batería LIB 18650, sin embargo la metodología es aplicable para cualquier batería tipo Ion Litio con los ajustes de acuerdo a su capacidad y condiciones de operación.

En resumen, la combinación de algoritmos de identificación de parámetros, datos sintéticos y RNAs representan un avance significativo en la estimación del SOH en LIB, con aplicaciones potenciales en sectores que van desde la electrónica de consumo hasta la movilidad eléctrica y almacenamiento de energía.

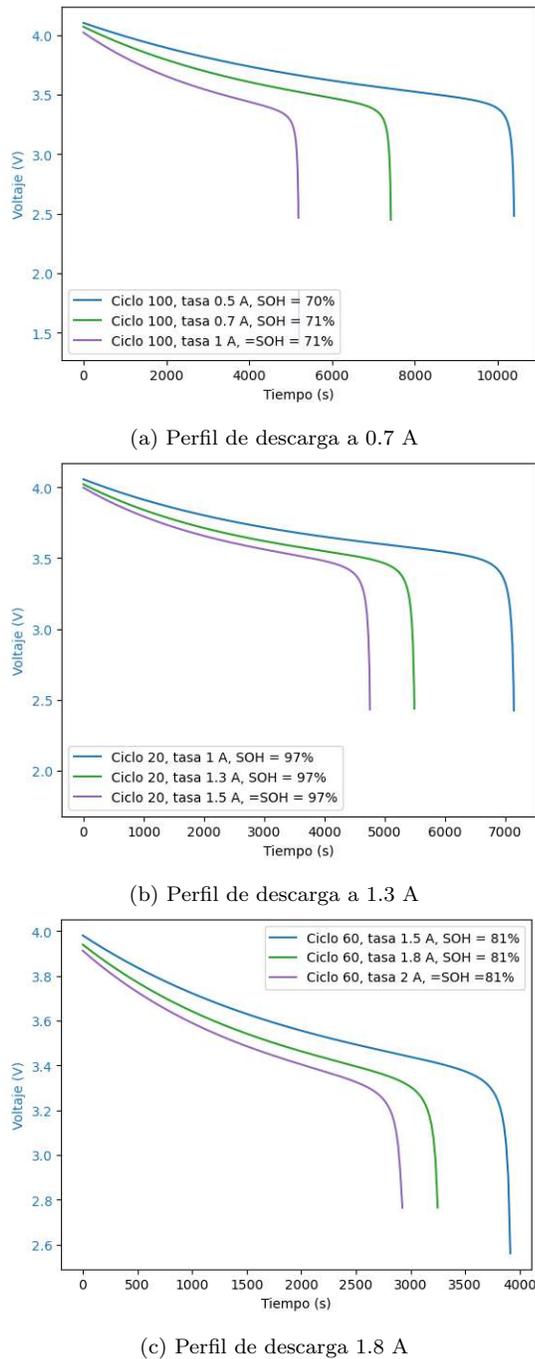


Fig. 9. Estimación de SOH

5. TRABAJOS FUTUROS

Desarrollar un sistema automático de adquisición de datos de descarga de baterías que incluya los sensores adecuados para el monitoreo de parámetros de voltaje, corriente y temperatura durante el proceso de descarga, así como el tiempo requerido a corriente variable. Los datos obtenidos se introducirán a la RNA para la estimación del SOH.

REFERENCIAS

- Benavides Flores, M.I. (2018). Determinación de indicadores de estado de salud de baterías de ion-litio, mediante el uso de espectroscopía de impedancia electroquímica y caracterización del término de la vida útil basado en la generación de calor dada por la impedancia interna.
- Campos-Taberner, M., García-Haro, F.J., Martínez, B., and Gilabert, M. (2020). Deep learning para la clasificación de usos de suelo agrícola con sentinel-2. *Revista de Teledetección*, (56), 35–48.
- Chen, T., Jin, Y., Lv, H., Yang, A., Liu, M., Chen, B., Xie, Y., and Chen, Q. (2020). Applications of lithium-ion batteries in grid-scale energy storage systems. *Transactions of Tianjin University*, 26(3), 208–217.
- Driscoll, L., de la Torre, S., and Gomez-Ruiz, J.A. (2022). Feature-based lithium-ion battery state of health estimation with artificial neural networks. *Journal of Energy Storage*, 50, 104584.
- Huang, S.C., Tseng, K.H., Liang, J.W., Chang, C.L., and Pecht, M.G. (2017). An online soc and soh estimation model for lithium-ion batteries. *Energies*, 10(4), 512.
- Juantorena Polo, A. (2023). Estudio de la degradación de las baterías de ion-litio (lib). NA.
- Li, P., Zhang, Z., Xiong, Q., Ding, B., Hou, J., Luo, D., Rong, Y., and Li, S. (2020). State-of-health estimation and remaining useful life prediction for the lithium-ion battery based on a variant long short term memory neural network. *Journal of power sources*, 459, 228069.
- Plakhtii, O., Nerubatskyi, V., Mashura, A., and Hordiienko, D. (2020). The analysis of mathematical models of charge-discharge characteristics in lithium-ion batteries. In *2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO)*, 635–640. IEEE.
- Saha, B. and Goebel, K. (2007). Battery data set. *NASA AMES prognostics data repository*.
- Song, S. (2019). Nasa. doi:10.21227/fhxc-ga72. URL <https://dx.doi.org/10.21227/fhxc-ga72>.
- Torfi, A., Fox, E.A., and Reddy, C.K. (2022). Differentially private synthetic medical data generation using convolutional gans. *Information Sciences*, 586, 485–500.
- Wang, D., Tan, D., and Liu, L. (2018). Particle swarm optimization algorithm: an overview. *Soft computing*, 22(2), 387–408.
- Zepeda H, J., Aguilar C, A., Hernández S, Á., and Salgado G, M. (2015). Metodología para determinar la topología, patrones de parámetros eléctricos y. *Revista Tecnología Digital*, 103–113.
- Zhang, D., Zhao, W., Wang, L., Chang, X., Li, X., and Wu, P. (2022). Evaluation of the state of health of lithium-ion battery based on temporal convolution network. *Frontiers in Energy Research*, 761.
- Zhu, F. and Fu, J. (2021). A novel state-of-health estimation for lithium-ion battery via unscented kalman filter and improved unscented particle filter. *IEEE Sensors Journal*, 21(22), 25449–25456.